國立交通大學

資訊科學與工程研究所

博 士 論 文

無線網路環境下網際網路電話之效能分析

VoIP Performance Analysis in the Wireless Environment

研 究 生：宋雅琴

指導教授：林一平　教授

中 華 民 國 九 十 九 年 六 月

無線網路環境下網際網路電話之效能分析

VoIP Performance Analysis in the Wireless Environment

研 究 生：宋雅琴　　Student：Ya-Chin Sung

指導教授：林一平　　Advisor：Yi-Bing Lin

國 立 交 通 大 學

資 訊 科 學 與 工 程 研 究 所

博 士 論 文

A Dissertation
Submitted to Institutes of Computer Science and Engineering
Department of Computer Science
College of Computer Science
National Chiao Tung University
in partial Fulfillment of the Requirements
for the Degree of
Ph. D.
in

Computer Science and Engineering

June 2010

Hsinchu, Taiwan, Republic of China

中華民國九十九年六月

無線網路環境下網際網路電話之效能分析

學生：宋雅琴　　指導教授：林一平　博士

國立交通大學資訊科學與工程研究所博士班

摘要

網際網路電話（VoIP）為目前 IP 網路上最具發展性的通訊方式之一。在無線行動環境下，頻寬資源有限，並且傳送之可靠性較有線環境為差。為了在無線行動網路中提供令人滿意的網際網路電話服務，必須保證網路的服務品質（QoS）。

隨著電波射頻技術的演進，頻寬大幅提升，各種無線網路下都可以提供網際網路電話服務。為達到電信等級的安全度，我們提出整合第三代行動網路與無線網路的認證方法，在無線網路環境下重複利用全球行動通訊系統（UMTS）的認證金鑰，並且針對 IEEE 802.1X 參數作效能分析以獲得較好的認證延遲效能。

此外，我們研究無線網路下網際網路電話封包的傳送效能。第三代行動合作通訊計畫（3GPP）標準文件允許被認證的無線網路透過無線接取閘道器/封包資料閘道器（WAG/PDG）接取第三代行動網路；然而，為確保電信等級的安全度，在行動終端設備（MS）和無線接取閘道器/封包資料閘道器間的網際網路電話封包必須以網際網路通訊協定安全性（IPsec）保護。我們分析在無線網路下以網際網路通訊協定安全性加密後的網際網路電話效能，檢視網際網路通訊協定安全性的額外負擔，包括：封包處理效能，封包遺失率，封包傳送延遲，以及傳送時間抖動。

最後，我們評估車用環境下網際網路電話的效能。基於 M 台灣計畫，我們在台北佈建全球互通微波存取（WiMAX）網路以支援高速行動寬頻服務，並在該網路環境中整合網際網路電話服務。我們研究在上述環境下網際網路電話的效能，其研究成果可作為未來其他先進的網路環境下分析網際網路效能時的參考。

關鍵字: 網際網路電話，服務品質，效能量測，平均意見分數，基於區域網路的延伸認證通訊協定，IEEE 802.1X，第三代行動合作通訊計畫，無線接取閘道器/封包資料閘道器，網際網路通訊協安全性，全球互通微波存取網路，寬頻無線通訊，電信服務，行動通訊

VoIP Performance Analysis in the Wireless Environment

Student: Ya-Chin Sung Advisor: Dr. Yi-Bing Lin

Institute of Computer Science and Engineering

National Chiao Tung University

Abstract

Voice over Internet Protocol (VoIP) is a promising low-cost voice communication over IP networks. In a mobile/wireless environment, the radio resource is restricted and the reliability of the wireless transmission is much poor than that of the wired environment. To provide satisfactory VoIP services in the mobile/wireless network, the *Quality of Service* (QoS) of the network should be guaranteed.

Due to the advances of various wireless technologies, the VoIP service is provisioned in different wireless networks. To achieve telecom grade security, we propose a WLAN-3G integrated security approach that reuses the UMTS authentication key in the WLAN environment and conduct a modeling study to tune the IEEE 802.1X parameters to yield better authentication delay performance.

Furthermore, we investigate the VoIP packet delivery efficiency in the wireless environment (i.e., WLAN network). In 3GPP specifications, the authenticated WLAN MS is allowed to access the 3G network through the WAG/PDG. However, to ensure telecom grade security, the VoIP traffic between the MS and the WAG/PDG must be protected with IPsec. We analyze the performance of IPsec-based VoIP service in a IEEE 802.11b WLAN environment. The IPsec overheads in terms of throughput, packet loss rate, latency, and jitter are investigated.

Finally, we present the VoIP performance in the vehicle environment. We conduct trials in the real WiMAX network which supports high-speed mobile broadband services and investigate the WiMAX-based VoIP of a *Mobile Taiwan* (M-Taiwan) funded program conducted during 2007-08 in the Taipei area. We investigate the VoIP performance in the wireless environment. These research results presented in this dissertation can be viewed as a useful foundation for further VoIP performance study in various advanced wireless environments.

Keywords: Voice over IP, Quality of Service, Performance Measurement, Mean Opinion Score, EAPOL, IEEE 802.1X, UMTS, WAG/PDG, IPsec, WiMAX, Broadband Wireless Communications, Telecommunication Services, Mobile Communications

Acknowledgement

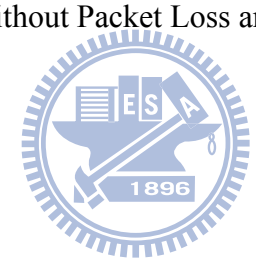
I would like to express my sincere thanks to my advisor, Prof. Yi-Bing Lin. Without his supervision and perspicacious advice, I can not complete this dissertation. Special thanks are due to my committee members: Dr. Chung-Hwa Rao, Dr. Sheng-Lin Chou, Prof. Han-Chieh Chao, Prof. Chu-Sing Yang, Prof. Ming-Feng Chang, and Prof. His-Lu Chao for their valuable comments and helpful discussions.

I also express my appreciation to all the faculty, staff and colleagues in the Department of Computer Science. In particular, I would like to thank Prof. Shun-Ren Yang, Prof. Pei-Chun Lee, Dr. Lin-Yi Wu, Prof. Sok-Ian Sou, Dr. Shih-Feng Hsu, Dr. Yung-Chun Lin, Prof. Meng-Hsun Tsai, Ms. Hsin-Yi Lee, Mr. Chien-Chun Huang-Fu, Mr. Pin-Jen Lin, Mr. Ren Huang Liou, Mr. Jer-Ming Lin, Mr. Chung-Hsiang Ou, and Mr. Kai-Qun Yang in the Laboratory 117 for their friendship and support in various ways.

Finally, I am grateful to my dear parents, my sisters, my brother, and my boy friend for their unfailing love and firmly support in these years.

Contents

Lists of Tables



List of Figures

Notations

The notation used in this dissertation is listed below.

- T_X: the associated timeout period of the message exchange X where $X = s$, a_1, a_2, or a_3
- t_X: the RTT of the message exchange X without waiting delay (i.e., the queueing at a network node) where $X = s$, a_1, a_2, or a_3
- τ_X: the response time of the message exchange X (i.e., the RTT of the message exchange X including the queuing delay) where $X = s$, a_1, a_2, or a_3
- $B_X(\cdot)$: the distribution function of the message exchange service time t_X
- $b_X(\cdot)$: the density function of the message exchange service time t_X
- $B_X^*(\cdot)$: the Laplace Transform of the message exchange service time t_X
- $F_X(\cdot)$: the distribution function of the message exchange response time τ_X
- $f_X(\cdot)$: the density function of the message exchange response time τ_X
- $F_X^*(\cdot)$: the Laplace Transform of the message exchange response time τ_X
- p_X: the timeout probability of the message exchange X where $X = s$, a_1, a_2, or a_3
- λ: the EAPOL message arrival rate to the AP
- p_f: the false failure detection probability of the IEEE 802.1X authentication procedure
- $E[\tau]$: the expected response time of the IEEE 802.1X authentication procedure
- N: the number of RTP streams
- R: the rating factor of E-Model

- R_o: the basic signal-to-noise ratio
- I_s: the simultaneous impairment factor
- I_d: the delay impairment factor
- $I_{e\text{-}eff}$: the effective equipment impairment factor
- A: the advantage factor

Chapter 1
Introduction

Voice over Internet Protocol (VoIP) is a promising low-cost voice communication over the wired or wireless Internet network. In the mobile/wireless environment, the radio resource is restricted and the reliability of the wireless transmission is much poor than that of the wired environment. To provide satisfactory VoIP services in the mobile/wireless network, the *Quality of Service* (QoS) of the mobile/wireless network should be guaranteed.

