Mobility Management and Congestion Control in Wireless Mesh Networks

by

Fawaz KHASAWNEH

THESIS PRESENTED TO
ÉCOLE DE TECHNOLOGIE SUPÉRIEURE
IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR
THE DEGREE OF DOCTOR OF PHILOSOPHY
Ph.D.

MONTREAL, JULY 15, 2017

ÉCOLE DE TECHNOLOGIE SUPÉRIEURE
UNIVERSITÉ DU QUÉBEC

THIS THESIS HAS BEEN EVALUATED

BY THE FOLLOWING BOARD OF EXAMINERS:

Mr. Michel Kadoch, Thesis Supervisor
Department of Electrical Engineering at École de Technologie Supérieure

Mr. Alain April, President of the Board of Examiners
Department of Software Engineering at École de Technologie Supérieure

Mr. François Gagnon, Member of the jury
Department of Electrical Engineering at École de Technologie Supérieure

Mr. Mohammed Lyakoubi, Independent External Evaluator
Freedom Mobile, Toronto, Ontario, Canada

THIS THESIS WAS PRENSENTED AND DEFENDED

IN THE PRESENCE OF A BOARD OF EXAMINERS AND PUBLIC

ON MAY 29, 2017

AT ÉCOLE DE TECHNOLOGIE SUPÉRIEURE

ACKNOWLEDGMENT

Firstly, I would like to thank God for his continuous guidance through my life. I would like to thank my supervisor Dr. Michel Kadoch, for his continuous support, valuable comments and recommendations during my PhD study period.

I would also like to convey my deepest appreciation to my parents, Amneh and Ali for their unconditional love and support throughout my entire life.

Last and not least, I would also like to thank my sisters, brothers, friends for everything they have done for me.

GESTION DE LA MOBILITÉ ET CONTRÔLE DE CONGESTION DANS SANS FIL RÉSEAUX MAILLÉS

Fawaz KHASAWNEH

RÉSUMÉ

Aujourd'hui, les réseaux maillés sans fil sont de plus en plus populaire. Afin d'être mieux adaptés au nombre croissant de services offerts en télécommunications, de nombreux problèmes de qualité de service (QoS) sont à l'étude. Certaines des questions importantes sont: le contrôle d'admission, le contrôle de congestion et la gestion de transfert du réseau. Cette recherche se concentre sur ces questions individuellement et les combiner ensemble afin de proposer des solutions pour améliorer la qualité du service fourni à chaque utilisateur tel que requis dans leur entente de service.

Un nouvel algorithme de contrôle d'admission basé sur la décision de Markov et de routage (MDACR) est proposé. L'algorithme MDACR trouve une solution sous-optimale en utilisant la méthode d'itération de valeur. Le taux d'admission augmente pour les deux types d'appels (transfert et de nouveaux appels), qui est visé par un algorithme et un routage proposé pour le transfert et les nouveaux appels. Cet algorithme associe l'utilisateur avec deux points d'accès différents. Ceci est avantageux dans un réseau très encombré, ce qui permet une nouvelle mesure de routage pour assurer un transfert transparent dans le réseau. L'algorithme MDACR trouve un itinéraire avec l'ancienne route pour l'appel de transfert, ce qui diminue le retard de transfert.

Un autre aspect est pris en compte afin d'améliorer la qualité de service dans WMN, qui est le contrôle de congestion. Une nouvelle approche proactive est proposée. Lorsqu'un modèle de prédiction d'ordre variable de Markov (VOM) est introduit pour prédire l'état d'encombrement dans chaque lien du réseau, un nouvel itinéraire est établi pour le trafic basé sur la sortie du modèle de VOM, et le taux de transmission est ajusté sur la base de l'état de congestion du lien pour augmenter la satisfaction globale des usagers. Un modèle sous-optimal est introduit et résolu en utilisant la méthode de Lagrange. Basé sur la prédiction de congestion du lien, l'algorithme de réacheminement est mis en œuvre afin d'assurer l'équilibrage de charge et d'atténuer la congestion sur le réseau WMN.

Le but ultime est d'améliorer la qualité de service dans WMN en traitant individuellement les problèmes énoncés ci-dessus en essayant de les combiner et de fournir le cadre de QoS qui traite nombreux types de services. Cette proposition a été simulée en utilisant MATLAB.

Mots-clés: WMN, Gestion de la mobilité, Contrôle de la congestion, Processus de décision Markovien, Contrôle de l'admission, Modèle Markov d'ordre variable, Optimisation Lagrange.

MOBILITY MANAGEMENT AND CONGESTION CONTROL IN WIRELESS MESH NETWORKS

Fawaz KHASAWNEH

ABSTRACT

Today, wireless mesh networks are increasingly popular. In order to be better adapted to the increasing number of offered services in telecommunications, many Quality of Service (QoS) problems are being considered. Some of the important issues are: admission control, congestion control, and handoff management of the network. Our research focuses on those issues individually and combining them together in order to find solutions to enhance the quality of service provided to each user as demanded in their SLA.

A novel Markov Decision-based Admission Control and Routing (MDACR) algorithm is proposed. The MDACR algorithm finds a sub-optimal solution using the value iteration method. Admission rate increases for both types of user associations (handoff and new user association request), which is addressed by a proposed multi-homing admission and routing algorithm. This algorithm associates the user with two different access points. This is beneficial in a highly congested network, which permits a new routing metric to assure seamless handoff in the network. When a user is moving, MDACR algorithm finds a maximally jointed route with the old route, which decreases the handoff delay.

Another aspect is considered in order to improve the QoS in WMN, which is the congestion control, a novel proactive approach is proposed. Where a Variable Order Markov (VOM) prediction model is introduced to predict the congestion status in each link in the network, a new route is established for the traffic based on the output of the VOM model, and the transmission rate is adjusted based on the link congestion status to increase the overall user satisfaction. Sub-optimal model is introduced and solved using Lagrange method. Based on the predicted link congestion, rerouting algorithm is implemented in order to insure load balancing and to mitigate congestion over WMN network.

Our ultimate goal is to improve the QoS in WMN by dealing individually with the issues stated above and try to combine them together and provide QoS framework which deals with many types of services.

Keywords: WMN, Mobility Management, Congestion Control, Markov Decision Process, Admission Control, Variable Order Markov Model, Lagrange Optimization.

TABLE OF CONTENTS

Page

CHAPTER 1 INTRODUCTION 1
1.1 Motivation 4
1.2 Problem Statement 5
1.2.1 Handoff and Mobility Management 6
1.2.2 Admission Control 7
1.2.3 Congestion Control 8
1.3 Objectives 8
1.4 Methodology Overview 9
1.5 Thesis Contribution 10
1.6 Thesis Outline 12

CHAPTER 2 RELATED WORK AND BACKGROUND INFORMATION 13
2.1 Introduction 13
2.2 Background on Prediction Algorithms 13
2.3 Related Work on Mobility Management, Congestion Control and Bandwidth Estimation 16
2.3.1 Mobility Management 17
2.3.2 Congestion Control 33
2.3.3 Bandwidth Estimation 45

CHAPTER 3 MOBILITY MANAGEMENT USING JOINT ROUTING ANDADMISSION CONTROL 49
3.1 Introduction 49
3.2 Our Proposed Mobility Management Algorithm 50
3.3 Problem Formulation and Objectives 55
3.4 Review of Performance Measurements 58
3.5 Markov Decision-based Admission Control and Routing (MDACR) 59
3.5.1 Markov Decision Process Model 61
3.6 System Model 69
3.7 Simulation Results 70

CHAPTER 4 ADAPTIVE TRANSMISSION PREDICTIVE CONGESTION CONTROL 81
4.1 Introduction 81
4.2 Problem Formulation 82
4.2.1 Variable Order Markov Congestion Prevention (VOMCP) 83
4.2.2 Routing Decision 84

4.2.3 Adaptive Transmission Rate Algorithm..87
4.3 Analysis of the Results...91
4.4 Conclusions..94

CONCLUSIONS AND FUTURE WORKS..98

APPENDIX I LIST OF PUBLICATIONS..100

APPENDIX II LIST OF PUBLICATIONS IN COLLABORATIONS WITH OTHER RESEARCHER..102

LIST OF BIBLIOGRAPHICAL REFERENCES...106

LIST OF TABLES

Page

Table 2.1 Mobility Management Algorithms Comparison......................................30

Table 3.1 Rewards Values...56

Table 3.2 Triggering Multi-homing and Routing Based on the Availability Bandwidth..60

Table 3.3 Simulation Parameters..70

Table 4.1 Prediction Probabilities..86

Table 4.2 Parameters Description..89

LIST OF FIGURES

Page

Figure 1.1	WMN Client/Backbone Network Architecture..................................	2
Figure 1.2	WMN Hybrid Network Architecture..	3
Figure 1.3	Our Proposed Congestion Control Algorithm in WMN.....................	9
Figure 1.4	Our Proposed Mobility Management Algorithm in WMN...............	10
Figure 2.1	Adjacent Relationship Table (ART)..	20
Figure 2.2	Relationship Scenarios..	21
Figure 2.3	Illustration for M^3 Scheme...	24
Figure 2.4	Proposed Handoff Scheme..	29
Figure 2.5	NICC Frames: Control Frame and the Data..	40
Figure 2.6	A Summary of ROCS Approach...	42
Figure 2.7	Classification of Bandwidth Estimation Algorithms........................	45
Figure 3.1	New Route Handoff Decision...	51
Figure 3.2	Handoff Management Sequence Diagram...	54
Figure 3.3	Markov Decision Process State Diagram...	61
Figure 3.4	System Model..	70
Figure 3.5	New Call Blocking Probability (NCBP)...	72
Figure 3.6	Handoff Call Blocking Probability (HCBP)..	73
Figure 3.7	Number of Hard Handoffs (NHH)..	74
Figure 3.8	Handoff Delay (HD)..	75
Figure 3.9	Reserved Resources for Joint and Disjoint Routes..............................	76
Figure 3.10	Handoff Delay vs. Population..	77

Figure 3.11	Packet Loss	78
Figure 4.1	Routing Example	86
Figure 4.2	Average End-to-End Delay	91
Figure 4.3	Average Packet Loss	92
Figure 4.4	Average User Throughput	93

LIST OF ALGORITHMS

Page

Algorithm 3.1 Handoff Management Scheme 53

Algorithm 4.1 Estimation Algorithm 83

Algorithm 4.2 Routing Algorithm 85

Algorithm 4.3 Scheduling Algorithm 90

LIST OF ABREVIATIONS

AC	Admission Control
AIMD	Additive-Increase/Multiplicative-Decrease
AOMDV	Ad hoc On-Demand Multipath Distance Vector-based
AP	Access Point
ART	Adjacent Relational Table
C3	Collective Congestion Control
CBR	Constant Bit Rate
CCNF	Congestion Control Notification Frame
CH	Clique Header
CLH	Caching-List based fast Handoff mechanism
CLM-TCP	Cross Layer Mesh-Transmission Control Protocol
CN	Correspondence Node
CNE	Congestion Notification Element
ConT	ConTention-based time division
CSMA-CA	Carrier Sense Multiple Access-Collision Avoidance
CTB	Collaborate Token Bracket
CTS	Clear to Send
CTW	Context Tree Weighting
DCC	Degree of Congestion Control
DHCP	Dynamic Host Control Protocol
DSR	Dynamic Source Routing

ECA-HWMP	Enhanced Congestion Avoidance - HWMP
ETT	Expected Time of Transmission
ETX	Expected Transmission Count
EWCCP	Explicit Wireless Congestion Control Protocol
FA	Foreign Agent
FPBR	Forward Pointer Based Routing
FTE	Frame Transmission Efficiency
GRAP	Gateway RAP
HA	Home Agent
HCBP	Handoff Call Blocking Probability
HCTB	Hierarchical Collaborative Token Bucket
HD	Handoff Delay
HLR	Home Location Register
HWMP	Hybrid Wireless Mesh Protocol
IMCC	Intra-Mesh Congestion Control
IoT	Internet of Things
IP	Internet Protocol
JMM	Joint Multi-channel and Multi-path routing
LAN	Local Area Network
LCQSR	Least Congestion QoS-aware Routing
LZ78	Lempel-Ziv 78
M^3	Mesh Mobility Management
MAC	Media Access Control

MC	Mesh Client
MCAP	Multi-Channel Adaptive Pacing
MD	Minimum Delay
MDACR	Markov Decision-based Admission Control and Routing
MDP	Markov Decision Process
MEMO	MEsh networks with MObility management
MH	Mobile Host
MH	Multi-Homing
MHWN	Multi-Hop Wireless Networks
MJ	Maximally Jointed
MJO	Maximally Jointed with the mobile user Old
MJTO	Maximally Jointed Trio paths with the mobile user Old
MLR	Mobile Location Register
MN	Mesh Node
MR	Mesh Router
MR-LQSR	Multi-Radio Link-Quality Source Routing
MS	Mobile Station
NAV	Network Allocation Vector
NCBP	New Call Blocking Probability
NDOUTE	Node-Disjoint rOUTE algorithm
NDP	Network Discovery Protocol
NHH	Number of Hard Handoffs

NIC	Network Interface Card
NICC	NeIghbourhood-based Congestion Control
NTP	Network Time Protocol
OSI	Open System Interconnection
PPM	Prediction by Partial Match
PST	Probabilistic Suffix Trees
QoE	Quality of Experience
QoS	Quality of Service
RAECA-HWMP	Rate Adaptive Enhanced Congestion Avoidance - HWMP
RA-OLSR	Radio-Aware Optimized Link State Routing
RAP	Routing Access Point
RCP	Rate Control Protocol
RREP	Route REPly
RREQ	Route REQest
RSSI	Received Signal Strength Indicator
RTS	Request to Send
SMesh	Seamless Wireless Mesh Network
TCP	Transmission Control Protocol
TDD	Time Division Duplex
TDM	Time Division Multiplexing
TMIP	Transparent Mobile IP
UDP	User Datagram Protocol
UGRS	User Group Representative Subgroup

VLR	Visitor Location Register
VMM	Variable order Markov Models
VoIP	Voice over IP
VOM	Variable Order Markov
VOMP	Variable Order Markov Prediction
WAHN	Wireless Ad-Hoc Network
WCETT	Weighted Cumulative ETT
WCP	Wireless Control Protocol
WCPCap	Wireless Control Protocol with Capacity estimation
WiFi	Wireless Fidelity
WiMAX	worldwide interoperability for microwave access
WMM	Wireless Mesh Mobility management
WMN	Wireless Mesh Network
WSNs	Wireless Sensor Networks
Xcast	eXplicit Multicast
XCP	eXplicit Control Protocol
XGRs	Xcast-based Group Routers

LIST OF SYMBOLS

$A_i(T)$	The number of users accepted at node i over a time period T
b_{al}	Available bandwidth in link l
b_{rk}	Minimum required bandwidth for user k
$f_{ij}(T)$	The weighted number of accepted users at node $i$over a period of time T
f_{ij}^{s}	Outgoing traffic
$g_s(x_s)$	Utility function
h_{ij}^{s}	Incoming traffic
$J^*(S_i)$	The expected future reward after moving to state i
$L(\ldots)$	Lagrange function
P_{ij}	Transition probability from state i to state j
P_{si}	The probability of being at state i
$\underline{P}[x]$	The matrix of all the transition probabilities from one state to another at time x
R_{ij}	The reward assigned in the case of moving from state i to state j
R_j	Reward assigned when selecting one path over the other
S	Set of traffics
w_{ijr}	The joint degree weight
x_s	Allocated bandwidth

CHAPTER 1

INTRODUCTION

Wireless Mesh Network (WMN) is a communication network which consists of two types of nodes, namely mesh clients and mesh routers. Mesh Clients (MC) can act as a small radio transmitter, with limited power. An MC could be a phone, PDA, laptop, or any other device with Network Interface Card (NIC) or Ethernet card. A wireless network can be formed by a group of MCs, organized in a mesh topology. Mesh clients can also be connected to a Mesh Router (MR). MRs has no restrictions with the power supply and is responsible for performing essentially routing, bridging between WMN and other types of networks such as Wireless Fidelity (WiFi), Wireless Sensor Networks (WSNs), worldwide interoperability for microwave access (WiMAX) and cellular networks. As previously mentioned, in case of emergency a wireless network can be formed by a group of MCs. MRs are the backbone of WMN network. It is worth mentioning that MRs have minimal mobility compared to MCs in the network (Akyildiz et al., 2005).

Wireless Mesh Network (WMN) can be categorized into three different architectures (Akyildiz et al., 2005; Xie et al., 2008). Firstly, the basic wireless infrastructure (or backbone) architecture is formed by the mesh routers as shown in Figure 1.1 (b), which allow wireless connectivity, routing and gateway functionality. In addition to the IEEE 802.11 radio protocol, various radio technologies may be integrated to improve radio communication within the infrastructure. In this architecture, the mesh routers create self-healing, self-configuring links among themselves for the provision of a wireless link. Conventional clients with a wired connection may connect to the wireless router via an Ethernet link. Having the same radio protocol, clients can directly access the wireless network, otherwise they access the network via their own base stations. Secondly, there is the client WMN architecture, where client nodes create a peer-to-peer network among client devices and maintain configuration and routing functions as shown in Figure 1.1 (a). Mesh routers are not required. The protocol used in this architecture is based on a single radio and requires more functionalities to be served by client devices compared to those in the backbone architecture.

Thirdly, WMN is a hybrid of both backbone and client WMN architectures as shown in Figure 1.2.

The advantages of this architecture are that it uses the backbone to integrate different types of radio technologies (WiMAX, Wi-Fi, cellular, etc.) into the same platform, and the client WMN architecture facilitates better connectivity between clients and wider coverage inside the network (Akyildiz et al., 2005; Xie et al., 2008).

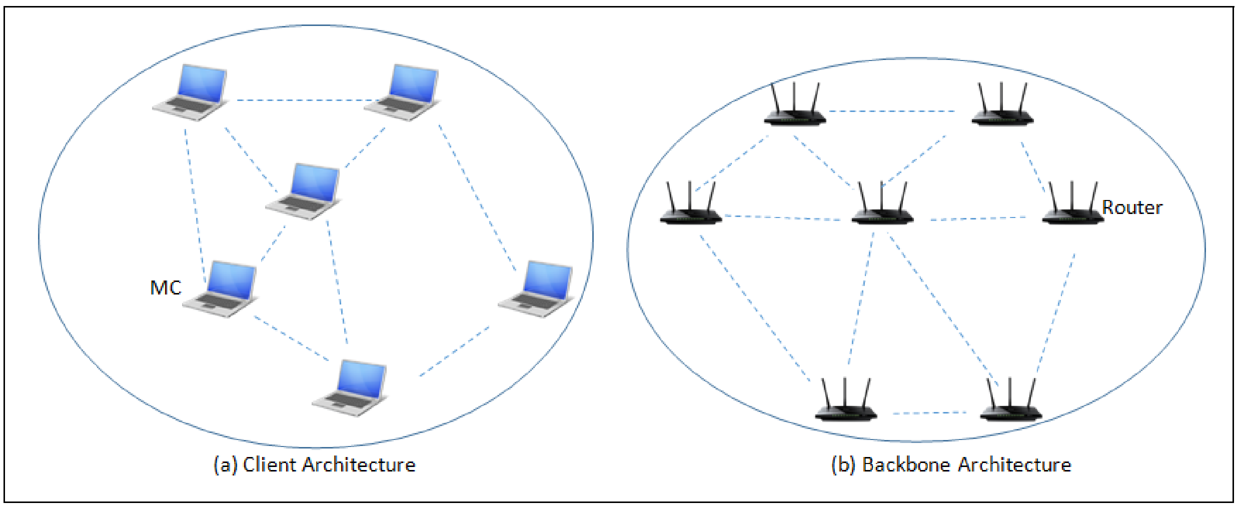

Figure 1.1 WMN Client/Backbone Network Architecture

WMN is used widely because it has high speed, ease of implementation and low setup cost. A most promising application is the provision of last mile wireless Internet access(Sun et al., 2009). Networking solutions rely on the versatility and flexibility of a WMN. However, a WMN has intrinsic performance and topology problems because its network is dynamic and unpredictable. These cause Quality of Service (QoS) and reliability issues, which need to be resolved. WMN performance can be improved, without compromising its best features through the design of a joint QoS aware admission and routing protocols. WMN requires a tree-based routing algorithm due to the static topology structure. The route discovery process involves large overhead. Many algorithms have been proposed that use different routing metrics to adjust and assess routing conditions in the WMN. Among others, these routing metrics have included interference, packet loss, link quality, congestion and hop count(Parissidis et al., 2009; Perkins, 1999; Navda et al., 2005). Routing issues can be solved by a single routing metric, however, there are trade-offs. WMN lacks a single routing algorithm capable of simultaneously using routing metrics to consider all the routing issues.

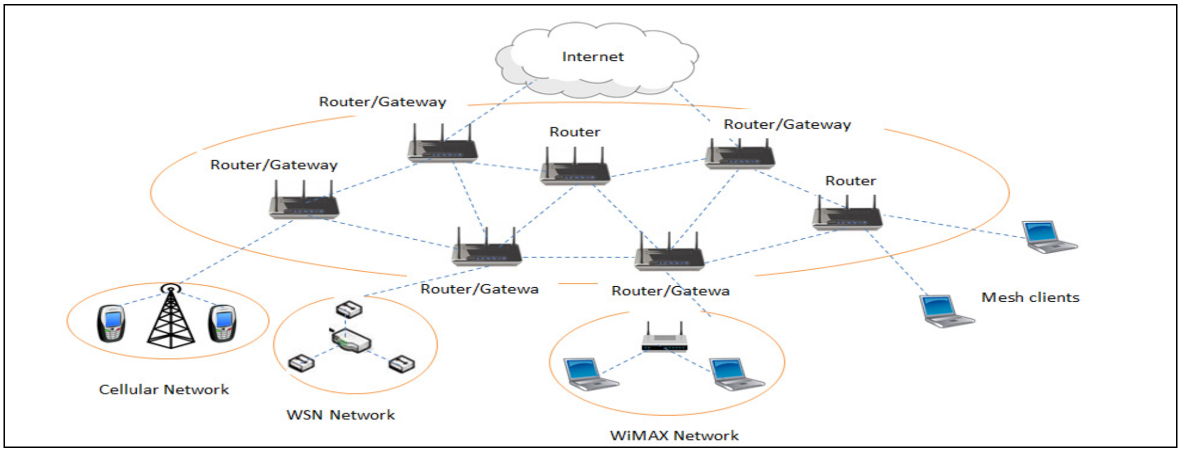

Figure 1.2 WMN Hybrid Network Architecture

QoS required by an application can be defined using the following performance metrics:

- Throughput: This can be defined as the number of packets received per time unit. It can be normalized by dividing the receiving packet rate over the sending packet rate. The objective is to improve/increase it as much as possible;
- Packet loss rate: it is the percentage of packets that are lost in the network due to congestion, link/node failure or any other factor. The objective is to improve/decrease it as much as possible;
- Delay: there are two types of delays in the network the first one is the end-to-end delay and the second one is the handoff delay which can be defined as the time it takes to transfer the control of the user traffic from one access point to another. The objective is to improve/decrease both delays as much as possible.

It is not necessary to sub-optimize all the factors mentioned above to support QoS for any kind of traffic. For example, it is obvious that the voice/video traffic is delay sensitive traffic but it can tolerate some packet loss in the network so delay has to be sub-optimized. On the other hand, the data traffic can tolerate delay but the packet loss rate is sub-optimized.

A lot of challenges have to be dealt with in order to provide QoS in WMN. Some of those challenges are limited capacity, inaccurate bandwidth estimation and node mobility. In order to have an effective mobility management and congestion control algorithms, those challenges should be encountered (Marwaha et al., 2008).The work introduced in this thesis could also be extended and adapted to provide a seamless handoff solution in multi-hop 5G heterogeneous networks. This will be done by taking into consideration our proposed algorithms to improve the QoS provided by existing handoff management and congestion control algorithms proposed in the literature.

1.1 Motivation

In the era of Internet of Things (IoT), many devices will be interconnected which is expected to reach 50 billion by 2020 (Jianguo 2014). This will generate a lot of traffic in the network and cause more congestion. On the other hand, the rapid increase in the number of applications which has a special QoS requirements has urged researchers to concentrate on improving QoS in WMN. QoS can be simply defined as the process of giving preferential treatment to some packets over the other in the network. WMNs have many advantages over other wireless networks which motivate us to provide the QoS for the applications that run over those networks. Those advantages can be summarized as follows(Karthika , 2016) :

- Low setup cost since it uses fewer wires. Only one node needs to be wired and this wired node will share its connectivity with other nodes in a wireless manner;
- Extended network coverage ranging from Personal Area Network (PAN) to Metropolitan Area Network (MAN). A lot of standards can be used such as IEEE 802.16a, IEEE 802.11 a,b and g, and IEEE 802.15.5, which makes the application usage for WMN architecture is scalable (Myung et al., 2006).

- It relies on IEEE 802.11a, b and g standards which are the same standards used in most wireless network types;
- Self-configured networks which means if any new node needs to be incorporated in the network this will be done automatically without any modifications done by the network administrator;
- Self-healing networks which means when one of the nodes fails all the packets that are going through this node will change their routes and finds the most reliable path to send the packets through.

Since WMN has many advantages over other wireless networks and getting more popular, providing QoS in those networks is an important topic that has to be addressed. A necessity to design a novel routing algorithm allowing for the integration of multiple routing metrics yields to a widely compatible routing protocol with consistent and reliable output motives us to study the routing algorithms for stationary and mobile nodes to assure better handoff delay, packet loss, throughput, etc.

In an era of technology, user satisfaction is one the most important key points that have to be assured when designing any QoS aware algorithm. Most of the users are not satisfied when a network is highly congested, QoS degradation is experienced by some or most of the users. This motivates us to design an algorithm dealing with this specific situation, trying to maintain the QoS provisioned for the existing users in the network as well as admitting new users into the network with no effect on the QoS provided for the existing users.

1.2 Problem Statement

WMN has many applications which require a specific QoS requirement. As more devices are connected together, and due to the nature of the mesh clients in WMN which are moving continuously. A necessity of designing a more robust mobility management with less handoff delay is crucial. In this thesis, the most important factor that we concentrate on is the user satisfaction. In order for the user to be more satisfied, a seamless handoff for the mesh clients should be performed, so that the user will not experience any interruption. To handle the mobility in WMN it should be done on three layers, which are application layer (to change the session itself), network layer (to change the route and IP address, if needed), and the data link layer (to change the channels). The layer that has the largest portion of the handoff delay is the network layer (Zhao et al., 2012). Our proposed algorithm deals with finding a new route, which is maximally jointed with the old primary route as explained in Chapter 3. Our algorithm outperforms other algorithms proposed in the literature.

