Quality of Service Differentiation for Multimedia Delivery in Wireless LANs

by

Zhenhui Yuan

A Dissertation submitted in fulfilment of the requirements for the award of Doctor of Philosophy (Ph.D.)

Dublin City University

School of Electronic Engineering

Supervisor: Dr. Gabriel-Miro Muntean

August, 2012

Declaration

I hereby certify that this material, which I now submit for assessment on the programme of study leading to the award of Ph.D is entirely my own work, that I have exercised reasonable care to ensure that the work is original, and does not to the best of my knowledge breach any law of copyright, and has not been taken from the work of others save and to the extent that such work has been cited and acknowledged within the text of my work.

Signed: _____

Candidate ID No: _____

Date: _____

To my dear parents and to my lovely fiancee

Acknowledgement

First of all, I want to thank to my dear parents. You gave me endless love during the last 26 years. Your attitude towards work, people, and life always inspired me so much. Special thanks to my younger brother, Pengqing Yuan. We have experienced so much happiness together and I am so lucky and proud to be your brother. What an amazing family I have! I love you all!

An extra special acknowledement goes to my dear fiancée, Miss. Xin Zhang. We have been together for five years. We shared, we share and we will share pain and joy, failure and success. I will not forget the hardest time during my first two years in Ireland. 8000 km distance and 8 hours time difference did not break our love. We learned how to bravely handle the difficulties and how to enjoy our current life. We are both getting stronger than ever. Life is tough sometimes, but it is also colourful. It is your deepest love and greatest encouragements that gave me the strength to successfully finish my PhD.

I would like to greatly appreciate the significant support of my principal supervisor, Dr. Gabriel-Miro Muntean. I am and will always be very proud of being your student. It would not have been possible to complete this doctoral thesis without your help, support and patience. There was a time that I really wanted to give up in the first year. I once seriously doubted myself and I did not know if I am qualified enough to continue the PhD study. I still remember that wonderful afternoon when we had a talk in your office. You encouraged me: "Zhenhui, trust yourself! Only the smartest guys in the world can do the PhD and you are already one of them. It is tough and that's also the reason why it is so precious." I am so happy I did not disappoint myself and you. Also, I would like to acknowledge the great help from my co-supervisor, Dr. Hrishikesh Venkataraman, who was invaluable at both academic and personal levels. I appreciate your constructive feedback.

Big thanks to Ramona, Sabine and Bogdan. You gave me many excellent ideas and feedback throughout the years. Everytime I disturbed you, I got surprising help. My heart-full thanks are given to my lovely colleagues, Irina, Aarthy, Jenny, Olga, Yang, Leijla, Seung-Bum, Kevin, Shengyang, Longhao, Ting, Ruiqi, Martin, Ronan, and Quang-Le. It was and is a wonderful time working with you.

Also I want to thank to the technical staff, especially to Robert Clare from School of Electronic Engineering, Dublin City University, Ireland. You are a professional advisor and a caring and warm-hearted person.

Special thanks to my old friends, Haoxuan Li, Wei Zhou, Ye Xiong, Chuanzhen Cai, Shuai Wan, Jiao Tian, and Sha Yu. Thank you for the constructive feedback and great help for my experimental tests. I am so proud of being your friend.

Last but not the least, I gratefully acknowledge Prof. Noel E O'Connor, Prof. Wout Joseph and Prof. Paul Whelan for organizing the PhD defense and giving me very useful comments.

Dublin, June 2012

Zhenhui Yuan

List of Publications

[Journals]

1. Z. Yuan, H. Venkataraman, and G.-M. Muntean, "A Prioritized Adaptive Scheme for Multimedia Services over IEEE 802.11WLANs", submitted to *IEEE Transactions on Multimedia*, Mar. 2012.
2. Z. Yuan, H. Venkataraman, and G.-M. Muntean, "Downlink/Uplink Fairness of VoIP for Mobile Consumer Devices in WLANs", submitted to *IEEE Transactions on Consumer Electronics*, Jan. 2012.
3. Z. Yuan, H. Venkataraman, and G.-M. Muntean, "A Model-based Achievable Bandwidth Estimation for IEEE 802.11 Data Transmissions", *IEEE Transactions on Vehicular Technology*, no. 5, vol. 61, pp. 2158-2171, Jun. 2012.

[Conference Papers]

1. Z. Yuan, H. Venkataraman, and G.-M. Muntean,"A Novel Bandwidth Estimation Algorithm for IEEE 802.11 TCP Data Transmissions", accepted by IEEE Wireless Communications and Networking Conference (WCNC) workshop on Wireless Vehicular Communications and Networks, Paris, France, Apr. 2012.
2. Z. Yuan, H. Venkataraman, and G.-M. Muntean,"iPAS:An User Perceived Quality-based Intelligent Prioritized Adaptive Scheme for IPTV in Wireless Home Networks", IEEE International Symposium on Broadband Multimedia Systems and Broadcasting, Shanghai, China, Mar. 2010, pp.1-6.
3. Z. Yuan, H. Venkataraman, and G.-M. Muntean, "iBE: A Novel Bandwidth Estimation Algorithm for Multimedia Services over IEEE 802.11 Wireless Networks", 12th IFIP/IEEE International Conference on Management of Multimedia and Mobile Networks and Services, Venice, Italy, Oct. 2009, pp. 69-80.
4. Z. Yuan, H. Venkataraman, and G.-M. Muntean,"Necessity for an Intelligent Bandwidth Estimation Technique over Wireless Networks", China-Ireland International Conference on Information and Communication Technologies, Maynooth, Ireland, Aug. 2009.

Abstract

Delivering multimedia content to heterogeneous devices over a variable networking environment while maintaining high quality levels involves many technical challenges. The research reported in this thesis presents a solution for Quality of Service (QoS)-based service differentiation when delivering multimedia content over the wireless LANs. This thesis has three major contributions outlined below:

1. A **Model-based Bandwidth Estimation algorithm (MBE)**, which estimates the available bandwidth based on novel TCP and UDP throughput models over IEEE 802.11 WLANs. MBE has been modelled, implemented, and tested through simulations and real life testing. In comparison with other bandwidth estimation techniques, MBE shows better performance in terms of error rate, overhead, and loss.

2. An **intelligent Prioritized Adaptive Scheme (iPAS)**, which provides QoS service differentiation for multimedia delivery in wireless networks. iPAS assigns dynamic priorities to various streams and determines their bandwidth share by employing a probabilistic approach-which makes use of stereotypes. The total bandwidth to be allocated is estimated using MBE. The priority level of individual stream is variable and dependent on stream-related characteristics and delivery QoS parameters. iPAS can be deployed seamlessly over the original IEEE 802.11 protocols and can be included in the IEEE 802.21 framework in order to optimize the control signal communication. iPAS has been modelled, implemented, and evaluated via simulations. The results demonstrate that iPAS achieves better performance than the equal channel access mechanism over IEEE 802.11 DCF and a service differentiation scheme on top of IEEE 802.11e EDCA, in terms of fairness, throughput, delay, loss, and estimated PSNR. Additionally, both objective and subjective video quality assessment have been performed using a prototype system.

3. A **QoS-based Downlink/Uplink Fairness Scheme**, which uses the stereotypes-based structure to balance the QoS parameters (i.e. throughput, delay, and loss) between downlink and uplink VoIP traffic. The proposed scheme has been modelled and tested through simulations. The results show that, in comparison with other downlink/uplink fairness-oriented solutions, the proposed scheme performs better in terms of VoIP capacity and fairness level between downlink and uplink traffic.

Table of Content

List of Figures

List of Tables

List of Abbreviations

3G: Third Generation Cellular Networks

3GPP: Third Generation Partnership Project

4G: Fourth Generation Cellular Networks

ARF: Auto Rate Fallback

AARF: Adaptive Auto Rate Fallback

AP: Access Point

AMP: Adaptive Media Player

AVC: Advanced Video Coding

BER: Bit Error Rate

BSC: Base Station Controller

BSS: Basic Service Set

BTS: Base Transceiver Station

CAC: Call Admission Control

CBR: Constant Bit Rate

CDMA: Code Division Multiple Access

CSMA/CA: Carrier Sense Multiple Access with Collision Avoidance

CN: Core Network

CSN: Connectivity Service Network

DCCP: Datagram Congestion Control Protocol

DCF: Distributed Coordination Function

DCT: Discrete Cosine Transform

DS: Distribution System

DVB: Digital Video Broadcasting

EDCF: Enhanced Distributed Coordination Function

EDCA: Enhanced distributed channel access

EDGE: Enhanced Data Rates for GSM Evolution

EGPRS: Enhanced GPRS

E-UTRAN: Evolved-UTRAN

FTP: File Transfer Protocol

GOP: Group of Pictures

GPRS: General Packet Radio Service

GSM: Global System for Mobile Communications

HD: High Definition

HSPA: High-Speed Packet Access

HSDPA: High-speed Downlink Packet Access

HSUPA: High-speed Uplink Packet Access

HTTP: Hypertext Transport Protocol

iBE: intelligent Bandwidth Estimation

IEEE: Institute of Electrical and Electronics Engineers

IEC: International Electro technical Commission

IETF: Internet Engineering Task Force

IMS: IP Multimedia Subsystem

IP: Internet Protocol

ISO: International Organization for Standardization

ITU: International Telecommunication Union

JPEG: Joint Photographic Experts Group

LAN: Local Area Network

LTE: Long-Term Evolution

MAC: Medium Access Control

MAN: Metropolitan Area Network

MBE: Model-based Bandwidth Estimation

MH: Mobile Host

MIHF: Media Independent Handover Function

MICS: Media-Independent Command Service

MIES: Media-Independent Event Service

MIIS: Media-Independent Information Service

MIMO: Multiple Input Multiple Output

MIP: Mobile IP

MIPv4: Mobile IPv4

MIPv6: Mobile IPv6

MN: Mobile Node

MOS: Mean Opinion Score

MPEG: Moving Picture Experts Group

OFDMA: Orthogonal Frequency-Division Multiple Access

PAN: Personal Area Networks

PCF: Point Coordination Function

PDA: Personal Digital Assistants

PDF: Policy Decision Function

PDN: Packet Data Network

PEVQ: Perceptual Evaluation of Vide Quality

PGM: Probe Gap Model

PRM: Probe Rate Model

PMIP: Proxy Mobile IP

PMIPv6: Proxy Mobile IPv6

PSNR: Peak-Signal-to-Noise-Ratio

PSTN: Public Switching Telephony Network

QoE: Quality of Experience

QoS: Quality of Service

RAN: Radio Access Network

RAT: Radio Access Technology

RRM: Radio Resource Management

RSS: Received Signal Strength

RTP: Real Time Transport Protocol

RTCP: RTP Control Protocol

RTMP: Real Time Messaging Protocol

RTMFP: Real-Time Media Flow Protocol

RTSP: Real Time Streaming Protocol

RTT: Round-Trip Time

SAMMy: Signal Strength-based Adaptive Multimedia Delivery

SCTP: Stream Control Transmission Protocol Mobile

ST-DEV: Standard Deviation

SIM: Subscriber Identity Module

SINR: Signal to Interference plus Noise Ratio

SIP: Session Initiation Protocol

SIR: Signal-to-Interferences Ratio

SMS: Standard Message Service

SNR: Signal to Noise Ratio

SS: Subscriber Station

SSIM: Structural Similarity Index

STA: Station

TDMA: Time Division, Multiple Access

TD-SCDMA: Time Division Synchronous CDMA

TFRC: TCP-Friendly Rate Control

TCP: Transmit Power Control

UDP: User Datagram Protocol

UMB: Ultra Mobile Broadband

UMTS: Universal Mobile Telecommunications System

UTRAN: UMTS Terrestrial Access Network

UWB: Ultra-Wideband

VoD: Video on Demand

VoIP: Voice over IP

VQM: Video Quality Metric

VSQI: Video Streaming Quality Index

VTQI: Video Telephony Quality Index

WAN: Wide Area Network

WCDMA: Wideband CDMA

Wi-Fi: Wireless Fidelity

WiMAX: Worldwide Interoperability for Microwave Access

WLAN: Wireless Local Area Networks

WMAN: Wireless Metropolitan Area Networks

WP: Weighted Product

WPAN: Wireless Personal Area Networks

WRAN: Wireless Regional Area Network

WSNR: Weighted-Signal-to-Noise-Ratio

WWAN: Wireless Wide Area Networks

CHAPTER 1

Introduction

The first chapter introduces the thesis and presents market trends, background technologies, challenges, and proposed solutions. The chapter starts with the description of the research motivation by analysing the current situation in the market of multimedia delivery, mobile devices, and wireless network technologies. Next section introduces the problem statement in the context of delivering multimedia content to mobile devices over wireless networks. It seems that the significant developments of multimedia applications, wireless networks and wireless devices come with several Quality of Service (QoS)-related issues, including service differentiation, bandwidth resource allocation, etc. Following section lists the primary contributions of the thesis in order to address these challenges.

1.1 Research Motivation

Along with the significant growth of broadband connectivity to home residences, the popularity of multimedia delivery services is also increasing. A study [1] involving 15 to 69 years old people in the US, UK, and Sweden shows that the population that uses the internet daily increased to 90 percent in 2010 from 36 percent in 2002. The study also indicates that people are passionate about using services which include information search, web browsing, e-mails, internet banking, on-line games, movies or TV shows, music, video conferencing, voice telephone, and social networking (e.g. Facebook[1], Twitter[2], Instagram[3], etc).

Fixed broadband has been developed over a long time. Recently, mobile broadband for computers and smartphones has grown significantly. According to the survey from Ericsson[4], the number of mobile broadband subscribers is expected to reach five billion by 2016 [2], as shown in Figure 1-1. The survey shows that mobile data traffic is expected to grow by around 60 percent a year from 2011 to 2016. Specifically, mobile broadband

[1] Facebook-http://www.facebook.com
[2] Twitter-http://www.twitter.com
[3] Instagram-http://www.instagram.com
[4] Ericsson Co. Ltd-http://www.ericsson.com

subscribers using video communication services are expected to be the primary drivers of the significant increase in mobile data traffic. Meanwhile, the number and type of devices receiving multimedia content over wireless networks has also increased significantly. These wireless networking-enabled devices, mostly mobile and hand-held, are highly heterogeneous in terms of processing capabilities, screen resolution, battery power, memory, etc. The Cisco white paper, *Cisco Visual Networking Index: Global Mobile Data Traffic Forecast for 2011 to 2016*[5] presents how laptops and smartphones will continue to generate the majority of traffic until 2016. Newer devices such as tablet PCs will start to generate a more significant portion of the traffic by 2016. For instance, it is expected that mobile-connected tablet PCs will generate almost as much traffic in 2016 as the entire global mobile network in 2012.

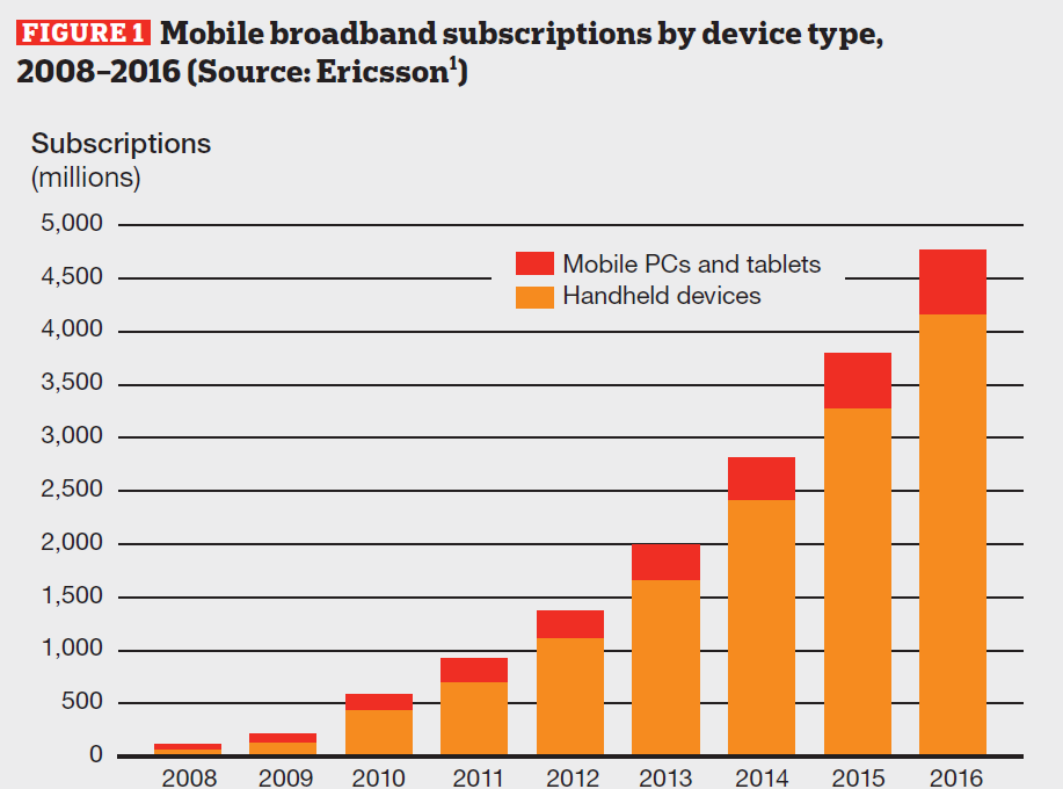


Figure 1-1Mobile broadband subscriptions by device type (source: Ericsson [2])

In terms of the network technologies, multimedia services can be delivered to the heterogeneous devices via different wireless networks, such as: Global System for Mobile Communications (GSM), Universal Mobile Telecommunications System (UMTS), High

[5] Cisco White Paper on global mobile data traffic, Feb 14, 2012- http://www.cisco.com/en/US/solutions/collateral/ns341/ns525/ns537/ns705/ns827/white_paper_c11-520862.pdf

Speed Packet Access (HSPA), Long Term Evolution (LTE), Worldwide Interoperability for Microwave Access (WiMAX), Wireless Personal Area Network (WPAN), Wireless Local Area Networks (WLAN), etc. Specifically, WLAN technologies such as IEEE 802.11 (or Wi-Fi) have been widely deployed due to the low cost, simplicity and convenience. According to the survey from *Informa Telecoms & Media*, growth of Wi-Fi in public spaces is expected to continue and the number of Wi-Fi hotspots will reach 5.8 million worldwide in 2015[6].

In conclusion, there is a continuous development and growth in the applications, devices, and networks for providing rich media services to increasingly large user population.

1.2 Problem Statement

The continuing growth of wireless devices and multimedia services create challenges for providing Quality of Service (QoS) support to users. The devices and internet connection (mobile and fixed) preferred by users depend on the context when availing from a specific service. There are several factors that influence users' decision: 1) **Time to access the content**. For instance, smartphone is preferred when it comes to take a picture and upload it to Facebook; 2) **Mobility**. Smartphone may be considered better than the laptop in terms of the mobility; 3) **Security**. When accessing the personal internet bank or shopping with a credit card, a laptop with a fixed line is preferred to a mobile deive over a wireless network, as the latter suffers in terms of security; 4) **Quality of Experience**. For instance, a movie played on a big screen device provides better viewing experience than on a smartphone; 5) **Context**. The user's decisions are also impacted by geographical factors, i.e., at home, at work, or waiting for a train; social factors, e.g. alone, with friends; service factors, e.g. car navigation, video on demand, video conferencing.

Figure 1-2 presents a common scenario inspired from the home wireless LAN (WLAN). A single IEEE 802.11 wireless router provides broadband services to multiple devices: XBOX[7], laptops, tablet PCs, and smartphones. Considering three scenarios in this home WLAN:

[6] White Paper on Wi-Fi CERTIFIED Passpoint™-
http://www.wi-fi.org/sites/default/files/uploads/20120229_wp_Wi-Fi_CERTIFIED_Passpoint.pdf
[7] XBOX player-http://www.xbox.com

1) The family members all watch video delivered via the wireless network. The kids are watching cartoons using the tablet PC, mom is watching a dance show from *Youtube*[8] using the laptop, and dad is watching living soccer game using his smartphone;

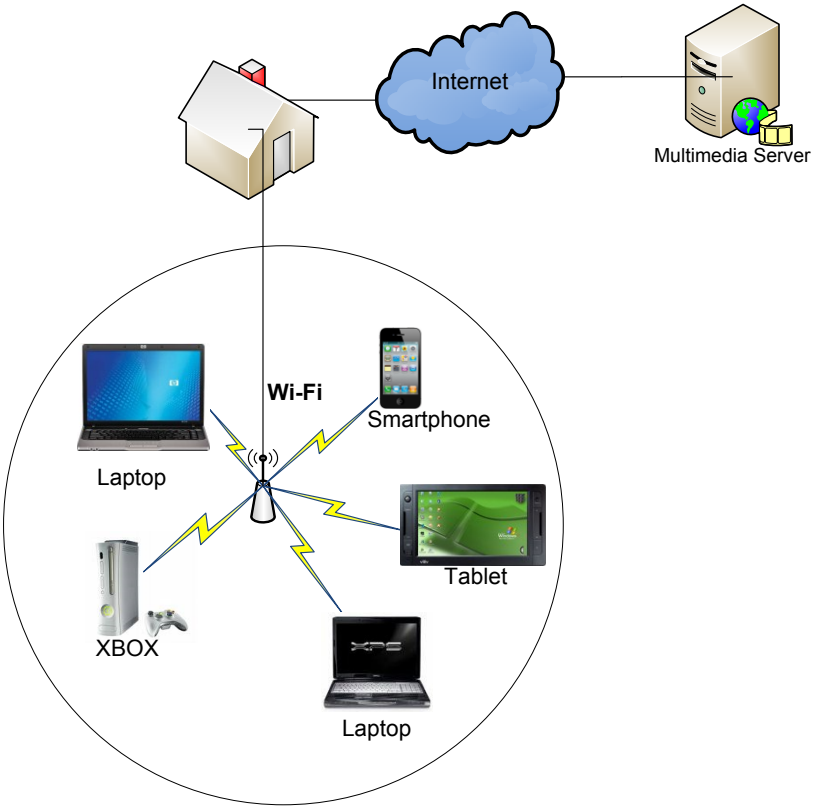


Figure 1-2 Home wireless environments with heterogeneous devices

2) The family members receive different multimedia services via the wireless network. The kids are playing an on-line video game via XBOX, mom is watching a video from *Youtube* using the laptop, dad is updating his *Facebook* profile with his smartphone;

3) Mom is having a Skype[9] video chat with friends and dad is having an important meeting using VoIP via Skype.

In the case of the first and second scenarios, the **downlink** traffic dominates the WLAN, while in the third scenario, the **downlink** and **uplink** traffic competes for the wireless channel. Current IEEE 802.11 networks can provide high speed access, i.e., up to 54Mpbs and 600Mbps for IEEE 802.11g and IEEE 802.11n, separately. However, the limitations of the original IEEE 802.11 protocols might affect the QoS in the three scenarios.

[8] Youtube-http://www.youtube.com
[9] Skype-http://www.skype.com

In the first scenario, video content is delivered to the laptop, tablet PC, and smartphone via the same wireless access router. The laptop might need higher bandwidth allocation than the tablet PC and the smartphone, due to more powerful data processing capability and larger screen resolution. In the second scenario, interactive video service (on-line game), video on demand service (*Youtube*), and best-effort service (*Facebook*) are delivered. There is a need to give the interactive video application users higher bandwidth than that for the best-effort service users, since video applications are more sensitive to delay and loss. However, according to original IEEE 802.11 protocols, the wireless bandwidth is equally shared by all downlink traffic. In the third scenario, interactive services (video chat and voice over IP) are delivered in the IEEE 802.11 WLAN. It is necessary to allow fair wireless channel access between downlink and uplink traffic in order to satisfy the end users. Nevertheless, according to original IEEE 802.11 protocols, the downlink flows obtain less channel access opportunity than the uplink flows due to the inherent contention mechanism of CSMA/CA. Consequently, the downlink traffic has lower priority in accessing the channel, despite much of the traffic being downlink and not uplink.

In this context, there is a need to develop an intelligent **QoS differentiation** solution when delivering multimedia content to different devices over the IEEE 802.11 networks. Such a QoS differentiation solution is expected to involve the following aspects:

1) **Device characteristics awareness**. For instance, when delivering video sequences to a laptop and smartphone which are connected to the same IEEE 802.11 access point, the laptop should be allocated higher bandwidth share than that of the smartphone;

2) **Multimedia services awareness.** For instance, bandwidth share allocated to the real-time traffic flows (i.e. video on demand, on-line gaming, streaming audio, etc) are expected to be higher than the best-effort or background traffic flows (i.e. e-mail service, web-browsing, etc);

3) **Fair traffic distribution between downlink and uplink**. For instance, when delivering VoIP services in the IEEE 802.11 networks, the downlink and uplink traffic flows are expected to achieve fair network access in terms of throughput, delay, and loss.

1.3 Contributions

The research work presented in this thesis contribute with the following issues to provide QoS differentiation for **last-mile** broadband access based on IEEE 802.11 networks, as illustrated in Figure 1-2.

Proposal of a novel **Model-based Bandwidth Estimation algorithm (MBE)** which:

- introduces novel TCP/UDP wireless traffic models;
- estimates the bandwidth based on the TCP/UDP throughput wireless throughput models;
- does not use the probing traffic and does not need modifications of IEEE 802.11 MAC protocol;
- provides good bandwidth estimation results for content delivery in conditions with different packet sizes, dynamic wireless link rate and different channel noise;
- provides lower overhead and lower error rate than other state-of-the-art bandwidth estimation techniques;
- extends a previous proposed bandwidth estimation algorithm, intelligent Bandwidth Estimation (iBE).

Proposal of an **intelligent Prioritized Adaptive Scheme (iPAS)** which:

- provides QoS differentiation for multimedia delivery in wireless networks;
- assigns dynamic priorities to streams and determines their bandwidth share by employing a probabilistic approach-which makes use of stereotypes. The priority level of individual streams is variable and dependent on stream-related characteristics (i.e. device resolution, device battery power left, and application type) and network delivery QoS parameters (i.e. delay, jitter, and packet loss rate);
- utilizes the estimated overall bandwidth based on MBE to allocate or re-allocate bandwidth based on stream's priority;
- has been evaluated based on a subjective video quality assessment, where 32 users are invited to rate the quality of the received video sequences in terms of Mean Opinion Score (MOS), continuity, blurred, and blockness.

Proposal of a **QoS-based downlink/uplink fairness scheme for VoIP in IEEE 802.11 networks** which:

- dynamically controls the contention window size at the wireless access point in order

to balance the downlink/uplink channel access opportunity;

- computes the optimum contention window size of the wireless access point based on the results of a stereotypes-based algorithm which utilizes the ratio between the QoS parameter values (i.e. throughput, delay, and loss) measured for the downlink and uplink traffic;
- improves the VoIP capacity of the IEEE 802.11 network;
- improves the fairness between downlink and uplink VoIP traffic, in terms of throughput, delay, and packet loss ratio.

1.4 Outline of the thesis

The thesis is scheduled as follows:

- **Chapter 1**- Introduction of the thesis including: research motivation, problem statement, contributions, and thesis outline.
- **Chapter 2**- Describes the background technologies of the thesis.
- **Chapter 3**- Describes the current research works regarding the following areas: bandwidth estimation, QoS-based multimedia delivery solutions, and mathematical theories in resource management.
- **Chapter 4**- Presents the architecture and overview of the proposed solutions.
- **Chapter 5**- Describes the proposed bandwidth estimation scheme, iBE and MBE.
- **Chapter 6**- Presents the principle of stereotype-based resource allocation, the architecture of iPAS.
- **Chapter 7**- Presents the QoS-based fairness scheme which provides fair channel access between the downlink and uplink VoIP traffic.
- **Chapter 8**- Experimental evaluation of MBE.
- **Chapter 9**- Experimental evaluation of iPAS
- **Chapter 10**- Experimental evaluation of the proposed QoS-based downlink/uplink fairness scheme for VoIP.
- **Chapter 11**- Presents the prototyping and result analysis of the objective and subjective video quality assessment for iPAS.
- **Chapter 12**- Presents the conclusions and future works of the thesis.
- **Appendix**- Presents the subjective test instructions and questionnaires.

CHAPTER 2

Background Technologies

This chapter introduces the background technologies related to our contributions: 1) intelligent bandwidth estimation (iBE) algorithm and model-based bandwidth estimation (MBE) algorithm; 2) intelligent prioritized adaptive scheme (iPAS); 3) IEEE 802.11 downlink/uplink fairness solution. These proposed solutions provide Quality of Service (QoS)/Quality of Experience (QoE) support for multimedia services in IEEE 802.11 networks. The concept of Radio Resource Management is presented here. Section 2.2 presents the evolution of cellular networks and briefly describes the widely used cellular techniques, e.g. GSM, UMTS, LTE, IMT-Advanced, and femtocell based network. Section 2.3 describes the mostly used broadband networks including IEEE 802.11, IEEE 802.15, IEEE 802.16, and IEEE 802.21. Additionally, QoS and QoE are discussed in Section 2.4. Section 2.5 introduces the functional characteristics and requirements of multimedia streaming systems.

2.1 Radio Resource Management

Radio resource management (RRM) refers to the control of radio transmission related parameters such as channel allocation, network selection, error encoding, etc [3]. The object of RRM is to improve the utilization efficiency of limited radio spectrum resources in wireless communication systems. According to the *Cisco Wireless LAN Controller Configuration Guide Release 7.0*[10], the RRM functions are embedded as software (i.e. *Cisco 5500 Series Wireless LAN Controllers*) in the IEEE 802.11 WLANs controller to provide real-time radio frequency management of the underlying wireless network. RRM periodically configures the wireless network for best efficiency and performs five functions as illustrated in Figure 2-1: Radio resource monitoring, Dynamic channel control, Transmit power control, Admission control, and Packet scheduling scheme.

[10] "Cisco Wireless LAN Controller Configuration Guide, Software Release 7.0"- http://www.cisco.com/en/US/docs/wireless/controller/7.0/configuration/guide/c70cg.pdf, June 2010.

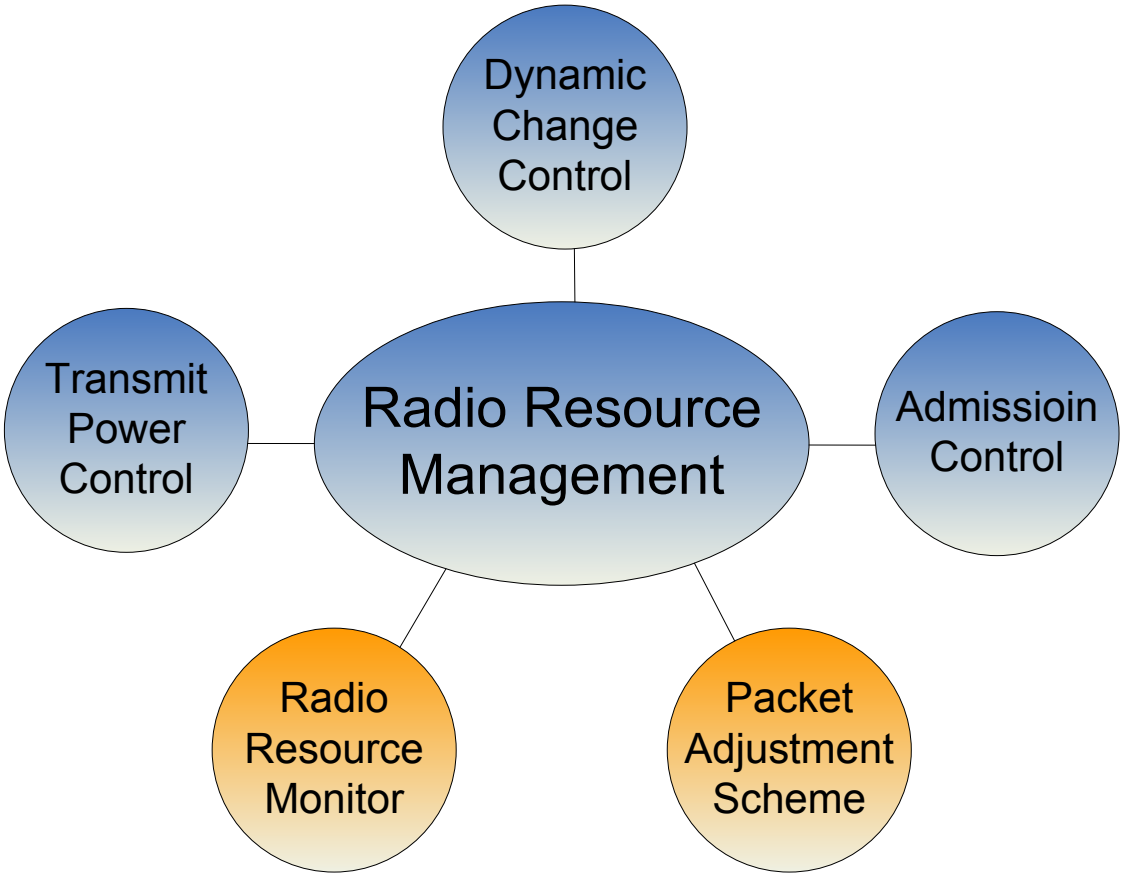


Figure 2-1 Block architecture of the Radio Resource Management

Radio resource monitoring function collects the wireless channel related information:

- Traffic load-measuring total bandwidth used (downlink and uplink) and **predicting the network growth ahead of client demand**.
- Interferences-crossing traffic coming from neighboring networks.
- Noise-the non-802.11 traffic that is interfering with the current channel. For example, access points go "off-channel" for a period not greater than 60ms to monitor all the channels for noise and interferences.
- Coverage-detecting the Received Signal Strength (RSSI) and Signal-to-Noise Ratio (SNR) for connected clients.

The above information can be utilized by the other four functions:

Transmit power control provides the ability to dynamically control the transmit power based on network conditions. In most instances, the transmit power is decreased to reduce interference and obtain more battery capacity (i.e. outdoor mobile access point).

Admission Control aims to limit the total traffic load of networks in order to maintain the QoS of existing traffic. Admission control is critical for real-time traffic such as VoIP and video conference which are vulnerable to fluctuating network conditions.

Dynamic Channel Control can be used to reduce the channel interference between adjacent networks or improve the network utilization. Take IEEE 802.11 network for instance, if the detected interference levels within one network exceed a predefined threshold, the RRM algorithm is used to rearrange the channel assignment. Additionally, RRM mechanism can set the adjacent IEEE 802.11 access points to use different channels, which increases network usage.

Packet Adjustment Scheme is one of the RRM functions to coordinate the sharing of resources. Extensive packet adjustment-based techniques have been developed at different OSI layers. For instance, at link layer, the packet scheduling scheme depends on the type of scheduling policies. Detailed introductions of packet adjustment schemes are presented in Chapter 3.

This thesis focuses on two critical issues of Radio Resource Management, **Radio Resource Monitoring** and **Packet Adjustment Scheme**. Specifically, two wireless bandwidth estimation schemes are proposed to monitor the available radio resource. An intelligent prioritized adaptive scheme and downlink/uplink fairness-oriented adaptive scheme are designed using packet adjustment-based techniques. Details of the proposed algorithms are discussed in Chapters 4, 5, and 6.

2.2 Cellular Networks

Nowadays, cellular networks have been widely implemented in mobile telephone systems. As shown in Fig. 2-2, a typical cellular network consists of a number of cells and each of them is served by at least one fixed-location transmitter-receiver, known as a “Base Station” (or Base Transceiver Station). Each cell is surrounded by 6 neighbouring cells. To avoid the interference, adjacent cells cannot use the same frequency. In practice, two cells using the same frequency must be separated by a distance of two or three times the diameter of the cell. As indicated in Figure 2-2, each colour represents a certain frequency and same colour cell are separated by more than two other coloured cells.

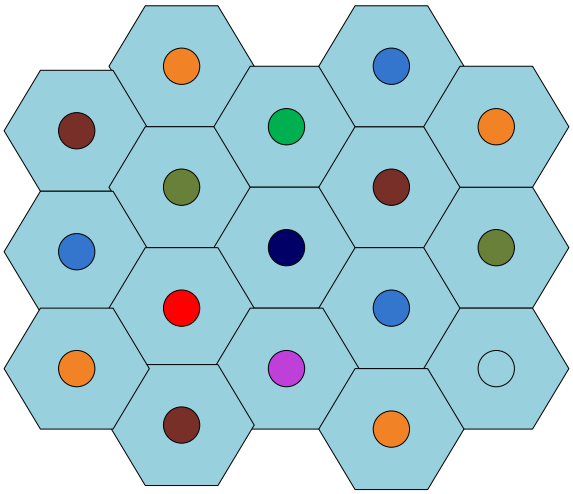


Figure 2-2 Basic Architecture of the Cellular Network

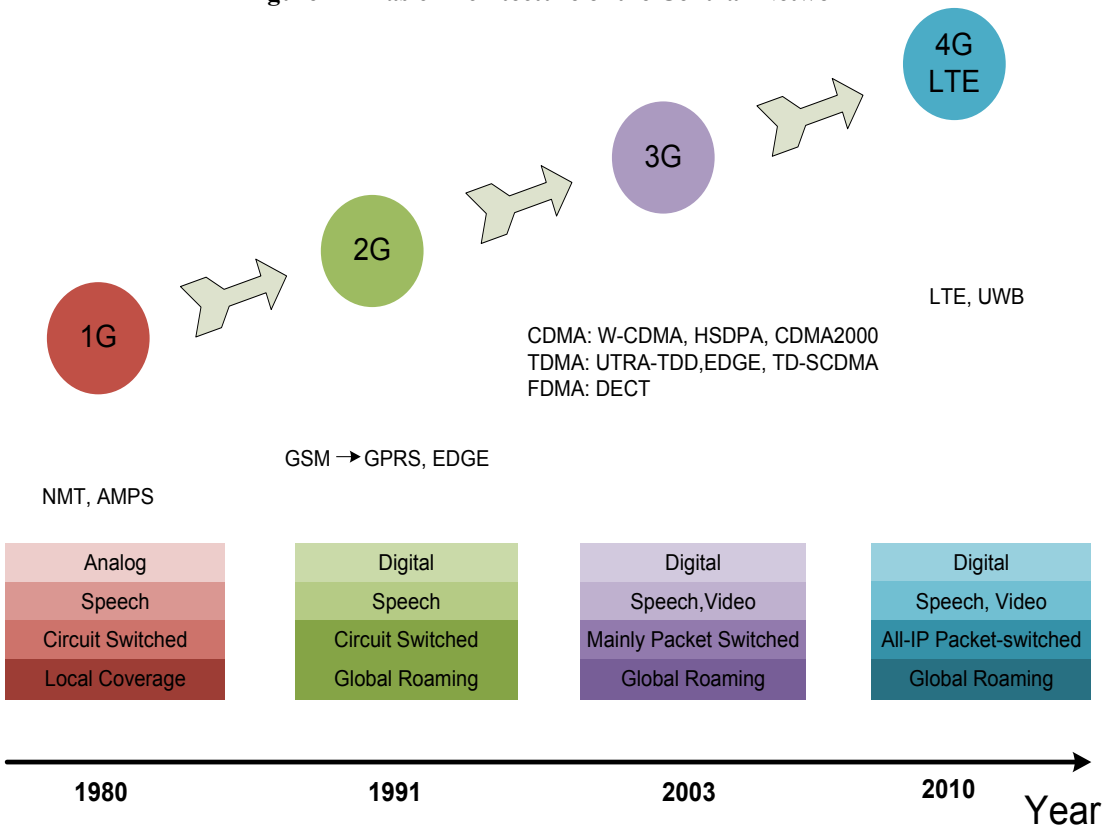


Figure 2-3 Evolution of the Cellular Network

The 3rd Generation Partnership Project[11](3GPP) organization has released a series of cellular communication standards during the last two decades which are summarized in Figure 2-3. There are four generations of cellular networks. The first generation network

[11] 3GPP-http://www.3gpp.org/.

(1980s), was developed for the telephone system using the analogue technique. The first generation phones support voice call only and signals were transmitted using frequency modulation. For instance, 25MHz frequency band might be allocated for both downlink and uplink communications. These bands are further split into a number of channels with a spacing of 30KHz. The first generation cellular network is circuit switched which means that the connection is maintained until the client (phone) stop the service. The second generation (2G) cellular network was launched on the *Global System for Mobile Communications (GSM)* standard [4] in Finland in 1991. Unlike the first generation network where radio signals are analogue, radio signals in the 2G network are delivered in digital mode. Many standards have been developed to improve the original 2G networks. For instance, *General Packet Radio Service (GPRS or 2.5G)* [5] provides data rates of up to 56-114 Kbps. *Enhanced Data Rates for GSM Evolution (EDGE or 2.75G)* [6] further improves the bandwidth to up to 236.8 Kbps with end-to-end latency of less than 150 ms. The third generation cellular network (3G), firstly offered in 2000s, refers to a series of standards for mobile telecommunications such as *Universal Mobile Telecommunications System (UMTS)* [7], *CDMA2000* [8], and the non-mobile wireless standards like *Digital European Cordless Telecommunication (DECT)* [9]. In comparison with first and second networks which are mainly designed for voice services and slow data communications, 3G aims to provide support for high-speed access to multimedia services. In ideal conditions, 3G networks can provide downlink data rates of up to 14.4 Mbps and uplink data rates of up to 5.8Mbps [10]. The fourth generation network (4G) is a successor to the current 2G/3G networks and is designed to provide support for comprehensive all-IP packet switch networks. Many 4G candidate systems have been designed and some already commercialized, such as *Worldwide Interoperability for Microwave Access (WiMAX)* or IEEE 802.16 [11], *Long Term Evolution Advanced (LTE Advanced)* [12], etc. The motivation of 4G is to improve the Quality of Service (QoS) and satisfy the high bandwidth requirements of multimedia services such as mobile broadband internet, mobile TV, etc. Theoretically, 4G networks support data rates of up to 1Gbps.

Table 2-1 summarized the best known cellular network standards. Following sections briefly describe the most widely deployed technologies of cellular networks in each generation 2G (GSM), 3G (UMTS), 3.75G (LTE) and 4G, respectively.

TABLE 2-1 SUMMARY OF CELLULAR NETWORKS[12]

Cellular Technology	Organization	Frequency bands (areas dependent)	Multiple Access	Data transfer rate		Latency
				Downlink	Uplink	
GSM	ITU	380MHz~1900MHz	TDMA FDMA	9.6Kbps	9.6Kbps	450ms
WCDMA (UMTS)	3GPP Release99	1885MHz~2025MHz 2110MHz~2200MHz	W-CDMA TDD TD-SCDMA	384Kbps	128Kbps	150ms
EDGE	3GPP Release7	GSM900 and GSM1800MHz	TDMA	472Kbps	70Kbps	100ms
HSPA/HSDPA /HSUPA	3GPP Release 5/6	2100MHz~900MHz	CDMA	14Mbps	5.7Mbps	100ms
LTE	3GPP Release 8	700MHz~1800MHz	OFDM SC-FDMA	300Mbps	50Mbps	10ms
LTE-Advanced	3GPP Release 10	700MHz~1800MHz	OFDM SC-FDMA	1Gbps	500Mbps	<5ms

2.2.1 GSM (2G)

GSM was standardized by the *European Telecommunications Standards Institute (ETSI)* [13] in 1989 and was seen as the replacement for the 1G analogue networks. GSM uses digital cellular technology to transmit voice and data services. Terrestrial GSM networks now cover more than 90% of the world's population representing over 4 billion individual subscribers.

GSM networks operate at different carrier frequency bands depending on different areas. For instance, in Europe, Asia, Middle East and Africa, most GSM networks use the 900 MHz or 1800 MHz bands; in Canada and the United States, the 850 MHz and 1900 MHz bands are used instead. In some other countries such as Tanzania, the 400MHz and 450 MHz frequency bands are supported. Take GSM-900MHz for instance, it provides 124 channels spaced at 200 KHz and each channel supports data rates around 270Kbps. Specifically, the 890MHz-915MHz frequency bands are designed to deliver the uplink information (i.e. from mobile devices to the base stations) and the 925MHz-960MHz frequency bands are used to send downlink information (i.e. from the base stations to the mobile devices).

[12] Statistics refer to 3GPP website-http://www.3gpp.org/specifications and GSM website-http://www.gsma.com
[13] ETSI-http://www.etsi.org

GSM technologies can be classified based on the access methods employed: *Time Division Multiple Access (TDMA)*, *Frequency Division Multiple Access (FDMA)* and *Code Division Multiple Accesses (CDMA)*. The term *Multiple Access* refers to the capability of allowing multiple transmitters to send information simultaneously over a single channel. *TDMA* allows several transmitters share the same frequency channel by dividing the access using time slots [4]. Similarly, *FDMA* divides the channel access using frequency [4]. *CDMA* is a form of spread-spectrum signalling and the information is delivered only between senders and receivers using the same coding scheme [4].

2.2.2 UMTS (3G)

UMTS [7] is a 3G system standardised by 3GPP along with other regional standards organisations. The motivation of UMTS is to provide high quality of voice and data services. UMTS provides data transfer rate of up to 384Kbps which is significantly faster than the 9.6Kbps of a single GSM channel. The first commercial usage of UMTS launched in 2002 with highlight applications such as mobile TV and video calling. UMTS system has been upgraded to *High Speed Downlink Packet Access (HSDPA)* [13] (3.5G) in many countries. *HSDPA* provides downlink rates of up to 21Mbps. Additionally, UMTS standard also specifies a protocol to improve the uplink transfer performance, *High Speed Uplink Packet Access (HSUPA)* [14], which supports uplink data transfer rates of up to 5.76Mbps.

UMTS presents a comprehensive network framework, consisting of *UMTS Terrestrial Radio Access (UTRA or terrestrial air interfaces)*, *UMTS Radio Access Network (UTRAN)* and *Core Network (CN)*. *UTRA* defines three types of air interfaces: *Wideband Code Division Multiple Access (W-CDMA)*, *Time Division Duplex High Chip Rate (TDD HCR),* and *Time Division Synchronous Code Division Multiple Access (TD-SCDMA)*. *W-CDMA*, maintained by 3GPP, uses the *Direct Sequence Code Division Multiple Access (DS-CDMA)* channel access method. It runs in a paired 5MHz bandwidth (separate frequency for downlink and uplink channel) and is capable of interworking with the *GSM* network. *W-CDMA* is criticized for the large spectrum usage. *TDD HCR* is standardized by 3GPP based on the combination of *TDMA* and *CDMA* (also known as *TD-CDMA*). In comparison with *W-CDMA*, it does not require separate frequency spectrum for downlink and uplink services, which benefits the deployment in tight frequency bands. *TD-SCDMA* utilizes the *TDMA* channel access method and the synchronous *CDMA* on the 1.6MHz carrier frequency bandwidth. *TD-SCDMA* project is funded by the government of People's Republic of China in order to avoid the dependence on the existing cellular technologies. *UTRAN* allows the

connection between user devices and the *Core Network*. UMTS and GSM can share a *Core Network* and they are often combined together according to the available coverage and service requirements.

2.2.3 LTE (3.75G)

The Long Term Evolution (LTE) [12] seen as 3.75G cellular network, is a wireless communication standard that evolved from GSM and UMTS. LTE is mainly built to significantly improve the capacity and data rates of cellular networks. Highlighted features of LTE includes an all-IP network architecture, QoS provisioning, peak download speed of 300Mbps and upload speed of 75Mbps, capacity exceeding 200 active users per 5MHz cell[14], wide range of bandwidth flexibility (i.e., 1.4MHz, 3MHz, 5MHz, 19MHz, 15MHz, and 20MHz), high mobile speed support to up to 500km/h, etc.

LTE uses *Orthogonal Frequency Division Multiplexing (OFDM)* and *Single Carrier Frequency Division Multiple Access (SC-FDMA)* as downlink and uplink multiple access methods, respectively. LTE supports both *Frequency Division Duplex (FDD)* and *Time Division Duplex (TDD)* multiplexing and scalable channel width of up to 20MHz. The Multiple Input Multiple Output (MIMO) technology is officially deployed for LTE antennas, providing downlink data rates of up to 100Mbps and uplink data rates of up to 50Mbps. LTE also enables seamless handoff to cell towers operating earlier cellular technologies such as GSM, CDMA2000, TD-SCDMA, WCDMA, HSDPA, etc.

The next stop of LTE is called LTE Advanced [15], as specified in 3GPP Release 10, which introduces new features such as carrier aggregation, downlink and uplink spatial multiplexing, etc.

2.2.3.1 Femtocell

According to the latest survey from Informa Telecoms & Media[15], eight out of the top ten mobile operator groups offer femtocell services. In June 2011, there are 31 commercial services and a total of 43 deployment committes. Current cellular networks suffer from limited wireless capacity and poor indoor coverage. Femtocell technology improves system capacity by reducing distance between transmitter and receiver. These problems can be alleviated by deploying the femtocell access points, as shown in Figure 2-4. A femtocell is a small user-installed cellular base station, typically deployed in home or company in order to

[14] "Evolution of LTE,", LTE World-http://www.lteworld.org
[15] Informa Telecoms & Media-http://www.informatandm.com/section/home-page

improve the indoor coverage and cellular signal strength. It is compatible with current cellular network technologies such as GSM, Wi-MAX and LTE. Femtocells are connected to the service operator's network through a broadband connection such as Digital Subscriber Line (DSL), television cable modem and fiber optic access. A typical femtocell station can provide coverage around 10 meters. *Berg Insight*[16] group estimates that the shipments of Femtocells will reach 12 million units in 2014. In summary, advantages of femtocell in comparison with traditional cellular technologies include: 1) higher capacity since more users can be served in the same channel due to reduced interferences; 2) higher Signal-to-Noise Ratio (SNR) as the distance between femtocell and clients is reduced; 3) lower cost because femtocell deployments require low operating and maintenance cost from the service provider.

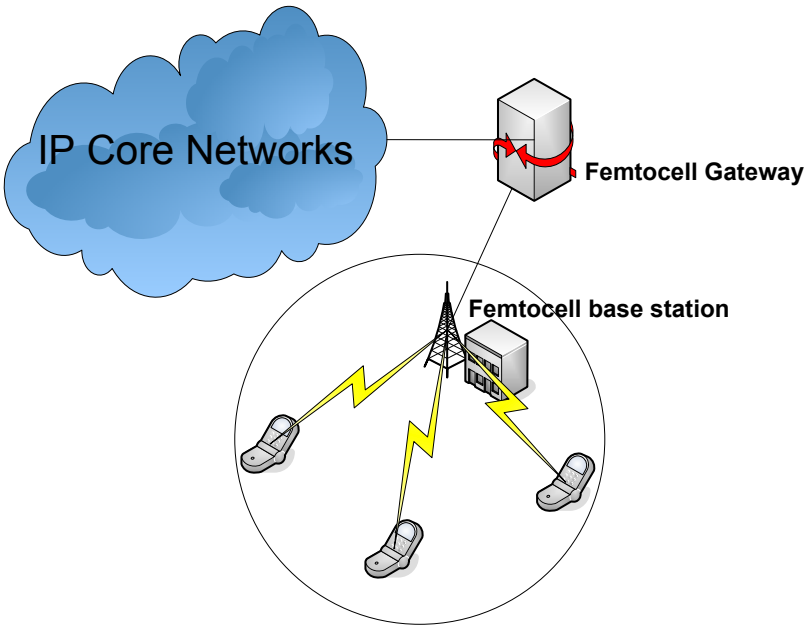


Figure 2-4 Femtocell network architecture

2.2.4 IMT-Advanced

On 18 January 2012, ITU announced two 4G network technologies: LTE Advanced and Wireless MAN-Advanced (IEEE 802.16m[17]). Both standards are known under the official designation of International Mobile Telecommunications-Advanced (IMT-Advanced). IMT-Advanced provides a worldwide platform for the next generations of mobile services-fast data access, unified messaging and broadband multimedia. It is designed based on all-IP

[16] Berg Insight- http://berginsight.com
[17] Wireless MAN-Advanced-http://www.itu.int/net/pressoffice/press_releases/2012/02.aspx

packet switched network paradigm and is able to interact with existing wireless networks, i.e. WLANs, digital video broadcasting systems. According to the new released specification[18], 4G devices are required to support 1Gbps for low mobility communications and 100Mbps for high mobility communications.

2.3 Broadband Networks

Figure 2-5 summarized the most widely deployed IEEE broadband networks. The latest group is referred as 802.23 which focus on the emergency services. Specifically, wireless standards such as IEEE 802.11, IEEE 802.15 and IEEE 802.16 are introduced in details. Additionally, IEEE 802.21 is described which present mechanism for seamless handover between different types of networks.

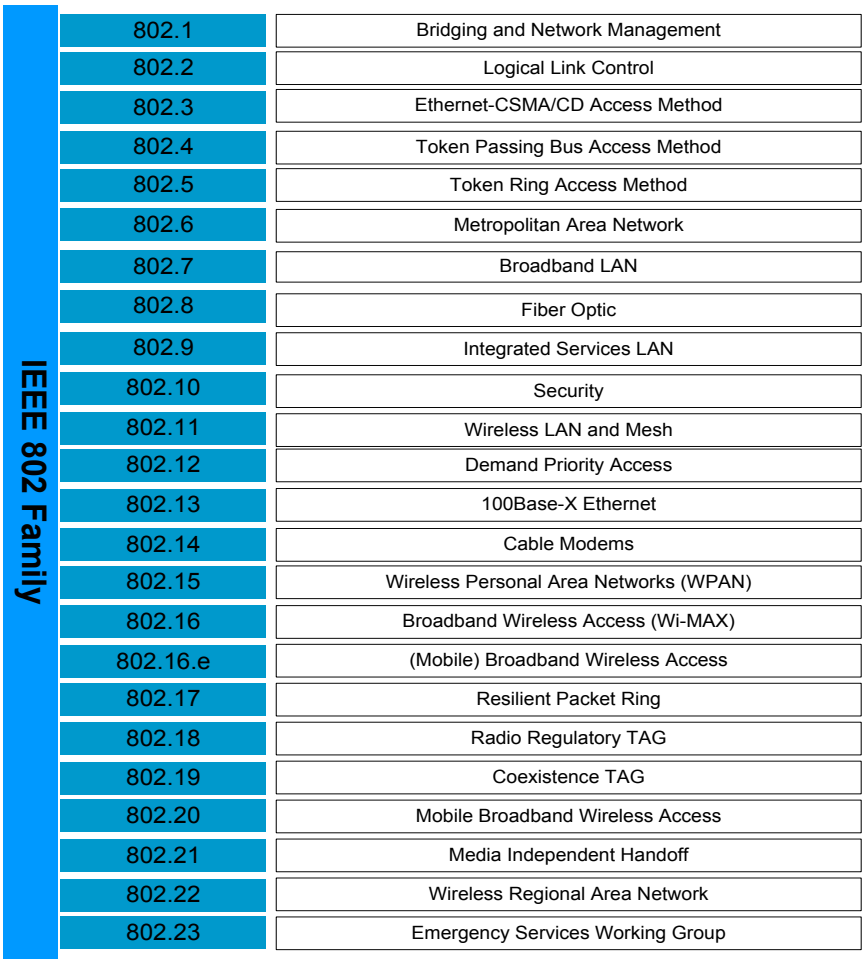


Figure 2-5 IEEE 802 Family Standards

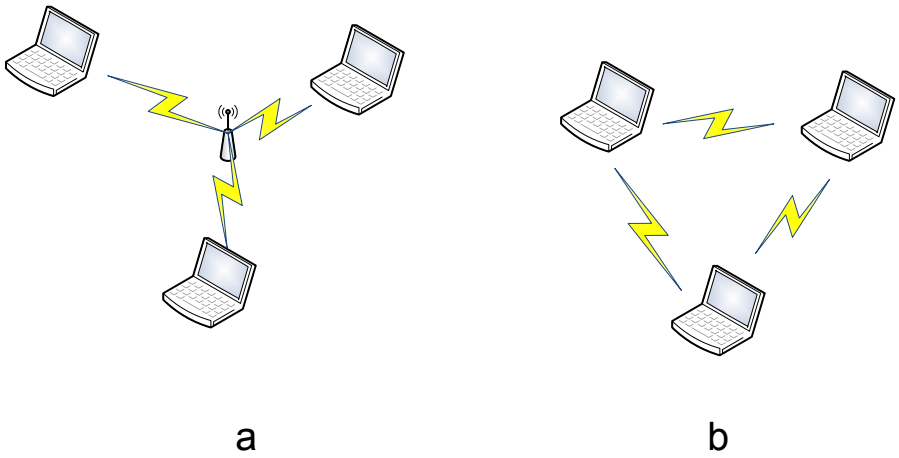
[18] ITU global standard for international mobile telecommunications ´IMT-Advanced´ ITU-R

TABLE 2-2 SHORT SUMMARY OF THE IEEE 802.11 STANDARD PROTOCOLS

Standard	Release	Frequency	Peak Data Rate (Physical Layer)	Modulation
802.11 legacy	1997	2.4GHz	2Mbps	DSSS, FHSS
802.11a	1999	5GHz	54Mbps	OFDM
802.11b	1999	2.4GHz	11Mbps	DSSS
802.11g	2003	2.4GHz	54Mbps	OFDM, DSSS
802.11n	2009	2.4GHz/5GHz	600Mbps	OFDM

TABLE 2-3 SUMMARY OF IEEE 802.11 AMENDMENTS

Amendment	Release	Description
802.11c	2001	Included in the IEEE 802.1D standard describing the bridging operation process between different 802 projects such as 802.3, 802.11 and 802.16
802.11d	2001	Supplements to 802.11-configuring devices with different RF regulation
802.11f	2003	Providing wireless access point communications among multivendor systems
802.11h	2004	Defines the spectrum and transmit power management for the 5GHz frequency band
802.11i	2004	Implemented as WPA2, specifying security mechanism of wireless network
802.11e	2005	Provides Quality of Service scheme for all 802.11 radio interfaces
802.11k	2008	Improves radio resource management, i.e., select the best available access point
802.11r	2008	Refers to Fast BSS Transition: allows seamless handoff between base stations for fast moving wireless devices
802.11w	2009	Improves the security of management frames of all 802.11 protocols
802.11p	2010	Provides Wireless Access in Vehicular Environments (WAVE), enabling data exchange between high-speed vehicles in the 5.9GHz
802.11z	2010	Under progress. Provides extensions to Direct Link Setup (DLS)
802.11s	2011	Defines the mechanism for wireless devices to form a WLAN mesh network
802.11u	2011	Intends to improve interworking with external networks, i.e., allow data access for a user who is not pre-authorized, keeping access when handoff from 3G to Wi-Fi
802.11v	2011	The next generation Wireless Network Management Standard. Allows client devices to exchange information about the network topology



(a) Infrastructure Mode (b) Ad hoc Mode

Figure 2-6 IEEE 802.11 network access topology

2.3.1 IEEE 802.11 (Wi-Fi)

In 1999, the IEEE 802.11 working group[19] was founded as the official organization to maintain the 802.11 standards (Wi-Fi) [16], as summarized in Table 2-2 and Table 2-3. These standards are supported by most mobile phones, laptops, and sensors due to the significant advantages in simple deployment, high data rate and low cost. The most popular 802.11 standards used today are 802.11b [17], 802.11g [18] and 802.11n [19]. The original IEEE 802.11 protocol adopts the CSMA/CA mechanism to manage the wireless channel access [4]. However, the 802.11 standard is only designed for best effort service. IEEE 802.11e [20] standard has also been introduced to provide Quality of Service (QoS) for multimedia applications by introducing support for differentiation between four different classes of traffic: voice, video, best effort and background. More details of 802.11 protocols are described in this section. There are two types of network topologies for IEEE 802.11: Infrastructure and ad-hoc. In the infrastructure topology, the channel access of all wireless devices is coordinated by a special device called the access point (AP). In the ad-hoc mode, wireless devices directly communicate with one another without an AP. Figure 2-6 illustrates the two topologies.

The data link layer of the 802 protocols is divided into two sub-layers: LLC (Logical Link Control) layer and the MAC (Medium Access Control) layer. The LLC provides quality oriented control mechanisms such as flow control, acknowledgement, and error detection. The MAC layer is responsible with the wireless channel access, i.e. which wireless client can

[19] IEEE 802.11 Working Group Setting the Standards for Wireless LANs-http://www.ieee802.org/11/.

get the opportunity to transmit. The 802.11 physical layer defines different physical interfaces for each 802.11 protocol. The initial version of IEEE 802.11 defined **Frequency Hopping Spread Spectrum (FHSS)** and **Direct Sequence Spread Spectrum (DSSS)**, which operate in the 2.4GHz frequency band with data rates up to 2Mbps. IEEE 802.11b operates in the 2.4GHz frequency with peak data rates of 11Mbps. IEEE 802.11b is designed based on direct-sequence modulation and the PHY is known as **High Rate Direct Sequence Spread Spectrum (HR/DSSS)**. IEEE 802.11g also operates in the 2.4GHz with peak data rates of 54Mbps. IEEE802.11g supports two types of PHY protocol: HR/DSSS and **Orthogonal Frequency Division Multiplexing (OFDM)**. IEEE 802.11a operates in the 5GHz frequency band with data rates of up to 54Mbps. IEEE 802.11a uses the OFDM as the PHY protocol. IEEE 802.11n can operate in both 2.4GHz and 5GHz frequency bands and support data rates of up to 600Mbps. IEEE 802.11n adopts an advanced PHY protocol called **Multiple Input and Multiple Output (MIMO)**, which uses multiple antennas at both the transmitter and receiver to improve signal transmission.

2.3.1.1 802.11 MAC Layer

IEEE 802.11 MAC layer controls the channel access for 802.11 enabled end points (i.e., wireless clients or access point). There are two access mechanisms: Distributed Coordination Function (DCF) and Point Coordination Function (PCF).

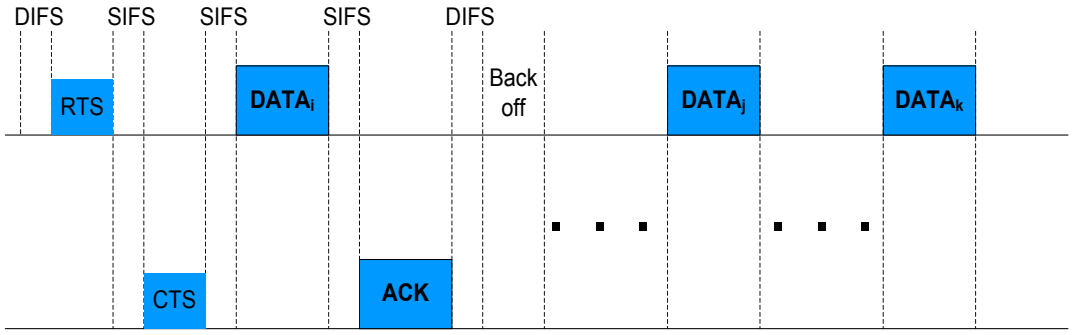


Figure 2-7 Event sequence of DCF for IEEE 802.11 MAC

DCF is employed by the majority of 802.11-enabled devices and uses the Carrier Sense Multiple Access with Collision Avoidance (CSMA/CA) mechanism. Figure 2-7 presents the event sequence of the DCF mechanism for 802.11 MAC protocol. In DCF, stations have to sense the channel status before any data packet ($DATA_i$) is transmitted. Additionally, DCF defines two medium access modes for stations to obtain the channel access opportunity: 1) basic access mode; 2) RTS/CTS (Require To Send / Clear To Send) mode. Take basic access mode for instance, if the channel is sensed idle for a time period

equal to a DCF Inter Frame Space (DIFS), the station is then allowed to transmit. If the channel is sensed busy, the station has to continue to monitor the channel until it has been idle for a DIFS period. In this case, the station has to wait for an extra duration called "backoff" in order to reduce the probability of collision. The "backoff" period is decreased by a time $T_{backoff.}$ DCF employs a Binary Exponential Backoff (BEB) scheme to determine the value of the backoff timer, $T_{backoff}$. $T_{backoff}$ value is decremented as long as the channel is sensed idle, and paused when the channel is sensed busy until the channel is idle for another DIFS. If the value of $T_{backoff}$ reaches zero, the station starts to transmit data. $T_{backoff}$ is randomly selected in the (0, CW-1) range. The parameter CW is called Contention Window, which is initialized with a value CW_{min} (minimum Contention Window). CW is doubled until it reaches CW_{max} (maximum Contention Window) at each unsuccessful transmission and is reset to CW_{min} after any successful transmission. CW_{min} and CW_{max} have static values initiated in the 802.11 specifications.

The RTS/CTS access mode [21] reserves the channel resource before DATA packets transmission. The purpose of using RTS/CTS is to avoid the hidden node problem due to a potential simultaneous channel access between multiple stations which do not directly sense each other's transmission as they are too far away. In the case of RTS/CTS access, if the channel has been idle for a period which equals DIFS, an RTS (Request to Send) frame is transmitted from the sender to receiver. If the RTS is received, the receiver then replies with a CTS (Clear to Send) frame after a SIFS (Short Inter-Frame Space). If the channel is idle for another SIFS, the station starts transmission. The backoff procedure is the same as in the basic access mode. However, the RTS/CTS mechanism introduces extra overhead and reduces the available bandwidth.

IEEE 802.11 PCF is a centralized MAC protocol to support collision free and time bounded services. In PCF, the access point (AP) grants the channel access opportunity to individual stations based on a polling mechanism during the Contention Free Period (CFP). Stations are only allowed to transmit when the AP polls them. The CFP alternates with a Contention Period (CP) in which data transmission is controlled using DCF. Generally, the CP value is set large enough to send at least one packet including RTS/CTS/ACK. Unlike in DCF where the overhead comes from collisions and backoff, over head in PCF are mainly caused by the polling mechanism.

Recent 802.11 standards, such as 802.11n and 802.11e, have different MAC layer and Physical layer protocols. Details of 802.11n and 802.11e are described separately in next sections.

2.3.1.2 IEEE 802.11n

IEEE 802.11n [19] is developed as an amendment to the IEEE 802.11-2007 standards in order to support higher data throughput and lower implementation costs. The major novelties of 802.11n are Multiple Input Multiple Output (MIMO), channel bonding, 2.4GHz and 5GHz band, and frame aggregation.

- Unlike in the Single Input Single Output (SISO) case where the channel is maintained by a single transmit antenna and a single receive antenna, MIMO uses multiple antennas to process more information. At the transmitter side, MIMO uses Space Division Multiplexing (SDM) to transmit data packets over different spatial channels on the same frequency, enabling an increased amount of data transmission. At the receiver side, MIMO allows combination of multiple signals, which increases the sigal strength and reduces the multipath fading effects.
- 802.11n operates at either 2.4GHz or 5GHz frequency bands, having each channel 40MHz wide, which doubles the 20MHz width existing in the previous 802.11 physical layer. The higher channel frequency allows for an increased physical data rate over a single channel. Due to the MIMO technique and the doubled physical channel frequency, the peak data rate supported by 802.11n is up to 600Mbps, in comparison with the 54Mbps and 11Mbps data rates for 802.11g and 802.11b, respectively.
- The MAC layer protocol of 802.11n extends the original DCF channel access mechanism. The difference from DCF is the introduction of frame aggregation, which allows sending multiple frames within a single frame. Frame aggregation significantly reduces the frame headers and the inter frame time space.

Recently, 802.11n-based local wireless networks have been widely deployed in companies[20] and universities[21]. IEEE 802.11n is backward compatible with the traditional 802.11a/b/g standards and provides significant improvements in both physical and MAC layers.

[20] White Paper, Intel, "Accelerating the Enterprise Network Using 802.11n Wireless," Jan.2010- http://www.intel.com/en_US/Assets/PDF/whitepaper/wp_IT_WirelessProtocol.pdf.
[21] "Next generation Wi-Fi technology for DCU students"- http://www.dcu.ie/news/2011/feb/s0211n.shtml.

TABLE 2-4 DEFAULT EDCA PARAMETERS FOR ACCESS CATEGORIES

Access Category	AIFS	CW_{min}	CW_{max}	TXOP*limit*
AC_Voice	2 μs	7	15	3.008 ms
AC_Video	2 μs	15	31	1.504 ms
AC_Best-Effort	3 μs	31	1023	0 ms
AC_Background	7 μs	31	1023	0 ms

2.3.1.3 IEEE 802.11e

IEEE 802.11e-2005 or 802.11e [20] is an approved amendment to the original IEEE 802.11 standard that supports Quality of Service (QoS) for multimedia services such as voice, video, best effort, and background. Many commercial access points provide 802.11e protocol as a software option.

802.11e modifies the original 802.11 Media Access Control (MAC) layer and enhances the DCF and the PCF mechanisms by developing a new coordination function: the Hybrid Coordination Function (HCF). HCF defines two medium access mechanisms: 1) a contention-based channel access method-Enhanced Distributed Channel Access (EDCA); 2) a polling-based HCF-controlled Channel Access method-HCCA. The novel idea for 802.11e is the introduction of access categories (AC) with different channel access opportunity. For instance, voice applications have the highest priority when accessing the channel, while background traffic-based applications are given the lowest priority. In EDCA, four ACs are defined to represent four types of services: voice, video, best-effort and background, as shown in Table 2-4. 802.11 EDCA assigns higher priority station with increased channel access opportunity. This is accomplished by using three parameters: Arbitration Inter-Frame Spaces (AIFS), Contention Window (CW) and Transmit Opportunity (TXOP). Before packet transmission, the station backs-off after the channel is being sensed idle for a time interval equal to AIFS. The backoff value is determined randomly in the [CW_{min}, CW_{max}] interval. Whenever collision occurs, CW increases up to CW_{max}, otherwise decreases to CW_{min}. After every successful transmission, the CW value is reset to CW_{min}. The term TXOP refers to a bounded time during which the station is able to transmit as many frames as possible. A TXOP is limited by TXOPlimit. As shown in Table 2-4, in EDCA, lower priority access categories have lower values of TXOPlimit and higher values of AIFS, CW_{max} and CW_{min}, than the higher priority access categories. Subsequently, higher priority traffic is given higher access to channel resources.

802.11e HCCA is based on a polling mechanism which is similar with the original 802.11 PCF protocol. In HCCA, during a beacon interval, an 802.11e enabled wireless station is allowed to send multiple contention-free packets called controlled access periods (CAPs) at any time after the channel is being sensed for a timer interval equals PCF Inter-Frame Space (PIFS). PIFS value is lower than DIFS and AIFS, therefore, a wireless station is given higher opportunity to use HCCA, in comparison with EDCA.

TABLE 2-5 SHORT SUMMARY OF IEEE 802.15 STANDARDS

Standard	Description
IEEE 802.15.1	Also called Bluetooth
IEEE 802.15.2	Intends to reduce interference between 802.11 and 802.15.1
IEEE 802.15.3	Provide high data rate in Ad hoc networks
IEEE 802.15.4	Enabling lower cost and provide low data rate
IEEE 802.15.5	Provides communication framework in wireless mesh network

2.3.2 IEEE 802.15 (WPAN)

IEEE 802.15, also known as *Wireless Personal Area Network (WPAN)*, refers to a series of communication listed in Table 2-5. The purpose of 802.15 is to enable the wireless interconnection between devices around individual's workspace. 802.15 enabled devices are expected to plug in to one-another in the same WPAN, provided they are within physical range of one another. Similar with the 802.11 standards, 802.15 protocols also operate at the 2.4 GHz frequency band. Following sections will discuss in more details the most widely used technologies such as Bluetooth [22] and Zigbee [23].

2.3.2.1 IEEE 802.15.1 (Bluetooth)

IEEE 802.15.1, or Bluetooth, is a wireless technology standard developed by Ericsson in 1994. It is designed for exchanging data over short distances (i.e. 10cm~10 m) with low implementation costs. Different Bluetooth devices can establish point-to-point connections or a Piconet through channel sharing. A Piconet consists of one master device and several slave devices using a star network topology. Multiple Piconets might be combined to form a Scatternet. The most recent versions of Bluetooth, Bluetooth v3.0 +HS, supports data rates of up to 24Mbps and Bluetooth v4.0 provides data rates of 25Mbps with lower energy consumption. Bluetooth uses multiple access methods, *Frequency-Hopping Spread Spectrum*, which fragments the data and transmits it by rapidly switching among up to 79 frequency channels (1 MHz each) within the unlicensed 2402-2480 MHz frequency band. The

frequency hopping technique effectively reduces the interference and enhances security; however, the bandwidth utilization is sacrificed.

2.3.2.2 IEEE 802.15.4 (ZigBee)

IEEE 802.15.4 specifies the physical and MAC layers for low cost and low data rate WPANs. ZigBee [23] extends the 802.15.4 standard by developing the upper layers. ZigBee specification is maintained by the ZigBee Alliance[22], an industry alliance consisting of a full spectrum of companies, ranging from Zigbee chip providers to solution providers. ZigBee supports a maximum data transfer rate of up to 250 Kbps for distances of up to 100 meters. It is designed to be simpler and less expensive than Bluetooth. ZigBee operates in the industrial, scientific and medical (ISM) radio bands, 868 MHz in Europe, 915 MHz in the USA and Australia, and 2.4 GHz in most jurisdictions worldwide. ZigBee can be activated in less than 15 ms, in comparison with the 3 seconds Bluetooth wake-up delay. ZigBee provides significant benefits for wireless sensor applications which require low data rate, long battery life, and secure networking. For non-commercial purposes, the ZigBee specification is available free.

TABLE 2-6 SUMMARY OF THE IEEE 802.16 STANDARDS [24]

Standards	Description
802.16-2001	Fixed Broadband Wireless Access (10–63 GHz)
802.16a-	Physical layer and MAC definitions for 2–11 GHz
802.16-2004	Air Interface for Fixed Broadband Wireless Access System
802.16f-	Management Information Base (MIB) for 802.16-2004
802.16e-	Mobile Broadband Wireless Access System
802.16k-	Bridging of 802.16 (an amendment to IEEE 802.1D)
802.16g-	Management Plane Procedures and Services
802.16-2009	Air Interface for Fixed and Mobile Broadband Wireless Access System

2.3.3 IEEE 802.16 (WiMAX)

IEEE 802.16 [24] specifies a series of wireless broadband standards. It has been commercialized under the name "WiMAX" (*Worldwide Interoperability for Microwave Access*) by the industry alliance, WiMAX Forum[23]. Table 2-6 shows the evolution history of

[22]Zigbee website-http://www.zigbee.org
[23] Wimax Forum-http:// www. Wimaxforum.org

the 802.16 standards since 2001. The Physical Layer (PHY) and the Media Access Control layer (MAC) of 802.16 are described in this section, separately.

2.3.3.1 IEEE 802.16 PHY Layer

IEEE 802.16 defines four types of PHY layer protocols: *Wireless MAN-SC (single carrier), Wireless MAN-SCa, Wireless MAN-OFDM (orthogonal frequency-division multiplexing)*, and *Wireless MAN-OFDMA (orthogonal frequency-division multiple access)*. WMAN-SC and WMAN-SCa operate in the 10–66 GHz frequency band and the last two PHY interfaces are designed for operating at frequency bands below 11 GHz. IEEE 802.16 supports both *time-division duplex* (*TDD*) and *frequency-division duplex* (*FDD*) operations for multiple accesses.

TABLE 2-7 IEEE 802.16E QOS CLASSES

Type	Service	Description
0	Unsolicited Grant Service	Real-time stream delivering constant-size packets at
1	Extended Real-time Polling	Real-time stream delivering variable-size packets on a
2	Real-time Polling Service	Real-time stream delivering variable-size packets at
3	Non-real-time Polling Service	Data stream with guaranteed minimum throughput
4	Best Effort	Data stream without minimum service level, like HTTP.

2.3.3.2 IEEE 802.16 MAC Layer

IEEE 802.16 provides two types of MAC layer protocols: *point-to-multipoint* (*PMP*) and *mesh* mode. In PMP, multiple nodes are used in an infrastructure mode which has one **base station (BS)** and several s**ubscriber stations (SSs)**. The channel frequencies are divided for uplink (from SS to BS) and downlink (from BS to SS) transmissions. In the mesh mode, similar with ad-hoc networking, each node acts as a relaying router in addition to their roles of sender and/or receiver.

Similar to the IEEE 802.11 standards, the 802.16 standards aims to provide high speed wireless communication. There are several differences. First, 802.16 protocols perform better in the outdoor environment and provide large coverage (i.e. 30-50km), while 802.11 is designed to be deployed mostly for indoor usage. Second, original 802.11 protocols do not support the Quality of Service (QoS) for multimedia communications. The 802.16-enabled base station uses a scheduling algorithm at MAC layer to control the delivered QoS of any subscriber station. For instance, the subscriber station (SS) can only start the transmission after allocated the required bandwidth resources by the Base Station.

2.3.3.3 IEEE 802.16e

IEEE 802.16e [24], also referred as *Mobile WiMAX*, is developed to support mobility for broadband wireless access and Quality of Service for multimedia services. 802.16e operates in the frequency band between 1.25 MHz and 20 MHz and uses the scalable OFDMA to transmit data. Both 802.16 and 802.16e standards support single carrier except that 802.16e delivers data using many sub-channels, i.e. up to 2048 sub-carriers. The 802.16e assigns each connection between SS and BS with a specific QoS class which is presented in Table 2-7. These QoS classes, combined with other parameters such as delay and throughput, are utilized by the base station to ensure support for the applications' QoS requirements. Unlike in 802.11, where the QoS is supported by adopting a distributed approach like 802.11e-EDCA, the per-flow QoS provisioning mechanism in 802.16 is centralized at the base station. Note that, 802.11 can also support centralized QoS in PCF and 802.11e HCCA; however, they are not widely implemented [25].

2.3.4 IEEE 802.21

The IEEE 802.21 [26] framework has been developed to improve user experience of mobile stations by enabling handovers between heterogeneous technologies such as Wi-Fi, WiMAX and 3G. IEEE 802.21 provides a mechanism that allows interaction between lower layers and network layer without dealing with specific technology. Figure 2-8 shows a logical diagram of the architecture of the different 802.21 enabled entities. Three types of 802.21 interfaces are presented, i.e., 802 network, mobile node, and 3G network. It can be observed from the figure that all the 802.21-compliant nodes have the same structure. IEEE 802.21 defines the Media Independent Handover Function (MIHF) that facilitates both mobile station and network initiated handovers. The MIHF logically serves as middle layer components between link layer and network layer. MIH User (MIHU) is introduced to represent the local entities (typically the mobile management protocols) that avail of MIHF services. MIH users use the MIHF functionality to control and gain handover-related information. MIH users locate at upper layers. The MIHF encompasses three types of communication services:

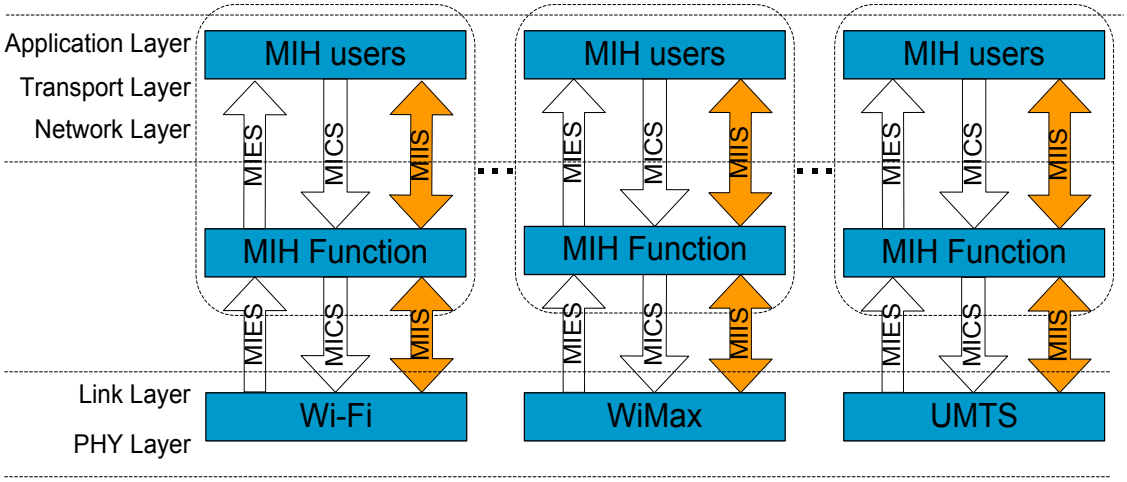


Figure 2-8 Communications between different 802.21interfaces

- *MIH Event Services (MIES)*. Events originate at the MIHF or lower layer. The destination of events is MIHF or upper layer entities. MIES provides event classification and reporting related to dynamic changes in link characteristics.
- *MIH Command Services (MICS)*. Commands originate at MIHF or upper layers. The destination of commands is MIHF or lower layer entities. MICS provides a set of commands to allow the MIH users to control the information from the lower layers. Any upper layers entities can register for an MIH event notification.
- *MIH Information Services (MIIS)*. Information originates at upper or lower layer with destination of upper or lower layer entities. MIIS presents a framework whereby the MIHF is able to acquire network and terminal information, such as network type, service provider identifier, QoS information, data rate, channel characteristics, vendor specifications, etc. MIIS specifies a standard format for this information, such as Extensible Markup Language (XML) or Type Length Value (TLV). They are transmitted through MIIS using query/response or broadcast/multicast mechanism.

These services are independent of each other and provide a unified interface for the upper layers.

2.4 Quality of Service (QoS) and Quality of Experience (QoE)

2.4.1 Introduction to QoS

Quality of Service (QoS) was first defined by a standardization body of ITU in 1994 [27] and referred to the ability to guarantee the quality of telephony communication. In general, the level of QoS is evaluated using metrics such as response time, signal-to-noise ratio, network capacity, etc. In the case of delivering multimedia content to Wireless LANs, the term QoS refers to the ability to provide quality provisioning and service differentiation among heterogeneous multimeida content and devices. The QoS level of the multimedia services are evaluated by a number of factors including end-to-end delay, jitter, throughput, and packet loss rate. These factors are described separately as follows.

1) End-to-end Delay

IETF RFC2679 [28] defines delay which refers to the time required for delivering a packet from the source to the destination in the IP-based networks. Delay consists of two parts: end-point delay and network delay.