Table 1.1: Interworking Scenarios and Service Capabilities

Service Capabilities	Scenario 1	Scenario 2	Scenario 3	Scenario 4	Scenario 5	Scenario 6
Common Billing	○	○	○	○	○	○
Common Customer Care	○	○	○	○	○	○
3G-based Access Control	×	○	○	○	○	○
3G-based Access Charging	×	×	○	○	○	○
Access to 3G PS Services	×	×	×	○	○	○
Service Continuity	×	×	×	○	○	○
Seamless Service Continuity	×	×	×	×	○	○
Access to 3G CS Services with Seamless Mobility	×	×	×	×	×	○

Furthermore, due to the advances of various wireless technologies, the VoIP service is provisioned in different wireless networks. These wireless networks can be integrated with different degrees. *Third Generation Partnership Project* (3GPP) Technical Report 22.934

conducts a feasibility study on integrating wireless access technologies with cellular (e.g., *Universal Mobile Telecommunication System*; UMTS) mobile system [1]. Six scenarios are proposed for incremental development of mobile and *Wireless Local Area Network* (WLAN) interworking. Each scenario enhances interworking functionalities over the previous scenarios as illustrated in Table 1.1. The service and operational capabilities of each scenario are described as follows.

Scenario 1 provides common billing and customer care for both WLAN and mobile operators. That is, a customer receives a single monthly billing statement combining both mobile and WLAN services. The customer also consults the same customer care center about problems with both services.

Scenario 2 re-uses cellular-access control and charging mechanisms for WLAN services. The WLAN customers are authenticated by the mobile core network without introducing a separate authentication procedure. In addition, the roaming mechanism between the cellular system and the WLAN is supported. In this scenario, users can access traditional Internet services but cannot access mobile services (such as *Circuit Switched* (CS) voice and GPRS data services) through the WLAN.

Scenario 3 allows a customer to access mobile *Packet Switched* (*PS*) services over the WLAN. The PS services include *Short Message Service* (SMS) [2], *Multimedia Messaging Service* (MMS) [3], and *IP Multimedia Core Network Subsystem* (IMS) Service [4]. Customers equipped with both a WLAN card and a cellular module can simultaneously but independently access WLAN and cellular networks.

Scenario 4 allows a customer to change access between cellular and WLAN networks during a service session. The system is responsible for reestablishing the session without user involvement. Service interruption during system switching is allowed in this scenario.

QoS is a critical issue for service continuity. Since cellular and WLAN networks have different capabilities and characteristics, the user would gain different QoS grades in different networks. Therefore, QoS adaptation is required during system switching.

Scenario 5 provides seamless service switching (that is, handoff) between the cellular system and the WLAN. Techniques must be developed to minimize the data lost rate and delay time during switching so that the customer does not experience significant interruption during handoff.

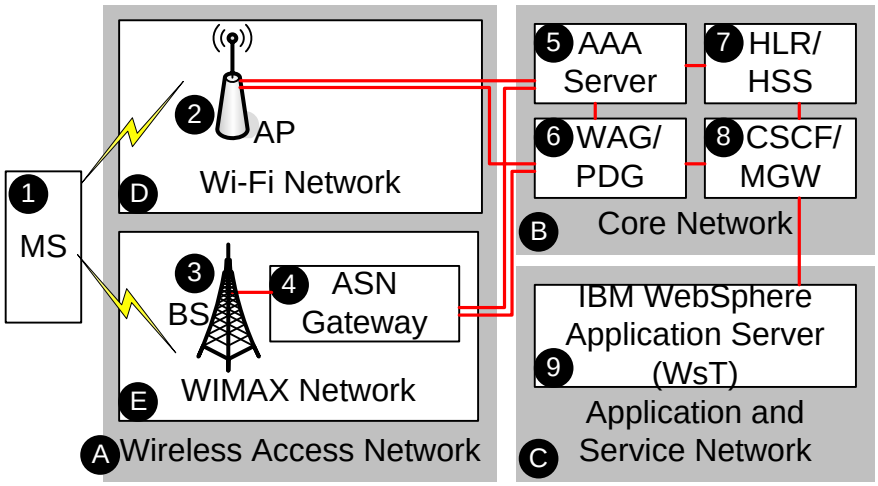
Scenario 6 supports mobile CS services in the WLAN environment. The seamless continuity feature described in Scenario 5 is also required to support CS services when customers roam between different networks.

Our survey with several mobile service providers indicates that the Scenario 3 features are essential for commercial operation of cellular/WLAN interworking in the first stage deployment. Depending on the business strategies, the Scenario 4 features may or may not be deployed in the long-term commercial operation. Scenarios 5 and 6 are typically ignored because the benefits of the extra features might not justify the deployment costs. This dissertation considers scenario 3 that allows a user to carry out VoIP authentication with 3G mobile core network through WLAN. Then we discuss the VoIP performance in the different wireless networks including the WLAN and the mobile *Worldwide Interoperability for Microwave Access* (WiMAX) network [5]. Figure 1.1 illustrates a typical mobile/wireless network model studied in this dissertation. According to the network functionalities, the mobile/wireless network model can be partitioned into three categories:

Wireless Access Network (Figure 1.1 (A)) provides wireless connectivity for the *Mobile Station* (MS; Figure 1.1 (1)).

Core Network (Figure 1.1 (B)) provides the various mobile services to the MS.

Application and Service Network (Figure 1.1 (C)) supports flexible and efficient approaches for mobile service development and deployment through a service platform.



AAA: Authentication Authorization, and Accounting
AP: Access Point
ASN: Access Service Network
BS: Base Station
CSCF/MGW: Call Session Control Function/Media Gateway
HLR/HSS: Home Location Register/Home Subscriber Server
MS: Mobile Station
WAG/PDG: WLAN Access Gateway/Packet Data Gateway
WAMAX: Worldwide Interoperability for Microwave Access
Wi-Fi: Wireless Fidelity
WLAN: Wireless Local Area Network
WsT: WebSphere software for Telecom

Figure 1.1 Mobile/Wireless Network Model

In the remainder of this chapter, we briefly introduce the wireless access network, core network, and application and service network. Then we present the VoIP performance issues concerned in the mobile/wireless network and describe the organization of this dissertation.

1.1 Wireless Access Network

Wireless access network provides wireless connectivity for the MS. In this dissertation, we elaborate on two popular wireless access networks: the WLAN and WiMAX. The WLAN radio network includes 802.11-based *Access Points* (APs; Figure 1.1 (2)) to provide radio

access for the MSs. In the WiMAX radio network, *Base Station* (BS; Figure 1.1 (3)) provides radio access for the MSs and connects to *Access Service Network Gateway* (ASN Gateway; Figure 1.1 (4)), which supports mobility and session management to MS.

To ensure secure communication, both WLAN AP and WiMAX ASN Gateway communicate with the *Authentication Authorization, and Accounting Server* (AAA Server; Figure 1.1 (5)) in the core network. An MS is required to perform the AAA procedure before associating to the wireless access network. The authentication messages between the MS and the AAA server are forwarded by the WiFi AP (or WiMAX ASN Gateway) when the MS attaches to the WLAN (or WiMAX network). To deliver user data in mobile/wireless network, both AP and ASN Gateway connect to *WLAN Access Gateway/Packet Data Gateway* (WAG/PDG; Figure 1.1 (6)) to transmit the packets between an MS and the core network. Details of the core network are described in Section 1.2.

1.2 Core Network

The core network provides various mobile services to the MS, such as SMS, MMS, IMS services, etc. In this dissertation, we focus on the IMS services. Specifically, we conduct performance studies on VoIP.

In the core network, AAA server is responsible to mutually authenticate with the MS when the MS associates to the WLAN and the WiMAX networks. WAG/PDG routes the packets received from/sent to the wireless access network. *Home Location Register/Home Subscriber Server* (HLR/HSS; Figure 1.1 (7)) is the master database containing all subscriber-related information. When receiving the authentication request from an MS, the AAA server retrieves the HLR/HSS to obtain the authentication information of the MS. After

the MS is authenticated, the packets between the wireless access network and the core network are routed through WAG/PDG.

The *Call Session Control Function* (CSCF) and *Media Gateway* (MGW; Figure 1.1 (8)) are in charge of delivering the call control signaling and the voice data respectively. An example of core network is *Chunghwa Telecom* (CHT) *Next Generation Network* (NGN) where IMS plays an important role to offer IP-based multimedia services. In current CHT NGN/IMS deployment, the CSCF is a Nokia Siemens Networks (NSN) CFX-5000, and MGW is an NSN hiG1100.

The core network is integrated with the application and service network by connecting the CSCF with the application server (e.g., IBM WebSphere Application Server; Figure 1.1 (9)) in the application and service network. The CSCF is responsible for processing the requests and responses between the core network and the application and service network. Details of the application and service network are elaborated in Section 1.3.