On the other hand, as the traffic volume in our network is growing exponentially and due to the popularity of WMN, vast applications and the limited resources in WMN (Jianguo 2014, Akyildiz et al., 2005). Resource management and utilization is another problem that has to be considered. Adaptive and predictive design is an essential thing to have in our congestion network to have better resource utilization. A lot of research has been done in the literature to handle the congestion in WMN. The congestion detection is the first step in those algorithms, which can be done implicitly or explicitly (Alnuem et al., 2007). Each approach has its own advantages and disadvantages, which will be discussed in more details in Chapter 2. In order to have an efficient detection algorithm, the available bandwidth should be accurately estimated. A taxonomy for mobility management, congestion control and bandwidth estimation is introduced in Chapter 2.

Due to the popularity of WMN, Quality of Service (QoS) is a substantial element that has to be provided for next-generation networks. As previously mentioned, many challenges have to be handled in order to provide a suitable QoS for each application in WMN. Below we will discuss the mobility management, admission control, and congestion control.

1.2.1 Handoff and Mobility Management

When a mobile node moves from one access point to another a mobility management algorithm or scheme should be applied to maintain the continuation of the connection. Many aspects should be considered to provide a good mobility management scheme. The first aspect is the handoff delay, which mainly consists of switching the channels if needed, changing the routing path/IP address for the mobile mesh client and changing the multimedia communication session in case of inter-gateway handoff (Zhao et al., 2012). The second aspect is the location management, in which the location of the mobile mesh node is updated. This can be done in different approaches, which will be discussed and categorized in details in Chapter 2. The challenge in the handoff and location management process is the minimization of the handoff latency and signalling overhead.

1.2.2 Admission Control

An admission control algorithm will check for the resources availability to admit or reject the incoming traffic based on the QoS required and the available bandwidth in the network. Two main challenges are encountered to implement an admission control in WMN. The first challenge is the bandwidth estimation. If the proposed algorithm underestimates the available bandwidth in the network, it might reject the new incoming traffic due to the lack of bandwidth while there is available bandwidth in the network and new traffic can be accepted without affecting the QoS provided for other users. The underestimation of the available bandwidth will decrease the utilization of the network resources. The accuracy of available bandwidth estimation is an important factor that has to be maintained in order to make the right decisions for better QoS (Rabia et al., 2011). On the other hand, if the available bandwidth is overestimated then new user's traffic will be accepted and it will affect other users and cause QoS degradation and network congestion. Another challenge is to provide a scalable admission control algorithm which can be applied in large network without making any kind of congestion since each node has to have a local/global knowledge on other nodes in the network to make its admission decision. These control messages can cause a signalling overhead if the scalability parameter is not considered when designing the admission control algorithm.

1.2.3 Congestion Control

When congestion occurs in the network it will affect the end-to-end delay, packet loss rate and throughput of the network. So traffic monitoring and congestion prediction algorithm should be designed to observe the traffic flows to avoid or predict any kind of congestion. There are two types of congestion control algorithms. The first type is the preventive type which expects the congestion before it really occurs in the network and use traffic engineering technique to change routes and prevent congestion. The other type is the one which deals with the congestion after occurring by adapting the flow rate or by applying the network maintenance concept and expand the network. The main challenge in the traffic monitoring algorithms is the minimization of the control messages overhead.

1.3 Objectives

In this thesis, we will concentrate on studying how the congestion and mobility are handled in wireless mesh network. Our main objective is to enhance the QoS provided for users in highly congested network situations, by designing a seamless handoff and mobility management algorithm with better handoff latency compared to the other algorithms in the literature. This thesis will also introduce a novel adaptive transmission rate predictive congestion control algorithm.

In order to achieve our goals, a Markov Decision-based Admission Control and Routing (MDACR) algorithm is proposed. Which proposed a routing algorithm for managing the mobility in WMN and applying a multi-homing concept when necessary. A multi-homing concept can be defined as associating the mobile mesh client with two access points at the same time for better network resources utilization and more user satisfaction. Another algorithm is proposed in this thesis, which is called Variable Order Markov Prediction (VOMP) algorithm, which predicts the congestion before it really happens in the network. The routing decision is made based on the congestion prediction and the transmission rate is also adjusted based on the predicted congestion status. Lagrange's method is used in order to calculate the new transmission rate based on the predicted congestion status.

1.4 Methodology Overview

In our methodology, we designed a Variable Order Markov (VOM) based congestion aware routing algorithm, which adjust the transmission rate based on the prediction performed by the VOM prediction model. The prediction output that comes from the VOM prediction model feeds our sub-optimal model with the new route and based on the predicted congestion status in the new route, a new transmission rate is applied as shown in the figure below. The solution for this sub-optimal model is found using Lagrange's method. The congestion history database is filled using the Explicit Congestion Notifications (ECNs) received (Lee et al., 2017).

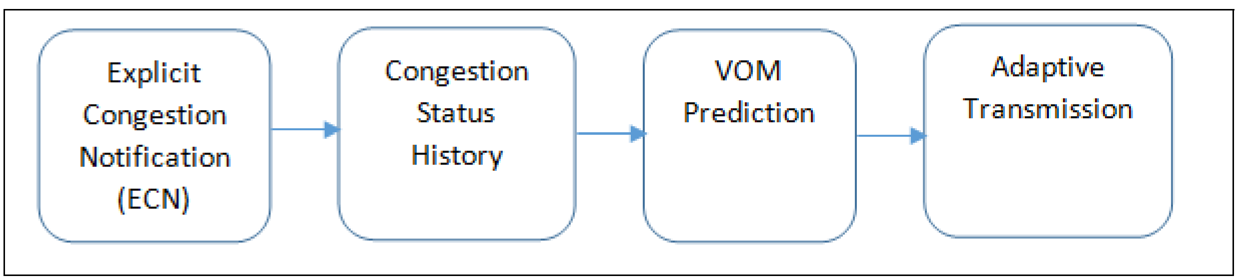

Figure 1.3 Our Proposed Congestion Control Algorithm in WMN.

On the other hand, another aspect should be considered to improve the QoS, which is the routing and controlling the admission control in our network, in order to maintain the QoS for the existing traffic. Whenever a new user's traffic arrives to a highly congested network, rejection is highly probable due to lack of bandwidth in the network. Our proposed algorithm will increase the admission rate by associating the new user's traffic with two access points, routing user traffic, from the source to the destination, through two different paths using multi-homing.

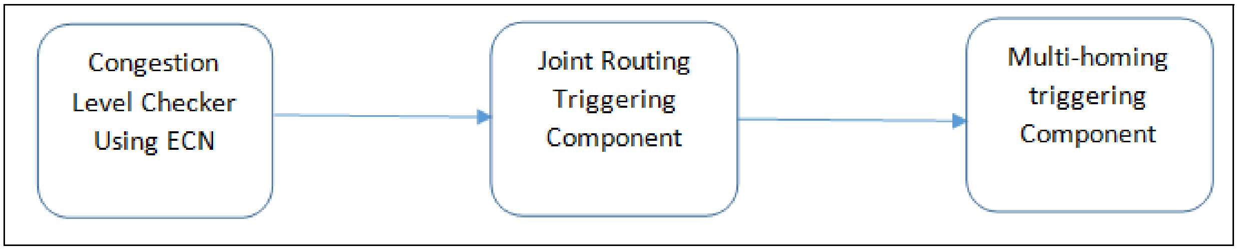

Figure 1.4 Our Proposed Mobility Management Algorithm in WMN.

Those two paths may be joint or disjoint. As shown in the results section, more reward will be assigned to the joint path, as this provides better QoS compared to the disjoint path. We design a Markov Decision Process (MDP) to make decisions on the selection of the route for the handoff, and new users. The benefits of our model on QoS arise from its use of joint routing and multi-homing over the disjoint routing and no multi-homing scenario in terms of QoS. As shown in the figure above, the multi-homing is triggered when the congestion level is high in the network and there is no available resources for the handoff or new user's traffic. The congestion level is detected through the Explicit Congestion Notifications (ECNs) received.

1.5 Thesis Contribution

In this thesis, two main contributions are introduced, which can be divided into multiple sub-contributions. Whenever a user is moving, handoff user's traffic have more priority compared to the new user's traffic. The first contribution in this thesis is the proposition of a new admission and routing algorithm, the algorithms proposed in the literature consider finding a completely disjoint path from the old path for the user handoff traffic, which makes the handoff process does not go smoothly and the end user could notice a degradation of QoS while moving. Our algorithm makes smoother handoff for the end user by proposing multi-homing and maximally jointed path for the handoff user's traffic. For the new user's traffic, the algorithms in the literature accept/reject the call based on the available resources. However, our algorithm accepts or rejects a new call based on the available resources in the

network, but it also tries to accept the new user's traffic by triggering the multi-homing in a highly congested network, which makes more rooms available for the acceptance of new traffic and utilize the network resources in a better way. This has been mathematically modelled using MDP, the sub-optimal solution is found using the value iteration method.
The other contribution is the proposition of a new congestion and routing algorithm. Which is an adaptive algorithm that can adjust the transmission rate based on the predicted congestion link status. The prediction phase allows us to route our packets in an efficient way, and the new route is used in our sub-optimal model in order to maximize the transmission rate. A new mathematical model to increase the user satisfaction for better user Quality of Experience (QoE) is introduced by defining the user satisfaction as part of the Heaviside and Sigmoid function.

Our proposed algorithm for mobility management outperforms other algorithms in the literature in terms of blocking probability for the new and handoff user's traffic, number of hard handoff, handoff delay and packet loss. The other proposed algorithm for the congestion control is also compared with other algorithms in the literature and shows betters performance in terms of delay, packet loss and throughput.

1.6 Thesis Outline

The rest of the thesis is organized as follows: in Chapter 2, we present some background information about Markov Decision Process (MDP), various prediction algorithms used in the literature, and Variable Order Markov (VOM) prediction. In the same chapter, related work is discussed about mobility management and congestion control techniques proposed by other researchers.

In Chapter 3, our proposed joint routing and admission control is presented along with problem formulation, simulation results. In Chapter 4, our proposed congestion prediction and adaptive transmission rate algorithm is introduced, which is based on VOM prediction algorithm. Results analysis and discussion are also introduced in this chapter. Finally, we conclude our work and discuss possible research directions in Chapter 5.

CHAPTER 2

RELATED WORK AND BACKGROUND INFORMATION

2.1 Introduction

In this Chapter, we explain some background information about the different prediction algorithms available and we introduce the literature review on handoff management and congestion control in wireless mesh network.

2.2 Background on Prediction Algorithms

Prediction in general is an important topic to study and has so many applications. This has urged researchers to study the performance of different prediction algorithms exist in the literature. In (Begleiter et al., 2004), six different prediction algorithms has been explained and their performance has been compared. In the following section we will explain some background information on different Variable order Markov Models (VMM) available.

VMM Prediction Algorithms

There is a subtle difference that exists between the different algorithms that have been tested in various studies. To understand the VMM predictions, it is important to describe in detail the six related algorithms. In this description we have taken into consideration the differences and the commonalities between these algorithms. The basis of learning the VMMs entails three main components. This includes smoothing, counting, and variable length modelling. Smoothing involves probabilities of the unobserved scenarios or events. Such kind of algorithms bases its probability estimates mostly on the number of counts of the symbols appearing in the training sequence. The numbers in this sequence generate predictor. On the other hand, the smoothing element establishes how to handle the different unobserved events in the series. The presence of these events can also be referred to as "zero frequency problem".

There are various ways in which the variable length modelling can be carried out. The algorithm described in this section only generates a single model while some algorithms construct several models and establish their average. To bind this model a pre-determined constant bound is introduced. This implies that the algorithms do not recognize the contexts that are greater than the limit. On the other hand, the models may not be delineated in the case where the maximal context size is data-driven.

There are various VMM prediction algorithms. It is important to note even any lossless compression algorithm can act as a predictor. To describe the six VMM prediction algorithms we have to include algorithms which are known to be top actors in lossless compression. The algorithms included are ‘Context Tree Weighting ‘(CTW) and (PPM) ‘Prediction by Partial Match'. Another prediction algorithm that was included is the Lempel-Ziv 78(LZ78) that is a component of the commercial applications used for compression (Begleiter et al., 2004).

Lempel –Ziv 78 (LZ78)

This is the most shared and most popular lossless compression algorithms (Begleiter et al., 2004; Langdon, 1983; Rissanen, 1983). This algorithm is mostly used as a component of the Unix compress utility and other prevalent archiving services for the PCs. Moreover, its performance is guaranteed especially with several analysis models. The compression algorithm has attracted massive attention and has significantly inspired the area of sequence prediction and lossless compression. (Rissanen,1983; Langdon,1983) described the prediction component of this algorithm for the first time. This algorithm begins with a dictionary that contains the empty phrase. The algorithm parses a new phrase at each step.

The algorithms can be discussed in a simple form in the binary case where § = {0, 1}, alternatively the algorithm can still be extended to alphabets of different sizes. In this algorithm the counter in an internal node can always be maintained. There are many performances that LZ78 compression has been proved and guaranteed (Begleiter et al.,2004), hence the Lz78 has been categorically shown to be a universal prediction algorithm for the large class of Ergodic Markov and stationary sources of finite order.

Prediction by Partial Match (PPM)

This algorithm is considered as one of the best existing lossless compressions(Cleary et al., 1984). It specifically requires an upper bound D, which is on the Maximal Markov order of VMM. PPM deals with the problem of zero frequency by using two mechanisms known as exclusion and escape. PPM has had variants that are distinguished by the establishment of the escape mechanism. The PPM variant determines the specific mass allocation for the "escape" and the particular weight distribution of the other symbols. The exclusion mechanism in this algorithm is used basically to enhance the escape estimation. The particular PPM option to be considered is referred to as the "Method C" (PPM –C). One of the applications of the PPM-C is mainly based on a tree data structure.

The Context Tree Weighting Method (CTW)

The CTW (Willems et al., 1995) is a strong lossless compression algorithm that is founded on a brilliant idea of merging several VMMs of bounded order exponentially. CTW estimates various sets of the prospect distributions for each model, each distribution is thereby a smoothed style of a maximum probability estimate that is centred on the training sequence. The fundamental idea of the CTW algorithm is to create a predictor that comprises all of the tree sources.

CTW for Multi-Alphabets

The CTW algorithm specifically for a binary alphabet can apparently be extended in a way that it naturally handles the sequences over a larger alphabet. Despite the easy attainment of these extensions, the resulting algorithm is said to perform poorly (Volf (2002)). Hence there is a huge challenge when it comes to extending the CTW algorithm. In (Tjalkens et al., 1997) they considered a more refined binary decomposition of CTW.

Probabilistic Suffix Trees (PST)

This prediction algorithm (Ron et al.,1996) tries to establish the single D-bounded VMM about the training sequence. The objective of the PST learning algorithm is to find out a good suffix set S for a PST tree and to allow a probability distribution.

LZ-MS: An Improved Lempel –Ziv Algorithm

The classic LZ78 compression algorithm has many variations. This algorithm has two main parameters that are M and S and therefore its acronym derived as LZ-MS (Nisenson et al., 2003).The main advantage of this algorithm is its speed. The speed is a result of compromising on the different systematic count of all sub-sequences.

2.3 Related Work on Mobility Management, Congestion Control and Bandwidth Estimation

Mobility management is composed of handoff management and location management and can be handled on three different layers, which are the application, network and the data link layers (Zhao et al.,2012). Another important topic that has been studied in this thesis is the congestion control. Since the cause of the packet loss and long delay is not only caused by congestion and could be caused by interference, node failure, hidden/exposed, etc (Islam et al., 2011). A novel algorithms should be introduced for better congestion detection. Both the mobility management and congestion control need an efficient bandwidth estimation algorithm in order to have a reliable and accurate algorithms, thus bandwidth estimation algorithms are also studied and discussed in this chapter.

2.3.1 Mobility Management

The mobility management algorithms proposed in the literature can be categorized using different parameters. As previously mentioned most of the mobility management algorithms proposed consist of two main steps, which are the handoff management and location management (Li Y. et al., 2012). In the handoff management, the transfer of control from one access point to another is handled. In the location management, the location of the user is updated using different approaches as discussed below. Some of the algorithms handle the location management alone, other algorithms handle the handoff management and some of them studies both issues. We classified the location management approaches into four main classifications, which are:

1) **Tunnelling Based approach (TB):** in this approach, the location updates are done explicitly by sending a location registration messages to a centralized node such as the gateway. This approach is not suitable for situations where the mesh clients are moving continuously in a dynamic and repetitive way, since it will cause a lot of overhead in the network (Li Y. et al., 2012).

2) **Routing Based approach (RB):** on the other hand, the routing approach is sending the location update messages implicitly by embedding this information in the header of the regular packet routing. This approach does not incur a high overhead compared to the tunnelling approach, but it is not suitable in situations where the user is moving but not sending any kind of traffic, in this case the location information is not instantly accurate (Li Y. et al., 2012).

3) **Hybrid Routing and Tunnelling Based approach (HRTB):** in this approach, if the user is moving in a small range, a routing approach is used and no need to send an explicit location update for the gateway. In this scenario, the gateway for example is sending the packets to the old serving mesh router and the old server mesh router knows that the mesh client has moved to a neighbour mesh router, old mesh router will check its routing table to find a forwarding pointer to route all the received packet to the new neighbour serving mesh router. But if a user moves outside a predefined range (bigger range), it means that the forwarding chain is long enough to make extra end-to-end delay. In this specific scenario, a tunnelling approach is used and an explicit location update message is sent to the gateway (Li Y. et al., 2012).

4) **Multicasting Based approach (MB):** in this approach, groups of mesh routers are formed and managed by special routers. Those special routers will be responsible on forwarding handoff and location information (Zhao et al., 2012).

Other parameters are considered in our mobility management classification, some algorithms are concerned about studying if the mobility is happening in the same domain, which will be shortened as IntraDH (Intra Domain Handoff), the mobility could also happens between different domains, which will be shortened as InterDH (Inter Domain Handoff) or a mobility management solutions which handle the inter and inter domain mobility. The necessity of IP Address Change (IPAC) of the mobile node is also another parameter we are considering, which is required when the mobile node is moving between different domains. The Number of Handoff Thresholds (NHT) to trigger the handoff, most of the algorithms have one handoff threshold except a solution proposed by (Jesus et al., 2016) which has two handoff thresholds for triggering the handoff. Joint Routing (JR) and Disjoint Routing (DR) for the mobile node is also considered. Multi-Homing Feature (MHF) is also an important parameter which is introduced in our proposed algorithm along with the maximally joint routing to manage the mobility in WMN. A detailed explanation of the mobility management algorithms in the literature are introduced below.

Tunnelling Based Mobility Management Approaches:

Since packets of data and packets of handoff signalling are sent on the same backbone channel in single-channel single-radio based WMNs, the channel conflict between data packets and signalling packets often creates long signalling packet queuing delay and channel access delay (Li et al., 2012).Therefore, the performance of handoff is largely determined by the volume of backbone data traffic. Most of existing handoff management solutions do not care about this important issue, in multi-hop WMNs. (Li et al., 2012) proposes scheme to reduce the delay in handoff in WMNs based on single-backbone channel from different viewpoints and a contention-based time division (ConT) scheme to shorten the channel access delay and queuing delay of packets of handoff signalling over multi-hop wireless paths.

The proposed methodology separate the data packets and the signalling packets and put them in a different queues. The data and signalling packets are also transmitted in a different times, the authors have specified a time slot for each type of packets. In this way, it is guaranteed that signalling will not be affected by the transmission of the long data packets transmitted in the network. Location update is done by sending the an explicit signalling messages to the gateway.

Nowadays, wireless mesh networks are developing as a quick-and-cheap method for wide wireless coverage. (Wang et al., 2007) proposes a network-based native mobility management technique for wireless mesh networks. When a Mobile Host (MH) connects to a network domain, firstly it focuses on the MAC-layer link with a Routing Access Point (RAP) in the network. By the virtue of this process of association, RAP becomes aware of the MAC address of the network interface card of the MH or, in other words, the MH-ID. As soon as RAP knows the ID, it transmits a message regarding location update with the acquired MH-ID, to the location server. Location server matches this ID with its database to find whether the MH is newly connected or existing one. If no entry is found, location server recognizes this MH as a new MH. Location server then requests for the assignment of new IP address to this new MH, to the DHCP server. The DHCP server sends a DHCP message following standard DHCP protocol, which contains information regarding newly assigned MH-IP. This MH-IP generally belongs to the same subnet group, within which the MH locates. The Gateway RAP (GRAP) becomes the default gateway. On the other hand, if the location server finds an entry of the MH in its database, it simply updates the location information and sends back an acknowledgement message to the RAP (Wang et al., 2007). Location update is also done in an explicit way by sending a location update message to the location server.

Routing Based Mobility Management Approaches:

In (Majumder et al., 2012; Majumder et al., 2014), a Forward Pointer Based Routing (FPBR) is proposed to manage the mobility in WMN. The first step in this scheme, is the formation of the Adjacent Relationship Table (ART). This table exists in each mesh router and gateway to define the relationship between two mesh routers in the network, which can be either parent, child or sibling. The fields of this table is defined in the figure below. This step is initiated by the gateway and goes down to each mesh router in the network.

Host IP	Relation	Interface	Link Capacity
….	….	….	….

Figure 2.1 Adjacent Relationship Table (ART)

The ART table formation can be done offline due to the limited mobility of the mesh routers and gateways. In this scheme, whenever the handoff is triggered the handoff process is gone through two steps which are: 1) Registration: in this step, a handoff notification message is sent to the old serving mesh router and a registration request message is sent to the new mesh router. Once receiving the handoff notification message by the old serving mesh router, it will be forwarded to the new candidate mesh router. After the new mesh router receiving both messages (the handoff notification and the registration request messages), it can finally register the mesh client. A handoff confirmation message and an acknowledgement message are sent to old mesh router and mesh client, respectively. 2) Route Optimization: by checking the ART table, the new Mesh Router (MR) can decide whether to send the Route Setup Message (RSM) or not based on three cases defined in the figure below. If the new mesh router is a child of old mesh router, then there is no need to send the route setup message. On the other side, if the new mesh router is the parent of old mesh router, then the new mesh router will send all the handoff information to the other siblings mesh routers (S1 and S2).If the ACK messages are received back from the sibling nodes, it means that a corresponding entry for the mesh client has been found in the routing tables of the sibling nodes and the next hop field is updated.RSM message is sent to the neighbouring parent in case that the new mesh router received only NACK (Negative ACK) from the sibling nodes(Majumder et al., 2012; Majumder et al., 2014).

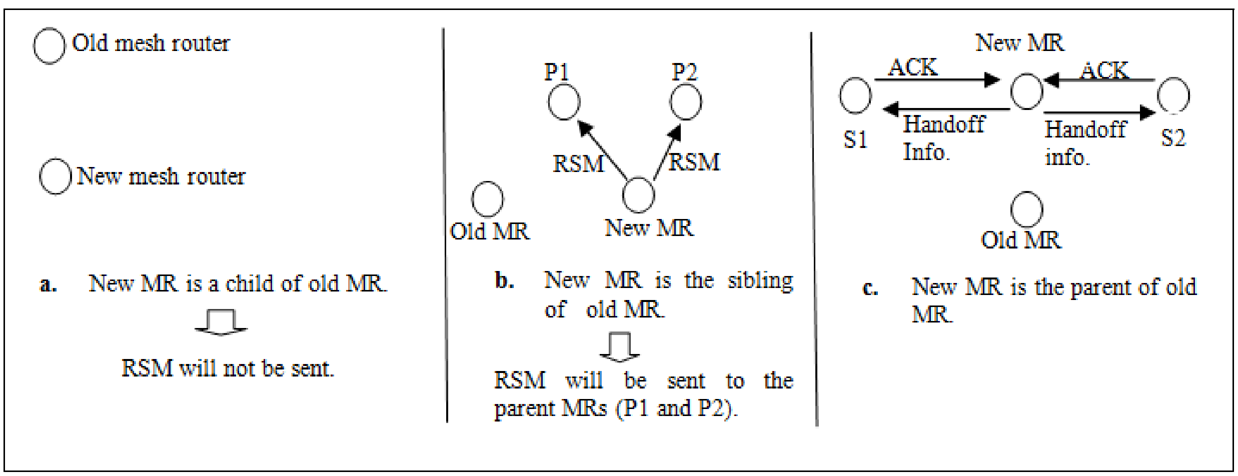

Figure 2.2 Relationship Scenarios

The last possible scenario, is that the new mesh router is the sibling of the old mesh router. In this scenario, the RSM messages will be sent to all the neighbouring parents (P1 and P2).

Another mobility management protocol for wireless mesh network called MEMO (MEsh networks MObility management) has been proposed by (Ren et al., 2007). MEMO supports transparency in client access and management for mobility at mobile nodes. The proposed design differs from the traditional design in the fact that its operation initiates by the MAC layer triggered mechanism and it makes use of a special routing algorithm namely AODV_MEMO that integrates both mobility management and routing solutions into a single package. This routing algorithm, AODV_MEMO, is advantageous because it does not necessitate any mobile node modification, it manages the mobility in a distributed way rather than in a centralized fashion, it offers low handoff latency of mobile nodes by facilitating cross layer communication between MAC and Routing Layer. Moreover, it offers an effective solution for the gateway functioning.

DHCP (Dynamic Host Configuration Protocol) (Droms 1997) server performs the IP addressing task. Unlike traditional approach which has a centralized DHCP server, DHCP server is installed on each Mesh Router (MR) that assigns IP address to a new Mesh Node (MN) based on its MAC address. A simple hash technique is deployed to allocate IP address to the new MN dynamically. Thus, every device is assigned with a unique IP address within the IP range 10.x.x.x based on its device MAC address. Where the last 24 bits of the IP address is the same as the last 24 bits of the MAC address and the first 8 bits of the IP address is always equal to 10, in this way the authors guarantees the seamless roaming in the network where the IP address change is not necessary when MN is moving from one domain to another. After assigning an IP, the MR selects a proper channel for the MN to connect to the network based on the best signal strength. Thus the MR in the proposed MEMO technique is self-configurable, which makes the deployment of the technique convenient, useful and error free.

A mesh backbone is configured in WMN to provide Internet connectivity, which takes the assistance of Mesh Access Points (MAPs) for the provision of wireless network connectivity to the Mobile Clients. These MAPs are static and are interconnected via the links of wireless mesh. To ensure correct and perfect delivery of packets from the MAPs to the Mobile Clients, mobility management issue gets highest priority because it is a service parameter for a good network. (Huang et al., 2008) proposes a unique mobility management technique called Wireless mesh Mobility Management(WMM).The underlying principle is based on the location-caching scheme; the mesh backbone and the MAPs are responsible for caching the location information of Mobile Clients to facilitate seamless routing functions.

In WMM technique, fixed IP addresses are assigned to Mesh Nodes (MNs) either by static configuration or by DHCP protocol. Each MN maintains both proxy table and routing table in its cache. The Mobile Station (MS) location information is contained in the packet headers and stored in the proxy table. There is no specific routing table maintenance protocol is proposed in WMM technique. In proxy table entry, MN maintains three information for MS: IP address of MS, IP address of Serving MAP (SMAP) of MS and the time when the MS is connected with the SMAP. To facilitate time related information storage and specific channel allocation, time synchronization protocol is required in WMM technique. Network Time Protocol (NTP) is applied here for the purpose of synchronization.