End-point delay is the delay introduced by end-point applications. For instance, streaming video applications require encoding and decoding processes at both server and client and introduce the codec delays; VoIP applications take time to analyse voice samples introducing packetization delays; some multimedia applications use jitter buffer to reduce delivery delay jitter and thus introduce buffering delays.

Network delay is defined as the time from delivering the first bit of a packet at the source until receiving the last bit of the packet at the receiver. Network delay could be further divided into three parts:

- **Transmission Delay** is the time taken to transmit a packet to the medium. Transmission Delay is insignificant for high-speed links. For instance, on a 622.080 Mbit/s STM-4 (Synchronous Transport Module) link delivering a 1500-byte packet takes 0.018ms. However, Transmission Delay is significant for low-speed access links such as 380Kbps Digital Subscriber Line (DSL), where the transmission of a 1500 byte packet cost 32ms.

- **Packet Processing Delay** is the time taken to process a packet at various network devices, i.e., delivering routing, queuing, etc.
- **Propagation Delay** is the time required to deliver a packet over the transmission medium. Propagation Delay varies for different medium, for instance, transmission bitrate in fiber (up to 10Gbps) is much larger than that of in copper (up to 15Mbps).

Different multimedia applications have various requirements regarding the delay, i.e. VoIP applications for very good quality require delays lower than 150ms, while the video conferencing applications need delays lower than 400ms for the same quality level.

2) Jitter

Jitter, or packet delay variation, is the difference between the delay of the current packet and the delay of the reference packet which generally refers to the packet with the lowest delay within a stream. Jitter is defined in IETF RFC3393 [29] and can be caused by many reasons:

- Different packets have different queuing delays and propagation delays, since they may travel via different network paths.
- Different packets can have different processing delays at the same network device. Such differences are significant in network devices that using software-based packet forwarding with caching mechanism.
- Different multimedia applications have different requirements on jitter. Generally, VoIP services with good quality level require jitter less than 50ms, while email services do not have specific jitter requirements at all.

3) Packet Loss Rate (PLR)

Packet Loss Rate, defined in IETF RFC 2680 [30], is the percentage of packet lost during transmissions. A packet might be dropped due to the following reasons:

- Buffer overflow caused by network link congestion.
- Network device failures.
- Fading effects due to the wireless characteristics.
- Collision occurrence in CSMA/CA based wireless networks.

Different applications have specific requirements on packet loss. In particular, packet loss ratios lower than 0.1 percent will provide good quality for most multimedia applications.

4) QoS Requirements for Multimedia Services

As discussed above, different applications have various requirements on delay, jitter and packet loss rates. Table 2-8 presents network performance objectives for IP-based applications, according to the ITU-Y.1541 standard [31]. The classes listed in the table indicated different levels of the network QoS-related parameters. It can be concluded that:

- Delays of 100ms meet the requirement of all multimedia services, and delays of 400ms are acceptable for interactive applications.
- Jitter below 50ms meet the need of all real-time applications. Many applications do not have specific requirements on jitter.
- Packet loss rates below 0.1 percent meet the need of most applications.

TABLE 2-8 IP NETWORK PERFORMANCE OBJECTIVES FOR DIFFERENT APPLICATIONS [31]

Network	Class 0	Class 1	Class 2	Class 3	Class 4	Class 5
Delay	100ms	400ms	100ms	400ms	1s	-
Jitter	50ms	50ms	-	-	-	-
Packet loss	$1x10^{-3}$	$1x10^{-3}$	$1x10^{-3}$	$1x10^{-3}$	$1x10^{-3}$	-
Applications	Real-time,	Real-time,	Transaction	Transaction	Low loss	Traditional
'-' refers to the unspecified detail in Y.1541						

2.4.2 Introduction to QoE

Quality of Experience (QoE) is not an objective metric, but a very subjective term reflecting the user's perception over the services they received (TV broadcasting, phone call, web browsing, video on demand, etc). Higher levels of QoE indicate that the users have better experience of the network and service. QoE is related to but differs from QoS. QoS indicates the capability of a network to provide certain level of quality to a service, with QoS measurement which are less familiar to customer. QoE is a subjective measure from the user's perception of the provided service. Although the purpose of QoS is to provide the end user with higher satisfaction level, the improvement of QoS cannot guarantee a better QoE.

Unlike the measurement of QoS, QoE assessment takes into account many factors that contribute to overall rating of user such as flexibility, security, cost, personalization, etc. Generally, QoS is evaluated by measuring the QoS-related parameters such as delay, jitter and loss and QoE can be assessed based on subjective tests. QoE can also be estimated using objective evaluation results, which predicts the QoE by combining weighted QoS parameters, such as delay, jitter, loss, bit-error rate and bandwidth.

2.4.3 QoS/QoE Evaluation Metrics

It is significant to measure the QoE despite its subjective nature. For instance, the service providers can optimize their commercial decisions based on the end user's perception of the provided service. Several objective and subjective metrics have been developed for evaluating QoS/QoE of multimedia delivery. These evaluation metrics are introduced in this section in terms of audio and video delivery services, respectively.

2.4.3.1 Objective Metrics

Objective assessment of multimedia services quality aim at determining the quality in the absence of the human viewer. Different principles are used such as, the comparison between reference and delivered sequences, statistical assessment of several analysed cases, etc. Objective metrics rely on predefined models or algorithms to quantify the quality of audio or video sequence.

A. Audio

ITU-T P.861 [33] document, released in 1997, defines a computational algorithm **Perceptual Speech Quality Measure (PSQM)** to objectively evaluate telephone voice (300-3400Hz) quality. PSQM algorithm converts physical signals into human perceptually psychoacoustic domain and analyses the difference between original and impaired voice signals. The comparison results are represented using PSQM values which range from 0 (non-degradation) to 6.5 (worst-degradation). In 2001, PSQM was replaced with **Perceptual Evaluation of Speech Quality (PESQ)** [34] which is standardised as **ITU-T P.862**. ITU-T P.862 maps PESQ values to MOS scores and supports measurements in common telephone band (300-3400Hz). In 2011, **ITU-T P.863** [35] is announced to estimate speech quality using digital speech analysis. P.863, also referred as **Perceptual Objective Listening Quality Assessment (POLQA)**, extends PESQ by supporting super-wideband speech signals (50-14000Hz).

TABLE 2-9 CATEGORIES OF SPEECH TRANSMISSION QUALITY [32]

Range of E-model Rating R	Speech transmission quality	User satifaction
90≤R<100	Best	Very Satisfied
80≤R<90	High	Satisfied
70≤R<80	Medium	Some users dissatisfied
60≤R<70	Low	Many users dissatisfied
50≤R<60	Poor	Nearly all users dissatisfied

E-Model [36] has been standardized by ITU as the recommendation **ITU-T G.107** and provides predication of the expected voice quality. The latest version of E-Model was released in Dec.2011 [37]. The E-model takes into account typical telephony-band impairments such as loss, noise, echo. It can be used for voice quality assessment in wired and wireless networks, based on circuit-switched and packet-switched technology. The E-model is developed by modelling a large amount of subjective testing results. The output value of the E-model is referred to "Transmission Rating Factor (R)" and R can be further mapped to the Mean Opinion Score (MOS). The E-model provides a mathematical algorithm that takes into account the combination effects of impairments, as given in equation (2-1):

$$R = Ro - Is - Id - Ie\_eff + A \tag{2-1}$$

where:

Parameter *Ro* means the basic signal-to-noise ratio, *Is* is a combination of all impairments that occur with the voice signal, such as, too loud speech level, non-optimum sidetone, quantization noise, etc. Factor *Id* represents the impairments caused by delay, *Ie_eff* is an "effective equipment impairment factor", which represents impairments caused by low bit-rate codecs. Factor *A* is an "advantage factor" that allows for compensation of impairment factors when the user benefits from other types of access to the user. The values of *R* range from 0 to 100, with higher values meaning higher speech quality. Table 2-9 relates *R* to speech transmission quality and user satisfaction.

B. Video

Peak-Signal-to-Noise-Ratio (PSNR) [38], measures the ratio between the maximum possible power of a signal and the power of any corrupting noise. Typical values for PSNR in image and video compression are between 30dB and 50dB, where higher PSNR indicates better received quality. PSNR is an approximation of the human perception of quality. In some cases, the reconstructed video may be very close to the original one, despite of a low PSNR calculated. The value of PSNR is derived by setting the mean square error (MSE) in relation to the maximum possible value of the luminance, as shown in equation (2-2):

$$MSE = \frac{1}{m \times n} \sum_{i=0}^{m-1} \sum_{j=0}^{n-1} [I(\mathrm{i}, \mathrm{j}) - \mathrm{K}(\mathrm{i}, \mathrm{j})]^2 \tag{2-2}$$

where *I(i,j)* is the original signal at pixel *(i, j)*, *K(i, j)* is the reconstructed signal, and $m \times n$ is the picture size. PSNR is then defined in equation (2-3):

$$PSNR = 10 \times \log(\frac{255^2}{MSE}) \tag{2-3}$$

Table 2-10 shows the mapping between PSNR and MOS values. PSNR is simple to compute and is one of the most widely used evaluation methods to analyse the difference between images. However, PSNR does not take into account the visual masking phenomenon impacted by a human viewer, that is, any pixel error leads to the decrease of the PSNR even if this error will not be perceived. This is as PSNR does not take human perceptual quality characteristics into consideration.

TABLE 2-10 PSNR-MOS MAPPING WITH THE EQUIVALENT ITU-T R. P.910 QUALITY AND IMPAIRMENT SCALE [38]

MOS	Impairment	PSNR(db)
5(Excellent)	Imperceptible	>37
4(Good)	Perceptible, not Annoying	31-37
3(Fair)	Slightly Annoying	25-31
2(Poor)	Annoying	20-25
1(Bad)	Very Annoying	<20

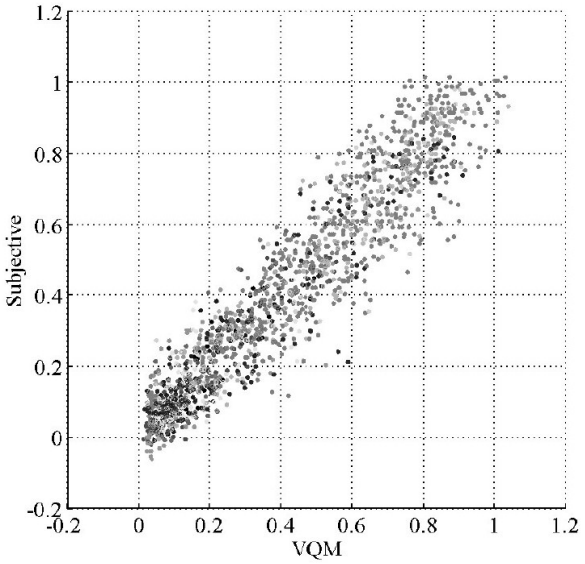


Figure 2-9 Clip subjective quality vs. clip VQMG [38]

Video Quality Metric (VQM) [39] provides an objective measurement for perceived video quality. VQM reflects perceptual effects of blurring, jerky/unnatural motion, global noise, block distortion and colour distortion. The computation of VQM applies four steps which take the original and transmitted video as input: 1) **Calibration**- it estimates and corrects the changed features from the transmitted video sequence, such as spatial and temporal shift, the contrast and brightness offset; 2) **Quality Features Extraction**- using a mathematical function, it extracts the VQM features (spatial gradients, chrominance, contrast, and absolute temporal information) that describe perceptual variation; 3) **Quality Parameters Calculation**- it computes the quality parameters that characterizes perceptual changes by comparing the original and transmitted video; 4) **VQM Calculation**- VQM is computed by linearly combining parameters obtained from the previous three steps. Testing results from [40] show how VQM has up to 95% correlation with subjective approaches and has been adopted by ANSI as an objective video quality standard. Figure 2-9 illustrates a summary of the test results from eleven experiments during 1992-1999 that were performed by Wolf and Pinson [41], which show the high correlation coefficient of 0.95 between subjective tests and the VQM general model (VQMG).

Moving Picture Quality Metric (MPQM) [42]**,** is an objective quality metric for moving pictures. In comparison with PSNR and VQM, MPQM takes into account human vision characteristics: contrast sensitivity and masking. The eye sensitivity varies for different spatial/temporal frequencies and a signal can be sensed by eye only if the contrast (of signal) is higher than a certain threshold. The masking phenomenon refers to the human response to combined signals, i.e. the foreground sensitivity might be impacted by the contrast of the background. There are three steps to implement MPQM-based assessment: 1) decompose the original sequence and distorted sequence into perceptual channels; 2) contrast sensitivity and masking are accounted using a channel-based distortion measure; 3) a mathematical data analysis is performed to compute the quality rating, which ranges from 1 (bad) to 5 (excellent).

Perceptual Video Quality Measure (PVQM) [43] uses the same approach for assessing video quality as standardised in ITU-T P.861 [44] for measuring speech quality. PVQM deals with spatial distortion, temporal distortion, and spatial-temporal-localised distortions found in error conditions. The method achieves a full reference metric and thus takes two video sequences as input (reference and delivered). The operation is based on the

fact that the Human Visual System is more sensitive to the sharpness of luminance component than that of chrominance components.

2.4.3.2 Subjective Metrics

The purpose of subjective evaluation is to obtain a close reflection of end user's perception of the delivered multimedia services quality. In this section, the most widely used subjective metrics are described in terms of audio and video services, respectively. Traditionally, many research works have been done in controlled laboratory settings. This type of research makes it convenient to investigate the influence of particular parameters on a user's perceptions without the influences of external environmental factors. However, subjective tests in living or semi-living labs are also attracting attention. Unlike lab tests, in semi-living tests, users are interacting with new technologies in (semi-) realistic contexts, for instance, interferences introduced by other networks, shadowing effects due to obstructions, brightness of the environment, users getting distracted by noise, etc. These realistic contexts might have significant impact on user's perception. Consequently, subjective-based evaluation research will result in more accurate results and have a higher environmental validity in comparison with controlled lab tests.

A. Audio

Due to the rapid deployment of modern telecommunication networks, there is an increasing need for evaluating the transmission characteristics of audio services. Generally, subjective methods for audio assessment require a group of listeners to evaluate the voice quality.

ITU-T P.800 [45] introduces methods for subjective determination of speech transmission quality. The recommendation presents the advice to administrations on implementing subjective tests of transmission quality in their own laboratories. ITU-T P.800 is developed intended to be generally applicable in order to cover different forms of degradation factors. The quality degradation might be caused by factors including: loss, circuit noise, transmission errors, environmental noise, sidetone, talker echo, non-linear distortion of various kinds including low bit-rate encoding, propagation time, harmful effects of voice-operated devices, distortions of the time scale arising from packet switching, and time –varying degradations of the communication channel. P.800 lists three primarily recommended methods as follows: 1) ***Conversation-opinion tests***, which aim to reproduce, in the laboratory situation, the actual service conditions experienced by telephone customers; 2) ***Listening-opinion tests***, which has less restrictions of realism as conversion-opinion tests,

but the artificiality has to be accepted; 3) ***Interview and survey tests***, which needs large amount of effort and includes the questions to be asked when interviewing customers. Additionally, several opinion scales are recommended by ITU-T. ***Mean Opinion Scale (MOS)*** defines five-point scales: "*Excellent*"=5, "*Good*"=4, "*Fair*"=3, "*Poor*"=2, and "*Bad*"=1. Such five-point scales can be applied to other situations. For instance, the question "*Please rating the loudness preference*" can be responded as follows: "*Much louder than preferred*"=5, "*Louder than preferred*"=4, "*Preferred*"=3, "*Quieter than preferred*"=2, and "*Much quieter than preferred*"=1. ***Difficulty scale*** is a binary response obtained from subjects. For instance, the answers "*Yes*" or "*No*" are used to response the question "*Did you or your partner have any difficulty in talking or hearing over the connection?*". Other opinion scales are also suitable according to the experimental methods [46].

Different with ITU-T P.800, **ITU-T P.835** [47] describes the subjective test methodology for evaluating speech communication systems that include ***Noise Suppression Algorithm (NSA)***. Typically, NSA attempts to reduce noise without adversely affecting the signal quality. However, higher levels of NSAs often adversely influence the speech component as more of the noise or background component is removed. In this case, the rating process becomes confusing. For instance, the background is improved due to less noise while the speech signal is degraded. To alleviate such problems, separate rating scales are used to independently estimate the subjective quality of the Speech Signal alone, the Background Noise alone, and Overall quality. The mean opinion score (MOS) used in ITU-T P.800 are adopted as the rating scale, for instance, *speech-MOS (S-MOS)* refers to speech signal quality, *noise-MOS (N-MOS)* refers to background noise level, and *Global-MOS (G-MOS)* is the overall quality level. Listeners shall complete the text instructions form to avoid ambiguity and differences across experiments. Examples of the rating scales are as follows: 1) ***Speech Signal***-"*Not distorted*"=5, "*Slightly Distorted*"=4, "*Somewhat Distorted*"=3, "*Fairly Distorted*"=2, and "*Very Distorted*"=1; 2) ***Background***-"*Not Noticeable*"=5, "*Slightly Noticeable*"=4, "*Noticeable But Not Intrusive*"=3, "*Somewhat Intrusive*"=2, and "*Very Intrusive*"=1; 3) ***Overall Speech Quality***-"*Excellent*"=5, "*Good*"=4, "*Fair*"=3, "*Poor*"=2, and "*Bad*"=1.

Other subjective methods for evaluating speech quality have also been developed by ITU. **ITU-T P.805** [48] describes the method to evaluate the effects of degradation (i.e. delay, echo, voice clipping, loss) on the quality of voice delivery. **ITU-T P.830** [49] introduces a testing method for assessing digital process and contains advice on the

performance evaluation of digital codecs. **ITU-T P.840** [50] describes a subjective test method to evaluate the speech quality of circuit multiplication equipment. **ITU-T P.851** [51] presents subjective evaluation methods providing information about the quality of telephone services based on spoken dialogue systems.

B. Video

The Human eyes are used for subjectively evaluating the delivered video quality. Two widely used subjective test methods, **ITU-T P.910** and **ITU-T P.911** are described in details.

ITU-T P.910 [52] defines non-interactive subjective assessment methods for evaluating the quality of **digital video images** for applications such as video telephony, video conferencing, and storage and retrieval applications. The proposed methods are also expected to be applied in other situations, i.e. ranking of video system performance, evaluation of the quality level during a video connection, etc. ITU-T P.910 specifies strict test conditions including: 1) ***Viewing Conditions***, such as, peak luminance of the screen (100-200cd/m), ratio of luminance of inactive screen to peak luminance (≤0.05), background room illumination (≤20lux), etc; 2) ***Video Content***, which should be stored in digital format in order to avoid distortion; 3) ***Number of observers*** in a viewing test is at least 15 and they should not be experienced assessors. Before starting the experiment, the observer should be given a description of test scenario and opinion scale. It must not be implied that the worst video quality is necessarily corresponds to the lowest subjective scale. Additionally, at least four different types of scenes should be selected in order to avoid boring the observers. After each presentation, the observers are required to evaluate the overall video quality using five-level scale: "*Excellent*"=5, "*Good*"=4, "*Fair*"=3, "*Poor*"=2, and "*Bad*"=1. Notably, a more level scale (i.e. seven-level, nine-level) might be used if high discriminative power is required.

ITU-T P.911 [53] describes non-interactive subjective evaluation methods for assessing one-way audiovisual quality for **multimedia** applications such as video conferencing, storage and retrieval applications, telemedical applications, etc. ITU-T P.911 outlines the characteristics of the source sequences to be selected including: 1) ***Viewing Conditions***, such as, peak luminance of the screen (100-200cd/m), ratio of luminance of inactive screen to peak luminance (≤0.05), background room illumination (≤20lux), etc; 2) ***Duration***, which should be about 10s, but not shorter than 8s. The termination of the scene should be a complete sentence or musical phrase. An initial and a final silent period with less

than 500ms can make a more natural sequence; 3) ***Content***, where the video and audio should be synchronized; 4) ***Number of sequences***, which should be defined according to experimental design. Generally, at least four different types of scenes should be selected. In order to measure the perceived quality of audiovisual sequences, subjective scaling methods are used including: 1) **Five-level scale for overall quality**: "*Excellent*"=5, "*Good*"=4, "*Fair*"=3, "*Poor*"=2, and "*Bad*"=1; 2) **Five-level scale for impairment level**: "*Imperceptible*"=5, "*Perceptible but not annoying*"=4, "*Slightly annoying*"=3, "*Annoying*"=2, and "*Very annoying*"=1. Additionally, the recommendation presents the relation between audio, video, and audiovisual quality. Highlighted conclusions are shown as follows: 1) Video dominates overall perception. The correlation between video and overall audiovisual quality is higher than the correlation between audio and overall audiovisual quality; 2) The one-way overall audiovisual quality can be predicted from the one-way audio and one-way video quality; 3) The mapping from separate audio and video quality to the overall audiovisual quality was found based on four sets of subjective experiments: $MOS_{AV} = \alpha + \beta \times MOS_A \times MOS_V$, where MOS_{AV} , MOS_A, and MOS_V refer to the quality of audiovisual sequence, audio sequence, and video sequence, separately. The recommended values are set to 1.3 for α and 1.1 for β. The correlation between predicted and measured overall audiovisual quality varied from 0.93 to 0.99.

Other subjective methods for evaluating speech quality have also been developed. **ITU-R BT.500-13** [54] provides methods for evaluating the quality of television pictures including setup of test, grading scales, and viewing condition. **ITU-T P.920** [55] describes evaluation methods for quantifying the impact of coding artifacts, transmission delay and transmission impairments (e.g. packet loss, cell loss, digital channel errors) on point-to-point or multipoint audiovisual communications. **Subjective Assessment Methodology for Video Quality (SAMVIQ)** [56], developed by *European Broadcasting Union (EBU[24])*, aims to evaluate the performance of video codecs for the internet. SAMVIQ takes into account the codec types such as image formats, bitrates, temporal resolutions, zooming effects, packet losses, etc.

2.4.3.3 Hybrid Model

Video quality is best assessed using subjective methods. However, subjective evaluation metrics are time consuming and cannot be implemented in real time. Recently, several new

[24] European Broadcasting Union-http://www.ebu.ch

quality metrics have been proposed that take into account both objective and subjective metrics. In this section, three objective and subjective correlation models are introduced.

Authors of [57] propose a mathematical equation describing the relationship between QoS and user QoE. The requirements of QoS parameters to satisfy QoE are different according to the characteristics of services. For video applications (IPTV, video conference), delay is the primary criteria. For email applications, packet loss is more significant than throughput and delay. The relation between QoS and QoE is given by equations (2-3) and (2-4):

$$QoS = A \times Delay + B \times Jitter + C \times Loss + D \times Error + E \times Bandwidth \tag{2-3}$$

$$QoE(QoS) = K\left\{\frac{(\mathrm{e}^{\mathrm{QoS}\text{-}\alpha} + e^{-QoS+\alpha})}{(\mathrm{e}^{\mathrm{QoS}\text{-}\alpha} + e^{-QoS+\alpha} + \beta)} + 1\right\} \tag{2-4}$$

The parameter α represents the QoS quality class of the network level and β is determined according to the class of service. QoE can be computed directly based on QoS parameters such as delay, jitter, loss, error and bandwidth (throughput). The weighted values (*A, B, C, D, E*) are determined by services, which indicate a Guaranteed Service, a Premium Service and Best Effort Service. However, real-life environment-based evaluation is needed to assess the performance of the model.

[58] proposes a novel video quality metric for bit rate control via joint adjustment of quantization and frame rate. It is demonstrated that the classical PSNR does not match subjective quality data. A correlation model between the PSNR and subjective metric is defined that accounts for both encoding parameters (quantization and frame rate), and intrinsic video sequence characteristics (motion speed). The average correlation coefficient tested is 0.93 for the proposed metric, in contrast with the PSNR's 0.70. The new quality metric is given by equation (2-5):

$$QM = PSNR + a \times m^b \times (30 - FR) \tag{2-5}$$

where a = 0.986 and b=0.378. *QM* represents the quality metric, *FR* indicates the frame rate and *m* denotes the normalized average magnitude of large motion vectors.

In [59], a new QoE estimation tool is proposed using temporal resolution, spatial resolution, and the Root Mean Square of the Error (RMSE) between the original image and the encoded one. RMSE measures the differences between values estimated by a model and

values observed. The tool is designed to be used at server side to adapt the encoding scheme to the three scalability factors. The objective metrics are converted into estimated subjective values, since objective measurement cannot accurately reflect the customer's perception. Equation (3-6) gives the QoE estimation function correlating the Human Vision System and RMSE variation: where Q is the overall QoE, *FrameRate* and *Definition* refer to the video frame rate and image definition, γ is used to maintain the resulting scale within the maximum

$$Q = \gamma \times e^{-\alpha \times (RMSE)^2} \times \left[\frac{\beta_1 + \beta_2 \times \log_{10}(FrameRate)}{1 + e^{-\delta_1 \times (Definition - \delta_2)}} \right] \quad (2\text{-}6)$$

variation of both original range and is set as 1.0747. Coefficient factors β_1 and β_2 are set as around 0.2827 and 0.4634, respectively, and δ_1 and δ_2 are set as around 1.186 and 1.819, respectively.

2.5 Multimedia Streaming Service

2.5.1 Introduction

Multimedia streaming refers to the continuous transmission of continued video, audio and text data. In contrast to other services in which users have wait for the entire data to be downloaded, multimedia streaming enables users watch the multimedia immediately after the transmission has started. Current multimedia streaming services can be categorized in two categories: video/audio-on-demand and live streaming. Typical video on demand (VOD) services include IPTV bring pre-recorded video content to television sets and computers; video conference and live sports broadcasting are important live streaming services that rely on real-time video delivery.

2.5.2 Features

2.5.2.1 Video/Audio Codec

At the sender side, multimedia stream is encoded from the video and audio signals with specific encoding algorithms. The receiving devices decode the multimedia data with the corresponding video/audio decoding algorithms. Table 2-11 lists the widely used video and audio encoding and decoding schemes. The ISO/IEC Moving Pictures Expert Group (MPEG)[25] and International Telecommunication Union (ITU)[26] are the primary bodies for

[25] ISO/IEC MPEG-http://mpeg.chiariglione.org/visions/mpeg/index.htm

developing the video and audio codecs. In this section, MPEG-2 and MPEG-4 compression algorithms are described in details since they have been widely used.

TABLE 2-11 WIDELY USED VIDEO/AUDIO CODEC

Type	Codec	Description
Video	**Lossless Video Codec**	HuffYUV: A very fast, lossless Win32 video codec written by Ben Rudiak-Gould.
	MPEG-2 [60]	Lossy data compression algorithm which is widely used for digital television signals broadcasted by cable, satellite TV, and distributed on DVD.
	MPEG-4 Part 2 [69]	• *DivX Pro Codec*: A MPEG-4 ASP codec made by DivX, Inc. • *Xvid*: Free/open-source implementation of MPEG-4 ASP • *Nero Digital*:MPEG-4 compatible codecs developed by *Nero AG*[27] and *Ateme*[28].
	MPEG-4 Part 10 (H.264/MPEG-4 AVC) [70]	*Nero Digital*: Commercial MPEG-4 ASP and AVC codecs developed by Nero AG *QuickTime H.264*: H.264 implementation released by Apple
	Microsoft Video Codec [71]	*Windows Media Video (WMV)*: Microsoft's video codec designs.
Audio	**Apple Lossless Audio Codec**[21]	Developed by Apple Inc. for lossless data compression of digital audio
	Advanced Audio Coding (AAC)	Standardized by ISO as part of MPEG-2 and MPEG-4 specifications. AAC is lossy compression and encoding scheme for digital audio.
	Dolby Digital[29]	Developed by Dolby lab as the audio compression technology.
	Windows Media Audio 9 Lossless [72]	Developed by Microsoft as the lossless audio codec.
	ITU Standards	• ITU-G.711, sampling with 8KHz and support 64Kbps • ITU-G.719, sampling with 48KHz and support up to 128Kbps • ITU-G.722, sampling with 16KHz and support up to 64Kbps • ITU-G.723.1, sampling with 8KHz and support up to 6.3Kbps, mostly used in VoIP service • ITU-G.729, sampling with 8KHz and support fixed bit-rate of 8Kbps, mostly used in VoIP service

[26] ITU-http://www.itu.int
[27] Nero-http://www.nero.com
[28] Ateme-http://www.ateme.com
[29] Dolby-http://www.dolby.com/digital

MPEG-2, was standardised by both ISO in ISO/IEC 13818 [60] and ITU in ITU-T H.262 [61]. ***MPEG-2 video*** targets high bit-rates of up to 20Mbps with full size pictures and high quality. It allows for high flexibility with the introduction of "profile" which defines subsets of the MPEG-2 syntax and semantics. ***MPEG-audio*** allows encoding the audio programs with multi-channels (i.e. up to 5.1 multi-channels). Additionally, ***MPEG-2 Transport Stream (MPEG-2 TS)*** is specified in MPEG-2 part 1 as a standard format for transmission and storage of audio and video data. Transport stream is designed for use in situations when errors are likely occurring, i.e. terrestrial and satellite broadcast. ***MPEG-2 Program Stream (MPEG-2 PS)*** is specified in MPEG-2 part 1 as a container format. MPEG-2 PS aims at combining several elementary streams, which have a common time base, into a single stream. Program stream is designed for use in relatively error-free environments and reliable media such as DVDs.

MPEG-4 [62], standardised as ISO/IEC 14496, defines compression algorithm of digital audio and visual data. It extends MPEG-2 by adding new features such as object-oriented composite files (including audio, video, and virtual reality modelling language), error resilience, digital rights management support, etc.MPEG-4 allows for transmission flexibility and provides support for objects with both natural and synthetic content. MPEG-4 is divided into a number of parts, in which the key parts to be aware of are MPEG-4 part 2 and MPEG-4 part 10. MPEG-4 part 2 is implemented in several popular codecs including *DivX*[30] and *Xvid*[31]. MPEG-4 part 10, or known as H.264/MPEG-4 AVC, is used by codecs such as *x264*[32], *QuickTime*[33], etc.

2.5.2.2 Synchronization

Synchronization between the video, audio, and media components of the stream is required to be maintained. The inter-media skew should be kept below 20ms according to the 3GPP specification[34]. 3GPP organization also suggests the minimum bandwidth of 14.4 Kbps for multimedia services with video and audio contents and the minimum bandwidth of 9.6 Kbps for audio-only streams. In the case when the minimum required bandwidth of multimedia cannot be satisfied, the streaming may continue but will be affected by blocking.

[30] DivX-http://www.divx.com/company/trademarks
[31] Xvid-http://www.xvid.org
[32] x264-http://x264.nl
[33] Quicktime-http://www.apple.com/quicktime
[34] Multimedia Streaming Services-Stage 1-http://www.3gpp2.org/public_html/specs/S.R0021-0_v2.0.pdf

2.5.2.3 Playout Delay/Jitter

Playout delay for multimedia streaming services is allowed to be longer because of the buffering process at terminal. The 3GPP recommended maximum playout delay is 30 seconds[35]. Additionally, the system should be able to operate under delay jitter of three times the radio link protocol retransmission time in the network.

2.5.2.4 Error Rate

According to 3GPP specification, the multimedia streaming services should operate over channels with end-to-end Bit Error Rate of the order of 10^{-3} (for circuit-switched network)[23] and Frame Error Rate in the order of 10^{-2} (for packet-switched network). These errors can be masked by employing applications at higher layers, e.g. error concealment [63] [64], error control [65] [66], buffering [67] [68], etc.

2.5.3 Streaming Protocols

Multimedia streaming protocols are required for setting up connections between different network and devices. Table 2-12 summarises the most popular multimedia streaming protocols and present their major characteristics. In particular, Real-time Transport Protocol (RTP) [73], Real Time Control Protocol (RTCP) [73], and Real Time Streaming Protocol (RTSP) [74] are briefly introduced, as they were mostly used in thesis.

Real-time Transport Protocol (RTP) [73] is both an IETF standard-RFC 1889 and an ITU standard-H.255.0 [75] and currently has been widely used for delivering real-time applications. RTP and RTCP are upper layer transport protocols. RTP provides end-to-end network transport functions suitable for applications transmitting real-time data, such as audio, video or simulation data, over multicast or unicast network services. RTP does not address resource reservation and does not guarantee quality-of-service for real-time services. RTCP is a companion protocol to RTP and allows for monitoring of the data delivery.

Real Time Streaming Protocol (RTSP) [74] is designed as a session control protocol. RTSP allows session establishment and control, as well as multimedia presentation. It does not typically transmit the streams itself. RTSP provides many benefits, for instance, enables full bidirectional delivery, ensures low overhead data delivery, provides high security

[35] 3GPP recommendation on multimedia streaming- http://www.3gpp2.org/public_html/specs/S.R0021-0_v2.0.pdf

TABLE 2-12 COMPARISONS OF THE WIDELY USED MULTIMEDIA STREAMING PROTOCOLS

		RTSP/TCP; RTSP/RTP/UDP	RTMP/TCP; RTMFP/UDP	HTTP/TCP
Organization		IETF (http://www.ietf.org)	Adobe (http://www.adobe.com)	IETF
Port Number		TCP port: 80 or 554 UDP ports:6970~9999	TCP port: 1935 UDP ports: 1024~65535	TCP port: 80
Description		A multimedia control protocol. Client can remotely control the streaming server by sending commands like 'play' and 'pause'.	RTMP server streams video bytes from any point of time according to client's request. Good for long content video	Progressive download: video is played back before completely downloaded
Cache		"cut-through", cache copies the streaming media as it arrives to client. No extra delay	Video is stored at flash player's memory	Video is stored at web browser's buffer before playback
Video Codec Suggestions		CBR preferable (http://lists.apple.com/archives/quicktime-users/2005/Mar/msg00020.html)	• Better use Constant Bit-rate (CBR) encoding; • Variable Bit-rate (VBR) encoded video might contain data spikes which can abruptly empty the flash player's temporal cache and result in annoying pause-play-pause experience;	Can safely use VBR encoding since video are progressively downloaded. Extreme data spikes might not impact during playback
Firewall		RTSP traffic and be encapsulated into inside TCP 80 and go through most firewalls. For RTP, UDP port should open.	RTMP takes two polices to deal with firewall: • If play video over port 1935 fails, RTMP over port 80 is used; • If port 80 still fails, RTMP packets are wrapped in HTTP packets;	In default, almost all the firewalls allow HTTP connections over port 80
Content Protection		Depending on the caching policy. Might need alternative solution.	Most users can only watch video content. (Of course the hackers can steal any video stream.) RTMP is useful for companies distributing movies	Users can fish the flash content out of the web browser cache
Hosting		Limited servers, e.g. RED5[36], Wowza[37]	Limited servers support RTMP services, e.g. RED5[20], FMS [38], Wowza[21]	Any HTTP web server such as Apache[39]
Cost	**Data Transfer**	Cost accounts the streamed video only	Cost accounts only the streamed video	Cost accounts since video starts downloading to cache and continues to download even users are not watching
	Implementation	Expensive	Expensive	cheap
Industry Deployment		*Quicktime (http://www.apple.com)* *Realmedia (http://www.real.com)*	*Warner Brothers (http://www.warnerbros.com/)* *HULU (http://www.hulu.com)*	*Youtube (http://www.youtube.com)* *Youku (http:/www.youku.com)*

[36] Red5-http://www.red5.org/
[37] Wowza-http://www.wowza.com
[38] Adobe-http://www.adobe.com
[39] Apache-http://www.apache.org

streaming, supports for intellectual property rights protection. RTSP works well both for large audiences and single-viewer media-on-demand.

2.5.4 Devices

Devices that receive the multimedia streaming services can be classified in terms of hardware-related parameters, such as CPU, memory, screen resolution, battery life, etc. Various server or client-based adaptation schemes might be adopted to improve streaming service experience. For instance, video server can send lower quality video to devices with smaller screen resolutions for bandwidth saving reasons; client device can appropriately request reductions in the received video quality in order to obtain longer battery life. Table 2-13 lists the most popular mobile devices on the market in 2012.

TABLE 2-13 DEVICES CLASSIFICATION

Type	Brand	CPU	Memory (RAM)	Screen		Wi-Fi Connectivity
				Resolution	Interactivity	
Smartphone	Blackberry9780	624MHz	256MB	480x360	Keyboard	IEEE
Smartphone	Google Nexus	1GHz	512MB	480x800	Touch	IEEE
Smartphone	HTC Evo 4G	1GHz	512MB	480x800	Touch	IEEE
Smartphone	HTC Desire	1GHz	576MB	480x800	Touch	IEEE
Smartphone	Samsung	1.2GHz	1GB	480x800	Touch	IEEE
Smartphone	Samsung i5500	600MHz	256MB	240x320	Touch	IEEE
Smartphone	HUAWEI	600MHz	256MB	340x480	Touch	IEEE
Smartphone	iPhone4s	800MHz	512MB	960x640	Touch	IEEE
Tablet PC	Sony Tablet S	1GHz	1GB	1280x800	Touch	IEEE
Tablet PC	Dell Streak	1GHz	512MB	800x480	Touch	IEEE
Tablet PC	New iPAD	1GHz	1GB	2048x1536	Touch	IEEE
Laptop	HP Pavilion	2.4GHz	4GB	1280x800	Keyboard	IEEE
Laptop	Dell Vostro	2.5GHz	4GB	1366x768	Keyboard	IEEE
Laptop	Thinkpad T400	2.4GHz	3GB	1280x800	Keyboard	IEEE

2.6 Summary

This chapter presented background knowledge on cellular networks, IEEE 802.x standard family, Quality of Service (QoS) and Quality of Experience (QoE), multimedia services and devices. Cellular networks play significant role in current telecommunication world and have evolved from the first generation techniques (i.e. analogue-based phone service) to the fourth generation techniques (i.e. LTE). A series of IEEE standards (802.11, 802.15, 802.16, and 802.21) have been released to provide different services for wireless communications. The concept of QoS and QoE as well as their evaluation metrics are introduced and discussed in this chapter. Multimedia streaming services including video and audio codecs, streaming protocols and devices are also presented.

In the next chapter, related works of our proposed schemes will be discussed. Bandwidth estimation, QoS-oriented multimedia delivery and resource allocation mechanisms will be presented.

CHAPTER 3

Related Works

The third chapter of this thesis presents research works related to the proposed solutions including bandwidth estimation techniques, QoS-based multimedia delivery solutions, and mathematical theories for resource management. In Section 3.1, current bandwidth estimation techniques are categorized as probing-based and cross-layer-based schemes. Section 3.2 discusses the existing QoS solutions for multimedia communications including packet adjustment-based techniques and admission control-based mechanisms. Section 3.3 introduces state-of-the-art mathematical theories used for resource management such as stereotype-based structure, fuzzy logic, clustering, and game theory. Finally, section 3.4 summaries the chapter.

3.1 Bandwidth Estimation

3.1.1 Introduction

Bandwidth estimation schemes are used by several solutions to improve the Quality of Service (QoS) of these applications [76]. Since the wireless channel capacity and the traffic characteristics change quickly, it is essential that the proposed bandwidth management scheme performs an instant and accurate estimation of the overall bandwidth. Shah et. al. [77] propose an admission control and dynamic bandwidth management scheme that provides fairness and a flexible rate guarantee. Li et. al. [78] developed a playout buffer and rate optimization algorithm for streaming in order to improve the streaming performance. It uses a bandwidth estimation solution designed for streaming networks and optimizes the streaming rate and initial buffer size based on the estimated wireless network bandwidth conditions. In an online learning system, bandwidth estimation is also required to adapt the delivery of online materials to available bandwidth [79] [80]. Research has shown that adaptation to available bandwidth reduces start-up delay [81] and consequently improves the quality of the learning experience.

TABLE 3-1 SUMMARY OF CURRENT BANDWIDTH ESTIMATION TECHNIQUES

Category	Principle	Algorithm	OSI layer	Parameters	Estimation Results	Networks	Test-bed	Ref
Packet Dispersion Technique	Measure the packet pair/train dispersion to estimate the bottleneck capacity	CProbe /bprobe	Application layer	Inter-arrival time of probing packet pair, packet size	Available Bandwidth	Wired	Real test	[83]
		Nettimer	Application layer	Inter-arrival time of probing packet pair, packet size	Capacity	Wired and WLAN	Real test	[84]
		Sprobe	Application layer	Time interval of TCP SYN packets, packet size	Capacity	Wired	Real test	[85]
		Pathrate	Application layer	Dispersion of packet train, probing packet size	Capacity	Wired	Real test	[86]
		WBest	Application layer	Capacity, interval of probing packet pair, packet size	Available Bandwidth	WLAN	Simulation and Real test	[88]
Probe Rate Model	Vary traffic load and measure the packet delay to estimate the available bandwidth of the bottleneck link	PathChirp	Application layer	Number of packet chirps, packet size, inter-spacing gap	Available Bandwidth	Wired	Simulation	[89]
		PathLoad	Application layer	Probing rate, packet size	Available Bandwidth	Wired	Simulation	[90]
		DietTOPP	Data link layer	Probe packet size, probing rate	Available Bandwidth	WLAN	Simulation	[91]
Probe Gap Model	Measure the dispersion of the packets gap to estimate the crossing traffic bandwidth, then estimate the available bandwidth. The bottleneck capacity is known	IGI/PTR	Application layer	Packet pair dispersion, single-hop gap model	Available Bandwidth	Wired	Simulation	[92]
		Spruce	Application layer	Capacity, packet pair time gap at sender/receiver	Available Bandwidth	Wired	Real test	[93]
		ProbeGap	Data link layer	Link idle time, capacity, packet pair gap	Available Bandwidth	Wired and WLAN	Real test	[94]
		AdhocProbe	Application layer	Capacity, probing packet pair dispersion	Capacity	WLAN	Real test	[95]
Cross layer-based Scheme	Monitor the channel status (idle/busy) at the data link layer and send the bandwidth information to the upper layer	IdleGap	Data link layer	Capacity, link idle duration	Available Bandwidth	WLAN	Simulation	[96]
		Shah, et.al.	Data link layer	Channel busy duration	Available Bandwidth	WLAN	Simulation	[97]

Table 3-1 summarizes the current bandwidth estimation techniques in terms of category, place in the OSI layer, objective, parameters needed, estimation results, network, test-bed type, and the reference. The next section categorizes and describes these techniques in details.

3.1.2 Probing-based Bandwidth Estimation

The common aspects of probing-based bandwidth estimation schemes are the usage of probing traffic. Probing-based schemes [82] can be categorized in the following classes: 1) packet dispersion-based techniques; 2) probe rate model-based schemes; 3) probe gap model-based techniques.

3.1.2.1 Packet Dispersion-based Techniques

Packet dispersion-based techniques estimate the bottleneck capacity by sending either packet pairs or packet trains as probing traffic:

1. **Packet Pair Probing**: the source sends multiple packet pairs **back to back** to the destination. Each packet pair consists of two packets of the same size. The dispersion of a packet pair at a link is computed as the time between the arrivals of last bits of the two packets. Assuming the probing packet size is L and the dispersed interval is R, the receiver will estimate the path capacity from equation (3-1):

$$C = L / R \tag{3-1}$$

Practically, the probing packet size L is set to the path Maximum Transmission Unit (*MTU*) size as higher values of L result in higher values of dispersion, which are easier to be measured. The packet pair probing technique was designed for the First-In-First-Out (FIFO) networks with the assumption that the two probing packets are sent out close enough in order to be queued together at the bottleneck link. Note that cross traffic might significantly impact the packet pair probing technique. For instance, the packets of cross traffic might queue between the two probing packets at the bottleneck link determining that the dispersion does not reflect the actual

bottleneck link capacity.

2. **Packet Train Probing**: packet train probing scheme extends the packet pair probing scheme by using multiple packet pairs to infer the capacity. The dispersion of a packet train at a link is calculated as the time difference between the first and last packets. The receiver measures the end-to-end dispersion R'. If N denotes the length of the packet train and L is the packet size, then the dispersion rate D is computed as in equation (3-2),

$$D = (N-1) \times L / R' \times N \tag{3-2}$$

In the case of no cross traffic, the dispersion rate D is equal to the path capacity.

Next section presents the existing bandwidth estimation schemes which use packet pair or packet train probing techniques.

Cprobe and **Bprobe** [83], developed by Carter and Crovella in 1996, are the earliest tools to measure the end-to-end available bandwidth based on the packet dispersion technique. The available bandwidth is determined by two factors: 1) the capacity of the underlying link between client and server; 2) the congestion condition of the network. Cprobe and Bprobe are used in combination to provide the available bandwidth to an application. Cprobe gives an estimation of the current congestion in the end-to-end path; and Bprobe provides the estimation of the uncongested bandwidth of a path. Cprobe estimates the available bandwidth based on the packet train dispersion technique by measuring the dispersion of a packet train with eight packets. The purpose of the Bprobe is to measure the transmission rate of the bottleneck link. The idea is to send a sequence of ICMP ECHO packets from the source to the destination and measure the inter-arrival times of the returning packets. More details of the algorithms for Cprobe and Bprobe are shown in [83],

Nettimer [84] is proposed to estimate the bottleneck link capacity using the packet pair dispersion technique introduced above. Unlike Cprobe/Bprobe approaches which can only measure bandwidth in one direction, Nettimer is able to measure bandwidth in one direction with one packet capture host and in both directions with two packet capture hosts. The capture hosts in Nettimer consist of packet capture servers and packet capture clients. The servers distribute the captured packet headers to clients and the clients perform the bandwidth estimation. The major contribution of Nettimer is the introduction of the filtering technique which can mitigate the impacts caused by cross traffic. The filtering algorithm

uses the *Kernel Density Estimation* technique which identifies the dominant factor in the distribution of packet pair dispersion and filters out packets that cause undesirable queuing. Experimental tests show that in most cases, Nettimer has less than 10% error for a variety of bottleneck link technologies such as 100Mbps Ethernet, 10Mbps Ethernet, 11Mbps WaveLAN, ADSL, CDMA cellular data, etc.

Sprobe [85] estimates the bottleneck bandwidth by utilizing the TCP protocol. The bandwidth estimation algorithm runs at the source side which sends a few TCP SYN packet pairs to the inactive port of the remote host. The receiver host replies with a TCP RST packet pair. Next on the source side, Sprobe uses the time dispersion of the received RST packet pair as an approximation to the time dispersion of the SYN packet pair. Consequently, the bottleneck capacity is estimated by using the packet pair dispersion technique. Regular SYN packets are 40 bytes without any payload data, having small probability of being appended by long overhead at the bottleneck link on the reverse path. The experiment tests involve a study of over 50, 000 uncooperative wide-area hosts comprising a variety of operating systems and networks, demonstrating that Sprobe performs well in accuracy, scalability, speed and practicality. However, Sprobe relies on correct TCP implementation at both sides.

Pathrate [86] estimates the capacity based on both packet pair and packet train techniques. It uses the research works from Paxon [87], who observe that the distribution of bandwidth measurements is multimodal. In general, packet-pair based bandwidth measurements follow a multimodal distribution and explain the causes of multiple local modes. Pathrate investigates the effects of network load, packet size variability of cross traffic, and probe packet size on the bandwidth distribution of packet pairs. It is concluded that the conventional suggestion of using MTU probing packet pair is not optimal for estimating the capacity of a path. Instead, the solution of using variable size of probing packet pairs is adopted. The proposed algorithm uses long packet trains to estimate the path average dispersion rate (ADR) which is shown as a lower bound of the capacity and an upper bound of the available bandwidth. Eventually, Pathrate estimates the capacity as the strongest local mode in the packet pair bandwidth distribution that is larger than ADR. Real test bed-based experiments show that Pathrate is quite accurate when the path capacity if not too high (below 500Mbps) and not heavily loaded.

WBest [88] employs the packet dispersion technique to estimate the capacity and the available bandwidth of the underlying wireless networks. It is designed for fast, non-intrusive, accurate estimation of available bandwidth in IEEE 802.11 networks. WBest

applies a two-step algorithm: 1) using the packet pair technique to estimate the effective capacity, (C_e), of the wireless networks; 2) using the packet train technique to estimate the achievable throughput and report the inferred available bandwidth. In the first step, n packet pairs are sent to estimate C_e, which represents the maximum capability of the wireless network to deliver traffic. The computation of C_e is given in equation (3-3), where L is the packet size, $T(t)$ is the packet dispersion at time t. For the second step, a packet train of length m is sent at rate C_e to estimate available bandwidth A, as given in equation (3-4), where R is the average dispersion rate at the receiver.

$$C_e = \frac{\int_{t_0}^{t_1} \frac{L}{T(t)} dt}{t_1 - t_0} \tag{3-3}$$

$$A = 2 \times C_e - \frac{C_e^{\,2}}{R} \tag{3-4}$$

WBest has been implemented and evaluated in the 802.11 wireless test bed. Comparison-based results demonstrate that WBest have higher accuracy, lower intrusiveness and faster convergence time.

3.1.2.2 Probe Rate Model-based Bandwidth Estimation

Probe Rate Model (PRM)-based bandwidth estimation, also referred as the Self-loading technique, estimates the bandwidth of the bottleneck link by varying the traffic load using probing traffic. Let C_{probe} denote the transmission rate of probing traffic sent from source and AB represents the available bandwidth of a path. The basic idea of PRM is to increase the probing rate until C_{probe} is higher than AB; in this case, the probing packets will be queued at the bottleneck link and introduce extra queuing delay. The value of the available bandwidth (AB) equals C_{probe} at the point where **queuing delay** starts increasing.

PathChirp [89] is a novel available bandwidth estimation tool based on a probing rate model-based technique. The term chirp refers to an exponential flight pattern of probes. PathChirp estimates the available bandwidth by sending a number of packet chirps from sender to receiver and the receiver conducts a statistical analysis to perform the estimation. A chirp consists of N exponential spaced packets with the same size. By investigating the ratio of successive packet inter-spacing times within a chirp, the packet k at which the queuing

delay starts increasing can be found. The instantaneous chirp rate E_k at packet k is considered as a simple estimated bandwidth, which can be computed based on the time difference between the arrival of packet k and packet $k+1$. PathChirp then takes a weighted average of all the instantaneous chirp rates to give a per-chirp available bandwidth. Finally, it estimates the available bandwidth by averaging across the chirp. The major advantage of PathChirp is that it uses less probing packets to estimate more accurate bandwidth, in comparison with packet pair dispersion techniques. For instance, a chirp of N probing packets is used instead of $2N-2$ packets in the packet pair technique. Additionally, by exponentially increasing the packet spacing, chirps probe the network for the range of rates $[G_1, G_2]$ Mbps using just $\log(G_2)-\log(G_1)$ packets.

PathLoad [90] estimates the available bandwidth based on probing rate methodology. The basic idea is that the one way delay of a periodic packet stream shows an increasing trend when stream bit rate is higher than the available bandwidth. PathLoad uses the Self-Loading Periodic Streams (SLoPS) to measure the available bandwidth. A periodic stream in SLoPS consists of N packets of size L, sent to the link at the rate R. If R is higher than the available bandwidth A, the one way delays of successive packet at the receiver shows an increasing trend. Such increasing trend is determined by constructing an iterative algorithm to analyse the one way delay of the stream. Let R_{max} and R_{min} represent the upper and lower bounds for A. Initially, $R_{min}=0$ and R_{max} is set sufficiently higher than A. If the rate of stream k is higher than A, (i.e. $R(k)>A$), the next SLoPS rate is reduced, (i.e. $R(k+1)<R(k)$); otherwise if the rate of stream k is lower than A, (i.e. $R(k+1)>R(k)$). The value of $R(k+1)$ is computed using equation (3-5).

$$\begin{cases} R_{\max} = R(k) & if \quad R(k) > A \\ R_{\min} = R(k) & if \quad R(k) < A \\ R(k+1) = (R_{\max} + R_{\min})/2 \end{cases} \tag{3-5}$$

The algorithm terminates when $R_{max} - R_{min}<\lambda$, where λ is user defined estimation resolution. In summary, pathload collects information for a range rather than making a single estimation. The average value of the range is the available bandwidth while the range indicates the

variation of available bandwidth. The major strength of PathLoad is that there is no significant increase in the network utilization, delay and loss.

DietTOPP [91] estimates the available bandwidth in wireless networks using probing rate-based and packet dispersion-based techniques. The authors show that the probing packet size affects the measured available bandwidth which is contrary to what is observed in wired networks. DietTOPP sends m probe packet trains with initial probing rate O_{min} and each train includes k probe packets with the same size. After the m packet trains have been transmitted, another set of probe trains are sent with a new probing rate, which is increased from O_{min} by ΔO. This process is repeated i times until the probing rate reaches the specified probing rate O_{max}. The receiver records time stamps of each arrived probe packets and measures the received probing rate m_i. DietTOPP then computes the ratio O_i/m_i for all i, where O_i is the probing rate of i^{th} probe train. The value of O_i/m_i equal 1 means unchanged dispersion of packet train and O_i/m_i higher than 1 indicates that link gets congested due to the increasing probing rate. If C represents link capacity and A is the available bandwidth, equation (3-6) is derived to infer A. Extensive tests have been performed in real test bed with different types of cross traffic, (i.e., CBR, bursty Pareto distributed traffic, etc).

$$O/m = (1 - A/C) + O/C \tag{3-6}$$

3.1.2.3 Probe Gap Model-based Bandwidth Estimation

The principle of Probe Gap Model-based (PGM) bandwidth estimation technique is that the source sends a probe packet-pair with **dispersion**, T_{in}. After the successful transmission, the terminal records a different dispersion time, T_{out}. The difference between T_{out} and T_{in} is supposed to be the time for transmitting the cross traffic under the assumption that there is a single bottleneck link. The cross traffic bit rate, R_{cross}, is then computed as $R_{cross} = (T_{out} - T_{in}) \times C/T_{in}$, where C is the capacity of the end-to-end links. Consequently, the estimated available bandwidth is $C - R_{cross}$. PGM assumes that the network capacity is known.

Initial Gap Increasing/Packet Transmission Rate (**IGI/PTR**) [92] estimates the available bandwidth according to the difference between the capacity and the cross traffic of the bottleneck link. The conventional packet pair mechanism is in general reliable for measuring the bottleneck link capacity; however, it performs poor when measuring the

available bandwidth in the presence of interferences of cross traffic. A two step-algorithm is proposed for the estimation process: 1) A *single-hop gap model* is developed to capture the relationship between the cross traffic throughput and the changes of the packet pair gap for a single hop network. The model is used to identify the conditions when packet pair gap accurately represents crossing traffic; 2) IGI/PTR are developed, based on the *single-hop gap model*, to characterize the available bandwidth. IGI/PTR determines an initial packet pair gap which can yield a high correlation between the cross traffic throughput and the packet gap. The cross traffic is then estimated by monitoring the packet pair dispersion gap after the probing packets pass through the bottleneck link. Additionally, a probing packet size around 700 Bytes proved to result in best results. Experimental results show that IGI/PTR techniques are much faster than existing methods such as PathLoad [90].

Spread PaiR Unused Capacity Estimate (**Spruce**) [93] is a tool for end hosts to estimate the available bandwidth based on the probe gap model. Similar with **IGI/PTR**, Spruce computes the available bandwidth according to the difference between the link capacity and the arrival rate at the bottleneck. Spruce requires three parameters to estimate the available bandwidth: link capacity (C), time gap of packet pair at sender (T_{in}), and time gap of packet pair at receiver (T_{out}). The value of C is assumed known. The value of T_{in} is set as the transmission time of 1500 bytes packet on the bottleneck link. The purpose is to improve the probability that the queue is not empty between the two probing packets in a pair. Spruce measures T_{out} at the receiver. The number of bytes that arrived at the queue before the transmission of second probe packet (or the cross traffic rate), is then calculated as $(T_{out} - T_{in}) \times C/T_{in}$. Consequently, the available bandwidth is estimated as the difference between capacity and the cross traffic rate. Spruce performs a sequence of probe-pair measurements and computes the average in order to improve the estimation accuracy. Experimental results demonstrate that Spruce is more accurate than PathLoad [90] and IGI/PTR. Pathload tends to overestimate the available bandwidth whereas IGI/PTR is insensitive when the bottleneck utilization is large.

ProbeGap [94] estimates the available bandwidth using the probe gap model. Existing techniques (e.g. PathLoad, Spruce) for estimating the capacity and available bandwidth always assume that the constrained link can be modelled as a fixed or well-defined raw bandwidth, with FIFO packet scheduling. However, these assumptions might break down in the context of broadband access networks such as cable model and 802.11 networks. For instance, the link bandwidth is not fixed because of token-bucket rate

regulation in cable modems[40] or dynamic link rate adaption schemes in 802.11[41]. Also, packet scheduling might not be FIFO due to the contention-based MAC of 802.11. ProbeGap aims to alleviates these problems. The basic idea is to estimate the fraction of idle time of the link by probing for "gaps" in the busy periods, and multiplying by the capacity to obtain an estimate of the available bandwidth. ProbeGap estimates the idle time fraction by gathering samples of one-way-delay (OWD) over the link. A sequence of Poisson-spaced probe packets, each with a 20 bytes payload containing a timestamp, are received by the end point to compute the OWD. The distribution of OWD samples indicates two conditions, the lower one corresponding to an idle channel and higher one corresponding to a busy channel. Therefore, the idle fraction can be identified by analysing the cumulative distribution function curve of the OWD distribution.

AdhocProbe [95] is a path capacity estimation tool based on the probing packet pair technique. It is designed for the multi-hop ad hoc wireless environment. Probing packet pairs of fixed size are sent from the sender to the receiver. The One Way Delay (OWD) is then computed at the receiver and the path capacity estimation is performed at the receiver and delivered back to the sender. Those packet pairs encountering no cross traffic are considered as the optimal samples, and the corresponding capacity is given by $C=P/T$, where P is the packet size and T is the dispersion of the packet pair. AdhocProbe has been evaluated in various test bed and shows that it is a useful and practical tool that can be deployed in real wireless networks.

3.1.3 Cross layer-based Bandwidth Estimation

The basic idea of cross layer-based bandwidth estimation scheme is to utilize the interaction between different OSI layers. Unlike probing-based techniques where the network conditions are generally predicted based on packet transmission (i.e. one way delay, packet loss, etc), cross layer-based solutions can directly obtain the channel status (via sensing). This provides faster and more accurate estimated bandwidth. The weakness of cross layer-based techniques is the requirement of modifications of standard protocols, as this might cause incompatibility when implementing in real-life environment.

[40] http://www.cablemodem.com
[41] http://standards.ieee.org/getieee802/802.11.html

IdleGap [96] is a cross layer-based bandwidth estimation tool for real-time system in wireless networks. A significant contribution is the independency from cross traffic. IdleGap estimates the available bandwidth via the ratio of free time in wireless links. To obtain the ratio of idle fraction time, an idle module located between the link layer and network layer is introduced. The idle module obtains the wireless link idle rate from the Network Allocation Vector (NAV) and sends it to the application layer. The link idle rate is computed based on the busy time of the link which can be estimated by adding up all the transactions of nodes in the network. Specifically, the transaction time of node *i* can be obtained via the sum of the differences between sending and receiving time at this node. The transaction time of other nodes is got based on idle time (OT_i) of node i. OT_i is obtained from the NAV of node *i*. Equation (3-7) computes the transaction time of all nodes, where ST_i the data sending time from node *i* and RT_i is data receiving time from other nodes to *i*.

$$T = ST_i + RT_i + OT_i \tag{3-7}$$

$$Idle\_rate = 1 - \frac{T}{total\_elapse\_time} \tag{3-8}$$

$$AB = C \times Idle\_rate \tag{3-9}$$

The available bandwidth (*AB*) is then calculated using link idle rate (*Idle_rate*) and the known capacity (C), as given in equation (3-9). *AB=C* x *Idle_rate*. In general, IdleGap accurately estimates the available bandwidth for all ranges of cross-traffic (100 Kbps ~1Mbps) with a very short observation time of 10 seconds.

Shah et al. [97] propose another estimation scheme the **Total Bandwidth Estimator (TBE)**. The basic idea is to capture the wireless channel characteristics at MAC layer by measuring the channel busy time, and uses it to infer the available bandwidth information. The TBE algorithm is located at the link layer of each 802.11 enabled node. TBE estimates the total network bandwidth perceived by each flow sourced at the node it locates. The total network bandwidth equals to the overall theoretical bandwidth (i.e. 5.5Mbps or 11Mbps for IEEE 802.11b) minus the loss due to inference and contention. The

loss is estimated from the transmission history recorded by each node. The TBE continuously measures the total bandwidth for each flow and sends it to the application layer admission control scheme.

In conclusion, probing-based bandwidth estimation schemes are not appropriate for usage in wireless networks. This is because probing traffic requires extra wireless bandwidth resources and therefore, useful data traffic might be negatively impacted due to less bandwidth being available. Additionally, current cross layer-based solutions require major modification of standard protocols which increases the implementation cost. In this thesis, I have proposed a model-based bandwidth estimation (MBE) algorithm to predict the available wireless bandwidth for TCP and UDP-based applications. MBE has two advantages in comparison with existing solutions as follows: 1) MBE does not use probing traffic and therefore does not introduce additional traffic; 2) MBE is located at the application layer and utilizes a middleware component, avoiding the requirement of modification of the MAC protocol. Details of MBE will be introduced in next chapters.

3.2 QoS-oriented Multimedia Delivery Solutions

3.2.1 Introduction

QoS is critical for providing satisfactory experience for end users and service providers. Specifically, multimedia services, such as VoIP and streaming video, require higher QoS levels than data services due to their vulnerability due to delay and jitter values. This section presents state-of-the-art research works that help increase QoS levels for multimedia services. In general, current QoS solutions can be categorized in two types: packet adjustment-based techniques and admission control-based mechanisms.

Table 3-2 and Table 3-3 provides a comparative summary of the QoS-oriented multimedia delivery solutions in terms of category, place in the OSI layer, objective, parameters needed, network, test-bed type, published year, and the reference. The next section categorizes and describes these techniques in details.

TABLE 3-2 SUMMARY OF THE STATE-OF-THE-ART QOS-ORIENTED MULTIMEDIA DELIVERY SOLUTIONS(1)

Category	Algorithm	OSI layer	Objective	Networks	Test-bed	Ref
Packet Adjustment -based Techniques	RAP	Application layer	Provide friendly TCP for real-time applications	Wired/WLAN	Simulation	[98]
	TFRCP	Application layer	Adapt transmission rate based on loss and RTT	Not specified	Real life	[100]
	TFRC	Application layer	Adapt transmission rate based on loss and RTT	Not specified	Simulation	[101]
	LDA+	Application layer	Adapt transmission rate to network condition	Not specified	Simulation	[102]
	LQA	Application layer	Adjust layered video quality using RAP	Not specified	Simulation	[105]
	QOAS	Application layer	Adapt multimedia delivery based on user perception	Wired	Simulation	[106]
	PR-SCTP	Transport layer	Improve throughput by setting retransmission threshold	Not specified	Not specified	[110]
	Shimonishi et al.	Transport layer	Adapt TCP parameters to avoid congestion	Wired	Simulation	[111]
	Lee et al.	Transport layer	Provide fair bandwidth share by stop greedy TCP	Satellite	Simulation	[112]
	Wakamiya et al.	Transport layer	Achieve fair bandwidth share between TCP and non-TCP flow	Wired	Simulation	[113]
	Iiiri et al.	Transport layer	A novel TCP window control scheme, consider channel status	WLAN	Real life	[114]
	IntServ	Network layer	Provide per-flow QoS provisioning in IP networks	Not specified	Not specified	[118]
	DiffServ	Network layer	Provide QoS per-class of traffic in IP networks	Not specified	Not specified	[119]
	Gorbil et al.	Network layer	A novel multi-hop routing protocol to enable QoS traffic	WLAN	Simulation	[120]
	Visoottiviseth et al.	Network layer	A fine-grained end-to-end QoS guarantee for handover	WLAN	Not specified	[121]
	Vaidya et al.	Data Link layer	A distributed fair scheduling algorithm at MAC layer	WLAN	Simulation	[126]
	Li et al.	Data Link layer	An error protection scheme to provide QoS for layered video	WLAN	Simulation	[128]
	Park et al.	Data Link layer	A fair QoS agent to provide per-class/per-station QoS	WLAN	Simulation	[130]
	Liu et al.	Data Link layer	A hybrid token CDMA protocol to support QoS	WLAN	Simulation	[131]
	Li et al.	Data Link layer	Provide QoS for multimedia using service differentiation	WLAN	Simulation	[132]

TABLE 3-3 SUMMARY OF THE STATE-OF-THE-ART QOS-ORIENTED MULTIMEDIA DELIVERY SOLUTIONS(2)

Category	Algorithm	OSI layer	Objective	Networks	Test-bed	Ref
Packet Adjustment-based Techniques	Zhu et al.	Cross layer	Use application/transport layer interaction to provide QoS for video streaming	Wired	Simulation	[139]
	Ferng et al.	Cross layer	Use application/link layer interaction to provide fairness and QoS provisioning	WLAN	Simulation	[140]
	Ozcelebi et al.	Cross layer	Use application and physical layer interaction to adapt video quality	WLAN	Simulation	[143]
	Xiao et al.	Cross layer	Prioritize video frame and control data bit-rate based on Application/MAC layer	WLAN	Simulation	[144]
	Chen et al.	Cross layer	Use application and MAC layer interaction to adapt retry limit and video quality	WLAN	Simulation	[145]
	Xia et al.	Cross layer	QoS support based on adaptive rate control and MAC/PHY layer interaction	WLAN	Simulation	[147]
Admission Control-based Techniques	Hadjadi-Aoul et al.	Physical layer	Alleviate the congestions in IP networks using fuzzy-based approach	Wired/ WLAN	Simulation	[153]
	Zhu et al.	Data link layer	Provide expected throughput and delay in IEEE 802.11e networks	WLAN	Simulation	[154]
	Assichadi et al.	Data link layer	Admission control based on retry limit, collision rate, TXOP	WLAN	Simulation	[155]
	Lin et al.	Data link layer	Admission control based on network conditions, error rate, retry limit	WLAN	Simulation	[156]
	Abdrabou et al.	Data link layer	Provide stochastic delay guarantees via distributed model-based admission control	WLAN	Simulation	[157]

3.2.2 Packet Adjustment-based Techniques

Many packet adjustment-based QoS solutions have been developed at different OSI layers: application layer, transport layer, network layer, link layer and cross layer. Next sections present these solutions in descending order of network layers.

3.2.2.1 Application Layer

The basic idea of application layer-based QoS solutions is to adjust the transmission rate or quality of multimedia traffic (i.e. video bit-rate) in order to better utilize the network in existing conditions. Application layer-based schemes avoid the modifications of existing protocols.

Rate-based Adaptation Protocol (**RAP**) [98] is an application layer-based congestion control mechanism for real time traffic. The motivation of RAP is to alleviate the unfairness problem between non-congestion controlled-based (i.e. UDP based service) real time applications and TCP-based applications. TCP traffic reduces its transmission rate when congestion is detected, resulting susceptibility to bandwidth occupancy by other non-congestion controlled applications. Therefore, "TCP-friendly" behaviour is significant for real time applications. According to RAP, the sender adapts the transmission rate based on *Additive Increase Multiplicative Decrease (AIMD)* algorithm. It has been shown that *AIMD* algorithm can efficiently converge to a fair state [99]. If congestion occurs, the transmission rate is reduced by half; otherwise the rate is increased by one packet per Round Trip Time (RTT). Additionally, RAP provides a fine-grained delay-based congestion avoidance mechanism using short-term and long-term RTT averages. Simulation-based experiments have demonstrated that bandwidth is fairly shared between TCP and RAP traffic. Additionally, the deployment of *RED (Random Early Detection)* queue protocol can significantly improve the fairness between RAP and TCP traffic.

TCP-Friendly Rate Control Protocol (**TFRCP**) [100] is an application layer-based solution that controls the transmission rate following the TCP approach. TFRCP determines the transmission rate based on the measured value of *loss rate* and *round-trip times (RTT)*. TFRCP consists of two sub-protocols: sender-side and receiver-side protocols. The sender computes the transmission rate with a certain time interval. In the beginning, a series of packets are sent and each packet carries a timestamp recording the sent time. The receiver acknowledges each packet with an ACK packet which also carries a bit vector of 8 bits to indicate whether the previous 8 packets were received. Packet loss is detected by checking

the sequence number and the timeout limit. The sending rate is doubled in the next round if no packets are lost; otherwise, the rate is computed based on a TCP throughput model which is the function of receiver's window size, round trip time, loss rate and timeout value. Experimental results show that TFRCP is fair to TCP and other TFRCP-enabled flows. More details of TFRCP are described in [100]

TCP-Friendly Rate Control (**TFRC**) [101] determines the transmission rate of traffic at application layer based on packet loss and round-trip time. Similar with **TFRCP**, TFRC utilizes the TCP throughput model to compute the transmission rate, but with more advanced methods to obtain the equation's parameters such as loss rate. A weighted average value of loss intervals is computed to allow for higher accuracy. The loss rate is then measured as the inverse of the weighted average loss interval. The weighted process avoids loss rate value from depending on single loss events or long time period of no loss. Additionally, TFRC provides delay-based congestion avoidance mechanism by adjusting the time between two consecutive packets. Experimental results show that TFRC provides more stable sending rate in comparison with TFRCP, while still owns high responsiveness to traffic conditions.

Enhanced Loss-Delay based Adaptation algorithm (**LDA+**) [102] adapts multimedia transmission in accordance with the network congestion state. LDA+ uses the Real-time Transport Protocol (RTP) [73] for data delivery and Real-Time Transport Control Protocol (RTCP) [73] for sending feedback (loss and delay) to multimedia senders. Similar with **RAP** [98], LDA+ also adopts *AIMD* algorithm for rate adjustment. In case no loss occurs, the additive increase value is set to the minimum of three values: ADD_m, ADD_{exp} and ADD_{TCP}. The purpose of using ADD_m is to smooth the additive value and allow fair share between flows, i.e. assigning higher bandwidth users with a lower additive value. ADD_m is determined by the bandwidth share of the sender. ADD_{exp} is used to prevent the transmission rate from exceeding to the bottleneck bandwidth. The value of ADD_{exp} converges to 0 as the bandwidth share of the flow converges to the bottleneck bandwidth. ADD_{TCP} is used to avoid an RTP flow not increasing its bandwidth share faster than a TCP connection sharing the same link. In case of loss situation, the rate r_m is decreased by r_{m-1} multiplied by $L^{1/2}$, where L is the loss fraction, but the final value should be lower than the TCP equation suggests [104]. Simulations and measurements over the internet show that LDA+ is efficient in terms of network utilization, congestion avoidance and fairness towards competing TCP connections.

Layered Quality Adaptation (**LQA**) [105] adjusts the quality of layered video to perform long-term coarse-grain adaptation, while using TCP friendly congestion control (RAP) to quickly respond to the congestion. LQA consists of two parts: 1) coarse-grain adding and dropping mechanisms; 2) fine-grain interlayer bandwidth allocation scheme. The server can perform coarse-grain adjustment on the total quantity of receiver-buffered data by adding or dropping certain layers of video stream. A fine-grain interlayer bandwidth allocation mechanism focuses on video layers themselves. When there is bandwidth available, the server increases the sending rate. If there is buffered data at receiver, then the server temporarily reduces the sending rate. LQA allows the server to trade short-term improvement for long-term smoothing of quality.

Quality-Oriented Adaptation Scheme (**QOAS**) [106] is an adaptive multimedia streaming mechanism designed for the application layer. It involves a server-located QOAS controller application and multiple instances of feedback-controlled QOAS client and server applications. The QOAS client application uses a Quality of Delivery Grading Scheme (QoDGS) to evaluate the delivery quality by monitoring the transmission related parameters (such as packet loss, delay, jitter, late packet for play out) and estimate the end user perceived quality. QoDGS considers both the short-term and long-term variation of monitored parameters in terms of estimated scores and regularly sends their weighted combined score to the server in the feedback messages. The QOAS server application uses a Server Arbitration Scheme (SAS) to analyse the received feedback reports and adjusts the delivery of video stream by varying its quality. Objective and subjective-based experiments have shown the significant performance achieved by QOAS, both in terms of the number of users and of end-user perceived quality.

Other application layer-based solutions have also been proposed to help support QoS. **Iqbal et al.** [107] have proposed a QoS scheme for multimedia multicast communications in wireless mesh networks. **Liang et al.** [108] have presented a framework for application level QoS management in services-oriented systems using AI techniques. Unfortunately, all these solutions, except QOAS which attempts to adapt quality at the users, do not consider end-user perceived quality issues include user's QoS requirements or expectations, device characteristics (screen resolution, battery life left), etc.

3.2.2.2 Transport Layer

Original TCP and UDP protocols do not support QoS for multimedia delivery. Therefore, the majority of QoS-oriented solutions at transport layer mainly optimize the TCP and UDP protocols. Several state-of-the-art schemes are introduced next.

The Internet Engineering Task Force (IETF) develops a novel transport layer protocol referred as ***Partial Reliable-Stream Control Transmission Protocol (PR-SCTP)*** [110]. It is an unreliable service mode extension of SCTP which differentiates retransmissions based on a reliability level that could be set dynamically. By using PR-SCTP, users can specify rules for data transmission. When a certain pre-defined threshold is reached, the sender abandons packet retransmission and sends the next incoming packet from the application layer. The reliability level is set based on different data types or the stream requirements.

Shimonishi et al. [111] have proposed **TCP-AV** to improve video streaming based on TCP. TCP-AV extends TCP-Reno and employs AIMD algorithm. The proposed scheme consists of two mechanisms: 1) dynamic adapt TCP parameters to stabilize TCP throughput; 2) properly reduce transmission rate to avoid congestion. In order to control the sending rate around the target rate, TCP-AV maintains a bucket counter which is an accumulation of the difference between sending rate and target rate. The congestion control parameters are tuned in order to make the bucket counter targets its specified value. To avoid congestion, TCP-AV monitors the frequency of retransmission timeout events and temporary reduces the target rate by reducing the bucket counter. Simulation tests show that TCP-AV provides better rate control for maintaining target rate, therefore better video quality.

Lee et al. [112] have introduced a **preferential suppression (PS)** scheme to suppress TCP flows which consume too much bandwidth, in order to provide fair share of reliable resources. Specifically, PS protects TCP traffic that traverses a satellite link and might drop packets of other flows. In PS, the target throughput for selected TCP flow is set as a reference value to meet QoS requirements. The probability of dropping packets is adjusted based on the bandwidth of the flows traversing a satellite link. Additionally, a TCP spoofing scheme is integrated with PS to reduce the round-trip time since lower responsive time contributes to a higher utilization of available bandwidth.The TCP spoofing mechanism locates in the congested edge router to mask the high latency of satellite links and, therefore,

increase the transmission rate of the sender. Simulation results show that target TCP flows are able to utilize a specified amount of bandwidth and thus achieve QoS objectives.

Wakamiya et al. [113] have proposed the **QoS-based TCP-Friendly Rate Control Protocol (Q-TFRCP)** to achieve the fair-share of link bandwidth between TCP and non-TCP traffic. Q-TFRCP extends the TCP-friendly rate control protocol (TFRCP) [100] in two steps: 1) estimate the application-level QoS; 2) determine the video transmission rate so that the estimated application-level QoS can be achieved. Application-level QoS consists of perceived video quality and file transfer delay. Video quality is estimated from the relationship between normalized SNR (Signal-Noise Ratio) and pre-determined QoS levels. The server estimates file transfer delay based on feedback from clients and video characteristics. It is assumed that the highest QoS is achieved when the TCP throughput is identical to the maximum rate of the video traffic. Finally, by applying Q-TFRCP where the video applications adjust their transmission rate with consideration of application-level QoS-based fairness, the QoS friendliness is improved.

Ijiri et al. [114] have developed a novel TCP window control mechanism by considering the channel occupancy status in WLANs. Traditionally, TCP congestion window size is increased gradually until congestion occurs, which aggressively obtains network bandwidth and affects all wireless stations belonging to the same access point. Each wireless station estimates the channel occupancy by analyzing the MAC layer information and delivers the channel condition to the TCP window control. The TCP window control mechanism then adjusts its traffic generation in accordance with the channel occupancy status. The traffic generation is controlled by the window flow control by adjusting the congestion window size and the advertised window size. Experimental results show that the proposed scheme reduces packet loss ratio of the CBR traffic up to 45% by empirical evaluations. However, the total throughput is slightly down.

Other solutions have been proposed to provide QoS at transport layer such as **TCP-Minimum Rate (TCP-MR)** [115], **Parallel Transport** [116], **Transport Layer Adaptable Rate Control (TARC)** [117]. However, they all require the modification of existing transport layer protocols which is complex and undesirable.

3.2.2.3 Network Layer

When the first TCP/IP protocol was standardized decades ago, most applications over the Internet were non-real-time (e.g. email, FTP). The first attempt to provide QoS at network layer started with Integrated Services and continues even today.

Integrated Services or **IntServ** [118], developed by IETF, presents a framework to provide per-flow QoS provisioning. It specifies a fine-grained QoS support system relying on **resource reservation, admission control,** and **packet QoS-based scheduling**. In IntServ-based systems, sufficient resources should be reserved at each network router for every application. Resource Reservation Protocol (RSVP) is used by network devices to explicitly notify the application's QoS requirements. If all the network devices are capable of reserving the necessary bandwidth, the applications are then allowed to transmit. Additionally, IntServ uses admission control to determine whether an incoming flow can obtain the requested QoS without affecting existing flows. When a router receives packets, the scheduler will dispatch the packets in a specific queue based on their QoS requirements. IntServ provides a tighter QoS mechanism for real-time traffic. However, the per-flow reservation processing at routers incur a significant overhead in large networks and was very difficult to be deployed.

Differentiated Services or **DiffServ** [119], developed by IETF, is a coarse-grained mechanism for classifying traffic and providing QoS per class of traffic in IP networks. Unlike in the IntServ architecture, where each flow notifies its QoS requirements to the routers, DiffServ requires each router setup with identical traffic classes to provide service differentiation. DiffServ does not rely on resource reservation and admission control, but depends on *packet prioritization*. The traffic class of each packet is marked in the **DS (Differentiated Service) field** in each packet header. These traffic classes might be extended to consider many parameters, such as IP address, application type, etc. Each DiffServ-enabled router implements **Per-Hop Behaviours (PHBs)** to determine the packet forwarding properties associated with a traffic class. In contrast to IntServ, DiffServ achieves better QoS scalability and requires no reservation and no negotiation for each flow. However, DiffServ cannot guarantee the QoS performance and it is less effective when there is large amount of high-priority traffic.

Gorbil et al. [120] have proposed a novel multi-hop hybrid routing protocol called Elessar to enable QoS traffic at network layer. Elessar combines link state topology

dissemination, source routing and on demand link cost dissemination. The source node selects the optimal routing path based on QoS requirements, network topology changes and network condition changes such as delay, loss, available bandwidth. The proposed solution does not require other OSI layers and is therefore compatible with existing networks. Simulation-based experiments show that Elessar can provide efficient QoS traffic support in small-to-medium sized mobile networks.

Visoottiviseth et al. [121] have introduced a fine-grained end-to-end QoS guarantee mechanism for handover in wireless networks. The solution uses IEEE 802.11e EDCA to obtain the traffic load of each access point and IEEE 802.11k to obtain a list of neighbour access points to reduce latency in scanning process. Fast Handovers for mobile IPv6 is considered in order to reduce packet loss and latency in handover. It is assumed that each access router adopts the Differentiated Service functions and all the APs are connected to one access router. Mobile nodes calculate scores of each access point based on both network layer status and link layer traffic load. Therefore, an optimal access point is selected for the handover. The proposed scheme can be applied to both predictive fast handover and reactive fast handover. However, the experimental evaluation is not provided.

Apart from the presented solutions, other mechanisms based on network layer have been proposed. **Wei et al.** [122] have developed an integrated QoS control scheme that combines IP layer features with reconfigurable optical layer in optical internet. **Wang et al.** [123] have proposed a QoS routing algorithm based on risk analysis suitable for IPv6. **Asokan** [124] reviews the QoS routing protocols for mobile ad hoc networks. **Dharmaraju et al.** [125] present a network layer QoS support mechanism that makes use of TORA (Temporally-Ordered Routing Algorithm) routing protocol for mobile ad hoc networks (MANET).

3.2.2.4 Link Layer

Generally, link layer-based QoS schemes in IEEE 802.11 WLANs focus on two issues: queuing and medium access control (MAC). Queuing schemes manage the packets in the queue with weighted factors or priorities. MAC-based approaches utilize the 802.11 MAC protocol to provide QoS-oriented channel access.

Vaidya et al. [126] have presented a distributed fair scheduling algorithm (DFS) based on IEEE 802.11 MAC. Similar with Self-Clock Fair Queuing (SCFQ) [127], DFS first transmits the packet with the smallest finish tag which associates with a certain backoff interval. The duration of the backoff interval is proportional to the scaling factor and inversely proportional to the weight factor of a flow. The scaling factor is used to calculate the time when a packet reaches the front of the queue, while the weight factor indicates the ability to obtain the bandwidth share. DFS performs in two steps: 1) DFS calculates the packet start and finish tag whenever a packet reaches the front of the queue. The start tag is equal to the current virtual time and the finish tag for the packet is set according to equation (3-10); 2) the packet with the smallest finishing tag is selected as the next packet to be transmitted.

$$finishtag = starttag + ScalingFactor \times packetLength / weight \tag{3-10}$$

The performance of DFS depends on the scaling factor and the weight value assigned to each flow. DFS selects larger contention window size for higher value of scaling factor and sets smaller contention window size for higher value of weight results. Therefore, the values for *scaling factor* and *weight* are adapted as a function of the contention for the channel. Experimental results show that DFS can allocate bandwidth according to the weights of the active flows.