1.3 Application and Service Network

Application and Service Network supports flexible and efficient approaches for mobile service development and deployment through a service platform. 3GPP defines several alternatives to construct the application and service network platform. In this dissertation, we use the IBM WebSphere application server to illustrate the concept of the application and service network [6]. IBM WebSphere application server is a web application server that offers a platform for creating various web applications. For telecom services, IBM implements *WebSphere software for Telecom* (WsT; Figure 1.1 (9)) to provide a platform for efficient telecom web service development without concerning the core network environment.

The IBM WsT (Figure 1.2 (B)) interacts with the CSCF in the core network (Figure 1.2 (A)) through the *IP Multimedia Subsystem Service Control interface* (ISC interface; Figure 1.2 (1)), which is defined in 3GPP TS 23.228 [4]. The ISC interface supports *Session Initiation Protocol* (SIP), which is a signaling protocol widely used for controlling multimedia communication sessions. In IBM WsT, ISC interface is supported in IMS Connector (Figure 1.2 (2)), which is in charge of processing the SIP request and responses between the IBM application server and the CSCF in the core network.

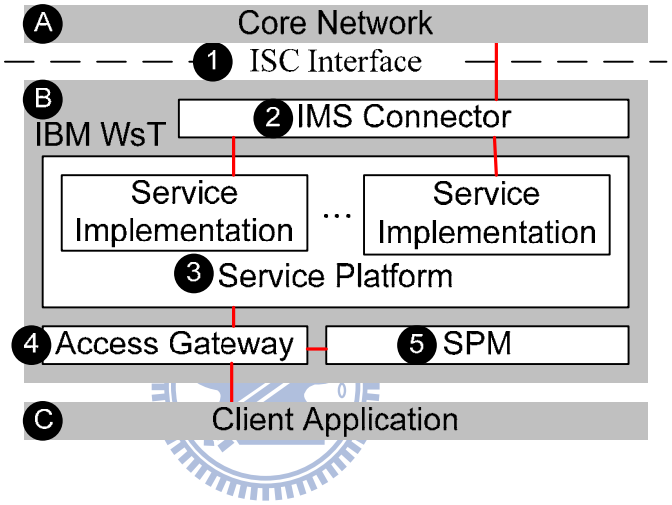


Figure 1.2: IBM WsT Architecture

IBM WsT provides a service platform (Figure 1.2 (3)) that allows third parties to implement their applications using Parlay X *Application Programming Interfaces* (APIs), which are web service APIs for developing telecommunication services. The service platform interacts with the client application (Figure 1.2 (C)) through the access gateway (Figure 1.2 (4)), which is in charge of service access control and user authorization based on the policies defined in the *Service Policy Manager* (SPM; Figure 1.2 (5)).

1.4 Dissertation Organization

Based on the mobile/ wireless network model described in this chapter, we investigate

the VoIP performance in the mobile/wireless network environment. In Chapter 2, we study the authentication delay for real-time applications such as VoIP. Before VoIP call setup, the MS (i.e., a call party) is required to attach to the mobile network and be authenticated. In our study, the authentication key in UMTS can be reused in the IEEE 802.1X authentication mechanism in WLAN environment to achieve telecom grade security. In this WLAN-3G integrated security approach, we conduct a modeling study to tune the IEEE 802.1X parameters to yield better authentication delay performance.

In Chapter 3, we investigate the VoIP packet delivery efficiency in the wireless environment (i.e., WLAN network). In 3GPP specifications, the authenticated WLAN MS is allowed to access the 3G network through the WAG/PDG. However, to ensure telecom grade security, the VoIP traffic between the MS and the WAG/PDG must be protected with IPsec. We analyze the performance of IPsec-based VoIP service in a IEEE 802.11b WLAN environment. The IPsec overheads in terms of throughput, packet loss rate, latency, and jitter are investigated.

In Chapter 4, we present the VoIP performance in the vehicle environment. We conduct trials in the real WiMAX network which supports high-speed mobile broadband services and investigate the WiMAX-based VoIP of a *Mobile Taiwan* (M-Taiwan) funded program conducted during 2007-08 in the Taipei area. For the worst-case-scenario, the tests were conducted under a stringent condition of both communicating devices, wirelessly connected to the same WiMAX base station under a heavy background traffic and interference, were experiencing simultaneous handovers during the communication. The results show excellent performance for the VoIP applications when both *Customer Premise Equipments* (CPEs) are stationary. An acceptable VoIP performance is observed when the CPEs move at the speeds up to 50 Km/h.

Finally, we conclude this dissertation in Chapter 5 by discussing our contributions and future work.

Chapter 2
Effects of the EAPOL Timers in IEEE 802.1X Authentication

This chapter studies IEEE 802.1X authentication for WLAN and cellular integration. In the IEEE 802.1X standard, several timeout timers are defined for message exchanges in the EAPOL protocol, where the same fixed value is suggested for these timeout timers. We observe that the delays for the EAPOL message exchanges may significantly vary. A modeling study is performed to tune the values of individual timers to yield better performance than that for the identical timeout period setting. Our study provides guidelines to select appropriate timeout values for IEEE 802.1X operation.

2.1 Introduction

The IEEE 802.1X standard specifies authentication and authorization for IEEE 802 LAN [7], which has also been widely adopted for mobile devices to access WLAN. Furthermore, if WLAN is integrated with cellular network (such as GSM or UMTS [8]), the SIM module (in the mobile device) and the *Authentication Center* (AuC) are utilized together with IEEE 802.1X for authentication. An example of WLAN and cellular integration (in terms of authentication) is illustrated in Figure 2.1.

In this figure, the HLR is a mobility database that stores and manages all mobile subscriptions of a specific operator. The AuC provides security data management for mobile subscribers. The AuC is typically collocated with the HLR. The AP provides radio access to a mobile device. Before a mobile device is authenticated, the AP only allows this mobile device to send

the IEEE 802.1X authentication messages. When the *Remote Authentication Dial In User Service* (RADIUS) server receives an authentication request from the mobile device, it retrieves the authentication information of the mobile device from the HLR/AuC. After the HLR/AuC returns the authentication information, the RADIUS server authenticates the mobile device following the standard GSM/UMTS authentication procedure [9].

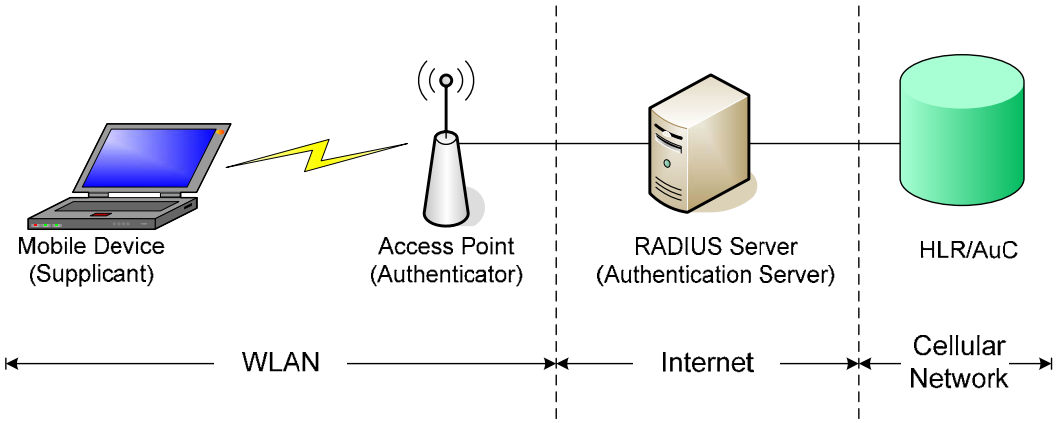


Figure 2.1: A WLAN and Cellular Integration Environment

Figure 2.2 illustrates the protocol stack for the WLAN and cellular integration system. In this figure, the mobile device to be authenticated is called a *supplicant*. The server (typically a RADIUS server) performing authentication is called the *authentication server*. The *authenticator* (e.g., a wireless access point) facilitates authentication between the IEEE 802.1X supplicant and the authentication server.

The integrated system utilizes the *Extensible Authentication Protocol* (EAP) to support multiple authentication mechanisms based on the challenge-response paradigm [10]. The IEEE 802.1X supplicant encapsulates the EAP packets in *EAP over LAN* (EAPOL) frames before they are transmitted to the authenticator. Upon receipt of an EAPOL frame, the authenticator decapsulates the EAP packet from the EAPOL frame. Then the EAP packet is sent to the authentication server using the RADIUS protocol [11]. Implemented on top of UDP, RADIUS provides mechanisms for per-packet authenticity and integrity verification

between the authenticator and the authentication server.