In (Li et al., 2013; Li et al., 2014), a Gateway Scheduling based (GaS) mobility management scheme is proposed. In this scheme, the location update is done by embedding the location update message in the packet header which will minimize the overhead in the network. The main idea in this scheme is to avoid the channel contention between the signalling and data packets. Another routing based mobility management algorithm is proposed in (Jesus et al., 2016), where two handoff thresholds are defined to allow the handoff to be triggered in two different situations. Whenever the signal strength is less than a predefined minimum threshold the handoff should be triggered in this scenario to avoid the bad link. On the other hand, whenever you are connected to an access point with a good and acceptable signal strength (above the minimum threshold) but there is a neighbouring access point with even better signal strength in this scenario the handoff is also can be triggered.

Hybrid Routing and Tunnelling Based Mobility Management Approaches:

Authors in (Li Y. et al., 2012) proposed a mobility management scheme named LMMesh (Location Management Mesh), it only handles the location management issue in mobility management and does not solve the handoff management issue. In this scheme, when a mobile node is moving the location management is handled on a different way using a hybrid routing and tunnelling based approach to exploit the advantages of both approaches. If the mobile node is moving in a small range, the number of hops (K) from the old serving mesh router to the new serving mesh router is not long ($K < 2$). In this case the routing based approach is used and a forwarding pointer is embedded in the routing table of the old mesh router to forward all the received packet to the new mesh router. In this way a lot of signalling overhead is avoided. If the number of hops to reach the new mesh router is more than or equal to 2, the tunnelling approach is used and an explicit location update is sent.

M^3 (Mesh Mobility Management) scheme is introduced by (Huang et al., 2007). The proposed protocol suggests 3 layers hierarchal structure as shown in the figure below. Where the top layer is the gateway, then the Superior Routers (SRs) connected to the gateway and below them are the regular access points. As soon as the mobile mesh clients powers up, a unique IP address is assigned to the mesh client by the gateway after the authentication, authorization and accounting information is done and stored in the gateway and in the mesh client's serving AP, this caching process in the serving AP is done to avoid the repetitive visits to the gateway database for information lookup. When a mesh client is moving and the signal strength is deteriorating, the handoff is triggered and the client sends a handoff request message to the new AP which contains the information of the old AP. This handoff request message is forwarded to the old AP and all the client's information is sent back to the new candidate AP.

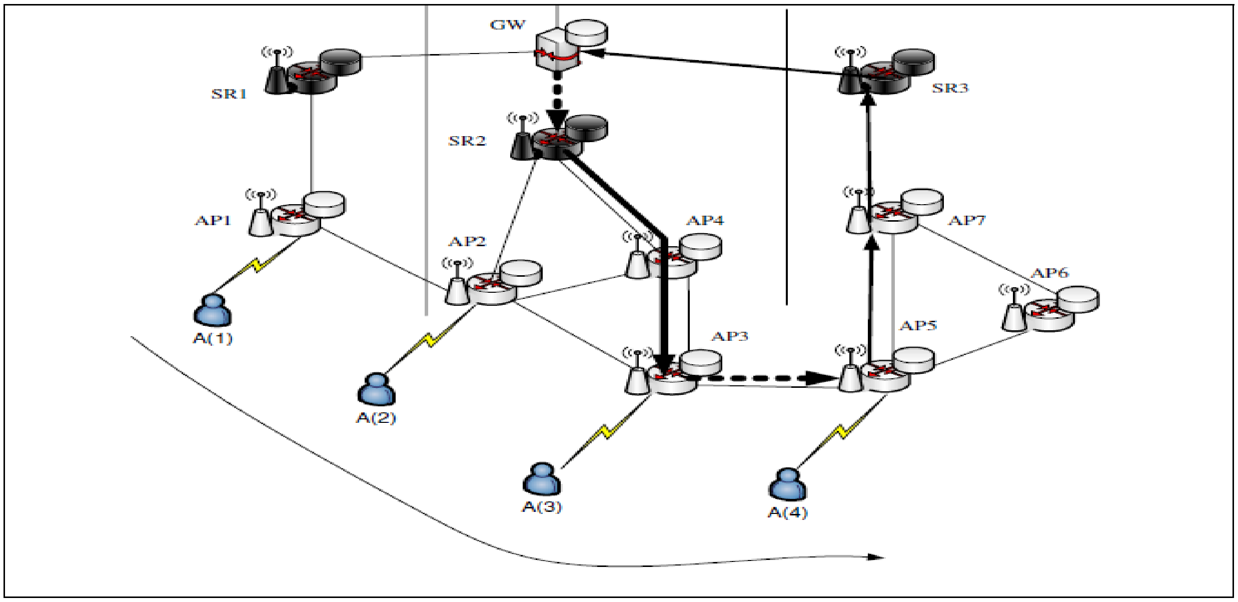

Figure 2.3 Illustration for M^3 Scheme

Taken from Huang et al. (2007)

Now suppose that a mesh client A is moving from position 3 to position 4 as shown in the figure above. if a downstream packet is being routed to the old AP (AP3) then the routing table should contain a forwarding pointer to the new AP (AP5) to reduce the packet loss. While mesh client is moving this chain of pointer is getting longer and longer, which introduced a triangular routing problem (Perkins 1997). To solve this problem, an explicit location update message is sent to the gateway after a specific time interval. In this scheme, and to make this explicit location update more efficient in terms of signalling overhead it will be done by the access point not by the mesh client itself. In this way an AP is sending a set of location update messages for all the newly associated mesh clients every period of time.

The explicit location update is different in the above two mentioned schemes and the comparison of the signalling overhead caused by a hop-based location management scheme (Li Y. et al., 2012) and the time-based location management scheme (Huang et al., 2007) is a good topic to study.

A new 802.11-based handoff management is introduced in (Navda et al., 2005) called iMesh. The proposed iMesh protocol handles the handoff in two layers which are the link layer and network layer. In the link layer handoff, a probe request message is broadcasted by the mobile mesh client and the new channel is selected based on the Signal-to-Interference-Noise-Ratio (SINR). The authentication and re-association is also accomplished in the link layer handoff. In the network layer handoff, route and IP address of the mobile mesh client could be necessary to be changed. In iMesh, the usage of Transparent Mobile IP (TMIP) protocol (Giovanardi et al., 1997) cancel the necessity to change the IP address and allow the mesh client to keep its old IP address in the new location it moves to. The TMIP protocol is standard-compliant Mobile IP protocol, where no modification should be done on the mobile mesh client.

In iMesh, there is a centralized server, namely Mobile Location Register (MLR), holds the responsibility to maintain protocol regarding handoff process. The MLR keeps information about home AP for each of the mobile stations. As soon as a mobile station roams from home AP to a foreign AP, the foreign AP requests a copy of mobile station's information to the MLR. The MLR takes a copy from home AP and forwards it to the foreign AP. Foreign AP, upon receiving required information regarding the visiting mobile station informs home AP with a message handshake and becomes new home of the visiting mobile station. All routing information is updated accordingly (Navda et al., 2005).

Multicast Based Mobility Management Approaches:

In (Zhao et al., 2010; Zhao et al., 2012), handoff delay composed of the handoff delay caused by the data link, network and application layers. The data link layer handoff is concerned about changing the channel that the mobile node is using. The channel switching process is gone through different steps such as channel scanning and selection based on a specific criteria, mutual authentication step to check the identity of the mobile node and the new access point based on a specific shared key authentication algorithm and the last step in the data link layer handoff is the association with the new access point. In the network layer handoff, a new route is found for the mobile mesh client and changing of the IP address is accomplished, if needed. The application layer handoff is to change the multimedia communication SIP (Session Initiation Protocol) session. In (Zhao et al., 2010; Zhao et al., 2012), two approaches are proposed, namely, Xcast and IMeX.

In Xcast protocol, WMN network zone is divided into multiple groups (subnets) and in each group there is a special mesh router named Xcast-based Group Router (XGR) which is responsible on multicasting and storing the data related to handoff and location information in small number of candidate access points, when the mobile node is moving from one subnet to another with the objective of minimizing the packet loss during the handoff process. A greedy algorithm is designed to deal with the problem regarding the optimum number of required special routers and their best placements. IMeX protocol is a cross layer handoff protocol which uses the same architecture used in the Xcast protocol. In the data link layer handoff the channel information for the candidate access points in the other subnets are collected in advance, the same process is followed in both the network and handoff layer where the routing discovery process is done in advance also to change the route and the session of the mobile node in a seamless manner. Routing discovery process is taking the largest percent from the network layer handoff delay. The IMeX protocol can be applied on both the intra and inter gateway handoff, where the IP address has to be changed in case of the inter-gateway handoff (Zhao et al., 2010; Zhao et al., 2012).

A unique Seamless wireless Mesh (SMesh) protocol is proposed by (Amir et al., 2006). The proposed protocol ensure the emergence of a new seamless, fast handoff network that supports time sensitive multimedia applications like VoIP, video conferencing and other real-time traffic from any device running 802.11 protocol. During the handoff process, signal quality between access point and the mesh client is continuously monitored by each access point, this information is shared between the access point and other access points in its range to cooperatively decide which access point is the most suitable one for the mobile mesh client. SMesh architecture composed of two main components which are the communication infrastructure and the interface with the mobile clients. SMesh protocol is using Spines messaging system (Amir et al., 2003). In Spines messaging system uni-cast, multicast and any-cast communication is allowed. Whenever a topology change in WMN is occurring, due to node failure, wireless mesh network expansion by adding new mesh router or due any other reason a link state routing information is shared to maintain an up to date network topology awareness. Each mesh node is associated with a specific group, so that the multicast group can be easily identified for the purposes of location and handoff management. The basic idea behind SMesh lies in the fact that, from the client's end, the entire topology is presented as a single, ubiquitous access point which accomplished by using DHCP (Dynamic Host Configuration Protocol) (Droms 1997) to provide the clients with the connectivity information.

Each client has a unique IP address which is issued using DHCP protocol. This IP address has a lease timer, when this timer is expired the client is no longer connected to any access point therefore a broadcast DHCP renewal request should be sent in a periodical manner to keep the connection up. The link quality information is calculated by each access point by monitoring the loss of the mesh client's DHCP request and the best quality link is chosen. To determine the most suitable access point for a client, the management of clients using multicast groups are accomplished in order to facilitate handoff and location management process.

Handoff Management Approaches:

In this subsection, we will explain the handoff management schemes where the location management are not part of the mobility management solution. Other mobility management schemes are also studied where the location update is done in a completely different way as proposed in (Li et al., 2012).

The process of handoff or transfer of connectivity from one access point to another in wireless mesh networks suffers from latency, which may be as long as several seconds and cause serious problems in time-sensitive traffic such as interactive video conferencing or voice over IP. To facilitate seamless connectivity for time-sensitive applications that are bounded by data transmission delay in Wireless Mesh Networks (WMNs), (Yin et al. (2013)) proposed a Caching-List based fast Handoff mechanism (CLH). The authors have divided the handoff process into three different phases which are: 1) channel probing phase (scanning phase): where the mobile host triggers the handoff when Received Signal Strength Indicator (RSSI) of the current access point reached a predefined threshold value, a scanning process of all the available channels for the mobile host is started by sending probe requests messages to all the neighbouring access points as shown in the figure below. The list of available mesh nodes are stored in the caching list; the list is used for the management of information regarding mesh nodes. The RSSI of a mesh node provides the information regarding its position in the list. The list of mesh nodes (neighbouring access points) is organized in order of the level of signal strength with the highest signal strength at the top; the advantage of such ordering is obvious, it facilitates fast switch-off to the most reliable node.2) Authentication phase: this phase starts whenever the best neighbouring access point is selected, which has the highest signal strength. After this, an authentication request message is sent to the selected access point. In this phase a shared key authentication mechanism is used. 3) Re-association phase: in this phase a re-association request is sent to the selected access point, and an Inter Access Point Protocol (IAPP) message is sent simultaneously with the re-association response message from the new access point to the old access point and mobile node, respectively(Yin et al. (2013)). IAPP is formally known as IEEE 802.11f, which is a protocol runs between the old and new access point to make sure all the information is passed to the right access point (Bob et al. (2005)).

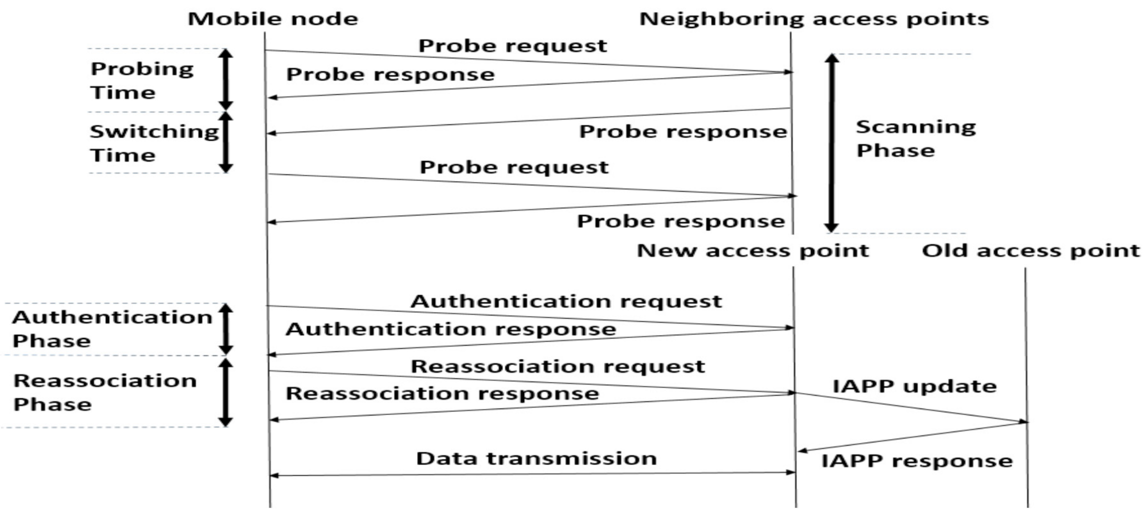

Figure 2.4 Proposed Handoff Scheme

Taken from Yin et al. (2013)

Another technique has been proposed in (Rezgui et al., 2009) that deals with the efficiency in the channel switching and efficient balancing of the load at the time of the process of handoff. To speak more specifically, following two things have been introduced in the paper- first, to handle the contention among neighbouring routers a clique concept is introduced to work in line with WMNs, to avoid neighbouring routers using the same channel at the same time a conflicting links are found and share the same channel using Time Division Multiplexing (TDM), where contending links share the same channel in a time division-multiplexing manner. Second, a dynamic load balancing policy during handoff has been proposed, which is supposed to combine the mechanisms of selection of mesh routers and admission control of the traffic flow.

When the user is moving it is obviously better to keep the same channel while moving, so the process of in-clique handoff gets preference over cross-clique handoff during handoff. A user keeps on using same channel at the time of handoff and after handoff.

For load balancing purpose, a dynamic load balancing based on the path load balancing using routing algorithm is proposed in this protocol. Faster handoff is facilitated by two approaches-MR selection and traffic admission control. A user detects available MR around and tries to connect to most appropriate MR by sending Clique Header (CH), which takes the responsibility to perform admission tests. If the result of the admission test is positive, handoff procedure initiates, otherwise next admission test starts to the next best MR. Two criteria dominates MR selection process: first, signal strength and second, the handoff clique. Generally, the MR with highest signal strength is selected and the MR within the same clique is preferred to keep the same channel and avoid the channel switching delay. On the other hand, the traffic admission control is performed by clique condition-a lazy clique gets preference over a busy clique to mitigate network congestion (Rezgui et al., 2009).

Summary and Conclusion

In the table below, we summarized the most important mobility and handoff management algorithms proposed in the literature. As noted in the table, none of the algorithms proposed in the literature are comparing the effect of using maximally jointed route with the old route for the mobile client. All the algorithms are finding a completely disjoint route with the old route. A maximally jointed route will enhance the QoS provided for the end users. In this thesis, we proposed a new mobility management scheme which finds the maximally jointed path and trigger the multi homing whenever it is necessary, which has never been done in the literature.

Table 2.1 Mobility Management Algorithms Comparison

Approach	**Parameters considered**										
	NHT	IntraDH	InterDH	JR	DR	MHF	IPAC	**Location Management**			
								RB	TB	MB	RTB
Jesus et al. 2016	2	✓			✓		No	✓			
Malikarjuna et al. 2016	1	✓			✓		No	✓			
Chaya et al. 2016	1	✓			✓		No				

Qin et al. 2015	1	✓			✓		No				
Mamidi et al. 2015	1	✓			✓		No	✓			
MajMuder et al. 2014	1	✓			✓		No	✓			
Khasawneh et al. 2014	1	✓		✓		✓	No	✓			
Yin et al. 2013	1	✓			✓		No				
Li et al. 2014	1		✓		✓		Yes	✓			
MajMuder et al. 2013	1	✓			✓		No		✓		
Zhao et al. 2012	1	✓	✓		✓		Yes/No			✓	
Li Y. et al. 2012	1										✓
Rezgui et al. 2009	1	✓	✓		✓		Yes/No				
Huang et al. 2008	1	✓			✓		No	✓			
Huang et al. 2007	1	✓			✓		No				✓
Ren et al. 2007	1	✓	✓		✓		No	✓			
Wang et al. 2007	1	✓			✓		No		✓		
Amir et al. 2006	1	✓			✓		No			✓	
Navda et al. 2005	1	✓			✓		No				✓
Zhang et al. 2010	1	✓	✓		✓		Yes/No	✓			

Other mobility management algorithms have been proposed in the literature such as (Chaya et al., 2016; Mallikarjuna et al., 2016; Mamidi et al., 2015; Qin et al.,2015; Hung et al.,2009), where the traffic for the handoff mobile client goes through another path which is completely disjoint with the older reserved path. This is opposite from our proposed algorithm as it will trigger the multi-homing feature, routing traffic into two maximally jointed paths, increasing the admission rate. For the traffic of the handoff mobile client in a congested network, a multi-homing feature locates maximally jointed paths, which are used to decrease the hard handoff ratio and decrease the handoff delay. This will be shown in the results in the next chapter.

Joint and disjoint routing algorithms are compared in (Khasawneh et al., 2014). In highly congested networks, joint routing alone is a poor solution. Our routing algorithm decision triggers multi-homing, (Khasawneh et al., 2015), whenever needed, where the traffic is routed through two different access points in order to utilize the limited resources in each link and increase the user admission rate and satisfaction. To verify our results and obtain a near optimal solution for our proposed algorithm we introduce a novel Markov decision process model, with its defined rewards, discount factor, transition probabilities.

Also our proposed mobility management algorithm is a routing based mobility management, so the location update messages are embedded in the packet headers. This is the most suitable solution because we are supposing that the network is highly congested and we do not want to increase the congestion with those signalling messages.

2.3.2 Congestion Control

The congestion control algorithms proposed in the literature can be categorized using different parameters. Some of the congestion control algorithms are dealing with a congestion in a reactive manner, where the congestion is dealt with after it happens. Other algorithms are proactively trying to handle the congestion before it really happens in the network in order to mitigate its negative effect. Each approach has its own advantages and disadvantages, some algorithms trying to handle the congestion in a hybrid manner by combining the reactive and proactive approach together to benefit from the advantages of both approaches. Another categorization parameter is based on the provision of the feedback in order to make the congestion control decision, if there is a feedback given on the level of congestion in the network this can be called a closed loop solution, otherwise it is called an open loop solution. This feedback can be done in an explicit way by sending an explicit messages about the congestion level, either locally or globally (end-to-end). An implicit feedback can be extracted from the level of delay in acknowledgment messages or the level of packet loss. In wired network, the main cause of the delay and packet loss is the congestion in the network so an implicit feedback could be suitable in such situations. Whereas in wireless network, the congestion is not the only cause of the delay and packet loss (Islam et al., 2011) so the explicit feedback is obviously better and more accurate compared with an implicit feedback approach.

Unlike wired networks, Multi-Hop Wireless Networks (MHWN) including Wireless Ad-Hoc Network(WAHN) and Wireless Mesh Networks (WMN) mostly share data with neighbouring local users as destinations, which is called a user locality property (Shifrin et al., 2010). With this assumption in (Shifrin et al., 2010), the TCP congestion control technique does not work well in this specific situation since it concerns about the end-to-end congestion signalling, which is obviously not the case in this algorithm. In the proposed algorithm, called Collective Congestion Control (CCC or C3), the multi-hop topology is divided into several groups and subgroups depending on their locality property. Each group is assigned with two Collaborate Token Bucket (CTB) for transmitting and receiving. The CTBs are managed and controlled by hierarchical Collaborative Token Bucket (HCTB) technique. To control inside group token generation rate, a unique User Group Representative Subgroup (UGRS) is selected for each groups whose function is to acquire the knowledge related to the buffer queues of the local buckets and distribute the knowledge to corresponding group members. At higher level UGRS also collects information about neighbouring groups. The reception and transmission rate are then adjusted based on the received information from the selected UGRS node (Shifrin et al., 2010).

Another algorithm is proposed by (Nascimento et al., 2008), which is called Cross Layer Mesh-Transmission Control Protocol (CLM-TCP). As the name implies, this algorithm is collecting information from different layers in order to make its congestion control decision and for the improvement of TCP performances in wireless mesh networks. The protocol is designed to collect information from layers below it, adjusts parameters for better performance and pass the information to the layers higher to it for necessary actions. For further improvement, this protocol incorporates Minimum Delay (MD) and Expected Transmission Count(ETX) metrics.

In multi-hop wireless channel link layer contention yields conflict problems due to exposed terminal and hidden terminal effects. A room of improvement still exists even though multiple multi-channel schemes are proposed in the literature (Ren et al., 2008; Draves et al., 2004), for better resource utilization in the network. Poor performance in multi-channel Wireless Mesh Networks resulting from congestion may be greatly reduced by Multi-Channel Adaptive Pacing (MCAP) scheme. This scheme utilizes a cross-layer congestion control scheme. This scheme is potentially immune to congestion as it coordinates and transmits packets in-group during high congestion situations and thereby improve multi-channel utilization (Ren et al., 2009).

Two proposed techniques are supposed to determine the protocol. Firstly, rate adaptive pacing where routers act like access controllers to control the flow rate based on the bandwidth information and congestion measurement. These routers are known as ingress routers. On the other hand, another technique is implemented at routers, which coordinates packet transmission locally to improve channel utilization (Ren et al., 2009). These routers are known as congested routers. The protocol suggests that the congested router should transmit packets for a certain period taking into account the channel reuse constraints and then remain silent for a period, known as forbidden period, to allow transmitted packets travel further away. WCETT routing metric (Draves et al., 2004) is incorporated with Adhoc On demand Distance Vector (AODV) protocol as multichannel routing protocol (Ren et al., 2009).

A unique, neural network based technique is proposed in (Islam et al., 2011) to create a new variant of TCP to address the problems associated with wireless mesh network and to ensure reliable performance. The new variant of TCP is named as intelligent TCP or iTCP. iTCP supposed to ensure better congestion control technique, enhanced performance and lower energy consumption per bit as compared to other TCP variance. As previously mentioned, the reason of packet loss and long delays in wireless mesh network is not necessarily caused by the congestion. The main idea in iTCP is the novelty of the congestion detection mechanism, which consider the consecutive data transmission failure caused by congestion. So, if a consecutive data transmission failures are experienced the flow rate is adjusted accordingly. Single data transmission failure might be resulted from the congestion but most probably it is resulted from any other cause if it is not followed by another failure (Islam et al., 2011).

Inputs to the artificial intelligence neural network proposed scheme are a number of consecutive timeouts, number of duplicate acknowledgements and existing congestion window size to compute next optimal congestion window size. The next window size may increase, decrease, or remain the same as compared to the current window size. Therefore, the technique uses neural network to compute the optimal window size based on previous window size (Islam et al., 2011).

An Additive-Increase/Multiplicative Decrease (AIMD) based rate-control protocol namely Wireless Control Protocol (WCP) is proposed in (Rangwala et al., 2011) to deal with the unfairness and starvation of conventional TCP over wireless multi-hop networks. The proposed protocol is designed with the main idea that congestion is a neighbourhood phenomenon, not a node-local phenomenon; it appropriately responds to such phenomenon. In order to take a good congestion control decision, the level of congestion should be estimated accurately, this can be done by the other scheme proposed in (Rangwala et al., 2011) called Wireless Control Protocol with Capacity estimation (WCPCap) that constantly measures the neighbourhood capacity available and distributes this capacity to the competing traffics; the scheme employs a distributed rate controller for better congestion control decision.

The WCP protocol incorporates 802.11 MAC protocol. The WCP protocol first observes and assesses the nature of congestion in Multi-hop Wireless Networks. It then shares the information regarding the congestion to all the adjacent and connected links. Finally, AIMD rate adaptation algorithm decides a comprehensive rate for transmission maintaining fair policy. WCPCap, a variant to WCP explicitly sends feedback to sources regarding the congestion information in the neighbourhood facilitates congestion estimation and thereby makes it possible to implement the protocol in distributed networks (Rangwala et al., 2011).

The eXplicit Control Protocol (XCP) (Dina et al., 2002) and the Rate Control Protocol(RCP) (Nandita et al., 2006) are two of the recent lighter and faster protocol to address the issues of the packet loss in a wired network. These protocols have been modified to work in wireless environment, since it has been shown in (Barreto et al., 2010) that the performance of those two algorithms is bad in wireless mesh networks. They fail to predict the link bandwidth properly in wireless environment. (Barreto et al., 2011) proposes new protocols based on XCP and RCP to efficiently control flows in wireless networks called XCP-Winf and RCP-Winf. A cross layer approach is implemented to determine the link bandwidth and capacity of the path in wireless networks effectively. The method of determination of these variables are based on the Carrier Sense Multiple Access-Collision Avoidance (CSMA-CA) scheme by checking the Network Allocation Vector (NAV) to calculate the idle and bust time and therefore calculates the available bandwidth.

XCP-Winf uses router information to determine congestion level status. This necessitates changes and modifications at routers and all end-devices. The proposed XCP-Winf network is composed of a sender host, a receiver host and intermediate nodes. The intermediate nodes are used to facilitate the queuing service for the incoming traffic from the sender to the receiver. The protocol uses feedback mechanism to inform the sender with information regarding network conditions including maximum goodput. Each packet includes a header that contains congestion information. The task of intermediate nodes is to update the congestion information in the header of the packet according to network conditions. The receiver copies that information and send them in outbound flows. The other protocol, Rate Control Protocol (RCP-Winf) is similar to XCP-Winf except for the fact that it deploys different approaches regarding the header information update. It updates the information regarding the capacity and end-system operating rate in the packet headers (Barreto et al., 2011).