Li et al. [128] have proposed an error protection mechanism to provide QoS for layered encoded video. The mechanism consists of two algorithms at the link layer: 1) prioritized queuing; 2) adaptive retry limit. Video packets are buffered in multiple queues and each queue is assigned with a specified priority level. Queues with lower priority are served only when all queues with higher priority become empty. Due to the dependency relationship among video frames, the loss of each video packet has a different impact on the received video quality [129]. Therefore, the proposed priority queuing scheme filters out the un-decodable packets. The values of retry limits are adapted depending on the wireless link conditions and traffic characteristics such as flow rate. Different video layers have associated different retry limits at MAC layer. By combinaing priority queuing and retry limit adaptation, video layers of different importance can receive appropriate delivery and error protection depending on channel conditions. The major limitation of the proposed solution is that it modifies the standard 802.11 protocols.

Park et al. [130] have designed a fair QoS agent (FQA) to simultaneously provide per-class QoS enhancement and per-station fair channel sharing in 802.11 WLAN. FQA includes two major components: service differentiator and service level manager. The former provides differentiated service through queue scheduling algorithm and the latter dynamically adjusts the service level of packets based on the estimated bandwidth share of each station. Packets are classified based on four traffic classes: voice, video, better-than-best effort and best effort. The voice class has the highest priority and the best effort class has the lowest priority. Before packets are queued, they are dropped with a probability according to packet delay/jitter requirements. Once packets are queued, they are served according to the service class, that is, higher priority packets are processed faster. FQA provides fairness in terms of channel access time instead of channel access opportunity or channel capacity, which proves effectiveness even when flow transmission rate is different from one station to another and channel capacity varies dynamically. The channel access time is estimated using the virtual carrier sensing mechanism. Finally, in order to balance the QoS provisioning and fairness, the service level of packets is adapted so that each user uses premium service up to its fair share. FQA is designed to be easy implementable as it does not require modifications to the MAC protocol. Simulation tests show that FQA can provide per-class QoS enhancement in terms of throughput and assure per-station fair sharing. However, other QoS factors are not studied such as delay, loss and jitter. Additionally, the authors only consider the downlink traffic fairness, and the unfairness between downlink and uplink traffic needs to be addressed.

Liu et al. [131] have presented a hybrid token-code division multiple access (CDMA) protocol to provide both QoS support and high network resource utilization in ad-hoc networks. The proposed scheme combines the CDMA mechanism and the guaranteed-access feature of the token-passing mechanism. Each network maintains one token consisting of a hop-leader address, a source address, a destination address, the number of codes available and other network parameters. The CDMA-based MAC scheme uses the token to perform the code allocation mechanism. Each station implements a modified leaky-bucket permit generation system and the queue of each station provides a permit buffer for storing the generated permits. The generation rate is determined based on a designated rate and a QoS parameter λ and is then assigned to a traffic class. The purpose of using λ is to provide fairness in the network by ensuring that all stations receive sufficient access to the network. The fairness is achieved by limiting the number of packets that can be sent and

varying the value of λ. Simulation results demonstrate that the proposed link layer scheme is effective in decreasing the packet delay and significantly shortens the length of the queue.

Li et al. [132] have proposed a solution to provide QoS support for multimedia traffic at the 802.11 MAC layer using service differentiations. The authors make two contributions: 1) analyse the optimal point where maximum throughput can be achieved; 2) propose a simple adaptive scheme that makes the system operate under the optimal operation point and, at the same time, achieve service differentiation. Service differentiation is achieved by allocating the bandwidth to individual traffic flows to satisfy a given target ratio. The proposed scheme controls the packet sending rate by adjusting the minimum contention window size in order to achieve the maximum throughput and target bandwidth allocation ratio. The optimal point where maximum throughput can be achieved is computed based on a throughput analysis model. The minimum contention window size is also computed according to the optimal point. However, the proposed solution relies on saturation and ideal channel conditions which cannot reflect situations in real world networks. Moreover, comparisons with other state-of-the-art service differentiation solutions were not presented.

Many other solutions have been proposed at link layer. **Chieochan et al.** [133] have developed a performance model for delay-sensitive multimedia streaming over a wired-cum-wireless network. **Banchs et al.** [134] have addressed the problem of providing throughput guarantees to EDCA stations in a WLAN in which EDCA and DCF stations coexist. **Meerja et al.** [135] have proposed enhanced collision avoidance (ECA) scheme for voice access category queues presented in 802.11e EDCA protocol. Other solutions [136] [137] [138] propose an extension of the DCF function of IEEE 802.11 to provide QoS support in wireless LAN.. All these solutions rely on modifications of existing 802.11 MAC protocols. Additionally, there is still a need to provide QoS-based service differentiation solutions considering both device characteristics and network QoS performance.

3.2.2.5 Cross Layer

Cross layer-based QoS solutions rely on the interaction between different OSI layers: application layer, transport layer, network layer, link layer and physical layer. For instance, channel conditions monitored at the physical layer can be used by the application layer to make more efficient QoS-based adaptations.

Zhu et al. [139] have proposed a cross layer-based QoS solution for video streaming using the interaction between application layer and transport layer. The QoS requirements of applications are translated into constraints of video encoding and sending rates. The proposed algorithm consists of three parts: 1) at the application layer, the source rate and sending rate constraints are derived based on a virtual network buffer management mechanism; 2) at the transport layer, a QoS-aware congestion control mechanism, called TCP-friendly rate control with compensation (TFRCC), is proposed to satisfy the sending rate; 3) a middleware component is designed between the application layer and transport layer at both sender and receiver. At the receiver, the middleware sends feedback information (e.g. the amount of received video frames) collected at the application layer and TFRCC to the sender. At the sender, the values source rate and the sending rate are determined within the middleware. The long-term TCP-friendliness is provided by a rate compensation algorithm. Comparison-based experiments show that the proposed approach can better support the QoS requirements of applications, and significantly improve the playback quality and improving quality smoothness.

Ferng et al. [140] have described the design of a cross layer-based scheme to provide both fairness and QoS provisioning in IEEE 802.11e WLANs. The proposed solution consists of four scheduling schemes: Enhanced Distributed Deficit Round Robin with Backoff Interval (EDDRR-BI), Enhanced Distributed Deficit Round Robin (EDDRR), Enhanced Distributed Elastic Round Robin with Backoff Interval (EDERR-BI) and Enhanced Distributed Elastic Round Robin (EDERR). In EDDRR-BI, three types of deficit count rather than a single type of deficit count as in DDRR [141] are defined for access categories of audio, video, and data. Deficit counts accumulate linearly and are proportional to the desired throughput within each segment. The authors introduce the idea of "deficit count" which records the transmission deficit for each flow and relates to backoff intervals for non-failure events which are events excluding collisions and failed transmissions, i.e., events of a busy medium and a successful transmission. EDDRR defines the same deficit counts as in EDDRR-BI. EDDRR works analogously to 802.11e EDCA except the cancellation of the backoff procedure for non-collision events. EDERR-BI-based scheduling scheme employs some elastic and adjustable amounts of traffic data allowed for transmission called *"allowances"* [142]. Three types of allowance are defined in EDERR-BI for the access categories audio, video, and data. Traffic associated with any particular allowance class can be consecutively transmitted until the amount of traffic data exceeds the allowance. EDERR works like 802.11e EDCA but removes the backoff procedure for non-collision

events and changes the IFS (i.e., AIFS) by using the mapping from the allowance to the IFS. Experimental results show that the proposed scheme performs better QoS at station-level and records improved flow-level fairness.

Ozcelebi et al. [143] have presented a cross layer-based multi-user video adaptation and scheduling scheme for wireless video communications. A multi-objective optimization (MOO) framework is proposed to consider both application layer and physical layer conditions. The available channel resources and the remaining video playback time for each flow are obtained based on feedback from physical layer and application layer. MOO schedules users that experience the best compromise between the shortest remaining playback time and the largest available video throughput enhancement, with the best video quality. Experiments demonstrate that the number of undesired pauses during playback and initial pre-roll delays are considerably reduced by this technique compared to the state-of-the-art schedulers at the same average video rate.

Xiao et al. [144] have proposed two QoS schemes to improve MPEG4 video transmission quality in WLANs: 1) a prioritized frame transmission scheme between the MAC layer and the application layer; 2) data transmission control scheme for IEEE 802.11e. The authors demonstrate that higher throughput does not always translate into a better quality of MPEG-4 video. Therefore, a cross layer-based approach using prioritized frames is designed to improve the received quality of MPEG-4 video. A five-steps algorithms is applied as follows: 1) P and B frames will be discarded at receiver's MAC layer if the corresponding I frame is lost; 2) MPEG-4 frames are prioritized at MAC layer so that I frames have higher priority than P frames, and P frames have higher priority than B frames; 3) frames should be deleted if the pre-defined time deadline is reached; 4) B frames will be dropped if the delay between an I frame and the next P frame is too long; 5) In the transmission queue of MAC layer, frames are reorganized according to the dependency and priorities. The proposed solution also attempts to reduce the collisions in order to control data transmission. The MAC layer parameters such as Arbitrary Inter Frame Space (AIFS), minimum and maximum Contention Window are dynamically adapted based on the observations of frame transmission behaviours so that when the number of active data stations is large, throughputs for voice and video flows are protected by increasing the initial contention window size and inter-frame space for the best-effort data traffic. The frame transmission behaviours can indicate the traffic conditions. For instance, consecutive successful transmissions mean that the traffic load condition is reasonably good, whereas

consecutive dropped frames mean that either the traffic condition is bad or the channel condition is bad. Simulation results have demonstrated the advantage of the proposed scheme in terms of throughput and MPEG-4 quality.

Chen et al. [145] have proposed a cross-layer based content-aware retry limit adaptation scheme for video streaming over IEEE 802.11 WLANs. The error propagation effects on each packet are estimated to determine the values of retry limits. Packets with higher loss impact are assigned higher retry limits. The impacts of packet loss are estimated using the pixel-level loss impact metric [146] which is the product of two parameters: pixel reference count (PRC) and pixel-wise concealment error (PCE). Additionally, the backoff time for each retransmission is estimated based on the Markov chain model [42]. A retransmission packet scheduler is proposed based on the backoff estimation to prevent useless waiting time. A packet will be discarded before a retry if the number of retries exceeds the retry limit or if the estimated arrival time of a packet is higher than the packet's presentation deadline. Experimental results show that the proposed scheme can effectively alleviate the error propagation and assure the on-time arrival of packets, so as to improve video quality.

Xia et al. [147] have developed a QoS enhancement scheme based on Adaptive Rate Control (WFS-ARC). WFS-ARC dynamically adjusts the data transmission rate at the sender's MAC layer, based on the channel condition information provided by the PHY layer. The adaptive rate control algorithm includes multi-rate retransmission which aims to quickly react to the short-term channel variations and reduce the data rate fluctuation in the long-term. On top of the MAC layer, a Logic Link Control layer-based scheduler is implemented to opportunistically schedule the packet transmission to the most promising user and satisfy the fairness constraints. The most promising user is selected by considering a general trade off model which maximizes a utilization function. Simulation based experiments show that the proposed scheme can significantly improve the system throughput.

Other solutions have been proposed to provide cross layer-based QoS enhancements. **Alonso-Zarate et al.** [148] have proposed an enhanced cross layer-based scheduling mechanism which employs a virtual priority function to reschedule transmissions. **Tang et al.** [149] have developed a cross-layer approach to investigate the impact of physical-layer infrastructure on data-link-layer quality-of-service (QoS) performance over wireless links in

[42] J. Bather, Decision Theory. New York: Wiley, 2000.

mobile networks. **Liu et al.** [150] have presented a cross layer-based solution for multi-user scheduling at the data link layer, with each user employing adaptive modulation and coding at the physical layer. Zhang et al. [151] review the study on the cross-layer paradigm for QoS support in multi-hop wireless networks.

3.2.3 Admission Control-based Techniques

In wireless networks, the purpose of admission control is to support QoS provisioning in terms of signal quality, call blocking and dropping probabilities, packet delay and loss rate, and transmission rate [152]. New calls are accepted or rejected to the network by the call admission scheme based on predefined criteria, most of the time considering the network conditions.

Hadjadj-Aoul et al. [153] have presented an admission control mechanism to alleviate the congestions in converged IP and broadcasting networks. The proposed call admission control scheme is based on an adaptive fuzzy-based approach, which aims to overcome issues related to variable link capacity, flow characteristics, high computational complexity, etc. The basic idea of the fuzzy control is to stabilize the router buffer utilization by blocking or accepting new incoming connections. The blocking probability is calculated based on queue status: 1) queue error which is the difference between actual queue length and the reference queue length; 2) the queue error variation. The values of the blocking probabilities range from 0 (smallest) to 8 (largest). Additionally, the fair bandwidth share among downlink traffic flows is guaranteed in the event of congestion. Simulation results show that the new admission control scheme can prevent downlink congestion and fairly allocate network resources among terminals.

Zhu et al. [154] have developed a novel call admission control mechanism to provide expected throughput and delay performance in IEEE 802.11e wireless networks. The proposed scheme sets the achievable per frame throughput and access delay as the admission decision criteria. There are four steps for the admission control algorithm: 1) Calculate the probabilities of collision and successful transmission for each new arrival call; 2) Calculate the maximum allowable average slot lengths such that the mean access delay is less than or equal to the delay bound; 3) Calculate the TXOP (Transmission Opportunity) value that needs to be allocated for each call; 4) Calculate the average slot length (using the assigned TXOP duration) of all admitted calls. Simulation results show that the proposed admission

control mechanism is effective in providing the expected throughput and mean access delay performance objectives for the admitted calls.

Assichadi et al. [155] have presented a novel flow-based admission control scheme which adjusts channel access parameters (CAPs) based on channel conditions. Three admission control related parameters are introduced: the retry limit value, the maximum tolerable collision rate (CR_{Max}) and transmission opportunity. The station computes the three parameters for every traffic stream based on its QoS requirement. The station then makes an admission decision whether to accept the new traffic by comparing its CR_{Max} with the current channel collision rate (CR_{cur}). If CR_{Max} is lower than CR_{cur}, the traffic required delay and dropping rate cannot be satisfied under the current channel conditions, and accordingly, the traffic request is rejected. Otherwise, if the request is admitted, the corresponding station will forward this request to the AP for further admission process based on the current available medium resources. The medium resources used by real-time streams can be calculated by the product between the TXOP requirement of each admitted flow and its current corresponding Surplus Bandwidth Allowance (SBA). SBA is a ratio computed based on the total allocated time period divided by the time period stated by the application and required for a successful transmission. The proposed algorithm has been compared with other admission control schemes, i.e. F-DAC and DAC. Tests results show that the proposed admission control scheme achieves better channel utilization in terms of the total number of admitted flows and the overall system throughput.

Lin et al. [156] have presented an efficient admission control algorithm for IEEE 802.11 DCF. In contrast to conventional approaches, both saturated and unsaturated networks are analysed and the impact of error rate and retry limit are considered. The term residual bandwidth is introduced which represents the difference between the saturated throughput and the unsaturated throughput. The effective bandwidth refers to the amount of bandwidth needed to satisfy the connection request in order to guarantee the pre-defined loss rate and bit rate. The effective bandwidth is calculated depending on connection distribution, buffer size, maximum tolerate packet loss rate and bit-rate. The proposed admission control algorithm is based on the residual bandwidth to admit the traffic in order to meet the QoS requirements. When a station connects to the Access Point Node (APN), it sends a request message including the QoS requirements. Upon receipt of the request message, the APN calculates the residual bandwidth to compare the effective bandwidth and the residual bandwidth, and if the latter is higher, APN respond to the mobile station with admission

approval. Otherwise, the APN rejects the request. Simulation-based experiments demonstrate that the admission control algorithm is efficient and determine better utilization of network resources.

Abdrabou et al. [157] have proposed a new approach to provide stochastic delay guarantees via fully distributed model-based call admission control for IEEE 802.11 single-hop ad-hoc networks. A stochastic link layer channel model is developed to mimic the variation of channel status based on a Markov-modulated Poisson process (MMPP). The call admission control algorithm consists of five-steps: 1) the new node attempting to join the network obtains the network information including the number of active nodes and traffic source; 2) the new node calculates its average traffic rate (λ); 3) the node calculates the service rate of the queue; 4) the node compares the value of λ and the value of λ_{sat} after its admission where λ_{sat} is the saturation traffic load. If $\lambda > 0.8\lambda^{sat}$, the node is not admitted; 5) Let μ representing the upper bound of QoS violation probability, D_{max} representing the delay bound among the different service classes, D_{act} denoting the delay that results in the violation probability less than or equal to μ, and N is the number of contending stations. Since all the nodes equally share the same channel, if $D_{max} \geq N x D_{act}$, the node is admitted to the network. Simulation results demonstrate that the MMPP link-layer model and the calculated effective capacity can be used effectively in allocating resources with stochastic delay guarantees.

Other QoS-oriented admission control solutions have been proposed. **Didi et al.** [158] have proposed a dynamic admission control algorithm for 802.11e networks, which adapts to the overall traffic load, number of best effort AC, and position of QoS-enabled stations. **Liu et al.** [159] have presented a dynamic admission control scheme based on the delay analysis model with the aim to guarantee the QoS of existing users and handoff users. **Dini et al.** [160] have developed a novel call admission control scheme taking into account the loss of channel time due to medium contention and the coexistence of VoIP traffic with background traffic such as TCP data flows. **Shin et al.** [161] have proposed a novel call admission control which looks at the Queue size in order to predict the traffic load.

In conclusion, current packet adjustment techniques and admission control solutions provide efficient QoS provisioning for multimedia applications. However, none of these solutions support both high QoS provisioning and QoS differentiation for delivering multimedia services to heterogeneous devices. Additionally, these solutions lack wireless network conditions awareness, as wireless communications are affected by interference, collisions, link rate adaptation, etc. In this thesis, an intelligent prioritized adaptive scheme

(iPAS) is proposed at the application layer. iPAS can provide both QoS provisioning and QoS differentiation by considering network conditions, service QoS requirements, and device characteristics (screen resolution, battery life left, etc.). Furthermore, due to the utilization of an effective bandwidth estimation algorithm, iPAS performs very well in wireless networks. More details of iPAS will be provided in the next chapters."

3.3 Mathematical Theories for Resource Management

3.3.1 Introduction

Resource is a very general concept, which can refer to natural resources (e.g. oil, water, electricity, cars, airplanes, etc.), human resources (e.g. workforce of an organization), computer hardware resources (CPU, memory, storage, DVD recorder, etc), and even virtual resources. In telecommunications world, the most expensive resource is the highly limited bandwidth. The significant increase in the amount of multimedia services, computers, mobiles, and Internet users has created an urgent requirement for rich network resource. In particular, the wireless LANs are under critical pressure to provide expected network bandwidth to the end users. Unlike in wired networks, wireless channel are prone to noise, interference, signal attenuation, etc, which further reduce the limited wireless network resources. One option to increase the network resources is to build more broadband networks and more powerful network servers. However, this relies on significant budgets. Another option is to use efficient resource management solutions in order to improve the network utilization. This section introduces the widely used mathematical theories that have been employed for resource management in IP-based networks. These techniques include stereotypes, fuzzy logic, clustering, and game theory.

Table 3-4 provides a comparative summary of the most widely used mathematical theories for resource management in terms of category, objective, parameters needed, published year, and the reference. The next section categorizes and describes these techniques in details.

TABLE 3-4 SUMMARY OF THE STATE-OF-THE-ART MATHEMATICAL THEORIES IN RESOURCE MANAGEMENT

Category	Solutions	Objective	Networks	Test-bed	Ref
Stereotypes	Muntean et al.	A QoS-aware adaptive system that controls web content based on user perception	Not specified	Simulation	[163]
	Amundsen et al.	A QoS-aware self-managed mobile computing system that can adapt to network loaded conditions	Wired/WLAN	Simulation	[164]
	Ortiz et al.	Develop a QoS model considering price, delay, response time, throughput	Not specified	Not specified	[165]
Fuzzy logic	Lo et al.	Propose a fuzzy channel allocation controller for hierarchical cellular system	Cellular	Simulation	[166]
	Chandramathi et al.	A dynamic bandwidth allocation scheme using fuzzy logic	Wired	Simulation	[167]
	Niyato et al.	A delay based admission control algorithm using fuzzy logic	Broadband Wireless	Simulation	[168]
	Todinca et al.	An admission control algorithm using fuzzy logic to achieve QoS differentiation	GPRS/EGPRS	Simulation	[169]
Clustering	Su et al.	A clustering-based multichannel communications scheme to improve security when delivering multimedia data	WLAN	Simulation	[174]
	Thenmozhi et al.	A clustering-based resource allocation in grid environment	WLAN	Simulation	[175]
	Hatoum et al.	A Femtocell cluster-based resource allocation using OFDMA technology	WiMAX, LTE	Simulation	[176]
	Cheng et al.	An intra-cluster resource allocation approach considering power allocation, subcarrier allocation, packet scheduling, and QoS support	Wireless Mesh Networks	Simulation	[177]
Game Theory	Wang et al.	Use game theory to deal with resource allocation of manufacturing resources	Not specified	Not specified	[181]
	Taleb et al.	A QoS negotiation scheme to provide efficient network utilization using auction theory, which is a subfield of game theory	Wireless	Not specified	[182]
	Berlemann et al.	A game theory-based resource allocation scheme for multiple wireless networks that are sharing frequency bands	WLAN	Simulation	[183]
	Tan et al.	Improves overall throughput using Nash equilibriums	WLAN	Simulation	[184]

3.3.2 Stereotypes

An effective bandwidth resource allocation scheme must be able to deal with uncertain and variable information related to the wireless channel, multimedia traffic and devices. These requirements can be satisfied by using a stereotypes-based resource allocation.

Stereotypes for managing groups were first introduced by Rich in the Grundy system [162] and they are still widely used by many QoS-oriented adaptive solutions. Stereotypes are defined as classes (groups) described by a set of features, which include attributes. Each stream will belong to every stereotype group with a certain probability depending on group features. The final resource management decisions are suggested by combining the features and the probabilities. Stereotypes make a powerful probabilistic analysing tool for dealing with a wide range of uncertain events, and are useful especially in the case of variable wireless environments.

Muntean et al. [163] have presented a QoS-aware adaptive web-based system (QoSAS) that controls web content based on user-perceived QoS. The basic idea of QoSAS is to use the stereotypes-based model which considers user-perceived performance in terms of different QoS metrics. Stereotype classes are used to construct and infer additional information about a user's perceived performance characteristics. A stereotype class contains characteristics of a group of users and associates with a list of features. These features include download time, client's throughput, round-trip time, HTTP version and the perceived-performance as suggested by user. Each feature has an occurrence probability according to Poisson distribution. Combining all the features and their occurrence probability results in the suggestions on the content constraint such as the number of embedded objects in the web page, the dimension of the based web page, and the total dimension of the embedded components. Simulation tests demonstrate that the proposed stereotypes-based model helps QoSAS improve end user's satisfaction. Also, a large improvement in QoS performance is achieved with minimal impact on content.

Amundsen et al. [164] have developed a QoS-aware self-managed mobile computing system that can adapt itself to the variance of traffic load, system failure and user demands. QoS characteristics and context properties are modelled using the Uniform Modelling Language (UML) which enables that models can be integrated into other UML compliant components. Two **UML-based stereotypes** are introduced: *QoSContext* represents quantifiable properties of the context elements and *Property Type* represents the quantification of one context property. The two stereotypes include a QoS prediction function which takes two inputs: context values and QoS values for the classified context

elements and resources. The Object Constraint Language (OCL) is used to specify the constraints (context dependencies and QoS values), and to correlate the QoS prediction functions to property types. By using these stereotypes constraints, the mapping from UML models to the service plans can be automatically created. Service plans provide four functions: 1) create a link between a service type and an implementation of the type; 2) define service composition and parameter configuration of the implementation; 3) illustrate dependencies to context elements; 4) introduce QoS characteristics of the implementation. The proposed solution significantly simplifies the software engineering of applications with QoS requirements and context dependencies.

Ortiz et al. [165] have developed a model-based approach for the implementation of QoS monitors by describing them as platform-independent models. The QoS criterion in terms of execution performance is focused, which includes execution price, latency, response time and throughput. In order to model the QoS criteria, the term QoS profile[43] is used. An abstract **stereotype QoS_Criterion** is defined to extend operation or interface meta-classes. To clarify the motivation, an example on university courses is introduced. The study consists of five web services: *PreregistrationService*, *RegistrationService*, *ExamOpportunityService*, *AcademicResultsService* and *TeacherService*. Additionally, two QoS criteria, latency and response time, are used in the study. In this regard, *NewPreregistration* in the interface offered by *PreregistrationService* and the interface offered by *RegistrationService* are stereotyped with the stereotype latency. In the client side model, the required interface *AcademicResultsService* is stereotyped with response time. Also, *bring Forward Exam* and *Cancel Exam* are stereotyped in the required interface *Exam Opportunity Service IF*, as well as *teacher Search* in *Teacher Service IF*. It is shown that the proposed approach enables the named code remains well structured and completely decoupled from the main functionality of the web service-based system.

3.3.3 Fuzzy Logic

Fuzzy logic is designed to get a subjective understanding of the way to best control the system. It is a mathematical theory that attempts to imitate the human decision logic and aims to model the imprecise concepts or rules. For instance, fuzzy logic uses the human knowledge-based fuzzy set of membership function {e.g., excellent, good, normal, bad} and

[43] OMG UML Profile for ModelingQoS and Fault Tolerance Characteristics and mechanisms, http://www.omg.org/technology/documents/formal/QoS_FT.htm

fuzzy rules {e.g., IF, THEN, ELSE, AND, OR, NOT}. The fuzzy set uses the concept of degree of membership, i.e., how much a factor is in the fuzzy set. In contrast with the traditional logic theories which use a two-valued logic true or false, Fuzzy logic employs a multiple-valued logic. For instance, the truth value of fuzzy logic variables ranges between 0 and 1, indicating ranges between false and true. There are three components in a fuzzy logic control system, as shown in Fig. 3-1: 1) **fuzzifier**, which is used to map the input variables into fuzzy sets; 2) **fuzzy logic controller**, which computes the output solution based on the redefined fuzzy rules; 3) **defuzzifier**, which transforms the output solution to the actual output expected. Fuzzy control rules do not require significant computational cost and tend to show a smoother response in comparison with conventional systems. Fuzzy logic has been employed by many mobile communication systems, as solution for resource management.

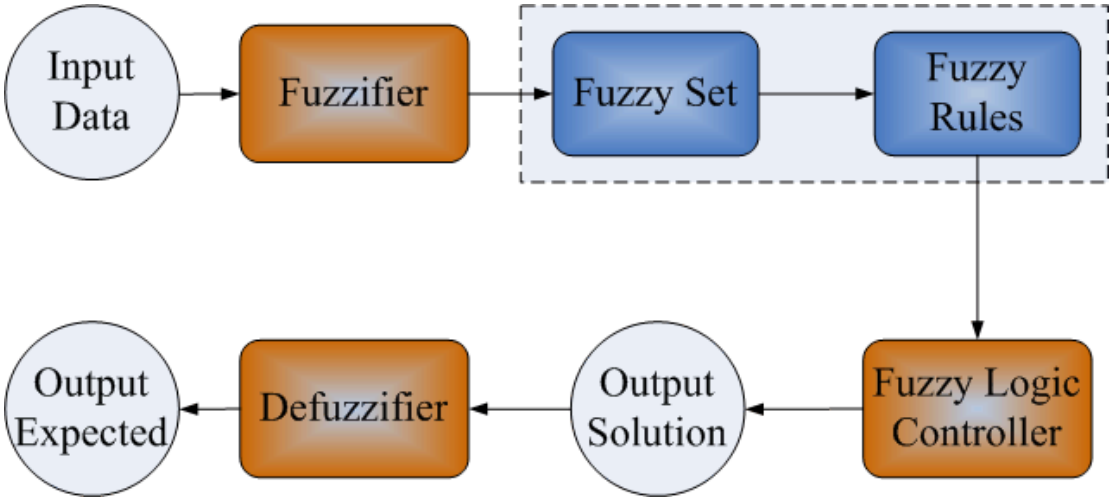


Figure 3-1 Components in Fuzzy logic

Lo et al. [166] have proposed a *fuzzy channel allocation controller* (FCAC) for hierarchical cellular systems. FCAC consists of a **fuzzy admission threshold estimator** and a **fuzzy channel allocator**. The purpose of the fuzzy admission threshold estimator is to determine the admission decision thresholds for the fuzzy channel allocator and guarantee the QoS requirement of handoff calls. The handoff failure probability is used as the evaluation metric of QoS and the resource availability is the input linguistic variables for the fuzzy set. The admission thresholds for macro cell *0* and micro cell *i* are sent to the fuzzy channel allocator when a new call comes. The fuzzy channel allocator takes three types of linguistic input variables to determine the channel allocation: the speed of user, the channel utilization in macro cell 0 or in micro cell *i*, and the resource availability in macro cell *0* or in micro cell *i*. The if-then fuzzy rules and the Sugeno's position-gradient type reasoning

method are employed by the fuzzy allocator, which produces two output instances, *O1* and *O2*. *O1* is represented as {accept, weak accept, weak reject, reject} and indicates whether the new call is admitted or not. *O2* is referred to {macro cell, micro cell} and indicates whether a macro cell or a micro cell is allocated. The utilization of fuzzy logic provides a flexible channel allocation decision, not only "accept" and "reject" but also "weak accept" and "weak reject". Simulation results show that FCAC can always guarantee the QoS requirement of handoff failure probability for all traffic loads, meanwhile, improve system utilization.

Chandramathi et al. [167] have presented a dynamic bandwidth allocation scheme using fuzzy logic for heterogeneous sources with multiple QoS requirements. The fuzzy logic controller is used to smoothly adapt to the variability of the traffic rate to the QoS requirements. Considering the ATM networks, the bandwidth to be allocated depends on the traffic arrival rate λ and the QoS requirements, both of which are stochastic. Therefore, the low value of λ might degrade the expected QoS. Fuzzy logic theory is employed to capture the uncertain characteristics of traffic such as λ, cell loss rate (CLR), cell transfer delay (CTD) and cell tolerance variation (CTV). In fuzzy logic, λ, CLR, CTD and CTV represent the fuzzy sets of the input linguistic variables and bandwidth is the single value in fuzzy set of the output linguistic variable. The fuzzy logic controller applies a two-step algorithm: 1) Three trapezoidal membership "LOW", "MEDIUM" and "HIGH" are defined to represent three QoS inputs (CLR, CTD, and CTV). The arrival rate-related trapezoidal memberships are defined using "VERY LOW", "LOW", "MEDIUM", "HIGH" and "VERY HIGH". Subsequently, the number of fuzzy rules needed is 3x3x3x5=135; 2) The output bandwidth membership functions for each traffic source is assumed to be Gaussian and the centre average de-fuzzy function estimates the crisp bandwidth values for each of the sources from these membership functions. The proposed bandwidth allocation scheme was validated in MATLAB and the results show that the required QoS can be satisfied by appropriately tuning the fuzzy logic controller.

Niyato et al. [168] have proposed a delay-based admission control algorithm using fuzzy logic for orthogonal frequency division multiple accesses (OFDMA) broadband wireless networks. The admission decision is made based on the traffic source parameters and delay requirements of the traffic. Whenever a new connection starts, four parameters, such as normal rate, peak rate, probability of peak rate, and traffic intensity, are delivered from the mobile stations to the base station. These parameters are fuzzified into fuzzy sets

("LOW","MEDIUM","HIGH") using the corresponding membership functions. The traffic source estimator estimates the traffic intensity as the output. The traffic intensity and the measured channel quality information (Signal-Noise Ratio) are used to obtain the number of sub-channels to be assigned. The number of allocated sub-channels and the fuzzified amount of load are used to determine whether an incoming connection can be accepted or not. Specifically, the results of admission processor is defuzzified so that the base station accepts the new connection with a probability. The performance of the proposed admission control algorithm is evaluated by simulation in terms of average number of connections and blocking probability.

Todinca et al. [169] have designed a new session admission control algorithm using fuzzy logic to achieve QoS differentiation and improve system utilization. The fuzzy logic controller takes two linguistic variables: network load and mobile node's precedence. The network load is defined as the sum of mobile node weights for all active users, divided to the number of channels available for data traffic. The mobile node's precedence depends on the sending delay which is the time necessary to transfer a file belonging to certain mobile node. Both network load and mobile node's precedence are mapped into the same fuzzy set {low, medium, high}. The admission decision is produced based on the output of fuzzy logic controller with linguistic terms **{reject, weak reject, weak admit, strong admit}**. When the network load is low, all users are accepted, when the network load is medium, then the low precedence mobile nodes are rejected, while when the network load is high, only the high precedence mobile nodes are accepted. Simulation results show that the proposed scheme is capable to maintain the QoS targets for users and to ensure QoS differentiation. The call dropping probability is zero and the call blocking probability for high precedence users is very low (less than 3%).

Other resource allocation solutions using fuzzy logic have been proposed such as [170] [171] [172] [173]. Although fuzzy logic based resource management has been applied to many areas; it is still controversial since the time it has been firstly used.

3.3.4 Clustering

The purpose of clustering is to group (or cluster) objects in such a way that objects in the same group resemble each other more than those objects in other groups. Clustering has been widely used in data analysis field, i.e. machine learning, image processing, bioinformatics, etc. Clustering itself is a process to be implemented rather than a specific algorithm. The

algorithms for solving a clustering process can be different significantly depending on the cluster model which defines clustering members and clustering rules. Typical cluster model include: connectivity models, centroid models, distribution models, density models, subspace models, hard clustering, soft clustering, overlapping clustering, etc. This section summarizes current resource management schemes utilizing clustering technology.

Su et al. [174] have proposed a clustering-based multichannel communications scheme to support delivery of both security information and multimedia applications. Each cluster selects one head node to manage the channel assignments for cluster member vehicles. The head node in one cluster delivers real time safety messages within its own cluster and forwards these messages to its neighbouring cluster head. The transmission of safety messages in one cluster uses contention free channel access in order to satisfy the real-time applications. Additionally, the communication between head nodes of different clusters uses the contention-based 802.11 MAC. Each cluster head employs a scheduling scheme over the Cluster Range Control (CRC) channel to send/receive safety messages and coordinate the cluster member vehicles. The CRC channel divides the regular time into time intervals with equal lengths. The length of each time interval (T) assigned to one cluster member is determined by both TDMA scheme and the number of cluster members. The TDMA scheme guarantees that each vehicle within a cluster has a chance to transmit data in every time interval T. Simulation results show that the proposed scheme can efficiently support the non-real-time traffic while guaranteeing the real-time delivery of the safety messages.

Thenmozhi et al. [175] have developed a cluster-based resource allocation scheme including resource monitoring and scheduling in grid environment. The *Mobile Grid* is divided into clusters and each cluster maintains one cluster head (CH). A master server (MS) is introduced to control all the clusters and update all the CH information. Each CH uses a monitoring agent (MA) to predict the mobility of the cluster nodes and monitor the resource availability. When MS forwards the job request of a user to the preferred CH, the CH schedules the jobs based on two conditions: the predicted time for resource availability and sufficiency of the resources. The job is allocated with the corresponding resources by the CH if it satisfies the above conditions. Otherwise, the job request will be forwarded to another CH. This process is repeated until the job is successfully assigned. The completed job is returned back to the server through the corresponding CH. The server then returns it to the requesting user. Simulations results show that the proposed scheme achieves good throughput with the reduced delay and energy consumption.

Hatoum et al. [176] have proposed a *Femtocell Cluster-based Resource Allocation (FCRA)* scheme in femtocell networks using OFDMA technology. The motivation is to associate the best spectrum set of frequency or time resources with each femtocell station while minimizing the interference. FCRA is designed based on the Min-Max optimization process and involves three steps: 1) femtocell clusters generation; 2) resource allocation within each cluster; 3) mitigate collisions among different clusters. Each femtocell maintains a specific one-hop neighbour list including its interfering femtocells. A femtocell is selected as the cluster head if it has the highest interference degree among its one-hop neighbours, and in this case, the associated neighbouring femtocells are the cluster members. Once femtocell stations are grouped into one cluster, the cluster head performs the resource allocation within each cluster in order to satisfy the femtocells' requirements while avoiding interferences. Additionally, the interference between cluster members located at the edge of neighbouring clusters is also reduced by a simple coordination mechanism. Users who are suffering from contention will send feedback to its associated femtocell and notify about the collision. Then each femtocell tries to resolve contention by sampling using a Bernoulli distribution. Consequently, the femtocell decides whether to remove the allocated resources from the attached user or not. Performance evaluation tests show that FCRA converges to the optimal solution in small-sized networks and outperforms two prominent related schemes (C-DFP and DRA) in large-sized ones.

Cheng et al. [177] have proposed an intra-cluster resource allocation approach taking into account the power allocation, subcarrier allocation, packet scheduling, and QoS support. The authors consider a wireless mesh network where the mesh routers are grouped into a number of clusters. In each cluster, one router is selected as the cluster head which provides timing information and performs resource allocation. It is assumed that each subcarrier can only be allocated to one transmission link in a cluster. In the case of joint allocation of transmit power, subcarriers, and timeslots, the achievable transmission rate is computed using the Shannon capacity formula. Additionally, the proposed resource allocation algorithm combines the Karush-Kuhn-Tucker (KKT)-driven approach and a genetic algorithm (GA)-based approach. The complexity of the Combined-KKT-GA is low, resulting in a preferred candidate for practical implementation. The novel resource allocation approach is shown to achieve good system performance in terms of throughput and packet dropping rate.

Other clustering-based resource allocation schemes have also been proposed. **Cheng et al.** [178] have proposed a node clustering algorithm with subcarrier allocation in wireless mesh networks with QoS support. **Bashar et al.** [179] have studied the admission control issue and resource (subcarriers, power and bit-loading) allocation by clustering subcarriers. **Wang et al.** [180] have consideried an adaptive multi-user resource allocation for the downlink transmission of multi-cluster multi-carrier Direct Sequence-CDMA networks.

3.3.5 Game Theory

Game theory applies to a regulated circumstance where a player's success is based on the choices of others. Each player acts by determining a best strategy to achieve his/her own interest. In this case, the outcome of the game can be predicted exactly or probabilistically. There are two types of strategies used by players: 1) a player makes decision deterministically; 2) at least one player makes decisions with a probability distribution. Many game theory models have been developed to describe different situations such as non-cooperative game/cooperative game, symmetric/asymmetric game, zero-sum/non-zero-sum game, simultaneous/sequential game, discrete/continuous game, etc. Game theory techniques have been applied for system optimization in wireless networks, such as radio resource management, admission control, etc. For instance, the players can be modelled as the end users, service providers, routers, applications, etc.

Wang et al. [181] have used the game theory approach to deal with the issue of product-mix resource allocation decision under limited manufacturing resources (i.e., investment fund and equipment) which are critical in optimizing manufacturing profit. In this paper, the manufacturing resources and strategy of profiles are represented using a quantification method. These quantified manufacturing data are then used to generate the game model including multidimensional factors, i.e., player, action, information, strategy, payoff, outcome and equilibrium. For instance, the products are selected as players of resource competition and player's interests include optimal resource allocation and maximum profits. By solving the game model, the Nash equilibrium solution is regarded as optimal resource allocation strategy. The desired equilibrium will satisfy each product by achieving optimal production status and simultaneously improve the overall resource utilization efficiency. However, the mix-strategy equilibrium solution is not considered in the paper.

Taleb et al. [182] have proposed a QoS negotiation scheme to provide efficient utilization of network resources using the auction theory, which is a subfield of game theory. The architecture consists of a number of access points which provide different domains and each domain is administrated by a Global Service Negotiation Manager (GSNM). The GSNM server maintains different service levels and each level is set with a minimum price and a maximum price. The price of the service level is determined by a function of the channel quality conditions and offered QoS metric. A user is subscribed to one of the service levels based its own budget. Higher service level indicates more bandwidth. According to the proposed auction-based resource allocation, the user with the highest budget is allocated its requested service level, and each of the user downgrades their requested service level if the available bandwidth is not enough. The authors demonstrate that when the three constraints (network utilization, fairness and revenue) are taken into consideration, the proposed resource allocation scheme provides a unique Nash equilibrium. It is shown that the proposed scheme provides a potential high degree of fairness, efficient utilization of network resources and improvement of the service provider's revenue.

Berlemann et al. [183] have proposed a game theory-based resource allocation scheme for multiple wireless networks that are sharing unlicensed frequency bands. Both single-stage and multi-stage games are used to study the QoS in terms of the usage of frequency spectrum. The radio systems are represented as players that compete for a shared resource. The QoS requirements of applications are modelled to build a multidimensional utility function which consists of throughput, channel access period length, delay and jitter. The strategy of players refers to the expectation of resource allocation. At each stage, players interact repeatedly by selecting their own behaviour, i.e., a selection of MAC parameters. After each stage, players estimate their opponent's behaviour which enables interaction based on punishment and cooperation, i.e., a handpicked allocation of the radio resource aiming at a specific intention. Nash equilibrium is considered as the solution of the game in a cooperative domain. Simulation results show that cooperation is an achievable equilibrium that often improves the overall spectrum efficiency.

Tan et al. [184] have improved the IEEE 802.11 MAC protocol by establishing independence between the allocation of channel resources and the transmission strategies of mobile nodes. The authors demonstrate that current 802.11 DCF can lead to undesirable Nash equilibriums and inefficient network utilization. The proposed solution is based on a non-cooperative game. The players are the mobile nodes and each player maintains two

parameters: data rate and average payload size. Considering two non-cooperative nodes, each sending UDP to a receiver. The utility of each player is the achieved UDP throughput. At each stage, each player sets its data rate and frame size. The purpose of each competing player is to employ the strategy that maximizes its achieved throughput given the other player's best transmission strategy. It is shown that the proposed scheme provides rational nodes with equilibriums and improves the throughput in comparison with the traditional DCF.

Other game theory technique-based resource allocation algorithms have also been proposed. **Niyato et al.** [185] have presented an adaptive bandwidth allocation and admission control scheme for polling service in IEEE 802.16 networks based on the non-cooperative game. **Zhang et al.** [186] have presented a dynamic subcarrier allocation algorithm using a suboptimal solution of Nash bargaining solution. **Seneviratne et al.** [187] have developed a non-cooperative spectrum sharing game theoretic approach for cognitive wireless sensor networks in order to determine the optimum spectrum demand.

In conclusion, current mathematical theories, such as stereotypes, fuzzy logic, clustering and game theory, can be used in efficient resource allocation algorithms. In the case of delivering multimedia content to heterogeneous devices over wireless networks, there is a need to deal with a wide range of uncertain events, e.g. variable network capacity, wireless interference, device characteristics, service types, etc. Consequently, stereotypes-based resource management is adopted in this thesis. Five stereotype classes are introduced to model five levels of stream priority, which eventually determine bandwidth share. More details on the usage of stereotypes are presented in the next chapters.

3.4 Summary

This chapter introduces related works regarding the proposed solutions in terms of bandwidth estimation, QoS-oriented scheme and mathematical theory-based resource management. Many bandwidth estimation techniques have been proposed to provide estimations in wired networks. However, bandwidth estimation in wireless networks is a more challenging issue due to flexible wireless conditions such as: increased and variable Packet Error Rate (PER), wireless link rate adaptation, signal fading, contention, transmission retries, etc. Most of the existing wireless bandwidth estimation solutions use probing-based techniques. Probing techniques introduce extra traffic which has a negative

influence on the multimedia applications. Additionally, cross layer based techniques have been proposed to estimate the wireless channel bandwidth. Unfortunately, the cross layer solutions require modifications of standard protocols which make it complex and not desirable. Current solutions for QoS provisioning or QoS differentiation can be categorized into two classes: packet adjustment-based techniques and admission control-based solutions. The proposed solution of this thesis focuses on the QoS differentiation issue. The basic idea of resource management schemes is to deal with the uncertainty events in order to allocate resources distributed across a heterogeneous environment. Several mathematical-based solutions are considered to optimize the resource allocation including: stereotypes, fuzzy logic, game theory, overlay network, etc. Stereotypes-based resource allocation is used as the proposed solution which will be discussed in next chapters.

CHAPTER 4

Architecture and Solutions Overview

This chapter presents a bird's eye view of the proposed solutions: 1) Model-based Bandwidth Estimation (MBE); 2) intelligent Prioritized Adaptive Scheme (iPAS); 3) QoS-based downlink and uplink fairness for VoIP in IEEE 802.11 Networks. The three solutions are described in terms of their positions in the network protocol stack and their interaction with other protocols. The architecture of each solution is briefly presented, and more details of the algorithms as well as the experimental tests are discussed in next chapters.

4.1 Introduction

IEEE 802.11 access points (AP) have been widely deployed to provide the last-mile Internet access due to the convenience and low implementation costs. Nowadays, many wireless devices (i.e., laptop, smartphone, tablet, etc) are equipped with IEEE 802.11 interfaces. These devices are used to receive different types of multimedia services such as web-browsing, E-mail, voice over IP, streaming video, etc. The research works in this thesis study the scenario where multiple wireless devices with potential different characteristics are connected to the same IEEE 802.11 AP. In this scenario, the limited wireless bandwidth resources are shared by these devices. Since the original IEEE 802.11 network does not provide any Quality of Service (QoS) support, the end user's experience and the quality of multimedia delivery, in particular, might get impacted negatively due to network congestion, low efficient bandwidth allocation scheme, unfair channel access, etc. Therefore, it is necessary to efficiently manage the bandwidth resources of the IEEE 802.11 network in order to provide high level of QoS. Additionally, it is important to provide a mechanism to distribute bandwidth resources according to the needs of the devices and characteristics of the content played. More discussions about the research motivation and problem statement have already been presented in chapter 1.

Three contributions have been made in the thesis: 1) **Model-based Bandwidth Estimation (MBE)-**estimates the available bandwidth resources of the IEEE 802.11 network;

2) **Intelligent Prioritized Adaptive Scheme (iPAS)**-allocates the estimated bandwidth for the contending streams based on stream's priority. A stereotype-based resource allocation scheme is used to relate the priority and bandwidth share; 3) **QoS-based downlink/uplink fairness scheme for VoIP** -provides fair channel access between downlink and uplink VoIP traffic. The overview architecture including these three solutions is presented next.

4.2 Architecture of the Proposed Solutions

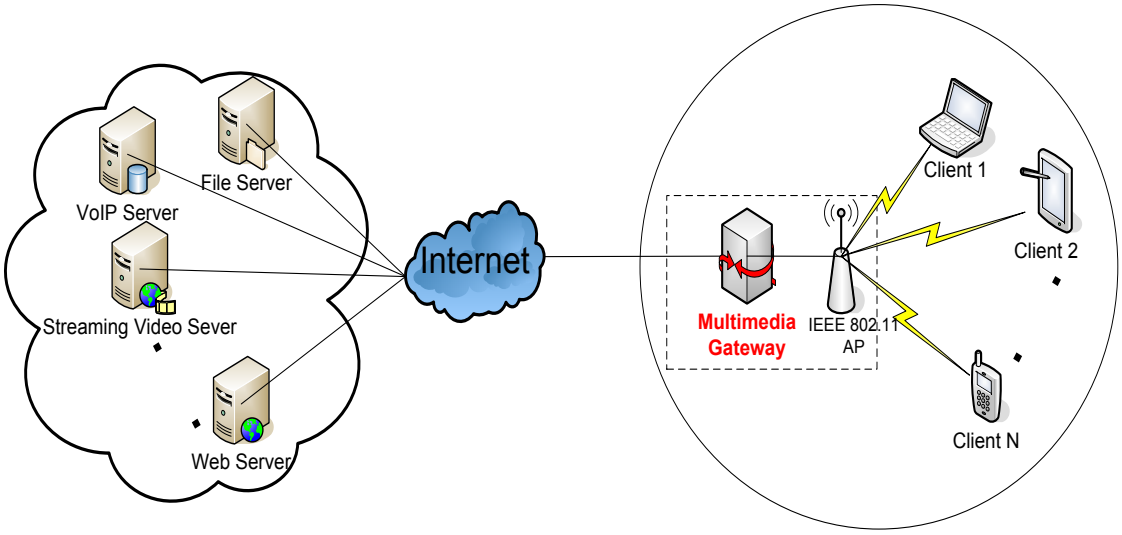


Figure 4-1 Overview architecture of the multimedia streaming system

Figure 4-1 illustrates the overview architecture of a multimedia streaming system. Different multimedia content originated from a group of servers presented including the File server, Voice over IP (VoIP) server, Streaming video server, Web server, etc. Heterogeneous wireless devices which receive these multimedia services are also illustrated including a laptop, tablet, smartphone, etc. These multimedia services are delivered to the end wireless clients via the IEEE 802.11 AP, which provides the last-mile access of the end-to-end link.

A *multimedia gateway* is implemented between the multimedia servers and the IEEE 802.11 AP. It deploys the proposed solutions in this thesis and aims to provide the following functions: 1) bandwidth estimation of the IEEE 802.11 network; 2) bandwidth allocation based on stream's priority which considers both transmission QoS and client device characteristics; 3) fair channel access between downlink and uplink VoIP traffic. Block-level structure, data communication, and the feedback mechanism employed by the proposed schemes are described in the next section.

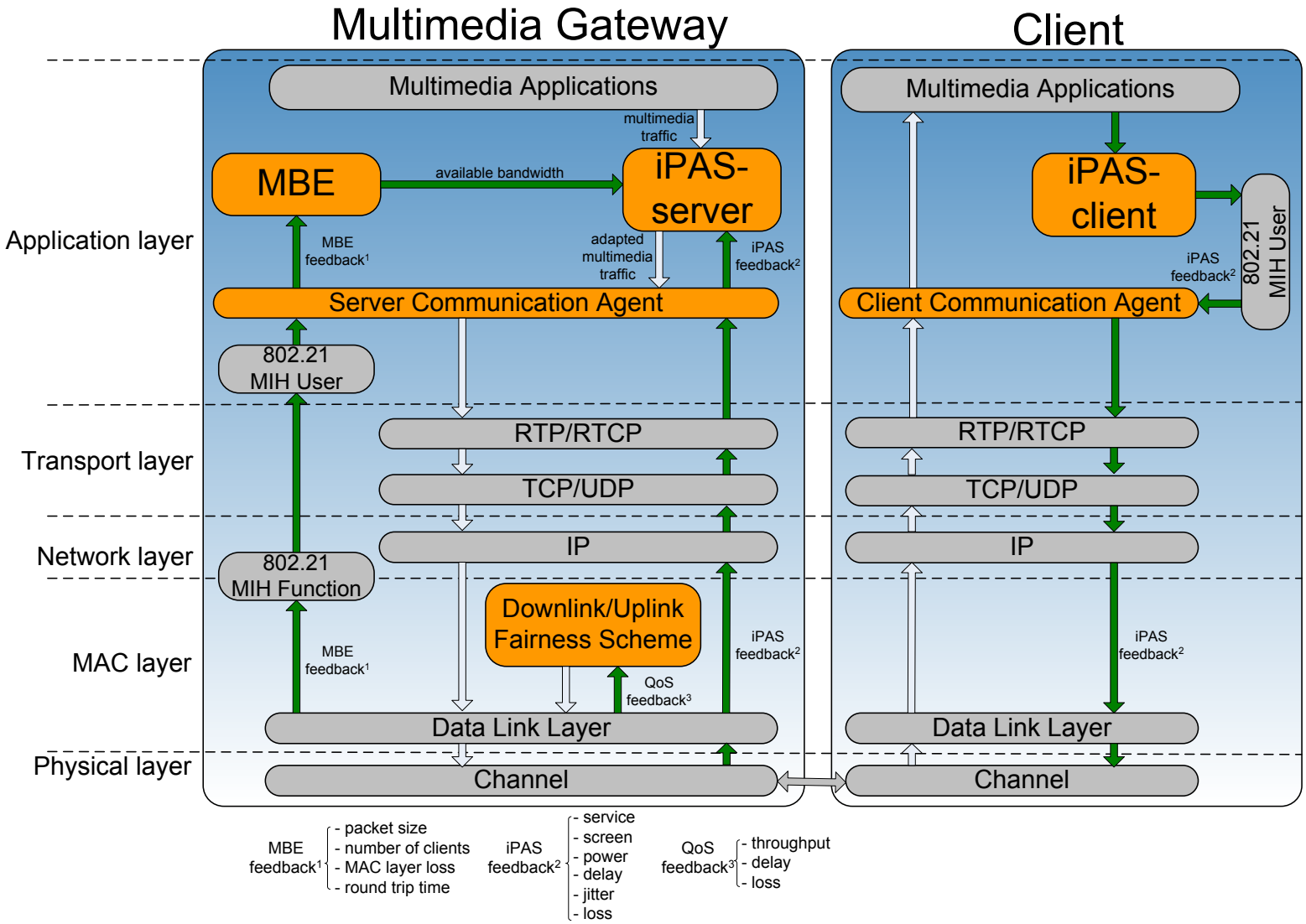


Figure 4-2 The Block structure of the multimedia gateway system

Figure 4-2 illustrates the TCP/IP protocol stack model-based architecture of the gateway system which consists of two main blocks: ***multimedia gateway*** and ***client***. The multimedia gateway block is designed to implement the three proposed schemes including: 1) *MBE*-estimates the available bandwidth resources of IEEE 802.11 network; 2) *iPAS server* - allocates bandwidth resources based on stream's priority over IEEE 802.11 network; 3) *Downlink/Uplink Fairness Scheme*-provides fair channel access between downlink and uplink traffic for IEEE 802.11 network. Client block uses an *iPAS client* module to collect information of stream preferences and delivered QoS parameters, which is sent as feedback to the multimedia gateway. The *Server Communication Agent* and *Client Communication Agent* modules situated at both sides of the two blocks, respectively, establish and maintain the communication link. *IEEE 802.21 MIH Function and MIH User* modules are utilized to gather feedback information from lower layer and upper layer of the multimedia gateway system.

4.2.1 Solutions Overview

The principles of the proposed schemes are described separately as follows:

1) The multimedia gateway deploys the novel **Model-based Bandwidth Estimation (MBE)** algorithm which was proposed to estimate the available bandwidth information of the IEEE 802.11 network.

   Deployment-MBE module is deployed at the application layer at the multimedia gateway and receives feedback information from the MAC layer of the gateway-side, in a cross layer approach.

   Algorithm-MBE involves two novel throughput models for TCP and UDP transmissions in IEEE 802.11 networks. The TCP wireless traffic model extends an existing TCP throughput model by considering the wireless characteristics such as transmission error, contention, and retry attempts. The UDP wireless traffic model is based on UDP packet transmission probability and dependent on IEEE 802.11 channel delay. MBE estimates the available bandwidth of the underlying IEEE 802.11 network based on the feedback delivered from the MAC layer. The feedback information includes packet size, the number of contending wireless stations, MAC layer-related packet loss, and round trip time. The feedback information is collected using *MIH Information Services (MISS)* provided by *IEEE 802.21 MIH Function* module [26], which is located between network layer and MAC layer.
2) The multimedia gateway also deploys the novel **intelligent Prioritized Adaptive**

Scheme (iPAS) which proposed to provide QoS differentiation for heterogeneous multimedia delivery over IEEE 802.11 networks.

Deployment-iPAS is deployed at the application layer at both multimedia gateway-side and client-side. The *iPAS client* module regularly sends feedback information to the *iPAS server*.

Algorithm-*iPAS server* module is responsible for managing bandwidth resources using a stereotype-based resource allocation mechanism and a bandwidth estimation scheme (MBE). *iPAS server* assigns dynamic priorities to various streams and determines their bandwidth share by employing a probabilistic approach-which makes use of stereotypes. The priority level of individual streams is variable and dependent on the feedback information from *iPAS client* which including: 1) stream-related characteristics, i.e. device resolution, device battery power left, and service type; 2) network delivery QoS parameters, i.e. delay, jitter, and packet loss rate. The feedback information is collected using *MIH Information Services (MISS)* provided by *IEEE 802.21 MIH User* module, which locates at application layer. Multimedia traffic is delivered using *Real-time Transport Protocol (RTP)* and feedback traffic is transmitted using *Real-Time Transport Control Protocol (RTCP)*.

3) A novel **QoS-based downlink/uplink fairness scheme** is proposed to achieve the downlink/uplink fairness of VoIP for Mobile Consumer Devices in WLANs.

Deployment-The proposed algorithm is deployed at the data link layer in the multimedia gateway. There is no feedback information needed from the client.

Algorithm-The proposed algorithm controls the contention window (CW) size at the AP to balance the downlink/uplink channel access opportunity. The optimum AP's CW size is computed based on the results of a stereotypes-based algorithm which utilizes the QoS-related information. This information is collected by monitoring the IEEE 802.11 link status and sent to the *Downlink/Uplink Fairness Scheme* module as QoS feedback. The QoS feedback includes: 1) throughput ratio between downlink and uplink-the ratio of the throughput at AP to the aggregation throughput of wireless stations. The throughput can be measured at the MAC layer of the AP; 2) delay ratio between downlink and uplink-the delay ratio between the downlink and uplink VoIP traffic is determined by computing the queue size and queue service rate at both AP and each wireless station. In particular, the queue size and queue service rate of each wireless station is estimated at AP and do not need any feedback from client. 3) packet loss rate ratio between downlink and uplink traffic-the packet loss

occurred at each wireless station is estimated at AP by analyzing the incoming client's packet sequence number.

4.2.2 Data Transmission and Feedback Mechanism

The block level architecture is presented in section 4.2.1 includes *Server Communication Agent (SCA)* and *Client Communication Agent (CCA)* modules which are responsible of managing the communication between the multimedia gateway and client. A double-channel is used for exchanging both *multimedia* data and *feedback* information, as illustrated in Figure 4-3. An unidirectional link from the multimedia gateway to the client delivers multimedia data to the latter. Another unidirectional link from the client to the multimedia gateway transmits feedback information to the former.

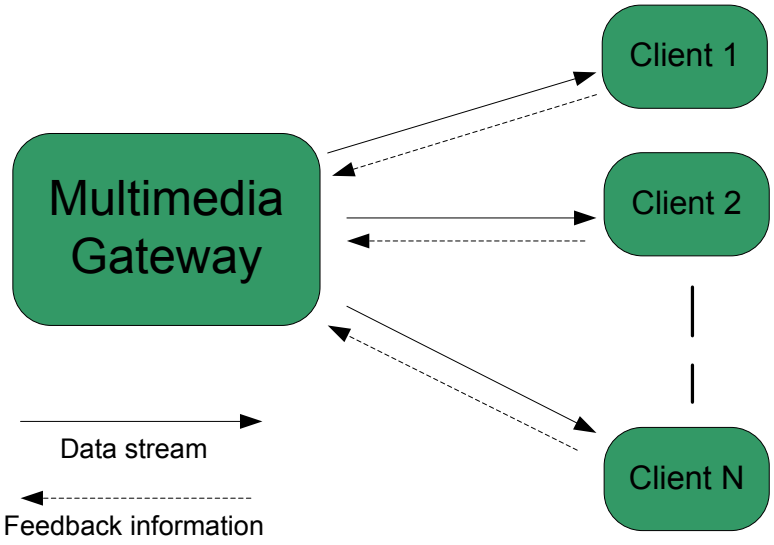


Figure 4-3 Multimedia data transmission and feedback information exchange between multimedia gateway and client applications

4.2.2.1 Data Transmission

The Real-time Transport Protocol (RTP) [73] is selected for multimedia data transport. It is designed as a higher-level transport protocol, which provides end-to-end transport functions suitable for real-time applications, such as audio, video, simulation data, over multicast or unicast network services. The data transport is augmented by Real-time Control Protocol (RTCP), which is a control protocol that allows for monitoring of the data delivery and provides minimal control and identification functionality. RTP and RTCP are designed to be independent of the underlying transport and network layers.

4.2.2.2 Feedback Collection and Transmission Mechanisms

An efficient feedback mechanism is critical in the multimedia gateway system. As shown in Figure 4-2, **the feedback information is collected using the IEEE 802.21 framework and delivered using RTCP**. The feedback traffic is needed by *MBE* and *iPAS* modules, in order to perform the bandwidth estimation and bandwidth allocation. The most significant reason to use IEEE 802.21 is to avoid modifying existing protocols. The principle of using IEEE 802.21 framework-based feedback collection and RTCP protocol-based feedback transmission is described as follows.

IEEE 802.21 framework introduced in chapter 2 provides a mechanism that allows interaction between lower layers (data link layer and physical layer) and upper layers (network layer, transport layer, and application layer). IEEE 802.21 defines *Media Independent Handover User (MIH User)* and *Media Independent Handover Function (MIHF)* entities, which provide three types of communication services including *MIH Event Services (MIES)*, *MIH Command Services (MICS)*, and *MIH information services (MIIS)*. *MIH Function* serves as middle layer components between link layer and network layer and *MIH User* locates at upper layers. *MIH Function* is a logical entity that provides services to the higher layers and obtains information from the lower layers through media specific interfaces. *MIH User* is abstraction of the functional entities that employ communication services (i.e. *MIIS*). In the proposed multimedia gateway system, *MIIS* is adopted as the communication service to collect feedback information needed by *MBE* and *iPAS*.

MIIS provides a framework for *MIH* entities (i.e. *MIH Function* and *MIH User*) to discover information of network conditions (i.e. network type, service provider identifier, QoS information, data rate, channel characteristics) and local properties (i.e. terminal information). *MIIS* specifies a standard format for this information, such as Extensible Markup Language (XML)[44] or Type Length Value (TLV)[45]. They are transmitted through *MIIS* using query/response or broadcast/multicast mechanisms. According to Figure 4-2, the *IEEE 802.21 MIH Function* in multimedia gateway collects the feedback needed by *MBE* (i.e. packet size, number of clients, MAC layer-related loss, and round trip time) using *MIIS* function which monitors the data link layer. The feedback information is then delivered to the Server Communication Agent (SCA) module in the upper layer. Additionally, the *IEEE 802.11 MIH User* module (implemented in *iPAS client* module as a functional entity) in

[44] XML-http://www.w3c.org/XML
[45] Type Length Value-http://en.wikipedia.org/wiki/Type-length-value#References

client side collects stream preference (i.e. service, screen size, power left, delay, jitter, loss) using *MIIS* function which interacts with the terminals and streams.

The *IEEE 802.21 framework* has been implemented as C++ objects in NS-2 under Linux environment based on the IEEE 802.21 specifications. Figure 4-4 presents the UML class relationship between *iPAS* components. IEEE 802.21 module is initialized from the *MIHAgent* class which extended from the NS2 Agent class. *MIHAgent* defines functions to get stream preference information (i.e. power left, device resolution, and service type) and QoS feedback (i.e. delay, loss, and jitter). The *iPASUser* class maintained a list of parameters representing the stream's information and delivered QoS. Each *iPASUser* object registers to the *MIHAgent* class. All of the initialized *iPASUSer* objects were managed by the class *iPASCtrl*, which deployed the stereotype-based bandwidth allocation algorithm.

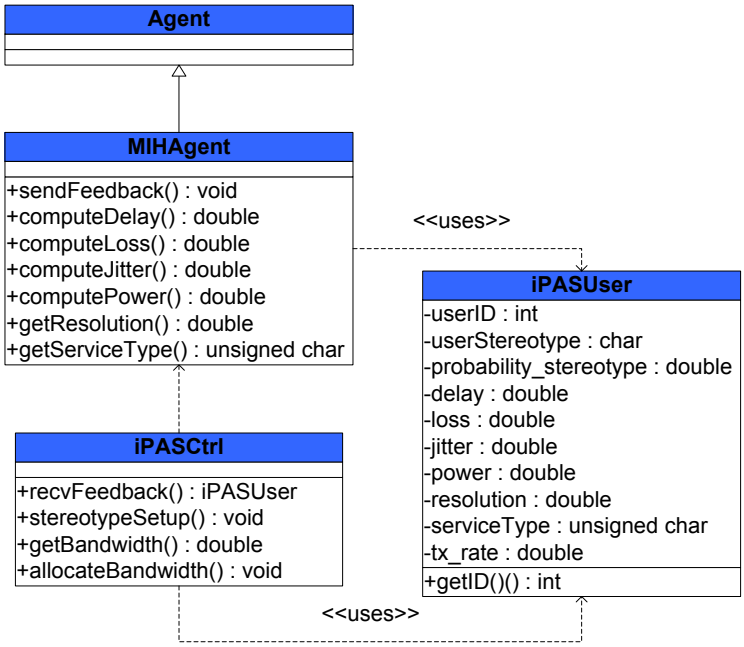


Figure 4-4 UML of classes for the IEEE 802.21 implementation in NS-2

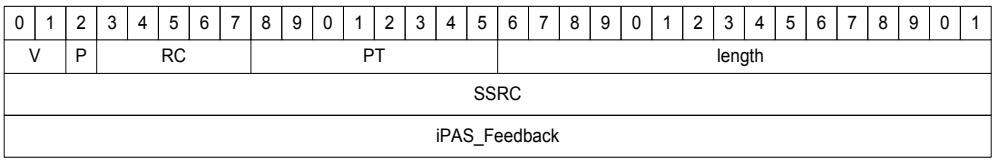


Figure 4-5 RTCP extension-iPAS_Feedback packet type

The RTCP standard allows for defining additional user defined packet types in the Receiver Report packet header. Therefore, a new RTCP packet type, *iPAS_Feedback*, is defined allowing RTCP be compatible with all existing systems. Figure 4-5 presents the structure of the new packet:

- *version* (V): 2 bits - which identifies the version of RTCP.
- *padding* (P): 1 bit - which indicates, if the padding bit is set, the packet contains one or more additional padding octets at the end which are not part of the payload. The last octet of the padding contains a count of how many padding octets should be ignored, including itself. Padding may be needed by some encryption algorithms with fixed block size or for carrying several RTP packets in a lower-layer protocol data unit.
- *reception report count* (RC): 5 bits – which contains the number of reception report blocks in this packet.
- *packet type* (PT): 8 bits – which identifies the packet type.
- *length*: 16 bits-which shows the length of this RTCP packet.
- *SSRC*: 32 bits – which is the synchronization source identifier for the originator of this SR packet.
- ***iPAS_Feedback***: 32 bits – which is the extension field that stores the feedback information. The feedback message contains six parameters including: service, screen size, power left, delay, jitter, and loss. As shown in Figure 4-2, the feedback is collected at *iPAS client* module using *IEEE 802.21 MIH User* function and sent to the gateway using RTCP protocol.

4.3 Summary

This chapter presents the overview architecture of the proposed solutions including MBE, iPAS, and the downlink/uplink fairness scheme. MBE and iPAS are proposed to provide QoS differentiation services for downlink traffic. The downlink/uplink fairness scheme adapts the contention window of the access point in order to achieve the fair channel access between downlink and uplink VoIP traffic. In a real life context, the proposed solution can be implemented as a middleware in a gateway, i.e. set-top box. This does not require major changes of existing systems and the middleware can be updated with low costs. In lab testing, an open source driver for the wireless interface card is needed to support MAC layer modifications. A widely used driver is *Madwifi*[46], which runs on a Linux platform and allows for 802.11 MAC protocol modifications. The place and principle of each contribution are described in terms of the TCP/IP protocol stack. Details of each contribution and evaluation are described in future chapters.

[46] Madwifi, http://www.madwifi-project.org

CHAPTER 5

Novel Wireless Bandwidth Estimation Schemes

This chapter presents the proposed bandwidth estimation schemes in IEEE 802.11 networks: intelligent Bandwidth Estimation (iBE) and Model-based Bandwidth Estimation (MBE). iBE uses the information provided by multimedia packets which are exchanged anyway to compute the available bandwidth without introducing extra probing traffic in 802.11 WLAN. MBE estimates the available bandwidth based on novel models for TCP and UDP traffic over IEEE 802.11 WLANs. They both utilize the cross layer-based information over the IEEE 802.11 networks and do not rely on any probing traffic. Details of performance evaluation will be introduced in chapter 7.

5.1 Introduction

Recently, the delivery of multimedia applications in IEEE 802.11 networks has grown significantly. Providing high quality of the received multimedia applications for multiple wireless clients becomes a challenging issue, due to the limited bandwidth resources offered by IEEE 802.11 network. Consequently, there is a need for an efficient technique that can estimate the available bandwidth of the wireless network in order to efficiently manage the bandwidth resources. Wireless network bandwidth estimation is a critical issue for Quality of Service (QoS) provisioning in IEEE 802.11 WLANs. A literature review of the existing bandwidth estimation techniques have been presented in chapter 2. Most bandwidth estimation solutions rely on probing traffic which introduces significant cost. Two novel bandwidth estimation schemes are prosed in this thesis: intelligent Bandwidth Estimation (iBE) [188] and Model-based Bandwidth Estimation (MBE) [189]. iBE makes use of the information provided by multimedia packets to compute the MAC layer-based available bandwidth. MBE estimates the available bandwidth uses the TCP and UDP throughput model, which takes the network condition-related parameters as input, i.e., packet loss, round

trip time, the number of contending wireless stations. Detailed algorithm of iBE and MBE are introduced next.

5.2 Intelligent Bandwidth Estimation (iBE)

5.2.1 Introduction

The principle mechanism of the proposed intelligent bandwidth estimation algorithm (iBE) is to make use of the differences between the packet's transmission time and reception time at MAC layer. IEEE 802.11 DCF protocol is used due to the widely deployment.

5.2.2 Assumptions

There are two assumptions for the iBE system. Firstly, iBE assumes that the last hop wireless network is the bottleneck link. Fig 5-1 presents the network paths with last hop wireless LAN. A pair of traffic server and client is introduced to generate cross traffic. Secondly, First Come First Serve (FCFS) queue is assumed as the default scheduler at the IEEE 802.11 access point (AP). This assumption applies for the original IEEE 802.11a/b/g protocols.

5.2.3 Algorithm

iBE utilizes IEEE 802.11 DCF protocol-related parameters such as **Short Interframe Space (SIFS)**, **DCF Interframe Space (DIFS)**, and **Request to send/Clear to send (RTS/CTS)** [21]. Figure 5-1 presents the packet processing sequence of IEEE 801.1 DCF. The sender firstly senses the status of wireless channel before transmission. When the channel is idle for a time equals DIFS, the station can then access the wireless medium. When a packet is transmitted correctly, the receiver immediately sends an acknowledgement packet after waiting for SIFS. DIFS and SIFS reduce the probability of conflict among different packets. RTS/CTS are optional mechanism of IEEE 802.11 protocols in order to overcome the hidden station problems. Station sends RTS frame before transmitting data. The destination then replies with a CTS frame. Any other station receiving RTS or CTS frame should wait for a given time in order to avoid collision. RTS/CTS packet size threshold is 0-2347 bytes. RTS/CTS frames are not initialized until the packet size exceeds this threshold.

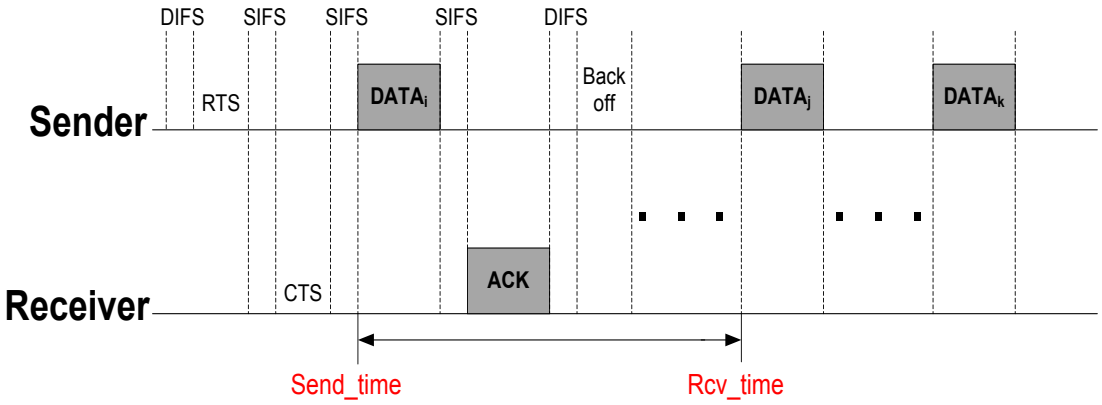


Figure 5-1 Packet sequence in 802.11 MAC Layer

iBE introduces two time stamps at MAC layer, *sent time* and *received time*, to record the packet sent and received time at sender and receiver, respectively. The time stamps are represented by 8-bit fields in the IP packet header. The values of the 802.11 MAC related parameters such as SIFS, DIFS, ACK, and RTS/CTS are configured according to the 802.11 specification. The instant available bandwidth can be computed using equation (5-1) to equation (5-3). A burst of multimedia packets is selected as a sample (S_i), where *i* implies the index of the picked sample. The sample size could be computed by equation (5-1), where $packet\_recvd_i$ is the number of data packets received within a sample at client MAC layer. PS_i is the size of data packet with MAC header.

$$S_i = packet\_recvd_i \times PS_i \tag{5-1}$$

$$\begin{aligned} Ti = {} & recv\_time_i - S\_time_i - (packet\_recvd_i - 1) \\ & \times (3 \times SIFS + 2 \times DIFS + Backoff_i + T_{ACK} + T_{RTS} + T_{CTS}) \end{aligned} \tag{5-2}$$

The time taken to transmit data packets (T_i) is calculated as defined in equation (5-2). The $recv\_time_i$ and S_time_i represent the received time of the last packet and the transmission time of the first packet in $sample_i$ respectively. Since the waiting time in the MAC layer does not reflect the actual bandwidth, the time cost due to MAC contention (i.e., DIFS and SIFS) is subtracted from sample transmission time. Additionally, T_{ACK}, T_{RTS} and T_{CTS} represent the time cost for processing and transmitting *ACK, RTS* and *CTS* packets. $Backoff_i$ indicates the duration (waiting time) between two consecutive packets. *Backoff* time depends on current contention window size. The subtraction of *ACK*, *RTS/CTS*, and *Backoff* contributes to the reality of bandwidth but might lead to the overestimation of bandwidth. Finally, the instant bandwidth (*instBW*) for a sample is calculated using equation (5-3),

$$instBW = S_i / T_i \quad (5\text{-}3)$$

5.2.4Limitations of iBE

Although iBE does not rely on the probing traffic for the bandwidth estimation, there are few limitations exist.

Firstly, iBE only considers the MAC layer bandwidth. However, the transport layer delay and the queuing delay also constitute the network bandwidth. Transport layer protocols such as TCP and UDP contribute different amount of delays due to the transport control mechanism. Only the delay for transmitting raw multimedia data is considered. Additionally, queuing delay cannot be ignored since the buffer plays a critical role in the MAC layer.

Secondly, the estimated bandwidth provided by iBE does not include some effects which occur in real circumstances such as shadowing, reflections, fading multipath, interference from other networks, etc. These factors will reduce the theoretical bandwidth resources. Therefore, in real life context, iBE might overestimate the wireless bandwidth.

iBE is not applicable when collisions and propagation errors in wireless networks are significant, as the two factors are not taken into account.

5.3 Model-based Bandwidth Estimation (MBE)

5.3.1 Introduction

This section introduces the analytical ***Model-based Bandwidth Estimation*** algorithm (*MBE*) to estimate the available bandwidth for data transmissions over IEEE 802.11 WLANs. Figure 5-2 shows the typical local and distributed network which consist of an application server, wired LAN, IEEE 802.11 access point (AP), and wireless devices. Three major contributions have been made. First, *MBE* relies on a novel TCP model for wireless data transmissions, which extends an existing TCP throughput model by considering the IEEE 802.11 WLAN characteristics (transmission error, contention, and retry attempts). Second, *MBE* makes use of a new UDP over wireless throughput model based on the UDP packet transmission probability and IEEE 802.11 channel delay model. Third, two models are proposed to estimate the bandwidth when TCP and UDP traffic co-exists in IEEE 802.11 networks. Note, unlike most existing estimation techniques, *MBE* neither requires modification of current transmission protocols, nor uses probing traffic. The following sections introduce these novel contributions and describe the algorithms in details.

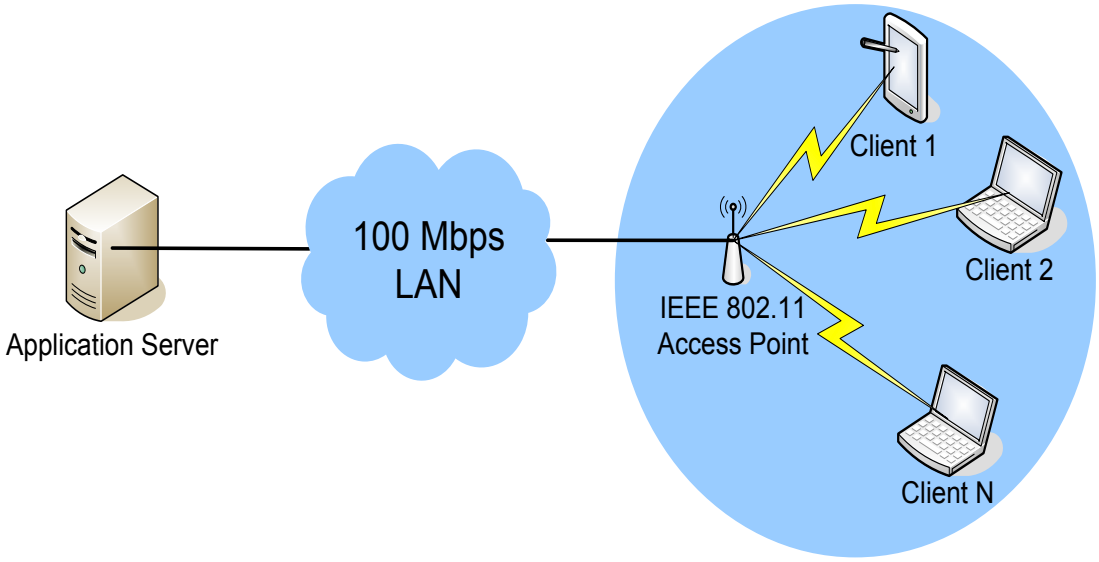


Figure 5-2 Network architecture of wireless bandwidth estimation

5.3.2 TCP Throughput Model and IEEE 802.11 Traffic Model

This section introduces the existing TCP throughput model and the IEEE 802.11 traffic model which are used by the new TCP throughput model for IEEE 802.11 networks.

Kurose et al. [190] have proposed an accurate TCP throughput model by capturing both TCP fast retransmission and time out mechanism. The model develops the steady-state sending rate of a bulk transfer TCP flow as a function of loss rate and round trip time (RTT). Kurose's TCP throughput model is described in equation (5-4), where B is the throughput received, MSS denotes the maximum segment size, RTT is the transport layer roundtrip time between sender and receiver, b is the number of packets that are acknowledged by a received ACK, P_{tcp} is the steady-state loss probability, and T_o represents the timeout value to trigger the retransmission.

$$B = \frac{MSS}{RTT \times \sqrt{\frac{2bP_{tcp}}{3}} + T_o \times \min\left(1,3\sqrt{\frac{3bP_{tcp}}{8}}\right) \times P_{tcp} \times \left(1 + 32P_{tcp}^{\ 2}\right)} \tag{5-4}$$

The IEEE 802.11 model was introduced by Chatzimisios, et.al [191]. The authors have extended Bianchi's IEEE 802.11 DCF Markov Chain model by taking into account packet retry limits, collisions and propagation errors (fading, interference). The key assumption of the model is that the transmission loss probability (P_{DCF}) of a transmitted

packet is constant and independent of the number of the collisions or errors occurred in the past. The probability P_{DCF} is given by equation (5-4), where N indicates the number of contending stations, BER is the bit error rate, L is the packet size, H is the packet header, and τ denotes the probability that a station transmits a packet in a randomly chosen slot time. The probability τ is given by equation (5-6), where W represents the initial contention window size and m is the retry limit. Chatzimisios has described a unique solution for equation (5-5) and equation (5-6) and has derived relation for the probability that at least one transmission occurs in a random time slot (P_{tr},). This could be written as shown in equation (5-7). When the retransmission reaches the retry limit m, the packet is dropped immediately. Consequently, we derived the drop probability P_{drop}, as presented in equation (5-8).

$$P_{DCF} = 1-(1-\tau)^{N-1} \times (1-BER)^{L+H} \tag{5-5}$$

$$\tau = \frac{2\times(1-2P_{DCF})\times(1-P_{DCF}{}^{m+1})}{W\times(1-(2P_{DCF})^{m+1})\times(1-P_{DCF})+(1-2P_{DCF})\times(1-P_{DCF}{}^{m+1})} \tag{5-6}$$

$$P_{tr} = 1-(1-\tau)^{N} \tag{5-7}$$

$$P_{drop} = P_{DCF}{}^{m+1} \tag{5-8}$$

The TCP throughput model does not offer accurate results in situations when TCP runs over IEEE 802.11 networks, since the wireless channel characteristics are not considered. For this reason, the thesis focus on extending and proposing a TCP throughout model by considering both TCP congestion control mechanism and 802.11 characteristics.

5.3.3 TCP Throughput Model for IEEE 802.11 Networks

The original TCP throughput model needs to be updated in order to consider wireless delivery conditions. These factors are highly different in wireless than in wired networks. In this context, we considered these factors and addressed them in their individual slips: 1) *packet loss probability update* (P_{tcp}); 2) *Round-trip Time* (*RTT*) *update*; 3) *Consideration of both TCP and 802.11 DCF models*.