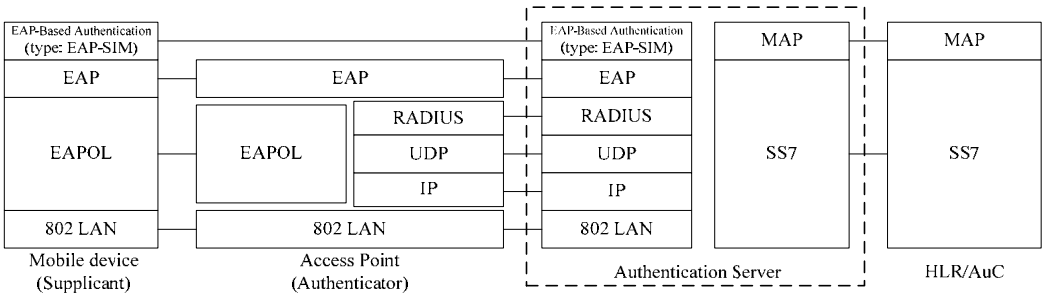


Figure 2.2: The Protocol Stack for WLAN and Cellular Integration

IEEE 802.1X authentication for the WLAN and cellular integration network has been investigated in [12], [13] and [14]. These studies focused on the design of the network integration architectures, and proposed IEEE 802.1X authentication procedures for the integration network. In [15], we proposed an integration solution for *Third Generation* (3G) and WALN services, called the *WLAN-based GPRS Support Node* (WGSN). WGSN re-uses 3G mechanisms for WLAN user authentication and network access without introducing new procedure and without modifying the existing 3G network components. In WGSN, the mobile device must obtain an IP address before it is authenticated by the HLR/AuC. This chapter describes IEEE 802.1X authentication that enhances the WGSN security by allowing a mobile device to be authenticated before it is assigned an IP address.

In our solution, the WLAN and cellular integration network in Figure 2.1 employs EAP-SIM authentication, which is an EAP-based authentication protocol utilizing the GSM *Subscriber Identity Module* (SIM) [16]. In GSM, a secret key *Ki* is stored in the HLR/AuC as well as in the SIM. The authentication server communicates with the HLR/AuC to obtain the GSM authentication information through the *Mobile Application Part* (MAP) implemented on top of the *Signaling System Number 7* (SS7) protocol [8]. In the EAP-SIM authentication, the MAP is responsible for retrieving the GSM authentication information in the HLR/AuC.

In the implementation of IEEE 802.1X authentication for WGSN, we observe that the elapsed times for authentication message pairs exchanged between the mobile device and the network are different. In IEEE 802.1X specification, the message pairs are associated with fixed timeout timers. We analyze the timeout timers used in IEEE 802.1X authentication and improve the performance of IEEE 802.1X authentication by selecting appropriate timer values.

2.2 SIM-based IEEE 802.1X Authentication

This section describes the SIM-based IEEE 802.1X authentication procedure. The authentication message flow is illustrated in Figure 2.3, which consists of the following steps:

Stap 1. The mobile device (the supplicant) sends the EAPOL-Start packet to the AP to initiate the IEEE 802.1X authentication.

Stap 2. The AP requests the identity of the mobile device through the EAP-Request message with type Identity. When the AP receives the EAP-Response/Identity message from the mobile device, it encapsulates this message in the Access-Request packet. Then the packet is sent to the RADIUS server.

Stap 3. Upon receipt of the Access-Request packet, the RADIUS server conducts the EAP-SIM authentication with the mobile device (the supplicant). Specifically, the RADIUS server generates an Access-Challenge packet that encapsulates the EAP-Request/SIM/Start in the EAP-Message attribute and sends it to the AP. This message requests the mobile device to initiate the EAP-SIM authentication. The AP decapsulates the EAP-Request from the Access-Challenge packet, and delivers it to the mobile device by an EAPOL packet.

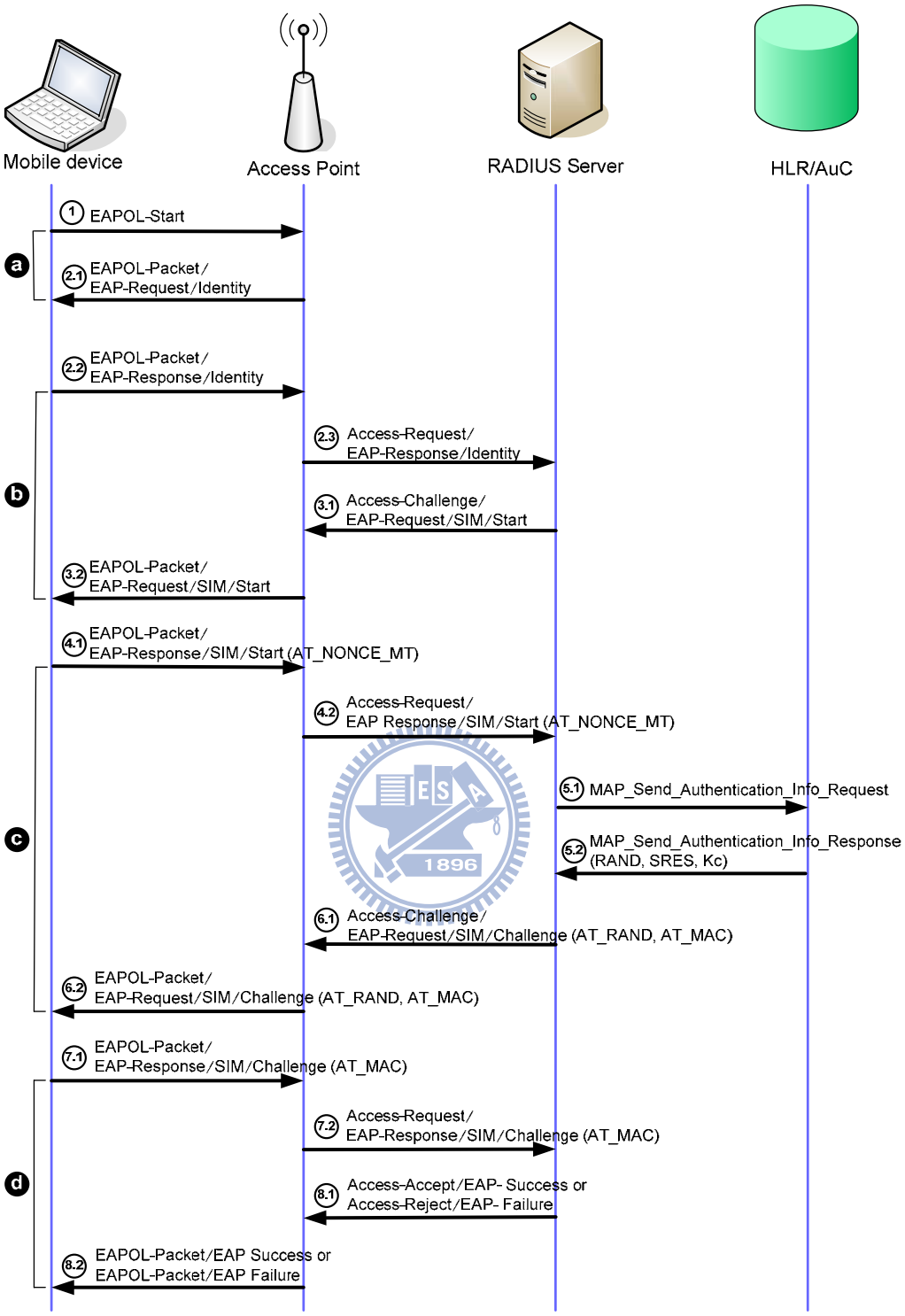


Figure 2.3: SIM-based IEEE 802.1X Authentication Message Flow

Stap 4. The mobile device responds the RADIUS server with the EAP-Response/SIM/Start message containing a random nonce AT_NONCE_MT. The random nonce will be used to generate the encryption key for data transmission between the mobile device and

the RADIUS server after the IEEE 802.1X authentication.

Stap 5. To obtain the authentication information of the mobile device, the RADIUS server sends the SS7 MAP_Send_Authentication_Info_Request message (with the argument IMSI) to the HLR/AuC.

The HLR returns the authentication vector (*RAND*, *SRES*, *Kc*) through the SS7 MAP_Send_Authentication_Info_Response message, where *RAND* is a random number generated by the HLR/AuC. By exercising the GSM authentication algorithm *A3*, both the SIM module and the HLR/AuC use the *RAND* and the secret key *Ki* to produce a *signed result SRES* (see Step 7). The RADIUS server will authenticate the mobile device by comparing the signed result *SRES** generated by the SIM module with the *SRES* generated by the HLR/AuC (see Step 8). If they are equal, the mobile device is successfully authenticated and an encryption key *Kc** is produced by the GSM encryption/decryption key generation algorithm *A8* (using *Ki* and *RAND* as inputs) [8].