One of the important developed congestion control mechanisms that are designed to exploit the advantages of Wireless Mesh Networks (WMN) is the Intra-Mesh Congestion Control (IMCC) mechanism. (Staehle et al., 2012) evaluates and compares three IEE 802.11s compatible variants of IMCC algorithms, namely Total Congestion Control (TCC) and Link Specific Congestion Control (LSCC), which are both introduced in (Desheng et al., 2010) and the newly proposed algorithm named Path Specific Congestion Control (PSCC) which is introduced in (Staehle et al., 2012).

The intra-mesh congestion control protocol is designed on the basic idea of decreasing intra-mesh packets. Nodes at the edge of the network has a higher possibility than other nodes in the center of the network in allocating idle channels and transmitting data since edge nodes are not like the center nodes which are responsible on forwarding the packets of other nodes and act as a router most of the time rather than being a transmission source. In case of the congestion, the overall network throughput will be dropped and a lot of time and resources will be wasted by the center nodes of forwarding packets to the other nodes. So the main objective in IMCC scheme is to decrease the sending of intra mesh packets (Staehle et al., 2012).

Each variant of IMCC protocol is designed by three basic building blocks: congestion detection and monitoring, congestion control signalling and local rate control. The congestion control signalling protocol detects congestion and sends a CCNF (Congestion Control Notification Frame) to adjacent mesh nodes. Each frame contains one or more CNE

(Congestion Notification Element) that carries the congestion notification duration timers, the station address field for destination mesh to determine where the local rate control have to be applied. The station address field could be the broadcast address, if this is the case the local rate is applied to all the flows (Staehle et al., 2012). The main difference between TCC, LSCC and PSCC is where to apply the congestion control which will be at the node level in TCC, link level in LSCC and path level in PSCC algorithm. Results show that PSCC is the best variant in term of traffic fairness in the network.

(Ma et al., 2013) investigates an optimization technique for the performance of wireless mesh networks by incorporating game theory optimization model with cross layer design. The resulting technique is Cross Layer Game Optimization. The technique is designed with an objective to limit congestion in wireless mesh networks. The proposed algorithm is designed to set the source node in charge of changing the path-sets and adjusting the congestion window size in accordance with the round-trip time to achieve a Nash equilibrium. Game theory helps in making complex decision regarding the selection of the best-suited path for transmission.

In this technique, TCP protocol is used in connection with cross layer technique and game theory. The randomness of wireless networks are taken care of with game theory, which helps the decision-making. Flow is divided into multiple paths and multipath TCP is employed to facilitate flow control and congestion detection. The main problem with multipath routing is associated with extra overhead. The load balancing is considered when selecting paths in multipath routing. Game theory in collaboration with cross layer information helps in the decision making process regarding the best suitable path that yields optimum load balancing and congestion control (Ma et al., 2013).

Unlike TCP that does not account for the neighbourhood link interdependency, the proposed NICC (NeIghbourhood-based Congestion Control) is a neighbourhood-based congestion control protocol (Masri et al., 2012; Masri et al., 2014), which is free from extra overhead and resolves the starvation issues without degrading network performance. The technique is designed to exploit some underexploited header in IEEE 802.11 frames to employ them in implicit multi-bit congestion feedback.

NICC is based on two objectives: fairness requirements and efficiency of bandwidth allocation. In wireless network a path or link's capacity or bandwidth is shared by neighbouring links and hence when a link is congested, its neighbour faces congestion. The proposed NICC protocol meets three set of conditions to clear a congestion issue in a link: first, it clearly identifies interference set, i.e. the set of adjacent links sharing common bandwidth, second, it sends a congestion signal to all members of interference set and finally, it sends a congestion notification message to the sources of data traversing the interference source to adjust their flow (Masri et al., 2012; Masri et al., 2014).

The NICC frames header as shown in Figure 2.5 comprises a 2 bytes field known as the packet control. This field has 2 bits to identify the type of frame and the 4 bits to determine the subtype structure. There are only three bits combinations of this type that are used: 01 for control frames, 00 for management frames, and 0 for data frames. The last combination of bits is not used (El Masri et al., 2014; El Masri et al., 2012).

Therefore, if we can make use of the combination of the bit then 11 of the type field and the 4 bits of the subtype field we would be able to obtain up to 16 new different IEEE 802.11 frames. NICC needs each node to maintain a three congestion tables (El Masri et al., 2014; El Masri et al., 2012). This entails the nodes table, links table, and the flows table. The links table is used for the purpose of keeping the list of congested outgoing or incoming links of the node. Nodes' table is used to uphold the list of the neighbouring nodes of node N. Lastly the flow tables are practically used to support the list of flows traversing node N.

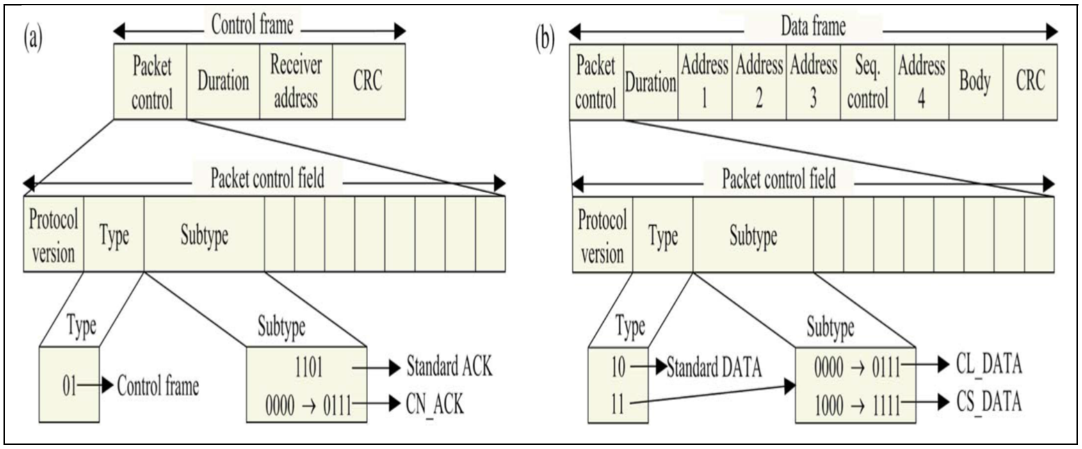

Figure 2.5 NICC Frames Control Frame and the Data
Taken fromEl Masri et al. (2014)

The acknowledgment frame (CN_ACK frames) in this design is applied to acquaint the upstream node that the consistent flow is crossing a congested neighbourhood. The new control frames embody a particular congestion level. The acknowledgment frame comprises the extent of the most congested neighbourhood across the flow path.

The data frames use the Congestion Signalling DATA(CS_ DATA), or Congested Link DATA (CL_DATA) frames to transmit data through node S. The data frames are also used to indicate congestion in a congested neighbourhood. It generates a standard DATA frame when it has no outgoing or incoming congested links. There are eight different CS_DATA and CL_DATA frames. The function of a CL_DATA frame reports the level of congestion of the outgoing congested link which is used to conduct data. On the other hand, the CS_DATA frames reports the highest existing congestion level of its outgoing or incoming congested links (El Masri et al., 2014; El Masri et al., 2012).

The congestion control process is made up of four phases: congestion signalling phase, congestion detection phase, congestion notification phase, and the rate of scrutiny period. In the congestion detection step, the node detects that one of its primary outgoing links is evenly congested. The second phase is the congestion signalling phase where congestion is signalled to every link in the neighbourhood of the full link that is either the receiver or the sender. The third step is the congestion notification where the sources of the flows that traverse congested localities are thereby notified about the existing congestion. The control rate phase is the last phase where the sources adjust the speed of sending their flows about the reported congestion levels (El Masri et al., 2014; El Masri et al., 2012).

Researchers in (El Masri et al., 2011) suggests a QoS technique for wireless mesh networks namely WiRS that supports real time data flow over wireless mesh networks. The protocol contains one admission control and two traffic regulation schemes. The admission control scheme exploits the unused resources from flows and assigns them to contending flows. On the other hand, the traffic regulation scheme regulates traffic for congestion avoidance and QoS maintenance. In short, the proposed scheme takes two-fold initiative to improve wireless mesh network performance-makes use of unused reserved bandwidth and regulates the flow for congestion control.

The proposed technique is based on the exploitation of reserved bandwidth at intermediate nodes along a traffic flow path and the regulation of network traffic. In the proposed protocol, the available network resources are measured by periodic UDP control packet transmission, where the requested bandwidth is embedded in this control packet, while the packet is moving from one intermediate node to another in the path from source to destination it checks the available bandwidth at each node and update the value stored in the control packet only if the available bandwidth is less than the requested bandwidth. Whenever the control packet reaches the destination the lowest available bandwidth that can be offered through the path is known and the source node is notified with this value. (Masri et al.,, 2011) also proposed two regulation schemes which changes the transmission rate based on the network conditions. The first regulation scheme is to decrease or increase the transmission rate based on the congestion status. The other regulation scheme is service oriented, in which some traffic such as video traffic will have a higher priority than the data traffic, so when there is a congestion in the network the transmission rate for the traffic with lower priority is decreased. (Masri et al., 2011).

Authors in (Qiu et al., 2014), proposed an algorithm called RObust joint Congestion control and Scheduling algorithm (ROCS) which is used to solve the joint congestion control and scheduling in time changing wireless networks.

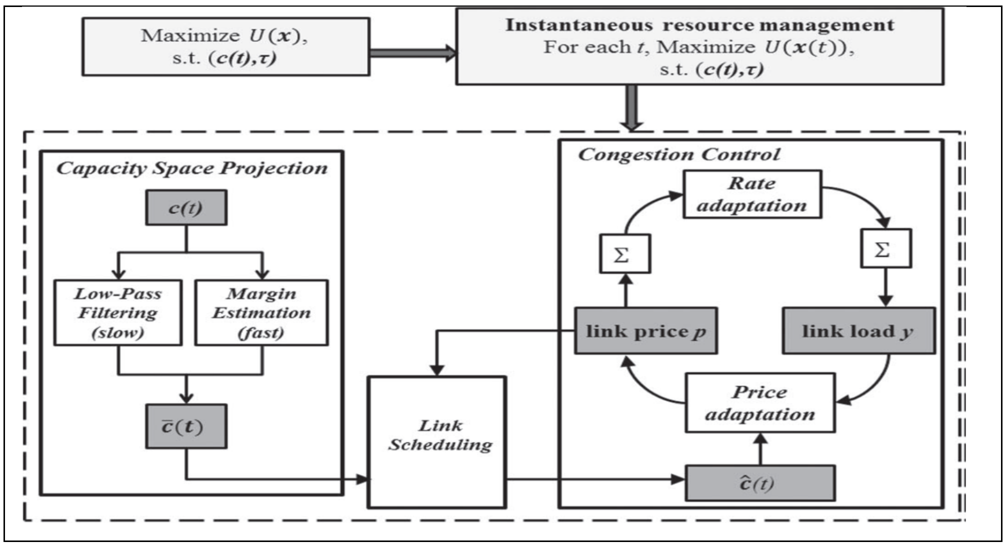

Figure 2.6 A Summary of ROCS Approach
Taken from Qiu et al. (2014)

The figure above presents an outline of the approach which comprises of three main components: link scheduling, capacity space projection, and congestion control. The design adapts to the time changing capacity space by maximizing the resource apportionment at each time instance. The idea of the design is a time decomposition to reduce the negative influence of delay and prevent resource over-provisioning (Qiu et al., 2014). The main idea of this algorithm solution is the Capacity Space Projection Algorithm (CSPA). The perturbation of the signal in the capacity link is linked with the fast time scale which can be harder to track. The implementation of CSPA is focused on the discussions of the slow time scale capacity margin estimation and capacity extraction. Each link autonomously executes the CSPA at a particular runtime and creates a virtual function.

IEEE 802.11s is providing its own way for frame retransmission via Automatic Repeat Request (ARQ). This also done by TCP protocol, which makes some kind of redundancy between layers so the protocol should be adjusted for better overall performance. In (Rethfeldt et al., 2016), cross layer approach is proposed in order to enhance the MAC layer ARQ when applying TCP over 802.11s.

Researchers in (Song et al., 2006) describe a congestion control routing algorithm, Least Congestion QoS-aware Routing (LCQSR). This algorithm eases global congestion by locating the best destination path. The algorithm is based on a metric called Degree of Congestion Control (DCC), exploring and exploiting the dynamic nature of two MAC layer statistical parameters: Frame Transmission Efficiency (FTE) and available Residual Bandwidth. The algorithm uses a cross layer design to balance routing performance and reliability of the system. There are disadvantages to this algorithm. Its performance deteriorates with dynamicity of the network. As well, in a congested network, where DCC is very high, the new call will be rejected, bandwidth being unavailable. Call rejection probability is reduced by our proposed algorithm, even in a congested network.

In the case of high congestion in the network, higher priority traffic is processed before the lower priority traffic which might cause an increased starvation for the lower priority traffic, authors in (Sheikh et al., 2015) proposed an algorithm that handles the fairness issue and the performance degradation. Researchers in (Komatsu et al., 2013) tried to mitigate the problem of buffer overflow in the access points.

IEEE 802.11s works on the data link layer of the OSI model and deals with the MAC address. Two route selection algorithms are being used in this standard, which are Radio-Aware Optimized Link State routing (RA-OLSR) protocol and the Hybrid Wireless Mesh Protocol (HWMP). This standard does not have the limitations of the previous standard. The default routing protocol used in this standard is HWMP, HWMP is a hybrid routing protocol which combines both reactive and proactive features in it, but it has some disadvantages. The most important one is the inefficiency when it comes to deal with congestion (Yang et al., 2009; Bari et al., 2012). An improvement to enhance the congestion response feature is introduced in (Khasawneh et al., 2015), the Enhanced Congestion Avoidance - HWMP

(ECA-HWMP) predicts the congestion before it really happens in the network. Another algorithm we have proposed which is called the Rate Adaptive Enhanced Congestion Avoidance - HWMP (RAECA-HWMP), which takes the output from the Variable Order Markov (VOM) prediction model and feed it to the routing algorithm to find the new route and adapt the transmission rate based on the congestion status.

In this thesis, a novel VOM prediction model is introduced to predict the congestion in the network. The output of this prediction model act as an input to find a new route for the packets to avoid delay caused by the potential congestion in the network and the transmission rate is also maximized in order to get better performance by the end user. The congestion status is stored in a database by sending explicit congestion notification message.

2.3.3 Bandwidth Estimation

The necessity of an efficient bandwidth estimation has urged authors in (Oriol et al., 2012) to classify the existing estimation algorithms in the literature. Under or over estimation of the available bandwidth could impose taking wrong decisions in admission control, congestion control or mobility management algorithms, which negatively affect the QoS. Bandwidth estimation algorithms can be categorized into passive (non-intrusive) or active (measurement based) approaches (Oriol et al., 2012). In passive approaches, local information is exchanged between neighbours in the network which might be embedded in the control messages of the used routing protocol. In the literature, two design options are used which are the cross layer design and the analytical modelling.

On the other hand, in active approaches end-to-end bandwidth estimation is performed by sending probe packets on a path using different rates and monitoring the packet inter-arrival times. The authors of (Gupta et al., 2009), has shown that the passive approaches have more accuracy compared with the active approaches. Active approach has also another disadvantage of the control overhead that should be incurred. However, passive approaches is more complex than the active approaches due to the modification that might be necessary to be done to obtain those local information by changing on the standard protocols. In the figure below are some schemes that have been proposed by researchers for the bandwidth estimation purposes for both active and passive approaches.

In (Gupta et al., 2009), an experimental comparison is implemented between active and passive approaches. For the passive approach, each node is listening to check if there is any packet is transmitted and total transmission time for a packet is calculated, therefore channel utilization can be easily calculated and the available bandwidth is determined by the following formula:

$$\text{Available-BW} = (1-\mu)*C \tag{2.1}$$

Where μ is the channel utilization and C is the capacity of the channel.

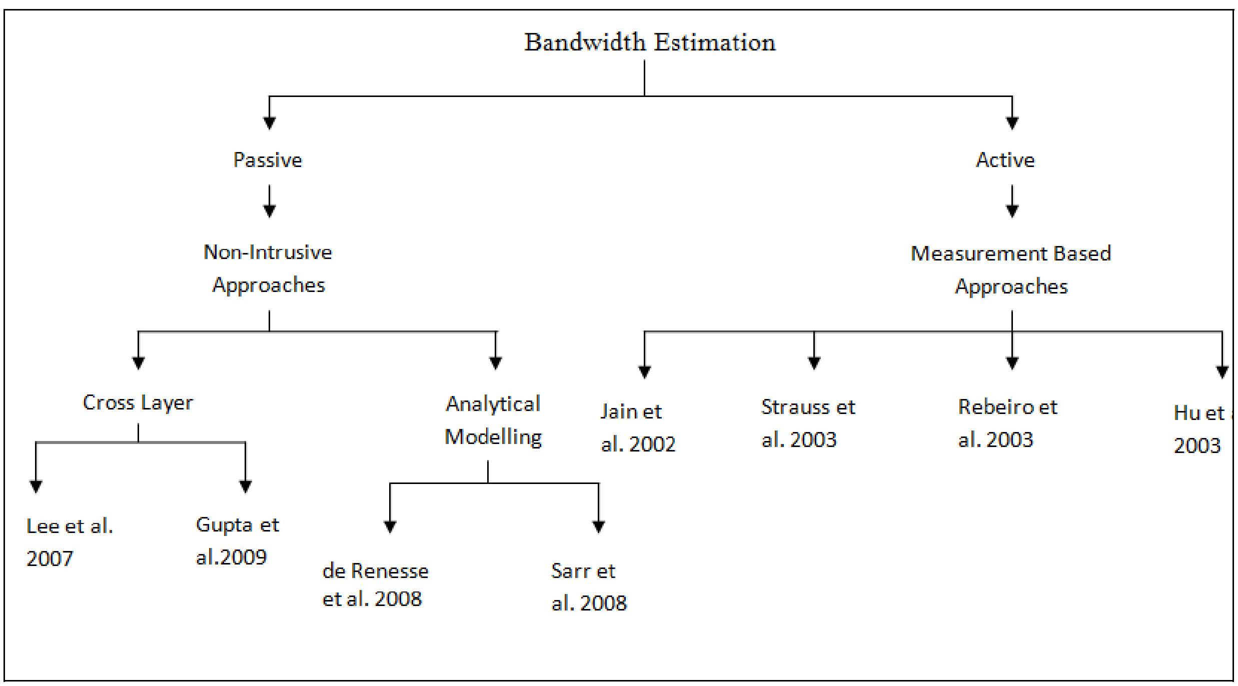

Figure 2.7 Classification of Bandwidth Estimation Algorithms

In this thesis, passive approach is used to estimate the available bandwidth. The estimated available bandwidth is calculated by using the formula above.

CHAPTER 3

MOBILITY MANAGEMENT USING JOINT ROUTING AND ADMISSION CONTROL

3.1 Introduction

Wireless Mesh Network (WMN) is becoming more and more popular day by day. However, it lacks development and robustness against adverse network conditions and high traffic demanding situations. One of the situations arises when real-time traffic enters wireless mesh network and over-consume bandwidth resources causing potential degradation in Quality of Service (QoS) of the network. To solve this issue, several researches have been conducted that are focused on advanced admission control techniques that control high bandwidth demanding applications to enter the network (De renesse et al., 2008; Dhurandher et al., 2015).

Admission Control (AC) is the key management function in Wireless Mesh Network, in order to satisfy Quality of Service (QoS) requirements, especially for multimedia applications. Without an effective admission control scheme in place, the network performance and QoS of Wireless Mesh Network (WMN) deteriorates over time because of uncontrolled admission of real-time traffic that is highly likely to consume large bandwidth and create network congestion (Rezgui et al., 2009).The main purpose of admission control is to make sure that a new flow can enter a network only if the QoS service requirements of existing flows can be met without any loss after that new flow begins.

The purpose of admission control management function is to provide uninterrupted and quality of service to all clients by restricting real-time applications from over-consuming network resources and making bandwidth unavailable for other services. The problem with early admission controls employed in multi-rate WMNs was that these schemes were designed to use a predefined and static rate, which lacks actual information regarding the network condition. On the other hand, few algorithms were designed with dependency on Media Access Control (MAC) layer for rate adaptation. Both these algorithms are inefficient because they cannot exploit the network condition efficiently and hence suffer from performance as well as QoS degradation.

The proliferation and widespread application of Wireless Mesh Network (WMN) necessitate the network to support large varieties of applications to facilitate one-stop solution to all networking needs. Voice Over Internet Protocol (VoIP), multimedia video and content sharing are few applications that require exclusive platform in WMN. One of the recent developments in WMN to support these sophisticated applications of wide variety is the service-oriented design, which unlike conventional method, seeks for the specification of the service user needs to use rather than destination address (Akyildiz et al., 2005).

Wireless Mesh Network (WMN) has an unpredictable medium and frequently changing networking topology characteristics. These characteristics make routing design a challenging job for the designer (Song et al., 2006). In our thesis, we proposed a joint admission and routing protocol that outperforms other methods proposed in the literature. Our proposed algorithm concerns about the handoff and new traffic coming into the network, giving the handoff traffic higher priority and trying to minimize the handoff delay as much as possible for smooth call experience for the end user.

3.2 Our Proposed Mobility Management Algorithm

When the mesh client is associated with an any access point, QoS profile information is shared and the session is established and resources are reserved to satisfy the QoS requirement for this session. Session can be established using any signalling protocol such as Session Initiation Protocol (SIP) (Rosenberg et al., 2002). As an example, when a session is established between terminal T1 and gateway (node 7) in Figure 3.1, a route ($1 \rightarrow 5 \rightarrow 6 \rightarrow 7$) will be reserved for this session only. Our handoff management scheme is based on selecting a new route for the Mobile Node (MN) before it moves and becomes disconnected from node 1. In order to reduce the handoff delay, this new route should be maximally jointed with the primary route.

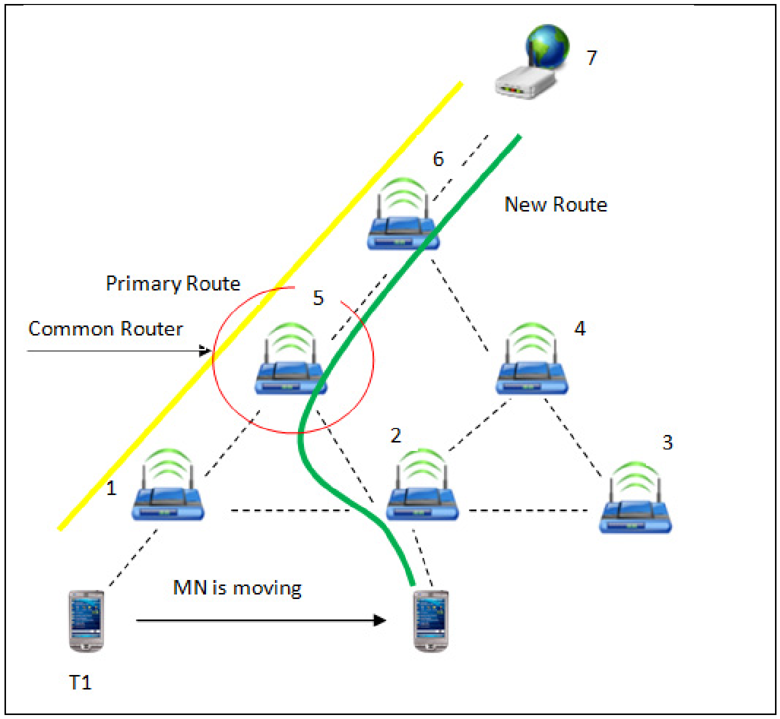

Figure 3.1 New Route Handoff Decision

Taken from Khasawneh et al. (2014)

Three timing thresholds are introduced in our algorithm. Whenever the mobile node is moving and the Signal to Noise Ratio (SNR) reached a predefined threshold. The handoff preparation process is triggered by scanning and collecting information about the channels and the access points around the mobile node. After finishing this step, our proposed algorithm is trying to find the maximally jointed route with the old primary route. As shown in the figure below the new route is found which is maximally jointed with the primary old route. Location update messages are embedded in the routing packet header to avoid making an overhead in the network, this is done at the third time threshold.

As previously mentioned, at the first time threshold, and when the mobile client has an afforded throughput less than the minimum throughput acceptable for the mobile client the handoff preparation process is triggered. In order to calculate the throughput Shannon's formula is used as indicated below:

$$THR_{AP} = W_{AP} \log_2(1 + SNR_{AP}) \tag{3.1}$$

Where the THR_{AP} is the maximum theoretical data rate for WLAN network, W_{AP} is the bandwidth for the WLAN link and SNR_{AP} is the Signal-to-Noise Ratio, which can be calculated as described below:

$$SNR_{APj,i} = \frac{G_{APj,i} P_{APj,i}}{P_B} \tag{3.2}$$

Where $G_{APj,i}$ is the channel gain between terminal i and the AP_j ; $P_{APj,i}$ is the power transmitted from the AP_j and terminal i; and P_B is the power of the noise at the terminal level.

As the radio wave is propagating over a distance, there is a path loss resulted which can be calculated by the following formula:

$$Pathloss(dB) = 20\log_{10}(d) + 20\log_{10}(f) + 32.44 - G_{tx} - G_{rx} \tag{3.3}$$

Where d is the distance between the sender and the receiver in meters, f is the signal frequency in MHz, G_{tx} and G_{rx} are the gains of the transmitter and receiver, respectively.

The newly selected route have to meet the QoS requirement for the mobile node session. When finding the maximally jointed path with the old primary path. This will save us time finding the complete route to the gateway which has to guarantee the same level of QoS and sometimes hard to find in case of highly congested network. The detailed mobility management scheme is given in the algorithm below.

Algorithm 3.1 Handoff Management Scheme

if $THR_i < THR_{i\,,\,thresold}$ **then**
    Trigger handoff preparation scanning
process ***at time threshold 1.***
    Find all APs in MN transmission range.
    Calculate the SNR for each AP and **store**
them in SNR_{APs} array.
end if

At time threshold 2
    for all AP(j)
    if $SNR_{APs}(j) >= SNR_{Acceptable}$ **then**
            Add AP(j) to the $AP_{Candidate}$ array.
    end if
    end for

    for all $AP_{Candidate}(k)$
    Find shortest path to the existing tree.
    Add P_k to the $P_{Candidate}$ array.
    end for

    for all $P_{Candidate}(t)$
        Find the number of common nodes
        between P_{prim} and $P_{Candidate}(t)$.
    end for

    Choose the optimal path with the maximum
    number of common nodes.

    if multiple paths are found with the same
        number of hops the existing tree **then**
        choose the one with the shortest path to
        the gateway.
    end if

At time threshold 3
    Perform handoff process where a new route
    satisfies QoS requirements for the session.
    Send Link layer association message from
    MN to the new AP.
    Update all the routing tables up to the
    common router implicitly by embedding the location update message in the routing

42.	packet.
43.	**Send** Link layer de-association message from MN to the old AP.