Packet Loss Update

There are two types of packet loss when transmitting TCP traffic over wireless: *congestion loss* (P_{cong}) and *transmission loss* (P_{DCF}). TCP assumes that all packet loss is caused by congestion and therefore reduces the congestion window. The value of P_{cong} depends on the queuing protocol. *MBE* considers the popular *Random Early Discard (RED*) queuing protocol proposed in RFC2309 [192]. *RED* determines the action of packet forwarding based on current queue size (q_{k+1}), and updates the average queue size ($\overline{q}_{k+1}$) for each arriving packet. The *RED* specification defines the average queue size, as given in equation (5-9), where w_q is the weight factor.

$$\overline{q}_{k+1} = (1 - w_q) q_k + w_q \times q_{k+1} \tag{5-9}$$

Both TCP and 802.11 MAC trigger a packet retransmission event when packet loss is detected. The packet loss can be caused by either queue congestion (P_{cong}), wireless transmission error (P_{DCF}) or retry-based drop (P_{drop}). The packet drop probability due to queue congestion (P_{cong}) is given in equation (5-10), where $q_{\min}$ and $q_{\max}$ denote the minimum and maximum threshold of the queue. P_{cong} is collected in the sender's queue.

$$P_{cong} = \begin{cases} 0 & if \quad \overline{q}_{k+1} \le q_{\min} \\ 1 & if \quad \overline{q}_{k+1} \ge q_{\max} \\ \dfrac{\overline{q}_{k+1} - q_{\min}}{q_{\max} - q_{\min}} & otherwise \end{cases} \tag{5-10}$$

P(drop|DCF) is the packet drop probability of IEEE 802.11 MAC layer, as shown in equation (5-11). The parameter P_{drop} is dependent on $P_{DCF,}$ since in the IEEE 802.11 MAC layer, the packet is dropped if the retransmission reaches the maximum number of attempts limit. The parameters P_{cong} and P_{DCF} are independent from each other, as they are determined by the queue status and wireless channel, respectively. Consequently, the conditional probability is used for drop probability. The probability *P(DCF|Drop)* equals 1, as this dependency always exists.

$$P_{retr}^{TCP} = P_{cong} + P_{DCF} + P(drop \mid DCF)$$

$$P(drop \mid DCF) = \frac{P(DCF \mid drop) \times P_{drop}}{P_{DCF}} \tag{5-11}$$

Consequently, the probability of successful transmission, P_{succ}^{TCP}, is written as shown in equation (5-12).

$$P_{succ}^{TCP} = 1 - P_{retr}^{TCP} \tag{5-12}$$

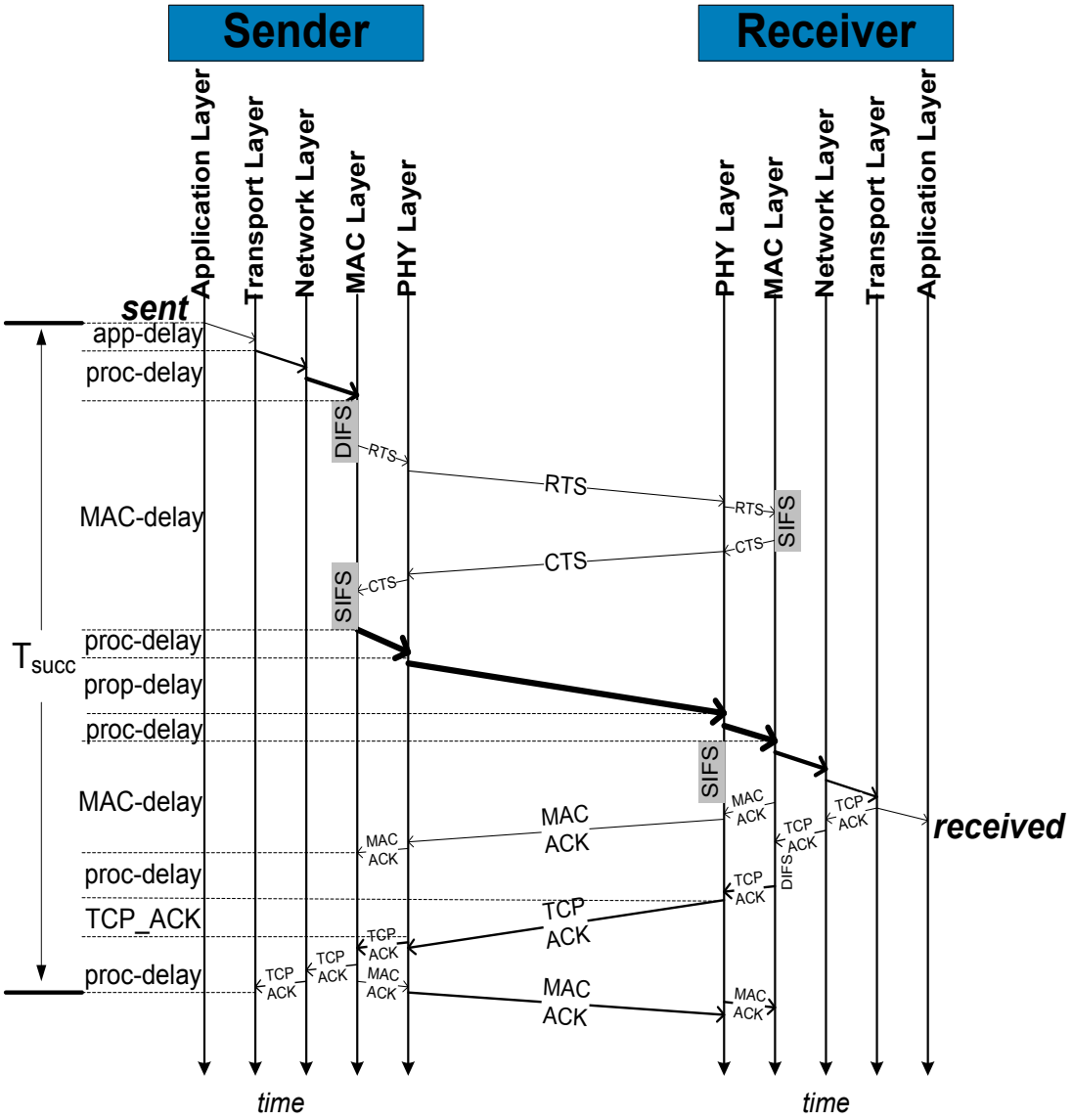


Figure 5-3 Successful transmission when TCP runs over 802.11 networks

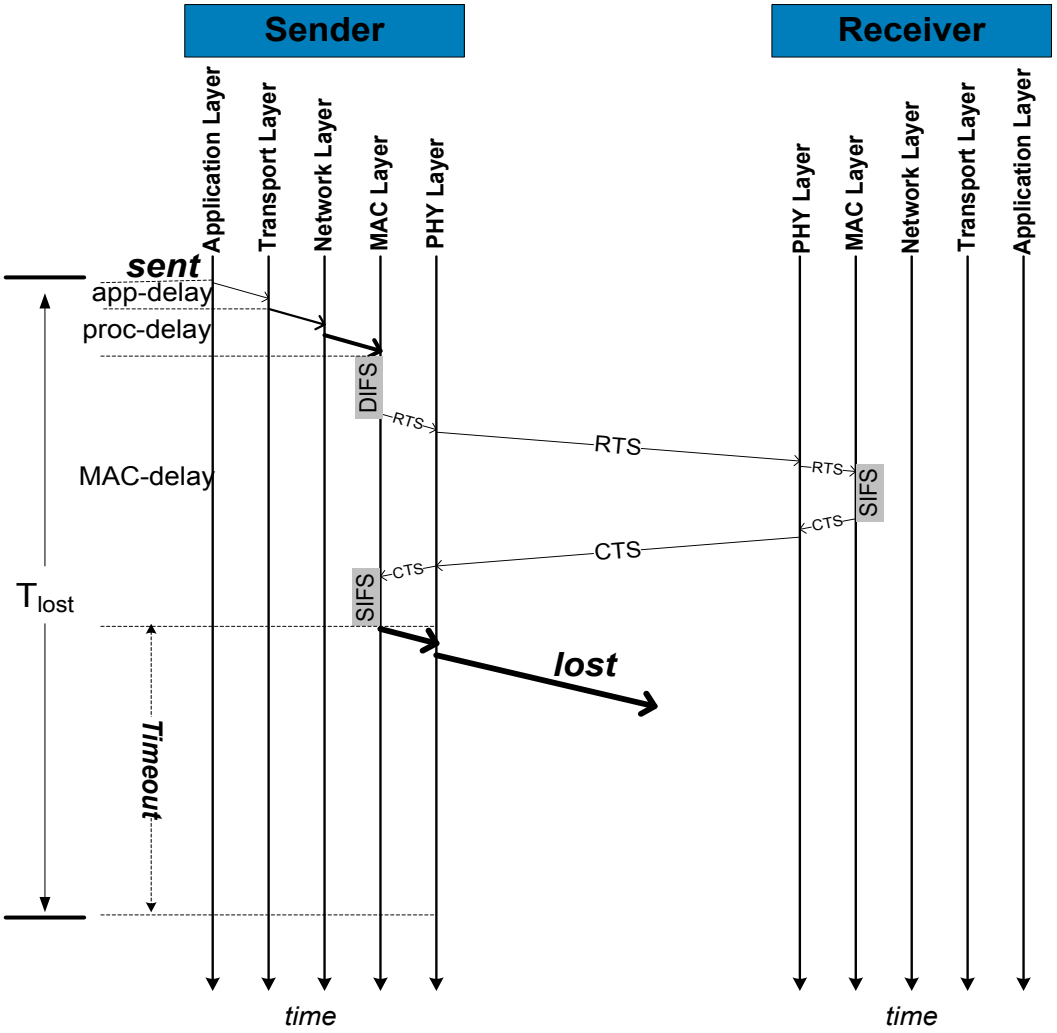


Figure 5-4 Packet loss when TCP runs over 802.11 networks

RTT Update

As shown in Figure 5-3 and Figure 5-4, the overall delay for transmitting the data can be decomposed into seven components based on the OSI model:

1) Application layer delay (*App_Delay*): the delay at application layer such as video encoding/decoding, etc.

2) Transport layer delay (*Transport_Delay*): the delay to implement transport protocol such as TCP congestion control

3) Network layer delay (*IP_Delay*): the delay at the IP layer for the routing protocol.

4) MAC layer delay (*MAC_Delay*): the delay caused by MAC contention.

5) Physical layer delay (*Phy_Delay*): the delay occurred at physical layer dependent on raw bits transmission type.

6) Propagation Delay (*Prop_Delay*): the data propagation delay on the channel medium.

7) Terminal processing delay (*Proc_Delay*): determined by terminal's processing ability such as CPU, memory, power mode, operating system, etc.

During the end-to-end round-trip time (*RTT*) in *MBE*, the receiver can be in one of the following states***: idle, successful transmission*** and ***retransmission***. The delay for successful transmission is denoted as T_{succ}. We derived equation (5-13) and equation (5-14) to present the 802.11 MAC layer delay for basic access mode (MAC_Delay_{basic}) and *RTS/CTS* mode (MAC_Delay_{RTS}), where *DIFS (Distributed Inter-Frame Space)* and *SIFS (Short Inter-Frame Space)* are contention control parameters defined in 802.11 MAC specifications. *MAC_ACK* represents the acknowledgment packet sent by the MAC receiver.

$$MAC\_Delay_{basic} = DIFS + SIFS + MAC\_ACK \tag{5-13}$$

$$MAC\_Delay_{RTS} = DIFS + 3 \times SIFS + RTS + CTS + MAC\_ACK \tag{5-14}$$

Combining equation (5-13) and equation (5-14), the delay for successful transmission is given by equation (5-15), where *TCP_ACK* represents the acknowledgment packet sent by the TCP receiver. Note that, the propagation delay represents the time taken to transmit data which includes the original data packet plus the stack protocol header: TCP/IP/MAC.

$$T_{succ}^{TCP} = APP\_Delay + \Pr oc\_Delay + \{MAC\_Delay_{basic}, MAC\_Delay_{RTS}\} + \Pr op\_Delay + TCP\_ACK \tag{5-15}$$

The TCP congestion control starts retransmission if any of the following two conditions occur: 1) Three duplicate ACKs are received at the sender as described in RFC 2581 [193]; 2) TCP sender does not receive an ACK after waiting a period equal with the timeout (T_O^{TCP}). RFC 2581 gives suggestions on how to calculate timeout, as shown in

equations (5-16), (5-17) and (5-18). In equation (5-16), the parameter β is a smoothing factor determining the weight given to the previous value of *RTT*, namely RTT'. The parameter *M* denotes the time taken for *ACK* to arrive. D_{RTT} is the estimation of the standard deviation of *RTT*. D'_{RTT} is the previous value of $D_{RTT.}$ Whenever an *ACK* is received, the difference between expected and measured values |RTT-M| is computed and D_{RTT} is updated as in equation (5-17). Subsequently, T_O^{TCP} is given by equation (5-18) based on dynamic timeout adjustment. A typical TCP implementation uses α=0.875 and β =0.75.

$$RTT = \beta \times RTT' + (1-\beta) \times M \tag{5-16}$$

$$D_{RTT} = \alpha \times D'_{RTT} + (1-\alpha) \times \left| RTT - M \right| \tag{5-17}$$

$$T_o^{TCP} = RTT + 4 \times D_{RTT} \tag{5-18}$$

Further, the delay (T_{lost}^{TCP}) caused by timeout is subsequently given by equation (5-19),

$$T_{lost}^{TCP} = \Pr oc\_Delay + MAC\_Delay + T_o^{TCP} \tag{5-19}$$

When three duplicate ACK packets are received at the sender, TCP protocol enters fast retransmission and the delay caused by the three ACK (T_{3ACK}) is T_{SUCC}^{TCP}. Consequently, the average retransmission delay T_{retr}^{TCP} is derived in equation (5-20). The retransmission delay could be T_{succ}^{TCP} or T_{lost}, depending on how the retransmission is triggered: three duplicate ACKs or the timeout.

$$T_{\text{retr}}^{\text{TCP}} = \left\{ T_{3ACK}, T_{lost}^{TCP} \right\} = \left\{ T_{succ}^{TCP}, T_{lost}^{TCP} \right\} \tag{5-20}$$

Combining TCP throughput and 802.11 DCF model

By combining equations (5-7), (5-11), (5-12), (5-15), and (5-20), the new Round-Trip Time (*MRTT)* is written as shown in equation (5-21).

$$MRTT = (1-P_{tr}) \times \sigma + P_{retr}^{TCP} \times T_{retr}^{TCP} + P_{succ}^{TCP} \times T_{succ}^{TCP} \tag{5-21}$$

The parameter σ is the MAC slot time. Note that P_{tr} defined in 802.11 MAC is adopted in the new model since it is independent of the protocols. It is necessary to use *MRTT*, as it considers the transmission and acknowledgement times contributed by both transport layer and MAC layer protocols. The *RTT* defined in Kurose's model (equation (5-4)) includes the time computed at transport layer only.

Based on equations (5-4), (5-11) and (5-21), the application layer throughput B^{TCP} for each TCP connection is described in equation (5-22), where *b* is the number of packets acknowledged by a received ACK, P_{retr} is the retransmission probability, *MRTT* is the updated round-trip time and *MSS* is maximum segmentation size.

$$B^{TCP} = \frac{MSS}{MRTT \times \sqrt{\frac{2bP_{retr}^{TCP}}{3}} + T_o \times \min(1,3\sqrt{\frac{3bP_{retr}^{TCP}}{8}}) \times P_{retr}^{TCP} \times (1 + 32P_{retr}^{TCP^2})} \quad (5\text{-}22)$$

If the network, device and the application service remains the same for a user, then the *MBE* would need to know values of only two types of parameters:

1) **Static parameters**: application delay, processing delay, 802.11 MAC configurations such as minimum contention window, *DIFS*, *SIFS*, slot time, retry limit and capacity.

2) **Dynamic parameters:** the number of contending stations, round trip time, packet loss and data packet size.

5.3.4 UDP Throughput Model for IEEE 802.11 Networks

As there is not any previously proposed model for UDP, we use the previously described TCP model and the particularity of UDP to propose the throughput estimation model for UDP over IEEE 802.11. Unlike TCP, the UDP protocol does not support packet retransmissions and therefore the UDP over WLAN throughput model should consider this. Hence, the terms P_{retr} and *MRTT* defined in equations (5-11) and (5-21) which consider TCP fast retransmission and timeout respectively, should be removed in UDP throughput model. By combining equations (5-5) and (5-8), the probability of retransmission when UDP traffic runs over 802.11 networks can be written as shown in equation (5-23).

$$P_{retr}^{UDP} = P_{DCF} + P_{drop} \tag{5-23}$$

Similar to the TCP transmission delay described in equation (5-15), the UDP transmission delay can be derived and is shown in equation (5-24) and equation (5-25), respectively.

$$T_{succ}^{UDP} = APP\_Delay + \Pr oc\_Delay + \{MAC\_Delay_{basic}, MAC\_Delay_{RTS}\} + \Pr op\_Delay \tag{5-24}$$

$$T_{o}^{UDP} = \Pr op\_ACK + \Pr op\_UDP + SIFS \tag{5-25}$$

Further, the retransmission delay triggered by 802.11 time out is given in equation (5-26):

$$T_{retr}^{UDP} = \Pr oc\_Delay + MAC\_Delay + T_{o}^{UDP} \tag{5-26}$$

Importantly, the average delay, *Delay_UDP*, for successfully transmitting the individual UDP packet could be written as in equation (5-27):

$$Delay\_UDP = (1 - P_{tr}) \times \sigma + P_{retr}^{UDP} \times T_{retr}^{UDP} + P_{succ}^{UDP} \times T_{succ}^{UDP} \tag{5-27}$$

The available bandwidth for UDP traffic over 802.11 WLANs is given in equation (5-28), where *Payload* is the total information in bytes at the transport layer, transmitted during one time period ($T_j - T_i$):

$$B^{UDP} = \frac{\int_{T_i}^{T_j} \frac{Payload}{Delay\_UDP} dt}{T_j - T_i} \tag{5-28}$$

There is no distinguishment for MBE between downlink and uplink traffic. In the case of non-interactive applications (e.g. video on demand, web browsing, file downloading, etc), downlink traffic dominates the wireless network. In the case of interactive applications (e.g. video conference, voice over IP, on-line game, etc), both downlink and uplink traffic require equal channel access. Notably, MBE can be applied in both types of applications. Each node is an independent MAC access entity, whose back-off period, collision probability, retransmission probability are computed using the same method as given by equation (5-22) and equation (5-28).

5.3.5 Co-Existing TCP and UDP Throughput Model for IEEE 802.11 Networks

This subsection introduces *MBE*, which considers the combined effect of TCP and UDP traffic over WLAN and makes use of the TCP and UDP over WLAN throughput models introduced before.

When TCP and UDP traffic are transmitted together, their throughputs are different than those when TCP and UDP are delivered alone. TCP adopts a congestion control mechanism to adjust the transmission rate to the available bandwidth. UDP is more aggressive and always takes as much bandwidth as possible, therefore affecting the TCP traffic. The major difference between the models for TCP and UDP is with regard to consideration of lost packet retransmissions. In order to address this effect of UDP on the TCP traffic, the weight *w* is introduced, as shown in Figure 5-5 and equation (5-29).

By combining the TCP and UDP over WLAN throughput models, the estimated aggregated throughput for co-existing TCP and UDP can be written as shown in equation (5-29). The parameter *w* is the bandwidth weight factor, *M* and *N* represents the total number of TCP and UDP flows, *i* and *j* are the index of TCP and UDP flows, respectively.

$$B^{\mathrm{TCP}+UDP} = w \times \sum_{i=1}^{M} B^{\mathrm{UDP}} + (1-w) \times \sum_{j=1}^{N} B^{TCP} \tag{5-29}$$

Next, we investigated the throughput performance when sending TCP and UDP flows together without any background traffic. When TCP and UDP traffic are transmitted together, their throughputs are different to those when TCP and UDP are delivered alone. This is mainly due to the fact that TCP adopts a fast congestion control mechanism to adjust the transmission rate based on packet loss. In order to address the influence of UDP over TCP, the weight *w* is introduced. A suggested value for *w* is given based on the research works by Bruno et al. [194], who have studied the throughput performance when delivering both TCP and UDP in IEEE 802.11 WLANs. Their study showed that, in saturated network conditions, *n* UDP flows obtain about *n* times the aggregate throughput achieved by the TCP flows, which is independent of the overall number of persistent TCP connections. Consequently, the value for *w* is given in equation (5-30).

$$w = \frac{M}{1+M} \tag{5-30}$$

Note that, the value of w suggested by equation (5-30) can well reflect the throughput relationship between TCP and UDP flows in saturated wireless network conditions. Future works will be conducted to suggest a more proper value of w in the case of the unsaturated network conditions.

5.4 Summary

This chapter introduces two proposed bandwidth estimation schemes, iBE and MBE. iBE provides a simple and effective bandwidth estimation scheme based on 802.11 MAC layer mechanism. MBE estimates the available bandwidth based on novel models for TCP and UDP traffic over IEEE 802.11 WLANs. The key novelty is that both iBE and MBE are server-side solutions that avoid the necessity to probe the network. MBE is designed to be implementation friendly and is an add-on component to the existing OSI model. However, iBE requires the modification of 802.11 MAC protocol. Next chapters will discuss the performance evaluation of iBE and MBE. Additionally, MBE is adopted as the preferred bandwidth estimation scheme to be included in our proposed prioritized adaptive scheme.

CHAPTER 6

Intelligent Prioritized Adaptive Scheme (iPAS)

This chapter introduces the intelligent Prioritized Adaptive Scheme (iPAS) which provides QoS differentiation between heterogeneous flows during multimedia delivery over wireless networks. iPAS uses the available bandwidth as estimated by MBE, which has been discussed in Chapter 4. iPAS relies on the IEEE 802.11 and IEEE 802.21 standards, which were described in Chapter 2.This chapter consists of four sections: 1) the background and motivation of the iPAS solution; 2) the system architecture for iPAS; 3) the key modules of iPAS-stereotype-based bandwidth allocation; 4) summary of the chapter.

6.1 Introduction

Recently, IEEE 802.11 WLANs have been used widely to deliver multimedia content: video, voice, graphics, financial, and medical data, etc [195]. At the same time, the number of mobile devices connected to the IEEE 802.11 networks has increased significantly.

Delivering multimedia content to heterogeneous devices over a variable networking environment while maintaining high quality levels involves many technical challenges [196] [197]. In order to support such high quality, different multimedia services have various delivery-related QoS requirements. For instance, real-time video traffic needs large bandwidth and is less tolerable to delay and jitter, in comparison with any best-effort service. Furthermore while delivering the same multimedia content, devices with high resolution and large battery power should benefit from a larger bandwidth share than that to be allocated to devices with small resolution and low battery power. Additionally, wireless multimedia delivery solutions should consider the unreliable and dynamic nature of the wireless channels, which impact on the quality of multimedia applications.

The original IEEE 802.11 protocol adopts the CSMA/CA mechanism to manage the wireless channel access [16]. However, the 802.11 standard is only designed for best effort services and incorporates limited QoS support with regard to multimedia applications and mobile devices. The IEEE 802.11e has been developed to overcome the QoS problem of traditional 802.11 networks, introducing support for differentiation between four different classes of traffic: voice, video, best effort and background [20]. However, when the 802.11e channel is occupied by high priority traffic, the low priority traffic might suffer from starvation due to the lower chance of channel access. Additionally, IEEE 802.11e cannot provide QoS differentiation between various devices, nor between services belonging to the same class. Recently, several solutions have been developed to optimize the original 802.11e focusing on starvation and overall throughput, such as [198], [199], and [200]. Other QoS-oriented solutions like TCP Friendly Rate Control (TFRC) [101], Quality Oriented Adaptive Scheme (QOAS) [106], [201], [202] [203], Partial Reliable-Stream Control Transmission Protocol (PR-SCTP) [110], etc are proposed at different layers of the protocol stack. Admission control-based schemes like [204], [205], [206] are designed to guarantee QoS levels of existing traffic. However, none of the above solutions provide support for both high QoS provisioning and QoS differentiation for delivering multimedia services to heterogeneous devices. Additionally, these solutions lack wireless network condition awareness, as wireless communications are affected by interference, collisions, link rate adaptation, etc.

This chapter introduces **the intelligent Prioritized Adaptive Scheme (iPAS)**, which provides QoS differentiation between multiple streams during wireless multimedia delivery. iPAS assigns dynamic priorities to various streams and determines their bandwidth share by employing a stereotypes-based approach. The priority level of individual streams is variable and dependent on stream-related characteristics (i.e. device resolution, device battery power left, and application type) and delivery-related QoS parameters (i.e. delay, jitter, and packet loss rate). This thesis considers the situation in which any user has assigned one device to which multiple streams can be delivered. iPAS is incorporated into the IEEE 802.21 framework [26] supporting the network information gathering as well as dissemination.

6.2 iPAS System Architecture

Figure 6-1 presents the iPAS system architecture which consists of two main building blocks: *iPAS server* side and *iPAS client* side modules. *iPAS server* is responsible with managing bandwidth resources using stereotype-based resource allocation and a bandwidth estimation scheme (MBE). iPAS client collects information of stream preferences and QoS, which is sent as feedback to iPAS server. Multimedia traffic is delivered using Real-time Transport Protocol (RTP) [73] and feedback traffic is transmitted using the IEEE 802.21 framework. Details of each sub-module in iPAS system are discussed next.

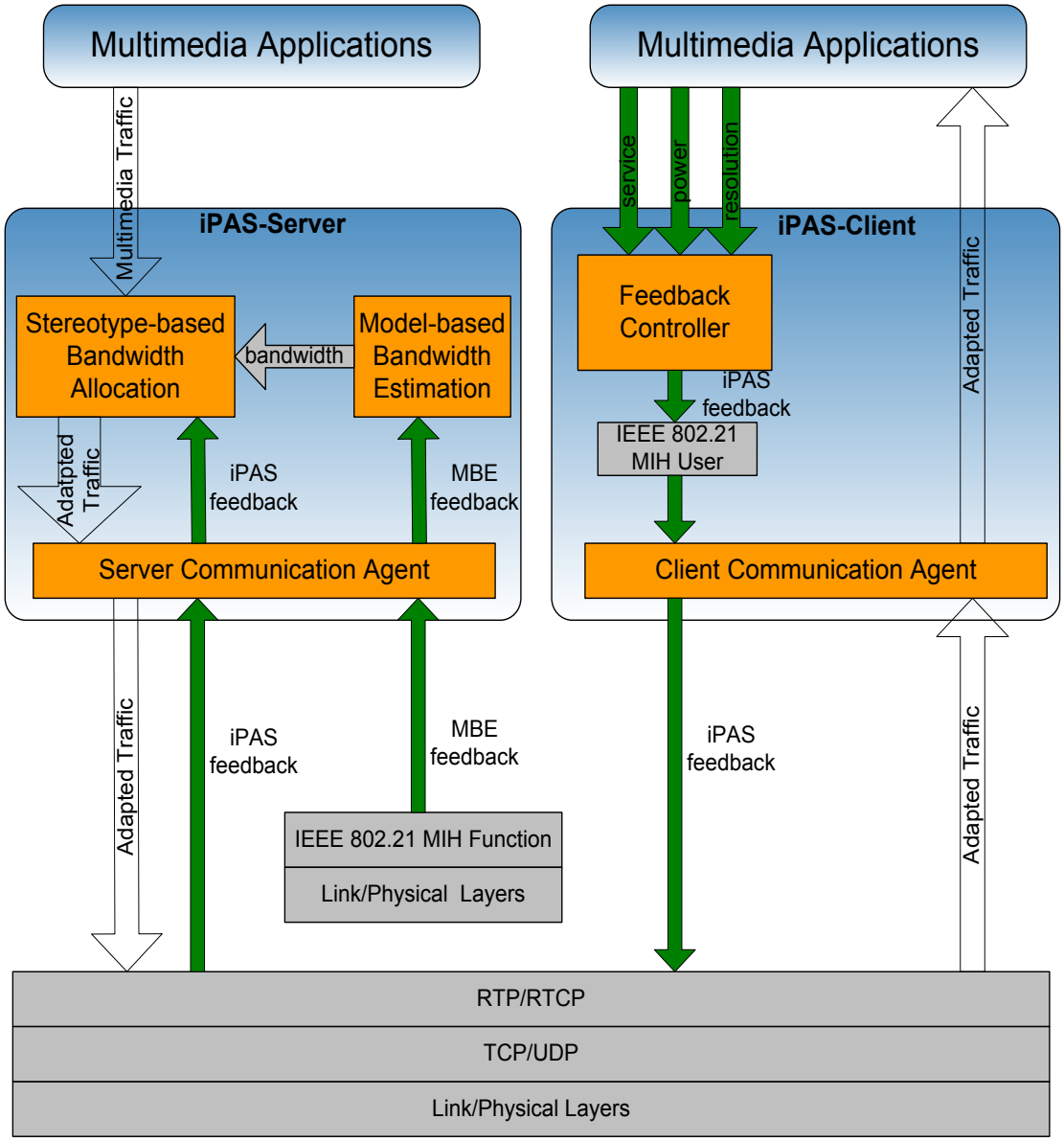


Figure 6-1 The Block Structure of the iPAS-based multimedia delivery system

6.2.1 Block-level Architecture and Principle

There are three steps in the functionality of iPAS: 1) Assign priority to each stream; 2) Allocate certain bandwidth share (expressed in percentage) for each stream with a probability value determined by the stereotype-based algorithm; 3) Allocate specific bandwidth among streams by combining the stream's bandwidth share and the estimated available bandwidth according to MBE.

The communication between the iPAS server and the iPAS client application makes use of a control communication link which is created when the client sends a request to the server. This request is used for the transmission of control messages including the feedback information. Subsequently, the multimedia communication link is established between the server and client allowing for multimedia data transmission.

Model-based Bandwidth Estimation (MBE)

The MBE block deploys the novel ***Model-based Bandwidth Estimation*** **(*MBE)*** algorithm for multimedia services over the IEEE 802.11 networks proposed in chapter 5. *MBE* considers the two basic, yet most widely used transport layer protocols, TCP and UDP, separately.

First, *MBE* models the TCP over WLAN throughput. The achievable bandwidth *B* for each TCP connection is described in equation (5-22), where *b* is the number of packets acknowledged by a received ACK, P_{retr}^{TCP} denotes the probability of packet retransmission, *MRTT* is the transport layer round-trip time between sender and receiver, and *MSS* is the maximum segment size. The proposed equation considers both TCP congestion control mechanism and wireless channel characteristics, i.e. transmission error, contention, and retries attempts.

Secondly, *MBE* approximates the UDP throughput by analyzing the UDP packet transmission probability and delay of the IEEE 802.11 channels. The maximum achievable bandwidth for UDP traffic over 802.11 WLANs is given in equation (5-28), where *Payload* is the total information transmitted during one time period from T_0 to T_1, and *Delay_UDP* denotes the average delay for successfully transmitting the individual UDP packet.

Additionally, *MBE* derives a formula predicting the achievable bandwidth when TCP and UDP co-exist in 802.11 networks, as shown in equation (5-29). The parameter *w* is the bandwidth weight factor since TCP and UDP have different bandwidth requirements.

Both extensive simulations and real tests which were performed demonstrate that *MBE* performs very well in conditions with variable packet size, dynamic wireless link rates and different channel noise.

Server Communication Agent (SCA) and Client Communication Agent (CCA)

The **Server Communication Agent (SCA)** and **Client Communication Agent (CCA)** situated at server and client, respectively, are responsible with establishing and maintaining the double-directional communicational link. The CCA component maintains the receiver buffer and forwards feedback messages from the **Feedback Controller (FC)** component to the 802.21 interface. The SCA component manages the sending buffer and forwards the feedback messages from the 802.21 interface to the **Module-based Bandwidth Estimation (MBE)** and **Stereotype-based Bandwidth Allocation (SBA)** component. iPAS makes uses of IEEE 802.21 framework (i.e. MIIS function) to transmit the control signals in order to prevent modifications of current communication protocols like 802.11. The multimedia traffic runs over the RTP/TCP/UDP protocol.

Stereotype-based Bandwidth Allocation (SBA)

The **Stereotype-based Bandwidth Allocation (SBA)** component located in the iPAS server is the center piece of the iPAS system. SBA is responsible for determining each stream's priority level and suggesting a proportional bandwidth share. Two types of control information, estimated bandwidth from MBE and feedback information from the SCA, are utilized by SBA for analysis. The MBE component estimates the available bandwidth using equations (6-1), (6-2), and (6-3) based on the feedback information (loss and transmitted data size) sent from SCA. The details of the stereotype-based process for bandwidth allocation are presented in the next section.

Feedback Controller (FC)

The main function of the **Feedback Controller (FC)**, located at the iPAS client, is to gather the feedback-related parameters from client applications and CCA and send the formatted feedback messages to the CCA. Two types of feedback related parameters are processed in the FC component:

1) *Stream characteristics related parameters* such as the application type, device resolution, and device power left. They are initialized by the client application process when sending the first request and updated whenever there is a change;

2) *Delivery QoS-related parameters* such as delay, jitter, and packet loss rate, which are extracted from the receiving buffer of CCA. The computation of the instant delay takes into consideration packets' timestamps as suggested in [27], the calculation of instant jitter is based on the computed delay as shown in [29], and the measurement of instant packet loss rate is done by analyzing the packets' sequence numbers as presented in [30]. These measured values are monitored by the FC and sent to the iPAS server as feedback messages. The instant values of the QoS parameters (delay, jitter, packet loss rate) are computed each time the multimedia packet arrives at the client. To alleviate the fluctuation of these QoS parameters values, average values $AVGQoS_{delay}$, $AVGQoS_{jitter}$, and $AVGQoS_{loss}$ are considered for which one of the QoS parameters: delay, jitter, packet loss rate, respectively. The incremental computation of the estimated average values is suggested in [207] and is given in equation (6-1) as an example for the *delay* parameter. α is a smoothing factor originally used in the TCP standard [207] when estimating the round trip time. Higher values of α give better smoothing and minimize the negative influence of sudden network changes. Experimentally, α=0.9 was demonstrated to be a good value and was used by TCP [207]. $AVGQoS_{delay}'$, $AVGQoS_{jitter}'$, and $AVGQoS_{loss}'$ are initialized with the first QoS parameter value available and updated at the end of each monitoring interval.

$$AVGQoS_{delay} = AVGQoS_{delay}' \times \alpha + QoS_{delay} \times (1-\alpha)$$

$$AVGQoS_{jitter} = AVGQoS_{jitter}' \times \alpha + QoS_{jitter} \times (1-\alpha)$$

$$AVGQoS_{loss} = AVGQoS_{loss}' \times \alpha + QoS_{loss} \times (1-\alpha) \qquad (6\text{-}1)$$

6.2.2 Block Sequence Diagram

Figure 6-2 shows the message exchange between iPAS blocks. Detailed explanation of the messages and process is presented next based on the time sequence of multimedia delivery.

- The process starts by the client making a request to the content server. This request starts by message 1 and is confirmed by message 8. The type of request messages depends on the transport protocol adopted, such as TCP or RTP/UDP. In the case of TCP, there is three-way handshake where the connection starts from message 1 and ends by message 12. Unlike in TCP, UDP allows application process directly talks with IP layer, and messages 9 to 12 would not be used. The stream preference

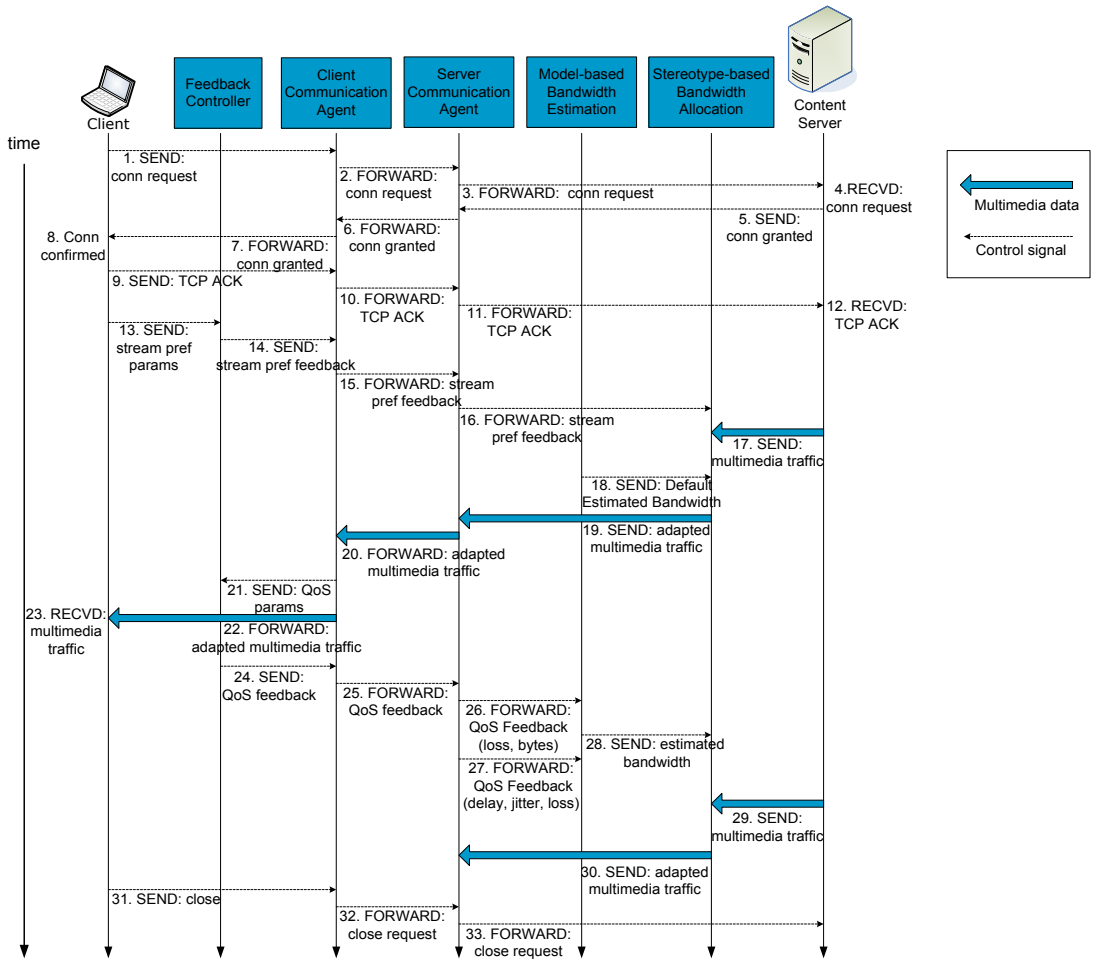


Figure 6-2 Block Sequence Diagram of iPAS system.

related parameters (device resolution, device power left, and application type) are sent to the FC component at the client side during the connection period (message 13). These parameters are encapsulated into the packet header defined by the employed transport protocol.

- Once an agreement between the client and the content server is established, the server starts to transmit multimedia traffic (message 17) which is processed by the Stereotype-based Bandwidth Allocation (SBA) model. SBA receives the default estimated available bandwidth from the Model-based Bandwidth Estimation (MBE) module (message 18) and the stream preference related feedback information (message 16) from the Server Communication Agent (SCA). Next, SBA allocates the bandwidth using the stereotype-based resource allocation algorithm based on the information provided by MBE and SCA. Finally, SBA outputs the adapted traffic to the SCA for lower layer communication. At the beginning of the transmission, QoS

feedback is missing as no multimedia traffic has been received. Therefore, SBA assigns priority based on the stream preference feedback which includes device resolution, device battery power left, and service type.

- When SCA receives the adapted multimedia traffic, as indicated by message 19, it forwards these data to the IP-based networks (message 20), i.e., Wireless LAN, Ethernet, etc. SCA maintains separate sender buffer and packet scheduler for each stream, consequently, each multimedia flow will be sent with the adapted transmission rate suggested by SBA.

- The Client Communication Agent (CCA) module of the client receives and stores the transmitted multimedia traffic in the receiving buffer. Meanwhile, CCA analyzes the delivered QoS for each stream, such as delay, jitter, packet loss rate and received bytes, and sends these information to the Feedback Controller module (message 21). The multimedia data stored in the receiver buffer will be forwarded to the client application process (message 22).

- After receiving the delivered QoS parameters, FC in client side formats these variables into the QoS feedback message (message 24), which are forwarded by CCA to the server (message 25).

- Once message 25 is received by the SCA in server side, the feedback information is extracted and forwarded into MBE (packet loss rate and delivered bytes, as indicated by message 26) and SBA (delay, jitter, and packet loss rate, as shown in message 27).

- The incoming multimedia traffic will be re-allocated by SBA as indicated by message 30. In contrast with message 19, current multimedia traffic is adapted based on the previous QoS feedback information as well as the stream preference related feedback.

- Finally, messages 31 to 33 close the multimedia transmission. Upon receiving the close command, each module in the iPAS system will release the resources occupied, such as buffers, queues, etc.

6.3 Stereotype-based Bandwidth Allocation

Stereotypes for managing groups have been widely used by many QoS-oriented adaptive solutions [163] [164] [165]. In the thesis, stereotypes are defined as stream classes (groups) described by a set of features, which include attributes. Each stream will belong to every stereotype group with a certain probability depending on stream features. These features include delay, jitter, packet loss rate, service, device resolution, and battery power left. iPAS utilizes the stereotypes to build stream profiles and then suggest a proper bandwidth share for each stream. The bandwidth share of a stream will be suggested by combining the features and the probabilities.

6.3.1 Principle of Stereotype-based Resource Allocation

In this thesis, five stereotypes classes (*Th*) are defined: *High Priority (HP), Medium to High Priority (MHP), Medium Priority (MP), Medium to Low Priority (MLP), Low Priority (LP)*. Each $Stream_i$ belongs to one of the five stereotypes with a certain probability. Each stereotype class *Th* consists of two components: a group of features $F= (F_1, F_2, ..., F_i, ..., F_m)$ describing the stereotype and a group of suggestions $S= (S_1, S_2, ..., S_j, ..., S_n)$ that should be performed to determine stream's bandwidth share. Each feature F_i has associated a list of linguistic terms $LF_i = (LF_{i1}, LF_{i2},...,LF_{iq})$. Each linguistic term LF_{iq} has a numeric value PF_{iq} between 0 and 1, representing the probability that the feature F_i equals the linguistic term LF_{iq} for this stereotype *Th*. The probability PF_{iq} indicates the degree of match between stream's characteristics and the stereotype. A similar structure is defined for each suggestion S_j, which has also associated the linguistic terms $LS_j = (LS_{j1}, LS_{j2},...,LF_{ip})$ probabilistic values PS_{jp}. Table 6-1 and Table 6-2 present the group of features and suggestions for a stereotype.

TABLE 6-1 GROUP OF FEATURES FOR A STEREOTYPE

Features	(Linguistic Term, Probability)
F_1	$(LF_{11}, PF_{11}), (LF_{12}, PF_{12}), ... , (LF_{1q}, PF_{1q})$
F_2	$(LF_{21}, PF_{21}), (LF_{22}, PF_{22}), ... , (LF_{2q}, PF_{2q})$
...	...
F_m	$(LF_{m1}, PF_{m1}), (LF_{m2}, PF_{m2}), ... , (LF_{mq}, PF_{mq})$

TABLE 6-2 GROUP OF SUGGESTIONS FOR A STEREOTYPE

Features	(Linguistic Term, Probability)
S_1	$(LS_{11}, PS_{11}), (LS_{12}, PS_{12}), \ldots, (LS_{1p}, PS_{1p})$
S_2	$(LS_{21}, PS_{21}), (LS_{22}, PS_{22}), \ldots, (LS_{2p}, PS_{2p})$
...	...
S_n	$(LS_{n1}, PS_{n1}), (LS_{n2}, PS_{n2}), \ldots, (LS_{np}, PS_{np})$

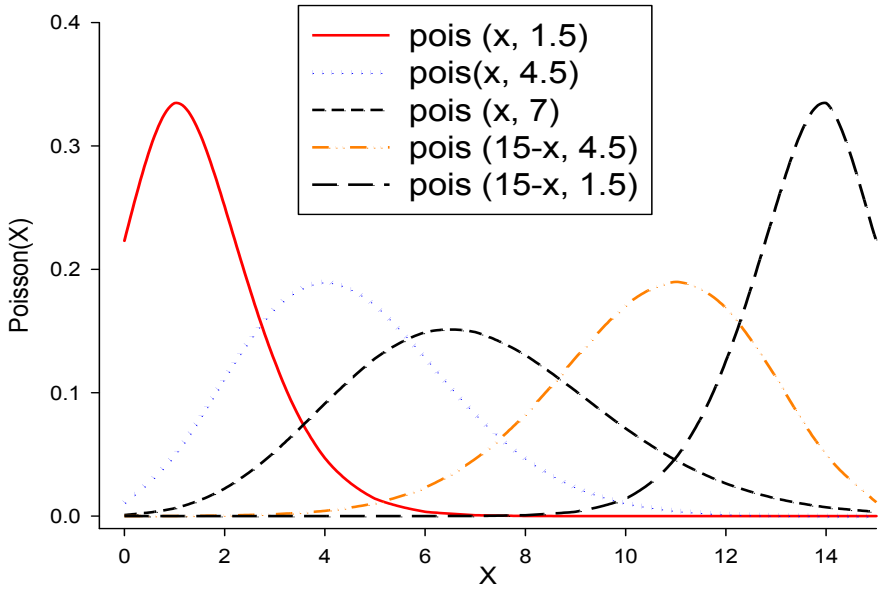


Figure 6-3 Poisson distribution for five stereotypes

The Poisson distribution is used to determine the probability associated with the linguistic terms. The Poisson distribution represents the probability of a given number of events occurring in a fixed interval of time. The occurrence of each event is independent of time of the last event. Equation (6-2) shows the Poisson distribution function where u is the shape parameter and indicates the mean and the variance of the distribution during a time interval. The integer value x (x=0, 1, 2, …, n) represents a particular event.

$$pois(x,u) = \frac{u^x \times \exp(-u)}{x!} \tag{6-2}$$

By analysing the shape of the Poisson function, a near normal distribution is obtained for u=7 across the [0, 15] interval. The selected value of u has also been used and validated in [163] for network parameters modelling. The maximum value of the normal distribution close to 0.15 (x=7, u=7) and the minimum value close to 0 (x=0 or x=15, u=7).

Consequently, the interval [0, 15] is considered for the computation of the Poisson function for all the stereotypes. Figure 6-3 shows the example when iPAS uses five stereotypes: *HP, MHP, MP, MLP,* and *LP*. It is noticed that each stereotype associates one Poisson distribution with a mean value u_k which is obtained by dividing the interval [0, 15] in five equal segments and considering their middle value. As shown in the Figure 6-3, the peak value of Poisson function increases when u_k gets closer to zero.

Considering a feature F_i has a list of linguistic terms, where the list length is q, the probabilistic values for each term PF_{ij} are computed as in equations (6-3), (6-4), and (6-5). The value i implies the index of feature and j is the index of linguistic term in feature i,

$$PFij = Average(pois(x_j, u_k)) \tag{6-3}$$

$$x_j \in [step \times (j-1), step \times j] \tag{6-4}$$

$$step = \left\lceil \frac{15}{q} \right\rceil \tag{6-5}$$

1) Stream Classification

The purpose of the stream classification is to determine the stereotype classes the stream belongs to and with what probability. iPAS describe a stream with the format shown in (6-6), where F_i is the name of the i_{th} feature and LF_iK_i represents the linguistic term of the i_{th} feature.

$$U = ((F_1, LF_1K_1), (F_2, LF_2K_2), ..., (F_m, LF_mK_m)) \tag{6-6}$$

A degree of match between a stream and each stereotype is computed in (6-7) based on probability theory.

$$\begin{aligned} M(Th) &= p(Th \mid F_1 = LF_1K_1, ..., F_m = LF_mK_m) \\ &= p(Th \mid F_1 = LF_1K_1) \times ... \times p(Th \mid F_m = LF_mK_m) \end{aligned} \tag{6-7}$$

The computation of each factor is computed using the Bayes rule, as shown in (6-8),

$$p(Th \mid F_i = LF_iK_i) = \frac{p(F_i = LF_iK_i \mid Th) \times p(Th)}{p(F_i = LF_iK_i)} = \frac{PF_iK_i \times p(Th)}{p(F_i = LF_iK_i)} \tag{6-8}$$

2) Suggestion Determination

The suggestion determination procedure is performed to get the bandwidth share for each stream.

First, for each stereotype class, the strength regarding each suggestion has to be re-computed by considering the probability with which the stream belongs to this class, as given in equations (6-9) and (6-10), where S_i means the name of the i_{th} suggestion and LS_iK_i represents the linguistic term of the i_{th} suggestion.

$$p'(S_i = LS_i k_i \mid Th) = p(S_i = LS_i k_i \mid Th) \times M(Th) \tag{6-9}$$

$$PS'_i k_i(Th) = PS_i k_i(Th) \times M(Th) \tag{6-10}$$

Second, the combination of all the stereotype suggestions produced from (6-10) can be calculated using probabilistic theory. An exemplification of how these equations are used is shown next.

6.3.2 Stereotype-based Resource Allocation for iPAS

The stereotype classes defined by iPAS are created based on six features: *delay, jitter, loss, power left, device resolution, and application type*. Delay, jitter, and loss are QoS parameters of the streams. Power left and device resolution indicate the client device conditions. The application type refers to one of the five widely used application types, which extend the 802.11e four class model: VoIP, Standard-Definition Video (SD-Video), High-Definition Video (HD-Video), Best-effort Service, and Background Traffic. Each feature is divided into five levels using threshold values, as shown in Table 6-3. These threshold values are suggested based on ITU-T Rec. G.1010 [208] and ITU-T Rec.Y.1541 [31]. Different applications have specific requirements on the QoS features. Take VoIP for example, a one way delay of less than 150ms indicates excellent quality, while delay higher than 400ms causes bad perceived quality. iPAS assigns higher priority to traffic which is sensitive to delay and jitter, i.e. voice, video.

All five stereotypes have the same structure: six features and each feature consist of five linguistic term-probability pairs. For the purpose of demonstration, groups of features and suggestions for the five priority stereotype (*MP*) are shown in Table 6-4 to Table 6-13.

TABLE 6-3 CLASSIFICATION OF FEATURES IN STEREOTYPE CLASSES

	Level 1	Level2	Level3	Level4	Level5
Delay	≤150ms	(150ms~400ms]	(400ms~1s]	(1s~5s]	>5s
Jitter	≤40ms	(40ms~50ms]	(50ms~60ms]	(60ms~70ms]	>70ms
Loss	$<10^{-5}$	10^{-5}~1%	1%~2%	2%~5%	>5%
Power Left	[100%~80%]	(80%~60%]	(60%~40%]	(40%~20%]	(20%~0]
Device Resolution	≥1024x768	(1024x768~768x480]	(768x480~480x360]	(480x360~320x240]	≤320x240
Application Type	VoIP	HD-Video	SD-Video	Best-effort	Background

TABLE 6-4 GROUP OF FEATURES FOR STEREOTYPE-HIGH PRIORITY

Feature	List
Delay (ms)	(≤150,0), ((150~400],0), ((400~1000],0.004), ((1000~5000],0.168), (>5000,0.828)
Jitter (ms)	(≤40, 0), ((40~50],0), ((50~60], 0.004),((60~70], 0.168), (>70ms,0.828)
Loss	($\leq 10^{-5}$, 0), ((10^{-5}~1%],0), ((1%~2%], 0.004),((2%~5%], 0.168), (>5%,0.828)
Power left	([0~20%], 0), ((20%~40%],0), ((40%~60%], 0.004),((60%~80%], 0.168), ((80%~100%],0.828)
Device Resolution	(≤320x240, 0), ((320x240~480x360],0), ((480x360~768x480], 0.004), ((768x480~1024x768],0.168), (>1024x768, 0.828)
Application type	(Background, 0), (Best-effort,0), (SD-Video,0.004), (HD-Video,0.168), (VoIP, 0.828)

TABLE 6-5 GROUP OF SUGGESTIONS FOR STEREOTYPE-HIGH PRIORITY

Feature	List
Bandwidth share	(0~20%, 0), (20%~40%, 0), (40%~60%, 0.004), (60%~80%, 0.168), (80%~100%, 0.828)

TABLE 6-6 GROUP OF FEATURES FOR STEREOTYPE-MEDIUM TO HIGH PRIORITY

Feature	List
Delay (ms)	(≤150,0.002), ((150~400],0.03), ((400~1000],0.212), ((1000~5000],0.497), (>5000,0.259)
Jitter (ms)	(≤40, 0.002), ((40~50],0.03), ((50~60], 0.212),((60~70], 0.497), (>70ms,,0.259)
Loss	($\leq 10^{-5}$, 0.002), ((10^{-5}~1%],0.03), ((1%~2%], 0.212),((2%~5%], 0.497), (>5%,,0.259)
Power left	([0~20%],0.002), ((20%~40%],0.03), ((40%~60%], 0.212),((60%~80%], 0.497), ((80%~100%],,0.259)
Device Resolution	(≤320x240, 0.002), ((320x240~480x360],0.03), ((480x360~768x480], 0.212), ((768x480~1024x768],0.497), (>1024x768, ,0.259)
Application type	(Background, 0.002), (Best-effort,0.03), (SD-Video,0.212), (HD-Video,0.259), (VoIP, ,0.497)

TABLE 6-7 GROUP OF SUGGESTIONS FOR STEREOTYPE-MEDIUM TO HIGH PRIORITY

Feature	List
Bandwidth share	(0~20%, 0.002), (20%~40%, 0.03), (40%~60%, 0.212), (60%~80%, 0.497), (80%~100%, 0.259)

TABLE 6-8 GROUP OF FEATURES FOR STEREOTYPE-MEDIUM PRIORITY

Feature	List
Delay (ms)	(≤150,0.062), ((150~400],0.317), ((400~1000],0.399), ((1000~5000],0.184), (>5000,0.038)
Jitter (ms)	(≤40, 0.062), ((40~50],0.317), ((50~60], 0.399),((60~70], 0.184), (>70ms,0.038)
Loss	(≤10^{-5}, 0.062), ((10^{-5}~1%],0.317), ((1%~2%], 0.399),((2%~5%], 0.184), (>5%,0.038)
Power left	([0~20%],0.062), ((20%~40%],0.317), ((40%~60%], 0.399),((60%~80%], 0.184), ((80%~100%],0.038)
Device Resolution	(≤320x240, 0.062), ((320x240~480x360],0.317), ((480x360~768x480], 0.399), ((768x480~1024x768],0.184), (>1024x768, 0.038)
Application type	(Background, 0.038), (Best-effort,0.062), (SD-Video,0.184), (HD-Video,0.317), (VoIP, 0.399)

TABLE 6-9 GROUP OF SUGGESTIONS FOR STEREOTYPE-MEDIUM PRIORITY

Feature	List
Bandwidth share	(0~20%, 0.062), (20%~40%, 0.317), (40%~60%, 0.399), (60%~80%, 0.184), (80%~100%, 0.038)

TABLE 6-10 GROUP OF FEATURES FOR STEREOTYPE-MEDIUM TO LOW PRIORITY

Feature	List
Delay (ms)	(≤150,0.259), ((150~400],0.497), ((400~1000],0.212), ((1000~5000],0.03), (>5000,0.02)
Jitter (ms)	(≤40, 0.259), ((40~50],0.497), ((50~60], 0.212),((60~70], 0.03), (>70ms, 0.02)
Loss	(≤10^{-5}, 0.259), ((10^{-5}~1%],0.497), ((1%~2%], 0.212),((2%~5%], 0.03), (>5%,0.02)
Power left	([0~20%],0.259), ((20%~40%],0.497), ((40%~60%], 0.212),((60%~80%], 0.03), ((80%~100%],0.02)
Device Resolution	(≤320x240, 0.259), ((320x240~480x360],0.497), ((480x360~768x480], 0.212), ((768x480~1024x768],0.03), (>1024x768, 0.02)
Application type	(Background, 0.002), (Best-effort,0.03), (SD-Video,0.212), (HD-Video,0.259), (VoIP, 0.497)

TABLE 6-11 GROUP OF SUGGESTIONS FOR STEREOTYPE-MEDIUM TO LOW PRIORITY

Feature	List
Bandwidth share	(0~20%, 0.259), (20%~40%, 0.497), (40%~60%, 0.212), (60%~80%, 0.03), (80%~100%, 0.02)

TABLE 6-12 GROUP OF FEATURES FOR STEREOTYPE-LOW PRIORITY

Feature	List
Delay (ms)	(≤150,0.828), ((150~400],0.168), ((400~1000],0.004), ((1000~5000],0), (>5000,0)
Jitter (ms)	(≤40, 0.828), ((40~50],0.168), ((50~60], 0.004),((60~70], 0), (>70ms,0)
Loss	(≤10^{-5}, 0.828), ((10^{-5}~1%],0.168), ((1%~2%], 0.004),((2%~5%], 0), (>5%,0)
Power left	([0~20%],0.828), ((20%~40%],0.168), ((40%~60%], 0.004),((60%~80%], 0), ((80%~100%],0)
Device Resolution	(≤320x240, 0.828), ((320x240~480x360],0.168), ((480x360~768x480], 0.004), ((768x480~1024x768],0), (>1024x768, 0)
Application type	(Background, 0), (Best-effort,0), (SD-Video,0.004), (HD-Video,0.168), (VoIP, 0.828)

TABLE 6-13 GROUP OF SUGGESTIONS FOR STEREOTYPE-LOW PRIORITY

Feature	List
Bandwidth share	(0~20%, 0. 828), (20%~40%, 0. 168), (40%~60%, 0. 004), (60%~80%, 0), (80%~100%, 0)

TABLE 6-14 PROBABILISTIC RESULTS INDICATING THE MATCH DEGREE AND BANDWIDTH SHARE

Stereotype	Probability-Stream$_1$	Probability-Stream$_2$
High Priority (HP)	0%	0%
Medium High Priority (MHP)	57.05%	0%
Medium Priority (MP)	42.95%	11.82%
Medium Low Priority (MLP)	0%	88.18%
Low Priority (LP)	0%	0%
Bandwidth Share	**Probability-Stream$_1$**	**Probability-Stream$_2$**
0~20%	17.44%	0.91%
20%~40%	41.97%	6.39%
40%~60%	29.23%	23.41%
60%~80%	9.61%	46%
80%~100%	1.75%	23.29%

6.3.3 Exemplification

With the stereotype classes proposed, we give an illustration of the bandwidth allocation process involving two streams: U_1 and U_2.The procedure includes the initialization phase (steps 1 to 3) and the update phase (step 4)

Step1: Collect stream related parameters regarding each feature F_i of the stereotype Th. These parameters include: delay, jitter, packet loss rate, power left, device resolution, and application type. For the purpose of demonstration, the two streams are first configured with the six features using static values. The features of the two streams are shown in equations (6-11) and (6-12).

$$U_1 = (delay, 150ms), (jitter, 50ms), (loss, 1\%), (power, 85\%), (resolution, 480 \times 360), (applicatio\,n, VoIP) \tag{6-11}$$

$$U_2 = (delay, 1500ms), (jitter, 70ms), (loss, 2\%), (power, 85\%), (resolution, 768 \times 480), (applicatio\,n, HTTP) \tag{6-12}$$

Step2: Determine the degree of match between the stream and each stereotype. This can be done using equations (6-6) to (6-8). After the normalization of the calculated values, we have probabilistic results indicating the match degree and bandwidth share suggestion, as shown

in Table 6-14. The two streams have different probabilities to belong to the stereotype classes, i.e., U_1 belongs to the MHP and MP stereotype with the probability of 57.05% and 42.95%, respectively and 0% to the other stereotype classes.

Step3: The maximum bandwidth share for U_1 and U_2 are denoted as B_1_MAX and B_2_MAX which can be calculated based on Table 6-14 and shown in equations (6-13) and (6-14):

$$\begin{aligned} B_1\_MAX &= 20\% \times 17.44\% + 40\% \times 41.97\% + 60\% \times 29.23\% \\ &\quad + 80\% \times 9.61\% + 100\% \times 1.75\% \\ &= 47.25\% \end{aligned} \tag{6-13}$$

$$\begin{aligned} B_2\_MAX &= 20\% \times 0.91\% + 40\% \times 6.39\% + 60\% \times 23.41\% \\ &\quad + 80\% \times 46\% + 100\% \times 23.29\% \\ &= 76.87\% \end{aligned} \tag{6-14}$$

By normalizing B_1_MAX and B_2_MAX, the bandwidth share for U_1 and U_2 are 38.07% and 61.93%, respectively. Consequently, the actual amount of bandwidth can be obtained by making use of the overall available bandwidth which is estimated by MBE.

Step4: Actual values for the QoS parameters during data transmission for each stream will be sent back regularly to the stereotype-based bandwidth allocation module in iPAS. Step 1 to step 3 are repeated to update the priority level and bandwidth share of each stream is re-evaluated. The probabilistic values PF_ik_i associated with the linguistic values LF_ik_i are recalculated.

The above four steps present the iPAS procedure for bandwidth allocation using stereotypes. The bandwidth share of certain stream depends on six features (delay, jitter, loss, power, resolution, and application) and the available bandwidth. The same procedure can be applied for any number of streams. Notably, our stereotype-based resource allocation model considers the same probability distribution, i.e., each stereotype class consists of six features and each feature is further divided into five levels. Different probability distributions can also be considered in this model based on the number of features and classification of feature's linguistic values.

6.4 Summary

This chapter introduces an intelligent Prioritized Adaptive Scheme (iPAS) to provide both QoS differentiation and high QoS levels for content delivery to heterogeneous devices over IEEE 802.11 networks. iPAS algorithm assigns a dynamic priority to each multimedia stream and suggests a proportional bandwidth share according to the results of a stereotype-based bandwidth allocation solution. This solution considers both QoS-related parameters such as delay, jitter, and packet loss rate and stream-related characteristics including device resolution, remaining device battery power, and application type. Performance evaluation, assessed in terms of six metrics: an inter-stream fairness index, throughput, packet loss rate, delay, video quality, and device resolution-awareness will be presented in a future chapter.

CHAPTER 7

QoS-based Downlink/Uplink Fairness for VoIP in Wireless LANs

This chapter presents a new scheme which can provide QoS-based fairness between downlink and uplink for Voice over IP (VoIP) services in IEEE 802.11 networks. The principle of the solution is to utilize the stereotypes to balance the achieved QoS parameters, i.e., throughput, delay, and loss. This chapter consists of four sections: 1) section 7.1 introduces the background and motivation of the proposed solution; 2) section 7.2 illustrates the details of the algorithm; 3) section 7.3 presents the simulation-based experiments and result analysis; 4) section 7.4 summarises the chapter.

7.1 Introduction

IEEE 802.11 wireless local area networks (WLANs) have been widely deployed for Internet access in homes, public institutions and companies [209]. Meanwhile, popular VoIP software products, such as Skype[47] and Viber[48], have been supported by the majority of mobile consumer devices and have attracted millions of users. A detailed investigation shows that more than 50 percent of voice calls originates from indoor WLANs [210]. As shown in Figure 7-1, there are diverse mobile devices connected to the access points (AP) in 802.11 WLANs. Mobile VoIP users in different WLANs communicate through remote VoIP servers. Nevertheless, the original 802.11 protocol does not support any Quality of Service (QoS) provisioning. Many solutions have been proposed to provide QoS for multimedia services in wireless networks, e.g. [211], [212], [213], [214], etc. However, these solutions do not consider the fairness between downlink and uplink traffic.

QoS of VoIP is affected by the fairness problem between the downlink and uplink traffic distribution in the IEEE 802.11 WLAN [215]. In IEEE 802.11 networks with the

[47] Skype-http://www.skype.com
[48] Viber-http://www.viber.com

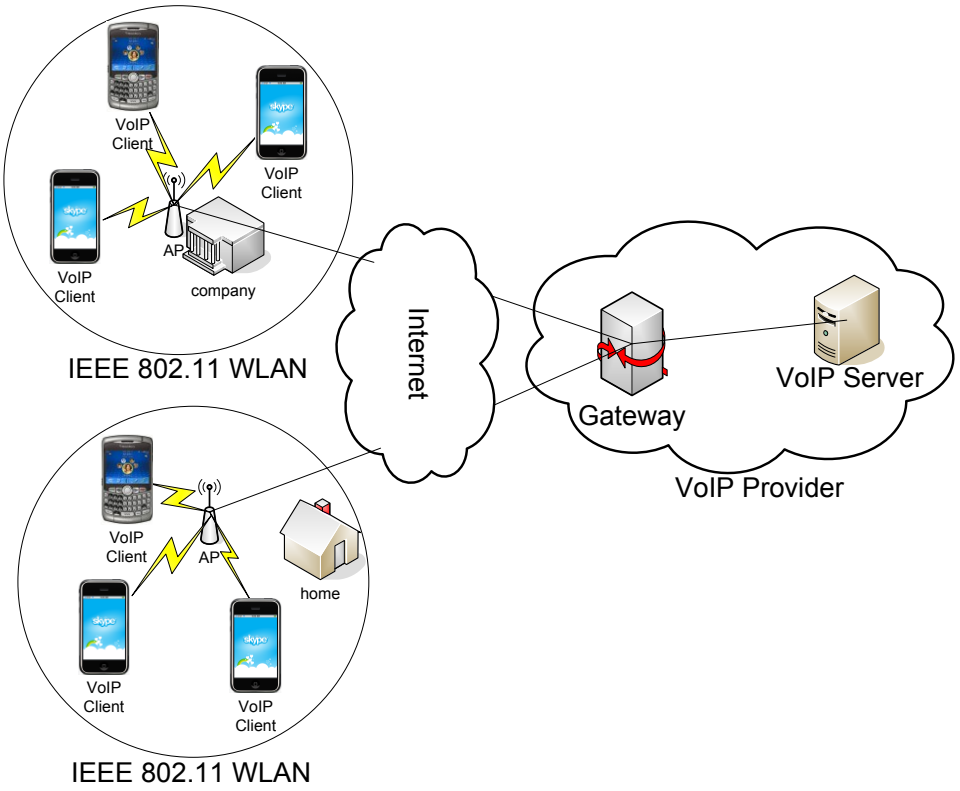


Figure 7-1 Architecture of IEEE 802.11-based VoIP Application

infrastructure mode, the downlink flows are given less channel access opportunity than the uplink flows due to the inherent contention mechanism of Carrier Sense Multiple Access with Collision Avoidance (CSMA/CA). Consequently, the downlink traffic has lower priority in accessing the channel, despite much of the traffic being downlink and not uplink. In [216], the downlink/uplink fairness is improved by controlling the minimum size of Contention Window (CW_{min}) parameter at AP according to a computed optimal ratio between the packet transmission rate of downlink flows and that of the uplink flows. A similar solution [217] achieves the downlink/uplink fairness by setting the CW_{max} size based on the channel occupancy time of aTCP traffic. In [218], the downlink/uplink fairness is provided in error-prone 802.11 WLANs by controlling both CW_{min} size and TXOP limit based on a channel error model which makes use of Markov chains. The solution proposed in [219] improves bandwidth share balance between downlink and uplink based on adaptive CW_{min}, and also, increases overall VoIP capacity using frame aggregation. However, very few of the previous research works [220] consider the QoS fairness between downlink and uplink traffic in 802.11 networks. In order to fairly balance QoS levels, it is necessary to make use of the values of several QoS-related parameters such as throughput, delay, and loss, since they are all critical for the VoIP traffic.

In order to have a fair QoS distribution between downlink and uplink, there is a need for an improved distribution of wireless channel access between the AP and the wireless stations. For instance, when the AP experiences poor QoS parameter levels in comparison with those measured at the wireless stations, there is a need to give the AP higher channel access opportunity. This can be done by considering two options: reduce CW size at AP or increase CW size at each wireless station.

This chapter proposes a QoS-based downlink/uplink fairness scheme for VoIP in IEEE 802.11 networks. The proposed algorithm dynamically controls the CW size at the AP in order to balance the downlink/uplink channel access opportunity. The optimum AP's CW size is computed based on the results of a stereotypes-based algorithm [221] which utilizes the ratio between the major QoS parameter values (i.e. throughput, delay, and loss) measured for the downlink and uplink traffic. Stereotypes for managing groups of users were first introduced by Rich in the Grundy system [162] and they are still widely used by many QoS-oriented adaptive solutions [222] [223]. The proposed algorithm collects QoS-related information at the AP and from the wireless stations via feedback.

7.2 QoS-based Downlink/Uplink Fairness Scheme

The principle behind the proposed mechanism is the utilization of stereotypes-based structure which has been introduced in chapter 6. This section illustrates the algorithm and how to use the stereotypes classes to adapt the contention windows size.

Figure 7-2 presents the system architecture of the proposed contention window adaptation scheme which is implemented at the MAC layer of the access point (AP). The scheme consists of two modules: **Contention Window Adaptation** and **QoS Monitor**. The Contention Window Adaptation module is responsible with controlling the contention window size of AP using a stereotype-based adaptive mechanism. The **QoS Monitor** module collects the QoS-related parameters by monitoring the downlink and uplink buffers maintained by AP. The QoS-related parameters for downlink and uplink traffic include throughput, delay, and loss, which are sent as feedback to the Contention Window Adaptation module. The details of the modules are presented next.

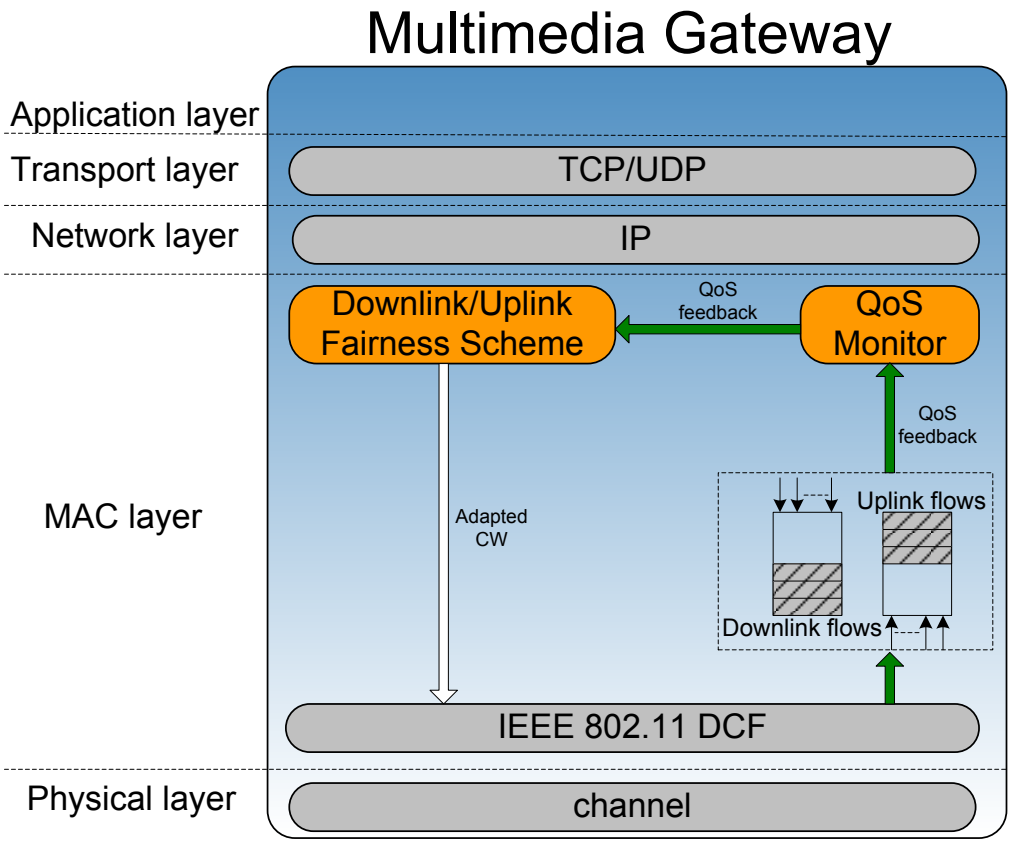


Figure 7-2 Block architecture the proposed scheme

7.2.1 QoS Monitor

There are two functions provided by the QoS Monitor module: 1) Collect QoS-related parameters (throughput, delay, and packet loss rate) for the downlink and uplink traffic; 2) Deliver the QoS parameters as feedback to the Contention Window Adaptation module. As shown in Figure 7-2, AP maintains downlink buffer and uplink buffer. Therefore, the throughput, delay, and loss for downlink and uplink flows can be measured by monitoring the buffer size. Details for computing the QoS parameters are discussed next.

Throughput Downlink/uplink Ratio

The throughput downlink/uplink ratio is considered fair when the throughput in both directions has equal values and therefore $Throughput_{down/up}$ equals one. Equation (7-1) illustrates how the downlink/uplink throughput ratio is computed by making use of the throughput at the AP ($Throughput_{AP}$) and the aggregation of the throughput at the wireless stations ($Throughput_{STAi}$), where i indicates the i^{th} wireless station. N is the number of

wireless stations. Both $Throughput_{AP}$ and the overall throughput $\sum_{i=1}^{N} Throughput_{STAi}$ of the wireless stations can be measured based on the downlink and uplink buffer at the MAC layer of the AP. For instance, the downlink/uplink throughput is the number of bytes leaves the downlink/uplink buffer during the sample time.

$$Throughput_{down/up} = \frac{Throughput_{AP}}{\sum_{i=1}^{N} Throughput_{STAi}} \tag{7-1}$$

Delay Downlink/Uplink Ratio

The downlink and uplink delay distribution is fair when the two communication directions will process the same amount of traffic during a sample interval. For example if the packet size is identical, it is fair to have the AP sending N packets to the stations (downlink) and have the N wireless stations sending N packets to the AP (uplink) in the same time period.

In order to achieve delay downlink/uplink fairness, it is noted that there are three types of delay for packet transmissions via wireless: 1) *Queuing delay*; 2) *MAC delay*; 3) *Propagation delay*. Queuing delay represents the total duration of time that packets have to wait in the queues. MAC delay is caused by the contention mechanism of CSMA/CA protocols, which may also include some uplink-downlink unfairness. However queuing delay is by far the largest delay that causes delay unfairness between downlink and uplink [219], [223] and therefore MAC delay is not considered in this paper. Propagation delay is dependent on the distance and signal propagation speed only. In the VoIP system, the wireless propagation delay is the same between downlink and uplink because they use the same medium and the packet size is the same, so it does not influence the downlink/uplink fairness.

The queuing delay is determined by many factors such as queue arriving rate, queue service rate, current queue size, etc. Since there is a desire that the proposed algorithm be deployed at the AP without modifying the wireless stations, the current queue size of the i^{th} wireless station $QSize_{STAi}$ is estimated using equation (7-2). The parameter $AvgPktSize_i$ is the average packet size received from the i^{th} wireless station during the sampled interval. λ_{STAi} represents the arrival rate of the packets entering the queue, and depends on the VoIP encoding scheme. For instance, a 64kbps VoIP traffic implies λ_{STAi}=64kbps. The VoIP

encoded information is delivered to AP via the feedback mechanism. $NumRcvdPkts_i$ is the number of packets received by the AP from the i^{th} wireless station. *Time* is the sampling time duration selected to investigate the queue state.

$$QSize_{STAi} = AVGPktSize_i \times (Time \times \lambda_{STAi} - NumRcvdPkts_i) \tag{7-2}$$

The burst arrival of packets results in a random queue size. If the sample interval is too small, it is possible that in it may be no packet transmissions due to the bursty nature of traffic. Otherwise, too large interval value reduces the update frequency and leads to inaccuracy in the queue size estimation. The sampling interval is selected based on the Nyquist theorem [224], as given in equation (7-3), where w is the signal frequency. The purpose of computing an optimal sample interval is to alleviate the aliasing phenomenon due to the traffic busrtiness.

$$Time = SamplingInterval = 1/(2 \times w) \tag{7-3}$$

$$Delay_{down/up} = \frac{QDelay_{AP}}{\sum_{i=1}^{N} QDelay_{STAi}} = \frac{QSize_{AP} / \mu_{AP}}{\sum_{i=1}^{N} QSize_{STAi} / \mu_{STAi}} \tag{7-4}$$

The delay downlink/uplink ratio is given in equation (7-4), where $QDelay_{AP}$ and $QDelay_{STAi}$ are the average queuing delay at the AP and i^{th} wireless station, respectively. $QSize_{AP}$ and $QSize_{STAi}$ are the number of bits waiting in the queue at the AP and i^{th} wireless station, respectively. μ_{AP} and μ_{STAi} are the service rate of the queue at the AP and i^{th} wireless station, respectively (i.e. the rate at which bits leave the queue). The values of $QSize_{AP}$, μ_{AP} and μ_{STAi} can be monitored based on the downlink and uplink buffer at AP. The values of $QSize_{STAi}$ are computed using equation (7-2) and equation (7-3).

Packet Loss Rate Downlink/uplink Ratio

Potential causes for packet loss during an end-to-end network data transmission are as follows:

1) *Queue Drop*: Packets can be dropped at the queue depending on the queuing management algorithms adopted. For instance, in the First Come First Service (FCFS) queues such as DropTail [225], packets are dropped when the queue has filled its capacity; in

Random Early Detection (RED) queues [226], packets are dropped with certain probability depending on the queue size;

2) *Channel Error*: packets can be dropped due to the wireless channel error;

3) *Retransmission Limit*: when packet retransmission reaches a retry limit defined by the 802.11 MAC, the packet is dropped;

4) *Collision*: collisions occur when multiple wireless stations (uplink) attempt to transmit the packets simultaneously; packets affected by collisions are dropped. There are no collisions among the downlink flows since the AP is the unique 802.11 DCF object generating traffic in the downlink mode. The packet loss rate downlink/uplink ratio is given in equation (7-5) and equation (7-6).

$$BERLoss_{AP} = BER \times \mu_{AP} \tag{7-5}$$

$$LossRate_{down/up} = \frac{QDropRate_{AP} + BERLoss_{AP} + RETRANLoss_{AP}}{\sum_{i=1}^{N} Loss_{STAi}} \tag{7-6}$$

The parameters $QDropRate_{AP}$, $BERLoss_{AP}$, and $RETRANLoss_{AP}$ represent the packet loss at AP caused by queue drop, channel error and retransmission limit, respectively. $QDropRate_{AP}$ and $RETRANLoss_{AP}$ are captured at the MAC layer of the AP, and $BERLoss_{AP}$ is computed using equation (7-7), where μ_{AP} is the service rate of the AP queue, M and N are the number of downlink and uplink flows, respectively. The parameter $Loss_{STAi}$ is the packet loss rate of the i^{th} wireless station which is measured at the AP based on packet sequence number in the uplink buffer.

$$FI_{down/up} = \frac{(\sum_{i=1}^{M} Q_D^i + \sum_{j=1}^{N} Q_U^j)^2}{(M+N)(\sum_{i=1}^{M} (Q_D^i)^2 + \sum_{j=1}^{N} (Q_U^j)^2)} \tag{7-7}$$

7.2.2 Contention Window Adaptation

Stereotypes classes are used to represent five fairness levels of traffic distribution between downlink and uplink. Theoretically, the number of fairness levels can have any value; in practice a not very high value is used to minimise the complexity, but allow for high enough adaptation granularity. In this paper, five fairness levels are selected in order to correspond to the five Mean Opinion Score levels in ITU-T Recommendations P.800 [45]: “Bad”, “Poor”, “Normal”, “Good”, and “Excellent”. Three QoS performance parameters,

$Throughput_{down/up}$, $Delay_{down/up}$, and $LossRate_{down/up}$ are modelled as stereotype features, representing throughput ratio, delay ratio and loss ratio between downlink and uplink communication channels, respectively. The computation of the downlink/uplink ratio for each of the QoS parameters is described in the next section. The linguistic terms for the features of each stereotype are denoted using five ranges representing the possible ratios between downlink and uplink as follows: ">1.5" (LF_{i1}), "1.2-1.4" (LF_{i2}), "0.9-1.1" (LF_{i3}), "0.6-0.8" (LF_{i4}), "<0.6" (LF_{i5}). The parameter i indicates the i^{th} stereotype.

The initial CW size for the AP and each wireless station is selected randomly from 0 to CW_{min} (equal to 15). The idea of adapting the AP's CW using the stereotype-based structure is to associate the CW size with stereotype suggestions. The CW range (i.e., 15-1023, as specified in standard) is equally divided into five levels representing five suggestion linguistic terms as follows: "0-15" (LS_{j1}), "15-267" (LS_{j2}), "267-519" (LS_{j3}), "519-771" (LS_{j4}), "771-1023" (LS_{j5}). The parameter j indicates the j^{th} stereotype. The adapted CW size of AP is then computed using the three step process introduced in the previous section: User Classification, Suggestion Determination and Update. Note that the original CW range can also be divided unequally.

7.2.3 Overhead Analysis

The overhead of the proposed scheme mainly come from the feedback traffic which is delivered from the QoS Monitor module to the Contention Window Adaptation module. The overhead consists of two aspects, feedback frequency and feedback packet length. Since the feedback is introduced and transmitted inside the AP other than through the wireless network, the overhead is ignored in this thesis due to the high processing ability of hardware.

7.3 Summary

This chapter proposes a new Contention Window (CW) adaptation scheme for mobile consumer devices using VoIP. The AP's CW size is dynamically changed according to the results of a stereotype-based adaptation. Performance of the proposed solution was evaluated in terms of two metrics: VoIP capacity and Jain's fairness index. Notably, the proposed scheme can also be applied for other interactive multimedia services (i.e. video conferencing) where downlink and uplink fairness is required. In chapter 10, the simulation results will show how the proposed algorithm improves the downlink/uplink fairness in comparison with sate of the art solution like 802.11 and dynamic CW.

CHAPTER 8

Experimental Evaluation of the Proposed Bandwidth Estimation Schemes

This chapter presents the experimental test scenarios and result analysis for the proposed bandwidth estimation schemes: intelligent Bandwidth Estimation (iBE) and Model-based Bandwidth Estimation (MBE). iBE tests were introduced under simulation-based test-bed. iBE performance was investigated by comparing with an existing bandwidth estimation scheme, Spruce. MBE performance was studied both via simulations and real life prototype-based tests. The robustness of MBE was investigated in terms of the impact of feedback frequency, packet size, packet error rate, and wireless link adaptation. Three scenarios were designed to assess the MBE performance in terms of error rate, overhead and loss. Additionally, three existing bandwidth estimation techniques were selected for comparison.