Stap 6. The RADIUS server sends the EAP-Request/SIM/Challenge (with the parameter *RAND*, which is encapsulated as AT_RAND) to the mobile device. To ensure the integrity of the challenge message, the message contains a *Message Authentication Code* (MAC) AT_MAC. Detailed usage of MAC can be found in [16], and will not be elaborated here.

Stap 7. After verifying AT_MAC received from the RADIUS server, the mobile device passes the random number *RAND* to the SIM module to perform GSM authentication. The SIM module computes its signed result *SRES** and the encryption key *Kc** based on the received *RAND* and the authentication key *Ki* stored in the SIM card. Then the mobile device sends *SRES** (encapsulated in AT_MAC) to the RADIUS server through

the EAP-Response/SIM/Challenge message.

Stap 8. The RADIUS server verifies AT_MAC and compares *SRES** calculated by the mobile device with the *SRES* received from the HLR/AuC. If the values are identical, the RADIUS server notifies the AP that authentication is successful through the EAP-Success message (encapsulated in the RADIUS Access-Accept packet). The AP passes the EAP-Success message to the mobile device. At this point, the mobile device is allowed to access the network through the AP. If the signed results are not the same, the RADIUS server notifies the AP that the authentication fails through the EAP-Failure message (encapsulated in the RADIUS Access-Reject packet).

2.3 EAPOL Timers

In the IEEE 802.1X supplicant (mobile device), three EAPOL timers are defined:

1. *startWhen* (associated with message pair ⓐ in Figure 2.3): When the IEEE 802.1X supplicant initiates the authentication, it sends EAPOL-Start to the authenticator and starts the *startWhen* timer. If the supplicant has not received any response from the authenticator after this timer expires, it resends EAPOL-Start. The supplicant gives up when it sends EAPOL-Start for n_1 times. In the IEEE 802.1X specification [7], the default n_1 value is 3. The default value of the *startWhen* timer is 30 seconds.

2. *authWhile* (associated with message pairs ⓑ, ⓒ, and ⓓ in Figure 2.3): Every time the supplicant sends an authentication message (Steps 2.2, 4.1, and 7.1 in Figure 2.3), it starts the *authWhile* timer. If the supplicant does not receive any response from the authenticator after this timer has expired, the supplicant sends an EAPOL-Start message to re-start the authentication procedure. The supplicant gives up after it has consecutively sent

EAPOL-Start for n_2 times. The default n_2 value is 3. The default value of the *authWhile* timer is 30 seconds.

3. *heldWhile* (associated with Step 8.2 message in Figure 2.3 if the client fails the authentication): If the IEEE 802.1X authentication fails, the supplicant has to wait for a period *heldWhile* before it re-starts the authentication procedure. The default value of the *heldWhile* timer is 60 seconds.

Selection of the EAPOL timer values is not trivial. If the timer value is too large, it will take long time before the mobile device detects the failure of the network (e.g., RADIUS server failure). If the timer value is too small, the timer may expire before the mobile device receives the response message. In this case, the mobile device needs to re-start the authentication process due to *false failure detection*.

Table 2.1 shows the *expected Round-Trip Times* (RTTs) of message exchanges that measured from the WGSN implemented in National Chiao Tung University. These measurements do not experience waiting delays due to queuing at the network nodes (i.e., AP, RADIUS server and HLR/AuC).

Table 2.1: Expected Round-Trip Times for EAP-SIM Authentication Messages (Without Queuing Delays)

Events occurring at the mobile device	Associated timer	RTT (sec.) (no queueing)
ⓐ in Figure 2.3	*startWhen*	0.005
ⓑ in Figure 2.3	*authWhile*	0.013
ⓒ in Figure 2.3	*authWhile*	1.087
ⓓ in Figure 2.3	*authWhile*	0.013

In our measurement, the mobile device and the AP are located in one subnet. The RADIUS server and the HLR are located in another subnet. The data rate of the fixed network is

100Mbps. It is observed that the RTT of a message exchange between the mobile device and the RADIUS server are much shorter than that of a message exchange between the mobile device and the HLR/AuC. This significant RTT discrepancy is due to the fact that accessing the HLR/AuC is much more time-consuming than accessing the RADIUS server. This phenomenon is especially true when the HLR/AuC is fully loaded by cellular user accesses and when the RADIUS server and the HLR/AuC are located at different cities or different countries. To reduce the false failure detection probability without non-necessary timer timeout delay, the values of the *startWhen* timer and *authWhile* timers should not be identical for all message exchanges in the IEEE 802.1X authentication. For example, the *authWhile* timer for ⓒ in Table 2.1 should be different from that for ⓑ and ⓓ.

2.4 Performance Modeling

We propose an analytic model to investigate the *false failure detection probability* p_f of the IEEE 802.1X authentication procedure and the expected elapsed (response) time E[τ] for executing the IEEE 802.1X authentication procedure. Input parameters and output measures used in the model are listed in Table 2.2.

In Table 2.2, t_X is the RTT of the message exchange without waiting delay (i.e., the queueing at a network node) where X = s, a_1, a_2, or a_3. This RTT is called "service time" in the queueing model. For our measurement in Table 2.1, E[t_S] =0.005 seconds, E[t_{a1}] = 0.013 seconds, E[t_{a2}] = 1.087 seconds, and E[t_{a3}] = 0.013 seconds. The response time τ_X of the message exchange is the RTT of the message exchange including the queuing delay. Since we cannot conduct large-scale service trial in our IEEE 802.1X prototype, τ_X are derived from the service times using the M/G/1 queuing model. Let EAPOL message arrivals to the AP be a Poisson stream with rate λ. The service time t_X of the message exchange has an arbitrary

distribution with the distribution function $B_X(\cdot)$, the density function $b_X(\cdot)$, and the Laplace Transform $B_X^*(\cdot)$. The response time of the message exchange is represented by the random variable τ_X, which has the distribution function $F_X(\cdot)$, the density function $f_X(\cdot)$ and the Laplace Transform $F_X^*(\cdot)$.

Table 2.2: Input Parameters and Output Measures

Input Parameters				
message pair	associated timeout period	service time of message exchange	response time of message exchange	timeout probability
ⓐ	T_S	T_S	τ_S	$p_S = \Pr[\tau_S \geq T_S]$
ⓑ	T_{a1}	t_{a1}	τ_{a1}	$p_{a1} = \Pr[\tau_{a1} \geq T_{a1}]$
ⓒ	T_{a2}	t_{a2}	τ_{a2}	$p_{a2} = \Pr[\tau_{a2} \geq T_{a2}]$
ⓓ	T_{a3}	t_{a3}	τ_{a3}	$p_{a3} = \Pr[\tau_{a3} \geq T_{a3}]$
Output Measures				
p_f	the false failure detection probability of the IEEE 802.1X authentication procedure; p_f = Pr[the mobile device has consecutively sent the EAPOL-Start frame for three times]			
$E[\tau]$	the expected response time of the IEEE 802.1X authentication procedure			

By applying the Pollaczek-Khintchine transform equation for the sojourn time, we can express $F_X^*(\cdot)$ as [18]:

$$F_X^*(s) = B_X^*(s) \sum_{k=0}^{\infty} (1-1\,E[t_X])\,(1\,E[t_X])^k \left(\frac{1-B_X^*(s)}{s\,E[t_X]}\right)^k, \qquad (2.1)$$

where $X = s$, a_1, a_2, or a_3. The density function $f_X(\cdot)$ can be obtained by inverting the Laplace Transform $F_X^*(\cdot)$ in (1). Therefore, we have

$$f_X(t_X) = b_X(t_X) + \sum_{k=0}^{\infty} (1-1\,E[t_X])\,(1\,E[t_X])^k\, \hat{b}_{(k)}(t_X), \qquad (2.2)$$

where $\hat{b}_{(k)}(t_X)$ is the k-fold convolution of the function $\frac{1-B_X(t_X)}{E[t_X]}$

Let T_X be the timeout period associated with the timer for the message pair X and p_X be the timeout probability of the message exchange.

$$p_X = \Pr[\ \tau_X \geq T_X\ ] = \int_{T_X}^{\infty} f_X(t)\,dt \tag{2.3}$$

The expected response time $E[\tau_X]$ of the message exchange can be obtained by differentiating the Laplace Transform $F_X^*(\cdot)$.

$$E[\tau_X] = p_X T_X + (1 - p_X) \int_0^{T_X} t f_X(t)\,dt \tag{2.4}$$

The probability transition diagram of the mobile device is illustrated in Figure 2.4. In IEEE 802.1X, the AP can also control the number of retransmissions for EAPOL-Start (① in Figure 2.3) sent from the mobile device to the AP. To simplify our discussion, we assume that the number of retransmissions is sufficiently large, so that the state diagram in Figure 2.4 is not affected.