Our handoff management sequence diagram is shown in the figure below. The Mobile Node (MN) has an established connection and a reserved resources through its route up to the destination gateway. This connection is passing through the old Access Router (AR). Whenever the predefined threshold is reached, a route discovery process is triggered based on our proposed algorithm, which will find the maximally jointed route with the old route and trigger multi-homing is needed. After the route is discovered, a link layer association is performed with the new AR selected as shown in the figure below.

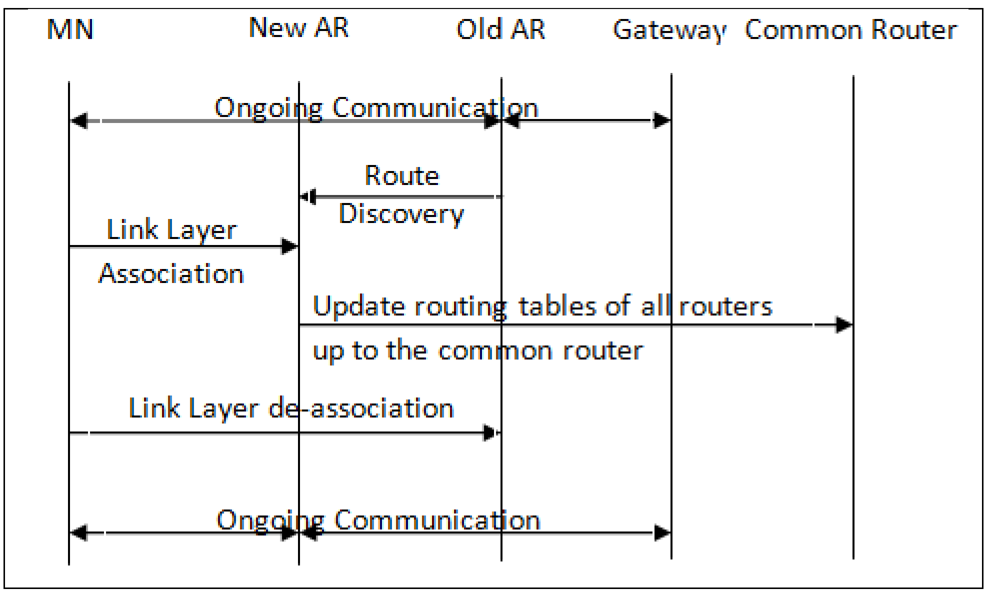

Figure 3.2 Handoff Management Sequence Diagram

Taken from Khasawneh et al. (2014)

The update process is taken place to update all the routing tables up to the first common router, which is the first router which is existing in both, the old and new selected route. All the uncommon routers in the old route should release the resources assigned to the user since the user traffic now is passing through different route, which is called the link layer de-association process. When this process is done all the upcoming traffic are going to be transferred through the new route. In some specific situations, where the congestion is at a high level, the mobile node does not find an access point to be associated with. In this case, multi-homing is triggered in order to avoid the hard handoff scenario. The multi-homing triggering conditions will be explained in more details in section 3.5 below.

3.3 Problem Formulation and Objectives

Problem formulation and the optimization model are presented in this section. Our proposed algorithm will be useful in a highly congested network. Let $G = (N,L)$ be an undirected graph. Where N is the set of nodes in the network and L is the set of links in the network. To each vertex, let the available bandwidth in each link, b_{al}, be $(1- \mu)*C$, where μ is the channel utilization as previously mentioned in Eq. 2.1. As previously mentioned, our proposed algorithm is more beneficial in the case of a highly congested network. The main objective of our algorithm is to increase the admission rate of the new and handoff calls by triggering joint routing and multi-homing whenever it is needed. In order to make sure that the joint routing and multi-homing are triggered at the right time, the following constraint is defined which guarantees that our algorithm is triggered when a handoff or a new call arrives into the network, with the minimum required bandwidth for user k, which can be denoted by b_{rk} , is not available. The first constraint in our problem that force us to trigger the multi-homing and associate the user with two access points is introduced below:

$$b_{al} \leq b_{rk} \tag{3.4}$$

Over a time T, our objective function is to increase the number of users accepted by the entire network. To increase the number of accepted users, multi-homing is triggered to accept the call via two or more access points. A single user is allowed to connect to only two access points as this is a simple solution that decreases the reordering complexity at the destination. For a new, or a handoff traffic, a reward function should be defined to distinguish between different paths that can be selected. In our design, when a new or handoff call arrives six route options can be selected for the new path:

- Arrival of a new user and routed its traffic to one path (a1).
- Arrival of a new user and routed its traffic to two disjoint paths (a2).
- Arrival of a new user and routed its traffic to two joint paths (a3).
- Arrival of a handoff user and routed its traffic to only one path which is maximally disjointed with the old primary path (a4).
- Arrival of a handoff user and routed its traffic to only one path which is maximally jointed with the old primary path (a5), as shown in Figure 3.1.
- Arrival of a handoff user and routed its traffic to maximally disjointed paths, old path and the two paths selected for the handoff call are all completely disjointed (a6).
- Arrival of a handoff user and routed its traffic to maximally jointed paths, old path and the selected one or two paths for the handoff call are all maximally jointed with each other (a7).

In a highly congested network environment routing possibilities will be chosen based on priority. Handoff user traffic and joint paths are assigned higher reward values. The reward parameter will equal 0 if a new, or handoff, call is rejected. This reward parameter will equal R_j, as shown in the equation below. When the new, or handoff, call is accepted in one of the above seven routing possibilities *(a1, a2, a3, a4, a5, a6, a7)*.

$$R_j = \begin{cases} r_j, & for\, some\, a_j \\ 0, & otherwise \end{cases} \tag{3.5}$$

The values of the rewards, r_j, are written in the Table 3.1 below:

Table 3.1 Reward Values

Reward	**Value**
r1	*p1*
r2	*p2*
r3	*p3*
r4	*p4*
r5	*p5*
r6	*p6*
r7	*p7*

Where p_i are reward's arbitrary values, in order to force our preferential algorithm to choose one path over the other, where *p1* has the minimum reward value and *p7* has the maximum reward value. Now we can define our objective function which will be to increase the number of calls accepted over a period of time *T*, the objective function could be expressed as:

$$f_{ij}(T) = A_i(T).\sum_{j=1}^{A_i(T)} R_j \tag{3.6}$$

Where $f_{ij}(T)$ represents the weighted number of accepted users at node *i,* based on our algorithm, after selecting our preferential more rewarded paths over a period of time *T*. $A_i(T)$ is the number of users accepted at node *i* over a time period *T*, and the summation is the rewards given to each user based on the selected route. R_j is the reward assigned for choosing one path over the other. This objective has to be reached under the constraint that the network should be highly congested, which means the summation for all the required bandwidths for all the accepted users (b_{rk}) should be less than the available bandwidth in link *l* (b_{al}).Hence, we can define our sub-optimization model as indicated below:

$$\max \frac{1}{T}\sum_{i=1}^{N}\left(A_i(T).\sum_{j=1}^{A_i(T)} R_j \right) \tag{3.7}$$

Subject to:

$$\sum_{k=1}^{A_i(T)} b_{rk} \le b_{al}$$

To solve the above sub-optimization problem, we propose a Markov Decision Process and value iteration method to find a near optimal solution.

3.4 Review of Performance Measurements

In this section, the performance measurements for blocking probability, average bandwidth utilization, and handoff delay are defined below:

Blocking Probability (B_P):

The call is blocked whenever there is no available bandwidth in the network and the queue is full. This can be expressed by the following formula (Sztrik (2011)):

$$P[C=c]=P_c=\frac{(1-\rho)\rho^c}{1-\rho^{c+1}} \tag{3.8}$$

Where *P[C = c]* is the probability that there are exactly c calls in the system, so if any other call arrives it will be blocked. The server utilization can be calculated by dividing the arrival rate over the service rate.

Handoff Delay ($D_{handoff}$):

Handoff delay is the time required for the mobile node to reserve new resources for the handoff call, this is divided to three sub-delays, which are T_{DL}, $T_{Network}$, and $T_{Application}$. As defined below (Khasawneh et al., 2014).

$$D_{handoff}=T_{DL}+T_{Network}+T_{Application} \tag{3.9}$$

Where T_{DL} is the handoff delay coming from the data link layer, $T_{Network}$ is the delay from the network layer and $T_{Application}$ is the delay from the application layer. Handoff delay is most affected by the network layer (Zhao et al., 2012; Xie et al.,2008), and is composed of delays related to discovering the route, associating the handoff call to the newly discovered route, and updating the location of the mobile node for the future routing decisions. We focus here on finding the handoff delay caused by the network layer. In this paper we will focus on finding the handoff delay caused by the network layer.

3.5 Markov Decision-based Admission Control and Routing (MDACR)

As we noted, new and handoff calls are the two types in a network. Whenever a new call arrives to a highly congested network, rejection is highly probable due to lack of bandwidth in the network. Our proposed algorithm will increase the admission rate by connecting the new call with two access points, routing user traffic, from the source to the destination, through two different paths using multi-homing. Those two paths may be joint or disjoint. As shown in the results section, more reward will be assigned to the joint path, as this provides better QoS compared to the disjoint path. We design a Markov decision process, in the mathematical model below, to make decisions on the selection of the route for the new, and handoff, calls. The benefits of our model on QoS arise from its use of joint routing and multi-homing over the disjoint routing and no multi-homing scenario in terms of QoS.

Our network topology two levels: network access and backhaul network levels. From Table 3.2 below, when a new call arrives, with bandwidth available in the two network levels, Multi-Homing (MH) is not triggered and one route is found for the new call traffic. However, MH is triggered by our proposed algorithm, if bandwidth is available at the access network level but not at the backhaul network level then our proposed algorithm will trigger MH feature in the access network level. This will split the new call traffic between two paths, increasing probability of new call acceptance, because the required bandwidth is reduced by half. In this scenario, it is better to find Maximally Jointed (MJ) paths to decrease the reordering complexity at the destination side. In the third case (C3), the bandwidth is not available at the access network level, but is available at the backhaul network level. In this case, MH will be triggered, increasing the admission rate by splitting the new call traffic into two different paths. MJ paths are also preferred for reduction of the reordering complexity at

Table 3.2 Triggering Multi-homing and Routing Based on the Availability Bandwidth

Type of Call	**Triggering Multi-homing and Routing Based on the Available Bandwidth.**			
	Case Number	**Access**	**Backhaul**	**MH**
New Calls	**C1**	Available	Available	Not Triggered
	C2	Available	Not Available	Triggered
	C3	Not Available	Available	Triggered
	C4	Not Available	Not Available	Triggered
Handoff Calls	**C5**	Available	Available	Not Triggered
	C6	Available	Not Available	Triggered
	C7	Not Available	Available	Triggered
	C8	Not Available	Not Available	Triggered

the destination side. In the last, worst, case (C4), there is no available bandwidth in the access network or backhaul network levels, triggers MH, and a MJ path is preferred.

There is no need to trigger Multi-Homing (MH) when a handoff call is arriving to available bandwidth in the two network levels: one route is found for the handoff call traffic, Maximally Jointed with the mobile user Old (MJO) path. However, if the bandwidth is available at the access network level but not at the backhaul network level then our proposed algorithm will trigger MH feature in the access network level. This splits the handoff call traffic between two paths, reducing required bandwidth in half, increasing the probability of handoff call acceptance. In this scenario, in order to decrease the reordering complexity at the destination side and to reduce the handoff delay, it is better to find a Maximally Jointed Trio paths with the mobile user Old (MJTO) path. In the C7 case, the bandwidth is not available at the access network level but is available at the backhaul network level. In this case, MH is also triggered. This increases the admission rate by splitting the handoff call traffic into two different paths. MJTO paths are also preferred to reduce the reordering complexity at the destination side and the handoff delay. In the last case (C8), the worst case scenario, bandwidth is unavailable either in the access network level or in the backhaul network level. In this case, MH is also triggered, and MJTO paths are preferred.

3.5.1 Markov Decision Process Model

Whenever an existing user call is moving, which is routed through the path Π1, a new path should be discovered, denoted Π2. If the two paths are joint, with common links between them, then this joint situation is denoted by J(Π1, Π2), otherwise it will be denoted by D(Π1, Π2).

In order to model our problem using MDP, states should be defined. In our situation, we have four different states: reject, one path state, joint paths state and disjoint paths state as shown in Figure 3.2 below. Two different types of calls are defined as new calls and handoff calls. The definition of each state is described below.

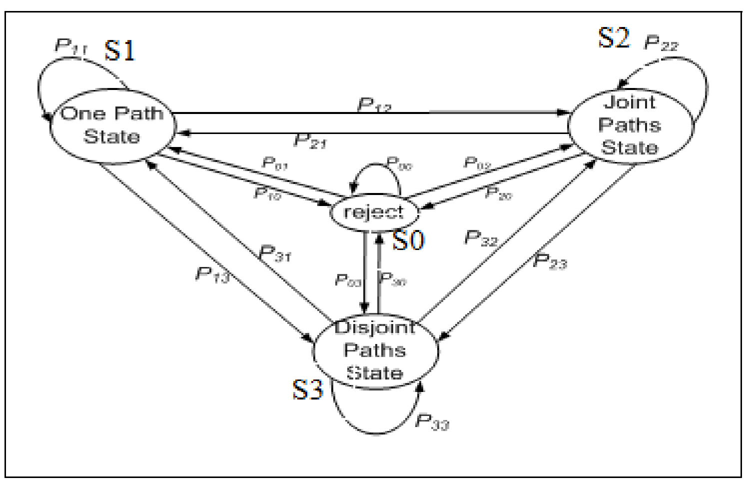

Figure 3.3 Markov Decision Process State Diagram
Taken from Khasawneh et al. (2016)

For applications that demand uninterrupted connectivity, users may suffer from connectivity issues during handoff of roaming mobile devices in a Wireless Mesh Networks. To resolve this issue and to serve a seamless operation by facilitating faster handoff management and lower packet loss, an intra-domain handoff management scheme is proposed in this paper for Wireless Mesh Networks. The proposed scheme is based on the selection of the shortest link between the source and destination of a mobile client, which are joint maximally with primary one.

The proposed technique is based on three timing thresholds. During the threshold-1, a network level handoff preparation process is initiated in advance with a view to reduce handoff delays. The preparation process includes scanning all the adjacent nodes near the terminal node for collecting required information. At the end of threshold-1, threshold-2 is initiated. During this threshold period, the joint route finding process is initiated and optimal path with best AP is selected. After the selection-process ends threshold-2 ends and threshold-3 starts. During this threshold, several signalling is performed. These signalling include disconnection from former AP router, connection to new AP router and update of all necessary location and subscription information.

Markov model states represent the connection of the mobile terminal with three possible types of paths through the network and are one path, joint, or disjoint, paths. A fourth is the reject state when the mobile terminal is not active, or not able to connect, to one of the access points due to the lack of bandwidth. Transition probabilities of the Markov Chain are defined to represent all possible transitions from one state to another as shown in Figure 3.2.

State 0 (Reject state): we have either a new, or a handoff, call. Whenever a new call cannot find available bandwidth for admission into the network, this call will be in the reject state and stay there until bandwidth is reserved for it. If, instead, we have handoff call that is moving to another highly congested part of the network, this call might be rejected into a reject state until some bandwidth is available. The transition probabilities from reject state to the other states are denoted by P_{01}, P_{02}, P_{03}. The last transition probability P_{00} happens when the reject state call has no bandwidth available.

State 1 (One path state): a new call arriving will be in state 0. Supposing there is enough available bandwidth in the network, then there is no need to trigger the multi-homing feature. The packets for this call are routed from source to destination through one route only. The call is transferred into state 1 with transition probability *P01*. However, when no bandwidth is available in the network, the multi-homing feature is triggered. More than one AP as needed is chosen and the call is transferred from state 0 to either state 2 or state 3 with transition probabilities *P02* and *P03*, respectively. The transition probabilities from one link state to the other states are denoted by P_{10}, P_{12}, P_{13}. The last transition probability P_{11} happens when the user call packets are routed though one path and the user is not moving.

State 2 (Joint paths state): a call will be in this state if the packets are routed through two different paths which are joint, denoted by J(Π1, Π2). This might involve new or handoff calls arriving into a highly congested network, which is obliged to connect to two different APs and make two joint paths. Transition probabilities from joint link state to the other states are denoted by P_{20}, P_{21}, P_{23}. The last transition probability P_{22} happens for a call in the joint link state and their remains no available bandwidth.

State 3 (Disjoint paths state): of this presented Markov model, occurs when the mobile terminal is connected to two access points that create disjoint paths. P_{33} is the probability that the mobile terminal is again in a disjoint paths state. P_{31} is the probability when the mobile terminal connected with just one access point. A cause of this event could be lost signal with one of the two connected APs. This happens when the mobile terminal can receive signal from just one AP. P_{32} is the probability for the transition from disjoint paths state to joint paths state. This transition probability might happen when, for example, with failure of some of the routers in the joint link and joint link cannot be established. P_{30} is the transition probability of the mobile terminal ending the active session when it is in the disjoint paths state.

Matrix P is developed from the above explained states and state transition probabilities. $P_{si}[x]$ is the steady-state probability at moment x-th. $P_{ij}[x]$ is used to denote the transition probabilities from state i to state j, where $i,j \in 0,1,\ldots,N_{state}$. In our case N_{state} is equal to 3. Using these terms the matrix P will be given by:

$$\underline{P}[x] = \begin{bmatrix} P_{s0}[x] \\ P_{s1}[x] \\ P_{s2}[x] \\ P_{s3}[x] \end{bmatrix} = \begin{bmatrix} P_{00}[x] & P_{10}[x] & P_{20}[x] & P_{30}[x] \\ P_{01}[x] & P_{11}[x] & P_{21}[x] & P_{31}[x] \\ P_{02}[x] & P_{12}[x] & P_{22}[x] & P_{32}[x] \\ P_{03}[x] & P_{13}[x] & P_{23}[x] & P_{33}[x] \end{bmatrix} \bullet \begin{bmatrix} P_{s0}[x-1] \\ P_{s1}[x-1] \\ P_{s2}[x-1] \\ P_{s3}[x-1] \end{bmatrix} \tag{3.10}$$

A new terminal connection changes the idle state 0 to some of the nonzero states. Transition to state 0 is accomplished if the terminal finishes the session while it is in some of the nonzero states, or whenever the mobile call is rejected. The proposed Markov model states are recurrent non-null and the equilibrium state probabilities can be determined to solve equation 3.7 . The sum of the steady-state probabilities is equal to 1, so :

$$\sum_{i=0}^{N_{state}} P_{si} = 1 \tag{3.11}$$

Coverage of the access points is the factor that has the greatest influence on the values of the steady-state probabilities $P_{si}[x]$. Steady-state probabilities $P_{s2}[x]$, $P_{s3}[x]$ are more probable if the mobile terminal, moving most of the time, is in overlapping coverage of two or more access points. So, here, $P_{s2}[x]$, $P_{s3}[x]$ would have higher values than $P_{s1}[x]$. Steady-state probabilities $P_{s1}[x]$, $P_{s0}[x]$ would have higher values if most often the mobile terminal moves in locations with coverage of just one access point. This is especially true for $P_{s0}[x]$ when the network is congested and the accessibility of the access points is low.

In the proposed model the Poisson process will be used to describe the arrival of the new sessions (packets or calls) in the coverage areas with mean arrival rates of $\lambda_{1\text{-}link}$, λ_{jl} (jl means joint link) and λ_{dl} (dl means disjoint link).

The duration of the session will be denoted with $T_{session}$. Probability of the session termination or completion will be denoted with $P_{termination}$. From the state diagram in Figure 3.3, it can be concluded that:

$$P_{termination} = P_{10} + P_{20} + P_{30} \tag{3.12}$$

Inter-arrival time will be denoted with T_{iar} and it is equal to:

$$T_{iar} = 1/\lambda_n; \qquad \lambda_n = \lambda_{1-link} + \lambda_{jl} + \lambda_{dl} \tag{3.13}$$

Since, we previously developed four states in the proposed Markov model for the proposed routing algorithm, it is obvious that:

$$\sum_{i=0}^{3} P_{si} = 1 \tag{3.14}$$

According to the above conditions, the probability that the mobile terminal is connected to the disjoint link will be:

$$P_{dl} = \sum_{i=0}^{3} P_{i3} \tag{3.15}$$

The probability that the mobile terminal is connected to the joint link will be:

$$P_{jl} = \sum_{i=0}^{3} P_{i2} \tag{3.16}$$

The probability that the mobile terminal is connected to just one link is going to be:

$$P_{1-link} = \sum_{i=0}^{3} P_{i1} \tag{3.17}$$

Hence, according to all previously developed equations related to the proposed Markov model, it is clear that the probability of the new session arrival with one link will be $P_{new1\text{-}link}$

$= P_{01}$, and to the joint link and disjoint link, $P_{newjl} = P_{02}$, $P_{newdj} = P_{03}$. We can conclude that the probability of the new session arrival will be expressed as:

$$P_{new} = P_{new1-link} + P_{newjl} + P_{newdl} = P_{01} + P_{02} + P_{03} \tag{3.18}$$

From the developed Markov model for the proposed routing scheme, after defining the states, probabilities and arrival rates of the sessions using One Link, Joint Link and Disjoint Link States, numerical analysis will be developed of the horizontal handover probabilities of the possible links. The overall horizontal handover probability is:

$$P_{ho}[x] = P_{s1}[x-1](P_{12} + P_{13}) + P_{s2}[x-1](P_{21} + P_{23}) + P_{s3}[x-1](P_{31} + P_{32}) \tag{3.19}$$

Every state that was previously explained in the Markov model includes a group of events that are denoted with S_i, because each of these three different types of links has a number of access points and routers, denoted with N_r. The probability of being in any of the three defined states $Ps1$, $Ps2$, $Ps3$ is specified with the probability of being connected with one link, joint links or disjoint links inside the network. Hence, we can write:

$$P_{si}[x] = \sum_{z=1}^{N_r} P_{siz}[x]; \quad i = 1,2,3 \tag{3.20}$$

$P_{siz}[x]$ is the probability that the mobile terminal is attached to the *z-th* access point or router from *i-th* link (the link could be One Link, Joint Link or Disjoint Link) at the *x-th* moment of time. Hence, the probabilities of the state transitions can be written as:

$$p_{ij}[x] = P\{S_j[x] \mid S_i[x-1]\} = \sum_{z=1}^{N_r} P\{S_{jz}[x] \mid S_{iz}[x-1]\} \tag{3.21}$$

Each of the probabilities $P\{S_{jz}[x]|S_{iz}[x\text{-}1]\}$ could be computed using the horizontal handover algorithm. The transition probability from state i to state j could be written in terms of the rewards and joint degree weight variables as follows:

$$p_{ijr}[x] = \frac{w_{ijr}.R_{ij}}{\sum\limits_{i,j,r} \left(w_{ijr} \cdot R_{ij}\right)} \tag{3.22}$$

Where w_{ijr} is the joint degree weight which depends on two main factors, the first one is the number of common links between the old primary path and the new candidate path r in the case of moving from state i to state j, and the other factor is the expected future discount value calculated after choosing the present state which is defined below. The multi-homed joint route has the highest joint degree weight in the case of a highly congested network. R_{ij} is the reward assigned in the case of moving from state i to state j in every possible route.

Discount Factor

In this section the reward and discount factor of the Markov Decision Process for the proposed model will be outlined. $RI\ (S_i,\ S_j)$ is denoted as immediate reward or expected immediate reward received after the transition to state S_j from state S_i. $\gamma \in [0,1]$ is defined as the discount factor, which presents the difference in importance between present, and future, rewards. States or actions (in our case denoted as transitions from one state to another) are not declared finite in the theory of Markov Decision processes. In our proposed model we assume that they are finite.

Each of the states in our model has an immediate reward denoted as $\{RI_0, RI_1, RI_2, RI_3\}$. On each time step, it is assumed that the state S_i is given an immediate reward RI_i and randomly moves to another state $P(NextState{=}S_j/ThisState{=}S_i) = P_{ij}$, with all future rewards being discounted by γ.

We will use $J^*(S_i)$ to define the expected discounted sum of future rewards starting with the state S_i. So we will have:

$$\begin{aligned} J*(S_i) &= RI_i + \gamma \times (\textit{Expected future rewards starting from our next state}) = \\ &= RI_i + \gamma(P_{i0}J*(S_0) + P_{i1}J*(S_1) + P_{i2}J*(S_2) + P_{i3}J*(S_3)) \end{aligned} \tag{3.23}$$

In vector notation:

$$\underline{J} = \begin{pmatrix} J*(S_0) \\ J*(S_1) \\ J*(S_2) \\ J*(S_3) \end{pmatrix} \qquad \underline{R} = \begin{pmatrix} R_0 \\ R_1 \\ R_2 \\ R_3 \end{pmatrix} \qquad \underline{P} = \begin{bmatrix} P_{00} & P_{01} & P_{02} & P_{03} \\ P_{10} & P_{11} & P_{12} & P_{13} \\ P_{20} & P_{21} & P_{22} & P_{23} \\ P_{30} & P_{31} & P_{32} & P_{33} \end{bmatrix} \tag{3.24}$$

A Markov model system has been proposed here, with 4 possible states: idle state, one path state, joint state and disjoint state. A reward or payoff is the cost of serving the user, which depends on whether it is a new session or a handover session. The purpose is to accept or reject users to increase the expected value of the rewards.

The joint degree weight (w_{ijr}) can be defined as:

$$w_{ijr} = n \cdot J*(S_i) \tag{3.25}$$

Where n is the number of common links between the old primary path and the new candidate path, and the $J^*(S_i)$ is the expected future reward if state S_i is chosen at present.

Our proposed model works in this way with call admission control for the new sessions. If the requested bandwidth of the new session is equal to or below the actual available bandwidth of the accessible access point, then the transition from State 0 to State 1 is more reasonable because one path is enough to fulfill the user's requirements. But, if the requested bandwidth of the new session is above the available bandwidth of the accessible access point (it happens when the network is congested), our proposed routing algorithm suggests a multi-homed joint path to be used for an improved network performance. The new session will be blocked if the requested bandwidth of the new session is still above the offered bandwidth of the access points that obtain joint link. Disjoint path state is the option used by algorithms proposed in the literature. Comparisons between multi-homed joint paths and non-multi-homed disjoint paths is introduced in the Results section.

Call admission control works for our proposed model similarly for the handoff sessions. If the mobile terminal is in State 1, connected to one path and it is in the handover phase, then if the requested bandwidth of the mobile terminal is equal or below the offered bandwidth from the new accessible access point, the best solution is again to have one path. So, in this case mobile terminal will stay in the same State 1. But, if the offered bandwidth from the new accessible AP is below that requested, then the mobile terminal will try a multi-homed joint link, so transition to State 2 will be made. Furthermore, if a joint link is not possible, then the handoff call will be blocked. A comparison between multi-homed joint paths and non-multi-homed disjoint paths is also discussed in the Results section. Those discussed will be the most important state transitions, other state transitions having been explained above. It is taken into consideration that the new call can be admitted into the network with the specified QoS required by the call, without affecting the QoS for the existing calls in the network (Dhurandher et al., 2015).