8.1 iBE Testing

8.1.1 Simulation Test-bed Setup

The principle mechanism of the proposed intelligent bandwidth estimation algorithm (iBE) [188] is to make use of the differences between the packet's transmission time and reception time at MAC layer. iBE has been modelled and evaluated using Network Simulator-2 (NS-2) version 2.29[49]. Figure 8-1 shows the simulation topology where servers sent multimedia and cross traffic to clients via a wired network as well as a last hop Wireless LAN (WLAN) implementing IEEE 802.11b. In the experiment, it was assumed that IEEE 802.11b WLAN was the bottleneck link on the end-to-end path. The multimedia traffic delivered is *The*

[49] NS-2-http://www.isi.edu/nsnam/ns/

Simpsons movie and the related trace file can be obtained from TKU website[50]. The video sequence is encoded using MPEG-4 with bit-rates of 1300Kbps and frame rates of 30frames/second. Video traffic and background traffic shared the bottleneck from the 802.11 access point (AP) to the wireless clients. CBR/UDP flows were introduced as the background traffic to vary the overall network load. The background traffic sent packets of 1500 bytes with bit-rates between 500Kbps and 1Mbps. Additionally, iBE was tested by comparing with the bandwidth estimation scheme, Spruce [93], which provides good guidelines for implementation and has been used widely. Table 8-1 summarizes the parameters used in NS-2. Two additional wireless update patches were deployed in the set-up: No Ad-Hoc (NOAH)[51] and Marco Fiero patch[52]. NOAH was used in order to allow direct communication between mobile users and the AP only. Marco Fiero's patch provided a realistic wireless network environment by adding realistic channel propagation, multi-rate transmission support and Adaptive Auto Rate Fallback (AARF) [227]. According to the default configuration of the access point, channel 7 was selected. The interference only comes from background traffic in the same channel. The Shadowing model is used by NS-2 as the channel model, which has been widely used to mimic the shadowing effect caused by obstacles.

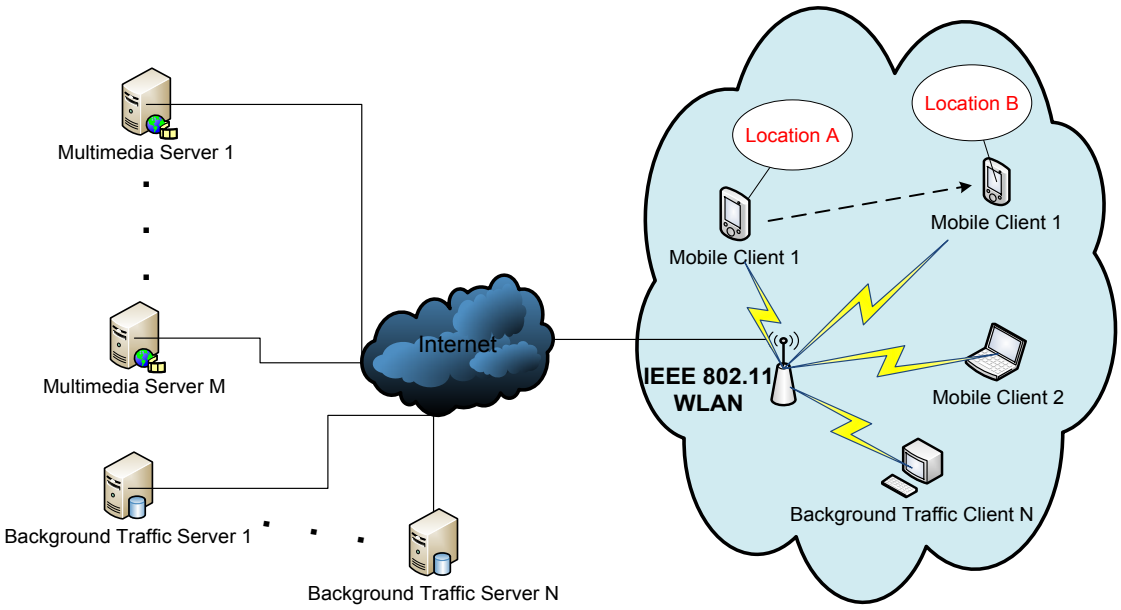


Figure 8-1 Simulation network topology

[50] Video sequence trace-http://www2.tkn.tu-berlin.de/research/trace/pics/FrameTrace/mp4/index28cd.html
[51] No Ad-hoc NS-2 extension, http://icapeople.epfl.ch/widmer/uwb/ns-2/noah/
[52] M. Fiore patch, http://www.telematica.polito.it/fiore/ns2_wireless_update_patch.tgz

TABLE 8-1 SIMULATION SETUP IN NS-2.29

Transport Protocol	UDP
Wireless protocol	802.11b
Routing protocol	NOAH
Error Model	Marco Fiero patch
Channel Model	Shadowing model
Wired Bandwidth	100 Mbps LAN
MAC header	52 bytes
W_{min}	31
W_{max}	1023
ACK	38 bytes
CTS	38 bytes
RTS	44 bytes
SIFS	10 μsec
DIFS	50 μsec
Basic rate	1 Mbps

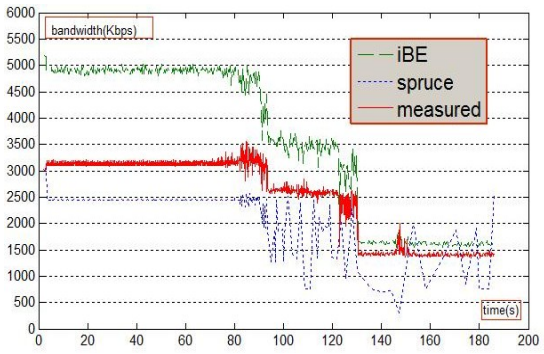
8.1.2 Scenarios

Six test cases were designed to evaluate iBE under different network conditions. All tests last 200 seconds. The setup of each test case was as follows: 1) in test case one, one multimedia server delivered the video sequence to one mobile client; 2) in test case two, two multimedia servers sent video sequences to two mobile clients, separately; 3) in test case three, two multimedia servers started delivering video at 2s and 30s, and one background traffic flow (500Kbps) started at 50s; 4) in test case four, two multimedia servers sent video sequences as in test case three and two background traffic flows (500Kbps and 1Mbps) started at 50s and 70s, separately; 5) test case five was based on test case four, and one additional background traffic flow (1Mbps) was added which starts at 80s; 6) in test case six, two multimedia servers delivered video to two mobile clients and three background traffic flows (each with 1Mbps) started at 50s. In all the six test cases, the mobile clients moved at 5s at the speed of 1 m/s (typical walking speed).

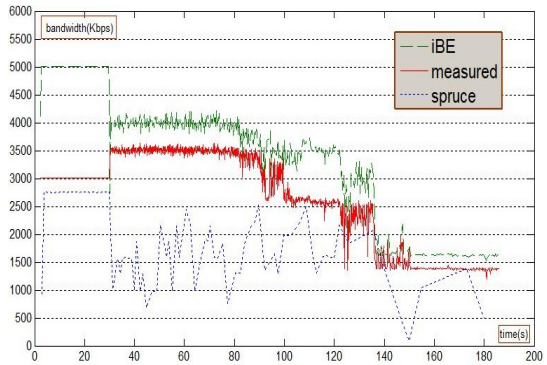
8.1.3 Performance Evaluation and Results Analysis

Figure 8-2 shows the comparison results between measured bandwidth (obtained from simulation) and estimated bandwidth based on iBE and Spruce, separately. Figure 8-2 (a) and Figure 8-2 (b) present the results of test case 1 and test case 2, separately.

In test case one, as shown in Figure 8-2 (a), it can be observed that the estimated available bandwidth drops when the client moves far away from AP, as the capacity of the wireless network decreases. Specifically, the bandwidth fluctuates at around t=80s and t=130s due to interference caused by the incoming background traffic. The average bandwidth estimated by iBE and Spruce are 3.52 Mbps and 1.51 Mbps respectively, both of which are different from the measured bandwidth value of 2.96 Mbps. The average bandwidth is estimated over an interval of 500ms, which is the default value of the beacon interval of most 802.11 access points. Synchronizing the bandwidth estimation interval with AP's beacon interval has been used in many previous works [93], [77]. However, in comparison with the measured bandwidth, iBE has lower errors (0.56) than that of Spruce (1.45). Testing results of test case two are shown in Figure 8-2 (b). In the case of iBE, the average difference between the estimated bandwidth and the measured bandwidth is 0.29, while in the case of Spruce, the same value is 1.63. Since test case two has higher network load than test case one, it can be concluded that iBE outperforms Spruce when two video flows are transmitted. Additionally, for test case one and test case two, iBE provides smoother estimated bandwidth, such as: 1) t=0 to t=80s, t=100s to t=120s and t=130s to t=200s in Figure 8-2 (a); 2) t=30s to t=90s and t=100s to t=120s in Figure 8-2 (b).



(a)

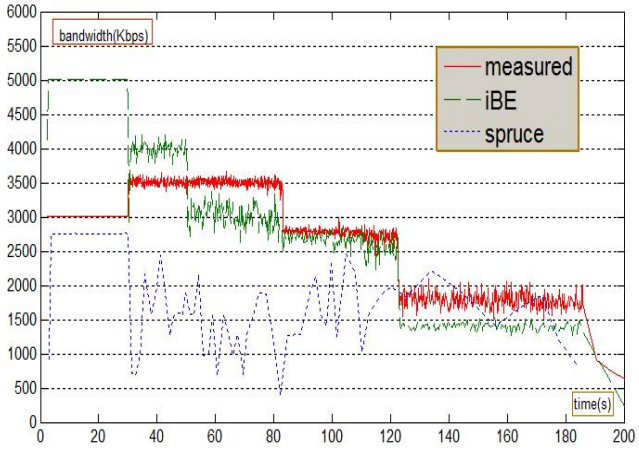


(b)

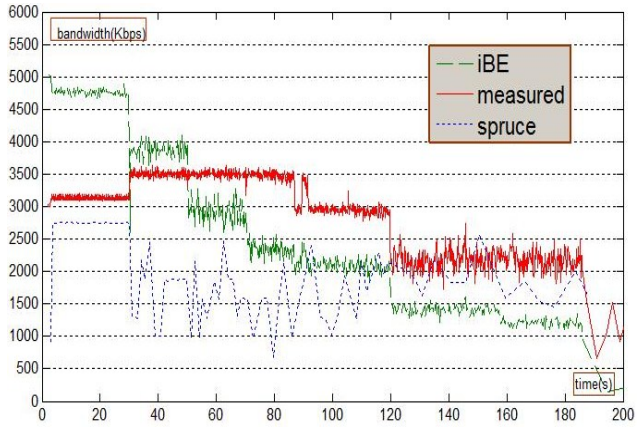
Figure 8-2 Comparison of estimated and measured bandwidth without cross traffic

Figure 8-3 presents the comparison results between iBE and Spruce when cross traffic is introduced to the wireless network. Figure 8-3 (a), (b), (c), and (d) present the results of test case three, four, five, and six, separately.

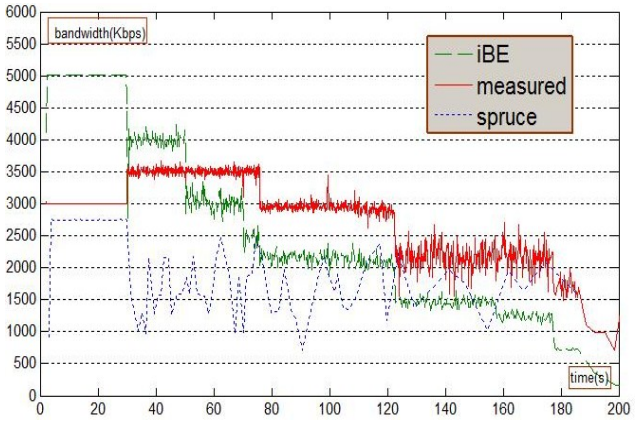
Table 8-2 presents the six test case results in detail as well as the average bandwidth based on iBE, Spruce and the actual measured bandwidth. The error column shows the relative deviation from the actual measured bandwidth. For instance, in test case four, bandwidth error of iBE is 0.38 Mbps whereas that of Spruce is 1.05 Mbps. It can be observed from Figure 8-3 (a)-(d) that iBE is significantly closer to the measured bandwidth than that of Spruce. For instance, in comparison with Spruce, the average difference between the bandwidth measured and estimated by iBE is lower by 95.5%, 63.8%, 74.2%, and 88.2%, in test case three, four, five, and six, separately. Figure 8-4 further shows that the average bandwidth estimated by iBE is closer to the measurement values than Spruce, for all the six test cases.



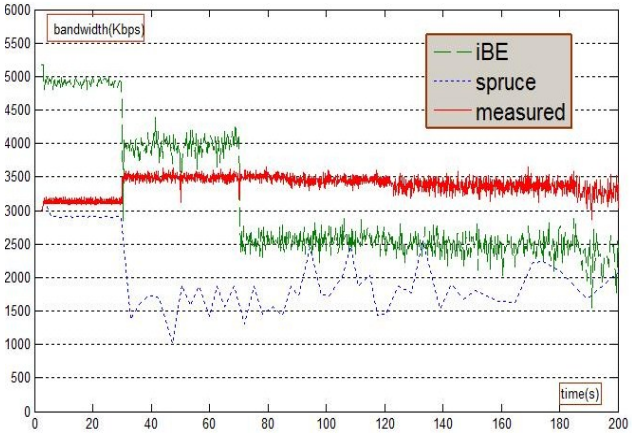
(a)



(b)



(c)



(d)

Figure 8-3 Comparison of estimated and measured bandwidth with cross traffic

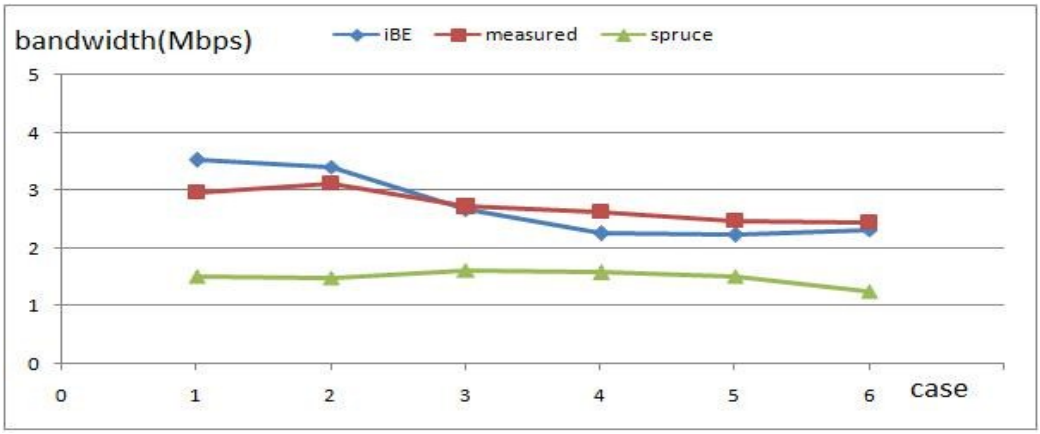


Figure 8-4 Average bandwidth between estimation and real measurement for six test cases

TABLE 8-2 BANDWIDTH ESTIMATION FOR THE SIX EXPERIMENTS

Test case	Video Clients	Cross Traffic	Bandwidth (Mbps)				
			Measured	iBE	Spruce	Error	
						iBE	Spruce
1	1	None	2.96	3.52	1.51	18.9%	49%
2	2	None	3.12	3.41	1.49	9.3%	52.2%
3	2	CBR/UDP 0.5Mb/s	2.72	2.67	1.62	1.8%	40.4%
4	2	CBR/UDP 0.5Mb/s CBR/UDP 1.0Mb/s	2.63	2.25	1.58	14.4%	39.9%
5	2	CBR/UDP 0.5Mb/s CBR/UDP 1.0Mb/s CBR/UDP 1.0Mb/s	2.48	2.23	1.51	10.1%	39.1%
6	2	CBR/UDP 1.0Mb/s CBR/UDP 1.0Mb/s CBR/UDP 1.0Mb/s	2.45	2.31	1.26	5.7%	48.6%

8.1.4 Limitations of the iBE Testing Results

Although iBE shows good bandwidth estimation performance in comparison with *Spruce* and *real measurement*, there are two limitations should be concerned. First, the background traffic consisted of UDP traffic only. TCP background traffic should also be included since TCP has flow control mechanism which affects the network load in a different way as UDP. Second, the error-prone feature of wireless channel was not considered. The bandwidth estimation performance might affected by high packet error rate of the wireless network.

8.2 Experimental Test for MBE

MBE [189] estimates the available bandwidth uses the TCP and UDP throughput model, which takes the network condition-related parameters as input, i.e., packet loss, round trip time, the number of contending wireless stations.

This section introduces the experimental setup and results analysis for MBE: 1) experimental setup including description of existing bandwidth estimation schemes MBE was compared against, evaluation metrics, and experimental scenarios; 2) evaluation of the robustness of MBE including the impact of feedback frequency, packet size, packet error rate, and wireless link adaptation; 3) evaluation of bandwidth estimation performance in terms of error, loss, and overhead.

8.2.1 Experimental Setup

MBE has been evaluated by using both ***modelling*** and ***prototyping***. Modelling and simulations were performed by employing the NS-2.33 network simulator. Prototyping and real life test involved the Candela Technologies' LANForge traffic generator V4.9.9-based network test bed. There are two assumptions for both simulation and real life tests. First of all, the application and hardware processing delays were assumed to be negligible. This is reasonable because the IP packet processing delay in terminals depends on CPU and memory specifications and these are state-of-the-art in our setup. This delay is very low and is in general negligible. Secondly, the IEEE 802.11 WLAN was assumed to be the bottleneck link. This was supported by connecting the IEEE 802.11 WLAN with a 100Mbps wired LAN. In this condition, the bandwidth estimation can closely reflect the wireless network capacity.

8.2.1.1 Simulation Test-bed Setup

Figure 8-5 (a) illustrates the wired-cum-wireless "**dumbbell**" network topology used in NS-2.33. Multiple wireless clients communicated with the servers via an IEEE 802.11b access point (AP). Each traffic connection consists of one server-wireless station pair. The wired link between the AP and server was set to 100Mbps with 2ms propagation delay. Two additional wireless update patches are deployed in the NS-2 set-up: NOAH and Marco Fiore patch. NOAH (No Ad-Hoc) was used for simulating the infrastructure WLAN environment, whereas Marco Fiore's patch provides a more realistic wireless network environment. The IEEE 802.11b protocol is configured according to the specification [17], as shown in Table 8-3, where *SIFS*=10μs, *PIFS*=30μs, *DIFS*=50μs, *slot time*=20μs, *PLCP preamble* =24bytes, *basic rate*=1Mbps, *minimum contention window (CWmin)*=31, *maximum contention window (CWmax)*=1023. *TCP/IP protocol header* = 40bytes, *UDP/IP protocol header* = 28bytes and *MAC protocol header* = 36bytes. The wireless access mode RTS/CTS was enabled to avoid the wireless hidden node problem. DropTail [225] was adopted as the default queue algorithm and the queue length was set to 50 in simulator. File Transfer Protocol (FTP) was used as application traffic over TCP which intended to use the entire wireless capacity. FTP divides a file into small parts and delivers them to a destination host. NS2 FTP module does not require an input file; instead, it informs an attached transport agent (TCP in this test) of file size in bytes. Constant Bit Rate application was used to run over UDP transport protocol. The length of TCP and UDP packet size were set to 1380 bytes and 1000 bytes.

TABLE 8-3 SIMULATION SETUP PARAMETERS IN NS-2.33

Experimental Input Parameters	Values
Queue	DropTail
Queue buffer	50 packets
Basic rate	1Mbps
Minimum contention window (CW_{min})	31
Maximum contention window (CW_{max})	1023
DIFS	50μs
SIFS	10μs
Slot time	20μs
TCP/IP header	40bytes
UDP/IP header	28bytes
MAC header	36bytes

8.2.1.2 Real Life Test-bed Setup

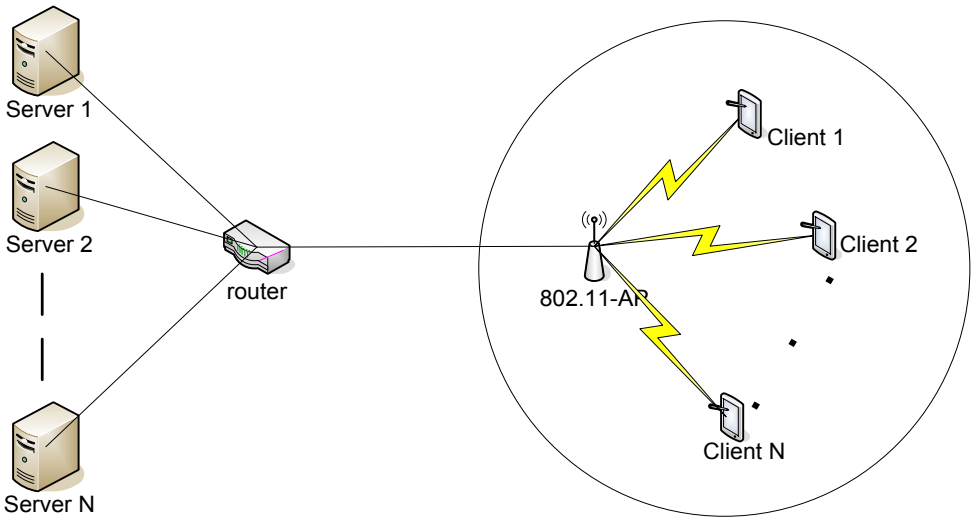
In the prototype-based test bed, as shown in Figure 8-5 (b), the LANForge traffic generator acts as a server which generates traffic transmitted via a 100Mbps Ethernet and a Linksys WRV210 access point to multiple virtual wireless stations. Multiple virtual servers were created to transmit FTP and UDP traffic to multiple virtual clients. The transmission power of AP is 20dBm through two omni-directional antennas. The wireless access mode RTS/CTS was enabled to avoid the wireless hidden node problem. The buffer at both server and client were set to 8K bytes. The length of TCP and UDP packet size were set to 1380 bytes and 1000 bytes. The measured bandwidth was obtained using the test bed as in Figure 8-5 (b). The measured bandwidth was obtained using a *LANForge Traffic Generator* and the *Wireshark*. The traffic generator sent variable numbers of TCP and UDP flows to multiple virtual wireless stations. The overall network bandwidth was then captured and analysed using *Wireshark*.

8.2.1.3 Other Bandwidth Estimation Techniques

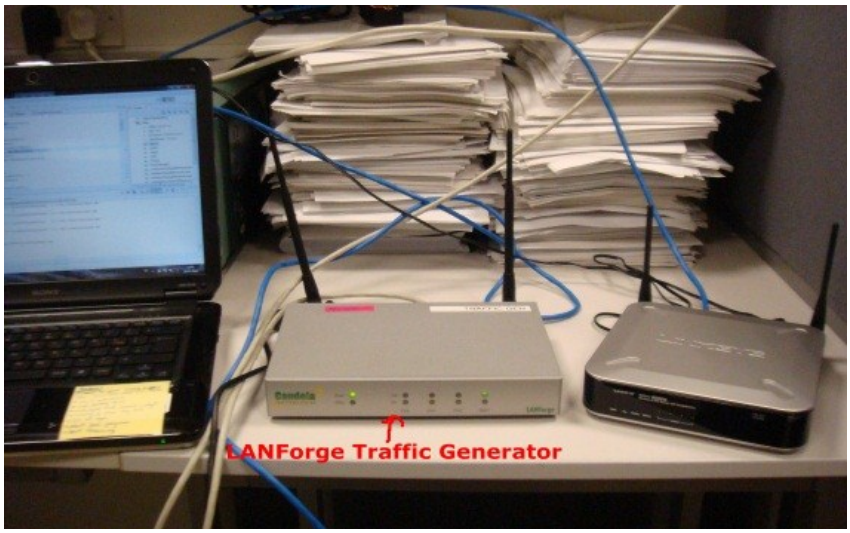
Three state-of-the-art bandwidth estimation schemes, which employ different types of techniques, were selected for comparison. The three techniques include the previously proposed non-probing technique-*iBE* [188], probing-based technique-*DietTOPP* [91], and cross-layer technique-*IdleGap* [96]. The principle of selecting these techiques followed two rules: 1) wide deployment; 2) availability of implementation details.

iBE was implemented at the 802.11 MAC layer. The 802.11 WLAN was assumed to be the bottleneck link in the end-to-end path. The feedback frequency of *iBE* client was set to

10ms as indicated in [188]. *RTS/CTS* function was enabled to achieve best performance of *iBE* in all conditions.



(a) Modeling and simulation test bed topology



(b) Real test bed including traffic generator and 802.11AP

Figure 8-5 Simulation and real life test-bed setup

DietTOPP relies on probe packet size and cross-traffic, with the condition that the wireless link is the bottleneck in the end-to-end path. Hence, 1500 bytes probing packets and 250Kbps cross-traffic were used to obtain better estimation of performance as indicated in [91]. *DietTOPP* was implemented real life test bed in C++ on Unix platform and is available online[53].

[53] A. Johnsson, "Diettopp implementation, beta version,"-http://www.idt.mdh.se/ãjn12/.

The *IdleGap* cross layer algorithm was implemented between the 802.11 link layer and network layer. The cross-traffic for *IdleGap* was set to 10Kbps as suggested in [96]. Application packet size was set to 700 bytes since *IdleGap* achieved good accuracy for packet size ranges from 512 bytes to 896 bytes. *RTS/CTS* function was also enabled. *IdleGap* was implemented under simulation test-bed using NS-2.

8.2.1.4 Evaluation Metrics

In order to evaluate the *MBE* performance, two estimation-based evaluation metrics were introduced: *error rate* and *overhead*. Error rate is defined as the difference (in percentage) between the estimation results and the measured result. Lower error rates indicate higher accuracy of bandwidth estimation. The error calculation is given in equation (8-1).

$$ErrorRate = \frac{|EstimatedBandwidth - MeasuredBandwidth|}{MeasuredBandwidth} \times 100\% \quad (8\text{-}1)$$

Overhead is defined as a ratio between the amount of overhead using bandwidth estimation algorithms and the amount of estimated bandwidth. For instance, the overhead load caused by *MBE* and *IdleGap* mainly come from the cross layer feedback traffic, the overhead using *DietTOPP* is caused by the probing traffic, and the overhead of using *iBE* started when receiving feedback from client. Lower overhead is critical for streaming applications over wireless networks, as they already put pressure on available bandwidth resources. The overhead computation is given in equation (8-2).

$$Overhead = \frac{OverheadLoad}{EstimatedBandwidth} \times 100\% \quad (8\text{-}2)$$

8.2.1.5 Scenarios

Two experiments were designed to study the performance of *MBE*. Their goals are as follows: 1) evaluate the robustness of *MBE* model; 2) evaluate bandwidth estimation quality.

The evaluation of robustness of MBE was designed as follows: 1) The impact of feedback frequency was studied using simulation test-bed, as the real life test did not involve the feedback traffic; 2) The impact of packet size was studied using both simulation and real life tests; 3) The impact of packet error rate was studied using both simulation and real life tests; 4) The impact of wireless link adaptation was studied using simulation test-bed, as the effects of wireless channel adaptation cannot be captured in the real life test-bed.

Additionally, the bandwidth estimation quality was studied by comparing *MBE* with *iBE*, *DietTOPP*, and *IdleGap*. *iBE*, *DietTOPP*, and *IdleGap* were implemented in the simulation test-bed. Error rate and overhead were used as the evaluation metrics, which have been described previously.

8.2.2 Evaluation of MBE Robustness

To study the MBE robustness in variable wireless network environment, separate tests were performed to study the impact of feedback frequency, packet size, packet error rate, and wireless link adaptation. For each test scenario, the variable-controlling method was adopted. Each scenario included a specific experimental setup which was based on the test bed described in section 8.2.1.

8.2.2.1 Impact of Feedback Frequency on MBE Performance

The purpose of this test was to investigate the impact of feedback traffic introduced by *MBE* and enabled to select a good feedback frequency for future tests. Too frequent feedback causes high overhead which reduces the performance of the multimedia traffic while too little feedback decreases the accuracy of estimated bandwidth. *MBE* uses RTCP Receiver Report to deliver the feedback (8 bytes-RTCP receiver report packet header, 8 bytes-UDP header, 20 bytes-IP header, and 4 bytes- feedback payload) due to the low cost and high reliability of this approach. Since the feedback size and the number of flows are relative static, the bandwidth taken by feedback relies on the inter-feedback interval. Most of the time, RTCP traffic uses UDP as the underlying transport protocol, so the single feedback packet size can be written as shown in equation (8-3). However, TCP can also be employed as transport layer protocol.

$$FeedbackSize = RTCPheader + UDPheader + IPheader + Payload \quad (8\text{-}3)$$

The value of feedback size is 40 bytes. Consequently, the feedback rate for each flow is given in equation (8-4).

$$FeedbackRate = FeedbackSize / FeedbackInterval \quad (8\text{-}4)$$

When the number of flows is N and the sampling time duration is T, the overhead load can be computed by equation (8-5).

$$OverheadLoad = FeedbackRate \times T \times N \quad (8\text{-}5)$$

Experimental Setup: **Simulation** tests were performed to study the impact of feedback frequency. As shown in Figure 8-5 (a), one server node started sending single 6Mbps CBR/UDP traffic with packet size set to 1000 bytes. Packet Error Rate (PER) was set to $1x10^{-5}$. One mobile node stayed close to AP at a distance smaller than 10m where the link data rate is 11Mbps. The duration of the experiment was 100s. The feedback interval was varied from 0.001s to 10.0s.

Experimental Result Analysis: Let α represent the ratio between the feedback rate and the channel bandwidth. *MBE* performance-related parameters in terms of mean estimation error rate, overhead and α are shown in Table 8-4. The RTCP standard recommends that the value of α has to account for less than 5% of the bandwidth in order to optimize the received quality of the application. By analyzing the results, the overhead introduced by *MBE* increases as the decrease in the inter-feedback interval. The mean error rate was affected by different inter-feedback interval. For instance, in the case the feedback interval is 1ms, the estimation overhead was 64Mb during 100s. This represents approximate 6.3% of the overall bandwidth and the relative mean error was as high as 31%. This suggests that too many feedback packets competing with application data packets might have caused higher packet loss.

Consequently, high packet loss ratio reduced the MBE bandwidth estimation accuracy and increased the estimation error. Subsequently, the optimal feedback frequency is selected by looking at Table 8-4. A good trade-off between the amount of overhead and mean error rate recommends an inter-feedback interval of 1.0s. This value will be used in the following tests.

TABLE 8-4 MEAN ESTIMATION ERROR, OVERHEAD AND A DEPENDENCY ON THE FEEDBACK INTERVAL. TIME DURATION=100S

Feedback interval (s)	Mean Error Rate (%)	OverheadLoad (MB)	α (%)
0.001	31	64	6.3
0.005	24	12.8	3.2
0.01	17	6.4	2.1
0.1	12	0.64	0.5
0.5	8	0.128	0.04
1.0	4	0.064	0.007
2.0	11	0.032	0.001
4.0	15	0.016	0.0005
6.0	19	0.0106	0.00009
8.0	22	0.008	0.00002
10.0	23	0.0064	0.000006

TABLE 8-5 EFFECT OF PACKET SIZE ON THE BANDWIDTH ESTIMATED BY MBE

Packet size (bytes)	MBE (Mbps)	Simulations (Mbps)	Real tests (Mbps)	Error rate (%)
100	0.82	0.84	0.92	10.9%
300	2.11	2.18	2.21	4.5%
500	3.19	3.21	3.25	1.8%
700	3.97	4.03	4.09	2.9%
900	4.61	4.68	4.71	2.1%
1000	4.9	4.95	4.97	1.4%
1100	4.81	4.89	4.93	2.4%
1300	4.73	4.7	4.82	1.95
1500	4.55	4.62	4.67	2.6%

8.2.2.2 Impact of Packet Size on MBE Performance

This section investigates the impact of packet size on the *MBE* estimation accuracy. The inter-feedback frequency suggested in section 8.2.2.1 was used in this test.

Experimental Setup: Both **simulation** and **real test** experiments were performed to study the packet size effects. A single CBR/UDP traffic with an average rate of 6Mbps was sent from the server to a mobile station via the 802.11b AP. The packet size was varied from 100 bytes to 1500 bytes (Ethernet MTU) with a step of 200 bytes every 20s. Feedback frequency was set to 1.0s. It was noticed that 6Mbps traffic was used to saturate the network so that the effect of packet size will be studied in a loaded network. The mobile node was placed close to AP at a distance less than 10m where the link data rate was 11Mbps. The experiment time duration was set to 160s.

Experimental Result Analysis: The estimation and measurement results of the packet size study are shown in Table 8-5 and further illustrated in Figure 8-6. It is shown that the available bandwidth increases along with the increase of packet size. Since smaller packet size leads to more frequent transmissions and higher packet overhead. Throughput is the highest when packet size is 1000 bytes, as 1000 bytes was the fragmentation threshold. Any packets larger than 1000 bytes get fragmented into multiple packets resulting a decrease in throughput. According to Table 8-5, following a two tailed T-test analysis it can be said with 95% confidence level that there is no statistical difference between the MBE results and

those of the real test. It can be concluded that *MBE* can provide high accurate estimated bandwidth with variable packet size.

8.2.2.3 Impact of Packet Error Rate on MBE Performance

In contrast with wired communications, wireless networks suffer from environmental noise e.g. building block interference or terminal generated noise e.g. thermal noise. These affect the communications and decrease the estimation accuracy. The purpose of this section was to study the performance of *MBE* when data transmission is affected over network with various PER.

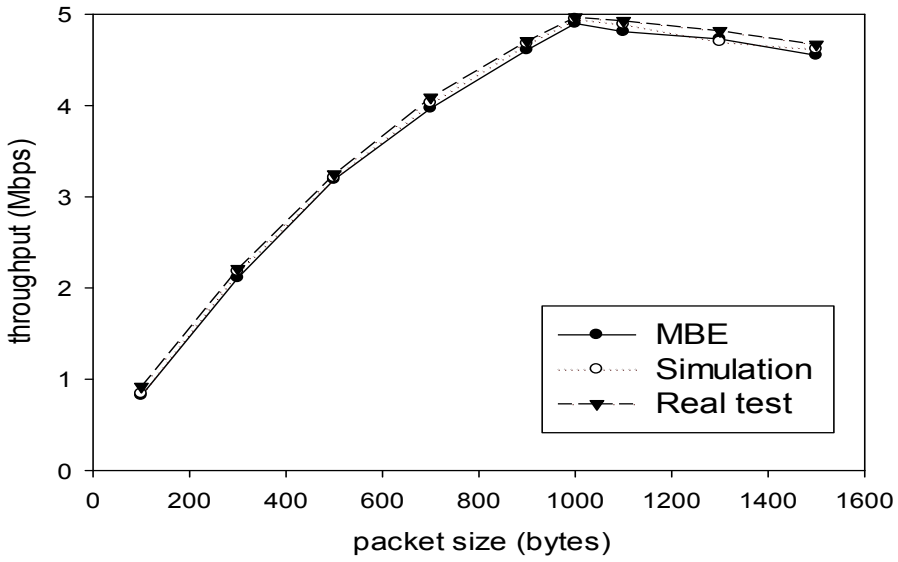


Figure 8-6 Comparison of bandwidth as estimated by MBE, measured by NS-2 simulations and obtained from the real-life tests for increasing packet size

Experimental Setup: The impact of *PER* was investigated in both **simulation** and **real test** environments. Similar with the previous tests, this experiment also transmitted a single CBR/UDP flow with packet size of 1000 bytes. Feedback frequency was set to 1.0s. NS-2 simulation tool provided functions to increase the *PER* from $1x10^{-8}$ to 1. For each given PER, there was a corresponding average packet loss ratio which was then imported to the *MBE* model to estimate the available bandwidth. In real life test, it is difficult to inject packet error into the wireless channel. An alternative solution is to adjust the AP transmitting power to mimic the effect of PER. As shown in Figure 8-5(b), we added the *Pascall*[54] signal manual attenuator between the Linksys WRV210 AP and an external N-type antenna. Since the maximum transmission power of AP is 20dBm, the attenuator gradually reduced the transmitting power with a 2dBm step. For both simulation and real test, the mobile node was

[54] http://www.pascall.co.uk

fixed close to AP at a distance smaller than 10m where the link data rate was 11Mbps. Experimental time duration was set to 100s. The bandwidth estimated by MBE is given based on the packet loss information under different simulation and real test conditions.

TABLE 8-6 EFFECT OF DIFFERENT PER ON THE BANDWIDTH ESTIMATED BY MBE, SIMULATION AND REAL-LIFE TEST. PER VARIATION IS SIMULATED BY ADAPTING TRANSMITTING POWER

Simulations			Real tests		
PER	**MBE (Mbps)**	**Measured (Mbps)**	**Transmitting Power (dBm)**	**MBE (Mbps)**	**Measured (Mbps)**
$1x10^{-8}$	5.45	5.15	20	5.31	5.49
$1x10^{-7}$	5.45	5.15	18	5.30	5.38
$1x10^{-6}$	5.45	5.15	16	5.30	5.38
$1x10^{-5}$	5.34	5.11	14	5.26	5.31
$1x10^{-4}$	5.23	5.01	12	4.11	4.23
$1x10^{-3}$	4.94	4.74	10	4.02	4.19
$1x10^{-2}$	3.96	4.52	8	2.15	3.23
$1x10^{-1}$	1.26	2.7	6	1.56	2.61
1	0	0	4	1.12	1.67

Experimental Result Analysis: Simulation and real test based results when PER varies are shown in Table 8-6 and are further illustrated in Figure 8-7 and Figure 8-8. It was noticed that the available bandwidth generally decreases along with the increase of PER. The bandwidth equals zero when PER equals one. This implies that unsuccessful transmission will be achieved even with maximum retry limit (number of retry limit =7). Specifically, there was an overestimation of MBE up to PER=10^{-3} and then an underestimation. The overestimation was caused as MBE igored application delay and processing delay, both of which were included in the simulation measurement. The bandwidth was underestimated when PER$\geq 10^{-3}$, which indicates that collisions started increasing significantly. Additionally, the available throughput decreased along with the reduction of transmission power. When the transmit power was lower than 10dBm, the throughput started decreasing significantly. This can be explained by that the receiving signal strength might lower than the receiving threshold defined at the AP. According to Table 8-6, the two taile t-test analysis presents with 90% confidence level that there is no statistical difference between MBE results and those of the real test. Hence it could be concluded that *MBE* can provide high accurate estimated bandwidth in network with variable PER.

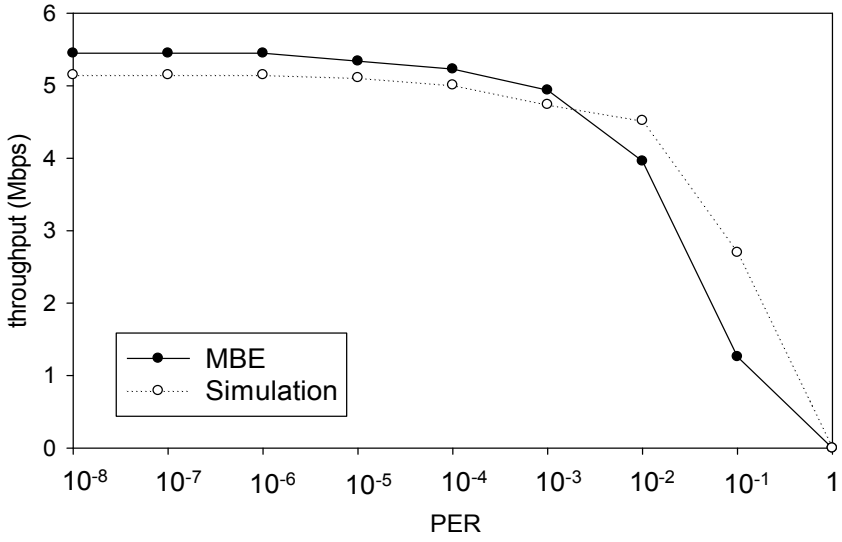


Figure 8-7 PER effect on throughput

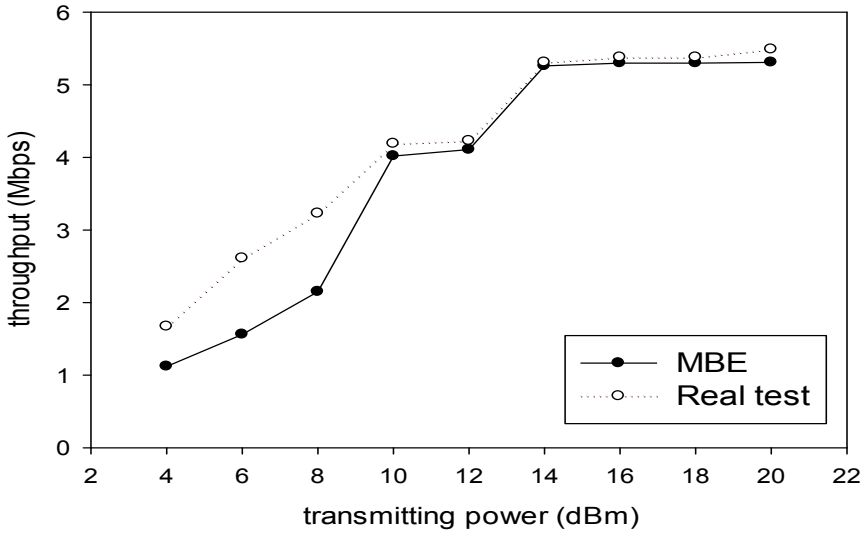


Figure 8-8 Transmitting power effect on throughput

8.2.2.4 Impact of Wireless Link Adaptation on MBE Performance

The goal of this section was to assess the performance of MBE under variable wireless link capacity. Unlike the wired networks, the capacity of wireless networks changes due to the link rate adaptation. The signal coverage of AP was divided into four sub-areas according to the link rate distribution defined in 802.11b, as shown in Figure 8-9. Darker colors indicate higher signal strength.

Experimental Setup: Three test scenarios were implemented in the **simulation** environment to study the impact of the wireless link adaptation on MBE performance. They are: 1) Single mobile nodes located in the areas labeled P1, P2, P3 and P4 in Figure 8-9, respectively. 2) Four mobile nodes evenly distributed around AP. 3) Multiple mobile nodes

located at random locations around AP. These tests used the same test bed as show in Figure 8-5 (a). The differences focused on the mobile node mobility, mobile node location and application traffic. The transmit power of the 802.11b AP in NS2 was set to 20dBm. According to the documentation of the Cisco Linksys WRV210 this can cover around 300 meters. NS2 provided methods to calculate the distance threshold for the signal change: 70m (P1-P2), 100m (P2-P3), and 130m (P3-P4), where P1, P2, P3, and P4 were four positions in each area.

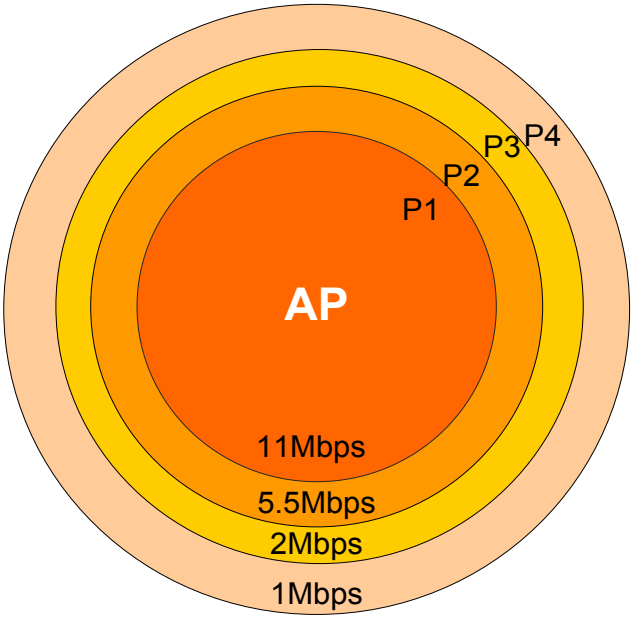


Figure 8-9 Theoretical wireless link capacity for IEEE 802.11b

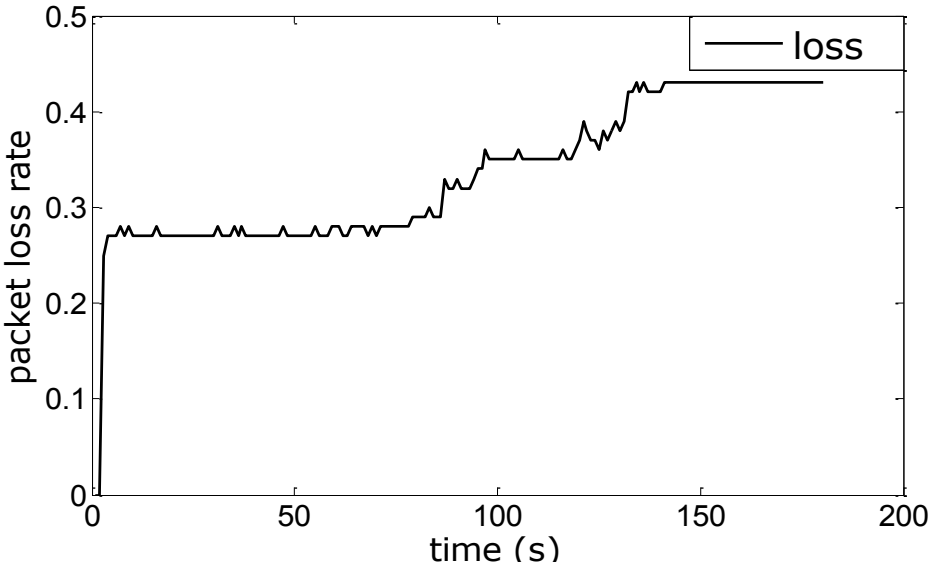


Figure 8-10 Packet loss rate variation while mobile node moves away from AP

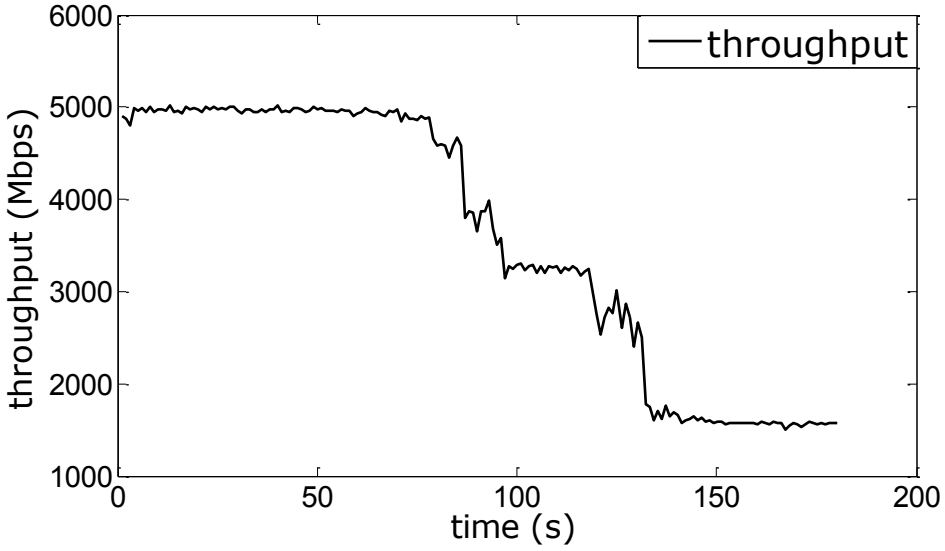


Figure 8-11 Throughput variation while mobile node moves away from AP

1) Single CBR/UDP Traffic to a Node Mobile from P1 to P4

Single CBR/UDP traffic with an average rate of 6Mbps was sent from the server to a mobile station. The mobility considered involved the mobile station moving away from AP towards P4 at the speed of 1m/s. Figure 8-10 and Figure 8-11 show the variations in throughput and packet loss during the transmission. Table 8-7 presents the comparison results between the simulation-based measured throughput and estimated bandwidth from *MBE*.

TABLE 8-7 IMPACT OF DISTANCE FROM AP

	11Mbps	5.5Mbps	2Mbps	1Mbps
Loss (Simulation)	0.27%	0.32%	0.36%	0.43%
Throughput (Simulation)	4.95Mbps	3.11Mbps	2.62Mbps	1.67Mbps
Bandwidth measured by MBE	5.01Mbps	3.08Mbps	2.58Mbps	1.62Mbps

Experimental Result Analysis: It is clear from Figure 8-10 and Figure 8-11 that there is significant packet loss increase and throughput decrease as the mobile node was moving away from AP. This is caused by the reduced transmission signal of AP. The two tailed t-test analysis is applied on the results from Table 8-7. It is shown that there is no statistical difference between MBE estimation results and the measured results under simulation with 95% confidence level.

2) Static Mobile Nodes within the Coverage of AP

FTP/TCP transmissions and 6Mbps CBR/UDP data traffic were considered in this scenario. Three test cases were considered in order to study the *MBE* performance in multiple stations conditions.

- Case 1: Four FTP/TCP flows were sent to four mobile stations and each mobile station was statically located at P1, P2, P3, and P4 respectively.
- Case 2: Four CBR/UDP flows were sent to four mobile stations and each mobile station was statically located at P1, P2, P3, and P4 respectively.
- Case 3: Two FTP/TCP flows were sent from mobile stations located at P1 and P3, and two CBR/UDP flows were transmitted from mobile stations located at P2 and P4.

Experimental Result Analysis: Table 8-8 presents the comparison results between the MBE estimated bandwidth and that measured in the simulation tests for all these three cases. Column "MBE" presents the overall bandwidth estimated by MBE when three test cases are considered. Column "Simulations" provides the overall bandwidth measured in NS-2 for the three test cases, respectively.

According to results of case 1 and case 2 presented in Table 8-8, UDP traffic achieves more throughput than FTP/TCP, since TCP uses the congestion control mechanism to reduce the sending rate when packet loss occurs or increases. Additionally, by comparing results of case 1 and case 2, the throughput of UDP traffic increases 47.9% compared with that of TCP traffic. In case 3, two TCP flows and two UDP flows are transmitted together, the overall throughput is lower than that of four UDP flows (case 2) and higher than that of four TCP flows (case 1). UDP traffic affects TCP traffic due to the aggressive nature on bandwidth cost. The two tailed t-test analysis present with 95% confidence level that there is no statistical difference between the MBE results and simulation results.

TABLE 8-8 IMPACT OF DISTANCE FOR MULTIPLE TCP AND UDP TRAFFIC

	P1	P2	P3	P4	MBE	Simulation
Case 1	1TCP	1TCP	1TCP	1TCP	1.86Mbps	1.99Mbps
Case 2	1UDP	1UDP	1UDP	1UDP	3.57Mbps	3.65Mbps
Case 3	1TCP	1UDP	1TCP	1UDP	2.58Mbps	2.69Mbps

TABLE 8-9 BANDWIDTH COMPARISON BETWEEN MBE AND SIMULATION

λ	MBE (Mbps)	Simulation (Mbps)
1	2.65	2.78
2	3.51	3.63
3	3.48	3.49
4	4.28	4.37
5	3.84	3.95

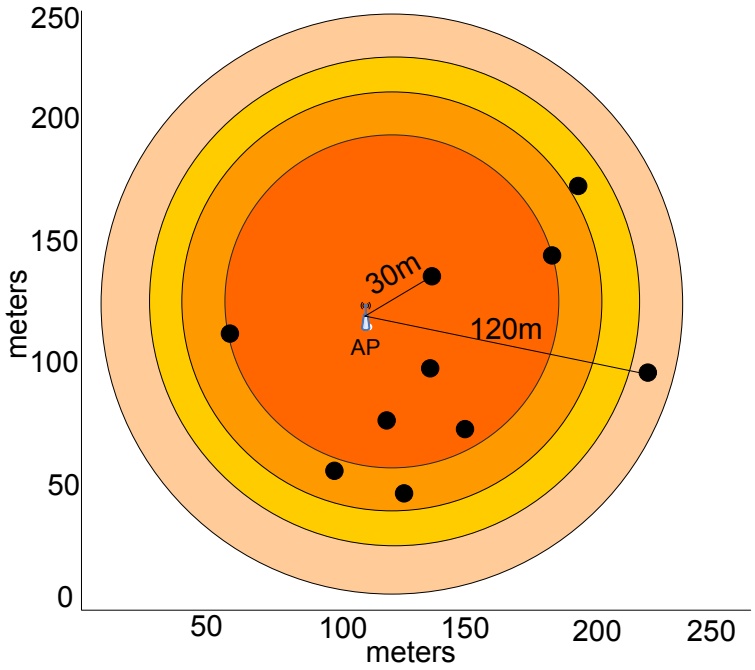


Figure 8-12 Random topology used in simulations.

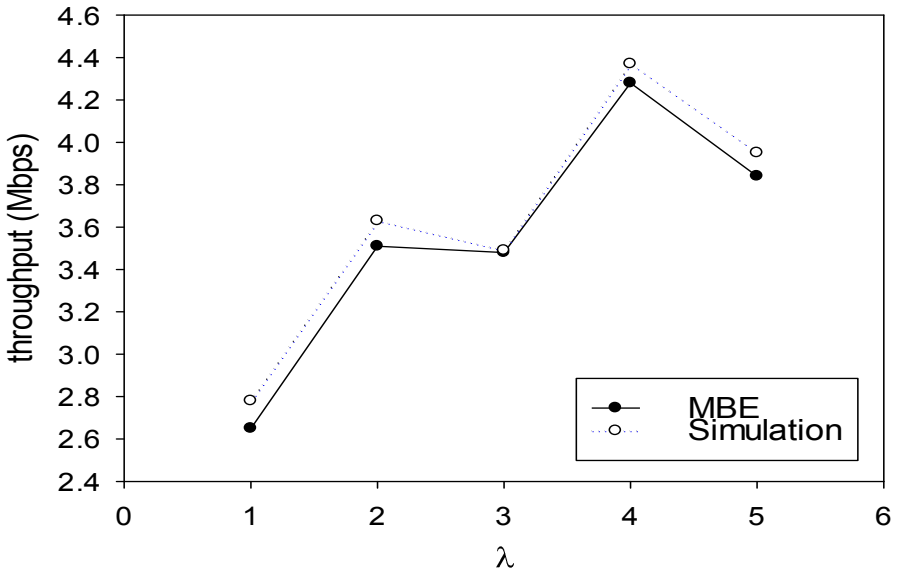


Figure 8-13 Bandwidth comparison between MBE and simulation when λ increases from 1 to 5

3) Multiple Mobile Nodes at Random Positions Within the Coverage of AP

In this scenario, FTP/TCP and 6Mbps CBR/UDP are sent. A 250m x 250m test topology was created in the simulation, as shown in Fig. 12. The position of AP is constant and wireless stations are located around AP with a random distance ranging from 30m to 120m. The number of TCP and UDP flows both equal λ which increases from 1 to 5. Hence the total number of contending stations ranges from 2 to 10, in steps of 2.

Experimental Result Analysis: The mean aggregate throughput was measured through simulation for the mobile nodes with random location. Table 8-9 and Figure 8-13 give the comparison results between the simulation-based measured throughput and estimated bandwidth from *MBE*. For λ smaller than 4, both estimated bandwidth and measured bandwidth increase with increasing number of flows, and the bandwidth starts decreasing when λ equals 5. The overall throughput of the application traffic close to the wireless capacity for λ equals 4, where the number of TCP and UDP flows was 8. The two tailed t-test analysis is used and shows a 95% confidence level that there is no statistical difference between *MBE* results and the simulation results.

Based on the test results from Table 8-7, Table 8-8, and Table 8-9, it is concluded that *MBE* shows high accurate estimated bandwidth with the variable wireless link capacity. This can be explained that the packet loss caused by the wireless link adaptation is used by *MBE* to infer the available bandwidth.

8.2.3 Evaluation of Bandwidth Estimation Performance

Three scenarios were designed to assess the *MBE* performance in terms of error rate, overhead and loss. *MBE* analytical model results are compared with simulation and real test results. Additionally the results of other bandwidth estimation techniques such as *iBE*, *DietTOPP*, and *IdleGap* were also considered.

8.2.3.1 Experimental Setup

Experimental Setup: Each scenario included 15 cases with variable number of FTP/TCP and 6Mbps CBR/UDP traffic load. Test case 1 to test case 5 transmitted TCP traffic only, test case 6 to test case 10 transmitted UDP traffic only while test case 11 to test case 15 sent TCP and UDP traffic simultaneously. In order to estimate the maximum bandwidth a network can support, it is necessary to use high traffic load in order to saturate the 802.11 channel. In a saturated network, any new incoming traffic will decrease the overall

throughput since the available throughput is higher than the network capacity. Based on tests scenarios in section 8.2.2, the feedback interval was set to 1.0s, packet size was 1000 Bytes and PER was set to 10^{-5}. The overall sending rate was greater than 6Mbps and less than 7Mbps. The mobile nodes are located close to AP at a distance smaller than 10m where the link data rate is 11Mbps. Testing time duration was 100s.

TABLE 8-10 COMPARISON OF BANDWIDTH ESTIMATED BY IBE, DIETTOPP, IDLEGAP, MBE AND BANDWIDTH MEASURED

Case	N (Number of flows)		iBE (Mbps)	DietTOPP (Mbps)	IdleGap (Mbps)	MBE (Mbps)	Simulation (Mbps)	Real Test (Mbps)
	TCP	UDP						
1	1	0	5.08	5.01	4.85	5.57	4.89	4.97
2	3	0	3.65	4.23	3.83	3.61	3.98	3.66
3	5	0	3.01	3.02	3.24	3.12	3.47	3.17
4	7	0	2.43	2.24	2.50	2.52	2.94	2.56
5	9	0	1.65	1.33	1.72	1.92	2.25	1.95
6	0	1	6.21	5.39	5.61	6.09	5.1	5.8
7	0	3	5.53	4.96	5.15	5.32	5.3	5.3
8	0	5	5.01	4.82	5.02	5.11	5.19	5.21
9	0	7	4.54	4.53	4.89	4.99	5.07	5.03
10	0	9	4.12	4.17	4.68	4.8	4.94	4.91
11	1	1	5.98	5.78	5.01	5.83	4.975	5.28
12	2	2	4.56	4.34	4.32	4.74	4.86	4.61
13	3	3	3.82	3.72	4.21	4.46	4.59	4.51
14	4	4	3.51	3.38	4.13	4.3	4.46	4.45
15	5	5	3.19	2.12	4.08	4.12	4.35	4.31

8.2.3.2 Estimation Error Rate Analysis

The purpose of this section is to study the estimation error rate which reflects the accuracy of MBE estimation. Table 8-10 shows the comparison results between bandwidth estimated and bandwidth measured. Real test and simulation results were obtained in the testing setup described in section 8.2.1.

Experimental Result Analysis: 15 test cases were implemented to study the error rate of *MBE* under variable traffic load. In single flow situations, such as case 6 and case 11, *IdleGap* provides better accuracy than *MBE* in comparison with the results from real test. From test case 1 to test case 5, the number of contending TCP flows increased from 1 to 9, in steps of 2. It is shown that the bandwidth estimated by the four algorithms and the bandwidth measured in simulations and real tests decrease as the overall traffic load increases. For test case 3 which transmits 5 TCP flows, the estimated bandwidth by *MBE* is 3.12 Mbps. Similarly, the impact of UDP traffic was studied, as shown from test case 6 to test case 10.

The number of UDP flows increased from 1 to 9 in steps of 2. Real test results show significant different throughput achieved between TCP and UDP traffic. When the number of TCP and UDP flows increased from 1 to 9 respectively, the throughput of TCP traffic decreased by 60.8% and the throughput of UDP traffic was reduced by 15.3%. The reason is that TCP flow can adjust the transmission rate using congestion control, and thus resulting in the significant throughput changes. Consequently UDP traffic obtains more bandwidth than TCP traffic which leads to unfair channel access. Although the 802.11b has 11Mbps capacity, the maximum aggregated bandwidth highly depends on the traffic conditions. The bandwidth supplied for single TCP and UDP was up to 5.8Mbps and 4.97 Mbps. Test case 11 to test case 15 give the scenario when TCP and UDP coexist sharing the wireless network. Due to the aggressive characteristic of UDP traffic, the total throughput achieved by TCP and UDP was higher comparing to TCP traffic only.

Iit was observed that, among *iBE*, *DietTOPP* and *IdleGap*, *DietTOPP* produced the highest error rate and *IdleGap* achieved the lowest error rate. Additionally, *MBE* achieved 47% less error rate than *IdleGap*. The two tailed t-test analysis shows that there is no significant statistical difference between *MBE* results and real test results with 95% confidence level. It can be concluded that MBE achieves the lowest error rate in comparison with other solutions.

Notice that the simulation and real test measured throughput was slightly higher in most cases. This can be explained by that, *MBE* model assumes that for each packet to be transmitted, the station invokes backoff mechanism and waits for a DIFS period. However, in simulation and real test, the packets might be transmitted immediately without the backoff delay when the channel is sensed idle.

8.2.3.3 Overhead Analysis

Similar with the experimental testing setup in section 8.2.3.2, this section also used 15 cases with different number of FTP/TCP and CBR/UDP flows. Test case 1 to test case 5 transmitted TCP traffic only, test case 6 to test case 10 transmitted UDP traffic only while test case 11 to test case 15 sent TCP and UDP traffic simultaneously. he overhead introduced by *MBE* come from the feedback traffic. Table 8-11 shows the comparison results between *MBE* and other bandwidth estimation techniques in terms of overhead.

Experimental Result Analysis: For all the 15 test cases, the overhead increases with the increasing number of contending flows. Among *iBE*, *DietTOPP* and *IdleGap*, *DietTOPP* introduced the highest average overhead and *iBE* introduced the lowest average

overhead. Because *DietTOPP* sent continuously probing traffic to detect changes of the receiving rate. *MBE* has lower overhead than *iBE*, as it relies on smaller feedback packets. For instance, in test case 5, 10, and 15, *MBE* decreases the overhead by 10.5%, 4.6%, and 28.3%, respectively, in comparison with *iBE*. It should be noted that applications using TCP traffic caused higher overhead than that of using UDP traffic. This might be explained by the fact that TCP acknowledges the successfully transmitted packets and retransmits the lost packets, which introduces additional overhead.

The mean and standard deviation of the error rate and overhead for all the test cases are shown in Table 8-12, and are further illustrated in Figure 8-14 and Figure 8-15, respectively. The results show that, among the existing bandwidth estimation algorithms, *MBE* achieved up to 89% lower standard deviation of *error rate* and up to 81% lower mean *error rate*, in comparison with *DietTOPP* Additionally, Furthermore, *MBE* obtained up to 84.4% lower standard deviation of *overhead*, and up to 79.4% lower mean *overhead*, in comparison with *DietTOPP*.

TABLE 8-11 COMPARISON OF THE OVERHEAD AMONG IBE, DIETTOPP, IDLEGAP, AND MBE

Case	N (Number of flows)		iBE (Mbps)		DietTOPP (Mbps)		IdleGap (Mbps)		MBE (Mbps)	
	TCP	UDP	Overhead Load (KB)	Overhead (%)	Overhead Load (KB)	Overhead (%)	Overhead Load	Overhead (%)	Overhead Load	Overhead (%)
1	1	0	6.1	1.0	118.1	19.0	7.6	1.3	7.2	1.0
2	3	0	20.2	4.4	138.7	26.2	33.7	7.0	21.5	4.7
3	5	0	31.4	8.3	142.2	37.7	60.4	14.8	35.4	9.0
4	7	0	42.6	14.0	151.4	54.0	86.2	27.6	45.2	14.3
5	9	0	58.8	28.5	168.8	101.5	101.2	47.1	61.2	25.5
6	0	1	7.25	0.9	127.5	18.9	7.5	1.1	8.3	1.1
7	0	3	22.5	3.3	161.2	26.0	38.4	6.0	26..4	3.9
8	0	5	32.1	5.1	163.2	27.2	77.5	12.4	41.5	6.5
9	0	7	47.5	8.4	170	30.0	106.2	17.4	52.2	8.4
10	0	9	65.8	12.8	173.1	33.3	113.4	19.4	73.6	12.3
11	1	1	7.8	1.0	151.5	20.9	10.4	1.6	8.8	1.2
12	2	2	26.25	4.6	163.1	30.2	41.2	7.6	23.7	4.0
13	3	3	36.3	7.6	170.4	36.6	81.4	15.4	38.5	7.0
14	4	4	53.8	12.3	171.2	40.5	106.27	20.6	57.7	10.7
15	5	5	68.9	17.3	177.5	67.0	123.9	24.3	63.9	12.4

TABLE 8-12 MEAN AND STANDARD DEVIATION OF ESTIMATION ERROR AND OVERHEAD FOR IBE, DIETTOPP, IDLEGAP, MBE

	iBE		DietTOPP		IdleGap		MBE	
	Mean	**STDEV**	**Mean**	**STDEV**	**Mean**	**STDEV**	**Mean**	**STDEV**
Estimation Error	9.7%	0.16	17.6%	0.26	6.2%	0.08	3.3%	0.03
Overhead	8.6%	7.2	37.9%	21.8	14.9%	12.24	7.8%	3.4

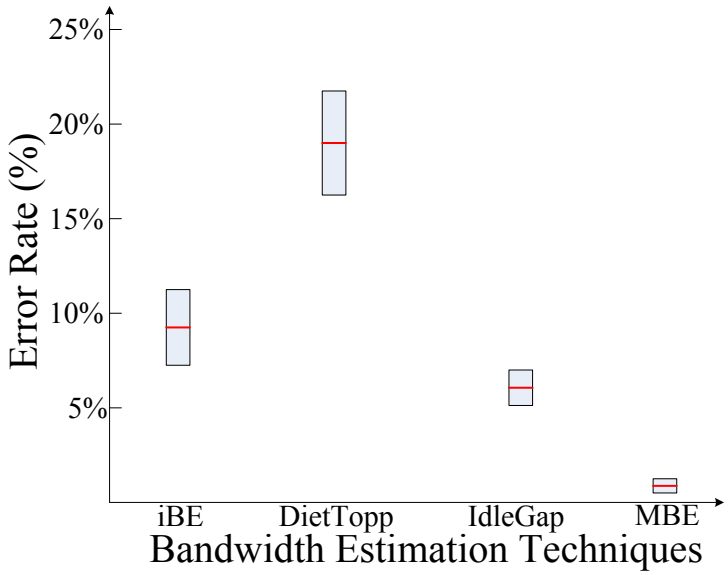


Figure 8-14 Mean value and standard deviation of error rate for iBE, DietTOPP, IdleGap and MBE

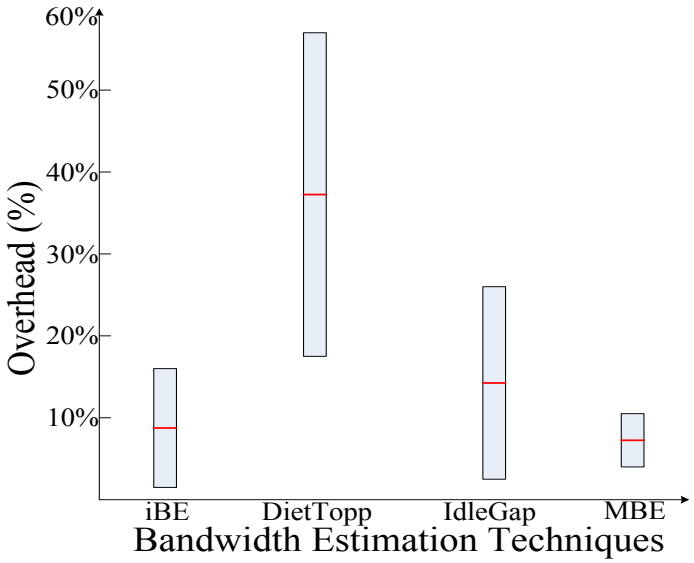


Figure 8-15 Mean and standard deviation of overhead for iBE, DietTOPP, IdleGap and MBE

8.2.3.4 Loss Analysis

The purpose of this section is to study the packet loss rate for different bandwidth estimation schemes. Figure 8-16 shows the results of the packet loss rate evolution with increasing number of 6Mbps CBR/UDP traffic flows when *iBE*, *DietTOPP*, *IdleGap* and *MBE* are used for bandwidth estimation, respectively.

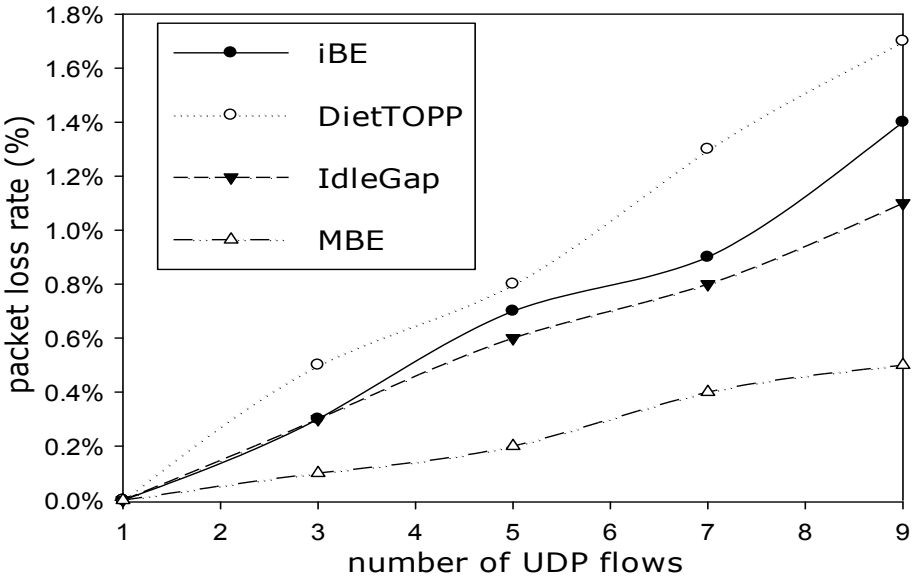


Figure 8-16 Packet loss rate of UDP traffic for iBE, DietTOPP, IdleGap and MBE

Experimental Result Analysis: The number of UDP flows was increased from 1 to 9, and the bandwidth was estimated by the four different bandwidth estimation schemes, *iBE*, *DietTOPP*, *IdleGap* and *MBE*. It is shown in Figure 8-16 that when using *DietTOPP* the highest packet loss rate of up to 1.7% for 9 UDP flows was recorded, as *DietTOPP* continuously sends probing traffic which contends with the UDP traffic. When using *MBE* to estimate the bandwidth, the packet loss rate was the lowest in comparison with all other solutions. For instance for 9 UDP flows when MBE was employed, the loss rate was only 0.4%. It is worth noting that in these conditions, when using *MBE* the packet loss rate decreased with up to 65% in comparison with that of *DietTOPP.* Also, *MBE* reduced packet loss with up to 56% in comparison with that of *iBE* and with up to 50% in comparison with that of *IdleGap.*

8.3 Summary

This thesis proposes a novel Model-based Bandwidth Estimation algorithm (*MBE*) to estimate the available bandwidth for data traffic over 802.11 WLANs. MBE is based on novel throughput models for TCP and UDP traffic over IEEE 802.11 WLANs. In contrast with current wireless bandwidth estimation techniques, *MBE* is fully compatible with the IEEE 802.11 standard protocol, has higher estimation accuracy and introduces lower overhead. *MBE* does not use additional probing traffic which would in turn reduce the already limited bandwidth resources.

iBE was first introduced as a bandwidth estimation scheme by utilizing the delivery of multimedia packets to infer bandwidth of network. iBE shows good bandwidth estimation performance, in comparison with Spruce and the measured bandwidth in IEEE 802.11 network. Additionally, it is shown that under high loaded network conditions, iBE presents significantly better performance than Spruce. For instance when there is background traffic, the difference between the bandwidth measured and estimated by iBE is up to 95.5% lower, in comparison with that of Spruce.

Experimental test results show that the *MBE* model is robust under different conditions: variant packet size, packet error rate and dynamic wireless link. It can be concluded that *MBE* provides the highest accurate bandwidth estimation with the lowest overhead in comparison with existing bandwidth estimation techniques such as *iBE*, *DietTOPP*, and *IdleGap*. Among the three compared techniques, *IdleGap* gives the smallest estimation error rate and *iBE* introduced the lowest overhead. *MBE* achieves 47% lower estimation error rate than *IdleGap* and 9.3% lower overhead than *iBE*. Additionally, *MBE* produces the lowest standard deviation and mean value for both error rate and overhead.

The results of MBE are expected to benefit solutions that provide QoS in wireless networks. Recently, a dynamic wireless resource allocation scheme is proposed in [228] which aims to fairly allocate the underlying network resources. Accurate estimation on available bandwidth is significant for the resource allocation scheme. MBE can also be utilized as part of prioritized adaptive bandwidth allocation scheme without using IEEE 802.11e. For instance, three wireless devices (with high, medium and small screen resolution, respectively) are connected to an IEEE 802.11 AP. Generally, higher resolution device requires more bandwidth than that of lower resolution device, for receiving the same multimedia service. By assigning higher resolution devices with higher priorities, the

bandwidth can be allocated to the three devices based on their priorities and the available bandwidth estimated using MBE. Such prioritized bandwidth allocation can improve the overall QoS distribution in terms of device requirements.

CHAPTER 9

Experimental Evaluation of iPAS

This chapter presents the experimental evaluation for the proposed intelligent Proritized Adaptive Scheme (iPAS). Simulation-based test bed is described and several experimental scenarios are introduced to study the performance of iPAS. These scenarios include the study of fairness, throughput, delay, packet loss rate, perceived video quality, and the device resolution awareness. The evaluation is performed in comparison with equal priority channel access mechanism of IEEE 802.11 DCF and prioritized channel access mechanism of IEEE 802.11e EDCA.

9.1 Simulation-based Experimental Setup

This section describes the simulation-based testing setup including the multimedia traffic characteristics, test-bed configuration, and evaluation metrics used.

9.1.1 Data Traffic

Four types of data traffic: voice, video, best-effort, and background were used for transmissions in the simulation setup. These are the same as the default traffic access categories in the IEEE 802.11e and were chosen for fair comparison. The characteristics of the four traffic classes used in the experiments are shown in Table 9-1.

TABLE 9-1 CHARACTERISTICS OF FOUR TRAFFIC CLASSES USED IN EXPERIMENTS

	Voice	Video	Best-effort	Background
Traffic	ITU-T G.711 CBR	H.264 CBR 25fps/CIF	Pareto distribution traffic model	Pareto distribution traffic model
Underlying Protocol	RTP/ UDP/IP	RTP/ UDP/IP	TCP/IP	RTP/ UDP/IP
Encoding Bit-rate	64Kbps	1000Kbps	128kbps	100kbps
Data packet size	100 bytes	1024bytes	512bytes	512bytes

The voice traffic uses ITU-T G.711 [229] which has been widely deployed in commercial products such as Skype[55]. Video traffic uses the MPEG-4 codec which is one of the most popular video codecs for video delivering over IP-based networks. Both best-effort traffic and background traffic were generated using Pareto distribution traffic models to mimic bursty traffic. The encoding bit-rate of voice and video data are set to typical values specified in the standard and industry, i.e., 64Kbps for voice traffic and 1000Kbps for video traffic. The bit-rate of the best-effort and background traffic was set to 128Kbps and 100Kbps, but they can be variable in real life. The overhead introduced by the underlying protocols RTP/UDP/IP and TCP/IP accounts for 40 bytes.

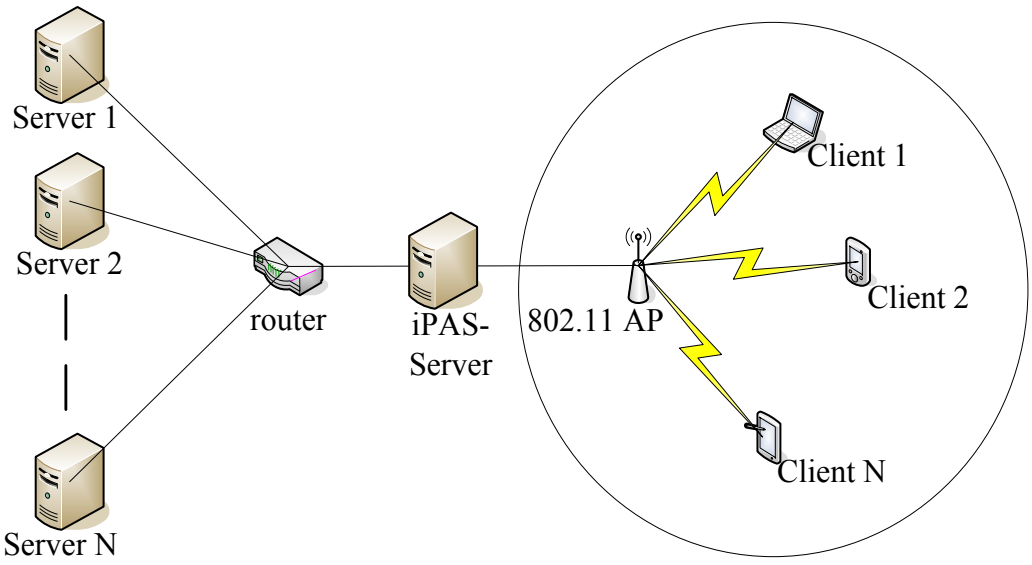


Figure 9-1 Simulation test-bed topology

9.1.2 Test-bed Setup

iPAS has been evaluated by using the NS-2.33[56] network simulator. The simulation topology is shown in Figure 9-1 includeing one iPAS server and N servers communicating with N clients, over an IEEE 802.11b wireless network. The original NS2 simulator was updated in the following aspects:

1) NS2 was extended to include the *IEEE 802.21 MIH* function based on the IEEE 802.21 specifications. The 802.21 MIH server and client were implemented as C++ objects in NS-2 under Linux environment;

[55] Skype-http://www.skype.com
[56] Network Simulator NS-2 [Online] http:// www.isi.edu/nsnam/ns/.

2) For the wireless environment, the No Ad-Hoc (NOAH[57]) patch was implemented in the NS-2 simulator in order to allow direct communication between mobile users and the AP only.

3) Marco Fiore patch[58]. Marco Fiore's patch was implemented to provide a more realistic wireless environment as the standard version of NS-2 does not consider the impact of interference and thermal noises. Marco Fiore patch added realistic channel propagation, multi-rate transmission support and Adaptive Auto Rate Fallback (AARF) [227]. The patch computes the Signal-to-Interference plus Noise Ratio (SINR) to add the effect of interference and thermal noises.

4) The IEEE 802.11e EDCA patch[59] for NS2 was also imported for the purpose of comparison.

9.1.3 Evaluation Metrics

Six evaluation metrics are used to assess the iPAS performance. Firstly, *throughput*, *delay*, *packet loss rate*, and *perceived video quality* (four metrics) are separately evaluated to study the effectiveness of both QoS differentiation and QoS provisioning. Secondly, the *fairness* between the demand and allocated bandwidth for all the traffic is studied. Additionally, the *impact of device resolution* on the bandwidth allocation is investigated. Next these metrics are described in details.

1) **Fairness**. In a system where streams make unequal demands for resources, one may want to measure fairness by closeness of the allocations to the respective demands. Jain's fairness index [230], as shown in equation (9-1) and equation (9-2), was selected to indicate the fraction of demand fairness.

$$f(x) = \frac{[\sum_{i=1}^{n} x_i]^2}{n \times \sum_{i=1}^{n} x_i^2}, \qquad x_i \geq 0 \tag{9-1}$$

[57] NOAH NS-2 extension, http://icapeople.epfl.ch/widmer/uwb/ns-2/noah/
[58] M. Fiore patch, http://www.telematica.polito.it/fiore
[59] 802.11e NS-2 patch, http://www.tkn.tu-berlin.de/research/802.11e_ns2/

$$x_i = \begin{cases} \frac{a_i}{d_i} & if\ a_i < d_i \\ 1 & otherwise \end{cases} \tag{9-2}$$

In the context of assessing the fairness of differentiated bandwidth allocation, the value d_i is the demand of i^{th} stream and a_i is the corresponding bandwidth allocation. The parameter n is the number of contending streams. This metric proposes a quantitative measurement-based criterion for resource allocation. The Jain's fairness index ranges between 0 and 1. For instance, a resource distribution algorithm with a fairness of 0.1 is unfair to 90% of the streams. It should be noted that allocating bandwidth more than the demand does not make any stream user happier. A higher value of Jain's fairness index indicates a closer relationship between demand and allocation, and therefore better QoS guarantee.

2) **Throughput**. The motivation of throughput investigation is to evaluate QoS provisioning. The throughput was studied in two aspects: per-class throughput and aggregated throughput. The analysis of the per-class throughput achieved by stations within each traffic class (voice, video, best-effort, background) indicates the effectiveness of QoS distribution. Additionally, the aggregate throughput presents the utilization of the limited wireless channel resources.

3) **Delay**. The transmission delay experienced by different stream reflects the effectiveness of the QoS differentiation and QoS provisioning. The delay represents the duration from the time when data packets are sent to the time when they are received. Multimedia applications such voice and video are sensitive to the delay, and lower delay contributes to better perceived quality. The instantaneous delay is computed for each arrived multimedia packet, as given in equation (9-3), where $Time_{rcvd}$ and $Time_{sent}$ represent the time stamp when the packet is received and sent for each flow, respectively.

$$Delay = Time_{rcvd} - Time_{sent} \tag{9-3}$$

4) **Packet Loss Rate**. The packet loss rate in wireless networks is due to three major causes: 1) **Signal attenuation** - Packets might be dropped due to the weak signal received; 2) **Collision** - When multiple stations try to access the shared wireless channels simultaneously, collision occurs. The packets are dropped and each station increases their contention window size; 2) **Retry attempts** - When the number of retransmission for lost packets exceeds the

retry threshold (the value equals 7 for 802.11), the packet is dropped. Higher packet loss rate indicates degraded received QoS. The calculation of packet loss rate, as given in equation (9-4), makes use of the total number of bytes sent by the server, *TotalSentBytes*, and the total number of bytes received by the client, *TotalRcvdBytes*.

$$LossRate = \frac{TotalSentBytes - TotalRcvdBytes}{TotalTxBytes} \tag{9-4}$$

5) **Perceived Video Quality**. The Peak Signal to Noise Ratio (PSNR) value has been widely used to objectively measure the video quality. Although PSNR cannot perfectly reflect the human perception, it performs well when using in the real-time adaptation [231]. The calculation of the estimated PSNR is supplied in equation (9-5), where *MAX_Bitrate* is the average bit-rate of stream, *EXP_Thr* is the average throughput expected to be achieved and *CRT_Thr* is the actual throughput measured. The higher the PSNR value obtained, the better video quality is.

$$PSNR = 20 * log_{10}\left(\frac{MAX\_Bitrate}{\sqrt{(EXP\_Thr - CRT\_Thr)^2}}\right) \tag{9-5}$$

6) **Device Resolution Awareness**. The impact of device resolution is studied to evaluate the performance of device awareness between iPAS, 802.11e EDCA, and 802.11 DCF. The device display resolution awareness is critical for streaming video applications. Typically, devices with higher screen resolution require more bandwidth allocated than those with lower screen resolution. The device screen resolution was mapped to the bandwidth requirement in a linear fashion. For instance, device with screen resolution of 480x360 (equals 172800 pixels) needs 21.7% lower bandwidth than that of 1024x768 (equals 786432 pixels).