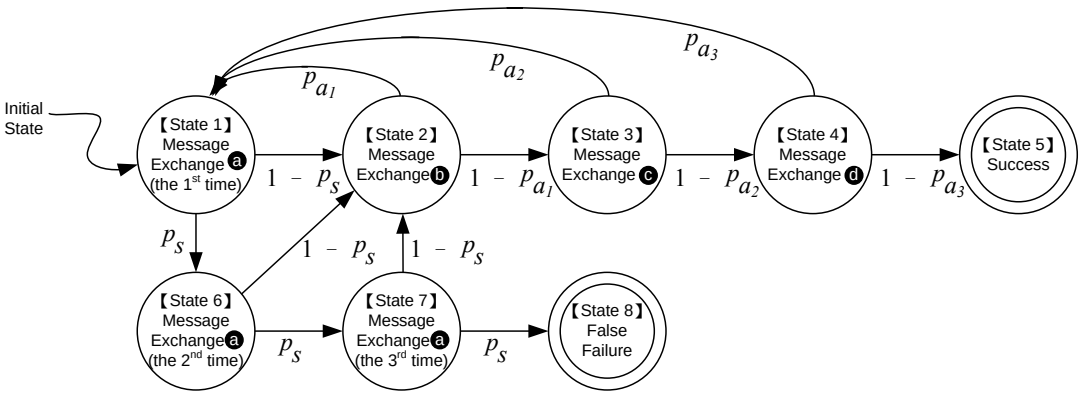


Figure 2.4: The Probability Transition Diagram of the IEEE 802.1X Authentication Message Exchange

During IEEE 802.1X authentication, the mobile device restarts the procedure (i.e., come back

to state ① again) whenever the *authWhile* timer (associated with message exchanges ❻, ❼, and ❽) expires. The authentication exits and is considered failed if the *startWhen* timer (associated with message exchange ❺) has consecutively expired for three times (i.e., the *finite state machine* (FSM) moves from state ①, ⑥, ⑦, to state ⑧).

Let x be the probability that the FSM starts from state ① and eventually comes back to state ① (i.e., state ① may be revisited zero or more times). All possible scenarios for the probability transitions in Figure 2.4 are described as follows:

Scenario I: From state ① (i.e., state ① may have been visited zero or more times), the *startWhen* timer consecutively expires for three times (i.e., the last transitions are ①→⑥→⑦→⑧). The probability for Scenario I is xp_S^3.

Scenario II: From state ①, the *startWhen* timer consecutively expires for two times, and the procedure successes at the third try (i.e., the last transitions are ①→⑥→⑦→②→③→④→⑤). The probability for Scenario II is $x\,p_S^2\,(1-p_S)(1-p_{a1})(1-p_{a2})(1-p_{a3})$.

Scenario III: From state ①, the *startWhen* timer expires once, and the procedure successes at the second try (i.e., the last transitions are ①→⑥→②→③→④→⑤). The probability for Scenario III is $xp_S(1-p_S)(1-p_{a1})(1-p_{a2})(1-p_{a3})$.

Scenario IV: From state ⊕, the procedure successes without incurring any timer expiration (i.e., the last transitions are ①→②→③→④→⑤). The probability for Scenario IV is $x(1-p_S)(1-p_{a1})(1-p_{a2})(1-p_{a3})$.

It is apparent that the false failure probability p_f is the probability that Scenario I occurs. The success probability $(1-p_f)$ is the probability that either Scenarios II, III, or IV occur. That is,

$$p_f = \Pr[\text{Scenario I occurs}] = x p_S^{\,3} \tag{2.5}$$

and

$$\begin{aligned} 1 - p_f &= \Pr[\text{Scenarios II, III, or IV occur}] \\ &= x(p_S^{\,2} + p_S^{\,1} + 1)(1 - p_S)(1 - p_{a1})(1 - p_{a2})(1 - p_{a3}) \\ &= x(1 - p_S^{\,3})(1 - p_{a1})(1 - p_{a2})(1 - p_{a3}) \end{aligned} \tag{2.6}$$

From (2.5) and (2.6), we have

$$p_f + (1 - p_f) = x\,[p_S^{\,3} + (1 - p_S^{\,3})(1 - p_{a1})(1 - p_{a2})(1 - p_{a3})] \tag{2.7}$$

By rearranging (2.7), we have

$$x = \frac{1}{p_S^{\,3} + (1 - p_S^{\,3})(1 - p_{a1})(1 - p_{a2})(1 - p_{a3})} \tag{2.8}$$

From (2.8) and (2.5),

$$p_f = \frac{p_S^{\,3}}{p_S^{\,3} + (1 - p_S^{\,3})(1 - p_{a1})(1 - p_{a2})(1 - p_{a3})} \tag{2.9}$$

By using (2.3) and (2.9), the value of p_f can be computed from $\lambda, f_S, f_{a1}, f_{a2}$, and f_{a3}.

Table 2.3: The p_X Values: Analysis Versus Simulation

($T_X = 10 \times E[t_X]$, $var[t_X] = E[t_X]^2$, and $X = s$, a_1, a_2, or a_3).

(UNIT: $\frac{1}{E[T_X]}$)	0.2	0.4	0.6	0.8
SIMULATION	0.0003	0.0025	0.0183	0.1353
ANALYTIC	0.0004	0.0027	0.0196	0.1271
ERROR	0.0001	0.0002	0.0013	0.0082

The above analytic model is validated against simulation experiments. The simulation model follows the discrete event approach [19], and the details are omitted. Table 2.3 indicates that the analytic and the simulation results are consistent (the errors are within 1%). Therefore, both the analytic model and the simulation implementation are validated.

2.5 Numerical Examples

This section uses numerical examples to investigate the impact of timers on the performance of IEEE 802.1X authentication. The input parameters and the output measures are listed in Table 2.2. The expected service times of the EAPOL message exchanges $E[t_S]$, $E[t_{a1}]$, $E[t_{a2}]$, and $E[t_{a3}]$ are obtained from the measurements as listed in Table 2.1. Other parameters include the EAPOL message arrival rate λ, and the variances of the EAPOL service times (i.e., $var[t_X]$, where $X = s$, a_1, a_2, or a_3). We have the following observations.

Table 2.4: Effects of T_X on p_f ($var[t_X] = 100 \times E[t_X]^2$; $X = s$, a_1, a_2, or a_3).

Timeout Timers (Unit: second)				Arrival Rate . $\left(\text{Unit: } \frac{1}{E[t_{a2}]}\right)$		
T_S	T_{a1}	T_{a2}	T_{a3}	0.900	0.950	0.975
5	10	10	10	0.0199	0.3473	0.9569
15	10	10	10	0.0000	0.0023	0.5000
10	**5**	10	10	0.0000	0.0574	0.8676
10	**15**	10	10	0.0000	0.0314	0.7510
10	10	**5**	10	0.0000	0.0416	0.8140
10	10	**15**	10	0.0000	0.0344	0.7801
10	10	10	**5**	0.0000	0.0582	0.8679
10	10	10	**15**	0.0000	0.0317	0.7527

Observation 1. The probability p_f is mainly affected by p_S.

From the transitions of the FSM in Figure 2.4, it is clear that if p_{a1}, p_{a2}, $p_{a3} > 0$, we have

$\lim_{p_s \to 1} p_f = 1$ and $\lim_{p_s \to 0} p_f = 0$.

Observation 2. The probabilities p_{a1}, p_{a2}, and p_{a3} have indirect impact on p_f.

Increasing p_{a1}, p_{a2}, and p_{a3} will increase the probability that the FSM moves toward state ① (i.e., loops among states ①, ②, ③, and ④) and thus decrease the probability that the FSM enters state ⑤. However, the effect in Observation 2 is not as significant as that in Observation 1.

Based on simulation experiments, Table 2.4 shows how timeout period T_X ($X = s$, a_1, a_2, or a_3) affects p_f when the variance of t_X is large (i.e., $var[t_X] = 100 \times E[t_X]^2$). These results in Table 2.4 are consistent with Observations 1 and 2. That is, T_S has more significant effect on p_f than T_{a1}, T_{a2}, and T_{a3} do, especially when the EAPOL message arrival rate λ is high. Specifically, when $\lambda = 0.975 \times E[t_{a2}]^{-1}$ and if T_{a1}, T_{a2}, and T_{a3} are fixed to 10 seconds, changing the value of T_S from 5 seconds to 15 seconds will reduce p_f from 95.69% to 50.00% (about 50% improvement). Conversely, changing any of the values for T_{a1}, T_{a2}, and T_{a3} from 5 to 15 seconds only insignificantly affects p_f (less than 13% improvement).