3.6 System Model

The system model consists of one gateway and 10 to 25 mesh routers. These are uniformly distributed to cover the whole simulation area of 400m x 400m to serve the arriving new and handoff calls. All mesh routers are connected to other neighbouring routers in their transmission range (see Figure 3.4 below).

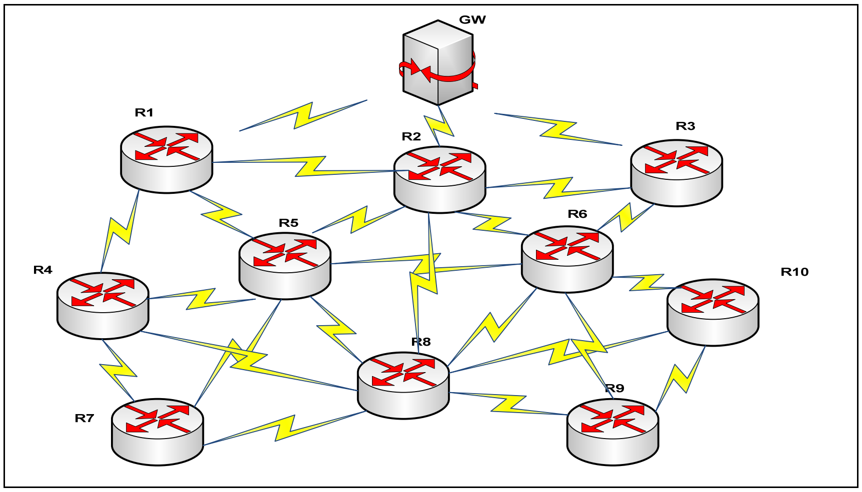

Figure 3.4 System Model

Horizontal intra-domain handoff describes this scenario, where users are moving randomly within the same domain.. Our proposed algorithm is applied when the network is highly congested with link bandwidth utilization equal to μ. QoS requirements for both new and handoff calls are ensured.

3.7 Simulation Results

For our simulation, the MAC layer protocol utilized between the mesh routers is 802.11a with frequency band of 5GHz, with eight non-overlapping channels available simultaneously without interference. Communication between mesh clients and routers is accomplished using 802.11b with a frequency band of 2.5 GHz. Only three orthogonal channels are available for use. In our simulation, the congestion is defined by 5 different levels and each congestion level has a range, which is uniformly distributed as a function of the occupied bandwidth in the network. Level 1 has a range of 0-0.2, which means that the occupied bandwidth in the network ranged from 0 to 20%. Level 1 is considered the lowest congestion level, level 2 has a range of 0.2-0.4, level 3 has a range of 0.4-0.6, level 4 has a range of 0.6-0.8 and the highest level of congestion, level 5, has a range 0.8-1. The list of simulation parameters are shown below in Table 3.3:

Table 3.3 Simulation Parameters

Parameter	Value
Mobility Model	Random Waypoint
Routers and Gateways	20-100 R, 1 GW
Simulator	Matlab
Simulation Area	400 m x 400 m
Transmission Range	75 m
Packet Size	1024 Bytes
Routing Algorithms	AODV, MDACR
Packets Arrival Rate	30 Packets/Sec
Radio Transmission Power	15dBm
Mobility Speed	10 m/s
Minimum Throughput Required	1Mbps
MAC Protocol	802.11a, 802.11b
Queuing Model	M/M/1/c

Below are the results, showing a comparison between our proposed algorithm and a disjoint routing algorithm used in the literature.

Figure 3.5 shows the New Call Blocking Probability (NCBP) when the congestion level increases in the network for both cases. Multi-homing and non-multi-homing scenarios are simulated. With congestion low and resources are available, multi-homing and non-multi-homing scenarios have the same blocking probability. When congestion increases the necessity of using multi-homing rises. This will reduce the blocking probability for the new call, as the traffic is divided between two different disjoint paths.

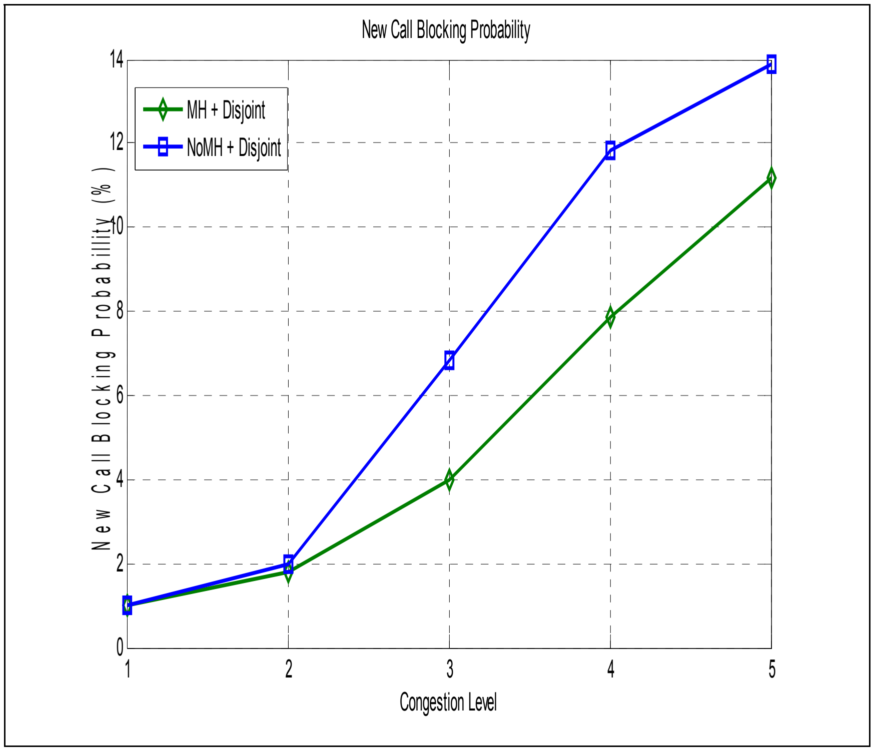

Figure 3.5 New Call Blocking Probability (NCBP)

Handoff Call Blocking Probability (HCBP) is calculated using the four possible scenarios shown in Figure 3.6 below. For the handoff calls, our proposed algorithm outperforms other algorithms in a congested network. A maximally jointed path with the old primary path provides better blocking probability compared to the disjoint routing since fewer resources are needed to connect the new path with the old path. More resources are needed with disjoint routing, so when congestion occurs the needed resources to establish the call might be unavailable, causing the call to be blocked. Multi-homing adds another level of assurance for reducing HCBP, where the required bandwidth for the handoff call is divided between two access points, maintaining its QoS requirements.

The Number of Hard Handoffs (NHH) is a count of the times a handoff call is disconnected due to lack of resources. Regardless of the applied algorithm, the number of hard handoffs in

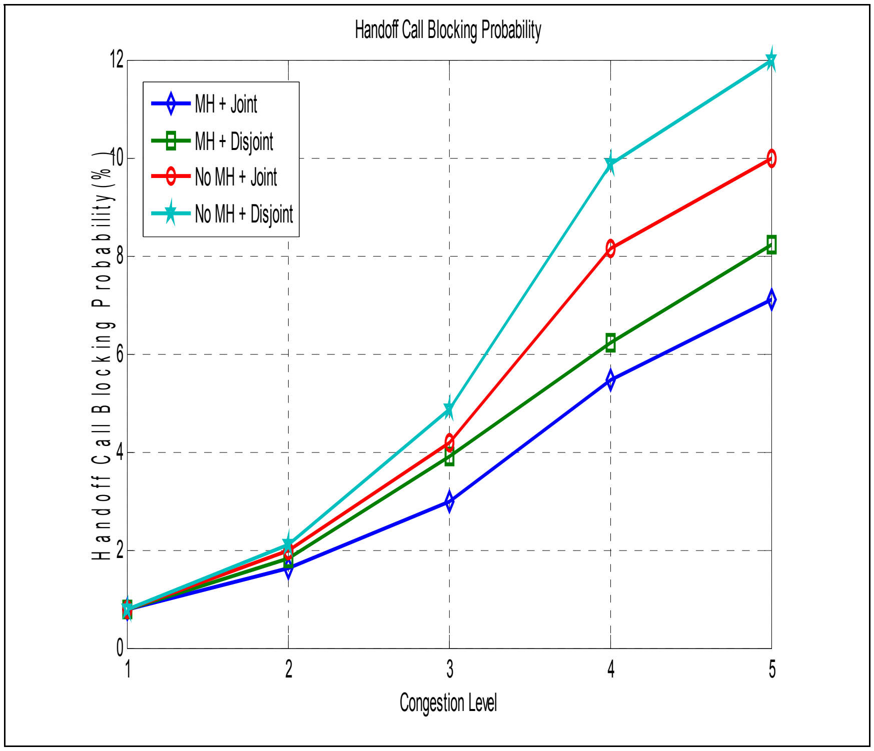

Figure 3.6 Handoff Call Blocking Probability (HCBP)

the system increases when the congestion rises. However, our proposed algorithm reduces NHH since it forces the handoff call to be routed to a maximally jointed path with the old reserved path. It also provides the feature of using multi-homing when congestion exists in the network . On the other hand, in the case of no multi-homing and a completely disjoint routing is selected, the number of hard handoff is the worst. Whereas on the other two cases the affection of using the joint routing alone or multi-homing feature alone is investigated. The simulation results show that the usage of MH feature alone is better than using the joint routing alone and will have a lower number of hard handoffs since if the MH is used the traffic is routed through two paths and the resource requirement for each path is less compared to the resource requirements when the traffic is routed through one path in the joint routing as shown in Figure 3.7.

In Figure 3.8, the Handoff Delay (HD) is calculated. As defined above, handoff delay is calculated on the three layers: application, network and data link. The main cause of handoff delay is located in the network layer, which is related to routing. As shown in Figure 3.8 below, when multi-homing is used with disjoint routing, the handoff delay is bigger than

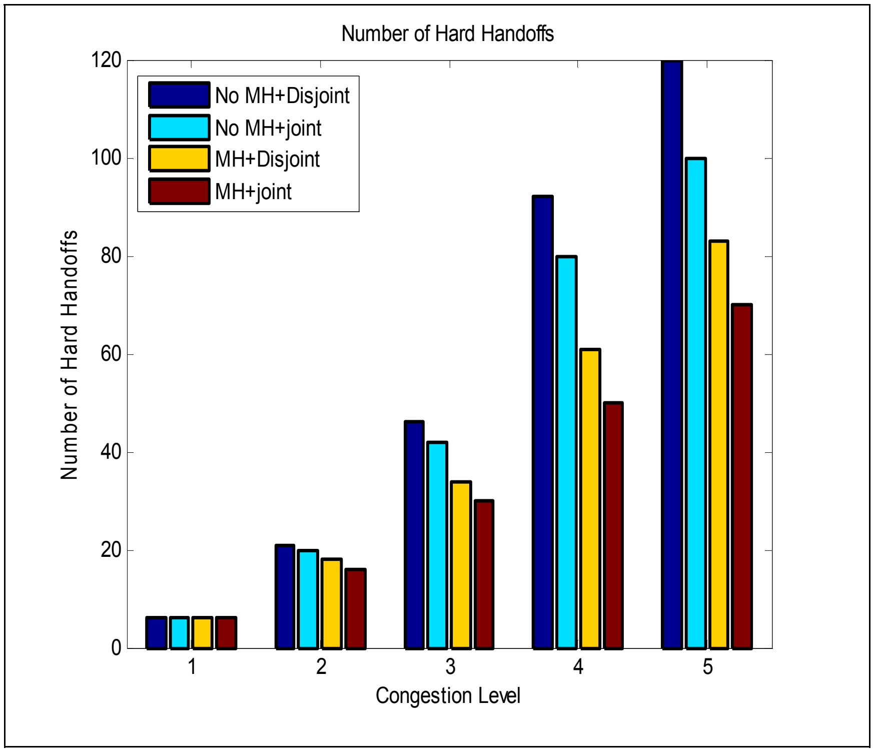

Figure 3.7 Number of Hard Handoffs (NHH)

when there is no multi-homing and disjoint routing. This is because with multi-homing two routes should be discovered. The greatest handoff delay is obtained when the multi-homing feature is disabled and joint routing is enabled, in this case only one route should be discovered which is maximally jointed with the old route.

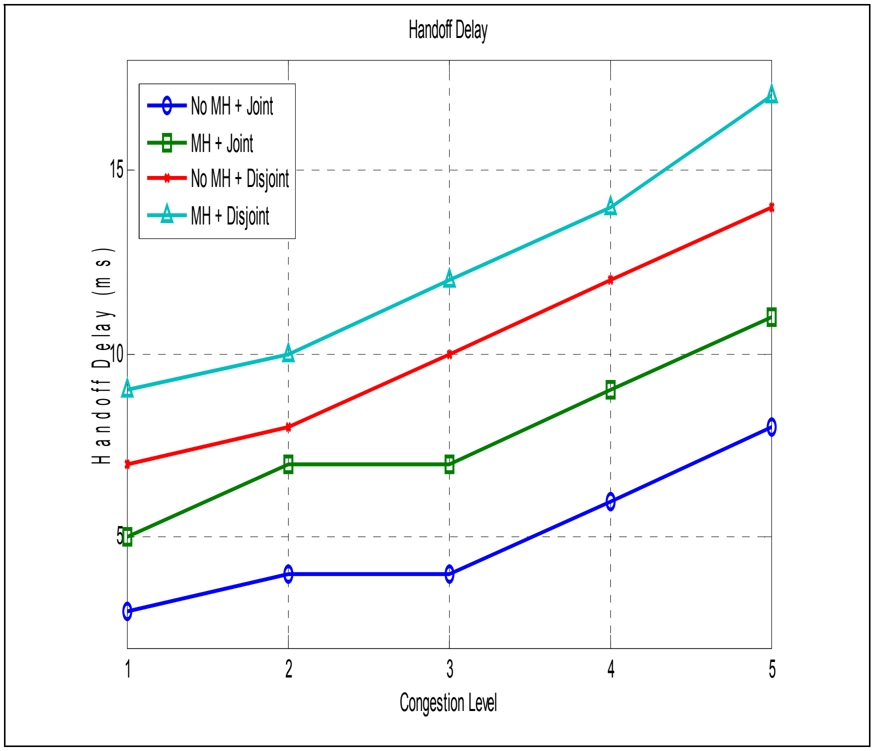

Figure 3.8 Handoff Delay (HD)

Figure 3.9 shows the relationship between the density of the network and the average number of reserved resources. The density of the network is termed as the population, which can be simply defined as the number of routers in the network. When a mobile node is moving, and applying our proposed handoff joint routing algorithm, the only resources that have to be newly reserved are the ones from the access point up to the first common router. After reaching the first common router there is no need to reserve any other kinds of resources since the same SIP session is still established and the same bandwidth are still allocated for the user engaged in the session. In case of highly congested network, this will reduce the probability of rejecting the handoff calls due to the lack of resources in the network. On the other hand, if the mobile node is moving and other algorithm is applied, which depends on finding a completely disjoint route for the handoff call. In this case, a new SIP session will have to be established and the new AP will ask for new resources to be reserved for this session from the source node up to the destination node. If the network is highly congested, the routers in the newly discovered disjoint path might not find available resources for this handoff call and it can be blocked. If the network is denser, it will decrease the average number of resources that have to be reserved since it will be more probable to find a shorter route to the destination.

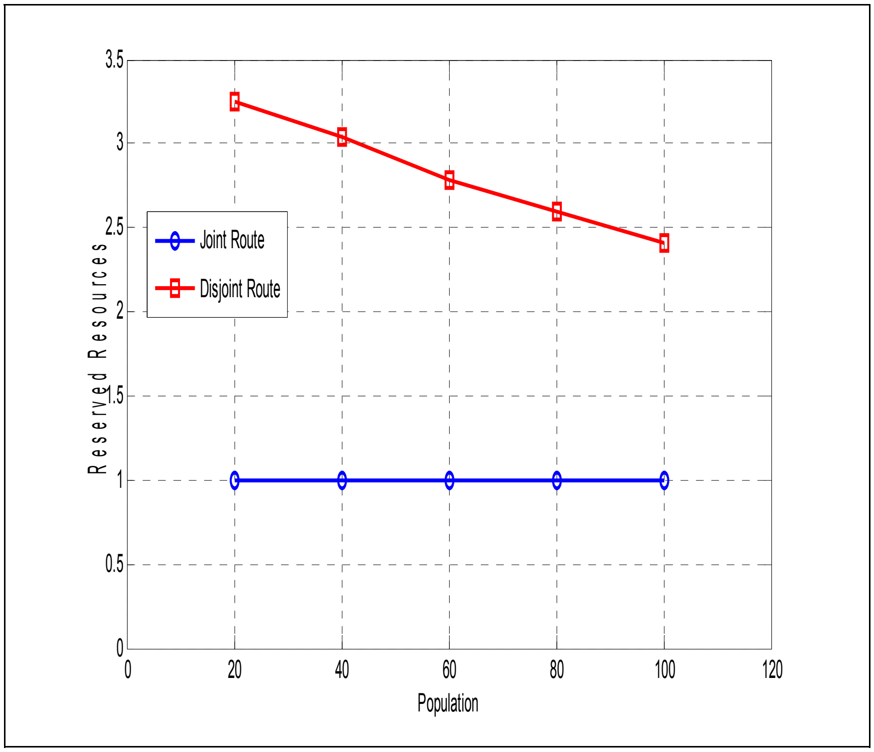

Figure 3.9 Reserved Resources for Joint and Disjoint Routes

Figure 3.10 below shows the relationship between the handoff delay and the density of the network. As previously explained, when the network is denser the probability to find a shorter path to the destination increases. When shorter paths are found, the time taken to reserve resources is less and thus the handoff delay is decreasing when the population increases. Handoff delay is decreasing since the route discovery, association and updating delays are also decreasing. If our algorithm is applied, most of the time the moving mobile node will find a one hop path connecting him to the old path and thus the handoff delay is almost the same and the increase in it is not noticeable when the population increases.

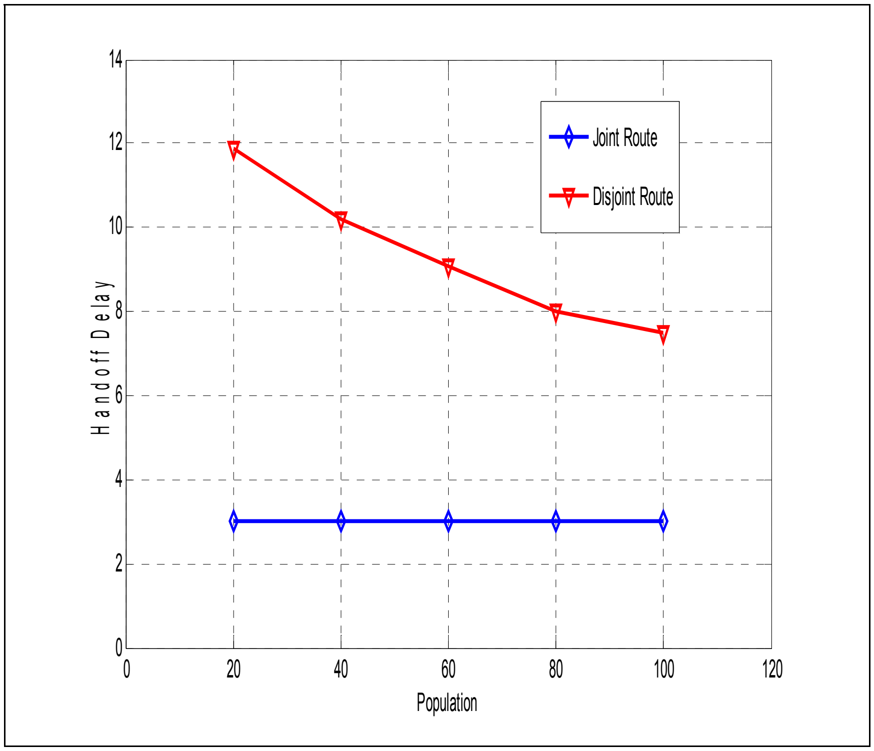

Figure 3.10 Handoff Delay vs. Population

Figure 3.11 shows the relationship between packet loss percentage and the population. Joint routing outperforms disjoint routing as the population increases. Since joint routing has less handoff delay and the average number of reserved resources are fewer in joint routing compared to the disjoint routing, the packet loss is also better in our proposed algorithm.

Before data transmission begins, we assume a new or handoff call reaching our network has established a connection. Uplink Constant Bit Rate (CBR) traffic is assumed to be sent from all the mesh clients to the destination gateway through one or more mesh routers. The packets reaching each mesh router is handled using Time Division Duplex (TDD). Our proposed routing algorithm is distributed, which means that each router is responsible for making the routing decision based on notifications received periodically from the neighbouring routers. These notifications provide the buffer occupancy, bandwidth availability in each candidate next mesh router, and the joint degree metrics. A random waypoint model is used to simulate the movement of the mobile users in the network. It is essential to our proposed protocol that congestion status be monitored in each link in the network.

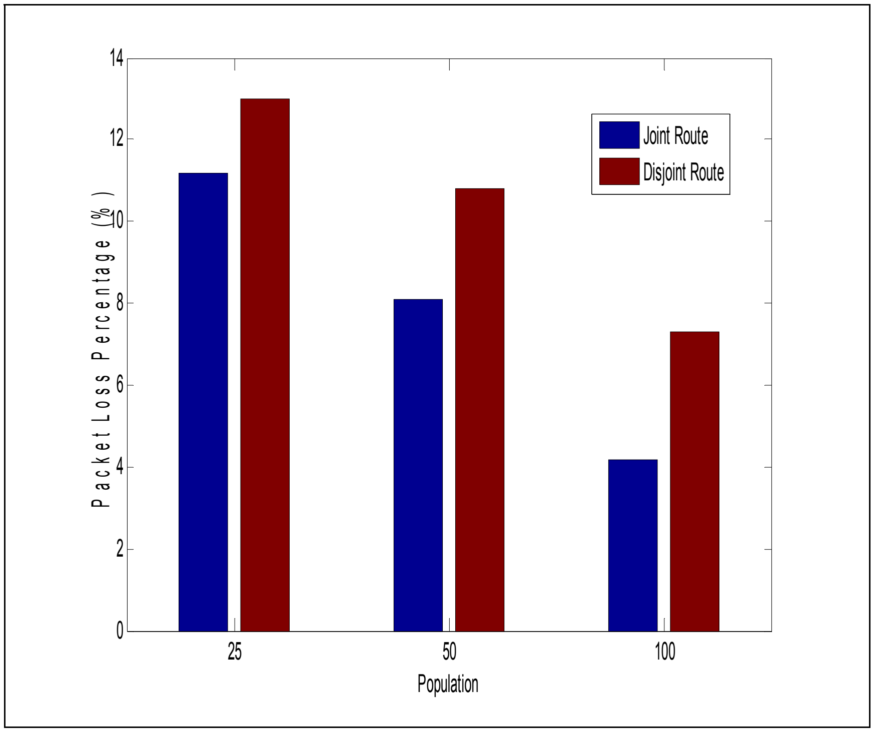

Figure 3.11 Packet Loss

The network traffic pattern is dynamic, so changes are a normal expectation. This leads to consideration of an enhancement of our proposed algorithm to design a network traffic pattern prediction algorithm, (Khasawneh et al., 2015), and use the result of the prediction in the routing decision for better QoS.

A lot of mobility management schemes have been discussed in Chapter 2 such as (Chaya et al., 2016; Mallikarjuna et al., 2016; Mamidi et al., 2015; Qin et al.,2015; Hung et al.,2009) and many others. None of them are evaluating the benefit of using maximally jointed for the user handoff traffic and none of them are adding the feature of multi-homing. Simulation results above shows that our proposed algorithm outperforms other algorithms.

CHAPTER 4

ADAPTIVE TRANSMISSION PREDICTIVE CONGESTION CONTROL

4.1 Introduction

Network congestion is a situation where a network node is carrying more data than which it can handle. Some of the network resources which should be considered in loading a network includes buffer memory and processing speed. Problems associated with congestion include long time wasted for the jobs and data in the queues, packet loss, as well as new connections that can also get blocked (Islam et al., 2014). The main steps that commonly exist in most of the congestion control algorithms proposed in the literature are congestion detection, congestion signalling and flow rate adjustment. However, the first step is the most important step which can be done in an explicit or implicit way. In implicit approaches, the node can implicitly determine that there is a congestion in the network whenever the data transmission takes longer what it usually takes to transmit the packet, packet loss or even not receiving acknowledgment packet. This approach is not suitable for wireless network as the delay and packet lost could be resulted from other causes. An explicit congestion notification is the approach that we are following in our proposed algorithm. Congestion status are stored when an explicit congestion notification are received. Those congestion history is used in our proposed Variable Order Markov prediction model.

The benefits and importance of studying congestion control are many. Through studying of congestion control, the researchers get to know various network policies which are likely to affect network schemes and network design. With this knowledge, the researcher or network designer is unlikely to make errors during the configuration of networks. The second importance of studying congestion is to know how to allocate dynamic network resources (Islam et al., 2014). When the researcher/network designer knows how and which and in which measure to allocate network resources, this is likely to solve the stubborn congestion problem. The other importance is that they will know various protocol design decisions.

One of the factors used in congestion mitigation is an explicit notification of congestion or feedback messaging. This is an extension on TCP/IP communication protocols whose role is to add the control to the flow of packets from one protocol level to another (Islam et al., 2014). Which provides various methods of dealing with network congestion control. The method used by this factor involves provision of different algorithms for dealing with congestion such as additive/multiplicative decrease (IAMD), which tend to reduce the number of network users and packets by certain factors, and also slow start as well as congestion window algorithms.

Congestion in a wireless network is different from that of a wired network in that since the neighbouring wireless nodes share the same channel. This implies that congestion in wireless networks is brought about by the traffic between the nodes on the network (Keerthana et al., 2015). The other characteristic of the wireless mesh network is that these networks lack or have very minimal mobility when it comes to routers, which forms the backbone for this kind of networks, but the mesh clients can be either mobile or stationary. The fact that the mesh clients are either mobile or stationary makes the network traffic very difficult to monitor (Kapadia et al., 2015). This has provided a very great challenge to the researchers to find ideas or reducing congestion from such networks.

4.2 Problem Formulation

In our methodology, we designed a Variable Order Markov (VOM) (Begleiter et al., 2004) based congestion aware routing algorithm, which maximize the transmission rate based on the prediction done by the VOM prediction model. The prediction output that comes from the VOM prediction model feeds our sub-optimization model with the new route and based on the predicted congestion status in the new route, a new transmission rate is applied. In the section below, our proposed Variable Order Congestion Prediction is described in more details along with our sub-optimization model and the solution for this sub-optimization model using Lagrange method is derived and explained.