9.1.4 Scenarios

In order to evaluate the performance of iPAS in terms of different evaluation metrics, three experimental scenarios were designed:

1) Scenario 1. This scenario aims to evaluate the fairness, throughput, transmission delay, and packet loss rate. AP and mobile station use the IEEE 802.11b mode. The scenario

includes four mobile stations, each transmitting a different traffic type (voice, video, best-effort, background). The number of mobile stations is increased from 4 to 32 in steps of 4 every 20s in order to increase the overall offered load. In all the tests, it was configured that the number of stations transmitting each traffic type is the same. Consequently, the ratio of the number of traffic flows in system is 1:1:1:1 for voice, video, best-effort and background. To efficiently analyze the iPAS performance under variable network conditions, the normalized offered load was used. The normalized offered load was computed as the absolute offered load divided by the channel capacity which is determined with respect to the theoretical maximum application-level throughput of the IEEE 802.11b, i.e. 7Mbps [232]. As the number of station increased, the corresponding normalized offered load increased from 20% to 160% (the channel is overloaded). The experimental time duration was set to 150s. Specifically, the normalized offered load achieved 100% when the number of stations exceeded 20 at around 80s. In order to collect the statistics data in stable conditions, all the measurements started 2s after the start of simulation.

2) Scenario 2.This scenario investigates the delivered video quality in terms of the PSNR value. The only difference from experimental scenario 1 is that, each mobile station received video traffic only.

3) Scenario 3. The purpose of this scenario is to study the effect of device awareness between iPAS, 802.11e EDCA, and 802.11 DCF. Unlike the setup in experimental scenarios 1 and 2, there were only two mobile stations receiving the same video traffic. The two stations were configured with different display resolutions, 480x360 and 1024x768. Unlike the previous scenarios, one server and client pair delivered the UDP background traffic as specified in Table 9-1. The background traffic was increased to vary the overall network load. The background traffic load increased from 20% to 160% of the wireless capacity .

For all of three experimental scenarios, the wireless access mode RTS/CTS was enabled to avoid the wireless hidden node problem. Moreover, the DropTail [224] was adopted as the default queue algorithm and the queue length was set to 50. In order to make the simulation close to a real environment, the Gaussian wireless error channel was introduced to provide a constant channel bit error rate around $1x10^{-7}$.The experiments analyze saturated network conditions where the active station has always a data frame to transmit.

9.2 Tests Results and Analysis

9.2.1 Fairness Study

In this simulation, the fairness performance of iPAS is compared with those of IEEE 802.11 DCF and IEEE 802.11e EDCA, which have been introduced in chapter 2. Jain's fairness index is applied to quantify the fairness level. The experimental scenario 1 was used.

Figure 9-2 and Table 9-2 show the fairness index for the three algorithms (iPAS, 802.11e EDCA, and 802.11 DCF) with various values of the offered load from 0% to 160%. All the stations are grouped to the four traffic types: voice, video, best-effort, and background. The figure shows that the fairness is good for low amount of offered load, e.g., F=0.9 (802.11 DCF), 0.91 (802.11e EDCA), and 0.93 (iPAS), when the offered load=20%. In the cases of 802.11 DCF and 802.11e EDCA, the fairness index decreases as the offered load increases. Specifically, the fairness decreases significantly for all traffic in 802.11 DCF and lower priority traffic in 802.11e EDCA. However, the fairness index of iPAS does not decrease significantly for any traffic class as the load increases. Take best-effort traffic for instance, when the offered load=100%, the fairness of 802.11 DCF and 802.11e EDCA decreases by about 21% and 40%, respectively, compared to the case of offered load =0%. In the case of iPAS, the decrease is around 5%. Additionally, iPAS achieves higher value of fairness index for certain traffic classes compared to 802.11 DCF and 802.11e EDCA for the entire range of offered load. Take video traffic for instance, when the network load is 80%, the fairness index achieved by using iPAS improved by 46.2% and 7.1%, in comparison with 802.11 DCF and 802.11e EDCA, respectively; when the network load is 120%, the fairness index achieved by using iPAS improved by 70.9% and 11.9%, in comparison with 802.11 DCF and 802.11e EDCA, respectively.

In conclusion, IEEE 802.11e EDCA shows the best performance for video and voice services when the network loads are lower than 60%. This is as IEEE 802.11e assigns very low contention window size for voice and video flows (i.e. 7-15 for voice and 15-31 for video), resulting in high channel access opportunity. At the same time, iPAS runs on top of the IEEE 802.11 protocol, where contention window size is defined by a higher range (i.e. 15-1023 for voice and video flows) than those of IEEE 802.11e. The lower the contention window range is, the higher the channel access opportunity obtained. Additionally, IEEE 802.11e performed the worst for best-effort and background services due to the starvation

issue. In general, IEEE 802.11 DCF provided the worst results as 802.11 DCF lacked service differentiation and network conditions adaptation mechanisms.

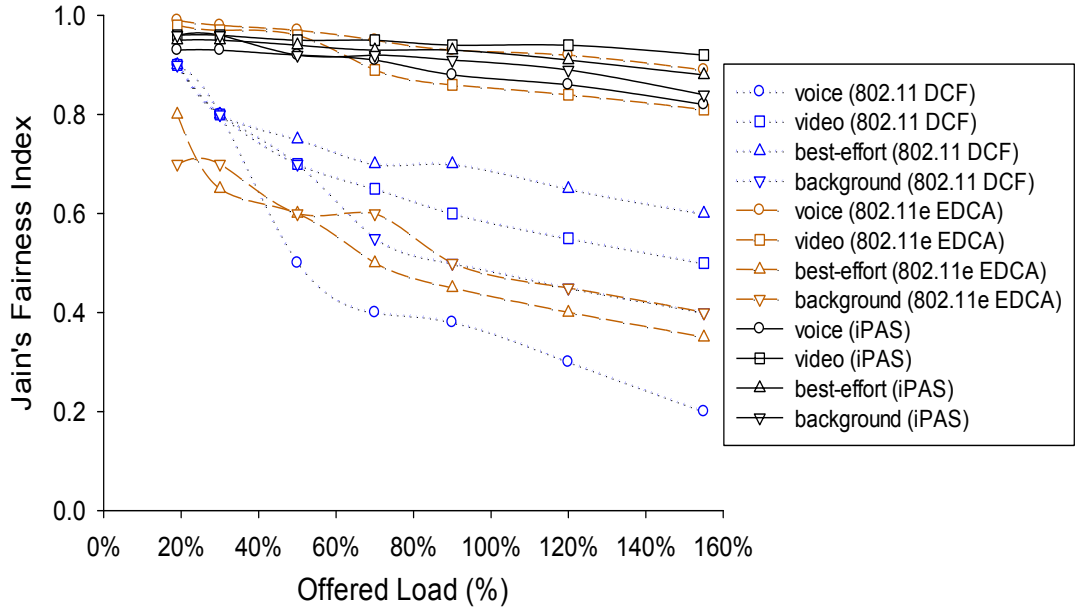


Figure 9-2 Jain's fairness index for different schemes delivering voice, video,best-effort and background traffic, with increasing amount of offered load

TABLE 9-2 JAIN'S FAIRNESS INDEX FOR DIFFERENT SCHEMES DELIVERING VOICE, VIDEO, BEST-EFFORT AND BACKGROUND TRAFFIC, WITH INCREASING AMOUNT OF OFFERED LOAD (VI-VIDEO, VO-VOICE, BE-BEST EFFORT, BG-BACKGROUND)

Load	IEEE 802.11 DCF				IEEE 802.11e EDCA				iPAS			
	VO	VI	BE	BG	VO	VI	BE	BG	VO	VI	BE	BG
20%	0.90	0.90	0.90	0.90	0.99	0.98	0.80	0.70	0.93	0.96	0.95	0.96
40%	0.80	0.80	0.80	0.80	0.98	0.97	0.65	0.70	0.93	0.96	0.95	0.96
60%	0.50	0.70	0.75	0.70	0.97	0.96	0.60	0.60	0.92	0.95	0.94	0.92
80%	0.40	0.65	0.70	0.55	0.95	0.89	0.50	0.60	0.91	0.95	0.93	0.92
100%	0.38	0.60	0.70	0.50	0.93	0.86	0.45	0.50	0.88	0.94	0.93	0.91
120%	0.30	0.55	0.65	0.45	0.92	0.84	0.40	0.45	0.86	0.94	0.91	0.89
140%	0.23	0.53	0.63	0.43	0.89	0.82	0.38	0.43	0.82	0.92	0.88	0.84
160%	0.21	0.51	0.61	0.41	0.85	0.81	0.37	0.42	0.81	0.91	0.84	0.81

The rapid decrease of fairness for traffic in 802.11DCF and low priority traffic in 802.11e EDCA is mainly due to the 802.11 CSMA/CA mechanism. The Contention Window (CW) size for any traffic in 802.11 DCF and lower priority traffic (best-effort and background) in 802.11e EDCA varies in the range of 15 to 1023 and 31 to 1023, respectively. This is as the higher priority traffic classes of 802.11e EDCA have lower CW ranges, i.e., 7

to 15 for voice traffic and 15 to 31 for video traffic. The increasing amount of offered load causes packet collisions, which determines the stations involved in collision enter the exponential backoff stage. The traffic in classes with higher CW ranges obtains less channel access opportunity and consequently lower bandwidth allocated. Therefore, the fraction of demand fairness decreases significantly for traffic with higher ranges of CW.

iPAS maintains very good fairness for all traffic types with the increasing network load. This can be explained by the fact that the throughput allocated to each stream is proportional to the stream priority and the wireless network conditions.

9.2.2 Throughput Study

This test investigated the throughput achieved by different traffic classes where 802.11 DCF, 802.11e EDCA, and iPAS are employed, respectively. The experimental scenario 1 was used.

Figure 9-3 presents the aggregate throughput received for voice and video traffic class, and Figure 9-4 shows the aggregate throughput experienced for best-effort and background traffic class. Table 9-3 presents the throughput values of Figure 9-3 and Figure 9-4. Each traffic class has the same number of flows or stations. When the total offered load is lower than 100%, (i.e., up to 20 stations), there is no significant difference between iPAS, 802.11 DCF and 802.11e EDCA for all traffic types, since there is enough bandwidth to transmit all of the traffic. It is observed that the aggregate throughput for the voice traffic class is the lowest among the four traffic classes for the three schemes due to the low bit-rate and packet size. In the case of 802.11 DCF and 802.11e EDCA, the throughput decreases significantly when the total offered load exceeds 100%. Note that the throughput experienced by the lower priority traffic (best-effort and background) in 802.11e EDCA drops more rapidly than the higher priority traffic (voice and video) with increasing amount of load. For instance, when the network load=140%, the aggregate throughput of the best-effort traffic class of 802.11e EDCA decreases by about 53%, compared to the case when load =100%. The aggregate throughput of video traffic in the 802.11e EDCA decreases with around 14%. This is because the traffic with lower priority has higher values for AIFS and contention window, meaning lower opportunity to obtain the channel access. Such phenomenon experienced in 802.11e EDCA is also called *starvation*. iPAS avoids the starvation problem of low priority traffic under high offered load. As shown in Figure 9-5, the aggregate throughput of best-effort traffic class and background traffic class decrease in a linear fashion due to the increase in the collision rate. Additionally, iPAS allocates the higher

throughput for both voice and video traffic in comparison to those of 802.11 DCF and 802.11e EDCA, demonstrating good QoS provisioning for multimedia services. When the total network load equals 100%, the available bandwidth estimated by *MBE* is around 6.5Mbps in these conditions. iPAS allocates 6% to voice, 73.8% to video, 12.7% to best-effort data, and 7.5% to background traffic. Take video traffic for instance, when the network load is 80%, the throughput achieved by using iPAS improved by 33.3% and 18.5%, in comparison with 802.11 DCF and 802.11e EDCA, respectively; when the network load is 120%, the throughput achieved by using iPAS improved by 31.2% and 17.1%, in comparison with 802.11 DCF and 802.11e EDCA, respectively.

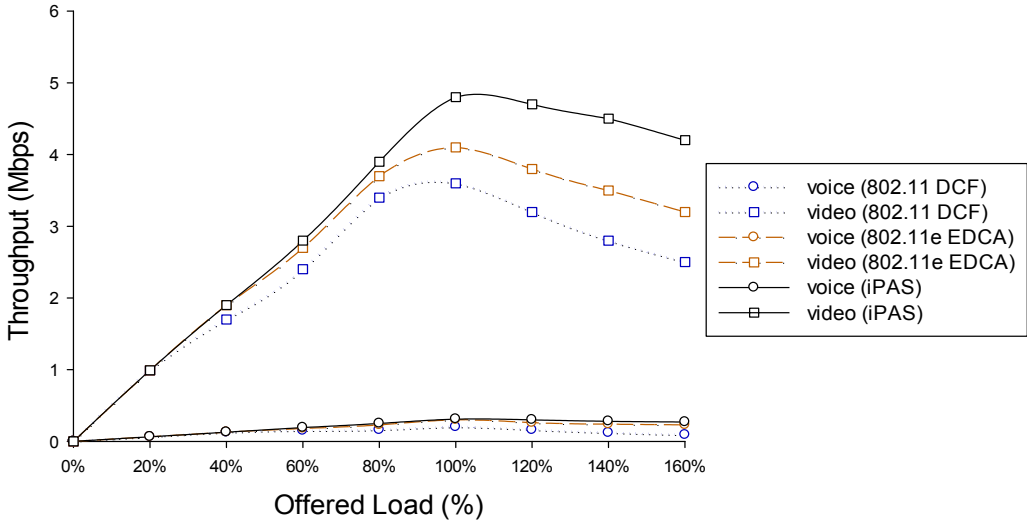


Figure 9-3 Aggregate Per-class throughput for different schemes delivering voice and video traffic with increasing amount of offered load

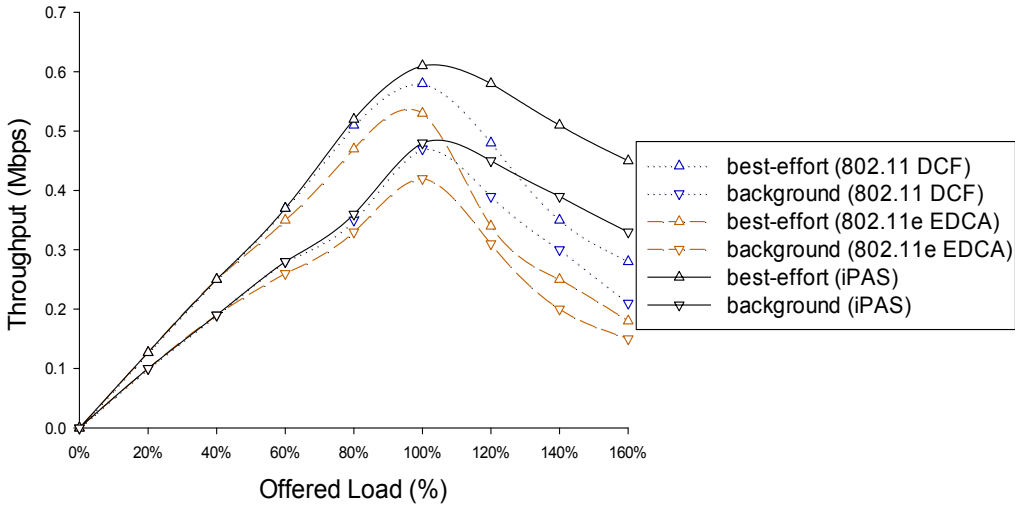


Figure 9-4 Aggregate Per-class throughput for different schemes delivering best-effort and background traffic with increasing amount of offered load

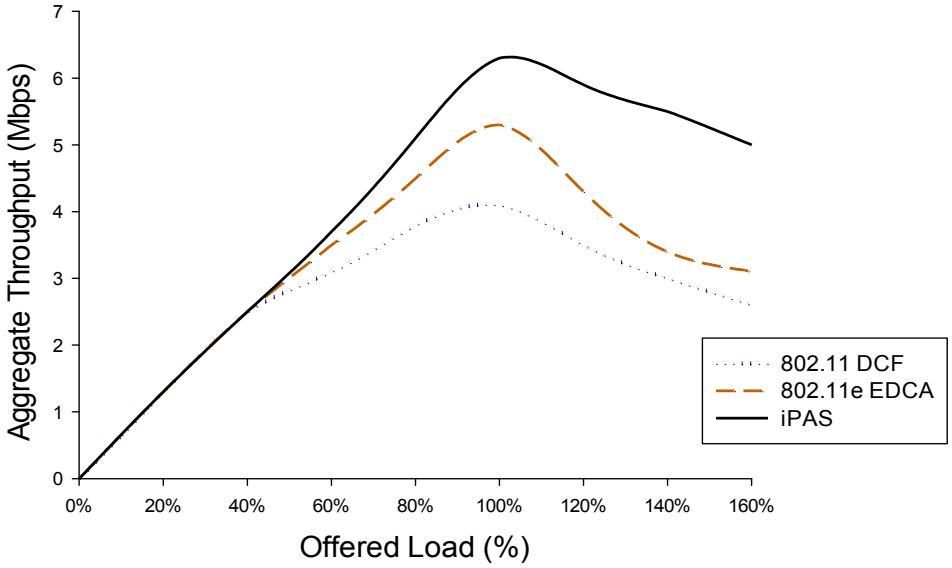


Figure 9-5 Aggregatte throughput for different schemes delivering voice, video, best-effort and background traffic, with increasing amount of offered load

TABLE 9-3 AGGREGATE PER-CLASS THROUGHPUT FOR DIFFERENT SCHEMES DELIVERING VOICE, VIDEO, BEST-EFFORT, AND BACKGROUND TRAFFIC WITH INCREASING AMOUNT OF OFFERED LOAD (VI-VIDEO, VO-VOICE, BE-BEST EFFORT, BG-BACKGROUND)

Load	IEEE 802.11 DCF (Mbps)				IEEE 802.11e EDCA (Mbps)				iPAS (Mbps)			
	VO	VI	BE	BG	VO	VI	BE	BG	VO	VI	BE	BG
20%	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
40%	0.06	0.99	0.13	0.10	0.06	0.99	0.13	0.10	0.06	0.99	0.13	0.10
60%	0.12	1.70	0.25	0.19	0.13	1.90	0.25	0.19	0.13	1.90	0.25	0.19
80%	0.15	2.40	0.37	0.28	0.18	2.70	0.35	0.26	0.19	3.20	0.37	0.28
100%	0.16	3.40	0.51	0.35	0.23	3.70	0.47	0.33	0.25	3.90	0.52	0.36
120%	0.20	3.60	0.58	0.47	0.30	4.10	0.53	0.42	0.31	4.80	0.61	0.48
140%	0.16	3.20	0.48	0.39	0.26	3.80	0.34	0.31	0.30	4.70	0.58	0.45
160%	0.12	2.80	0.35	0.30	0.24	3.50	0.25	0.20	0.28	4.50	0.51	0.39

TABLE 9-4 AGGREGATTE THROUGHPUT FOR DIFFERENT SCHEMES DELIVERING VOICE, VIDEO, BEST-EFFORT AND BACKGROUND TRAFFIC, WITH INCREASING AMOUNT OF OFFERED LOAD

load	IEEE 802.11 DCF(Mbps)	IEEE 802.11e EDCA (Mbps)	iPAS(Mbps)
20%	1.30	1.30	1.30
40%	2.50	2.50	2.50
60%	3.10	3.50	3.70
80%	3.80	4.50	5.10
100%	4.10	5.30	6.30
120%	3.50	4.30	5.90
140%	3.00	3.40	5.50
160%	2.60	3.10	5.00

Figure 9-5 and Table 9-4 show the aggregate throughput for all traffic classes which is a critical performance metric for QoS enhancement. It is observed that iPAS obtains higher aggregate throughput than both 802.11 DCF and 802.11e EDCA with increases of 38% and 20% respectively, for the entire network load variance. Furthermore, under high traffic load (i.e., offered load>100%), the aggregate throughput of iPAS and 802.11 DCF decreases linearly while that of 802.11e EDCA decreases abruptly. Take video traffic for instance, when the network load is 80%, the aggregate throughput achieved by using iPAS improved by 34.2% and 13.3%, in comparison with 802.11 DCF and 802.11e EDCA, respectively; when the network load is 120%, the aggregate throughput achieved by using iPAS improved by 68.6% and 37.2%, in comparison with 802.11 DCF and 802.11e EDCA, respectively. The high aggregate throughput of iPAS is also due to the performance of the MBE bandwidth estimation algorithm which gives more accurate information of current available bandwidth in the wireless network.

The study of throughput shows that both iPAS and 802.11e EDCA provide bandwidth differentiation for different traffic types. Moreover, iPAS offers better throughput provisioning than 802.11e EDCA and 802.11 DCF for increased traffic load.

9.2.3 Delay Study

This simulation compared the performance of iPAS with that of 802.11 DCF and 802.11e EDCA, with respect to the transmission delay. The experimental scenario 1 was used in this test.

To begin with, the focus is on delay differentiation performance. Figure 9-6 presents the average transmission delay experienced by voice, video, best-effort and background traffic in iPAS, 802.11 DCF and 802.11e EDCA. Since 802.11 DCF does not differentiate traffic based on the traffic classes, the delay of voice and video are significantly higher than those of 802.11e EDCA and iPAS. In the case of both voice and video traffic, the average delay experienced by 802.11e EDCA is slightly better than that of iPAS when traffic load is lower than 80%. This is as the contention window sizes of voice and video traffic in 802.11e are lower than that of iPAS. Figure 9-5 also shows that, in the cases of iPAS and 802.11e-EDCA, traffic with higher priority (voice and video) experienced significantly lower delay than traffic with lower priority (best-effort and background). This phenomenon confirms that both iPAS and 802.11e EDCA provide QoS differentiation for different traffic types.

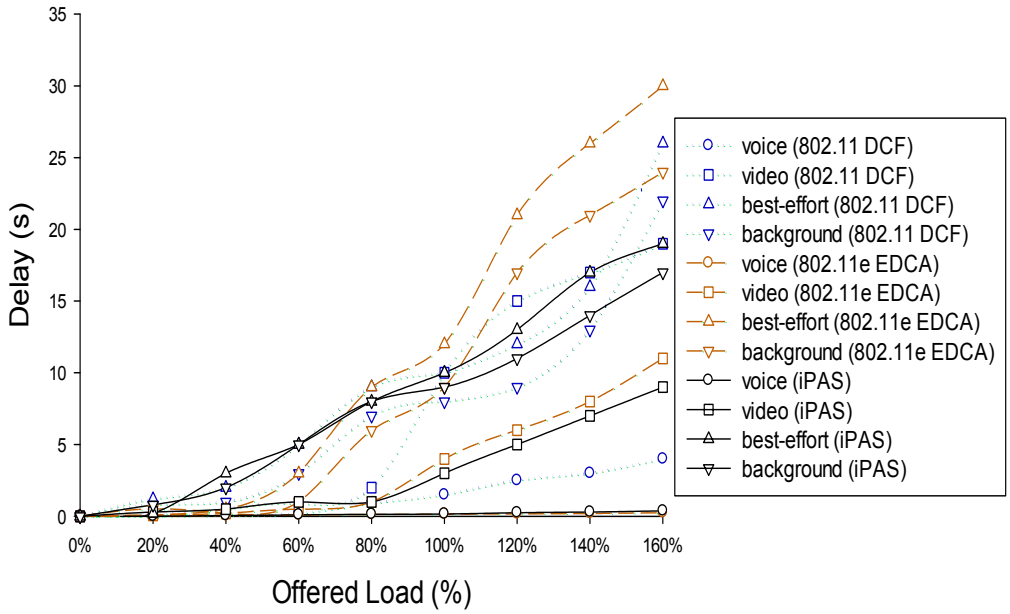


Figure 9-6 Average delay for different schemes delivering voice, video, best-effort and background traffic, with increasing amount of offered load

TABLE 9-5 AVERAGE DELAY FOR DIFFERENT SCHEMES DELIVERING VOICE, VIDEO, BEST-EFFORT AND BACKGROUND TRAFFIC, WITH INCREASING AMOUNT OF OFFERED LOAD (VI-Video, VO-Voice, BE-Best Effort, BG-Background)

Load	IEEE 802.11 DCF (s)				IEEE 802.11e EDCA (s)				iPAS (s)			
	VO	VI	BE	BG	VO	VI	BE	BG	VO	VI	BE	BG
20%	0.01	0.10	1.20	0.80	0.01	0.06	0.09	0.50	0.01	0.33	0.21	0.81
40%	0.07	0.50	2.00	1.20	0.05	0.20	0.50	0.20	0.05	0.50	3.30	2.03
60%	0.20	0.90	5.00	3.30	0.09	0.50	3.46	1.04	0.10	0.80	5.20	5.04
80%	0.90	2.00	9.40	7.40	0.13	1.40	9.45	6.02	0.14	0.80	9.42	6.34
100%	1.50	10.35	10.50	8.30	0.15	4.00	12.6	9.50	0.16	3.10	10.34	9.40
120%	2.50	15.30	12.07	9.60	0.17	6.90	21.70	17.60	0.25	4.80	13.30	11.30
140%	3.20	17.30	16.34	13.01	0.20	8.40	26.40	21.24	0.30	7.04	17.53	14.40
160%	4.05	19.43	26.63	22.1	0.25	11.30	30.10	24.16	0.38	9.r0	19.34	17.06

Next, the effect on delay of the increasing amount of traffic load is studied. It is observed from Figure 9-6 and Table 9-5 that, both 802.11e EDCA and iPAS perform effectively in satisfying QoS of voice and video traffic when the total network load was below 120% (i.e., the delay for voiceand video is lower than 400ms and 5s, respectively), indicating a fair quality level. Furthermore, the best-effort traffic experienced higher delay than background traffic, due to the TCP congestion control mechanism adopted by the best-

effort service. Despite 802.11e EDCA providing low delay for both voice and video traffic under low network load (i.e., traffic load<80%), the delay experienced by best-effort and background traffic in 802.11e increases dramatically under heavy offered load. Take video traffic for instance, when the network load is 80%, the delay achieved by using iPAS decreased by 60% and 42.9%, in comparison with 802.11 DCF and 802.11e EDCA, respectively; when the network load is 120%, the delay achieved by using iPAS decreased by 68% and 30.4%, in comparison with 802.11 DCF and 802.11e EDCA, respectively.

Consequently, both iPAS and 802.11e-EDCA can provide delay differentiation and delay provisioning for the four traffic types. Additionally, iPAS introduces low and smooth delay for the entire range of network load variance.

9.2.4 Packet Loss Rate Study

In this simulation, the packet loss rate for delivering voice, video, best-effort, and background traffic in iPAS, 802.11 DCF, and 802.11e EDCA was separately investigated. The experimental scenario 1 was used. Figure 9-7 shows the results, where the network load (X-axis) represents the overall load produced by the data traffic.

It is observed that iPAS provides the lowest packet loss rate for the entire range of network load variance. Under low traffic load (i.e., load<30%), the difference between the packet loss rate for the three schemes is not significant. In case of heavy traffic load (i.e., offered load>120%), the packet loss rate of 802.11 DCF and 802.11e EDCA increased significantly. For load from 100% to 160%, iPAS obtained packet loss rates lower with 18% and 34%, compared to 802.11e EDCA and 802.11 DCF, respectively. When the offered load is 80%, the loss produced by iPAS decreased by 30% and 22.2%, in comparison with 802.11 DCF and 802.11e EDCA, respectively; when the offered load is 120%, the loss produced by using iPAS decreased by 38.9% and 31.3%, in comparison with 802.11 DCF and 802.11e EDCA, respectively. This is because iPAS utilizes the wireless channel more efficiently due to an increase in the accuracy of estimated bandwidth by MBE.

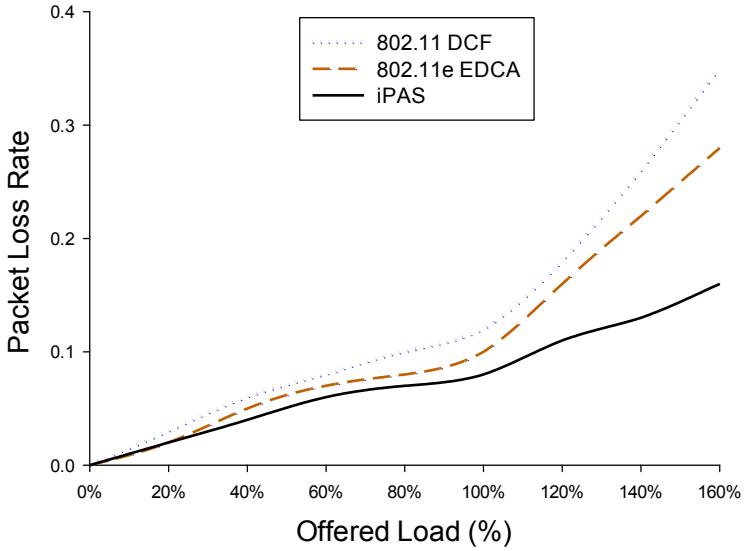


Figure 9-7 Average packet loss rate for 802.11DCF, 802.11e EDCA, and iPAS

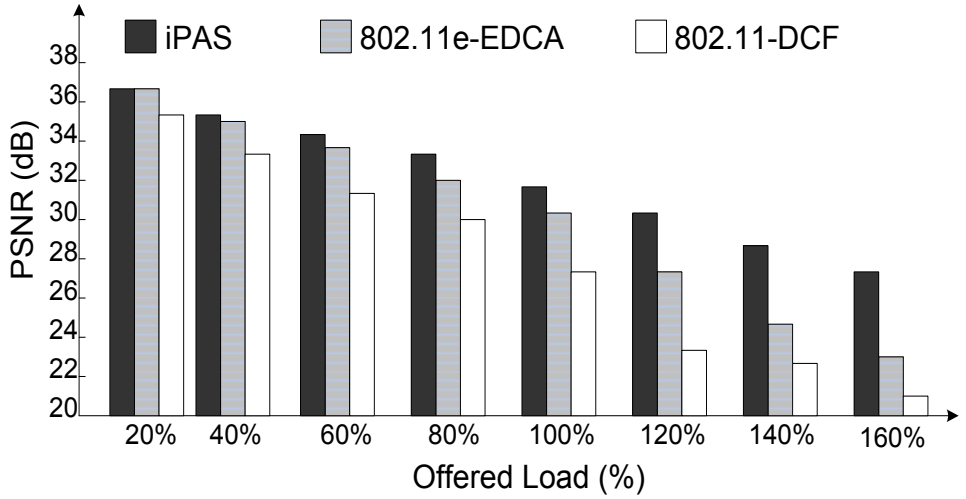


Figure 9-8 PSNR values with increasing amount of background traffic load for 802.11DCF, 802.11e EDCA, and iPAS

TABLE 9-6 AVERAGE PACKET LOSS RATE FOR 802.11DCF, 802.11E EDCA, AND IPAS

load	IEEE 802.11 DCF	IEEE 802.11e EDCA	iPAS
20%	0.03	0.02	0.02
40%	0.06	0.05	0.04
60%	0.08	0.07	0.06
80%	0.10	0.09	0.07
100%	0.12	0.10	0.08
120%	0.18	0.16	0.11
140%	0.26	0.22	0.13
160%	0.35	0.28	0.16

TABLE 9-7 PSNR VALUES WITH INCREASING AMOUNT OF BACKGROUND TRAFFIC LOAD FOR 802.11DCF, 802.11E EDCA, AND IPAS

load	IEEE 802.11 DCF (dB)	IEEE 802.11e EDCA (dB)	iPAS (dB)
20%	36.2	36.2	35.7
40%	35.2	35	33.1
60%	34.6	33.8	31.5
80%	33.5	32.2	30.7
100%	32.4	31.8	27.9
120%	31.2	27.4	23.3
140%	28.7	24.6	21.9
160%	27.7	23.1	21.2

9.2.5 Perceived Video Quality Study

This simulation investigated the received video quality for iPAS, 802.11 DCF, and 802.11e EDCA in terms of PSNR.

Figure 9-8 and Table 9-7 show that with the increase in the traffic load, the PSNR value of video traffic in 802.11 DCF decreases rapidly. For low traffic load (i.e., load<100%), the value of PSNR achieved by IEEE 802.11e EDCA and iPAS are both acceptable (PSNR>30). Although 802.11e EDCA showed better performance than iPAS in terms of delay as indicated in Figure 9-6, iPAS here performed better than 802.11e EDCA since the network load was increased by the background traffic instead of video traffic.

However, when the traffic load is higher than 100%, the value of PSNR for 802.11e EDCA decreases significantly while iPAS achieves high PSNR values for the whole range of the traffic load tested. For instance, when offered traffic load=160%, the PSNR value of video traffic in 802.11e EDCA decreases by about 32%, compared to the case of load =100%, whereas the PSNR of video traffic in iPAS decrease by 12% only. When the network load is 80%, the PSNR achieved by using iPAS improved by 8.4% and 4.7%, in comparison with 802.11 DCF and 802.11e EDCA, respectively; when the network load is 120%, the PSNR achieved by using iPAS improved by 25.3% and 14.9%, in comparison with 802.11 DCF and 802.11e EDCA, respectively. The reason is that iPAS can adapt the transmission rate of video traffic based on the accurate bandwidth estimation algorithm, which efficiently reduces the packet loss probability and improve the throughput.

9.2.6 Device Resolution Awareness Study

Experimental scenario 3 is used to conduct the performance comparison in terms of the device resolution among iPAS, 802.11 DCF, and 802.11e EDCA.

Each access scheme includes two video streams with device resolution equal 480x360 and 1024x768, respectively. It is observed from Figure 9-9 and Table 9-8 that there is no throughput differentiation for both 802.11 DCF and 802.11e EDCA (their plots overlap in the figure), since they do not support QoS differentiation based on device resolution. In the case of iPAS, the throughput for the device with lower resolution (480x360) is always lower than that for the device with higher resolution (1024x768), and the throughput ratio is around 2:3. This can be explained by the fact that the stereotypes-based resource allocation assigns a higher priority level to the device with higher resolution, therefore a higher bandwidth share is allocated. Figure 9-9 also shows that the throughput decreases linearly with the increasing amount of background traffic load, as expected. Take device resolution of 1024x768 for instance, when the offered load is 80%, the throughput achieved by using iPAS improved by 38.5% and 14.1%, in comparison with 802.11 DCF and 802.11e EDCA, respectively; when the offered load is 120%, the throughput PSNR achieved by using iPAS improved by 60% and 45.5 %, in comparison with 802.11 DCF and 802.11e EDCA, respectively.

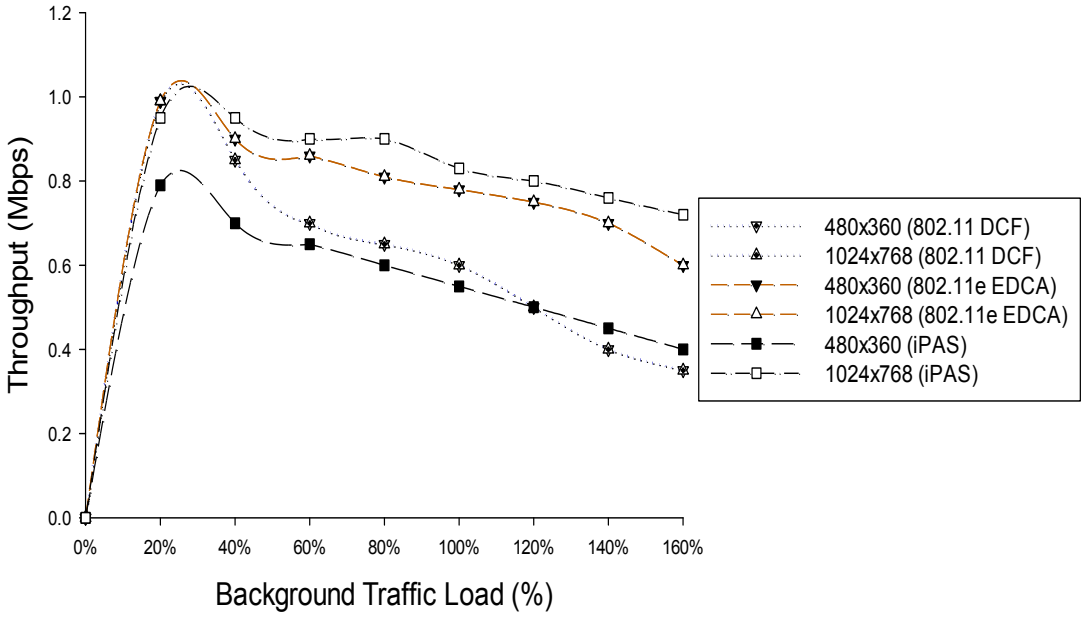


Figure 9-9 Throughput obtained by devices with different resolution for 802.11DCF, 802.11e EDCA, and iPAS

TABLE 9-8 THROUGHPUT OBTAINED BY DEVICES WITH DIFFERENT RESOLUTION FOR 802.11DCF, 802.11E EDCA, AND IPAS

load	IEEE 802.11 DCF (Mbps)		IEEE 802.11e EDCA (Mbps)		iPAS (Mbps)	
	480x360	1024x768	480x360	1024x768	480x360	1024x768
20%	0.99	0.99	0.99	0.99	0.79	0.95
40%	0.85	0.85	0.90	0.90	0.70	0.95
60%	0.70	0.70	0.86	0.86	0.65	0.90
80%	0.65	0.65	0.81	0.79	0.60	0.90
100%	0.60	0.60	0.78	0.76	0.55	0.83
120%	0.50	0.50	0.75	0.75	0.50	0.80
140%	0.40	0.40	0.70	0.70	0.45	0.76
160%	0.35	0.35	0.55	0.60	0.40	0.72

9.3 Summary

The proposed intelligent Prioritized Adaptive Scheme (iPAS) was introduced to provide both QoS differentiation and high QoS levels for content delivery to heterogeneous devices over IEEE 802.11 networks. iPAS assigns dynamic priority to each data stream and suggests a proportional bandwidth share according to the results of a stereotypes-based bandwidth allocation solution. This solution considers both QoS-related parameters such as delay, jitter, and packet loss rate and stream-related characteristics including device resolution, remaining device battery power, and application type. Performance evaluation was assessed in terms of six metrics: an inter-stream fairness index, throughput, packet loss rate, delay, video quality, and device resolution-awareness.

The following conclusions have been reached. 1) iPAS achieves higher and more stable fairness index than 802.11 DCF and 802.11e EDCA for four different service types (voice, video, best-effort, and background) with increasing network load; 2) iPAS can differentiate the bandwidth share among different streams according to the priority level. Noting that iPAS allocates higher throughput for both voice and video traffic in comparison with 802.11 DCF and 802.11e EDCA, demonstrating good QoS support for multimedia services. It is also observed that iPAS achieves the highest aggregate throughput for the entire range of network loads test-bed. The aggregate throughput of iPAS is higher than that of 802.11 DCF and 802.11e EDCA with up to 38% and 20%, respectively; 3) iPAS and 802.11e EDCA both provide delay differentiation for the four service types (i.e., voice traffic

experience the lowest delay and best-effort traffic achieve the highest delay); 4) iPAS obtains the lowest packet loss rate for the entire range of network loads. In the case of heavy traffic load, packet loss rate was lower with 18% and 34%, for iPAS than those of 802.11 DCF and 802.11e EDCA, respectively; 5) With increasing load, iPAS maintains higher and smoother PSNR values for video traffic, compared to 802.11 DCF and 802.11e ECCA; 6) Unlike 802.11 DCF and 802.11e EDCA, iPAS considers device resolution and therefore devices with higher screen resolution obtain more bandwidth share, which they require in terms of providing good user perceived quality.

CHAPTER 10

Experimental Evaluation of the Downlink/Uplink QoS Fairness Scheme

This chapter presents the experimental evaluation for the proposed QoS fairness scheme between downlink and uplink traffic in IEEE 802.11 network. The simulation-based test bed is described and the scenario focuses on VoIP traffic only. The performance evaluation studies VoIP capacity and fairness between downlink and uplink traffic. The VoIP capacity supported by the proposed scheme is measured. The QoS-based fairness between downlink and uplink traffic is assessed using Jain's fairness index.

10.1 Simulation-based Test bed Setup

The proposed scheme has been modelled and evaluated using Network Simulator-2 version 2.29[60]. The test bed setup is described in this section including network topology, test scenarios, assessment metrics, and setup of the schemes to be compared, *Dynamic-CW* and unfair channel access scheme between downlink and uplink provided by IEEE 802.11.

10.1.1 Network Topology

The simulation test bed used the wired-cum-wireless "**dumbbell**" topology, as shown in Figure 10-1. Multiple wireless clients received and sent unicast VoIP traffic from/to a group of servers via an IEEE 802.11b access point (AP). The servers were connected to the AP throughput one router and the wired link between the AP and router was set to 100Mbps with 2ms propagation delay. It was assumed that the IEEE 802.11b WLAN was the bottleneck link on the end-to-end path.

[60] NS-2.29-http://www.isi.edu/nsnam/ns

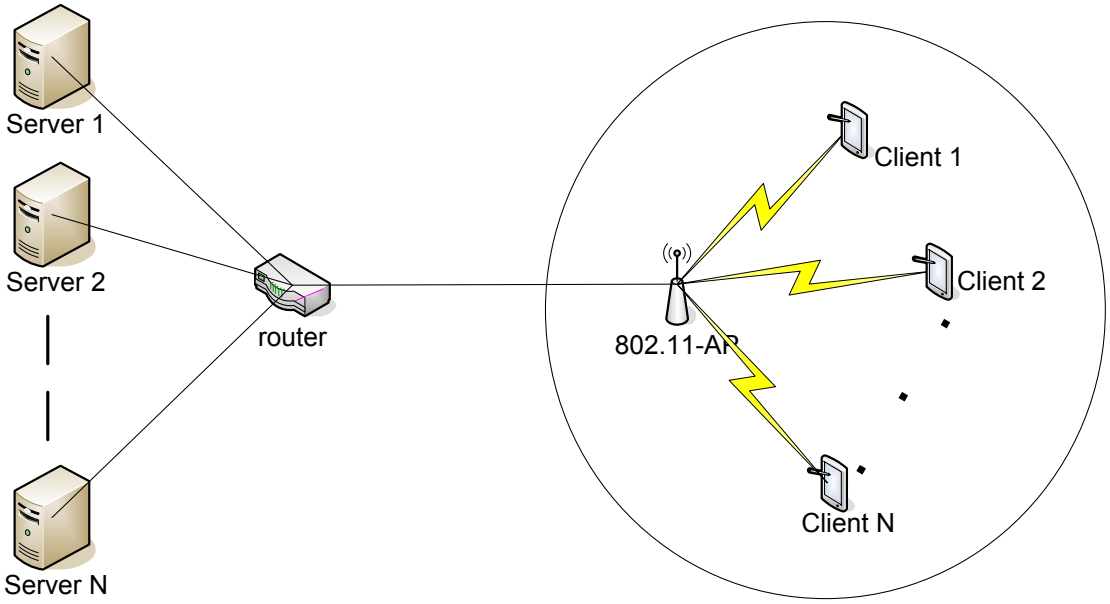


Figure 10-1 Downlink-Uplink Fairness Assessment Test bed topology

10.1.2 VoIP Traffic and Scenarios

Constant Bit Rate (CBR) VoIP traffic was generated using the ITU-T Rec. G.711 codec [229], with payloads of 160 bytes/packet. The bit-rate of CBR was set to 64kbps representing an inter packet transmission interval of 20ms. DropTail [224] queue with a limit of 100 packets was set to each wireless station. Two separate test scenarios were designed to study the VoIP capacity and fairness level between downlink and traffic, which are affected by network loaded conditions. For instance, higher network load might lead to higher collision rate and one-way delay than that of lower network load. Therefore, the VoIP capacity and fairness level would be degraded under high network loaded conditions. In order to vary the wireless network, the number of wireless stations and servers increased from 0 to 20 with steps of 2 every 30 seconds. In each scenario, the test time was set to 300 seconds.

10.1.3 Assessment Metrics

Two assessment metrics were used to evaluate the proposed downlink/uplink fairness scheme: VoIP capacity and the downlink/uplink fairness level.

The one-way end-to-end delay was used to estimate the VoIP capacity [233]. The acceptable delay for VoIP traffic is required to be less than 150ms according to the ITU-T G.114 [234]. It is assumed that the codec delay is 30-40ms, considering the 20ms bottleneck

delay, therefore the wireless network delay should less than 60ms[61]. Consequently, the VoIP capacity is defined as the maximum number of wireless stations whose delay (both uplink and downlink) is lower than 60ms. The 90th percentile value, instead of mean value, is used to measure the delay experienced by VoIP flows, because it better reflects the end user perceived quality of experience [235].

The fairness level between the downlink and uplink traffic is measured using the Jain's fairness index [230]. Let Q_D^i (i=1, 2, …, M) and Q_U^j (j=1, 2, … , N) represent the QoS parameters (throughput, delay, packet loss rate) of the i^{th} downlink flow and j^{th} uplink flow, respectively. For instance, $Q_{throughput}^i$, Q_{delay}^i, and Q_{loss}^i represent the throughput, delay, and packet loss rate of the i^{th} downlink flow. The parameter M and N are the numbers of downlink and uplink flows, respectively. The Jain's fairness index in terms of QoS of the downlink and uplink traffic is given in equation (10-1), where the parameter $FI_{down/up}$ ($0 < FI_{down/up} \leq 1$) is computed in the case of throughput, delay, and packet loss rate, separately. The fairness index ranges from 0 to 1. The closer the fairness index to 1, the higher level fairness of the system is.

$$FI_{down/up} = \frac{(\sum_{i=1}^{M} Q_D^i + \sum_{j=1}^{N} Q_U^j)^2}{(M+N)(\sum_{i=1}^{M} (Q_D^i)^2 + \sum_{j=1}^{N} (Q_U^j)^2)} \quad (10\text{-}1)$$

10.1.4 Setup of Dynamic-CW scheme and IEEE 802.11

The proposed scheme was compared with an existing downlink/uplink fairness improvement scheme called *Dynamic-CW* [233]. Additionally, the two schemes were compared with the unfair channel access scheme provided by IEEE 802.11 protocol.

Dynamic-CW scheme was proposed to consider the fairness issue between uplink and downlink flows in IEEE 802.11 wireless LANs, where uplink flows dominate over downlink flows in terms of wireless bandwidth usage. *Dynamic-CW* modifies the IEEE 802.11 MAC protocol at access points (APs) by dynamically controls the minimum

[61] CISCO, Understanding Delay in Packet Voice Networks, http://www.cisco.com/application/pdf/paws/5125/delay-details.pdf,2009

contention window size at APs. The authors of the *Dynamic-CW* scheme have evaluated the algorithm using IEEE 802.11b protocol on NS2 platform. In this chapter, *Dynamic-CW* is setup according to the specifications of the original paper [233]. The buffer size of each STA is set to be 100 packets. IEEE 802.11b with RTS/CTS mechanism [17], [21] is used, where *SIFS*=10μs, *PIFS*=30μs, *DIFS*=50μs, *slot time*=20μs, *PLCP preamble* =24bytes, *basic rate*=1Mbps, *data rate*=11Mbps, CW_{min}=31, CW_{max}=1023. All wired links have bandwidth of 100Mbps and propagation delay of 2ms. The test scenarios conducted for the proposed scheme was implemented to *Dynamic-CW*.

Original IEEE 802.11b protocol provides unfair channel access between downlink and uplink traffic. In particular, uplink traffic obtains the majority bandwidth of the wireless network. The IEEE 802.11b protocol is configured according to the specification [17]. The RTS/CTS mechanism was enabled to avoid the hidden node problem. The MAC layer parameters were set according to the 802.11b specifications, where *SIFS*=10μs, *PIFS*=30μs, *DIFS*=50μs, *slot time*=20μs, *PLCP preamble* =24bytes, *basic rate*=1Mbps, *data rate*=11Mbps, CW_{min}=31, CW_{max}=1023.

10.2 Performance Evaluation

10.2.1 VoIP Capacity Study

Figure 10-2 and Table 10-1 show the 90 percentile delay of per-flow VoIP traffic. In the case of the *IEEE 802.11* protocol, when the number of wireless stations (*N*) is higher than 12, the delay of both downlink and uplink VoIP traffic is higher than 60ms. In the case of *Dynamic-CW* scheme, when $N>14$, the delay of both downlink and uplink VoIP traffic is higher than 60ms. In the case of the proposed scheme, when $N>17$, the delay of both downlink and uplink VoIP traffic became higher than 60ms. Consequently, it can be concluded that the VoIP capacity using the proposed QoS-fairness scheme increases by 42% (i.e., from 12 calls to 17 calls) and 21% (from 14 calls to 17 calls), in comparison with *IEEE 802.11* and *Dynamic-CW*, respectively. Additionally, the downlink delay in the proposed algorithm is reduced at the cost of the uplink delay, in comparison with that of *IEEE 802.11*. For instance, the delay of downlink VoIP traffic is lower than that of *IEEE 802.11* protocol while the delay of uplink VoIP traffic is higher than that of *IEEE 802.11* protocol. Also the gap between the uplink delay and downlink delay is lower than that of *IEEE 802.11* and

Dynamic-CW when the number of stations is higher than 12. For instance, when *N*=16, the gap between downlink and uplink traffic decreases by 87.4% and 80%, in comparison with *IEEE 802.11* and *Dynamic-CW*, respectively. Notably, it is important to have a low uplink and downlink delay difference. In practical communication, lower delay difference indicates fairer service quality between downlink and uplink traffic.

10.2.2 Downlink/Uplink Fairness Study

Figure 10-3 and Table 10-2 present the Jain's fairness index in terms of QoS parameters, delay, throughput and loss. The fairness level of throughput, delay, and loss achieved by the proposed scheme is higher than both *IEEE 802.11* and *Dynamic-CW* as the number of VoIP stations increases from 0 to 20. For instance, when *N*=10, the delay fairness of the proposed scheme increases by 16.7% and 8.2%, the throughput fairness of the proposed scheme increase by 7.2% and 6.2%, and the loss fairness of the proposed scheme increase by 11.5% and 6.3%, in comparison with *IEEE 802.11* and *Dynamic-CW*, respectively. When *N*=20, the delay fairness of the proposed scheme increases by 54.5% and 48.9%, the throughput fairness of the proposed scheme increase by 31.5% and 14.3%, and the loss fairness of the proposed scheme increase by 30.1% and 24.7%, in comparison with *IEEE 802.11* and *Dynamic-CW*, respectively. Along with the increasing number of VoIP stations, the QoS fairness of *IEEE 802.11* and *Dynamic-CW* decreases significantly, while the proposed scheme is not impacted as much. For instance, as the number of wireless stations (*N*) increases from 2 to 20, the fairness level of delay, throughput, and loss decrease by 59.6%, 39.4%, and 34.3% for *IEEE 802.11*; 54.5%, 21.2%, and 29.3% for *Dynamic-CW*; and 11.1%, 9.1%, and 7.1% for the proposed scheme.

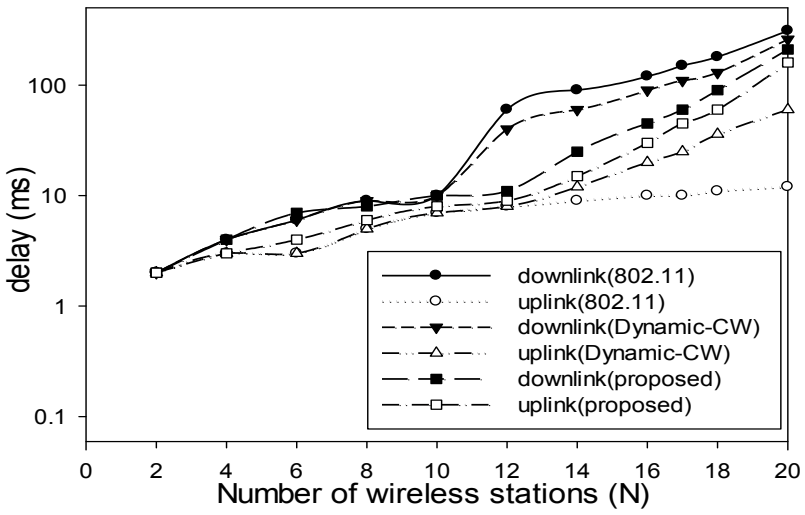


Figure 10-2 Ninetieth percentile delay of VoIP

TABLE 10-1 DELAY OF DOWNLINK AND UPLINK TRAFFIC FOR IEEE 802.11, DYNAMIC-CW AND THE PROPOSED SCHEME

N	IEEE 802.11		Dynamic-CW		Proposed Scheme	
	Downlink (ms)	Uplink (ms)	Downlink (ms)	Uplink (ms)	Downlink (ms)	Uplink (ms)
2	2	2	2	2	2	2
4	4	3	5	3	3	3
6	6	3	5	3	7	4
8	9	5	11	4	8	6
10	11	7	10	6	10	8
12	63	8	40	5	11	9
14	87	9	56	12	25	15
16	121	10	91	21	44	30
18	149	10	104	25	63	45
20	178	11	122	36	94	60

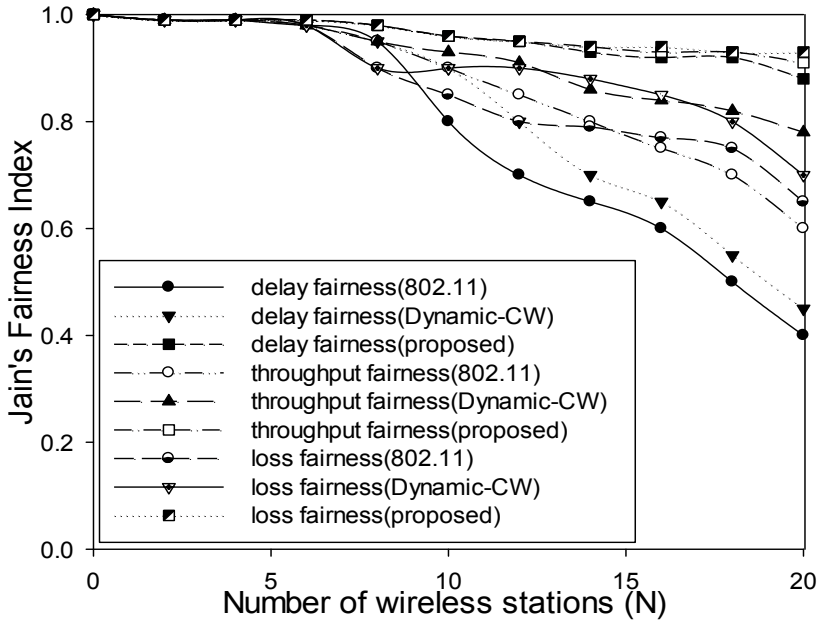


Figure 10-3 Jain's fairness index in terms of delay, throughput, and loss

TABLE 10-2 JAIN'S FAIRNESS INDEX IN TERMS OF DELAY, THROUGHPUT, AND LOSS ACHIEVED BY IEEE 802.11, DYNAMIC-CW, AND THE PROPOSED SCHEME

N	delay			throughput			loss		
	IEEE 802.11	Dynamic-CW	Proposed Scheme	IEEE 802.11	Dynamic-CW	Proposed Scheme	IEEE 802.11	Dynamic-CW	Proposed Scheme
2	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99
4	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99
6	0.98	0.98	0.99	0.98	0.98	0.99	0.98	0.98	0.99
8	0.95	0.95	0.98	0.95	0.95	0.98	0.90	0.90	0.98
10	0.80	0.90	0.98	0.90	0.91	0.97	0.85	0.90	0.96
12	0.70	0.80	0.95	0.85	0.91	0.95	0.80	0.90	0.95
14	0.65	0.70	0.93	0.80	0.86	0.94	0.79	0.88	0.94
16	0.60	0.65	0.92	0.75	0.84	0.93	0.77	0.85	0.94
18	0.50	0.55	0.92	0.70	0.82	0.93	0.75	0.80	0.93
20	0.40	0.45	0.88	0.60	0.78	0.9	0.65	0.70	0.92

10.3 Summary

This chapter proposes a new Contention Window (CW) Adaptation Scheme for mobile consumer devices using VoIP. The AP's CW size is dynamically changed according to the results of a stereotype-based adaptation. Performance evaluation was assessed in terms of two metrics: 90th percentile delay of VoIP and Jain's fairness index. Simulation results show how the proposed algorithm improves the downlink/uplink fairness for three QoS parameters such as delay, throughput and loss in comparison with sate of the art solution like *IEEE 802.11* and *dynamic CW*. The following conclusions have been reached. 1) By using the proposed scheme, the VoIP capacity increases by 42% and 21%, respectively, in comparison with *IEEE 802.11* and *Dynamic-CW*; 2) The gap between the uplink delay and downlink delay is lower than that of *IEEE 802.11* provided that the number of stations is higher than 12; 3). The fairness level provided by the proposed scheme is higher than both *IEEE 802.11* and *Dynamic-CW* with up to 132%, 52% and 37%, in terms of delay, throughput and packet loss rate, respectively.

CHAPTER 11

Prototyping and Result Analysis

This chapter presents the iPAS prototyping and related experimental results. The purpose of iPAS is to improve the original IEEE 802.11-based protocols by allocating bandwidth resources based on stream's priority. iPAS has been first deployed in a network simulation environment and the same topology was implemented in real life test-bed which will be described in this chapter. iPAS is evaluated both objectively and subjectively by comparing with the classic equal priority channel access scheme of IEEE 802.11 protocol. The experimental test results are presented and commented on.

11.1 Introduction

The intelligent Prioritized Adaptive Scheme (iPAS) provides QoS differentiation between multiple streams during wireless multimedia delivery. iPAS assigns dynamic priorities to various streams and determines their bandwidth share by employing a stereotypes-based approach. Previous simulation-based tests have assessed the performance of iPAS in terms of fairness, throughput, delay, packet loss rate, video quality, and device-awareness. Although the estimated Peak Signal to Noise Ratio (PSNR) was used as the video quality metric in the simulation test, actual measured PSNR as well as user perceptual evaluation are required in real life tests. Prototyping iPAS and the related real life test-bed setup are presented in this chapter.

The real life test focuses on evaluating the performance of iPAS in terms of the delivered video quality. For instance, in the real life test, MPEG-4 encoded videos are transmitted to three devices (laptop, tablet, and smartphone) over IEEE 802.11g network. The delivered video clips are recorded on each device. Additionally, the same MPEG-4 video sequences are delivered in the simulation-based test with the same test topology. The delivered video clips from both real life test-bed and simulation test-bed are saved and then evaluated using objective and subjective video quality metrics.

This chapter is structured as follows. Firstly, the implementation details related to the real life test-bed and simulation test-bed are presented, such as test topology, equipment and software used video content, background traffic configuration, and experimental scenarios. Secondly, the delivered video sequences from real life test and simulation are compared in terms of objective video quality metrics such as throughput and PSNR. Finally, the transmitted video clips from real tests and simulations are compared using subjective video quality metrics based on ITU-T R. P.911 [53].

11.2 Real Life Test-bed Setup

11.2.1 Test Topology

The real life test-bed topology is shown in Fig.11-1, and consists of: a multimedia server, a traffic generator (with traffic generator controller), an IEEE 802.11g wireless router, a network monitor, and an Android smartphone, a laptop, and a tablet PC. Figure.11-2 further presents the photo of the test-bed based on the topology in Fig. 11-1.

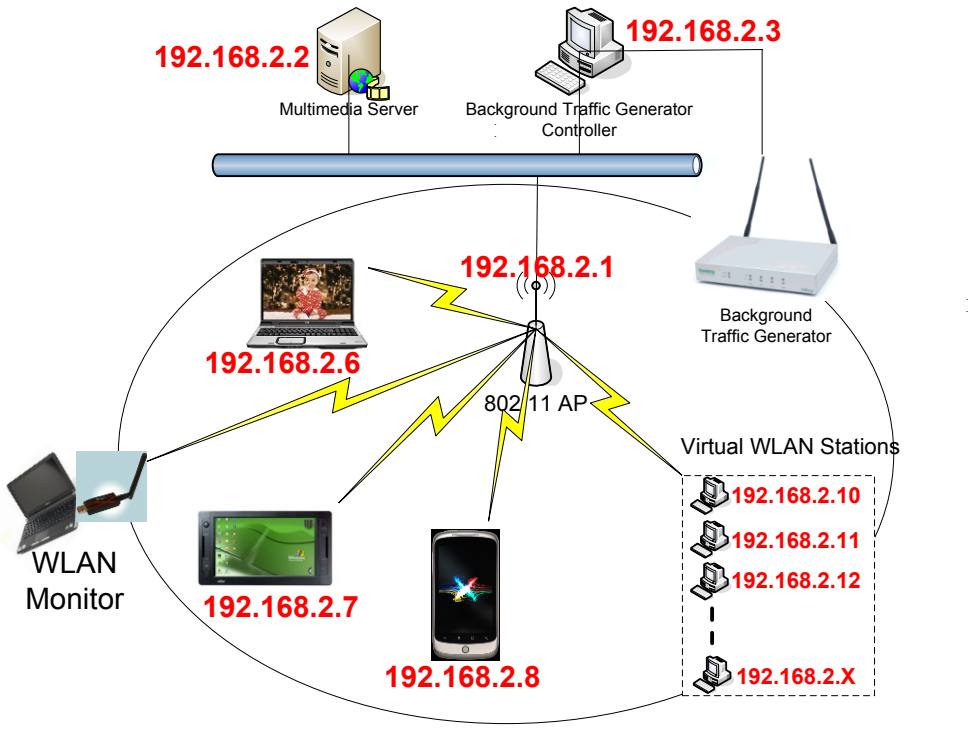


Figure 11-1 Real life test-bed topology

Figure 11-2 Photo of the real test bed environment

11.2.2 Equipment and Software Specifications

The equipment and software involved in the test-bed are described below:

Multimedia Server

The multimedia server runs on a HP Pavillion dv3 laptop with Microsoft Windows 7 Home Edition x64, Intel Core 2 Duo T6600 at 2.2GHz and 4GB RAM. The multimedia software used on the laptop is the *Wowza Media Server 3*[62] which supports live or on-demand streaming to computers, mobile devices and IPTV endpoints. *Wowza Media Server 3* enables multiple streaming protocols such as Real Time Messaging Protocol (RTMP), Real Time Messaging Protocol Tunnel (RTMPT), HTTP streaming, Real Time Streaming Protocol (RTSP), Real-time Transport Protocol (RTP), MPEG transport stream (MPEG-TS), etc. As shown in Fig.11-1, the IP address of multimedia server is 192.168.2.2.

Traffic Generator

The traffic generator used is the LANForge-WiFIRE 802.11a/b/g from Candela Technologies[63], which supports creating up to 32 virtual wireless stations. The traffic

[62] Wowza Media Server 3-http://www.wowza.com
[63] LANForge-WiFIRE, Candela Technologies-http://www.candelatech.com/lanforge_v3/ct520_product.html

generator is capable to generate more than 45Mbps traffic by using various of protocols such as TCP/IP, UDP/IP, FTP, HTTP, etc. A separate computer is needed to run the LANForge management software. As shown in Fig.11-1, the IP address of traffic generator is 192.168.2.3. The IP address of the virtual WLAN stations starts from 192.168.2.10.

Wireless Router

Belkin N Wireless Router[64] is used to provide the local wireless network. The router is configured to run on channel 6 with no other networks running on the same channel, in order to avoid the interference. IEEE 802.11g protocol is enabled at 2.437GHz frequency. The multimedia server and the LANForge traffic generator are connected to the wireless router via Unshielded Twisted Pair (UTP) cable. As shown in Fig.11-1, the IP address of the wireless router is 192.168.2.1.

Network Monitoring Devices

Two pieces of hardware equipment are used to monitor the IEEE 802.11g network: Wi-Spy DBx65 and AirPcap Nx66.They are capable of monitoring the interference levels and enable analysing the network traffic characteristics such as bandwidth, retransmission, frame size distribution, etc. The Wi-Spy DBx includes Chanalyzer 467 software and AirPcap comes with WiFi Pilot 2.468 and Wireshark69 software. Both Wi-Spy DBx and AirPcap Nx are connected to a Sony VAIO VGN-CS11S laptop running Microsoft Windows 7 Enterprise x86 with Intel Core 2 Duo P8400 processor at 2.26GHz and 4GB RAM.

Client Devices

There are three wireless devices used to receive the multimedia streams. The hardware specifications are summarized in Table 11-1. These devices are connected to the Belkin wireless router via IEEE 802.11g protocol. As shown in Fig 11-1, the IP address allocated for the laptop, tablet and smartphone are 192.168.2.6, 192.168.2.7, and 192.168.2.8.

[64] Belkin N Wireless Router-http://www.belkin.com
[65] Wi-Spy DBx-http://www.metageek.net/products/wi-spy/
[66] AirPcap Nx-http://www.metageek.net/products/airpcap/
[67] Chanalyzer 4-http://www.metageek.ent/products/chanalyzer-4
[68] WiFi Pilot 2.4-http://www.metageek.net/products/wifipilot
[69] Wireshark-http://www.wireshark.org

TABLE 11-1 CLIENT DEVICES USED FOR RECEIVING MULTIMEDIA STREAMS

Devices	Model	Screen Resolution	Memory (RAM)	CPU	WLAN	Operating System
Laptop	HP Pavilion dv3	13.3'' Brightview (1280x800)	4GB	Intel Core 2 Duo T6600@2.2GHz	IEEE 802.11b/g/n	Windows 7
Tablet	Viliv X70	7'' WSVGA (1024x600)	1GB	Intel Atom @1.33GHz	IEEE 802.11b/g	Windows XP
Smartphone	Google Nexus One	3.7''WVGA (800x480)	512MB	Qualcomm Snapdragon@1GHz	IEEE 802.11b/g	Android 2.2

iPerf and JPerf

iPerf[70] is used to measure the available bandwidth of the wireless network. iPerf software is installed in client mode in the Android smartphone and works in conjunction with Jperf[71] installed on the server side.

MSU Video Quality Measurement Tool

MSU Video Quality Measurement Tool[72] software is used for assessing the objective video quality. It provides functionality for both full-reference (two videos are examined) and single-reference (one video is analyzed) comparisons. The tool supports most of the objective video quality metrics such as PSNR, Aligned PSNR (APSNR), Video Quality Model (VQM), Structural Similarity (SSIM), Mean Square Error (MSE), MSU Blurring Metric, MSU Blocking Metric, MSU Brightness Flicking Metric, MSU Drop Frame Metric, MSU Noise Estimation Metric, etc. The tool requires the original and delivered video as input, in order to perform the video quality measurement.

VLC Media Player

The Video LAN Client (VLC)[73] Version 2.0.1 is an open-source video player. It can be deployed on various platforms such as Microsoft Windows, Linux, Mac OS X, Unix, etc.

[70] iPerf for Android-http://www.appbrain.com/app/iperf-for-android/com.magicandroidapps.iperf
[71] JPerf-http://sourceforge.net/projects/jperf/
[72] MSU Video Quality Measurement Tool
http://compression.ru/video/quality_measure/video_measurement_tool_en.html
[73] VideoLAN Client Version 2.0.1-http://www.videolan.org/vlc/

VLC supports most codecs with no codec packs needed: MPEG-2, H.264, DivX, MPEG-4, WebM, WMV palyer, etc. Additionally, VLC provide "convert/save" function which can record the playback video. VLC media player is installed on the two client devices: HP Pavilion dv3 laptop and the Viliv tablet PC.

MoboPlayer

MoboPlayer[74] is used as the video player on the Android platform since VLC player is not supported by the Android smartphone. MoboPlayer supports multi-audio streams and multi-subtitles playlists and continuous play on same type files Videos streamed through HTTP, RTSP protocols. The current version of MoboPlayer needs to run on Android 1.6 or later. It supports almost all Android devices with an ARM architecture.

11.2.3 Wireless Environment

The IEEE 802.11g wireless network is setup by using the Belkin N Wireless Router. The basement of the Electronic Engineering building in Dublin City University was selected to deploy the wireless network in order to reduce the interferences, since there is a significant number of wireless networks on campus whose signal strength are very much reduced in the basement. Wi-Spy DBx spectrum analyser software is used for monitoring the surrounding wireless networks. Fig 11-3 illustrates the wireless networks monitored in the test room. The SSID of the studied network is "iPAS" which is running on channel 6 (frequency 2.437GHz) with no other networks running on the same or adjacent channels.

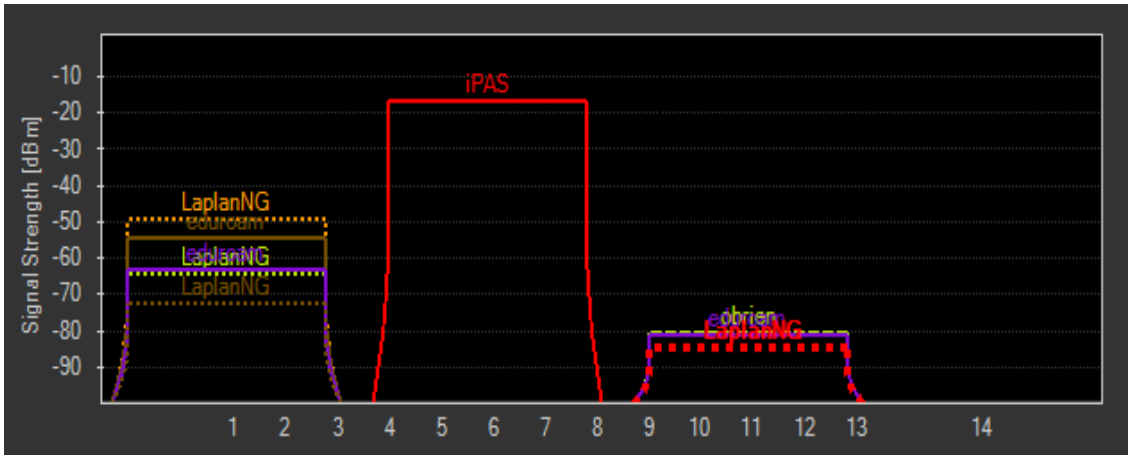


Figure 11-3 Surrounding wireless networks

[74] MoboPlayer-http://www.moboplayer.com/

11.2.4 Video Sequences

A cartoon sequence, *The Simpsons Movie*[75] , is used for real life test as it has been widely used in many other video tests. Additionally, video trace files for *The Simpsons Movie* are provided on the *TKN* website[76] and can be readily used for simulations. The movie is encoded into three different quality levels according to Adobe's mobile video encoding recommendations[77]. The visual differences of the three video sequences are clear, since the cartoon consists of images with sharp edges which are very sensitive to bit-rate and resolution modifications. The encoding characteristics of the three video sequences for real test are presented in Table 11-2. MPEG-4 AVC video compression and AAC audio compression are used together with an MP4 container[78].

TABLE 11-2 VIDEO CLIPS USED FOR REAL LIFE TEST-BED

Video Clip Quality Level	Codec	Overall Bit-rate (Kbps)	Resolution	Frame rate (fps)	Duration (minutes)	Audio Codec
High	MPEG-4 (Base media/version2)	1300	800x448	30	20	AAC 25Kbps 8KHz
Medium	MPEG-4 (Base media/version2)	420	512x288	25	20	
Low	MPEG-4 (Base media/version2)	230	320x176	20	20	

The bit-rate and resolution of each video clip are varied together in order to maintain a consistent level of compression quality. For instance, the high quality video uses a bit-rate of 1300Kbps and resolution of 800x448. If the low quality video uses the same resolution of 800x448 while being encoded with lower bit-rate of 230Kbps, the video quality would decrease due to compression effects, such as blockness, blurring, colour smearing, etc. According to the encoding considerations from Adobe[79], the aspect ratio of the original video should be maintained. The source movie is encoded at 16:9 which is supported by most of the wide screen devices.

[75] *The Simpsons Movie* website: http://www.simpsonsmovie.com/
[76] http://www2.tkn.tu-berlin.de/research/trace/trace.html
[77] Adobe mobile encoding guide:
http://www.adobe.com/devnet/devices/articles/mobile_video_encoding.html
[78] MP4 multimedia container-http://en.wikipedia.org/wiki/MPEG-4_Part_14
[79] Adobe encoding guide for android phone:
http://download.macromedia.com/flashmediaserver/mobile-encoding-android-v2_7.pdf

Fig.11-4 shows the screenshot pictures of the video sequences with different quality levels. The picture of the low quality video (Fig.11-4 c) presents some blurry aspects in comparison with that of the high and medium quality video (Fig.11-4 and Fig.11-b).

(a) High quality video (1300Kbps/800x448) **(b) Medium quality video (420Kbps/512x288)**

(c) Low quality video 230Kbps/320x176

Figure 11-4 Video clips with different quality

11.2.5 Background Traffic Setup

Background traffic is introduced to the wireless network in order to evaluate the impact of network load on video transmission. According to a survey provided by Cisco[80], the ratio of downlink to uplink traffic could rise to 10:1 over the next five years. Therefore, the background traffic for real life test-bed is generated based on this traffic estimation (10:1). In order to load the IEEE 802.11g network gradually, the available bandwidth is measured using JPerf at the server side and IPerf at the client side. IPerf measures the available bandwidth of a certain path by generating probing traffic. The estimated available bandwidth

[80] Cisco Systems, Capacity, Coverage, and Deployment Considerations for IEEE 802.11g, White Paper, [Online]. Available: http://www.sparcotech.com/Cisco%20Capacity%20and%20Coverage%20with%20G.pdf

is in the range of 21Mbps of 23Mbps. To study the impact of background traffic on video delivery, the overall background traffic load is gradually increased up to around 19Mbps~20Mbps. The selected background traffic load provides a high load, but does not overload the network.

Table 11-3 illustrates the background traffic characteristics. TCP and UDP traffic are used for both downlink and uplink traffic. The packet size of the downlink TCP flow ranges between 100bytes and 1472bytes with a transmission rate between 56Kbps and 1.5Mbps, which covers many widely used services. For instance, *Skype*[81] voice call uses a private protocol over TCP/IP with data frame size ranges between 110bytes and 150bytes. Video content providers also select TCP as the primary protocol, i.e., *Youtube*[82] uses HTTP over TCP and *WarnerBros*[83] uses RTMP over TCP. The video data frame size ranges from 1300bytes to 1500bytes. The packet size of the downlink UDP flow is 1472bytes with transmission rate 1Mbps.The uplink TCP and UDP traffic uses smaller packet size (i.e. 60bytes-640bytes) and lower transmission rate (i.e. 56Kbps-512Kbps) in comparison with that of downlink traffic.

Table 11-4 shows the number of background traffic flows. The video transmission time is set to 320s. During the first 20s, there is no background traffic. From 20s to 320s, the number of background flows increases from 6 to 30 with 6 new flows added every 60s. For instance, during the 140s-200s time intervals, there are 6 TCP downlink flows, 6 UDP downlink flows, 3TCP uplink flows, and 3 UDP uplink flows. Consequently, the overall background traffic load is up to 17.3Mbps based on Table 11-4. For every 60s (starting from 20s), the number of downlink flows is higher than that of uplink flows, in order to maintain a high ratio of downlink to uplink. For instance, during 140s-200s, the ratio of downlink to uplink traffic load is around 7.1, and during 260s-320s, the ratio is around 8.2.

Figure 11-5 presents the background traffic configuration page from the LANForge-WiFIRE traffic manager. The number of running traffic is set up based on Table 11-5.

[81] Skype-http://www.skype.com
[82] Youtube-http://www.youtube.com
[83] WarnerBros-http://www.warnerbros.com

TABLE 11-3 DOWNLINK AND UPLINK TRAFFIC CHARACTERISTICS

	TCP		UDP	
downlink	Packet size	Transmission rate	Packet size	Transmission rate
	100bytes-1472bytes	56Kbps-1.5Mbps	1472bytes	1Mbps
uplink	60bytes-300bytes	56Kbps-256Kbps	640bytes	512Kbps

TABLE 11-4 THE NUMBER OF BACKGROUND TRAFFIC FLOWS

Time	Number of flows	Number of downlink flows		Number of uplink flows	
		TCP	UDP	TCP	UDP
0s~20s	0	0	0	0	0
20s~80s	6	2	2	1	1
80s~140s	12	4	4	2	2
140s~200s	18	6	6	3	3
200s~260s	24	6	10	4	4
260s~320s	30	6	16	4	4

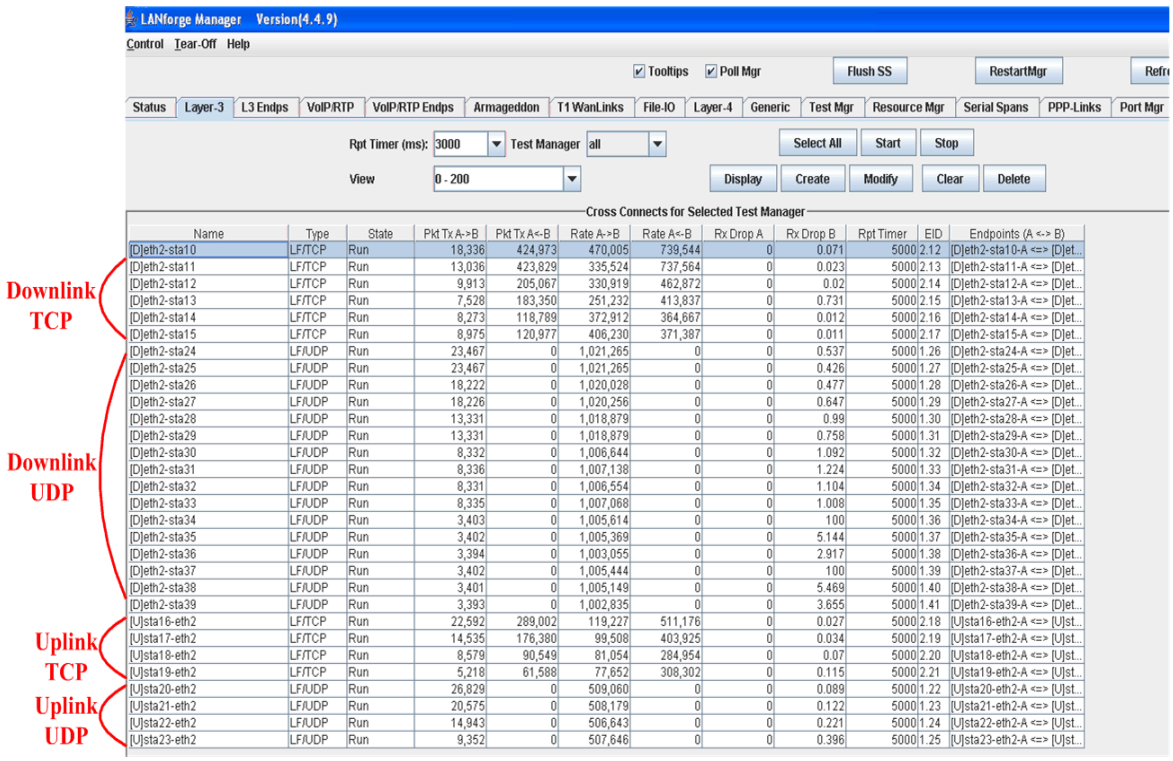


Figure 11-5 Screenshot of the background traffic configuration software

11.3 Simulation Test-bed Setup

11.3.1 Test Topology

The simulation test-bed for evaluating iPAS has been described in chapter 9, and the test topology was shown in Figure 9-1. The difference was that, the wireless network used in this chapter supported IEEE 802.11g instead of IEEE 802.11b.

11.3.2 Video Sequences

The simulation test used the same video content (*The Simpsons Movie*) with that of real life test, in order to avoid the impact of variable video content on the perceived quality. The video trace file was obtained from the TKU[84] website.

Table 11-5 shows the format information of the video trace files used. Three quality levels of the same video content were encoded with MPEG-4. The video encoded bit-rates are 1300 Kbps, 420 Kbps, 230 Kbps, for the high, medium, and low quality video clip, respectively. These bit-rates are consistent with that of used in real life test. Figure 11-6 presents part (first twenty frames) of the trace files relating to the three quality video sequences. Each trace file includes four types of frame information: frame number, frame type (I, P, or B), time, and frame length.