When $var[t_X]$ is small (e.g., $var[t_X]$ is less than $E[t_X]^2$), the service time t_X does not significantly vary, and the IEEE 802.1X authentication message exchange is more likely to complete if the value for T_X is set larger than the expected service time $E[t_X]$ of the corresponding EAPOL message pair exchange. Therefore, we have the following observation.

Observation 3. When $var[t_X]$ is small and T_X is sufficiently large, changing T_X only insignificantly affects p_f.

Table 2.5 shows how T_X and $var[t_X]$ affect $E[\tau]$ when the EAPOL message arrival rate λ is set

to 0.5 × E[t_{a2}]$^{-1}$. From this table, we have the following observations.

Observation 4. When *var*[t_X] is small (i.e., *var*[t_X] = E[t_X]^2), E[τ] is insignificantly affected by the change of T_X if T_X is larger than the expected response time of the EAPOL message exchanges. This result is similar to that in Observation 3.

Table 2.5: Effects of T_X and *var*[t_X] on E[τ] ($X = s$, a_1, a_2, or a_3; λ = 0.5 × E[t_{a2}]$^{-1}$).

Timeout Timers (Unit: second)				*var*[t_X] = E[t_X]^2		*var*[t_X] = 100 × E[t_X]^2	
T_s	T_{a1}	T_{a2}	T_{a3}	E[τ]	effects	E[τ]	effects
5	10	10	10	2.24	decreasing 0.45%	11.00	decreasing 0.09%
15	10	10	10	2.23		10.99	
10	**5**	10	10	2.24	decreasing 0.45%	10.96	decreasing 0.36%
10	**15**	10	10	2.23		11.00	
10	10	**5**	10	2.25	decreasing 0.89%	7.26	increasing 96.01%
10	10	**15**	10	2.23		14.23	
10	10	10	**5**	2.24	decreasing 0.45%	11.17	decreasing 1.70%
10	10	10	**15**	2.23		10.98	

We also observe two effects on T_X.

Effect 1. When T_X is increased, the probability p_X is decreased, and the number of looping among states ①, ②, ③, and ④ is reduced.

When the variance of service time *var*[t_X] is increased, more large t_X periods are expected. Therefore, even if T_X is increased, these long EAPOL message delays still result in timeouts, and p_f cannot be significantly reduced. Therefore, we have

Observation 5. Effect 1 is more significant when *var*[t_X] is small than when *var*[t_X] is large.

Effect 2. When T_X is increased, if timeout does occur, the non-necessarily waiting for timeout is increased and $E[\tau]$ is increased.

Based on the above discussion, we examine the results in Table 2.5 as follows.

Observation 6. If $var[t_X]$ is large and T_X is not much larger than $E[t_X]$ (e.g., 5 seconds < T_X < 10 seconds), then increases T_X only insignificantly decreases p_f(indicated in Observation 5), but significantly increases extra waiting period for timeout described in Effect 2. Therefore $E[\tau]$ is significantly increased. This phenomenon occurs for changing T_{a2} in Table 2.5, where $E[T_{a2}] = 1.087$ seconds.

When $E[t_X] << T_X$, then even if $var[t_X]$ is large, it is still likely that $t_X < T_X$, and both Effects 1 and 2 are not significant. Instead, the situation is similar to that in Observation 4. That is, $E[\tau]$ is insignificantly affected by $var[t_X]$ and T_X when $E[t_X] << T_X$. The phenomena occur for changing T_S, T_{a1}, and T_{a2} in Table 2.5, where these timeout periods T_X are larger than 500 × $E[t_X]$.

A comparison of different timers settings is shown in Figure 2.5. In Cases 1, 2, and 3, the timers in different message pairs are set to identical values, which means that a fixed *authWile* timer is used in message pairs ⓑ, ⓒ, and ⓓ (as suggested in IEEE 802.1X standard). In Case 4, we adjust the values of the timers according to the previous discussions to obtain better performance. For the illustration purpose, It is assumed that the variance of service time $var[t_X]$ is 100 × $E[t_X]^2$. Similar results are observed for different variances and are not presented in this study. Figure 2.5 indicates that the false failure detection probability p_f is zero if the timeout period T_S is larger than 10 seconds and the EAPOL message arrival rate λ, is below 0.925 × $E[t_{a2}]^{-1}$. Theses figure indicate that when λ < 0.925 × $E[t_{a2}]^{-1}$, Case 4 has the same p_f performance as Cases 1 – 3, but has much better $E[\tau]$ performance than these

three Cases. When $\lambda > 0.925 \times E[t_{a2}]^{-1}$, Case 4 improves $E[\tau]$ at the cost of degrading p_f as compared with Cases 1 and 2. For all λ values, Case 1 outperforms Case 2, and Case 2 outperforms Case 3 in terms of the p_f measure. For the $E[\tau]$ performance, the result reverses. In Case 3, the total timeout value $T_S + T_{a1} + T_{a2} + T_{a3} = 40$ seconds. For Case 4, the total timeout value is 55 seconds. It is interesting to note that for all values, Case 4 outperforms Case 3 for both p_f and $E[\tau]$ performances. Also note that when $\lambda > 0.925 \times E[t_{a2}]^{-1}$, the system is saturated, and will not occur in most commercial operations.

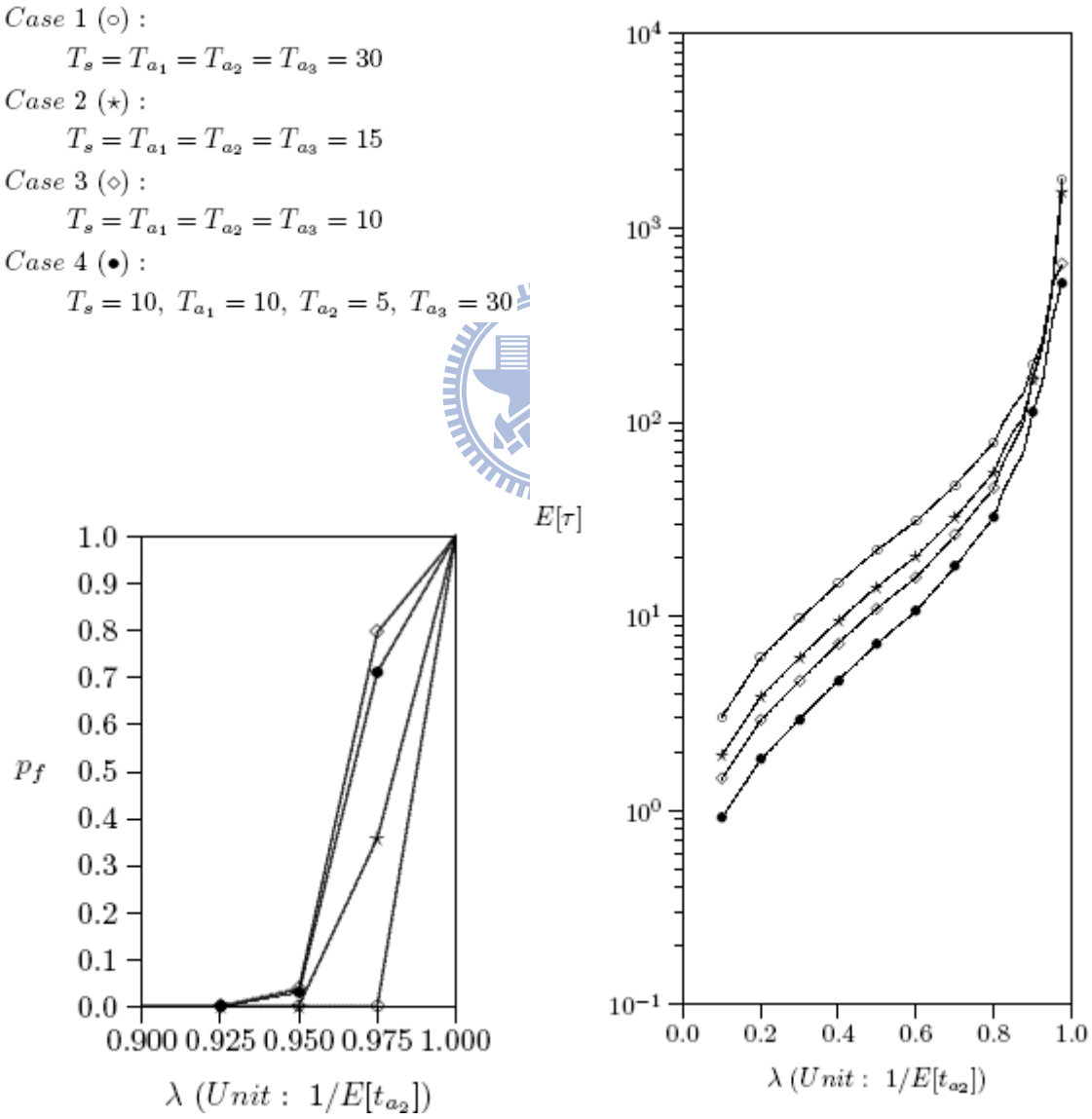


Figure 2.5: Effects of different timeout period settings when $var[t_X] = 100 \times E[t_X]^2$ ($X = s$, a_1, a_2, or a_3).