4.2.1 Variable Order Markov Congestion Prevention (VOMCP)

Let *G = (N,L)* be an undirected graph. Where *N* is the set of nodes, *L* is the set of links. Each vertex $v \in N$ we assign a probabilistic finite automaton A with a finite alphabet $\sum = \{a, b, c, d, e\}$. Where *a* is considered the least congested link and *e* is the most congested link. Let $w = w_1 w_2, ..., w_n$ be a word (training word) of length n with $w_i \in \sum$ and $w_i w_{i+1}$ is the concatenation of w_i and w_{i+1}. The goal is to learn A that provides a conditional probability assignment for any future outcome based on past observations. Let $v = v_1 v_2 ... v_m$ be a test word of length *m*. Prediction performance is measured with using average log-loss:

$$l(P, y) = -\frac{1}{m}\sum_{i=1}^{m} logP(v_i \mid v_1 v_2 ... v_{i-1}) \tag{4.1}$$

We refer to *P* as a conditional consistent distribution of the test sequence *y*. For the prediction algorithm we choose LZ-MS algorithm (Nisenson et al., 2003) which is an improved version of LZ78 compression algorithm.

Algorithm 4.1 Estimation Algorithm

1. Input: *G = (N,L),* $w_l = w_1 w_2 ... w_n$, , $l \in L$
2. Output: Distribution of congestion for each link $l \in L$
3. **For** each l, w_l :
4. *distr(l) = LZMS(l,*w_l*)*

This algorithm is responsible on finding the congestion status on each link and store it in an array.

4.2.2 Routing Decision

For the routing decision to be taken accurately, the prediction should be performed to know the future congestion state of the link. In our proposed algorithm, some arithmetic operator is overloaded, which means it has a different definition and behaviour other than the normal behaviour of those operators. Below is the new operator definition or behaviour that is used in our proposed algorithm.

Let $(H,\oplus,\otimes)$ with $H \neq \varnothing$, $\oplus : H \times H \to H$, $\otimes : H \times H \to H$ such that the following axioms are satisfied:

1. $\forall o,u,r \in H : (o \oplus u) \oplus r = o \oplus (u \oplus r)$
2. $\forall o,u,r \in H : (o \otimes u) \otimes r = o \otimes (u \otimes r)$
3. $\forall o,u \in H : o \oplus u = u \oplus o$
4. $\forall o \in H : o \oplus o = o$
5. $\exists \varepsilon \in H : \forall o \in H : o \oplus \varepsilon = \varepsilon \oplus o = o$
6. $\exists e \in H : o \otimes e = e \otimes o = o$
7. $\forall o \in H : o \otimes \varepsilon = \varepsilon \otimes o = \varepsilon$
8. $\forall o,u,r \in H : o \otimes (u \oplus r) = o \otimes u \oplus o \otimes r$
9. $\forall o,u,r \in H : (o \oplus u) \otimes r = o \otimes r \oplus u \otimes r$

Let $H = [0,1]$. Set $o \oplus u = \max(o,u)$, $o \otimes u = ou$ with $o,u \in [0,1]$, $\varepsilon = 0$ and $e = 1$.

Let W be the weighted adjacency matrix of G . Let R_{st}^{max} be the probability of the most reliable st -path,then:

$$W_{st}^{*} = (\bigoplus_{k \geq 0} W^{k})_{st} = R_{st}^{max} \tag{4.2}$$

$$e_s = (1, 0, ..., 0) \tag{4.3}$$

Given the vector z and the vector of length $n = |N|$ we present an iterative method to solve W^*, where $z^{(0)} = e_s = (1,0, \ldots ,0)$.

$$z^{(k+1)} = (z^k W) \oplus e_s \tag{4.4}$$

where zW is usual matrix multiplication w.r.t $(H, \oplus, \otimes)$. After *(n-1)* stages the method terminates.

The vector z contains probabilities of the most reliable path to all other vertices. By setting $o \oplus u = \{argmax(p_o, p_u)\}$ and $o \otimes u = ou$ with $o,u \in [0,1]$ we get the links of corresponding most reliable path leading from start node to end node.

Algorithm 4.2 Routing Algorithm

1. Input: *G(V,E)* , link distribution, source node (s).
2. Output: Probability of the most reliable path.
3. **Read** at most *D* previous states.
4. **Construct** the test word *v* and weight the adjacency matrix W of G with the predicted probabilities.
5. **Set** $e_s = (1,0,\ldots,0)$ and $z^0 = e_s$
6. **While** $k < n$ **do:**
7. $z^{(k+1)} = (z^k W) \oplus e_s$
8. **end**
9. The *j*-th entry of *z* is the probability of the most reliable *sj*-path.

In order to make our model more understandable, the following example will go through our prediction model step by step, in order to explain how it works.

Example:

Assume we want to route video traffic from vertex 1 to other vertices. The preferred states are {a,b}, since they are the least congested link states.

Consider following graph in the figure below:

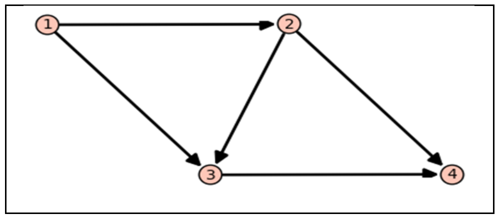

Figure 4.1 Routing Example

Table 4.1 below shows prediction results and edge weights:

Table 4.1 Prediction Probabilities

$l \in L$	Training	Test y =	$Pr(X_i = a$ or $b \mid y)$
{1,2}	Aaccdccacce	cc	1/3
{1,3}	Aacceccaa	cc	1/2
{2,3}	Abbccabba	abb	1/2
{2,4}	Abbccde	bcc	0
{3,4}	Dccdccb	cc	1/2

The conditional probabilities can be calculated using the trainingd and test words and apply them on our compression algorithm LZ78.

Now using routing algorithm:

$$e_s = (1,0,0,0)$$

$$z^0 = e_s$$

$$W = \begin{pmatrix} 0 & \frac{1}{3} & \frac{1}{2} & 0 \\ 0 & 0 & \frac{1}{2} & 0 \\ 0 & 0 & 0 & \frac{1}{2} \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

$$z^1 = (z^0 W) \oplus e_s = (1,0,0,0)W \oplus (1,0,0,0) \; z^1 = (0,\frac{1}{3},\frac{1}{2},0) \oplus (1,0,0,0) = (1,\frac{1}{3},\frac{1}{2},0)$$

$$z^2 = (z^1 W) \oplus e_s = (1,\frac{1}{3},\frac{1}{2},0)W \oplus e_s = (1,\frac{1}{3},\frac{2}{3},\frac{1}{4})$$

$$z^3 = (z^2 W) \oplus e_s = (1,\frac{1}{3},\frac{2}{3},\frac{1}{4})W \oplus e_s = (1,\frac{1}{3},\frac{2}{3},\frac{1}{3})$$

$$z^4 = (z^3 W) \oplus e_s = (1,\frac{1}{3},\frac{2}{3},\frac{1}{3})W \oplus e_s = (1,\frac{1}{3},\frac{2}{3},\frac{1}{3})$$

We obtain z^j , based of the calculated probability distribution vector z^j , traffic will be forwarded from node 1 to node 3, since it has the highest probability to get one of the preferred congestion states {a,b}. The output of VOM prediction model, which is the new route which has the highest probability to get the least congestion state. The new route is fed to the following sub-optimization model with the objective to maximize the allocated bandwidth x_s , and to maximize the user satisfaction as described in the section below.

4.2.3 Adaptive Transmission Rate Algorithm

As an utility function we propose:

$$g_s(x_s) = \frac{1}{1 + \delta e^{-\beta \frac{x_s}{x_{req}}}} \tag{4.5}$$

The utility function is the sigmoid function, we chose this specific function in order to define the user satisfaction, which is simply described using Heaviside function, but since the Heaviside function is not differentiable we approximated it using the sigmoid function defined above, user satisfaction can be introduced as following:

$$UserSat = \frac{x_s}{x_{req}} \tag{4.6}$$

User satisfaction is maximized whenever the allocated bandwidth x_s is maximized. x_{req} is a required bandwidth by the user. Sigmoid function is a continuous alternative for a Heavyside function. This means that we count the number of data traffic from the set, where the users were satisfied with the amount of incoming traffic. The parameters δ and β controls the slope of the function. We can assign different parameters for different data traffic types to represent video, voice or text traffic and also to assign the requirement when the data traffic is classified as satisfied by users.

$$\max_{x_s} \sum_{s \in S} g_s(x_s) \tag{4.7}$$

Subject to:

$$f_{ij}^s - h_{ij}^s = x_s, \forall s \in S, i = source(s) \tag{4.8}$$

$$f_{ij}^s - h_{ij}^s = 0, \forall s \in S, \forall (i, j) \in L, i \neq source(s) \tag{4.9}$$

$$\sum_{s \in S} f_{ij}^{\,s} \leq h_{ij}, \forall (i, j) \in L \tag{4.10}$$

The first constraint is a flow conservation law. The second constraint means that the link transmission rate should be larger than the incoming traffic. Where the notations used in the model above are defined in the table 4.2 below:

Table 4.2 Parameters Description

Parameter	Description
$g_s(x_s)$	Utility function
x_s	Allocated bandwidth
S	Set of traffics
f_{ij}^s	Outgoing traffic
h_{ij}^s	Incoming traffic
$\mathrm{L}(x,h,\mu,\gamma)$	Lagrange function
h_{ij}	Maximum link capacity

Using Lagrange relaxation we can solve the problem with respect to the desired variables and by setting $\alpha_s = \mu_{ij}^s$ and *i = source(s).*

$$\mathrm{L}(x,h,\mu,\gamma) = \sum_{s\in S} g_s(x_s) - \sum_{s\in S, i=source(s)} \alpha_s (f_{ij}^s - h_{ij}^s - x_s)$$

$$- \sum_{(i,j)\in L, i\neq source(s)} \sum_{s\in S} \mu_{ij}^s f_{ij}^s + \sum_{(i,j)\in L, i\neq source(s)} \sum_{s\in S} \mu_{ij}^s h_{ij}^s$$

$$- \sum_{(i,j)\in L} \sum_{s\in S} \gamma_{ij} f_{ij}^s + \sum_{(i,j)\in L} \gamma_{ij} h_{ij} \tag{4.11}$$

$$\mathrm{L}(x,h,\mu,\gamma) = \sum_{s\in S} (g_s(x_s) + \alpha_s x_s) + \sum_{(i,j)\in L} \sum_{s\in S} \mu_{ij}^s f_{ij}^s$$

$$- \sum_{(i,j)\in L} \sum_{s\in S} \mu_{ij}^s h_{ij}^s - \sum_{(i,j)\in L} \sum_{s\in S} \gamma_{ij} f_{ij}^s + \sum_{(i,j)\in L} \gamma_{ij} h_{ij}$$

By setting $f_{ij}^s = h_{ij}^s = x_s$, since the routing decision is known by applying VOM prediction algorithm, we get:

$$\mathrm{L}(x,h,\mu,\gamma) = \sum_{s\in S} (g_s(x_s) + \alpha_s x_s) - \sum_{(i,j)\in L} \sum_{s\in S} \gamma_{ij} x_s$$

$$+\sum_{(i,j)\in L}\gamma_{ij}h_{ij} \tag{4.12}$$

Flow control:

$$\frac{\partial \mathrm{L}\left(x,h,\mu,\gamma\right)}{\partial x_s}=0$$

$$0=g_s^{'}(x_s)+\alpha_s-\sum_{(i,j)\in L}\gamma_{ij}$$

$$\xrightarrow{yields}\quad g_s^{'}(x_s)=-\alpha_s+\sum_{(i,j)\in L}\gamma_{ij}$$

$$x_s=g_s^{'-1}(-\alpha_s+\sum_{(i,j)\in L}\gamma_{ij})$$

$$x_s^{*}=\frac{x_{req}\ln\left(-\alpha_s+\dfrac{\alpha_s\beta}{2x_{req}\delta}+\dfrac{\sqrt{-4\beta x_{req}\delta+\beta^2\alpha_s}}{2x_{req}\delta}\right)}{\beta} \tag{4.13}$$

The scheduling part of the Lagrange function reduces to the following problem:

$$\max_h \sum_{(i,j)\in L}\gamma_{ij}h_{ij} \tag{4.14}$$

Algorithm 4.3 Scheduling Algorithm

1. Input: $v_l\, p_l\, R_{ij}$
2. Output: Scheduling decision
3. **At** node i , **at** time t:
4. **Calculate** $w_{ij}(t)= v_l\, p_l\, R_{ij}$
5. **Determine** $l'=argmax_{l\in N(i)}\{w_{ij}(t)\}$
6. **Transmit packets in link** l'

To solve the scheduling part we represent each link as a queue and introduce the quantity $w_{ij}(t)= v_l\ p_l\ R_{ij}$,which weights the links and proceeds the transmission at each node with respect to the maximum weight. Where, R_{ij} is the link rate, P_l is the priority, and $V_l$is the link layer queue size.

4.3 Analysis of the Results

In this section, results are introduced, which shows that our proposed algorithms outperform other algorithms in the literature in terms of end-to-end delay, packet loss and throughput. In our simulation, we have used the same system model shown in Figure 3.4 in the previous chapter. The simulation parameters are also the same as the ones introduced in Table 3.3, except that there is no mobility and the routing of the packets are done based on the proposed algorithm in this chapter. In Figure 4.2, average end-to-end delay is calculated for three algorithms, HWMP (Yang et al., 2009; Bari et al., 2012), ECA-HWMP (Khasawneh et al., 2015), and our proposed algorithm RAECA-HWMP. HWMP is known for its inefficient response to congestion (Yang et al., 2009; Bari et al., 2012). Our algorithm outperforms other algorithms since the congestion is being avoided before it really happens by our VOM prediction model and route the packet through less congested route and also adjusting the transmission rate based on the congestion status. As mentioned earlier, HWMP is not dealing with congestion in an efficient way, there is no prediction or rate adjustment. On the other hand ECA-HWMP, does not have the rate adjustment capability, so our algorithm has an average less delay compared to the other algorithms.

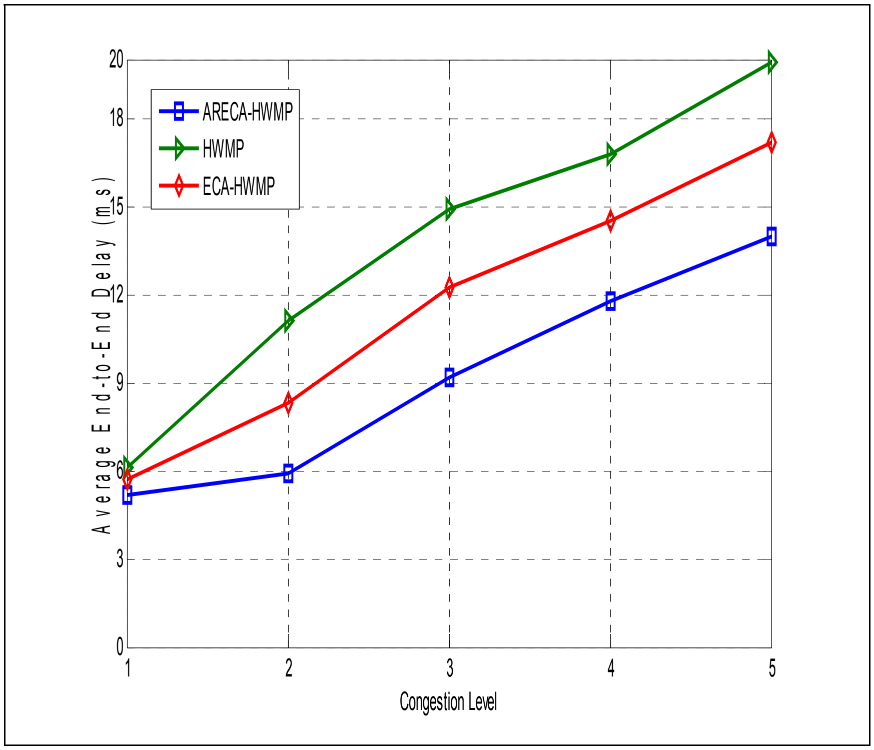

Figure 4.2 Average End-to-End Delay

Figure 4.3 shows the average packet loss, packet loss percentage is less than the other algorithm in the literature. Our algorithm deals better than the other algorithm in the literature in a highly congested situation, by using our proactive approach and adapting the transmission rate based on the predicted congestion status.

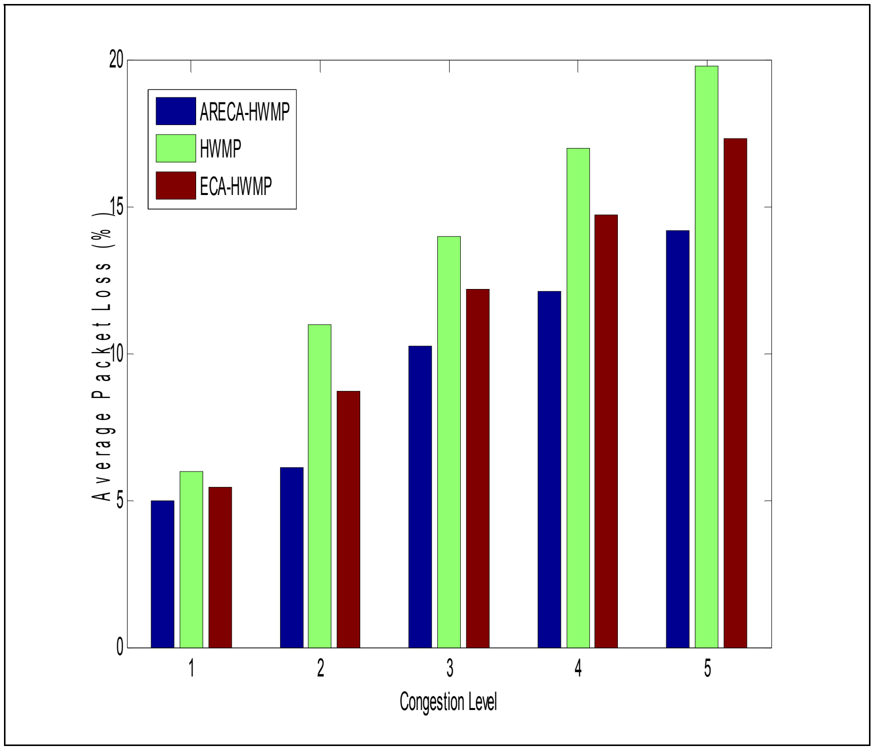

Figure 4.3 Average Packet Loss

Figure 4.6 shows the average user throughput, throughput is maximized after getting the congestion level from our proposed VOM prediction model. The proactive congestion prediction and the adjustment of the allocated bandwidth for each user based on the predicted link congestion status. This makes our proposed algorithm outperform the other algorithm in the literature.

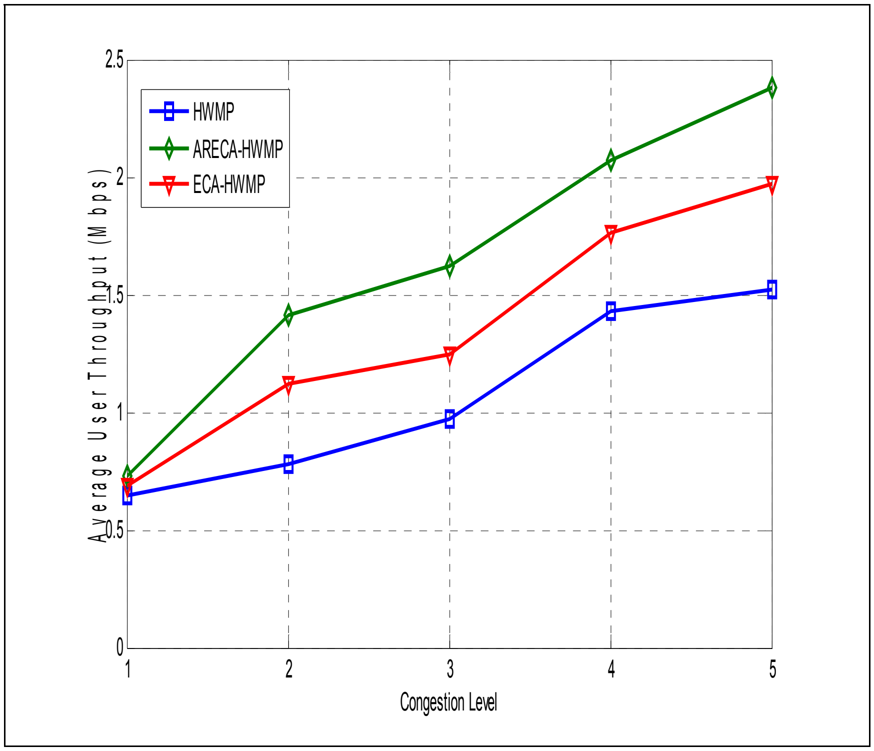

Figure 4.4 Average User Throughput

Our proposed algorithm can be easily applied to handle the handoff problem and predict the congestion before it really happens in the network. In this way, we can avoid the ping pong effect that might be caused in a highly congested network. A better future route can be selected for the handoff call.

4.4 Conclusions

We dealt with traffic congestion in wireless mesh network to improve the Quality of Service (QoS) provided for mesh clients. a novel VOM prediction proactive approach is proposed to predict the congestion status in each link in the network, new route is discovered, and the transmission rate is adjusted based on the link congestion status to maximize the overall user satisfaction. Sub-optimization model is introduced and solved using Lagrange method. Based on the predicted link congestion, rerouting algorithm is implemented in order to assure the load balancing and to mitigate congestion over WMN network. Simulation results show that our proposed algorithm outperforms other algorithms in the literature in terms of throughput, end-to-end delay, and packet loss.

Our proposed algorithm outperforms HWMP (Bari et al., 2012) and ECA-HWMP (Khasawneh et al., 2015) protocol in terms of end-to-end delay, packet loss and throughput. As previously mentioned in our proposed protocol, handoff traffic has higher priority than the new user traffic. This might cause a starvation problem for the new users traffic, (Sheikh et al., 2015) handles this specific problem.

CONCLUSIONS AND FUTURE WORKS

Conclusions

Many industrial applications are nowadays accessing the Internet through WMN. Many of those applications have very strict QoS requirements. A smooth handoff is a compulsory requirement for moving users. In order to provide the smoothness for the moving user and maintain the provided QoS for the users a lot of challenges need to be dealt with such as the minimization of the handoff delay to maintain the continuation of the connection, location management and the packet loss during the handoff process.

Routing and admission control are studied jointly in this thesis, where a novel routing protocol is merged with multi-homing to get better QoS for new and handoff calls in WMN. Handoff algorithms proposed in the literature find a completely disjointed path from the old one, and reject those calls in highly congested networks. Our proposed algorithm introduces a new routing approach for new and handoff calls. It is based on locating a maximally jointed path with the old primary path. The multi-homing feature is triggered when the network is congested. Our proposed algorithm is compared with four different scenarios where multi-homing and joint routing is enabled or disabled. A MDP Markov model is introduced and solved using a value iteration method to find a near optimal solution. Our proposed algorithm, MDACR, outperforms other algorithms proposed in the literature in terms of handoff delay, new and handoff calls blocking probability, and number of hard handoffs, especially in highly congested networks. MDACR provides the mobile user a seamless handoff approach with higher admitted users into the network.

Congestion control is also studied in this thesis, we dealt with traffic congestion in wireless mesh network to improve the Quality of Service (QoS) provided for mesh clients. a novel VOM prediction proactive approach is proposed to predict the congestion status in each link in the network, new route is discovered, and the transmission rate is adjusted based on the link congestion status to maximize the overall user satisfaction. Sub-optimization model is introduced and solved using Lagrange method. Based on the predicted link congestion, rerouting algorithm is implemented in order to assure the load balancing and to mitigate congestion over WMN network. User satisfaction is mathematically defined using Heaviside and Sigmoid function. Simulation results show that our proposed algorithm outperforms other algorithms in the literature in terms of throughput, end-to-end delay, and packet loss.

Future Works

We believe that there will always be a room for improvement and things to be done in the future, some of the potential future research directions are:

- As this thesis concentrate on improving the QoS in WMN, network security is one of the aspects that we should not be overlooked. As many attacks could degrade the QoS provided to the end users. A novel security aware handoff and congestion approaches should be developed.
- Since we are in the era of Internet of Things (IoT), huge traffic is passing through the network, since all the devices are connected together. This will cause a lot of congestion problems and will degrade the QoS and decrease the user satisfaction and the Quality of Experience (QoE), unless a novel algorithms are proposed to solve this kind of problems. Another useful and beneficial research direction would be to concentrate on the data compression, which should be efficient and fast to make the communication easier for the end users.
- Our proposed algorithms could also be adapted to be useful in heterogeneous 5G networks.

APPENDIXI

LIST OF PUBLICATIONS

Journals:

Accepted and Published

Khasawneh, F.A., Benmimoune, A. and Kadoch, M. (2016) "Joint Routing and Admission Control in Wireless Mesh Network,".Int. J. Communications, Network and System Sciences, 9, 311-325. http://dx.doi.org/10.4236/ijcns.2016.98028.

Accepted

Khasawneh, F. A., Kadoch, M. (2017) "Adaptive Transmission Rate Congestion Aware Routing Algorithm in Wireless Mesh Network,". Wireless Personal Communication, Springer.

Submitted

Khasawneh, F.A., Benmimoune, A. and Kadoch, M. (2017) "Quality of Service Provisioning in Wireless Mesh Network,". IEEE Access Journal, IEEE.

Conferences:

Accepted and Published

Khasawneh, F. A., BenMimoune, A., Kadoch, M., Alomari, A. and Al-Khrayshah, M., "Multihoming admission and mobility management in wireless mesh network," *Computer, Information and Telecommunication Systems (CITS), 2015 International Conference on*, Gijon, 2015, pp. 1-5.IEEE.

Khasawneh, F. A., Benmimoune, A., Kadoch, M., and Khasawneh, M. A., "Predictive Congestion Avoidance in Wireless Mesh Network," Future Internet of Things and *Cloud (FiCloud), 2015 3rd International Conference on*, Rome, 2015, pp. 108-112. IEEE.

Khasawneh F. A., BenMimoune A., Kadoch M., Osama S. B., "Intra-Domain Handoff Management Scheme for Wireless Mesh Network," 8th Conference on Circuits, Syst, Signal and Telecom, Tenerife, Spain, 2014.

APPENDIX II

LIST OF PUBLICATIONS IN COLLABORATION WITH OTHER RESEARCHERS

Journals:

Accepted and Published

Abderrahmane BenMimoune, Fawaz A. Khasawneh, Bo Rong, & Michel Kadoch. (2015). Dynamic Joint Resource Allocation and Relay Selection for 5G Multi-hop RelaySystems. Telecommunication Systems, Springer.