TABLE 11-5 VIDEO TRACE FILES USED FOR SIMULATION TEST-BED

Video Clip Quality Level	Codec	Mean Bit-rate (Kbps)	Number of Frames	Duration (minute)
High	MPEG-4	1300	30334	20
Medium	MPEG-4	420	30335	20
Low	MPEG-4	230	30335	20

[84] Trace file for *The Simpsons Movie*: http://www2.tkn.tu-berlin.de/research/trace/trace.html

Frame No.	Frametype	Time [ms]	Length [byte]
1	I	0	13599
2	P	120	1545
3	B	40	11929
4	B	80	832
5	P	240	716
6	B	160	567
7	B	200	555
8	P	360	3329
9	B	280	1214
10	B	320	1932
11	I	480	5179
12	B	400	2870
13	B	440	2969
14	P	600	5059
15	B	520	3818
16	B	560	3856
17	P	720	5399
18	B	640	3668
19	B	680	4016
20	P	840	5590

(a) Trace file for high quality video(1300Kbps)

Frame No.	Frametype	Time [ms]	Length [byte]
1	I	0	6885
2	P	120	503
3	B	40	3848
4	B	80	176
5	P	240	126
6	B	160	42
7	B	200	46
8	P	360	625
9	B	280	206
10	B	320	146
11	I	480	830
12	B	400	531
13	B	440	206
14	P	600	1150
15	B	520	745
16	B	560	686
17	P	720	1317
18	B	640	513
19	B	680	628
20	P	840	1259

(b) Trace file for medium quality video(420Kbps)

Frame No.	Frametype	Time [ms]	Length [byte]
1	I	0	6885
2	P	120	472
3	B	40	1219
4	B	80	240
5	P	240	207
6	B	160	86
7	B	200	79
8	P	360	501
9	B	280	195
10	B	320	120
11	I	480	830
12	B	400	329
13	B	440	78
14	P	600	854
15	B	520	428
16	B	560	369
17	P	720	942
18	B	640	334
19	B	680	333
20	P	840	888

(c) Trace file for low quality video(230Kbps)

Figure 11-6 Trace files for the MPEG-4 video clips used in simulation

11.3.3 Background Traffic Setup

The background traffic was generated according to Table 11-3 and Table 11-4, which were used for the real life test-bed. In real life test-bed, the available bandwidth estimated by iPerf and JPerf was in the range of 21Mbps of 23Mbps, which were selected as the estimated capacity of simulation test-bed as well. The video transmission time is set to 320s. During the first 20s, there is no background traffic. From 20s to 320s, the number of background flows increases from 6 to 30 with 6 new flows added every 60s. The background traffic consisted of UDP and TCP flows which were implemented by UDP and TCP agents provided by NS2, respectively. UDP agents were attached to the Constant Bit Rate (CBR) application and the TCP agents were attached to the File Transfer Protocol application modules, which are provided by NS2. As the transmission bit-rate of TCP flows was variable according to Table 11-3, the TCP sending rate was adjusted by changing the size of the receiving window.

11.4 Experimental Scenarios

The real test assesses the video delivery performance of the equal channel access mechanism of original IEEE 802.11 g protocol, while the simulation test evaluates the video transmission performance of iPAS over IEEE 802.11g. Therefore, there are four test cases (A, B, C, and D) designed for both real life and simulation tests. The test case configuration is listed in Table 11-6. The symbol “Y” is the abbreviation of “YES” meaning the video is delivered to certain device. For instance, in test case A, the high quality video sequence is delivered to the laptop, tablet, and smartphone. Similarly, in test case B and C, the medium quality video and low quality video sequences are delivered to the three devices. In test case D, laptop receives high quality video, tablet receives medium quality video, and smartphone receives low quality video. Each test case is repeated twice by delivering video using both HTTP/TCP and RTSP/RTP/UDP protocols, in order to analysis the impacts of streaming protocols on both equal channel access mechanism of IEEE 802.11g and iPAS over IEEE 802.11g.

TABLE 11-6 FOUR TEST CASES FOR THE VIDEO DELIVERY

Test case	Video quality	Device		
		Laptop	Tablet	Smartphone
A	High Quality	Y	Y	Y
B	Medium Quality	Y	Y	Y
C	Low Quality	Y	Y	Y
D	High Quality	Y	-	-
	Medium Quality	-	Y	-
	Low Quality	-	-	Y

11.5 Objective Video Quality Assessment

11.5.1 Objective Video Quality Metrics

PSNR is used to measure the delivered video quality in each test case. In the real life test-bed, the received video on each device is recorded and compared with the original video sequence using MSU Video Quality Measurement Tool. The quality degradation is mostly related to data loss. Figure 11-7 presents the frame loss effect in MSU tool. The MSU provides PSNR result visualizations as shown in Figure 11-8. Note that, high throughput does not guarantee better PSNR, however, in general the higher the throughput is, the better video quality can be received.

(a) Source video frame **(b) Destination video frame with loss**

Figure 11-7 An example of data loss effect in frame

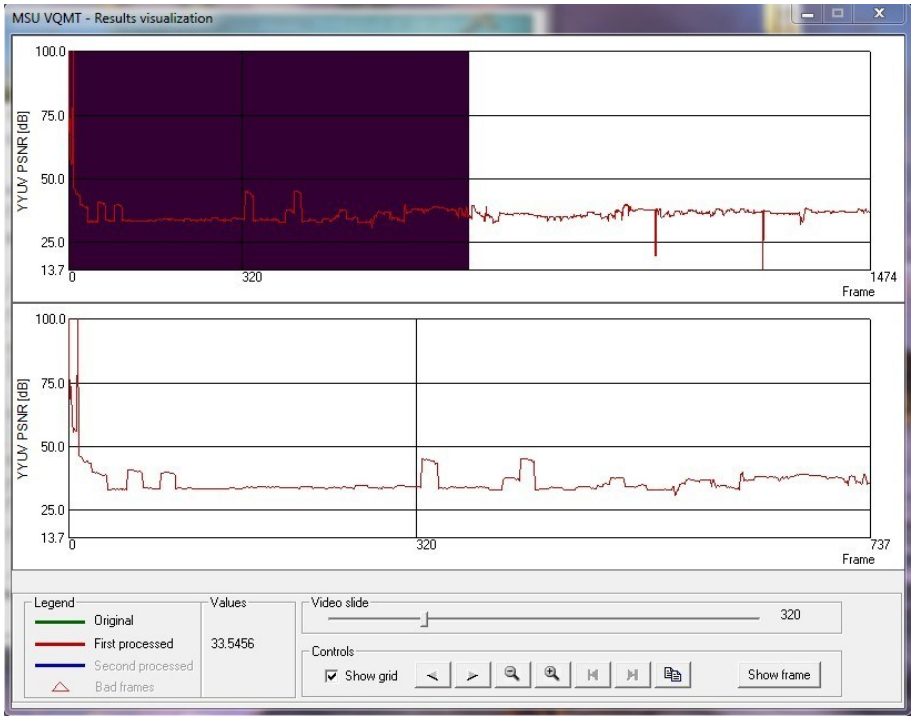


Figure 11-8 MSU-based PSNR result visualization

11.5.2 Results Analysis

Test cases A, B, C, and D described in Table 11-6 deliver high quality video, medium quality video, and low quality video to the laptop, tablet PC, and smartphone, respectively. Figure 11-9, Figure 11-10, Figure 11-11, and Figure 11-12 show the PSNR measured for test cases A, B, C, and D during video delivery. Additionally, details of the four figures are summarized in Table 11-7, Table 11-8, Table 11-9, and Tablet 11-10. The number of background stations is denoted by *N*, which increases from 0 to 30 with step of 6. Each test case lasts 320s.

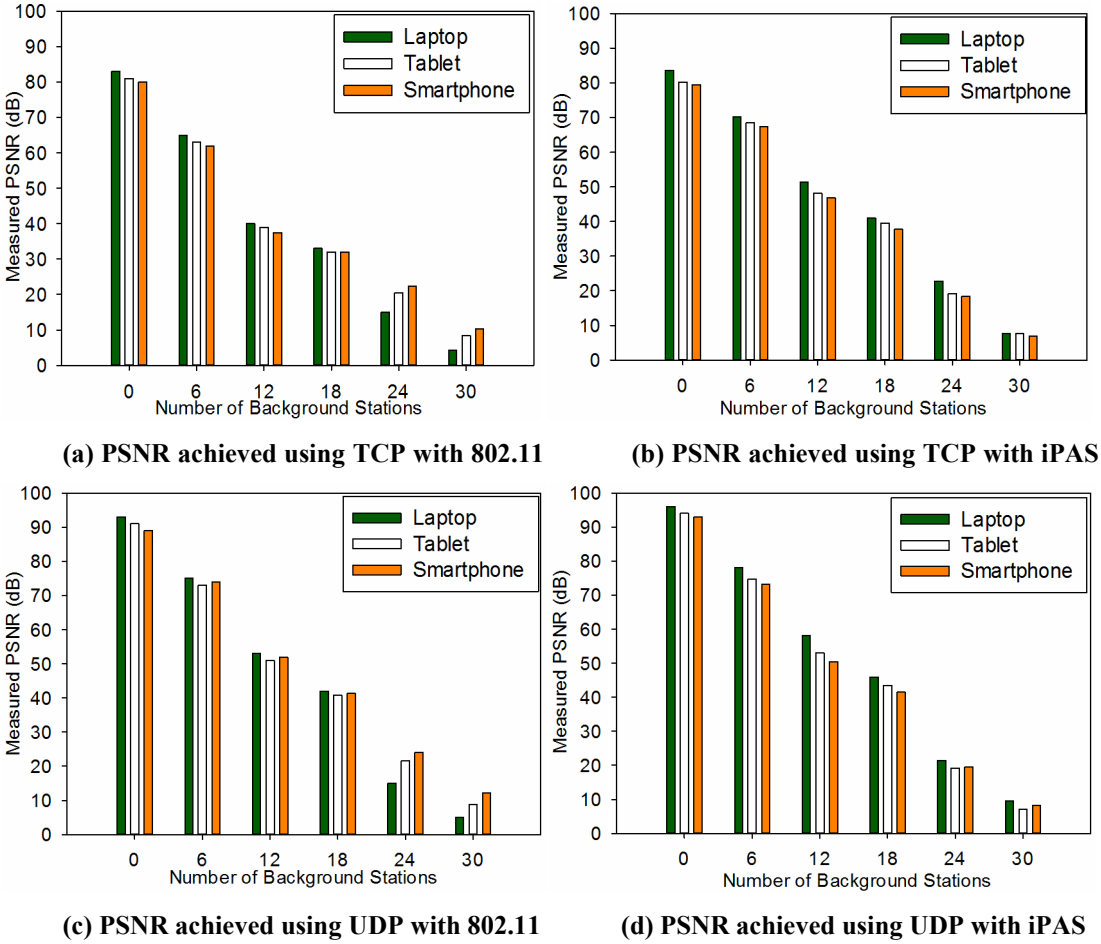
A. Test Case A

In test case A, the high quality video sequence (bitrate: 1300Kbps, frame rate: 30fps and resolution: 800x448) is delivered to the laptop, tablet PC, and smartphone, which are connected to the same wireless router. Figure 11-9 and Tablet 11-7 present the compared PSNR values measured between using IEEE 802.11 and iPAS in test case A.

Figure 11-9 (a) and Figure 11-9 (b) present the PSNR values measured during video delivery using TCP with IEEE 802.11 and iPAS, separately. In low loaded (N=0, 6, 12) and average loaded (N=18) network conditions, video delivered using iPAS has higher PSNR than that of IEEE 802.11. For instance, when N=12, PSNR values measured at laptop, tablet

PC, and smartphone using iPAS increase by 28.6%, 23.6%, and 25.1%, compared to the case of IEEE 802.11. The reason is that iPAS can adapt the transmission rate of video traffic based on the accurate bandwidth estimation algorithm, which efficiently reduces the packet loss probability and improves the received video quality. Additionally, in Fig.11-9 (a), under high loaded (N=24) and overloaded (N=30) network conditions, PSNR measured at the laptop is lower than that of tablet PC and smartphone. This can be explained by that laptop has more powerful data process capability than tablet and smartphone and requires more bandwidth to play the video. Since IEEE 802.11 allocates fair channel access for the three devices, laptop suffers more quality degradations, in comparison to the tablet PC and smartphone. Figure 11-9 (b) presents the PSNR measured at the three devices using TCP with iPAS. It is shown that, in comparison with the equal channel access mechanism of IEEE 802.11, iPAS over IEEE 802.11improves the PSNR at the laptop by re-allocating certain bandwidth share from the tablet PC and smartphone: 1) when N=24, PSNR value measured at the laptop using iPAS increases by 52% and PSNR values measured at the tablet PC and smartphone using iPAS decrease by 6.8% and 17.9%; 2) when N=30, PSNR value measured at laptop using iPAS increases by 76.7% and PSNR measured at the tablet PC and smartphone using iPAS decrease by 8.3% and 33%, respectively.

Figure 11-9 (c) and Figure 11-9 (d) illustrate the PSNR values measured when UDP video is delivered via the equal channel access mechanism of IEEE 802.11 and iPAS over IEEE 802.11, separately. It is shown that, generally, UDP traffic can result in higher PSNR than when TCP is used. The primary reason is that TCP uses flow control which causes retransmission delay and thus degraded the video quality. Also, TCP protocol has much higher overhead than that of UDP. For instance, when N=18, PSNR values measured at the laptop, tablet PC, and smartphone using UDP with IEEE 802.11 increase by 27.3%, 27.5%, and 29.1%, in comparison with those of using TCP with IEEE 802.11. Additionally, similar with TCP traffic as shown in Figure 11-9 (a) Figure 11-9 (b), UDP video traffic delivered using iPAS has higher PSNR than that of IEEE 802.11. Figure 11-9 (d) presents the PSNR measured at the three devices when delivering UDP traffic with iPAS. It is shown that, in comparison with IEEE 802.11, iPAS improves the PSNR measured at the laptop by dividing certain bandwidth share from the tablet PC and smartphone: 1) when N=24, PSNR value measured at laptop using iPAS increases by 43.3% and PSNR measured at the tablet PC and smartphone using iPAS decrease by 11.1% and 18.3%; 2) when N=30, PSNR value measured at laptop using iPAS increases by 88.2% and PSNR measured at the tablet PC and smartphone using iPAS decrease by 19.3% and 31.4%, respectively.



(a) PSNR achieved using TCP with 802.11 **(b) PSNR achieved using TCP with iPAS**

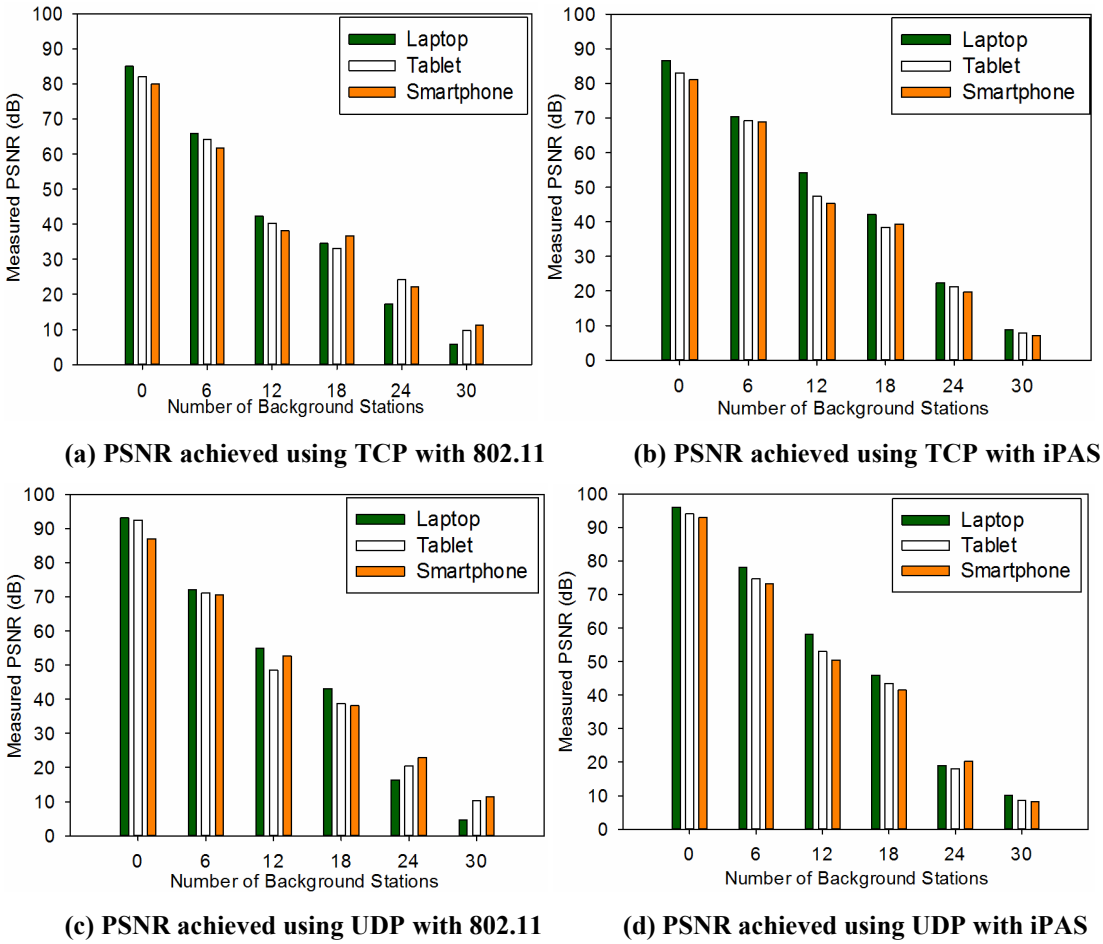
(c) PSNR achieved using UDP with 802.11 **(d) PSNR achieved using UDP with iPAS**

Figure 11-9 Comparison of PSNR between 802.11 and iPAS in test case A

B. Test Case B

In test case B, the medium quality video sequence (bitrate: 420Kbps, frame rate: 25fps and resolution: 512x288) is delivered to laptop, tablet, and smartphone. Figure 11-10 and Table 11-8 present the PSNR values measured when using the equal channel access mechanism of IEEE 802.11 and iPAS over IEEE 802.11 in test case B.

Figure 11-10 (a) and Figure 11-10 (b) present the PSNR values measured during video delivery using TCP with IEEE 802.11 and iPAS, separately. Similar with test case A, under low loaded (N=0, 6, 12) and average loaded (N=18) network conditions, video delivered using iPAS has higher PSNR than that of IEEE 802.11, due to the accurate bandwidth estimation algorithm used. For instance, when N=12, PSNR values measured at the laptop, tablet PC, and smartphone using iPAS increase by 28.4%, 17.6%, and 19.2%, compared to the case of IEEE 802.11. Additionally, in Figure 11-10 (a), under high loaded



(a) PSNR achieved using TCP with 802.11

(b) PSNR achieved using TCP with iPAS

(c) PSNR achieved using UDP with 802.11

(d) PSNR achieved using UDP with iPAS

Figure 11-10 Comparison of PSNR between 802.11 and iPAS in test case B

(*N*=24) and overloaded (*N*=30) network conditions, PSNR measured at the laptop is lower than that of tablet PC and smartphone. Figure 11-10 (b) presents the PSNR measured at the three devices using iPAS. It is shown that, in comparison with IEEE 802.11g, iPAS improves the PSNR measured at the laptop by dividing certain bandwidth from the tablet PC and smartphone: 1) when *N*=24, PSNR value measured at the laptop using iPAS increases by 29.7% and PSNR measured at the tablet PC and smartphone using iPAS decrease by 12% and 10.9; 2) when *N*=30, PSNR value measured at the laptop using iPAS increases by 54.4% and PSNR measured at the tablet PC and smartphone using iPAS decrease by 20.4% and 36.6%, respectively.

Figure 11-10 (c) and Figure 11-10 (d) illustrate the PSNR values measured during video delivered using UDP with IEEE 802.11 and iPAS, separately. Similar with the test case A, video over UDP can result in higher PSNR than that when TCP is employed. For instance, when *N*=18, PSNR values measured at the laptop, tablet PC, and smartphone using

UDP with IEEE 802.11 increase by 24.9%, 16.9%, and 3.8%, in comparison with that of video using TCP over IEEE 802.11. Additionally, similar with video over TCP as shown in Figure 11-10 (a) and Figure 11-10 (b), UDP-based video traffic delivered using iPAS has higher PSNR than that of IEEE 802.11. Figure 11-10 (d) presents the video PSNR measured at the three devices using UDP with iPAS. It is shown that, in comparison with IEEE 802.11, iPAS improves the PSNR measured at the laptop by dividing certain bandwidth shaer from the tablet PC and smartphone: 1) when N=24, PSNR value measured at the laptop using iPAS increases by 15.9% and PSNR measured at the tablet PC and smartphone using iPAS decrease by 11.7% and 11.4%; 2) when N=30, PSNR value measured at the laptop using iPAS increases by 94.2% and PSNR measured at the tablet PC and smartphone using iPAS decrease by 16.5% and 28.7%, respectively.

TABLE 11-7 PSNR MEASURED WITH 802.11 AND iPAS IN TEST CASE A

Load	Number of stations	TCP						UDP					
		Average PSNR of 802.11 (dB)			Average PSNR of iPAS (dB)			Average PSNR of 802.11 (dB)			Average PSNR of iPAS (dB)		
		Laptop	Tablet	Smart phone	Laptop	Tablet	Smart phone	Laptop	Tablet	Smart phone	Laptop	Tablet	Smart phone
Low loaded	0	83.0	81.0	80.0	83.6	80.2	79.5	93.0	91.0	89.0	96.0	94.0	93.0
	6	65.0	63.0	62.0	70.1	68.4	67.3	75.0	73.0	74.0	78.0	74.6	73.2
	12	40.0	39.0	37.5	51.3	48.2	46.9	53.0	51.0	52.0	58.2	53.1	50.5
Average loaded	18	33.0	32.0	32.0	41.0	39.5	37.7	42.0	40.8	41.3	46.0	43.5	41.5
High loaded	24	15.0	20.5	22.3	22.8	19.1	18.3	15.0	21.6	24.0	21.5	19.2	19.6
Over loaded	30	4.3	8.4	10.3	7.6	7.7	6.9	5.1	8.8	12.1	9.6	7.1	8.3

TABLE 11-8 PSNR MEASURED WITH 802.11 AND iPAS IN TEST CASE B

Load	Number of stations	TCP						UDP					
		Average PSNR of 802.11 (dB)			Average PSNR of iPAS (dB)			Average PSNR of 802.11 (dB)			Average PSNR of iPAS (dB)		
		Laptop	Tablet	Smart phone	Laptop	Tablet	Smart phone	Laptop	Tablet	Smart phone	Laptop	Tablet	Smart phone
Low loaded	0	85.0	82.0	80.0	86.6	83	81.1	93.1	92.4	87.0	96.0	94.0	93.0
	6	65.8	64.2	61.7	70.3	69.2	68.9	72.0	71.2	70.5	78.0	74.6	73.2
	12	42.2	40.3	38.1	54.2	47.4	45.4	55.0	48.6	52.6	58.2	53.1	50.5
Average loaded	18	34.5	33.1	36.7	42.1	38.4	39.3	43.1	38.7	38.1	46.0	43.5	41.5
High loaded	24	17.2	24.2	22.1	22.3	21.3	19.7	16.4	20.5	22.9	19.0	18.1	20.3
Over loaded	30	5.7	9.8	11.2	8.8	7.8	7.1	5.2	10.3	11.5	10.1	8.6	8.2

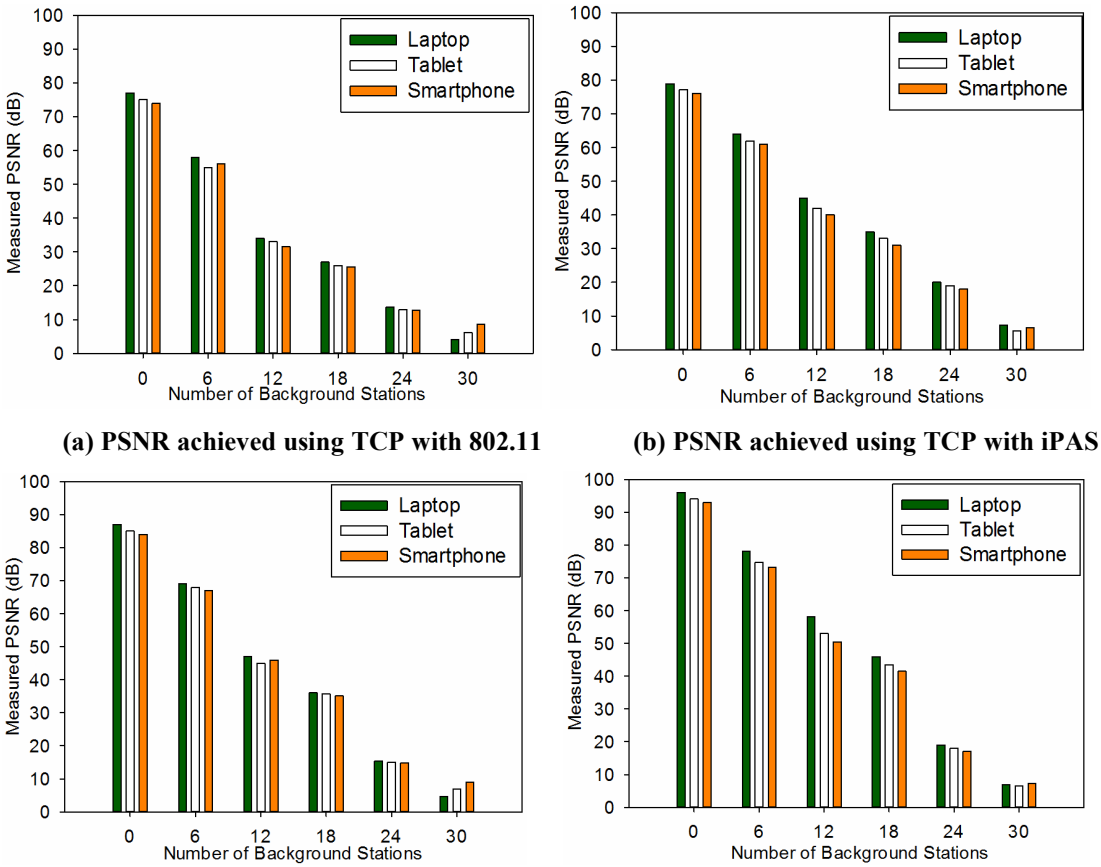
C. Test Case C

In test case C, the low quality video sequence (bitrate: 230Kbps, frame rate: 20fps and resolution: 320x176) is delivered to the laptop, tablet PC, and smartphone. Figure 11-11 and Tablet 11-9 present the PSNR values measured when using the equal channel access mechanism of IEEE 802.11 and iPAS over IEEE 802.11 in test case C.

Figure 11-11 (a) and Figure 11-11 (b) present the PSNR values achieved during video delivery using TCP with IEEE 802.11 and iPAS, separately. Similar with test cases A and B, under low loaded (N=0, 6, 12) average loaded (N=18), and high loaded (N=24) network conditions, video delivered using iPAS has higher PSNR than that of IEEE 802.11, due to the accurate bandwidth estimation algorithm used and iPAS intelligence. For instance, when N=12, PSNR values measured at the laptop, tablet PC, and smartphone using iPAS increase by 32.4%, 27.3%, and 27%, compared to the case of IEEE 802.11. Additionally, in Fig.11-11 (a), overloaded (N=30) network conditions, PSNR measured at the laptop is lower than that of tablet PC and smartphone. Figure 11-11 (b) presents the PSNR measured at the three devices using TCP traffic with iPAS. This shows that, in comparison with IEEE 802.11g, iPAS improves the PSNR of laptop by dividing certain bandwidth share from the tablet PC and smartphone: 1) when N=24, PSNR value measured at the laptop increases by 46.2% and PSNR measured at the tablet PC and smartphone decrease by 11.6% and 20.5%; 2) when N=30, PSNR value measured at the laptop using iPAS increases by 75.6% and PSNR measured at the tablet PC and smartphone decrease by 9.7% and 23.5%, respectively.

Figure 11-11 (c) and Figure 11-11 (d) illustrate the PSNR values measured during delivering video using UDP with IEEE 802.11 and iPAS, separately. Similar with test cases A and B, UDP traffic can result in higher PSNR than that of TCP traffic. For instance, when N=18, PSNR values measured at the laptop, tablet PC, and smartphone using UDP with IEEE 802.11 increase by 21.2%, 22.8%, and 19.7%, in comparison with those of using TCP with IEEE 802.11. Additionally, similar with TCP traffic as shown in Figure 11-11 (a) and Figure 11-11 (b), UDP video traffic delivered using iPAS has higher PSNR than that of IEEE 802.11. Figure 11-11 (d) presents the PSNR measured at the three devices using UDP with iPAS. This shows that, in comparison with IEEE 802.11, iPAS improves the PSNR measured at the laptop by dividing certain bandwidth share from the tablet PC and smartphone: 1) when N=24, PSNR value measured at the laptop using iPAS increases by 36.1% and PSNR measured at the tablet PC and smartphone decrease by 11.9% and 14.7%;

2) when N=30, PSNR value measured at the laptop using iPAS increases by 66.5% and PSNR measured at the tablet PC and smartphone decrease by 18.4% and 29.3%, respectively.



(a) PSNR achieved using TCP with 802.11

(b) PSNR achieved using TCP with iPAS

(c) PSNR achieved using UDP with 802.11

(d) PSNR achieved using UDP with iPAS

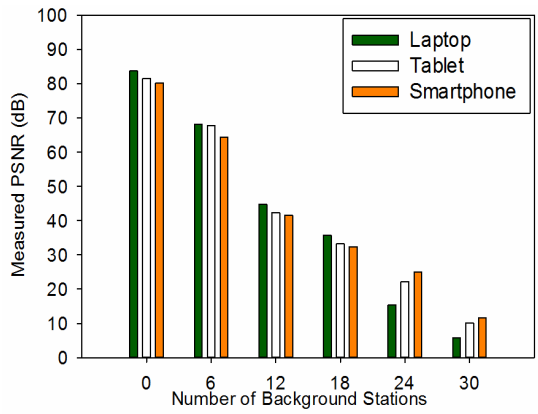
Figure 11-11 Comparison of PSNR between 802.11 and iPAS in test case C

D. Test Case D

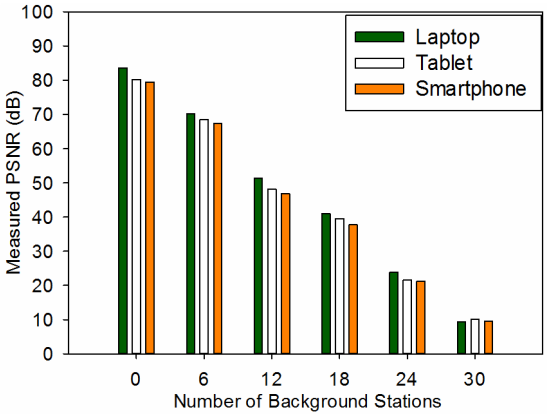
In test case D, the high quality video sequence (bitrate: 1300Kbps, frame rate: 30fps and resolution: 800x448) is delivered to the laptop, the medium quality video sequence (bitrate: 420Kbps, frame rate: 25fps and resolution: 512x288) is delivered to the tablet, and the low quality video sequence (bitrate: 230Kbps, frame rate: 20 and resolution: 320x176) is delivered to the smartphone. Figure 11-12 and Tablet 11-10 show the PSNR values measured when using the equal channel access mechanism of IEEE 802.11 and iPAS over IEEE 802.11 in test case D.

Figure 11-12 (a) and Figure 11-12 (b) present the PSNR values measured during video delivery using TCP with IEEE 802.11 and iPAS, separately. Similar with test case A

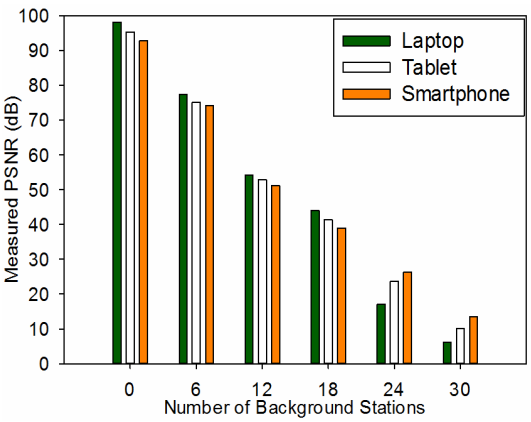
and B, under low loaded (*N*=0, 6, 12) and average loaded (*N*=18) network conditions, video delivered using iPAS has higher PSNR than that of IEEE 802.11, due to the accurate bandwidth estimation algorithm used and iPAS intelligence. For instance, when N=12, PSNR values measured at the laptop, tablet PC, and smartphone using iPAS increase by 20.3%, 20.9%, and 21.1%, compared to the case of IEEE 802.11g. Additionally, in Fig.11-12 (a), under high loaded (N=24) and overloaded (N=30) network conditions, PSNR measured at the laptop is lower than that of tablet PC and smartphone. Figure 11-12 (b) showsthe PSNR measured at the three devices using TCP with iPAS. It is shown that, in comparison with IEEE 802.11g, iPAS improves the PSNR measured at the laptop by dividing certain bandwidth share from the tablet PC and smartphone: 1) when N=24, PSNR value measured at the laptop increases by 54.2% and PSNR measured at the tablet PC and smartphone decrease by 4.6% and 24.2%; 2) when N=30, PSNR measured at the laptop increases by 59.8 % and PSNR measured at the tablet PC and smartphone decrease by 5.9% and 19.5%, respectively.



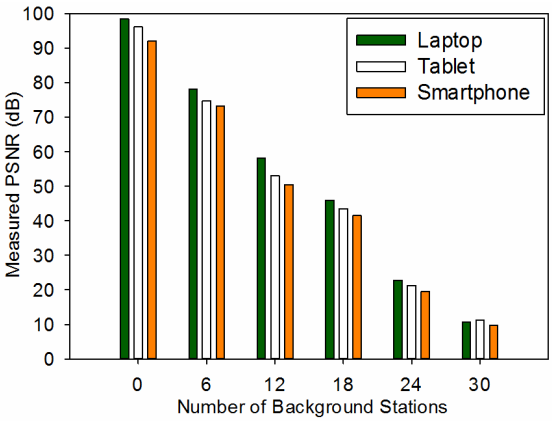
(a) PSNR achieved using TCP with 802.11



(b) PSNR achieved using TCP with iPAS



(c) PSNR achieved using UDP with 802.11



(d) PSNR achieved using UDP with iPAS

Figure 11-12 Comparison of PSNR between 802.11 and iPAS in test case D

Figure 11-12 (c) and Figure 11-12 (d) illustrate the PSNR values measured during video delivery when using UDP with IEEE 802.11 and iPAS, separately. Similar with test cases A, B, and C, UDP traffic can result in higher PSNR than that of TCP traffic. For instance, when *N*=18, PSNR values measured at the laptop, tablet PC, and smartphone using UDP with IEEE 802.11 increase by 18.7%, 19.4%, and 19.1%, in comparison with that of using TCP with IEEE 802.11. Additionally, similar with the TCP case as shown in Figure 11-12 (a) and Figure 11-12 (b), UDP video delivered using iPAS has higher PSNR than that of IEEE 802.11. Figure 11-11 (d) presents the PSNR measured at the three devices using UDP with iPAS. It is shown that, in comparison with IEEE 802.11, iPAS improves the PSNR measured at the laptop by dividing certain bandwidth share from the tablet PC and smartphone: 1) when N=24, PSNR value measured at the laptop using iPAS increases by 40.3% and PSNR measured at the tablet PC and smartphone using iPAS decrease by 15.9% and 18.7%; 2) when N=30, PSNR value measured at the laptop using iPAS increases by 64.1% and PSNR measured at the tablet PC and smartphone using iPAS decrease by 8.4% and 25.6%, respectively.

In test cases A, B, C, and D, when the laptop experiences significantly quality degradation under high traffic load conditions, some bandwidth resources allocated for the tablet PC and smartphone are transferred to the laptop by iPAS, in order to provide a fair video delivery. However, PSNR cannot reflect the impact of reduced PSNR on user's perception of the received video. Consequently, subjective video quality assessment is performed in next section.

TABLE 11-9 PSNR MEASURED WITH 802.11 AND IPAS IN TEST CASE C

Load	Number of stations	TCP						UDP					
		Average PSNR of 802.11 (dB)			Average PSNR of iPAS (dB)			Average PSNR of 802.11 (dB)			Average PSNR of iPAS (dB)		
		Laptop	Tablet	Smart phone	Laptop	Tablet	Smart phone	Laptop	Tablet	Smart phone	Laptop	Tablet	Smart phone
Low loaded	0	77.0	75.0	74.0	78.9	77.1	76.1	87.0	85.0	84.0	96.0	94.0	93.0
	6	58.0	55.0	56.0	64.0	62.0	61.0	69.0	68.0	67.0	78.0	74.6	73.2
	12	34.0	33.0	31.5	45.0	42.0	40.0	47.0	45.0	46.0	58.2	53.1	50.5
Average loaded	18	27.0	26.0	25.6	35.0	33.0	31.0	36.0	35.7	35.1	46.0	43.5	41.5
High loaded	24	13.7	12.7	12.5	20.0	18.2.0	18.0	15.3	15.0	14.8	19.0	18.0	17.0
Over loaded	30	4.1	6.2	8.5	7.2	5.6	6.5	4.6	6.9	8.9	6.9	6.5	7.3

TABLE 11-10 PSNR MEASURED WITH 802.11 AND IPAS IN TEST CASE D

Load	Number of stations	TCP						UDP					
		Average PSNR of 802.11 (dB)			Average PSNR of iPAS (dB)			Average PSNR of 802.11 (dB)			Average PSNR of iPAS (dB)		
		Laptop	Tablet	Smart phone	Laptop	Tablet	Smart phone	Laptop	Tablet	Smart phone	Laptop	Tablet	Smart phone
Low loaded	0	83.7	81.4	80.1	83.6	80.2	79.5	98.1	95.3	92.8	98.4	96.1	92.0
	6	68.1	67.7	64.3	70.1	68.4	67.3	77.4	75.1	74.2	78.0	74.6	73.2
	12	44.7	42.3	41.5	51.3	48.2	46.9	54.1	52.8	51.1	58.2	53.1	50.5
Average loaded	18	35.7	33.2	32.4	41.0	39.5	37.7	44.0	41.3	39.0	46.0	43.5	41.5
High loaded	24	15.3	22.1	24.9	23.8	21.6	21.3	17.1	23.7	26.2	22.8	21.3	19.5
Over loaded	30	5.7	10.5	11.7	9.4	10.1	9.6	6.1	10.1	13.5	10.6	11.2	9.8

11.6 Subjective Video Quality Assessment

Although PSNR is one of the most widely used video quality metrics, however, PSNR values do not correlate perfectly with human perceived visual quality since the human visual system behaves non-linearly. In this section, the performance of iPAS is studied by comparing it with that of the original IEEE 802.11 protocol using subjective video quality assessment. ITU-T Rec.P.911 [53] is selected for measuring the subjective video quality. The same video (*The Simpsons Movie*) is transmitted to three devices (laptop, tablet, and smartphone) using two schemes separately: 1) the equal channel access mechanism of IEEE 802.11g protocol; 2) IEEE 802.11g with iPAS. The delivered video clips are obtained based on the four test cases which are the same with those described in the objective video quality assessment in section 10.1.

11.6.1 Subjective Test Setup

The subjective test was conducted in a separate room in the lab without any outside disturbance. 32 users were invited to complete four video quality test cases and each test case was finished by 8 users. The users included 18 men and 14 women who worked in various areas, e.g. students, workers, scientists, engineers, etc. All users were subjected to a pre-test in order to be familiar with the operations. Each user was asked to watch unique video sequences taken from *The Simpsons Movie* on the three devices separately. Users then rated the quality of the video based on the following metrics: Mean Opinion Score (MOS), continuity, blurring, and blockiness. The questionnaire was presented to the subjects on paper and was given to the users before the tests.

Note that, the number of users was selected by considering both testing costs and result accuracy. A large number of users can provide good average values, however, the testing cost also increases significantly. Based on our study in each test case, the average MOS value computed on the results provided by 8 users was not very much different to that computed for 9 users. Consequently, eight users were invited for each test case. The quality scale for MOS is presented in Table 2-10 in chapter 2, where the value 5 indicates the excellent quality and 1 describes the bad quality. Appendix-A illustrates the details of the test instructions and the questionnaire.

The following rules have been setup based on the experimental design guidelines specified in the ITU-T Rec.P.911.

1. Video content presented to any user will never repeat to the same user in order to avoid viewer boredom.
2. The order of devices held by users is changed in order to give a random sample. For instance, the first user views content in order on the laptop, tablet PC, and smartphone; the second user on the tablet PC, smartphone, and laptop; the third user on the smartphone, laptop, and tablet, etc.
3. Any one person is assigned an unique test case in order to avoid the effects of video quality variation on MOS. Considering the case where *User-1* first takes test case A and then takes test case B, *User-1* watches the same video clip with different qualities. This would affect the user mean opinion score assessment for test case B.
4. Each test case is taken by many users. The purpose of inviting more users to take each test case is to get better average values.

Based on the above regulations, video clips obtained in both real test and simulations are evaluated by users. Different parts of the movie have been delivered using the two schemes based on

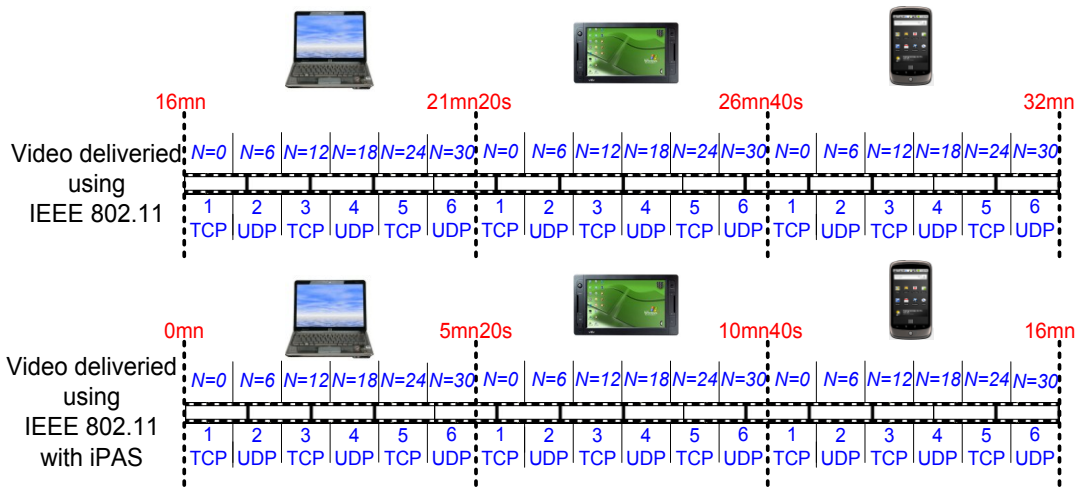


Figure 11-13 Video clips presented for users

the experimental scenarios in section 10.1.7. Figure 11-13 illustrates an example of what video clips one user watches for a certain test case. A 32 minutes long video is cut and delivered as follows: 1) Sequence from t=0 minutes to 16 minutes of the movie is delivered using IEEE 802.11g with iPAS in simulation test bed; 2) t=15 minutes to t=31minutes of the movie is delivered using the original IEEE 802.11g in the real test bed. Considering *User-1* is assigned to take test case A, in which the high quality video is delivered to the laptop,

tablet PC, and smartphone. There are six steps to be completed by *User-1*. The test lasts around 30minutes including video play-back time and question answering time.

1) *User-1* watches the video delivered using **IEEE 802.11** on the **laptop**. The video starts from 15mins and lasts 320s. According to the background traffic setup described in section 10.1.6, there is no background traffic during the first 20 seconds. Starting from t=20 seconds, the number of background stations increases with steps of 6 for every 60seconds. Consequently, the 320s long video clip is further divided into six shorter clips: the first clip is 20s long and the remaining five clips are 60s each. Each short clip is related to a certain background traffic condition. Additionally, these clips are delivered using either TCP or UDP. As shown in Figure 11-13, for instance, video clips 1, 3, and 5 are transmitted using TCP and video clips 2, 4, and 6 are transmitted using UDP. As User-1 watches each of the six clips, the questionnaire asks the user to indicate the video quality in terms of MOS, blurriness, and continuity, etc.
2) *User-1* watches the video delivered using **IEEE 802.11** on the **tablet PC**. The procedures are similar with those in step 1. The difference is that the video content is changed to avoid boring the user.
3) *User-1* watches the video delivered using **IEEE 802.11** on the **smartphone**. The procedures are similar with that in step 1 and step 2. The difference is that the video content is changed to avoid boring the user.
4) *User-1* watches the video delivered using **IEEE 802.11 with iPAS** on the **laptop**. The procedures are similar with that in step 1. The difference is that the video content is changed to avoid boring the user.
5) *User-1* watches the video delivered using **IEEE 802.11 with iPAS** on the **tablet PC**. The procedures are similar with that in step 1. The difference is that the video content is changed to avoid boring the user.
6) *User-1* watches the video delivered using **IEEE 802.11 with iPAS** on the **smartphone**. The procedures are similar with that in step 1. The difference is that the video content is changed to avoid boring the user.

Based on the test rules, seven additional users will be invited to complete test case A. Each of the users follows the similar steps as finished by *User-1*. The difference is the order of devices selected (refers to 2nd rule). For users who are invited to finish test cases B, C, and D, the test instructions are the same with those in the test case A.

11.6.2 Result Analysis

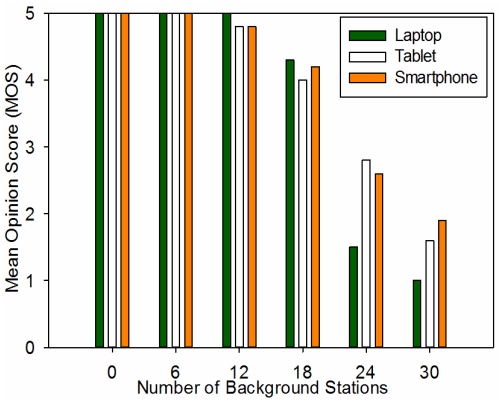
Test cases A, B, C, and D deliver the high, medium, and low quality video sequences to the laptop, tablet PC, and smartphone. Figure 11-14, Figure 11-15, Figure 11-16, and Figure 11-17 show the MOS score rated by users for test cases A, B, C, and D, respectively. Additionally, details of the four figures are summarized in Table 11-11, Table 11-12, Table 11-13, and Table 11-14. The number of background stations is denoted by *N*, which increases from 0 to 30 in steps of 6. Each test case is completed by eight users and the average values of MOS are computed.

A. Test Case A

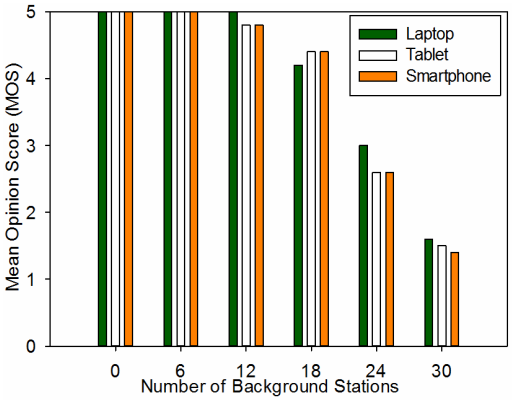
In test case A, the high quality video sequence (bitrate: 1300Kbps, frame rate: 30fps and resolution: 800x448) is delivered to the laptop, tablet PC, and smartphone, which are connected to the same wireless router. Figure 11-14 and Tablet 11-11 present the MOS values when using IEEE 802.11 and iPAS in test case A, respectively.

Figure 11-14 (a) and Figure 11-14 (b) present the MOS values achieved when delivering video on top of TCP with IEEE 802.11 and iPAS, separately. Under low loaded (*N*=0, 6, 12) and average loaded (*N*=18) network conditions: 1) the video quality delivered at the three devices are all above 4.0 (good quality); 2) there is no significant difference between the video delivered using IEEE 802.11 and iPAS in terms of MOS. In the case of sending TCP with IEEE 802.11 as presented in Figure 11-14 (a), under high loaded (*N*=24) and overloaded (*N*=30) network conditions, the MOS values measured at the laptop are significantly lower than those measured at the tablet PC and smartphone. For instance, when *N*=24, MOS at the laptop decreases 0.35by 86.7% and 73.3%, in comparison with the tablet PC and smartphone, separately. This phenomenon is similar with the PSNR values measured for video over TCP with IEEE 802.11, which are shown in Figure 11-9 (a). The reason is that the laptop requires more bandwidth than the tablet PC and smartphone, due to the more powerful processing ability and larger screen size. Figure 11-14 (b) presents the MOS measured at the three devices using TCP traffic with iPAS. It is shown that, in comparison with IEEE 802.11, iPAS improves the MOS of the video delivered to the laptop by reducing certain bandwidth share allocated to the tablet PC and smartphone: 1) when *N*=24, video MOS measured at the laptop using iPAS increases from 1.5 (bad quality) to 3.1 (fair quality), video MOS measured at the tablet PC decreases from 2.8 (poor to fair quality) to 2.6 (poor to fair quality), and video MOS measured at the smartphone stays at 2.6 (poor to fair quality); 2)

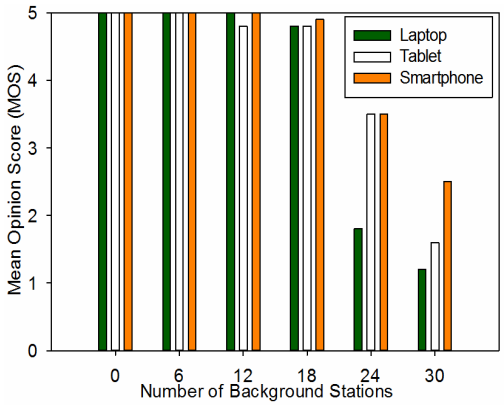
when *N*=30, video MOS value measured at the laptop using iPAS increase from 1.0 (bad quality) to 1.6 (bad to poor quality), MOS measured at the tablet PC decreases from 1.6 (bad to poor quality) to 1.5 (bad quality), and MOS measured at the smartphone decreases from 1.8 (bad to poor quality) to 1.4 (bad quality).



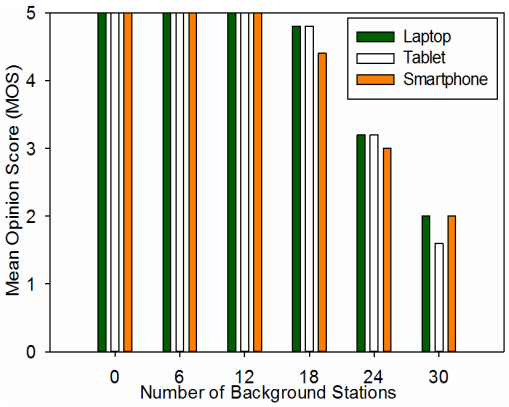
(a) MOS for video over TCP with 802.11



(b) MOS for video over TCP with iPAS



(c) MOS for video over UDP with 802.11



(d) MOS for video over UDP with iPAS

Figure 11-14 Comparison of PSNR between 802.11 and iPAS in test case A

Figure 11-14 (c) and Figure 11-14 (d) present the MOS values measure when delivering video using UDP with IEEE 802.11 and iPAS, separately. In low loaded (*N*=0, 6, 12) and average loaded (*N*=18) network conditions: 1) the video quality delivered at the three devices are all above 4.5 (good quality); 2) there is no significant difference between the video delivered using iPAS and IEEE 802.11 in terms of MOS. Similar with the TCP case, video delivering using UDP with IEEE 802.11 also results in significantly video quality degradation at the laptop in high loaded network conditions. Take *N*=24 for instance, as shown in Figure11-14 (c), MOS of video delivered to the laptop decreases by 15.4% and 38.9%, in comparison with the tablet PC and smartphone. By comparing Figure 11-14 (c) and Figure 11-14 (d), iPAS improves the video quality at the laptop by reducing the

bandwidth allocation for tablet and smartphone: 1) when *N*=24, video MOS measured at the laptop using iPAS increases from 1.8 (bad to poor quality) to 3.2 (fair quality), MOS measured at the tablet PC decreases from 3.3 (fair quality) to 3.2 (fair quality), and MOS measured at the smartphone decreases from 3.5 (fair quality) to 3.0 (fair quality); 2) when *N*=30, MOS measured at the laptop using iPAS increases from 1.2 (bad quality) to 2.0 (poor quality), MOS measured at the tablet PC decreases from 1.6 (bad to poor quality) to 1.5 (bad quality), and MOS measured at the smartphone decreases from 2.5 (poor quality) to 2.0 (poor quality).

It can be concluded based on test case A results that, the reduction of bandwidth allocation for the tablet PC and smartphone according to iPAS can result in the following benefits, in comparison with using IEEE 802.11: 1) there is none or little impact in terms of video quality degradation for the video clips played on the tablet and smartphone; 2) there is an increased video quality for the video clips played on the laptop. For instance, when *N*=24, the MOS values at the laptop using TCP and UDP with iPAS increase by 106% and 77.8%, in comparison with those when IEEE 802.11 is used; 3) there is a decrease in standard deviation in the average MOS for laptop, tablet PC, and smartphone. For instance, when *N*=24, the standard deviation in MOS using TCP and UDP with iPAS decreases by 57.1% and 87.8%, separately, in comparison with the values of IEEE 802.11.

B. Test Case B

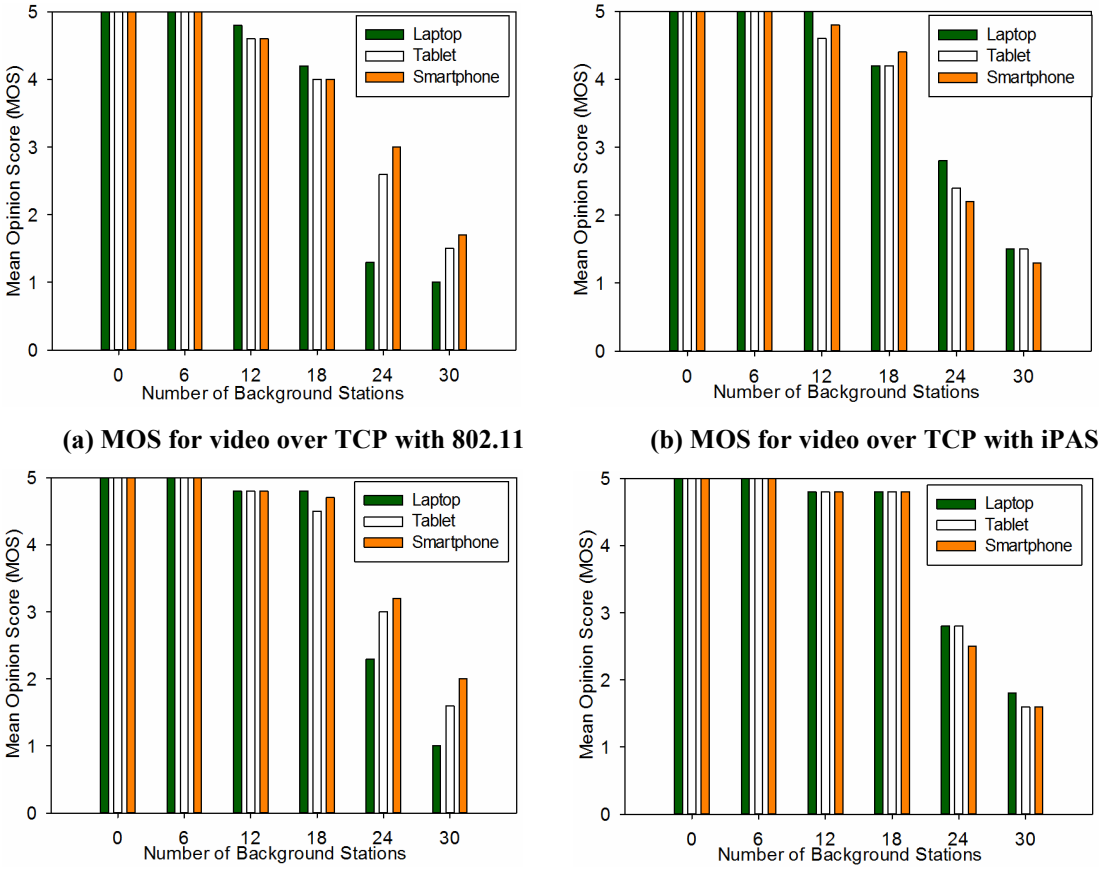
In test case B, the medium quality video sequence (bitrate: 420Kbps, frame rate: 25 fps and resolution: 512x288) is delivered to the laptop, tablet PC, and smartphone. Figure 11-10 and Tablet 11-12 present the MOS values.

Figure 11-15 (a) and Figure 11-15 (b) present the MOS values measured during video delivery using TCP with IEEE 802.11 and iPAS, respectively. Generally, the MOS values measured at the three devices are lower when *N*>12, in comparison with test case A. Since the quality of the video sequence in test case B is decreased. For instance, when *N*=12, the MOS values measured at the laptop, tablet PC, and smartphone using TCP with IEEE 802.11 decrease by 4%, 6.5%, and 6.3%, in comparison with test case A. In the case of video delivery using TCP with IEEE 802.11 and iPAS, in low loaded (*N*=0, 6, 12) and average loaded (*N*=18) network conditions: 1) the video quality delivered at the three devices are all above 4.0 (good quality); 2) there is no significant difference between the video delivered using IEEE 802.11 and iPAS in terms of MOS. In the case of TCP traffic with

IEEE 802.11 as presented in Figure 11-15 (a), in high loaded (N=24) and overloaded (N=30) network conditions, the MOS values measured at the laptop are significantly lower than that of the tablet PC and smartphone. For instance, when N=24, MOS measured at the laptop decreases by 66.7% and 80%, in comparison with the tablet PC and smartphone, respectively. This phenomenon is similar with the PSNR values measured using TCP with IEEE 802.11, which is shown in Figure 11-10 (a). The reason is that the laptop requires more bandwidth than those of the tablet PC and smartphone. Figure 11-15 (b) presents the MOS measured at the three devices using TCP with iPAS. It is shown that, in comparison with IEEE 802.11, iPAS improves the MOS of the video clips played on the laptop by reducing certain bandwidth share allocated for the tablet PC and smartphone: 1) when N=24, MOS measured at the laptop using iPAS increases from 1.3 (bad quality) to 2.8 (poor to fair quality), MOS measured at the tablet PC decreases from 2.6 (poor to fair quality) to 2.4 (poor quality), and MOS measured at the smartphone decreases from 3.0 (fair quality) to 2.2 (poor quality); 2) when N=30, MOS value measured at the laptop using iPAS increase from 1.0 (bad quality) to 1.5 (bad quality), MOS measured at the tablet PC stays with 1.5 (bad quality), and MOS measured at smartphone decreases from 1.7 (bad quality) to 1.2 (bad quality).

Figure 11-15 (c) and Figure 11-15 (d) present the MOS values measured when delivering video using UDP with IEEE 802.11 and iPAS, separately. Generally, the MOS values measured at the three devices are lower when N>12, in comparison with test case A. Since the quality of the video sequence in test case B is decreased. For instance, when N=12, the MOS values measured at the laptop, tablet, and smartphone using TCP with IEEE 802.11 decrease by 4%, 2%, and 2%, in comparison with test case A. In the case of delivering video using UDP with IEEE 802.11 and iPAS, in low loaded (N=0, 6, 12) and average loaded (N=18) network conditions: 1) the video quality delivered at the three devices are all above 4.5 (good quality); 2) there is no significant difference between the video delivered using iPAS and IEEE 802.11 in terms of MOS. Similar with the case of TCP, video delivery using UDP with IEEE 802.11 also results in significantly video quality degradation at the laptop in high loaded network conditions. Take N=24 for instance, as shown in Figure11-15 (c), MOS measured at the laptop decreases by 30.4% and 39.1%, in comparison with those of the tablet PC and smartphone. By comparing Figure 11-15 (c) and Figure 11-15 (d), iPAS improves the video quality at laptop by reducing the bandwidth allocation for tablet and smartphone: 1) when N=24, MOS measured at the laptop using iPAS increases from 2.3 (poor quality) to 2.8 (poor to fair quality), MOS measured at the tablet PC decreases from 3.0 (fair quality) to 2.8 (poor to fair quality), and MOS measured at the smartphone

decreases from 3.2 (fair quality) to 2.5 (poor quality); 2) when *N*=30, MOS value measured at the laptop using iPAS increase from 1.0 (bad quality) to 1.8 (bad to poor quality), MOS measured at the tablet PC stays with 1.6 (bad to poor quality), and MOS measured at the smartphone decreases from 2.0 (poor quality) to 1.6 (bad to poor quality).



(a) MOS for video over TCP with 802.11

(b) MOS for video over TCP with iPAS

(c) MOS for video over UDP with 802.11

(d) MOS for video over UDP with iPAS

Figure 11-15 Comparison of MOS between 802.11 and iPAS in test case B

It can be concluded based on test case B that, the reduction of bandwidth allocation for the tablet PC and smartphone according to iPAS can result in the following benefits, in comparison with using IEEE 802.11: 1) there is none or little impact of video quality degradation for video clips played on the tablet PC and smartphone; 2) there is an increase in video quality for the video clips played on the laptop. For instance, when *N*=24, video MOS measured at the laptop using TCP and UDP with iPAS increase by 66.7% and 38.1%, in comparison with those when IEEE 802.11 is used; 3) there is an decrease in standard deviation in the average MOS for laptop, tablet PC, and smartphone. For instance, when *N*=24, the standard deviation in MOS using TCP and UDP with iPAS decreases by 80.1% and 64.4%, separately, in comparison with the values of IEEE 802.11.

TABLE 11-11 MOS MEASURED WITH 802.11 AND IPAS IN TEST CASE A

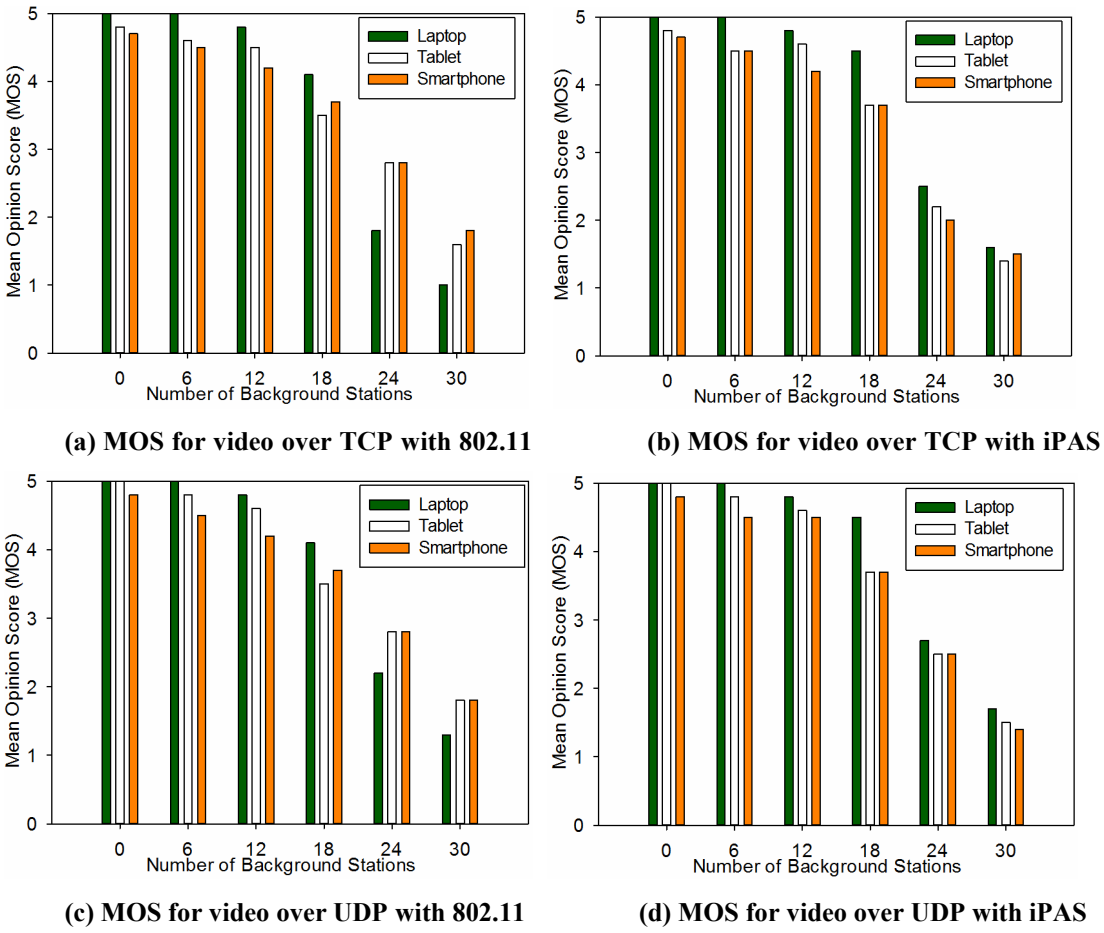
Load	Number of stations	TCP						UDP					
		Average MOS of 802.11			Average MOS of iPAS			Average MOS of 802.11			Average MOS of iPAS		
		Laptop	Tablet	Smart phone	Laptop	Tablet	Smart phone	Laptop	Tablet	Smart phone	Laptop	Tablet	Smart phone
Low loaded	0	5.0	5.0	5.0	5.0	5.0	5.0	5.0	5.0	5.0	5.0	5.0	5.0
	6	5.0	5.0	5.0	5.0	5.0	5.0	5.0	5.0	5.0	5.0	5.0	5.0
	12	5.0	4.6	4.8	5.0	4.8	4.8	5.0	4.9	5.0	5.0	5.0	5.0
Average loaded	18	4.3	4.0	4.2	4.2	4.4	4.4	4.8	4.8	4.9	4.8	4.8	4.5
High loaded	24	1.5	2.8	2.7	3.1	2.6	2.6	1.8	3.3	3.5	3.2	3.2	3.2
Over loaded	30	1.0	1.6	1.8	1.6	1.5	1.4	1.2	1.6	2.5	2.0	1.5	2.2

TABLE 11-12 MOS MEASURED WITH 802.11 AND IPAS IN TEST CASE B

Load	Number of stations	TCP						UDP					
		Average MOS of 802.11			Average MOS of iPAS			Average MOS of 802.11			Average MOS of iPAS		
		Laptop	Tablet	Smart phone	Laptop	Tablet	Smart phone	Laptop	Tablet	Smart phone	Laptop	Tablet	Smart phone
Low loaded	0	5.0	5.0	5.0	5.0	5.0	5.0	5.0	5.0	5.0	5.0	5.0	5.0
	6	5.0	5.0	5.0	5.0	5.0	5.0	5.0	5.0	5.0	5.0	5.0	5.0
	12	4.8	4.3	4.5	5.0	4.6	4.8	4.8	4.8	4.8	4.8	4.8	4.8
Average loaded	18	4.2	4.0	4.0	4.2	4.2	4.4	4.8	4.5	4.7	4.8	4.8	4.8
High loaded	24	1.5	2.6	2.8	2.5	2.4	2.4	2.1	3.0	3.1	2.9	2.8	2.7
Over loaded	30	1.0	1.5	1.7	1.5	1.5	1.3	1.0	1.6	2.0	1.8	1.6	1.6

C. Test Case C

In test case C, the low quality video sequence (bitrate: 230Kbps, frame rate: 20fps and resolution: 320x176) is delivered to the laptop, tablet PC, and smartphone. Figure 11-16 and Tablet 11-13 present the MOS values measured when delivering video using IEEE 802.11 and iPAS in test case C.



(a) MOS for video over TCP with 802.11

(b) MOS for video over TCP with iPAS

(c) MOS for video over UDP with 802.11

(d) MOS for video over UDP with iPAS

Figure 11-16 Comparison of MOS between 802.11 and iPAS in test case C

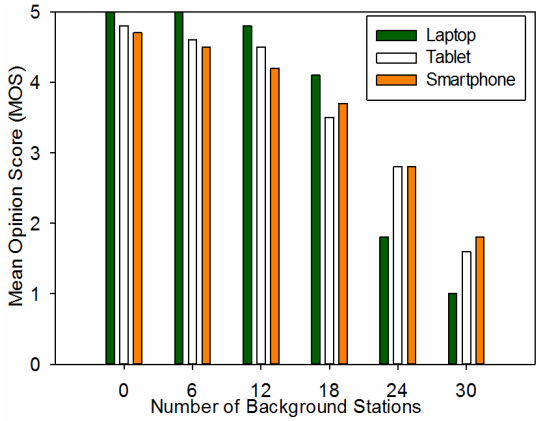
Figure 11-16 (a) and Figure 11-16 (b) present the MOS values measured when delivering video using TCP with IEEE 802.11 and iPAS, separately. Generally, the MOS values measured at the three devices are lower when N>12, in comparison with test case A and B. Since the quality of the video sequence in test case C is decreased. For instance, when N=12, the MOS values measured at the laptop, tablet PC, and smartphone using TCP with IEEE 802.11 decrease by 6.3%, 2.3%, and 6.7%, in comparison with test case B. In the case of video delivery using TCP with IEEE 802.11 and iPAS, in low loaded (N=0, 6, 12) and average loaded (N=18) network conditions: 1) the video quality delivered at the three devices

are all above 3.5 (fair quality); 2) there is no significant difference between the video delivered using IEEE 802.11 and iPAS in terms of MOS. In the case of TCP traffic with IEEE 802.11 as presented in Figure 11-16 (a), in overloaded (N=30) network conditions, the MOS values measured at the laptop users are significantly lower than those of the tablet PC and smartphone. Note that, in test case A and B, the video quality at the laptop starts decreased significantly since N equals 24. For instance, when N=30, MOS at the laptop decreases by 80%, in comparison with both the tablet PC and smartphone. This phenomenon is similar with the PSNR values measured using TCP with IEEE 802.11, which is shown in Figure 11-11 (a). The reason is that the laptop requires more bandwidth than the tablet PC and smartphone. Figure 11-16 (b) presents the MOS measured at the three devices using TCP with iPAS. It is shown that, in comparison with IEEE 802.11, iPAS improves the MOS measured at the laptop by reducing certain bandwidth share allocated for the tablet PC and smartphone. For instance, when N=30, MOS value measured at the laptop using iPAS increase from 1.0 (bad quality) to 1.6 (bad to poor quality), MOS measured at the tablet PC decreases from 1.8 (bad to poor quality) to 1.5 (bad quality), and MOS measured at the smartphone decreases from 1.8 (bad to poor quality) to 1.5 (bad quality).

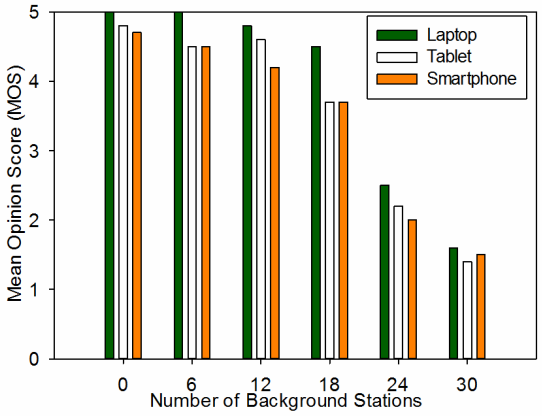
Figure 11-16 (c) and Figure 11-16 (d) present the MOS values measured using UDP with IEEE 802.11 and iPAS, separately. Generally, the MOS values measured at the three devices are lower when N>12, in comparison with test case A and B. Since the quality of the video sequence in test case B is decreased. For instance, when N=12, the MOS values measured at the laptop, tablet PC, and smartphone using TCP with IEEE 802.11 decrease by 6.3%, 12.5%, and 6.3%, in comparison with test case A. In the case of video delivery using UDP with IEEE 802.11 and iPAS, in low loaded (N=0, 6, 12) and average loaded (N=18) network conditions: 1) the video quality delivered at the three devices are all above 3.5 (fair quality); 2) there is no significant difference between the video delivered using iPAS and IEEE 802.11 in terms of MOS. Similar with the TCP case, video delivery using UDP with IEEE 802.11 also results in significantly video quality degradation at the laptop in high loaded network conditions. Take N=30 for instance, as shown in Figure11-16 (c), MOS measured at the laptop decreases by 80% and 100%, in comparison with the tablet PC and smartphone. By comparing Figure 11-16 (c) and Figure 11-16 (d), iPAS improves the video quality at the laptop by reducing the bandwidth share allocated to the tablet PC and smartphone. For instance, when N=30, MOS value measured at the laptop using iPAS increase from 1.0 (bad quality) to 1.8 (bad to poor quality), MOS measured at the tablet stays

with 1.7 (bad quality), and MOS measured at the smartphone decreases from 1.8 (bad to poor quality) to 1.7 (bad to poor quality).

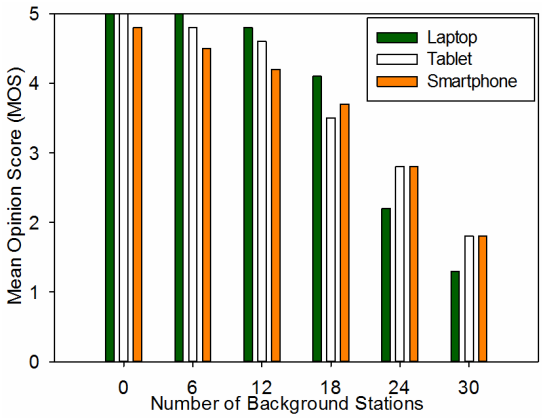
It can be concluded based on test case C that, the reduction of bandwidth share allocated to the tablet PC and smartphone according to iPAS can result in the following benefits, in comparison with using IEEE 802.11: 1) there is none or little impacts of video quality degradation for the tablet PC and smartphone; 2) there is an increase in the video quality measured at the laptop. For instance, when N=30, the MOS at the laptop using TCP and UDP with iPAS both increase by 80%, in comparison with that of IEEE 802.11; 3) there is a decrease in standard deviation in the average MOS for the laptop, tablet PC, and smartphone. For instance, when N=30, the standard deviation in MOS using TCP and UDP with iPAS decreases by 87% and 86.4%, separately, in comparison with that of IEEE 802.11.



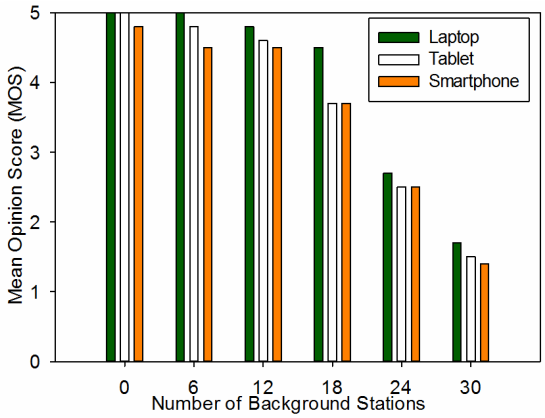
(a) MOS for video over TCP with 802.11



(b) MOS for video over TCP with iPAS



(c) MOS for video over UDP with 802.11



(d) MOS for video over UDP with iPAS

Figure 11-17 Comparison of MOS between 802.11 and iPAS in test case D

D. Test Case D

In test case D, the high quality video sequence (bitrate: 1300Kbps, frame rate: 30 fps and resolution: 800x448) is delivered to the laptop, the medium quality video sequence (bitrate: 420Kbps, frame rate: 25 fps and resolution: 512x288) is delivered to the tablet PC, and the low quality video sequence (bitrate: 230Kbps, frame rate: 20 fps and resolution: 320x176) is delivered to the smartphone. Figure 11-17 and Tablet 11-14 show the MOS values measured during video delivery using IEEE 802.11 and iPAS in test case D, respectively.

Figure 11-17 (a) and Figure 11-17 (b) present the MOS values measured using TCP with IEEE 802.11 and iPAS, separately. In the case of delivering video using TCP with IEEE 802.11 and iPAS, in low loaded (N=0, 6, 12) and average loaded (N=18) network conditions: 1) the video quality delivered at the three devices are all above 3.5 (fair quality); 2) there is no significant difference between the video delivered using IEEE 802.11 and iPAS in terms of MOS. In the case of video delivery using TCP with IEEE 802.11 as presented in Figure 11-17 (a), in high loaded (N=24) and overloaded (N=30) network conditions, the MOS values measured at the laptop users are significantly lower than that of the tablet PC and smartphone. For instance, when N=24, MOS measured at the laptop decreases by 55.6%, in comparison with both the tablet PC and smartphone. This is similar with the PSNR measured using TCP with IEEE 802.11, which is shown in Figure 11-12 (a). The reason is that the laptop requires more bandwidth than the tablet PC and smartphone due to the more powerful processing ability and larger screen size. Furthermore, in test case D, the laptop receives the highest quality video sequence and thus need much more bandwidth. Figure 11-17 (b) presents the MOS measured at the three devices when delivering video using TCP with iPAS. It is shown that, in comparison with IEEE 802.11, iPAS improves the MOS measured at the laptop by reducing certain bandwidth share allocated to the tablet PC and smartphone: 1) when N=24, MOS measured at the laptop using iPAS increases from 1.8 (bad to poor quality) to 2.5 (poor quality), MOS measured at the tablet PC decreases from 2.8 (poor to fair quality) to 2.2 (poor quality), and MOS measured at the smartphone decreases from 2.8 (poor to fair quality) to 2.0 (poor quality); 2) when N=30, MOS value measured at the laptop using iPAS increase from 1.0 (bad quality) to 1.6 (bad to poor quality), MOS measured at the tablet decreases from 1.6 (bad to poor quality) to 1.4 (bad quality), and MOS measured at the smartphone decreases from 1.8 (bad to poor quality) to 1.5 (bad quality).

Figure 11-17 (c) and Figure 11-17 (d) present the MOS values measured using UDP with IEEE 802.11 and iPAS, separately. In the case of delivering video using UDP with IEEE 802.11 and iPAS, in low loaded (N=0, 6, 12) and average loaded (N=18) network conditions: 1) the video quality delivered at the three devices are all above 3.7 (fair to good quality); 2) there is no significant difference between the video delivered using iPAS and IEEE 802.11 in terms of MOS. Similar with the TCP case, video delivery using UDP with IEEE 802.11 also results in significantly video quality degradation at the laptop in high loaded network conditions. Take N=24 for instance, as shown in Figure11-17 (c), MOS measured at the laptop decreases by 27.3% in comparison with both the tablet PC and smartphone. By comparing Figure 11-17 (c) and Figure 11-17 (d), iPAS improves the video quality at the laptop by reducing the bandwidth share allocated to the tablet PC and smartphone: 1) when N=24, MOS measured at the laptop using iPAS increases from 2.2 (poor quality) to 2.7 (poor to fair quality), MOS measured at the tablet PC decreases from 2.8 (poor to fair quality) to 2.5 (poor quality), and MOS measured at the smartphone decreases from 2.8 (poor to fair quality) to 2.5 (poor quality); 2) when N=30, MOS value measured at the laptop using iPAS increase from 1.3 (bad quality) to 1.7 (bad to poor quality), MOS measured at the tablet PC decreases from 1.8 (bad to poor quality) to 1.5 (bad quality), and MOS measured at the smartphone decreases from 1.8 (bad to poor quality) to 1.4 (bad quality).

It can be concluded based on test case D that, the reduction of bandwidth share allocated to the tablet PC and smartphone according to iPAS can result in the following benefits, in comparison with using IEEE 802.11: 1) there is none or little impacts of video quality degradation for the tablet PC and smartphone; 2) there is an increased video quality for the laptop. For instance, when N=24, the MOS measured at the laptop using TCP and UDP with iPAS increase by 38.9% and 22.7%, in comparison with that of IEEE 802.11; 3) there is a decrease in the standard deviation in the average MOS for the laptop, tablet, and smartphone. For instance, when N=24, the standard deviation in MOS using TCP and UDP with iPAS decreases by 73.8% and 48.3%, separately, in comparison with that of IEEE 802.11.

TABLE 11-13 MOS MEASURED WITH 802.11 AND IPAS IN TEST CASE C

Load	Number of stations	TCP						UDP					
		Average MOS of 802.11			Average MOS of iPAS			Average MOS of 802.11			Average MOS of iPAS		
		Laptop	Tablet	Smart phone	Laptop	Tablet	Smart phone	Laptop	Tablet	Smart phone	Laptop	Tablet	Smart phone
Low loaded	0	4.5	4.4	4.4	4.7	4.7	4.7	4.8	4.8	4.8	4.8	4.8	4.8
	6	4.4	4.5	4.5	4.4	4.5	4.5	4.7	4.7	4.7	4.7	4.6	4.7
	12	4.5	4.2	4.2	4.5	4.5	4.5	4.5	4.2	4.5	4.4	4.2	4.5
Average loaded	18	3.8	3.5	3.8	3.8	3.6	3.8	3.8	3.5	3.8	3.8	3.5	3.8
High loaded	24	2.4	2.2	2.5	2.6	2.4	2.8	2.6	2.5	2.6	2.6	2.5	2.6
Over loaded	30	1.0	1.8	1.8	1.6	1.5	1.5	1.0	1.7	1.8	1.8	1.7	1.7

TABLE 11-14 MOS MEASURED WITH 802.11 AND IPAS IN TEST CASE D

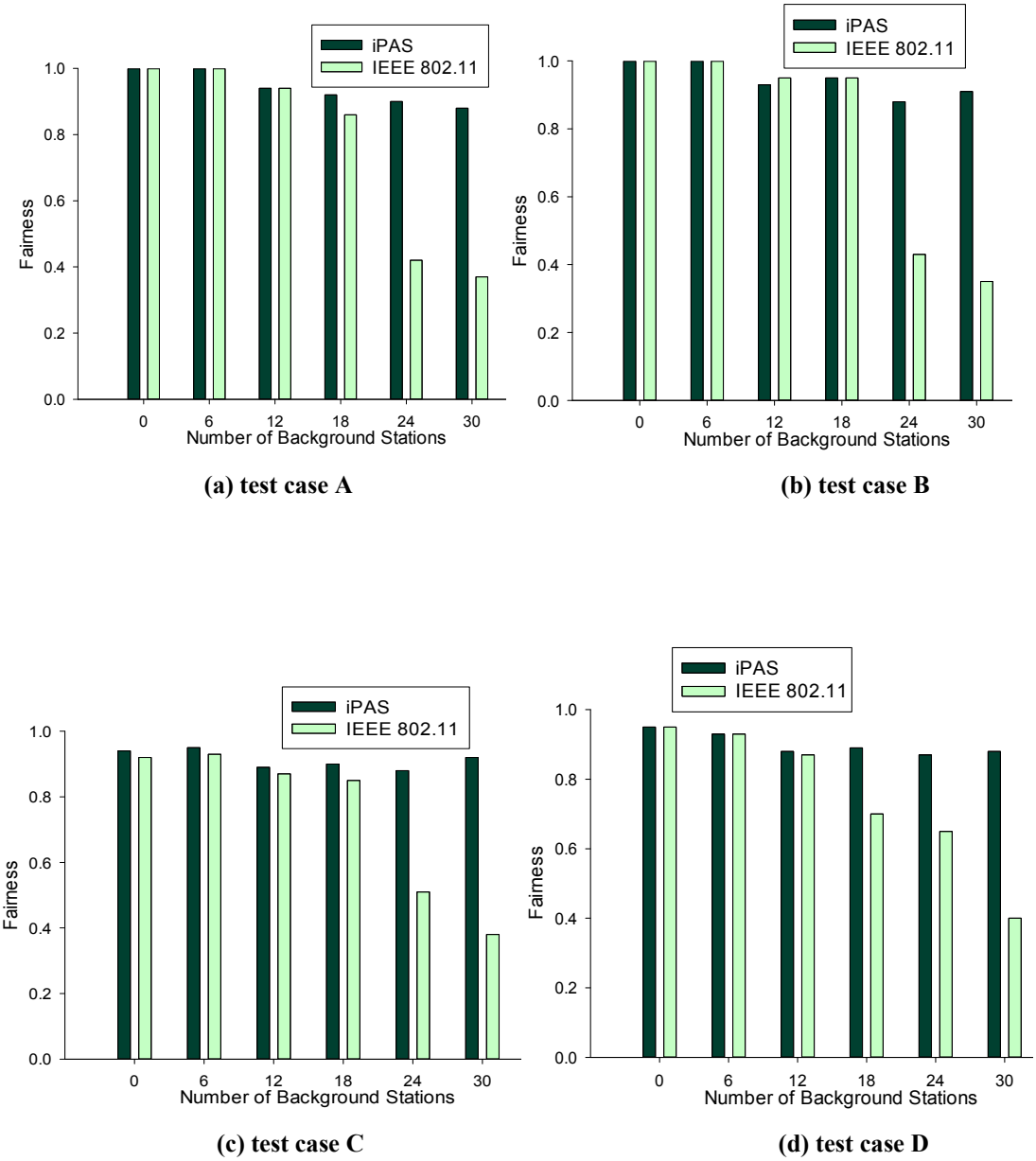
Load	Number of stations	TCP						UDP					
		Average MOS of 802.11			Average MOS of iPAS			Average MOS of 802.11			Average MOS of iPAS		
		Laptop	Tablet	Smart phone	Laptop	Tablet	Smart phone	Laptop	Tablet	Smart phone	Laptop	Tablet	Smart phone
Low loaded	0	5.0	4.8	4.7	5.0	4.8	4.7	5.0	5.0	4.8	5.0	5.0	4.8
	6	5.0	4.6	4.5	5.0	4.5	4.5	5.0	4.8	4.5	5.0	4.8	4.5
	12	4.8	4.5	4.2	4.8	4.6	4.2	4.8	4.6	4.2	4.8	4.6	4.5
Average loaded	18	4.1	3.5	3.7	4.5	3.7	3.7	4.1	3.7	3.8	4.5	3.7	3.7
High loaded	24	1.8	2.6	2.6	2.5	2.3	2.2	2.2	2.8	2.8	2.7	2.5	2.5
Over loaded	30	1.0	1.6	1.8	1.6	1.4	1.5	1.2	1.8	1.8	1.9	1.5	1.4

11.6.3 Fairness Analysis

In order to study the benefits of iPAS based on the subjective video quality assessment, the Jain's fairness index [230], as shown in equation (9-1), is selected to indicate the **device fairness**-fairness level of the video quality among laptop, tablet, and smartphone. The value x is the MOS value of certain device. The parameter i indicates the i^{th} device, n is number of devices, n=3 as there are three devices involved: laptop, tablet PC, and smartphone. The Jain's fairness index ranges between 0 and 1. For instance, a fairness index equals to 1.0 indicates that the three devices (laptop, tablet PC, and smartphone) have the same video quality rated by the users.

Figure 11-18 presents comparison of device fairness between iPAS and IEEE 802.11 using **TCP** for test case A, B, C, and D. It can be observed that, when the number of background stations (N) is lower than 18, there is no significant difference between iPAS and IEEE 802.11 in terms of device fairness. However, when N=24 or N=30, iPAS achieves better device fairness than that of IEEE 802.11. For instance, when N=24, the device fairness provided by iPAS improves by 62%, 51%, 47%, and 37% for test case A, B, C, and D, separately, in comparison with IEEE 802.11; when N=30, the device fairness provided by iPAS improves by 65%, 71%, 53%, and 49% for test case A, B, C, and D, separately, in comparison with IEEE 802.11.

Figure 11-19 presents comparison of device fairness between iPAS and IEEE 802.11 using **UDP** for test case A, B, C, and D. Similar with that of TCP traffic, when the number of background stations (N) is lower than 18, there is no significant difference between iPAS and IEEE 802.11 in terms of device fairness. However, when N=24 or N=30, iPAS achieves better device fairness than that of IEEE 802.11. For instance, when N=24, the device fairness provided by iPAS improves by 23%, 21%, 26%, and 25% for test case A, B, C, and D, separately, in comparison with IEEE 802.11; when N=30, the device fairness provided by iPAS improves by 27%, 31%, 25%, and 24% for test case A, B, C, and D, separately, in comparison with IEEE 802.11.



(a) test case A **(b) test case B**

(c) test case C **(d) test case D**

Figure 11-18 Comparison of device fairness between 802.11 and iPAS using TCP

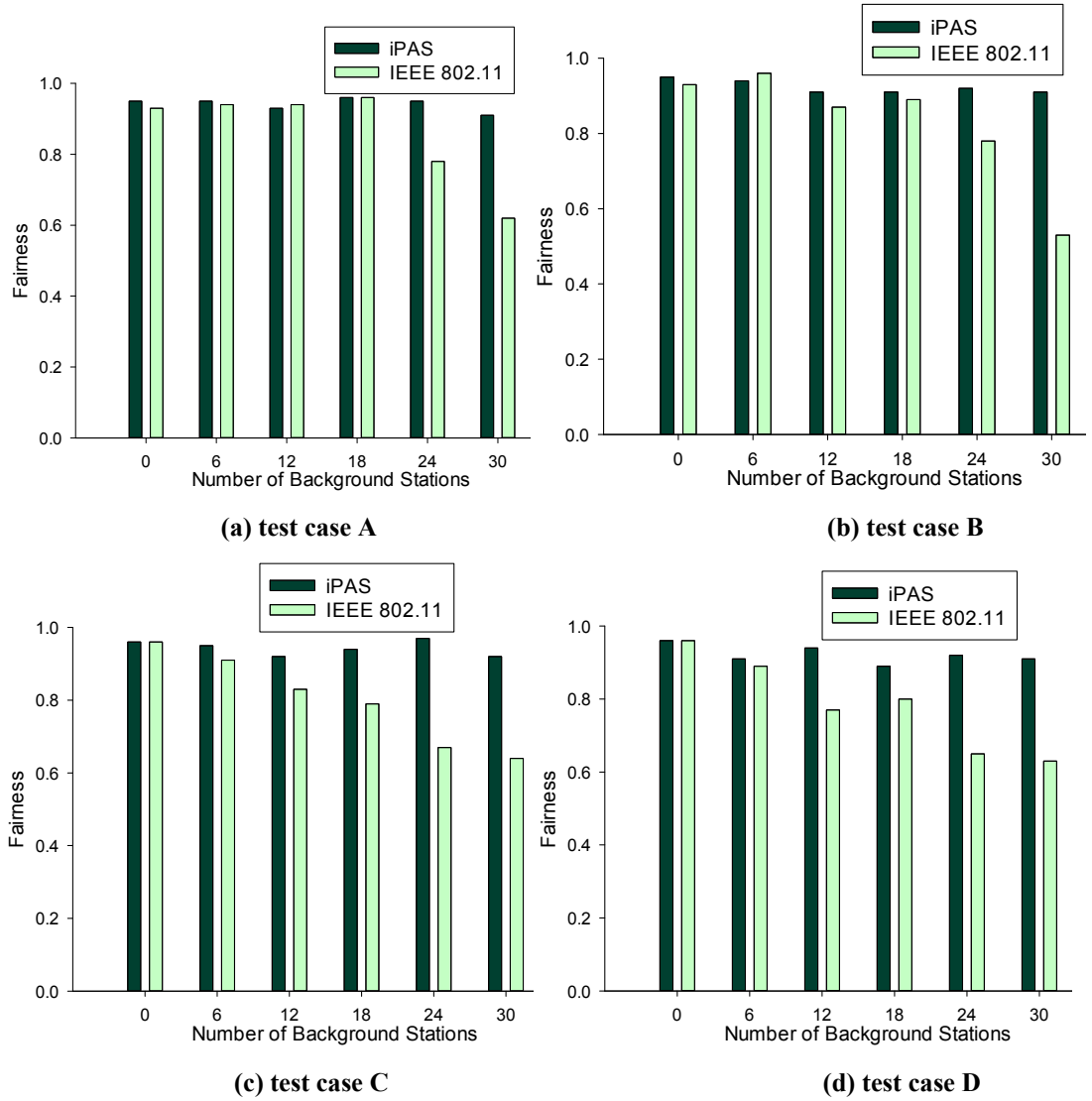


Figure 11-19 Comparison of device fairness between 802.11 and iPAS using UDP

11.7 Results Comparison

The video quality has been studied in four test cases (test case A, B, C, and D) using both objective and subjective metrics. In these cases, iPAS shows better perceived video quality than the equal channel access mechanism of IEEE 802.11, when delivering TCP or UDP flows. Additionally, the following conclusions can be drawn: 1) Generally, UDP video delivery provides higher video quality than TCP in terms of both PSNR and mean opinion score; 2) Lower delivered quality video results in lower user perceived quality. For instance, the perceived video quality achieved the highest in test case A and the worst in test case C, as test case A delivered the highest quality video sequence and test case C - the lowest. 3)

Under low and average loaded network conditions, there is no significant difference in the video quality measured at the laptop, tablet PC, and smartphone between using iPAS and original IEEE 802.11 protocol; 4) Under high and overloaded network conditions, iPAS can improve the video quality at the laptop without significantly decreasing the video quality at the tablet PC and smartphone; 5) Under high and overloaded network conditions, iPAS improves the device fairness with up to 71% and 31% for video delivery over TCP and UDP, respectively, in comparison with IEEE 802.11.

11.8 Summary

This chapter presents the performance comparison-based video transmission between using the equal channel access mechanism of IEEE 802.11 and iPAS over IEEE 802.11. Real life test-bed is described for assessing the performance of video delivery using IEEE 802.11 protocol. The simulation test-bed is also introduced for evaluating the performance of iPAS. Both test beds include equipment and software specifications, wireless environment configuration, video content, background traffic, and experimental scenarios. The delivered video sequences based on IEEE 802.11 protocol and iPAS are assessed in terms of both objective and subjective metrics. In the case of objective video quality assessment, PSNR values are measured at the laptop, tablet PC, and smartphone. In subjective video quality assessment, MOS values obtained at the laptop, tablet PC, and smartphone are generated based on feedback from 32 users.

CHAPTER 12

Conclusions and Future Works

This chapter summarizes the research work reported in this thesis, and highlights the significant contributions which include Model-based Bandwidth Estimation algorithm (MBE), intelligent Prioritized Adaptive Scheme (iPAS), and QoS-based downlink/uplink fairness scheme for VoIP communications over IEEE 802.11 networks. Several potential future works are listed at the end.

12.1 Problems and Solutions Overview

The purpose of this research is to find ***a solution for delivering high quality and fair multimedia services to heterogeneous devices in IEEE 802.11 networks*** that would be beneficial to both consumers and service providers. Existing solutions involve complicated and expensive implementations, have poor performance in loaded network conditions, and/or do not consider consumer devices characteristics and end-user perceived quality. This thesis proposes a ***Model-based Bandwidth Estimation algorithm (MBE)*** which supports accurate bandwidth estimation for IEEE 802.11 networks. MBE was validated in various conditions including different packet sizes, dynamic wireless link rate, different channel noise, and variable network loaded conditions. The ***intelligent Prioritized Adaptive Scheme (iPAS)*** was also proposed to deliver differentiated services according to content and device characteristics. iPAS also improves fairness, delay, throughput, loss, and user perceived video quality. The ***QoS-based downlink/uplink fairness scheme*** for VoIP was introduced to provide fair wireless channel access between downlink and uplink VoIP flows and was assessed with increasing number of VoIP stations.

12.2 Contributions to the State of the Art

This thesis presents the ***principles, mechanisms, experimental evaluation, and result analysis*** for the proposed schemes, MBE, iPAS, and the QoS-based downlink/uplink fairness

scheme. The performance of *MBE* and *iPAS* are evaluated under both simulation-based test-bed and real life test-bed. Specifically, subjective video quality assessment is conducted to investigate the effect of video delivery using *iPAS* on the perception of end users. The performance of the *QoS-based downlink/uplink fairness scheme for VoIP* is studied under simulation-based test-bed by importing real VoIP traffic. The novel contributions and significant results of the proposed scheme are summarized as followed.

1) Unlike most existing bandwidth estimation schemes that send probing traffic to estimate the network conditions, *MBE* proposes a server-side based TCP/UDP throughput model to estimate the available bandwidth, based on feedback information from the MAC layer such as packet loss and round trip time. *MBE* relies on two novel ***throughput models for TCP and UDP traffic over IEEE 802.11 networks***.

Experimental results show that *MBE* model is robust under different conditions: variant packet size, packet error rate and dynamic wireless link. Additionally, *MBE* provides accurate bandwidth estimation with low overhead in comparison with existing bandwidth estimation techniques such as *iBE*, *DietTOPP*, and *IdleGap*. Among the three compared techniques, *IdleGap* achieved very good estimation error rate and *iBE* introduced the lowest overhead. However, *MBE* achieved 47% less estimation error rate than *IdleGap* and 9.3% lower overhead than *iBE*. *MBE* also reduces packet loss with up to 56% in comparison with that of *iBE* and with up to 50% in comparison with that of *IdleGap*. Additionally, *MBE* produces the lowest standard deviation and mean value for both error rate and overhead. Furthermore, the two tailed t-test statistical analysis shows that there is no significant statistical difference between MBE and real test results with 95% confidence level.

2) *iPAS* used the mathematical theory-***stereotypes, which prioritized individual stream according to stream-related characteristics (i.e. device resolution, device battery power left, and application type) and network delivery QoS parameters (i.e. delay, jitter, and packet loss rate).*** The original IEEE 802.11 protocol is designed for best effort service and incorporates limited QoS support with regard to multimedia applications and mobile devices. IEEE 802.11e supports the differentiation of traffic based on type only (voice, video, best effort and background) and does not consider network conditions, nor device characteristics.

Simulation-based tests demonstrated how better results are obtained when employing iPAS than when either 802.11 DCF or 802.11e EDCA mechanisms are used.

iPAS performance benefits are as follows: 1) achieves better fairness in bandwidth allocation; 2) achieves higher throughput than 802.11 DCF and 802.11e EDCA with up to 38% and 20%, respectively; 3) enables definite throughput and delay differentiation between streams; 4) reduces packet loss rate with 34% and 18%, in comparison with 802.11 DCF and 802.11e EDCA, respectively; 5) improves delivered video quality. Notably, iPAS does not require any MAC layer modifications.

A real life test-bed has been used to study objective and subjective video quality when using *iPAS over IEEE 802.11* and the original *equal channel access mechanism of IEEE 802.11g*. PSNR and MOS values are used as the video quality metric in the objective and subjective tests. It has been concluded that: 1) under high and overloaded network conditions, iPAS can improve the video quality at laptop without significantly decreasing the video quality at tablet and smartphone; 2) under high and overloaded network conditions, iPAS improves the device fairness up to 71% and 31% for TCP and UDP, respectively, in comparison with IEEE 802.11.

3) The ***QoS-based downlink/uplink fairness scheme*** for VoIP used the ***stereotypes-based structure to balance the QoS parameters (throughput, delay, and loss) between downlink and uplink VoIP traffic***. The original IEEE 802.11 protocol does not support fair traffic distribution between downlink and uplink. Additionally, none of the existing fairness-based downlink/uplink schemes consider the three QoS parameters (throughput, delay, and loss). Three QoS performance parameters, $\text{Throughput}_{\text{down/up}}$, $\text{Delay}_{\text{down/up}}$, and $\text{LossRate}_{\text{down/up}}$ are modelled as stereotype features, representing throughput ratio, delay ratio and loss ratio between downlink and uplink communication channels, respectively.

Based on the simulation tests, the following conclusions have been reached. 1) by using the proposed scheme, the VoIP capacity increases by 42% and 21%, respectively, in comparison with 802.11 and Dynamic-CW; 2) the gap between the uplink delay and downlink delay is lower than that of 802.11 provided that the number of stations is higher than 12; 3). The fairness level provided by the proposed scheme is higher than both 802.11 and Dynamic-CW with up to 132%, 52% and 37%, in terms of delay, throughput and packet loss rate, respectively.

In conclusion, the proposed solutions in this thesis provide service differentiation in terms of network QoS and device characteristics when delivering multimedia content over

the IEEE 802.11 networks. Additionally, the VoIP delivery in IEEE 802.11 networks has been optimized by providing fair traffic distribution between downlink and uplink.

12.3 Future Works

The following potential future works are considered based on the proposed solutions described in this thesis:

Bandwidth Estimation for IEEE 802.11e and IEEE 802.11n Networks

MBE can be extended in IEEE 802.11e and IEEE 802.11n networks. IEEE 802.11e provides multimedia QoS support by introducing traffic access categories and a block acknowledgement mechanism at MAC layer. IEEE 802.11n network provides significant data transfer rate with up to 600Mbps. 802.11n improves the multimedia transmission quality by using group-based frame at MAC layer and MIMO technique at PHY layer. Since MBE is developed based on the original 802.11 DCF, and 802.11e and 802.11n are also based on the 802.11 DCF protocol, MBE will also work in 802.11e and 802.11n. Future works will report the results of MBE in 802.11e/n networks.

Bandwidth Estimation for Admission Control in IEEE 802.11 Networks

IEEE 802.11 networks have been widely adopted to offer high-speed data access. One critical challenge is that too many wireless stations might compete for the wireless channel and thus cause congestion or even system collapse. Additionally, real time applications such as voice over IP (VoIP) and video chatting are more sensitive to delay which is mainly caused by the network congestion. Admission control scheme is designed to alleviate the network congestion and improve the network resource utilization by accepting or rejecting an incoming flow. MBE algorithm can be used to serve for an admission control mechanism in IEEE 802.11 networks. An incoming flow is accepted or rejected can be determined based on the estimated available bandwidth of the underlying wireless network. For instance, if the acceptance of a new flow would cause the wireless network overloaded, such flow will be rejected immediately; otherwise, flow will be accepted.

Bandwidth Estimation for Load Balancing in IEEE 802.11 Networks

In the infrastructure-mode of IEEE 802.11 networks, a wireless station connects to the access points (APs) to access the wired worlds. Considering the campus network environment where there lots of APs deployed in order to provide large area access. Since the wireless client independently selects AP to connect, traffic loads might be unevenly distributed among the available APs. This load-balancing problem can lead to network congestion and overloaded network condition. Efficient bandwidth estimation scheme like MBE can provide a good load-balancing mechanism for the available APs. MBE can be implemented at AP and the available bandwidth provided by each AP is estimated. Whenever a wireless station starts a connection request to certain AP, this AP first estimates the available bandwidth it can provide. If the network is already loaded above a pre-defined threshold, the connection request would be rejected and/or forwarded to the adjacent AP. In this case, the value of the threshold is critical in order to avoid the network congestion.

Development of Home Smart Gateway based on iPAS

It is a common scenario that multiple wireless devices access internet via the wireless access point (AP) in the home. Family members might using their own mobile device to receive different services, for example, on-line video game console with its own IEEE 802.11 enabled interface, video on demand service using laptop, web-browsing service using smartphone, etc. A resource allocation scheme is needed to manage these services and devices in order for each user to obtain high quality of services, while the overall bandwidth resources are used efficiently. iPAS has been evaluated via simulations and has shown good performance in providing differentiated services. In future, iPAS could be implemented in the home smart gateway which monitors the available bandwidth and allocates the available resources based on the stream priority.

QoS-based Downlink/Uplink Fairness Scheme for Video Conferencing

The proposed QoS-based downlink and uplink traffic fairness scheme aims to improve the QoS of VoIP service. The principle idea is to use the mathematical theory-stereotypes to dynamically control the QoS parameters of downlink and uplink traffic. The proposed scheme can be extended to provide downlink/uplink fairness for interactive video application such as video conference over IEEE 802.11 network. Different with VoIP service, live video streaming services consume more bandwidth due to three primary reasons: 1) video packet size is generally higher than that of VoIP packet size, i.e., typical video packet ranges from 1000-1500 bytes while VoIP packet size ranges from 100-300 bytes; 2) video streams are

encoded at higher bitrates; 3) video is continuous: in a video conference scenario, the bandwidth is used as long as the video camera is on; while in VoIP, bandwidth is consumed only when a user is talking.

Appendix

Perceptual Test Instructions

Acknowledgement

Appreciate your support for the video perceptual test organized by the Performance Engineering Lab, School of Electronic Engineering, Dublin City University, Ireland.

Test Motivations

The objective of the perceptual test is to compare the performance of original IEEE 802.11 protocol and the proposed resource allocation solution, intelligent Prioritized Adaptive Scheme (iPAS).

Test Guidelines

To complete the perceptual tests, you are encouraged to finish one test case which takes around 30 minutes. You are also recommended to start the test according to the following guidelines:

1. Filling the **Personal Information Form**.
2. Filling the **Questionnaire** while watching video clips. First, you will watch three different video sequences on laptop, tablet, and smartphone, separately. Afterwards, you will watch another three different video sequences on the same laptop, tablet, and smartphone, separately. Note that, each video sequence played on certain device is further divided into 6 shorter clips, with the first clip of 20s and the remaining five clips of 60s each. During the playback of each short video clip, you are asked to rate the video quality in terms of MOS, continuity, and blurry. Each user will fill totally six questionnaires after the whole test.
3. The rating is supposed to be finished immediately when one video clip ended.
4. You are not allowed to change the distance from the device screen too much.

Personal Information Form

Please check "√" for your choice **User No: _____**

Gender:	(A) Male			(B) Female			
Age:							
Working type:	(A) Computer Science	(B) Engineering	(C) Education	(D) Finance	(E) others		
Do you use glasses?	(A) Yes			(B) No			
Do you have visual disabilities, such as colour blindness and colour weakness?	(A) Yes			(B) No			
How often do you watch streaming videos?	(A) Every day	(B) Twice/week	(C) Once/week	(D) Once/month	(E) Never		
Which device is your favourite to watch video?	(A) Laptop		(B) Tablet		(C) Smartphone		
How familiar are you with wireless multimedia delivery?	(A) Expert		(B) Familiar		(C) Not familiar		
How familiar are you with video quality assessment?	(A) Expert		(B) Familiar		(C) Not familiar		
Which network do you use most to watch video?	(A) Wi-Fi		(B) 3G/4G		(C) Ethernet with cable		
Which is your favourite movie type?	(A) Action	(B) Horror	(C) Comedy	(D) Cartoon	(E) Crime	(F) Romance	(G) Other

Guidlines for Rating the Video Quality

Rating	Description			
	Mean Opinion Score	Continuity	Blurred	Blockness
1	Bad	Bad	Significantly blurred	Too much Blockness
2	Poor	Poor	Much blurred	Much blockness
3	Fair	Fair	Few blurred	Few blockness
4	Good	Good	Very few blurred	Very few blockness
5	Excellent	Excellent	Not blurred	No blockness

Example of questionnaire for one user (1/6)

User No: —————— **Test Case: ——————**

Device: Laptop

Video Clip	MOS	Continuity	Blurred	Blockness
1	1	1	1	1
	2	2	2	2
	3	3	3	3
	4	4	4	4
	5	5	5	5
2	1	1	1	1
	2	2	2	2
	3	3	3	3
	4	4	4	4
	5	5	5	5
3	1	1	1	1
	2	2	2	2
	3	3	3	3
	4	4	4	4
	5	5	5	5
4	1	1	1	1
	2	2	2	2
	3	3	3	3
	4	4	4	4
	5	5	5	5
5	1	1	1	1
	2	2	2	2
	3	3	3	3
	4	4	4	4
	5	5	5	5
6	1	1	1	1
	2	2	2	2
	3	3	3	3
	4	4	4	4
	5	5	5	5

Example of questionnaire for one user (2/6)

User No: —————— **Test Case: ——————**

Device: Laptop

Video Clip	MOS	Continuity	Blurred	Blockness
1	1	1	1	1
	2	2	2	2
	3	3	3	3
	4	4	4	4
	5	5	5	5
2	1	1	1	1
	2	2	2	2
	3	3	3	3
	4	4	4	4
	5	5	5	5
3	1	1	1	1
	2	2	2	2
	3	3	3	3
	4	4	4	4
	5	5	5	5
4	1	1	1	1
	2	2	2	2
	3	3	3	3
	4	4	4	4
	5	5	5	5
5	1	1	1	1
	2	2	2	2
	3	3	3	3
	4	4	4	4
	5	5	5	5
6	1	1	1	1
	2	2	2	2
	3	3	3	3
	4	4	4	4
	5	5	5	5

Example of questionnaire for one user (3/6)

User No: —————— **Test Case: ——————**

Device: Tablet

Video Clip	MOS	Continuity	Blurred	Blockness
1	1	1	1	1
	2	2	2	2
	3	3	3	3
	4	4	4	4
	5	5	5	5
2	1	1	1	1
	2	2	2	2
	3	3	3	3
	4	4	4	4
	5	5	5	5
3	1	1	1	1
	2	2	2	2
	3	3	3	3
	4	4	4	4
	5	5	5	5
4	1	1	1	1
	2	2	2	2
	3	3	3	3
	4	4	4	4
	5	5	5	5
5	1	1	1	1
	2	2	2	2
	3	3	3	3
	4	4	4	4
	5	5	5	5
6	1	1	1	1
	2	2	2	2
	3	3	3	3
	4	4	4	4
	5	5	5	5

Example of questionnaire for one user (4/6)

User No: —————— **Test Case: ——————**

Device: Tablet

Video Clip	MOS	Continuity	Blurred	Blockness
1	1	1	1	1
	2	2	2	2
	3	3	3	3
	4	4	4	4
	5	5	5	5
2	1	1	1	1
	2	2	2	2
	3	3	3	3
	4	4	4	4
	5	5	5	5
3	1	1	1	1
	2	2	2	2
	3	3	3	3
	4	4	4	4
	5	5	5	5
4	1	1	1	1
	2	2	2	2
	3	3	3	3
	4	4	4	4
	5	5	5	5
5	1	1	1	1
	2	2	2	2
	3	3	3	3
	4	4	4	4
	5	5	5	5
6	1	1	1	1
	2	2	2	2
	3	3	3	3
	4	4	4	4
	5	5	5	5

Example of questionnaire for one user (5/6)

User No: —————— **Test Case: ——————**

Device: Smartphone

Video Clip	MOS	Continuity	Blurred	Blockness
1	1	1	1	1
	2	2	2	2
	3	3	3	3
	4	4	4	4
	5	5	5	5
2	1	1	1	1
	2	2	2	2
	3	3	3	3
	4	4	4	4
	5	5	5	5
3	1	1	1	1
	2	2	2	2
	3	3	3	3
	4	4	4	4
	5	5	5	5
4	1	1	1	1
	2	2	2	2
	3	3	3	3
	4	4	4	4
	5	5	5	5
5	1	1	1	1
	2	2	2	2
	3	3	3	3
	4	4	4	4
	5	5	5	5
6	1	1	1	1
	2	2	2	2
	3	3	3	3
	4	4	4	4
	5	5	5	5

Example of questionnaire for one user (6/6)

User No: —————— **Test Case: ——————**

Device: Smartphone

Video Clip	MOS	Continuity	Blurred	Blockness
1	1	1	1	1
	2	2	2	2
	3	3	3	3
	4	4	4	4
	5	5	5	5
2	1	1	1	1
	2	2	2	2
	3	3	3	3
	4	4	4	4
	5	5	5	5
3	1	1	1	1
	2	2	2	2
	3	3	3	3
	4	4	4	4
	5	5	5	5
4	1	1	1	1
	2	2	2	2
	3	3	3	3
	4	4	4	4
	5	5	5	5
5	1	1	1	1
	2	2	2	2
	3	3	3	3
	4	4	4	4
	5	5	5	5
6	1	1	1	1
	2	2	2	2
	3	3	3	3
	4	4	4	4
	5	5	5	5

References

[1] Ericsson connectivity report, http://www.ericsson.com/res/thecompany/docs/corpinfo/reports/ericsson_connectivity_report_feb2011.pdf, Feb. 2011.

[2] Ericsson Report, http://www.ericsson.com/res/docs/2012/ER-WiFi-Integration.pdf, Feb. 2011.

[3] H.Chen, L.Huang, S.Kumar, and C.C.Jay Kuo, “Radio Resource Management for Multimedia QoS Support in Wireless Networks,” Springer, 2004, pp.4-6.

[4] M. Uhlirz, “Concept of a GSM-based Communication System for High-speed Trains”, IEEE 44th Vehicular Technology Conference, Stockholm, Sweden, 1994.

[5] Digital cellular telecommunications system (Phase 2+); General Packet Radio Service (GPRS); Overall description of the GPRS radio interface; Stage 2 (3GPP TS 03.64 version 8.9.0 Release), 1999.

[6] S.G. Zhao, X.W. Zhou, Enhanced Data rate for Global Evolution-EDGE, *Publishing House of Electronic Industry*, ISBN:9787121077630, 2009.

[7] ETSI Special Mobile Group, “Universal Mobile Telecommunications Systems: Objectives and Overview (UMTS 01.01),” ETSI Technical Report 217, European Telecomm. Standards Inst., Sophia-Antipolis, France, http://www.etsi.fr/, 1996.

[8] D. Knisely, Q. Li, and N. Ramesh, "cdma2000 - Evolution of cdmaOne to IMT-2000," to appear, Bell Labs Tech. J., 1998.

[9] ETSI: "ETS 300 175 - Radio Equipment and Systems (RES); Digital European Cordless Telecommunications (DECT); Common interface; Parts 1 to 9”, European Telecommunication Standards Institute, Valbonne Cedex, France, Oct. 1992.

[10] D.Barrett and T.King, Computer Networking Illuminated, Jones & Bartlett Learning, 2000, pp.180-183.

[11] IEEE 802.16-2004, IEEE Standard for Local and Metropolitan Area Networks Part 16: Air Interface for Fixed Broadband Wireless Access Systems, Jul. 2004.

[12] D. Astely, E. Dahlman, A. Furuskar, Y. Jading, M. Lindstrom, S. Parkvall, “LTE: the evolution of mobile broadband-[LTE part II: 3GPP release 8],” *IEEE Comm. Magazine*, vol.47, issue.4, pp.44-51, May. 2009.

[13] 3GPP, Technical Specification Group Radio Access Network: Physical Layer rocedures (FDD) (3GPP TS 25.214 v7.7.0), Nov. 2007. [Online]. Available: http://www.3gpp.org/ftp/Specs/html-info/25214.html.

[14] 3GPP TS 25.309 V6.6.0, “Group Radio Access Networks; FDD Enhance Uplink; Overall Description,” 2006.

[15] B. A. Bjerke, “LTE-advanced and the evolution of LTE deployments,” *IEEE Wireless Communications*, vol. 18, no. 5, pp. 4-5, Oct. 2011.

[16] IEEE Std. 802.11,“IEEE Std. 802.11-2007, Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) specifications,” 2007.

[17] IEEE 802.11b, part 11: Wireless LAN medium access control (MAC) and physical layer (PHY) specifications: higher-speed physical layer extension in the 2.4GHz band, supplement to IEEE 802.11 Std., 1999.

[18] IEEE Std. 802.11g/D1.1-2001, Part11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) specifications: Further Higher-Speed Physical Layer Extension in the 2.4 GHz Band.

[19] Wireless LAN MAC and PHY Specifications: Enhancements for Higher Throughput, IEEE Draft P802.11n/D.2.00, Feb. 2007.

[20] “Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications Amendment-Quality of Service Enhancements,” IEEE 802.11e, IEEE Standard for Information Technology, 2005.

[21] K. Xu, M. Gerla, S.Bae, “How Effective is the IEEE 802.11 RTS/CTS Handshake in Ad hoc Networks,” *IEEE Global Telecommunications Conference,* Taiwan, China, Nov. 2002.

[22] Bluetooth Special Interest Group, Specification of the Bluetooth system Version 1.1B, specification Vol. 1&2, Feb. 2001.

[23] Zigbee Alliance, "Zigbee specification: Zigbee document 053474r13 Version 1.1," 1 Dec. 2006. Web site: http://www.zigbee.org.

[24] IEEE Standard for Local and Metropolitan Area Networks—Part 16: Air Interface for Fixed and Mobile Broadband Wireless Access Systems. Amendment 2: Physical and Medium Access Control Layers for Combined Fixed and Mobile Operation in Licensed Bands, IEEE Std 802.16e™-2005 and IEEE Std 802.16™-2004/Cor 1-2005.

[25] Todor Cooklev. Wireless Communication Standards: A Study of IEEE 802.11, 802.15, and 802.16. IEEE Press, USA, 2004.

[26] IEEE 802.21-2008, Standard for Local and Metropolitan Area Networks-Part 21: Media Independent Handover Services, IEEE Computer Society, Jan. 2009.

[27] S.Amante et al, “Inter-provider Quality of Service”, Quality of Service Working Group, MIT Communications Futures Program (CFP), Nov. 2006.

[28] G. Almes, S. Kalidindi and M. Zekauskas, “A One-way Delay Metric for IPPM,” RFC 2679, Sept.1999.

[29] C. Demichelis and P. Chimento, “IP Packet Delay Variation Metric for IP Performance Metric (IPPM),” RFC 3393, Nov. 2002.

[30] G. Almes, S. Kalidindi and M. Zekauskas, “A One-way Packet Loss Metric for IPPM,” RFC 2680, Sept.1999.

[31] ITU-T Y.1541, “Network Performance objectives for IP-based Services”, Dec. 2011.

[32] "E.800: Terms and definitions related to quality of service and network performance including dependability," ITU-T Recommendation. Aug. 1994.

[33] ITU-T Rec. P.861, “Objective Quality Measurement of Telephone-Band (300-3400 Hz) Speech Codecs,” Geneva, 1996.

[34] “Perceptual evaluation of speech quality (PESQ): An objective method for end-to-end speech quality assessment of narrow-band telephone networks and speech codecs,” ITU-T Rec. P.862, 2001.

[35] ITU-T Draft Rec. P. 863 “Perceptual Objective Listening Quality Assessment (POLQA)”, International Telecommunication Union, Geneva, 2010.

[36] ITU-T Recommendation G.109, “Definition of Categories of Speech Transmission Qualit”, Sept. 1998.

[37] ITU-T G.107, The E-Model-http://www.itu.int/rec/T-REC-G.107-201112-P/en, Dec. 2011.

[38] Wang, Y. “Survey of Objective Video Quality Measurements,” Tech report, Worcester Polytechnic Institute, Jun. 2006.

[39] G. Hauske, T. Stockhammer, and R. Hofmaier, “Subjective Image Quality of Low-rate and Low-Resolution Video Sequences”, Proc. International Workshop on Mobile Multimedia Communication, Munich, Germany, Oct. 2003.

[40] O.Nemetheova, M.Ries, E.Siffel, M.Rupp, “Quality Assessment for H.264 Coded Low rate and low resolution Video Sequences,” Proc. of Conf on Internet and Inf. Technologies (CIIT), St.Thomas, US Virgin Islands, pp. 136-140, 2004.

[41] M. Pinson and S. Wolf. “A New Standardized Method for Objectively Measuring Video Quality,” IEEE Trans. Broadcasting, vol. 50, issue.3, pp. 312-322, Sept. 2004.

[42] C. J. Branden Lambrecht and O. Verscheure. “Perceptual Quality Measure using a Spatio-Temporal Model of the Human Visual System,” Proc. SPIE Vol. 2668, p. 450-461, March, 1996.

[43] A.P. Hekstra, J.G. Beerends, D. Ledermann, F.E. de Caluwe, S. Kohler, R.H. Koenen, S. Rihs, M. Ehrsam, and D. Schlauss, “PVQM-A Perceptual Video Quality Measure,”

Journal of Signal Processing: Image Communications," vol.17, no.10, 2002, pp. 781-798.

[44] ITU-T P.861, "Objective Quality Measurement of Telephone-band (300-3400 Hz) peech Codecs," Feb 1996.

[45] ITU-T Recommendation P.800, "Methods for Subjective Determination of Transmission Quality," Aug 1996.

[46] STEVENS (S.S): Psychophysics-Introduction to its perceptual, neural and social prospects, John Wiley and Sons, 1975.

[47] ITU-T Recommendation P.835, "Subjective Test Methodology for Evaluating Speech Communications Systems that Include Noise Suppression Algorithm," Nov. 2003.

[48] ITU-T Recommendation P.805, "Subjective Evaluation of Conversational Quality," Apr. 2007.

[49] ITU-T Recommendation P.830, "Subjective Performance Assessment of Telephone-based and Wideband Digital Codecs," Feb. 1996.

[50] ITU-T Recommendation P.840, "Subjective Listening Test Method for Evaluating Circuit Multiplication Equipment," Nov. 2003.

[51] ITU-T Recommendation P.851, "Subjective Quality Evaluation of Telephone Services based on Spoken Dialogue Systems," Nov. 2003.

[52] ITU-T Recommendation P.910, "Subjective Video Quality Assessment Methods for Multimedia Applications," Apr. 2008.

[53] ITU-T Recommendation P.911, "Subjective Audiovisual Quality Assessment Methods for Multimedia Applications," Dec. 1998.

[54] ITU-T Recommendation BT.500-13, "Methodology for the Subjective Assessment of the Quality of Television Pictures," Jan. 2012.

[55] ITU-T Recommendation P.920, "Interactive Test Methods for Audiovisual Communications," May. 2000.

[56] F.Kozamernik et al, "Subjective quality of internet video codecs ---phase II evaluations using SAMVIQ," European Broadcasting Union (EBU) Technical Review, No. 301, Jan. 2005.

[57] Hyun Jong Kim, Dong Hyeon Lee, Jong Min Lee, Kyoung Hee Lee, Won Lyu, Seong Gon Choi, "The QoE Evaluation Method through the QoS-QoE Correlation Model", Networked Computing and Advanced Information Managemet(NCM), Gyeongju, Korea , Sept. 2008.

[58] Rosario.F, Filippo.S, Demin.W, Andre.V, "Video Quality Metric for Bit Rate Control

via Joint Adjustment of Quantization and Frame Rate," IEEE Trans. Broadcasting, vol.53, no.1, Mar. 2007.

[59] J.M. N, Monteiro, and S. Mario, "A Subjective Quality Estimation Tool for the Evaluation of Video Communication Systems", 12th IEEE Symposium on Computers and Communications (ISCC), Aveiro, Portugal, Jul. 2007.

[60] ISO/IEC International Standard 13818, "MPEG-2-Generic Coding of Moving Pictures and Associated Audio Information", Nov. 1994.

[61] ITU-T Recommendation H.262, "Information Technology-Generic Coding of Moving Pictures and Associated Audio Information: Video", Feb. 2000.

[62] ISO/IEC 14496 (MPEG-4), "Information Technology-Coding of Audio-Visual Objects", International Standards Organization, 2001.

[63] M.M. Hannuksela, V. Varsa, A. Hourunranta, and M. Gabbouj, "The Error Concealment Feature in the H.26L Test Model," IEEE International Conference on Image Processing, Rochester, New York, USA, September 22-25, 2002, pp.729-732.

[64] J.-Y. Pyun, "Error concealment aware streaming video system over packet-based mobile networks," IEEE Transactions on Consumer Electronics, vol. 54, no. 4, pp. 1705 –1713, Nov. 2008.

[65] D. Silva, F. R. Kschischang, and R. Koetter, "A Rank-Metric Approach to Error Control in Random Network Coding," IEEE Transactions on Information Theory, vol. 54, no. 9, pp. 3951 –3967, Sep. 2008.

[66] M. C. Vuran and I. F. Akyildiz, "Error Control in Wireless Sensor Networks: A Cross Layer Analysis," IEEE/ACM Transactions on Networking, vol. 17, no. 4, pp. 1186 – 1199, Aug. 2009.

[67] L. Tseng, H.-C. Chuang, C. Y. Huang, and T. Chiang, "A Buffer-Feedback Rate Control Method for Video Streaming over Mobile Communication Systems," International Conference on Wireless Networks, Communications and Mobile Computing, Maui, HI, 13-16 Jun.2005, pp. 1266 –1270.

[68] S.-Y. Lee and M.-K. Chang, "New Results on Single-Step Power Control System in Finite State Markov Channel: Power Control Error Modelling and Queueing Variation Modelling," IEEE International Conference on Communications, 2007, pp. 205 –210.

[69] ISO/IEC 14496-2, "Information technology - Coding of Audio-Visual Objects-Part 2: Visual", 1999.

[70] ISO/IEC 14496-10, "Information technology - Coding of Audio-Visual Objects - Part 10: Advanced Video Coding", 1999.

[71] Windows Media Codec-Working with video, http://msdn.microsoft.com/en-us/library/ff819508(v=VS.85).aspx#WindowsMediaVideo9VCM, last access on 9 May 2012.

[72] Windows Media Codec-Working with audio, http://msdn.microsoft.com/en-us/library/ff819484(v=vs.85).aspx, last access on 9 May. 2012.

[73] H. Schulzrinne, S. Casner, R. Frederick, V. Jacobson, "RTP: A Transport Protocol for Real-time Applications," RFC1889, January 1996, http://www.ieft.org/rfc/rfc1889.txt.

[74] H. Schulzrinne, A. Rao, R. Lanphier, "Real Time Streaming Protocol (RTSP)," RFC2326, April 1998, http://www.ieft.org/rfc/rfc2326.txt.

[75] ITU-T Recommendation H.225.0, "Call Signalling Protocols and Media Stream Packetization for Packet-based Multimedia Communication Systems," Nov. 2000.

[76] R.S. Prasad, M. Murray, C. Dovrolis, and K.C. Claffy, "Bandwidth Estimation: Metrics, Measurement Techniques, and Tools," IEEE Network, vol. 17, no. 6, pp. 27-35, Dec. 2003.

[77] S. Shah, K .Chen and K. Nahrstedt, "Dynamic Bandwidth Management in Single-Hop Ad Hoc Wireless Networks," Journal. Mobile Networks and Applications, vol. 10, no. 1-2, pp. 199-217, Feb. 2005.

[78] M.Li, M.Claypool, and R.Kinicki, "Playout Buffer and Rate Optimization for Streaming over IEEE 802.11 Wireless Networks," ACM Trans. Multimedia Computing, Communications and Applications, vol. 5, no. 3, article 26, Aug. 2009.

[79] D.C.A. Bulterman and L. Rutledge, "SMIL 3.0 – Flexible Multimedia for Web, Mobile Devices and Daisy Talking Books," Springer, Heidelberg, 2009.

[80] C.H. Muntean, "Improving Learner Quality of Experience by Contenet Adaptation based on Network Conditions," Journal of Computers in Human Behavior, vol. 24, no. 4, pp. 1452-1472.

[81] S. Moebs, J. McManis, "From Delphi to Simulations: How Network Conditions Affect Learning" IEEE Learning Technologies Newsletter, vol. 12, issue. 4, 2010.

[82] J. Strauss, D.Katabi and F.Kaashoek, "A Measurement Study of Available Bandwidth Estimation Tools", the 3rd ACM SIGCOMM conference on Internet measurement, Miami Beach, FL, USA, October, (2003).

[83] R. L. Carter and M.E. Crovella, "Measuring Bottleneck Link Speed in Packet-switched Networks", Performance Evaluation, Elsevier Science Publishers B. V., vol.27-28, no.8, pp.297-318, Oct.1996.

[84] K.Lai, M. Baker, "Nettimer: A Tool for Measuring Bottleneck Link Bandwidth",

USENIX Symposium on Internet Technologies and Systems, San Francisco, CA, Mar. 2001.

[85] S. Saroiu, P. K. Gummadi, and S. D. Gribble, "Sprobe: A Fast Technique for Measuring Bottleneck Bandwidth in Uncooperative Environments," Aug. 2001, Online: http://sprobe.cs.washington.edu/.

[86] C. Dovrolis, P. Ramanathan, and David Moore, "Packet-dispersion Techniques and A Capacity-Estimation Methodology," IEEE/ACM Trans. Networking, vol. 12, no. 6, pp. 963-977, Dec. 2004.

[87] V. Paxson, "Measurements and Analysis of End-to-end Internet Dynamics," Ph.D. dissertation, University of California, Berkeley, Apr. 1997.

[88] M. Li, M. Claypool, and R. Kinicki, "WBest: A Bandwidth Estimation Tool for IEEE 802.11 Wireless Networks", IEEE Int. Conf. Local Computer Networks (LCN), Montreal, Canada, Oct.2008, pp. 374-381.

[89] V. Ribeiro, R.Riedi, R.Baraniuk, J.Navratil, and L. Cottrell, "pathchirp: Efficient Available Bandwidth Estimation for Network Paths", Passive and Active Measurement Workshop, La Jolla, CA, USA, Apr. 2003.

[90] M. Jain and C. Dovrolis, "End-to-end Available Bandwidth: Measurement Methodology, Dynamics, and Relation with TCP Throughput," IEEE/ACM Trans. in Networking, , no. 295-308, Aug. 2003

[91] A.Johnsson, B.Melander, andM.Bjo¨rkman,"DietTOPP: A First Implementation and Evaluation of A Simplified Bandwidth Measurement Method," Swedish National Computer Networking Workshop, 2004.

[92] Ningning Hu and Peter Steenkiste, "Evaluation and Characterization of Available Bandwidth Probing Techniques," IEEE Journal. Selected Areas. Communications, vol. 21, no. 6, pp. 879-894, Aug. 2003.

[93] J. Strauss, D.Katabi and F.Kaashoek, "A Measurement Study of Available Bandwidth Estimation Tools", 3rd ACM SIGCOMM Conf. Internet Measurement, Miami Beach, FL, USA, Oct. 2003, pp. 39-44.

[94] K. Lakshminarayanan, V. N. Padmanabhan, and J. Padhye, "Bandwidth Estimation in Broadband Access Networks," Proc. ACM SIGCOMM Conf. Internet Measurement, Sicily, Italy, Oct. 2004, pp. 314-321.

[95] T. Sun, G. Yang, L. Chen, M.Y. Sanadidi, and M. Gerla, "A Measurement Study of Path Capacity in 802.11b based Wireless Networks," Proc. International Workshop on Wireless Traffic Measurements and Modeling, Seattle, WA, USA, Jun. 2005, pp. 31-37.

[96] H.K.Lee, V.Hall, K.H.Yum, K.I.Kim and E.J.Kim, “Bandwidth Estimation in Wireless LANs for Multimedia Streaming Services,” IEEE International Conf. Multimedia and Expo (ICME), Toronto, Canada, Jul. 2006, pp. 1181-1184.

[97] S. Shah, K .Chen and K. Nahrstedt, “Dynamic Bandwidth Management in Single-hop Ad Hoc Wireless Networks,” Journal. Mobile Networks and Applications, vol. 10, no. 1-2, pp. 199-217, Feb. 2005.

[98] R. Rejaie, M. Handley and D. Estrin, “RAP: An End-to-End Rate based Congestion Control Mechanism for Real-time Streams in the Internet,” Proceedings of IEEE INFOCOM. Vol. 3, 1999, pp 1337-1345.

[99] D. Chiu and R. Jain. Analysis of the Increase and Decrease Algorithm for Congestion Avoidance in Computer Networks. Journal of Computer Networks and ISDN, pp. 1-14, Jun. 1989.

[100] J.Padhye, J.Kurose, D.Towsley and R.Koodli, “A Model Based TCP Friendly Rate Control Protocol,” Proceedings International Workshop on Network and Operating System Support for Digital Audio and Video-NOSSDAV, 1999.

[101] M. Handley, S. Flyod, J. Padhye, J. Midmer, “TCP Firendly Rate Control (TFRC): Protocol Specification,” RFC 3448, Jan. 2003.

[102] D. Sisalem and A. Wolisz, “LDA+: A TCP Friendly Adaptation Scheme for Multimedia Communication,” Proceedings of the International Conference on Multimedia & Expo, vol. 3, Jul – Aug. 2000, pp. 1619-1622.

[103] H.Schulzrinne, S.Casner, R.Frederick, V.Jacobson, “RTP: A Transport Protocol for Real-Time Applications,” RFC 1889, Jan. 1996.

[104] J. Padhye, V. Firoiu, D. Towsley, and J. Kurose, “Modeling TCP Throughput: A Simple Model and its Empirical Validation,” ACM conference on Applications, technologies, architectures, and protocols for computer communication (SIGCOMM), Vancouver, Canada, Setp. 1998, pp. 303-314.

[105] R.Rejaie, M.Handley and D.Estrin, “Layered Quality Adaptation for Internet Video Streaming,” IEEE Journal on Selected Areas of Communications (JSAC), Special Issue on Internet QoS, vol. 18, no. 12, pp. 2530-2543, Dec. 2000.

[106] G.-M. Muntean, P. Perry and L. Murphy, "A New Adaptive Multimedia Streaming System for All-IP Multi-Service Networks," IEEE Transactions on Broadcasting, vol. 50, no. 1, pp. 1-10, Mar. 2004.

[107] M. Iqbal, X. Wang, S. Li, and T. Ellis, “QoS Scheme for Multimedia Multicast Communications over Wireless Mesh Networks,” IET Communications, vol. 4, no. 11, pp. 1312-1324, Jul. 2010.

[108] Qianhui Liang, HoongChuin Lau, and Xindong Wu, "Robust Application-Level QoS Management in Service-Oriented Systems," in IEEE International Conference on e-Business Engineering, Xi'an, China, 2008, pp. 239-246.

[109] T. Ozcelebi, M. O. Sunay, M. R. Civanlar, and A. M. Tekalp, "Application-Layer QoS Fairness in Wireless Video Scheduling," in 2006 IEEE International Conference on Image Processing, 2006, pp. 1673-1676.

[110] R.Stewart, M. Ramalho, Q. Xie, M. Tuexen, and P. Conrad, "SCTP Partial Reliability Extension," IETF RFC 3758, May. 2004.

[111] H. Shimonishi, T. Hama, and T. Murase, "TCP Congestion Control Enhancements for Streaming Media," in 4th IEEE Consumer Communications and Networking Conference, Las Vegas, United States, Jan. 2007, pp. 303-307.

[112] Unghee Lee and S. F. Midkiff, "Quality of Service for TCP over Satellite Links in Congested Networks," in IEEE Wireless Communications and Networking Conference, Chicago, United States, Sept. 2000, pp. 1515- 1520.

[113] N. Wakamiya, M. Murata, and H. Miyahara, "On TCP-friendly Video Transfer with Consideration on Application-level QoS," in IEEE International Conference on Multimedia and Expo, New York, United States, Jul. 2000, vol. 2, pp. 843-846.

[114] K. Ijiri, S. Ohzahata, and K. Kawashima, "TCP Window Control for QoS Improvement based on Channel Occupancy Information over Wireless LAN," in 8th IEEE International Conference on Industrial Informatics (INDIN), Osaka, Japan, Jul. 2010, pp. 1016-1021.

[115] I.Kim, Y.Kim, M.Kang, J.Mo, and D.Kwak, "TCP-MR: Achieving End-to-end Rate Guarantee for Real-time Multimedia," in 2nd International Conference on Communications and Electronics, HoiAn, Vietnam, Jun. 2008, pp. 80-85.

[116] Hung-Yun Hsieh and R. Sivakumar, "Parallel Transport: A New Transport Layer Paradigm for Enabling Internet Quality of Service," IEEE Communications Magazine, vol. 43, no. 4, pp. 114- 121, Apr. 2005.

[117] R. Y. Brockett and I. Rubin, "Transport Layer Adaptable Rate Control (TARC): Analysis of a Transport Layer Flow Control Mechanism for High Speed Networks," in 7th Annual Joint Conference of the IEEE Computer and Communications Societies, 1998, vol. 2, pp. 540-547.

[118] J. Wroclawski, "The Use of RSVP with IETF Integrated Services," RFC 2210, Sep. 1997.

[119] S. Blake, D. Black, M. Carlson, E. Davis, Z. Wang, W. Weiss, "An Architecture for Differentiated Services" RFC 2475, Dec. 1998.

[120] G. Gorbil and I. Korpeoglu, “Supporting QoS traffic at the Network Layer in Multi-hop Wireless Mobile Networks,” in International conference on Wireless Communications and Mobile Computing Conference (IWCMC), 2011, pp. 1057-1062.

[121] V. Visoottiviseth and S. Siwamogsatham, “End-to-end QoS-aware Handover in Fast Handovers for Mobile IPv6 with DiffServ using IEEE802.11e/IEEE802.11k,” in 10th International Conference on Advanced Communication Technology, vol. 3, pp. 1548-1553, 2008.

[122] Wei Wei, QingjiZeng, Yong Ouyang, and D. Lomone, “Differentiated Integrated QoS Control in the Optical Internet,” IEEE Communications Magazine, vol. 42, no. 11, p. S27- S34, Nov. 2004.

[123] L. Wang, M. Gao, and C. Ji, “Research of QoS Routing Algorithm Suitable for IPv6,” in 2010 International Conference on Computer Design and Applications (ICCDA), 2010, vol. 5, pp. V5-267-V5-269.

[124] R. Asokan, “A Review of Quality of Service (QoS) Routing Protocols for Mobile Ad Hoc Networks,” in International Conference on Wireless Communication and Sensor Computing, Tamil Nadu, India, Jan. 2010, pp. 1-6.

[125] D. Dharmaraju, A. Roy-Chowdhury, P. Hovareshti, and J. S. Baras, “INORA-A Unified Signalling and Routing Mechanism for QoS Support in Mobile Ad Hoc Networks,” in International Conference on Parallel Processing Workshops, Taipei City, Taiwan, Sept. 2002, pp. 86- 93.

[126] N. H. Vaidya, A. Dugar, S. Gupta, and P. Bahl, “Distributed Fair Scheduling in a Wireless LAN.” IEEE Transactions on Mobile Computing, vol. 4, no. 6, pp. 616–629, 2005.

[127] S.J. Golestani, “A Self-Clocked Fair Queueing Scheme for Broadband Applications,” IEEE International Conference on Computer Communications (INFOCOM), Toronto, Canada, Jun. 1994, pp. 636-646.

[128] Qiong Li and M. van der Schaar, “Providing Adaptive QoS to Layered Video Over Wireless Local Area Networks Through Real-time Retry Limit Adaptation,” IEEE Transactions on Multimedia, vol. 6, no. 2, pp. 278- 290, Apr. 2004.

[129] P. A. Chou and Z. Miao, “Rate-distortion Optimized Sender-driven Streaming over Best-effort Networks,” in IEEE Workshop on Multimedia Signal Processing, Cannes, France, Oct. 2001, pp. 587–592.

[130] Eun-Chan Park, Dong-Young Kim, Chong-Ho Choi, and Jungmin So, “Improving Quality of Service and Assuring Fairness in WLAN Access Networks,” IEEE

Transactions on Mobile Computing, vol. 6, no. 4, pp. 337-350, Apr. 2007.

[131] I-Sheng Liu, F. Takawira, and Hong-Jun Xu, "A Hybrid Token-CDMA MAC Protocol for Wireless Ad Hoc Networks," IEEE Transactions on Mobile Computing, vol. 7, no. 5, pp. 557-569, May. 2008.

[132] B. Li, R.Battiti, and Y.Fang, "Achieving Optimal Performance by Using the IEEE 802.11 MAC Protocol With Service Differentiation Enhancements," IEEE Transactions on Vehicular Technology, vol. 56, no. 3, pp. 1374-1387, May. 2007.

[133] S. Chieochan and E. Hossain, "Wireless Fountain Coding with IEEE 802.11e Block ACK for Media Streaming in Wireline-cum-WiFi Networks: A Performance Study," IEEE Transactions on Mobile Computing, vol. 10, no. 10, pp. 1416-1433, Oct. 2011.

[134] A. Banchs, P. Serrano, and L. Vollero, "Providing Service Guarantees in 802.11e EDCA WLANs with Legacy Stations," IEEE Transactions on Mobile Computing, vol. 9, no. 8, pp. 1057-1071, Aug. 2010.

[135] K. A. Meerja and A. Shami, "Analysis of Enhanced Collision Avoidance Scheme Proposed for IEEE 802.11e-Enhanced Distributed Channel Access Protocol," IEEE Transactions on Mobile Computing, vol. 8, no. 10, pp. 1353-1367, Oct. 2009.

[136] A. Banchs and X. Perez, "Distributed Weighted Fair Queuing in 802.11 Wireless LAN." in Proceedings of IEEE International Conference on Communications (ICC), NY, USA, Apr-May.2002, pp. 3121–3127.

[137] M. Shreedhar and G. Varghese, "Efficient Fair Queueing Using Deficit Round Robin," in Proceedings of ACM SIGCOMM, Massachusetts, USA, Sept.1995, pp. 231–242.

[138] M. van der Schaar, Y. Andreopoulos, and Zhiping Hu, "Optimized Scalable Video Streaming over IEEE 802.11 a/e HCCA Wireless Networks Under Delay Constraints," IEEE Transactions on Mobile Computing, vol. 5, no. 6, pp. 755- 768, Jun. 2006.

[139] P. Zhu, W. Zeng, and C. Li, "Joint Design of Source Rate Control and QoS-Aware Congestion Control for Video Streaming Over the Internet," IEEE Transactions on Multimedia, vol. 9, no. 2, pp. 366-376, Feb. 2007.

[140] Huei-Wen Ferng and Han-Yu Liau, "Design of Fair Scheduling Schemes for the QoS-Oriented Wireless LAN," IEEE Transactions on Mobile Computing, vol. 8, no. 7, pp. 880-894, Jul. 2009.

[141] W. Pattara-Atikom, S. Banerjee, and P. Krishnamurthy, "Starvation Prevention and Quality of Service in Wireless LANs," in International Symposium on IEEE Wireless Personal Multimedia Communications, Honolulu, Hawaii, Oct. 2002, pp.

1078-1082.

[142] H.W. Ferng, C.F. Lee, J.J. Huang, and G.M. Chiu, “Designing aFair Scheduling Mechanism for IEEE 802.11 Wireless LANs,” IEEEComm. Letters, vol. 9, no. 4, pp. 301-303, Apr. 2005.

[143] T. Ozcelebi, M. O. Sunay, M. R. Civanlar, and A. M. Tekalp, “Application-Layer QoS Fairness in Wireless Video Scheduling,” in 2006 IEEE International Conference on Image Processing, 2006, pp. 1673-1676.

[144] Y. Xiao, M. Nolen, X. Du, and J. Zhang, "Simulating MPEG-4 over the IEEE 802.11 Wireless Local Area Networks," IEEE Wireless Communications and Networking Conference (WCNC), Hong Kong, China, Mar. 2007, pp. 2660-2664.

[145] C. M. Chen, C. W. Lin, and Y. C. Chen, “Cross-Layer Packet Retry Limit Adaptation for Video Transport Over Wireless LANs,” IEEE Transactions on Circuits and Systems for Video Technology, vol. 20, no. 11, pp. 1448-1461, Nov. 2010.

[146] S. Rane, P. Baccichet, and B. Girod, “Systematic Lossy Error Protection of Video Signals,” IEEE Trans. Circuits Syst. Video Technol., vol. 18, no.10, pp. 1347–1360, Oct. 2008.

[147] Q. Xia, X. Jin, and M. Hamdi, “Cross Layer Design for the IEEE 802.11 WLANs: Joint Rate Control and Packet Scheduling,” IEEE Transactions on Wireless Communications, vol. 6, no. 7, pp. 2732-2740, Jul. 2007.

[148] J. A-Z, C. Verikoukis, E. Kartsakli, A. Cateura, and L. Alonso, “A Near-optimum Cross-layered Distributed Queuing Protocol for Wireless LAN,” IEEE Wireless Communications, vol. 15, no. 1, pp. 48-55, Feb. 2008.

[149] Jia Tang and Xi Zhang, “Cross-layer Modelling for Quality of Service Guarantees over Wireless Links,” IEEE Transactions on Wireless Communications, vol. 6, no. 12, pp. 4504-4512, Dec. 2007.

[150] Qingwen Liu, Shengli Zhou, and G. B. Giannakis, “Cross-layer Scheduling with Prescribed QoS Guarantees in Adaptive Wireless Networks,” IEEE Journal on Selected Areas in Communications, vol. 23, no. 5, pp. 1056- 1066, May. 2005.

[151] Qian Zhang and Ya-Qin Zhang, “Cross-Layer Design for QoS Support in Multihop Wireless Networks,” Proceedings of the IEEE, vol. 96, no. 1, pp. 64-76, Jan. 2008.

[152] M.H. Ahmed, “Call Admission Control in Wireless Networks: a Comprehensive Survey,” IEEE Communications Surveys & Tutorials, no. 1, vol. 7, pp. 49-68, May. 2005.

[153] Y. Hadjadj-Aoul and T. Taleb, “An Adaptive Fuzzy-based CAC Scheme for Uplink and Downlink Congestion Control in Converged IP and DVB-S2 Networks,” IEEE

Transactions on Wireless Communications, vol. 8, no. 2, pp. 816-825, Feb. 2009.

[154] J. Zhu and O.F. Abraham, "A New Call Admission Control Method for Providing Desired Throughput and Delay Performance in IEEE 802.11e wireless LANs," IEEE Transactions on Wireless Communications, vol. 6, no. 2, pp. 701-709, Feb. 2007.

[155] M. AssiChadi, Agarwal Anjali, and Yi Liu, "Enhanced Per-Flow Admission Control and QoS Provisioning in IEEE 802.11e Wireless LANs," IEEE Transactions on Vehicular Technology, vol. 57, no. 2, pp. 1077-1088, Mar. 2008.

[156] L. Lin, H. Fu, and W. Jia, "An Efficient Admission Control for IEEE802.11 Networks based on Throughput Analysis of Unsaturated Traffic," in Proc. IEEE Globecom'05, Dec. 2005, pp. 3017–3021.

[157] A. Abdrabou and W.Zhuang, "Stochastic Delay Guarantees and Statistical Call Admission Control for IEEE 802.11 Single-Hop Ad Hoc Networks," IEEE Trans. Wireless Communications, vol. 7, no. 10, pp. 3972-3981, Otc. 2008.

[158] F. Didi, M. Feham, H. Labiod, and G. Pujolle, "Dynamic Admission Control Algorithm for WLANs 802.11," 3rd International Conf. Information and Communication Technologies: From Theory to Applications, 2008.

[159] J. Liu and Z. Niu, "A Dynamic Admission Control Scheme for QoS Supporting in IEEE 802.11e EDCA," IEEE Wireless Communications and Networking Conferencde (WCNC), Hong Kong, China, Mar. 2007, pp. 3700-3705.

[160] P. Dini, N. Baldo, J. Nin-Guerrero, J. Mangues-Bafalluy, S. Addepalli, and L.L. Dai, "Distributed Call Admission Control for VoIP over 802.11 WLANs Based on Channel Load Estimation," IEEE International Conf. Communications (ICC), Cape Town, South Arica, May.2010, pp. 1-6.

[161] S. Sangho and H. Schulzrinne, "Call Admission Control in IEEE 802.11 WLANs Using QP-CAT," 27th IEEE Conf. Computer Communications (INFOCOM), Apr. 2008, pp. 726-734.

[162] E. Rich, "User Modeling via stereotypes," Cognitive Science Journal, Vol. 3, No. 4, 1979, pp. 329-354.

[163] C. H. Muntean and J. McManis, "A QoS-aware adaptive Web-based system," IEEE International Conf. Communications (ICC), Paris, France, Jun. 2004, pp. 2204-2208.

[164] S.L. Amundsen and F. Eliassen, "A Resource and Context Model for Mobile Middleware," Journal of Personal and Ubiquitous Computing-Special Issue: Selected Papers of the ARCS06 Conference, vol. 12, no. 2, pp. 143-153, Jan. 2008.

[165] G. Ortiz and B. Bordbar, "Model-Driven Quality of Service for Web Services: An Aspect-Oriented Approach," IEEE International Conference Web Services, Beijing,

China, pp. 748-751, Sept. 2008.

[166] Kuen-Rong Lo, Chung-JuChang, and C. B. Shung, "A QoS-Guaranteed Fuzzy Channel Allocation Controller for Hierarchical Cellular Systems," IEEE Transactions on Vehicular Technology, vol. 49, no. 5, pp. 1588-1598, 2000.

[167] S. Chandramathi and S. Shanmugavel, "Adaptive Allocation of Resources with Multiple QoSHeterogenous Sources in ATM Networks - A Fuzzy Approach," IEEE International Conference on Communications, New York, USA, Apr. 2002, pp. 1279-1283.

[168] D. Niyato and E. Hossain, "Delay-Based Admission Control Using Fuzzy Logic for OFDMA Broadband Wireless Networks," IEEE International Conference on Communications, Istanbul, Turkey, Jun. 2006, pp. 5511-5516.

[169] D. Todinca, "Novel admission control algorithm for GPRS/EGPRS based on fuzzy logic," 2004, 5st IEE International Conference on 3G Mobile Communication Technologies, London, UK, Oct. 2004, pp. 342-346.

[170] P. Kejik and S. Hanus, "Fuzzy logic based call admission control in UMTS system," 20th International Conference on Radioelektronika, 2009, pp. 375-378.

[171] M. Ditze and M. Grawinkel, "Fuzzy Logic Based Admission Control for Multimedia Streams in the UPnP QoS Architecture," 21stInternational Conference on Advanced Information Networking and Applications Workshops, Ontario, Canada, May. 2007, pp. 970-976.

[172] C. Khuen, C. Yong, and F. Haron, "Multi-Agent Negotiation System Using Adaptive Fuzzy Logic in Resource Allocation," 2nd International Conference on Distributed Frameworks for Multimedia Applications, Penang, Malaysia, May. 2006, pp. 1-7.

[173] M. Jamshidi, Large-Scale Systems: Modeling, Control and Fuzzy Logic. Prentice Hall PTR, 1997.

[174] Hang Su and Xi Zhang, "Clustering-Based Multichannel MAC Protocols for QoS Provisioning Over Vehicular Ad Hoc Networks," IEEE Transactions on Vehicular Technology, vol. 56, no. 6, pp. 3309-3323, Nov. 2007.

[175] S. Thenmozhi and A. Tamilarasi, "A Cluster based Resource Allocation Architecture for Mobile Grid Environments," International Conference on Computing Communication and Networking Technologies, Zurich, Switzerland, July. 2010, pp. 1-5.

[176] A. Hatoum, N. Aitsaadi, R. Langar, R. Boutaba, and G. Pujolle, "FCRA: Femtocell Cluster-Based Resource Allocation Scheme for OFDMA Networks," in IEEE International Conference on Communications (ICC), Kyoto, Japan, Jun. 2011, pp. 1-

6.

[177] Ho Ting Cheng and Weihua Zhuang, “Novel Packet-level Resource Allocation with Effective QoS Provisioning for Wireless Mesh Networks,” IEEE Transactions on Wireless Communications, vol. 8, no. 2, pp. 694-700, Feb. 2009.

[178] Ho Ting Cheng and WeihuaZhuang, “Pareto Optimal Resource Management for Wireless Mesh Networks with QoS Assurance: Joint Node Clustering and Subcarrier Allocation,” IEEE Transactions on Wireless Communications, vol. 8, no. 3, pp. 1573-1583, Mar. 2009.

[179] S. Bashar and Zhi Ding, “Admission Control and Resource Allocation in A Heterogeneous OFDMA Wireless Network,” IEEE Transactions on Wireless Communications, vol. 8, no. 8, pp. 4200-4210, Aug. 2009.

[180] Zhuwei Wang, Qihang Peng, and L. B. Milstein, “Multi-User Resource Allocation for Downlink Multi-Cluster Multicarrier DS CDMA System,” IEEE Transactions on Wireless Communications, vol. 10, no. 8, pp. 2534-2542, Aug. 2011.

[181] Y. Wang, Y. Hu, Y. Jia, and R. Zhang, “A Game Theoretic Approach to Product-mix Resource Allocation,” 5th IEEE Industrial Electronics and Applications (ICIEA), Taiwan, China, Jun. 2010, pp. 2142-2147.

[182] T. Taleb and A. Nafaa, “A Fair and Dynamic Auction-Based Resource Allocation Scheme for Wireless Mobile Networks,” IEEE International Conference on Communications (ICC), Beijing, China, May.2008, pp. 306-310.

[183] L. Berlemann et al., “Radio Resource Sharing Games: Enabling QoS Support in Unlicensed Bands,” IEEE Network, vol. 19, no. 4, July–Aug. 2005, pp. 59–6.

[184] G. Tan and J. Guttag, “The 802.11 MAC Protocol Leads to Inefficient Equilibria,” Proc. IEEE INFOCOM ’05, vol.1, Mar. 2005, pp. 1–11.

[185] Dusit Niyato and Ekram Hossain, “Radio Resource Management Games in Wireless Networks: An Approach to Bandwidth Allocation and Admission Control for Polling Service in IEEE 802.16 [Radio Resource Management and Protocol Engineering for IEEE 802.16],” IEEE Wireless Communications, vol. 14, no. 1, pp. 27-35, Feb. 2007.

[186] Tiankui Zhang, Zhimin Zeng, Chunyan Feng, Jieying Zheng, and Dongtang Ma, “Utility Fair Resource Allocation Based on Game Theory in OFDM Systems,” in Proceedings of 16th International Conference on Computer Communications and Networks, 2007. ICCCN 2007, 2007, pp. 414-418.

[187] C. Seneviratne and H. Leung, “A Game Theoretic Approach for Resource Allocation in Cognitive Wireless Sensor Networks,” in 2011 IEEE International Conference on Systems, Man, and Cybernetics (SMC), 2011, pp. 1992-1997.

[188] Z. Yuan, H. Venkataraman, and G.-M. Muntean, "iBE: A Novel Bandwidth Estimation Algorithm for Multimedia Services over IEEE 802.11 Wireless Networks," 12th IFIP/IEEE International Conference on Management of Multimedia and Mobile Networks and Services, Venice, Italy, Oct.2009, pp. 69-80.

[189] Z. Yuan, H. Venkataraman, and G.-M. Muntean, "A Model-based Achievable Bandwidth Estimation for IEEE 802.11 Data Transmissions," to appear in IEEE Transactions on Vehicular Technology, 2012.

[190] J. Padhye, V. Firoiu, D. Towsley, and J. Kurose, “Modeling TCP Reno Performance: A Simple Model and its Empirical Validation,” IEEE/ACM Trans. Networking, vol. 8, no. 2, pp. 133-145, Apr. 2000.

[191] P.Chatzimisios, A. C. Boucouvalas, and V. Vitsas, “Performance Analysis of IEEE 802.11 DCF in Presence of Transmission Errors,” IEEE Int. Conf. Communications (ICC), Paris, France, Jun. 2004, pp.3854-3858.

[192] B. Braden, D. Clark, J. Crowcroft, B. Davie, S. Deering, D. Estrin, S. Floyd, V. Jacobson, G. Minshall, C. Partridge, L. Peterson, K. Ramakrishnan, S. Shenker, J. Wroclawski, L. Zhang, RFC 2309: Recommendations on Queue Management and Congestion Avoidance in the Internet, April 1998.

[193] W. Stevens, M. Allman, and V. Paxson, "TCP Congestion Control", RFC 2581, April 1999.

[194] R.Bruno, M.Conti, and E.Gregori, “Throughput Analysis and Measurements in IEEE 802.11 WLANs with TCP and UDP Traffic Flows,” IEEE Trans. Mobile Computing, vol. 7, no. 2, pp. 171-186, Feb. 2008.

[195] G. Venkatesan, “Multimedia Streaming over 802.11 Links [Industry Perspectives],” IEEE Wireless Communications, vol. 17, no. 2, pp. 4-5, 2010.

[196] M. Lopez-Benitez and J. Gozalvez, “Common Radio Resource Management Algorithms for Multimedia Heterogeneous Wireless Networks,” IEEE Transactions on Mobile Computing, vol. 10, no. 9, pp. 1201-1213, Sep. 2011.

[197] J. C. Fernandez, T. Taleb, M. Guizani, and N. Kato, “Bandwidth Aggregation-Aware Dynamic QoS Negotiation for Real-Time Video Streaming in Next-Generation Wireless Networks,” IEEE Transactions on Multimedia, vol. 11, no. 6, pp. 1082-1093, Oct. 2009.

[198] S. Chieochan and E. Hossain, “Wireless Fountain Coding with IEEE 802.11e Block ACK for Media Streaming in Wireline-cum-WiFi Networks: A Performance Study,” IEEE Transactions on Mobile Computing, vol. 10, no. 10, pp. 1416-1433, Oct. 2011.

[199] A. Banchs, P. Serrano, and L. Vollero, “Providing Service Guarantees in 802.11e

EDCA WLANs with Legacy Stations," IEEE Transactions on Mobile Computing, vol. 9, no. 8, pp. 1057-1071, Aug. 2010.

[200] K. A. Meerja and A. Shami, "Analysis of Enhanced Collision Avoidance Scheme Proposed for IEEE 802.11e-Enhanced Distributed Channel Access Protocol," IEEE Transactions on Mobile Computing, vol. 8, no. 10, pp. 1353-1367, Oct. 2009.

[201] G.-M. Muntean, P. Perry and L. Murphy, "A New Adaptive Multimedia Streaming System for All-IP Multi-Service Networks," IEEE Transactions on Broadcasting, vol. 50, no. 1, pp. 1-10, Mar. 2004.

[202] H. Venkataraman and G.M. Muntean, "Dynamic Time Slot Partitioning (DTSP) for Multimedia Streaming in Two-Hop Cellular Networks", IEEE Trans. Mobile Computing, vol. 10, no. 4, pp. 532-543, Apr. 2011.

[203] C. H. Muntean and J. McManis, "Fine Grained Content-based Adaptation Mechanism for Providing High End-user Quality of Experience with Adaptive Hypermedia Systems," Proceedings of the 15th International Conference on World Wide Web , pp. 53-62, 2006.

[204] Y. Hadjadj-Aoul and T. Taleb, "An Adaptive Fuzzy-based CAC Scheme for Uplink and Downlink Congestion Control in Converged IP and DVB-S2 Networks," IEEE Transactions on Wireless Communications, vol. 8, no. 2, pp. 816-825, Feb. 2009.

[205] J. Zhu and O.F. Abraham, "A New Call Admission Control Method for Providing Desired Throughput and Delay Performance in IEEE 802.11e wireless LANs," IEEE Transactions on Wireless Communications, vol. 6, no. 2, pp. 701-709, Feb. 2007.

[206] M. Assi Chadi, Agarwal Anjali, and Yi Liu, "Enhanced Per-Flow Admission Control and QoS Provisioning in IEEE 802.11e Wireless LANs," IEEE Transactions on Vehicular Technology, vol. 57, no. 2, pp. 1077-1088, Mar. 2008.

[207] W. Richard Stevens, "TCP/IP Illustrated Volume 1: the Protocols," Addison-Wesley, 1994.

[208] ITU-T G.1010, "End-user multimedia QoS categories," 2001.

[209] G. Venkatesan, "Multimedia Streaming over 802.11 Links [Industry Perspectives]," IEEE Wireless Communications, vol. 17, no. 2, pp. 4-5, Apr. 2010.

[210] Presentations by ABI Research, Picochip, Airvana, IP.access, Gartner, Telefonica Espana, 2nd Int'l. Conf. Home Access Points and Femtocells;http://www.avrenevents.com/dallasfemto2007/purchase_presentations.htm, 2007.

[211] H. Park, S. Pack, and C. H. Kang, "Dynamic adaptation of contention window for consumer devices in WiMedia home networks," IEEE Transactions on Consumer

Electronics, vol. 57, no. 1, pp. 28-34, Feb. 2011.

[212] P. Corcoran, C. Iancu, F. Callaly, and A. Cucos, “Biometric Access Control for Digital Media Streams in Home Networks,” IEEE Transactions on Consumer Electronics, vol. 53, no. 3, pp. 917-925, Aug. 2007.

[213] V. Miguel, J. Cabrera, F. Jaureguizar, and N. Garcia, “Distribution of high-definition video in 802.11 wireless home networks,” IEEE Transactions on Consumer Electronics, vol. 57, no. 1, pp. 53-61, Feb. 2011.

[214] Z. Yuan, H. Venkataraman, and G.-M. Muntean,"iPAS: An User Perceived Quality-based Intelligent Prioritized Adaptive Scheme for IPTV in Wireless Home Networks," IEEE International Symposium on Broadband Multimedia Systems and Broadcasting, Shanghai, China, Mar.2010, pp. 1-6.

[215] P. Lin, W-I Chou, and T. Lin, “Achieving airtime fairness of delay-sensitive applications in multirate IEEE 802.11 wireless LANs,” IEEE Communications Magazine, vol. 49, no. 9, pp. 169-175, Sep. 2011.

[216] B. Hirantha Sithira Abeysekera, T. Matsuda, and T. Takine, “Dynamic Contention Window Control Mechanism to Achieve Fairness between Uplink and Downlink Flows in IEEE 802.11 Wireless LANs,” IEEE Transactions on Wireless Communications, vol. 7, no. 9, pp. 3517-3525, Sep. 2008.

[217] K. Kashibuchi, A. Jamalipour, and N. Kato, “Channel Occupancy Time Based TCP Rate Control for Improving Fairness in IEEE 802.11 DCF,” IEEE Transactions on VehicularTechnology,vol.59, no. 6, pp. 2974-2985, Jul. 2010.

[218] W-S Lim, D-W. Kim, and Y-J. Suh, “Achieving Fairness Between Uplink and Downlink Flows in Error-Prone WLANs,” IEEE Communications Letters, vol. 15, no. 8, pp. 822-824, Aug. 2011.

[219] S. Yun, H. Kim, H. Lee, and I. Kang, “100+ VoIP Calls on 802.11b: The Power of Combining Voice Frame Aggregation and Uplink-Downlink Bandwidth Control in Wireless LANs,” IEEE Journal on Selected Areas in Communicaitons, vol. 25, no. 4, pp. 689-698, May. 2007.

[220] B. A. Hirantha Sithira Abeysekera, T. Matsuda, and T. Takine, "Dynamic Contention Window Control Scheme in IEEE 802.11e EDCA-Based Wireless LANs ," IEICE Transactions on Communications, vol. E93-B, no.1, pp.56-64, Jan. 2010.

[221] C.H. Muntean and J. McManis, “End-user Quality of Experience oriented adaptive e-learning system,” Journal of Digital Information, Special Issue on Adaptive Hypermedia, vol. 7, no. 1, 2006.

[222] S.L. Amundsen and F. Eliassen, “A Resource and Context Model for Mobile

Middleware," Journal of Personal and Ubiquitous Computing-Special Issue: Selected Papers of the ARCS06 Conf., vol. 12, no. 2, pp. 143-153, Jan. 2008.

[223] Q. Xia, X. Jin, and M. Hamdi, "Active Queue Management with Dual Virtual Proportional Integral Queues for TCP Uplink/Downlink Fairness in Infrastructure WLANs," IEEE Transactions on Wireless Communications, vol. 7, no. 6, pp. 2261-2271, Jun. 2008.

[224] S. Haykin, Communication Systems, John Wiley & Sons, 2000.

[225] Douglas E. Comer, Internetworking with TCP/IP (5 ed.), Prentice Hall: Upper Saddle River, NJ, 2006.

[226] S. Floyd, V. Jacobson, "Random Early Detection gateways for congestion avoidance," IEEE/ACM Transactions on Networks, vol. 1, no. 4, pp.397–413, Aug. 1993.

[227] Y. Xi, B.-S. K. J. Wei, and Q.-Y. Huang, "Adaptive Multirate Auto Rate Fallback Protocol for IEEE 802.11 WLANS," IEEE Communications Conference in Military (MILCOM), 2006, pp. 1 –7.

[228] T. Taleb and A. Nafaa, ―A Fair and Dynamic Auction-Based Resource Allocation Scheme for Wireless Mobile Networks, IEEE International Conference on Communications, Beijing, China, May.2008, pp. 306-310.

[229] ITU-T G.711, "Pulse coded modulation (PCM) of voiceb frequencies," Nov. 1988.

[230] R. K.Jain, D. W.Chiu, W.R.Have "A Quantitative Measure of Fairness and Discrimination for Resource Allocation in Shared Computer Systems," DEC Research Report TR-301, 1984.

[231] N.Thomos, N.Boulgouris, and M.Strintzis, "Optimized Transmission of JPEG2000 Streams over Wireless Channels," IEEE Transactions on Image Processing, no. 1, vol.15, pp.54-67, Jan. 2006.

[232] K. Chebrolu, B. Raman, and S. Sen, "Long-Distance 802.11b Links: Performance Measurements and Experience," 12th Annual International Conference on Mobile Computing and Networking (MOBICOM), California, USA, Sept. 23-26, 2006.

[233] B. H. S. Abeysekera, T. Matsuda, and T. Takine, "Dynamic Contention Window Control Mechanism to Achieve Fairness between Uplink and Downlink Flows in IEEE 802.11 Wireless LANs," IEEE Transactions on Wireless Communications, vol. 7, no. 9, pp. 3517-3525, Sept. 2008.

[234] ITU-T Recommendation G.114, "One-Way Transmission Time", Feb. 1996.

[235] S. Wanstedt, M. Ericson, L. Hevizi, J. Pettersson, and J. Barta, "The Effect of F-DPCH on VoIP over HSDPA Capacity," Proc. IEEE Vehicular Technology Conference, pp. 410-414, May. 2006.