Abderrahmane BenMimoune, Fawaz A. Khasawneh, & MichelKadoch. (2015). User Association-Based Joint Admission and Power Control for Heterogeneous Wireless Networks. Wireless Personal Communication, Springer.

Conferences:

Accepted and Published

Addali K. M., BenMimoune A., Khasawneh F. A., Saied A. M. and Kadoch M., "Dual-Backhaul Links in LTE - A Mobile Relay System for High-Speed Railways," *2016 IEEE 4th International Conference on Future Internet of Things and Cloud Workshops (FiCloudW)*, Vienna, 2016, pp. 98-102.

Abderrahmane BenMimoune, Fawaz A Khasawneh, Michel Kadoch, & Bo Rong.(6-10 Dec.2015).Resource Allocation Framework in 5G Multi-hop Relay System. Paper presented at the Global Communications Conference (GLOBECOM), 2015 IEEE.

Abderrahmane BenMimoune, Fawaz A. Khasawneh, Michel Kadoch, Sun Songlin, & RongBo. (21-26 Sept. 2014). Inter-cell handoff performance improvement in LTE-a multihoprelay networks. Paper presented at the Proceedings of the 12th ACM international symposium on Mobility management and wireless access, Montreal, QC, Canada. 2014 IEEE.

Abderrahmane BenMimoune, Fawaz A. Khasawneh, Michel Kadoch. (24-26 Aug. 2015). User Association for HetNet Small Cell Networks. Paper presented at the 3rd International Conference on Future Internet of Things and Cloud, Rome, Italy. 2015 IEEE.

LIST OF BIBLIOGRAPHICAL REFERENCES

Akyildiz, I.F. and Wang, X.D. (2005) A Survey on Wireless Mesh Networks. IEEE Communications Magazine, Vol. 43, Issue 9, pp. S23-S30.

Alnuem, M. A., Mellor, J. E. and Fretwell, R. J., "Explicit Congestion Notification for error discrimination A practical approach to improve TCP performance over wireless networks," *2007 Wireless Telecommunications Symposium*, Pomona, CA, 2007,pp.1-9.doi: 10.1109/WTS.2007.4563297

Amir, Y., Danilov, C., Hilsdale, M., Musˇaloiu-Elefteri, R., and Rivera, N.,"Fast handoff for seamless wireless mesh networks," in MobiSys '06: Proc. 4th international conference on Mobile systems, applications and services. New York, NY, USA: ACM, 2006, pp. 83–95.

Amir, Y. and Danilov C., "Reliable communication in overlay networks," *2003 International Conference on Dependable Systems and Networks, 2003. Proceedings.*, 2003, pp. 511-520.doi: 10.1109/DSN.2003.1209961

Bari, S. M. S., AnwarF. and MasudM. H., "Performance study of hybrid Wireless Mesh Protocol (HWMP) for IEEE 802.11s WLAN mesh networks," *2012 International Conference on Computer and Communication Engineering (ICCCE)*, Kuala Lumpur, 2012,pp.712-716.doi: 10.1109/ICCCE.2012.6271309

Barreto, L. and Sargento, S., "TCP, XCP and RCP in Wireless mesh networks: An evaluation study," *The IEEE symposium on Computers and Communications*, Riccione, Italy, 2010, pp. 351-357.doi: 10.1109/ISCC.2010.5546550

Barreto, L. and Sargento, S., "XCP-Winf and RCP-Winf: Congestion Control Techniques for Wireless Mesh Networks," *2011 IEEE International Conference on Communications(ICC)*, Kyoto, 2011, pp. 1-6.doi: 10.1109/icc.2011.5963464

Begleiter, Ron, Ran El-Yaniv and Golan Yona. (2004) "On prediction using variable order Markov models." Journal of Artificial Intelligence Research, Vol. 22, Issue 1, pp. 385-421.

BenMimoune, A., Khasawneh, F.A., Kadoch, M. and Rong, B. (2015) Resource Allocation Framework in 5G Multi-Hop Relay System. 2015 IEEE Global Communications Conference (GLOBECOM), San Diego, 6-10 December 2015, pp. 1-6.

Bob, O'Hara; Al Petrick, "IEEE 802.11F Inter Access Point Protocol (IAPP)," in IEEE 802.11 Handbook:A Designer's Companion , 1, Wiley-IEEE Standards Association, 2005, pp.217-219doi: 10.1109/9781118098851.ch8

Chaya, Devi, R. , Basavaraju, T. G., A High Throughput cross Layer based Mobility Management for Wireless Mesh Networks, International Journal of Networks and Communications, Vol. 6 No. 4, 2016, pp. 65-71. doi: 10.5923/j.ijnc.20160604.01

Cleary, J. and Witten, I., "Data Compression Using Adaptive Coding and Partial String Matching," in *IEEE Transactions on Communications*, vol. 32, no. 4, pp. 396-402, Apr 1984.doi: 10.1109/TCOM.1984.1096090

De Renesse, R., Friderikos V., and Aghvami A. H., ”Cross-layer cooperation for accurate admission control decisions in mobile ad hoc networks”, IEEE Transactions on Mobile Computing, vol. 7, no. 10, pp. 1228-1241, 2008.

Desheng, Fu, Barbara, Staehle, Rastin Pries, and Dirk Staehle. 2010. On the potential of IEEE 802.11s intra-mesh congestion control. In *Proceedings of the 13th ACM international conference on Modeling, analysis, and simulation of wireless and mobile systems* (MSWIM '10).ACM, New York, NY, USA, 299-306. DOI=http://dx.doi.org/10.1145/1868521.1868569

Dina, Katabi, Mark Handley, and Charlie Rohrs. 2002. Congestion control for high bandwidth-delay product networks. In *Proceedings of the 2002 conference on Applications, technologies, architectures, and protocols for computer communications* (SIGCOMM '02). ACM, New York, NY, USA, 89-102. DOI=http://dx.doi.org/10.1145/633025.633035

Dhurandher, S.K., Woungang, I., Obaidat, M.S., Kumar, K., Joshi, M. and Verma, M. (2015) A Distributed Adaptive Admission Control Scheme for Multimedia Wireless Mesh Networks.IEEE-Systems-Journal, Vol. 9, Issue 2, pp. 595-604. http://dx.doi.org/10.1109/JSYST.2013.2296336

Draves, R., Jitendra, P., and Brian, Z. 2004. Routing in multi-radio, multi-hop wireless mesh networks. In *Proceedings of the 10th annual international conference on Mobile computing and networking* (MobiCom '04). ACM, New York, NY, USA, pp. 114-128. DOI=http://dx.doi.org/10.1145/1023720.1023732

Droms, R., "Dynamic Host Configuration Protocol,"RFC2131, Mar 1997.

Giovanardi, A. and Mazzini, G., "Transparent mobile IP: an approach and implementation," *GLOBECOM 97. IEEE Global Telecommunications Conference. Conference Record*, Phoenix, AZ, 1997, pp. 1861-1865 vol.3. doi: 10.1109/GLOCOM.1997.644594

Gupta, D., Wu D., Mohapatra, P. and Chuah, C. N., "Experimental Comparison of Bandwidth Estimation Tools for Wireless Mesh Networks," *IEEE INFOCOM 2009*, Rio de Janeiro, 2009, pp.2891-2895.doi: 10.1109/INFCOM.2009.5062253

Hu, N. and Steenkiste, P., ”Evaluation and characterization of available bandwidth probing techniques”, Selected Areas in Communications, IEEE Journal on, 2003.

Huang, D. W., Lin, P. and Gan, C. H., "Design and Performance Study for a Mobility Management Mechanism (WMM) Using Location Cache for Wireless Mesh Networks," in *IEEE Transactions on Mobile Computing*, vol. 7, no. 5, pp. 546-556, May2008.doi: 10.1109/TMC.2007.70745

Huang, R.S., Zhang, C. and Fang, Y.G. (2007) A Mobility Management Scheme for Wireless Mesh Networks. IEEE Global Telecommunications Conference, GLOBECOM’07, pp. 5092-5096. http://dx.doi.org/10.1109/glocom.2007.965

Hung, T.C., Phuc, L., Duy, N.N., Quynh, D.N.T., Cuong, L.Q., Jung, B.K. and Kang, Y. (2009) The Proposed Improvement 3-Layer Mobility Management Scheme for Wireless Mesh Networks Using IP Prefix Mechanism. 11th International Conference on Advanced Communication Technology, ICACT 2009, pp. 953-958.

Islam, A. B. M. A. A.and Raghunathan, V., "End-to-end congestion control in wireless mesh networks using a neural network," *2011 IEEE Wireless Communications and Networking-Conference*,Cancun,Quintana-Roo,2011,pp.677-682.doi: 10.1109/WCNC.2011.5779214

Islam, M., Lutfar Rahman, M. and Mamun-Or-Rashid, M., "IL-CD: Interference and load aware congestion detection and bandwidth efficient rate allocation for wireless mesh networks," *The 2014 International Symposium on Networks, Computers and Communications*, Hammamet, 2014, pp. 1-5. doi:10.1109/SNCC.2014.6866530

Islam, M., Rahman, M. L. and Mamun-Or-Rashid, M., "Load adaptive congestion control and rate readjustment for wireless mesh networks," 2014 IEEE 5th International Conference on Software Engineering and Service Science, Beijing, 2014, pp. 1088-1092.doi: 10.1109/ICSESS.2014.6933755

Jesus, A., Fotouhi, H. , Almeida, L. , and Villarroel, J. L., "DoTHa - a Double-Threshold Hand-off algorithm for managing mobility in wireless mesh networks," *2016 IEEE 21st International Conference on Emerging Technologies and Factory Automation (ETFA)*, Berlin, 2016, pp. 1-8.doi: 10.1109/ETFA.2016.7733511

Jianguo, Ma, "Internet-of-Things: Technology evolution and challenges," *2014 IEEE MTT-S International Microwave Symposium (IMS2014)*, Tampa, FL, 2014, pp.1-4. doi: 10.1109/MWSYM.2014.6848429

Kapadia, K. N. and Ambawade, D. D., "Congestion aware load balancing for multiradio Wireless Mesh Network," *2015 International Conference on Communication, Information & Computing Technology (ICCICT)*, Mumbai, 2015, pp. 1-6.doi: 10.1109/ICCICT.2015.7045750

Karthika, K. C, "Wireless mesh network: A survey," *2016 International Conference on Wireless Communications, Signal Processing and Networking (WiSPNET)*, Chennai, 2016, pp.1966-1970.doi: 10.1109/WiSPNET.2016.7566486

Keerthana, P. and ChandraMohan, S., "Adaptive approach based joint scheduling and congestion control in wireless networks," *2015 International Conference on Innovations in Information, Embedded and Communication Systems (ICIIECS)*, Coimbatore, 2015, pp. 1-5.doi: 10.1109/ICIIECS.2015.7193254

Khasawneh, F.A., BenMimoune, A., Kadoch, M. and Osama, S.B. (2014) Intra-Domain Handoff Management Scheme for Wireless Mesh Network. Proceedings of the 12th ACM International Symposium on Mobility Management and Wireless Access, Montreal, 21-26 September 2014, pp. 9-13.

Khasawneh, F.A., BenMimoune, A., Kadoch, M., Alomari, A. and Al-Khrayshah, M. (2015) Multihoming Admission and Mobility Management in Wireless Mesh Network. 2015 International Conference on Computer, Information and Telecommunication Systems (CITS), Gijon, 15-17 July 2015, pp. 1-5. http://dx.doi.org/10.1109/CITS.2015.7297763

Khasawneh, F.A., Benmimoune, A., Kadoch, M. and Khasawneh, M.A. (2015) Predictive Congestion Avoidance in Wireless Mesh Network. 3rd International Conference on Future Internet of Things and Cloud (FICloud), Rome, 24-26, pp.108-112.

Komatsu, J., SakataS., KomuroN., ShiodaS. and MuraseT., "Relay access point congestion control scheme using buffer control for wireless LAN mesh networks," *2013 16th International Symposium on Wireless Personal Multimedia Communications (WPMC)*, Atlantic City, NJ, 2013, pp. 1-5.

Langdon, G., "A note on the Ziv - Lempel model for compressing individual sequences (Corresp.)," in *IEEE Transactions on Information Theory*, vol. 29, no. 2, pp. 284-287, Mar 1983.doi: 10.1109/TIT.1983.1056645

Lee, C. , Park, C., Jang, K., Moon, S. and Han, D., "DX: Latency-Based Congestion Control for Datacenters," in *IEEE/ACM Transactions on Networking*, vol. 25, no. 1, pp. 335-348, Feb. 2017.doi: 10.1109/TNET.2016.2587286

Lee, H. K., Hall V., Yum K. H., Kim K. I., and Kim E. J., "Bandwidth Estimation in Wireless Lans for Multimedia Streaming Services", Advances in Multimedia, vol. 2007, no. 1, January 2007.

Li, H. and Xie, J. (2014) GaS: A Gateway Scheduling-Based Beamforming Scheme for Handoff Support in Wireless Mesh Networks. IEEE Transactions on Vehicular Technology, Vol. 63, Issue 9, pp. 4560-4573.

Li, H. and Xie, J. (2013) GaS: A Gateway Scheduling-Based Handoff Scheme in Single-Radio Infrastructure Wireless Mesh Networks. Proceedings IEEE of INFOCOM, Turin, pp.1860-1868. http://dx.doi.org/10.1109/infcom.2013.6566985

Li, H. and XieJ., "ConT: A contention-based time division scheme for handoff support in single- radio wireless mesh networks," *2012 IEEE Global Communications Conference(GLOBECOM)*,Anaheim,CA,2012,pp.2024-2029.doi:10.1109/ GLOCOM.2012.6503413

Li, Y.N. and Chen, I.-R. (2012) Mobility Management in Wireless Mesh Networks Utilizing Location Routing and Pointer Forwarding. IEEE Transactions on Network and Service Management,Vol. 9, Issue 3, pp. 226-239.

Marwaha, S., Indulska J., and Portmann M., "Challenges and recent advances in QoS provisioning in wireless mesh networks," in Proceedings of the IEEE 8th International Conference on Computer and Information Technology (CIT '08), pp. 618– 623, July 2008.

Ma, X., Xu L. and Min G., "Congestion Control Based on Cross-Layer Game Optimization in Wireless Mesh Networks," *2013 IEEE 9th International Conference on Mobile Ad-hoc and Sensor Networks*, Dalian, 2013, pp. 41-46.doi: 10.1109/MSN.2013.49

Mallikarjuna, Rao Yamarthy, SubramanyamM.V., SatyaK. Prasad, A Multi-layer Routing Protocol for Mobility Management in Wireless Mesh Networks, Procedia Computer Science,Vol. 89,2016,pp. 51-56,ISSN-1877-0509, http://dx.doi.org/10.1016/j.procs.2016.06.008.

Mamidi, A.V., Babu, S. and Manoj, B.S. (2015) Dynamic Multi-Hop Switch Handoffs in Software Defined Wireless Mesh Networks. 2015 IEEE International Conference on Advanced Networks and Telecommunications Systems (ANTS), Kolkata, 15-18 December 2015, pp. 1-6. http://dx.doi.org/10.1109/ANTS.2015.7413638

Majumder, A. and Roy, S. (2012) A Forward Pointer Based Mobility Management Scheme for Multi-Hop Multi-Path Wireless Mesh Network. International Conference on-Data-Science&Engineering(ICDSE),pp. 194-197. http://dx.doi.org/10.1109/icdse.2012.6281899

Majumder, A., Roy, S. and Kumar Dhar, K. (2013) Design and Analysis of an Adaptive Mobility Management Scheme for Handling Internet Traffic in Wireless Mesh Network. Annual International Conference on Emerging Research Areas and International Conference on Microelectronics, Communications and Renewable Energy(AICERA/ICMiCR),pp.1-6.

Majumder, A. and Roy, S. (2014) A Tree Based Mobility Management Scheme for Wireless

Mesh Network. 2014 2nd International Conference on Emerging Technology Trends in Electronics, Communication and Networking (ET2ECN), Surat, 26-27 December 2014, pp. 1-8. http://dx.doi.org/10.1109/et2ecn.2014.7044983

Masri, A., SardoukA., KhoukhiL., HafidA. and GaitiD., "Neighborhood-Aware and Overhead-Free Congestion Control for IEEE 802.11 Wireless Mesh Networks," in *IEEE Transactions on Wireless Communications*, vol. 13, Issue. 10, pp. 5878-5892, Oct. 2014.doi: 10.1109/TWC.2014.2349898

Masri, A. E., Sardouk, A., Khoukhi, L. and Gaiti, D., "An efficient and fair congestion control protocol for IEEE 802.11-based Wireless Mesh Networks," 2012 IEEE 23rd International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC),Sydney, NSW,2012, pp.738-744. doi: 10.1109/PIMRC.2012.6362882

Masri, A., Khoukhi, L., Sardouk, A., Mansouri, M. and Gaiti, D., "WIRS: Resource Reservation and Traffic Regulation for QoS Support in Wireless Mesh Networks," *2011 IEEE Global Telecommunications Conference - GLOBECOM 2011*, Houston, TX, USA, 2011, pp. 1-5.doi: 10.1109/GLOCOM.2011.6133579

Nandita, Dukkipati and Nick McKeown. 2006. Why flow-completion time is the right metric for congestion control. *SIGCOMM Comput. Commun. Rev.* 36, 1 (January 2006), 59-62. DOI=http://dx.doi.org/10.1145/1111322.1111336

Nascimento, V., Gomes, C., Moreira, W. and Abelém, A., "CLM-TCP: TCP Congestion Control Mechanism Using Cross Layer for Wireless Mesh Networks," *2008 14th IEEE International Conference on Parallel and Distributed Systems*, Melbourne, VIC,2008,pp.885-890.doi: 10.1109/ICPADS.2008.34

Navda, V., Kashyap, A. and Das, S. R., "Design and evaluation of iMesh: an infrastructure-mode wireless mesh network," *Sixth IEEE International Symposium on a World of Wireless-Mobile-and-Multimedia-Networks*,2005,pp.164-170.doi: 10.1109/WOWMOM.2005.35

Nisenson, M., Yariv, I., El-Yaniv, R., & Meir, R. (2003). Towards behaviometric security systems: Learning to identify a typist. In The 7th European Conference on Principles and Practice of Knowledge Discovery in Databases

Oriol, D., Nguyen T. M. T. and Pujolle G., "Available bandwidth estimation in Wireless Mesh Networks using neural networks," *2012 Third International Conference on The Network of the Future(NOF)*,Gammarth,2012,pp.1-5.doi: 10.1109/NOF.2012.6463990

Parissidis, G., Karaliopoulos M., Spyropoulos T., Baumann R., and Plattner B.. Routing metrics for wireless mesh networks. In Guide to Wireless Mesh Networks, pages 199–230. Springer London, 2009.

Jain, M. and Dovrolis C., ”Pathload: A measurement tool for end-to-end available bandwidth”, in Passive and Active Measurements workshop (PAM), 2002.

Perkins, C. Ad hoc on demand distance vector (AODV) routing. Internet-Draft, draft-ietf-manet-aodv-04.txt, Oct. 1999.

Perkins, C. "Mobile IP," in *IEEE Communications Magazine*, vol. 35, no. 5, pp. 84-99,May1997.doi: 10.1109/35.592101

Qin, L. and Zhao, D. (2015) Channel Time Allocations and Handoff Management for Fair Throughput in Wireless Mesh Networks. IEEE Transactions on Vehicular Technology, Vol. 64, Issue 1, pp. 315-326.

Qiu, F. and XueY., "Robust Joint Congestion Control and Scheduling for Time-Varying Multi- Hop Wireless Networks With Feedback Delay," in *IEEE Transactions on Wireless-Communications*,vol.13,no.9,pp.5211-5222,Sept.2014.doi: 10.1109/TWC.2014.2339277

Rabia, Ali, and Dr. Fareeha Zafar, “Bandwidth Estimation in Mobile Ad-hoc Network (MANET)”, International Journal of Computer Science, vol. 8, no. 5, 2011, pp. 331-337.

Rangwala, S., Jindal, A., Jang, K. Y., Psounis, K. and Govindan, R., "Neighborhood-Centric Congestion Control for Multihop Wireless Mesh Networks," in *IEEE/ACM Transactions on Networking*, vol. 19, no. 6, pp. 1797-1810, Dec. 2011.doi: 10.1109/TNET.2011.2146272

Ren, M.S., Liu, C., Zhao, H.Z., Zhao, T. and Yan, W. (2007) MEMO: An Applied Wireless Mesh Network with Client Support and Mobility Management. IEEE Global Telecommunications-Conference,GLOBECOM’07,pp. 5075-5079. http://dx.doi.org/10.1109/glocom.2007.962

Ren, J., Xu H. and Qiu Z., "Cross-Layer Congestion Control for Multi-Channel Wireless Mesh Networks," *2009 5th International Conference on Wireless Communications, Networking-and-Mobile-Computing*,Beijing,2009,pp.1-4.doi: 10.1109/WICOM.2009.5303206

Ren, Juan and Qiu, Zhengding, "Centralized quasi-static channel assignment in multi-radio Wireless Mesh Networks," *2008 11th IEEE Singapore International Conference on Communication Systems*, Guangzhou, 2008, pp. 1149-1154. doi: 10.1109/ICCS.2008.4737363

Rethfeldt, M., Danielis P., Beichler B., Konieczek B., Uster F. and TimmermannD., "Evaluating Cross-Layer Cooperation of Congestion and Flow Control in IEEE 802.11s Networks," *2016 IEEE 30th International Conference on Advanced*

Information Networking and Applications (AINA), Crans-Montana, 2016, pp. 181-188.doi: 10.1109/AINA.2016.12

Rezgui, J., Hafid A., Gendreau M. and Rong B., "Congestion-Aware Clique-Based Handoff in Wireless Mesh Networks," *2009 Fifth International Conference on Mobile Ad-hoc and Sensor Networks*, Fujian, 2009, pp. 388-395.doi: 10.1109/MSN.2009.49

Ribeiro, V. J., Riedi R., Baraniuk R., Navratil J., and Cottrell L., ”PathChirp: Efficient available bandwidth estimation for network paths”, in Passive and Active Measurements Conference (PAM), 2003.

Rissanen, J., "A universal data compression system," in *IEEE Transactions on Information Theory*, vol. 29, no. 5, pp. 656-664, Sep 1983.doi: 10.1109/TIT.1983.1056741

Ron, D., Singer, Y., & Tishby, N. (1996). The power of amnesia: Learning probabilistic automata with variable memory length. Machine Learning, Vol. 25, Issue (2–3), pp. 117–149.

Rosenberg, J. , Schulzrinne, H., Camarillo, G., Johnston, A., Peterson, J., Sparks, R. and Schooler, E. SIP: Session Initiation Protocol. RFC 3261, IETF, June 2002.

Sarr, C., Chaudet C., Chelius G., et Lassous I. G., ”Bandwidth Estimation for IEEE 802.11-Based Ad Hoc Networks”, Mobile Computing, IEEE Transactions on, vol. 7, no. 10, pp. 1228 -1241, October 2008.

Sheikh, S., Wolhuter R. and Engelbrecht H. A., "An Adaptive Congestion Control and Fairness Scheduling Strategy for Wireless Mesh Networks," *2015 IEEE Symposium Series on Computational Intelligence*, Cape Town, 2015, pp. 1174-1181.doi: 10.1109/SSCI.2015.169

Shifrin, M. and Cidon I., "C3: Collective congestion control in Multi-Hop Wireless Networks," *2010 Seventh International Conference on Wireless On-demand Network Systems-and-Services(WONS)*,Kranjska-Gora,2010,pp.31-38.doi: 10.1109/WONS.2010.5437137

Song, W. and Fang, X.M. (2006) Routing with Congestion Control Using Cross-Layer Design in Wireless Mesh Network. IET International Conference on Wireless, Mobile and Multimedia Networks, pp. 1-3.

Staehle, B., Bahr M., Fu D. and Staehle D., "Intra-Mesh Congestion Control for IEEE 802.11s Wireless Mesh Networks," *2012 21st International Conference on Computer Communications-and-Networks(ICCCN)*,Munich,2012,pp.1-7.doi: 10.1109/ICCCN.2012.6289277

Strauss, J., Katabi D., and Kaashoek F., ”A measurement study of available bandwidth estimation tools”, in the 3rd ACM SIGCOMM IMC, 2003.

Sun, X.-M. and Lv, X.-Y. (2009) Novel Dynamic Ant Genetic Algorithm for QoS Routing in Wireless Mesh Networks. 5th International Conference on Wireless Communications, Networking and Mobile Computing,WiCom’09, PP. 1-4. http://dx.doi.org/10.1109/wicom.2009.5305851

Sztrik, J., Basic Queueing Theory. University of Debrecen: Faculty of Informatics, 2011, p. 32.

Tjalkens, T., & Willems, F. (1997). Implementing the context-tree weighting method: Arithmetic coding. In International Conference on Combinatorics, Information Theory and Statistics, pp. 83-90.

Volf, P. (2002). Weighting Techniques in Data Compression Theory and Algorithms. Ph.D. thesis, Technische Universiteit Eindhoven.

Wang, H., Huang, Q., Xia, Y., Wu, Y.C. and Yuan, Y.G. (2007) A Network-Based Local Mobility Management Scheme for Wireless Mesh Networks. IEEE Wireless Communications-and-Networking-Conference,WCNC,pp. 3792-3797. http://dx.doi.org/10.1109/wcnc.2007.694

Willems, F. M. J., Shtarkov, Y. M. and Tjalkens, T. J., "The context-tree weighting method: basic properties," in *IEEE Transactions on Information Theory*, vol. 41, no. 3, pp. 653-664, May 1995.doi: 10.1109/18.382012

Xie, J. and Wang, X. (2008) A Survey of Mobility Management in Hybrid Wireless Mesh Networks.IEEE Network, Vol. 22, Issue 6,PP.34-40.

Yang, K., Ma J. f. and Miao, Z. h., "Hybrid Routing Protocol for Wireless Mesh Network," *2009 International Conference on Computational Intelligence and Security*, Beijing, 2009, pp.547-551.doi: 10.1109/CIS.2009.48

Yin, H. , Haibo Ni, Liang Sun, Mingxuan Wang and Xun Zhou, "A caching-list based fast handoff mechanism in wireless mesh networks," 2013 International Conference on-ICT- Convergence(ICTC),Jeju,2013,pp.402-407.doi:10.1109/ICTC.2013.6675382

Zhao, W. and Xie, J. (2012) IMeX: Intergateway Cross-Layer Handoffs in Internet-Based Infrastructure Wireless Mesh Networks. IEEE Transactions on Mobile Computing, Vol. 11, Issue 10, pp. 1585-1600. http://dx.doi.org/10.1109/TMC.2011.192

Zhao, W. and Xie J., "A Novel Xcast-based Caching Architecture for Inter-gateway Handoffs in Infrastructure Wireless Mesh Networks," *2010 Proceedings IEEE INFOCOM*,SanDiego, CA, 2010, pp. 1-9.doi: 10.1109/INFCOM.2010.5462108