In these numerical examples, we demonstrate that appropriate T_X values can be selected through our modeling study to yield better performance than the fixed T_X value setting.

2.6 Summary

This chapter described IEEE 802.1X authentication for WLAN and Cellular integration. We presented the protocol stack and the authentication message flow, and measured the response times of all EAPOL message exchanges in the IEEE 802.1X authentication for the integrated system implemented in NCTU.

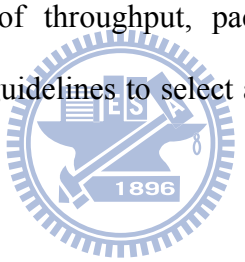
In the IEEE 802.1X standard, a fixed-value timer is used in all authentication message exchanges, which does not reflect the real network operation. A modeling study was presented in this chapter to tune the values of individual timers, which yields better performance than the fixed timeout period setting.

Our study provides guidelines to select appropriate timeout periods for corresponding authentication message exchanges. For example, comparing with the fixed timeout periods setting where T_X are set to 10 seconds, the suggested setting for the timeout periods (i.e., T_S = 10 seconds, T_{a1} = 10 seconds, T_{a2} = 5 seconds, and T_{a3} = 30 seconds) decreases the false failure detection probability p_f and significantly improves the expected response time $E[\tau]$ of the IEEE 802.1X authentication procedure.

Chapter 3

IPsec-Based VoIP Performance in WLAN Environments

3GPP specifies 3G-WLAN interworking that allows a *Mobile Station* (MS) to access the 3G network through a *WLAN Access Gateway/Packet Data Gateway* (WAG/PDG), and specifies that the packets delivered between the MS and the WAG/PDG should be protected with IPsec. This chapter studies the performance of IPsec-based VoIP service in a WLAN environment. The IPsec overheads in terms of throughput, packet loss rate, latency, and jitter are investigated. Our study provides guidelines to select appropriate system parameter values for VoIP over WLAN.



3.1 Introduction

The 3GPP Technical Specification 23.234 [20] specifies 3G-WLAN interworking that extends 3G services to the WLAN environment. In 3G-WLAN interworking, a *Mobile Station* (MS) in the WLAN accesses the 3G core network through a *WLAN Access Gateway/Packet Data Gateway* (WAG/PDG). The security requirements are enforced such that between an MS and the WAG/PDG, IPsec security association is established and the transmitted packets are protected by IPsec with *Encapsulating Security Payload* (ESP) in the tunnel mode [21]. In an IPsec tunnel, an original IP packet, including the header and the payload, is encrypted and authenticated, and a new IP header is added to route the packet between the MS and the WAG/PDG. Before sending an IP packet, the MS checks the security policies applied to the

packet and performs IPsec encapsulation according to the methods defined by the security association. When receiving an IPsec packet, the WAG/PDG validates and decapsulates the packet according to the security information of the corresponding security association. By exercising IPsec encapsulation, the sizes of the packets transmitted between the MS and the WAG/PDG increase and the performance degrades.

Voice over Internet Protocol (VoIP) provides voice communications over Internet [15], where *Session Initiation Protocol* (SIP) [22] is often used to control the call and *Real-time Transport Protocol* (RTP) [23] is used to deliver the voice data. Several codecs can be used in the RTP calls to meet the bandwidth restrictions. In a 3G-WLAN interowking environment, the VoIP performance may be degraded due to IPsec encapsulation. This chapter investigates the performance of IPsec-based VoIP in the WLAN environment, and is organized as follows. The related works are elaborated in Section 3.2. Our experimental environment is described in Section 3.3, and the performance measures (throughput, packet loss rate, latency, and jitter) are reported in Section 3.4. Based on the performance measurements, Section 3.5 suggests proper system parameter setups for IPsec-based VoIP in the WLAN environment.

3.2 Related Works

The IPsec performance for VoIP in the wireless environments has been investigated in [24, 25, 26, 27]. The latency performance for IPsec security association was investigated in [25]. The study in [24] presented the IPsec performance for VoIP by evaluating speech quality with G.711 codec. The speech quality was evaluated by a subjective method called *Mean Opinion Score* (MOS). The study in [25] measured IPsec-based VoIP over 3G-WLAN interworking system with different types of codecs. The studies in [26, 27] investigated the VoIP performance with mathematical analysis and simulation experiments of network capacity (in

terms of the number of supported RTP streams).

In the previous experimental studies, the performance measurement tools were run on the MS, which may affect the accuracy of the reported results. In [24], the QoS measurement tool developed by the authors was executed on the MS. In [25], four tools were executed on the MS, including RTP Tools to send and receive the RTP packets, network monitor tools pktstat and Netperf to measure the network traffic and collect the performance data, and Ethereal to record network packet events on the MS. In our experiments, all performance data are collected and measured by Smartbit. The MS only processes the VoIP packets as in the normal operation mode, and its computing power is not consumed for measurement.

The performance results presented in the previous studies are quite different due to different experimental setups and the output measures defined. The MOS measure considered in [24] provides useful insight for voice quality. However, it does not reflect the effect of delays. Also the IPsec performance in terms of packet loss, latency, and jitter were not presented. The analytical studies in [26, 27] showed the IEEE 802.11b AP capacity for plain VoIP without packet loss. They did not consider the relationships between the throughput and the VoIP traffic load. The performance results in [25] are not consistent with that in [26, 27]. On the other hand, we conduct our measurements by setting the same experimental parameters to those in [26, 27], and obtain the consistent results

All these previous studies did not consider heavy VoIP traffic issues that are further elaborated in this chapter. Heavy traffic analysis provides useful insight to a VoIP operator to determine what kinds of codec and packet loss concealment techniques should be employed. We also elaborate on the IPsec overhead in terms of latency, and compare the jitter performance for VoIP with and without IPsec. These aspects were not investigated in the previous studies.

3.3 IPsec-based VoIP Experimental Environment

Figure 3.1 illustrates a simplified 3G-WLAN integration system, where the MS (Figure 3.1 (1)) is a laptop equipped with Pentium M 1.3GHz CPU and 256MB memory. The WAG/PDG (Figure 3.1 (2)) is a laboratory prototype implemented in a PC equipped with Pentium IV 3.0 GHz CPU and 1GB memory. The MS communicates with the WAG/PDG via a D-Link DL-524 IEEE 802.11b *access point* (AP; Figure 3.1 (3)). The AP connects to the WAG/PDG through the Ethernet where the peak rate is 10Mbps. The Smartbit (Figure 3.1 (4)) is utilized for performance measurement [28], and is connected to the MS and the WAG/PDG using CAT 5 cables with the RJ-45 interface.

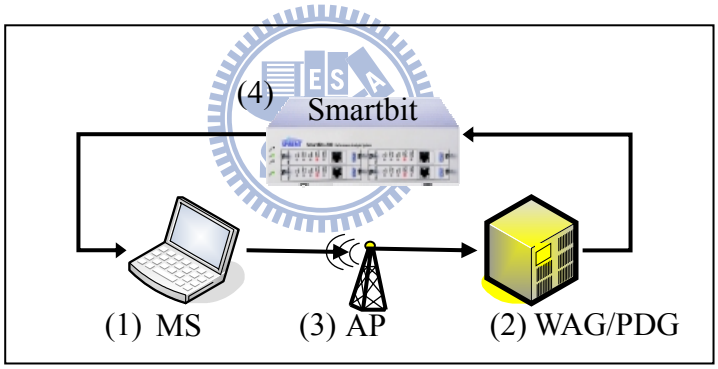


Figure 3.1 The Experimental Environment

Table 3.1: The Codec Attributes

Codec	G.711	G.729
Bit Rate	64 Kbps, sampling at an 8 KHz rate with 8 bits per sample	8 Kbps, sampling at a 1 KHz rate with 8 bits per sample
Sample Period	20 ms	10 ms
RTP Payload Length (Sample Rate x Sample Period)	160 Bytes, one frame per RTP packet	10*2 Bytes, two frames are combined into one RTP packet
RTP Packet Rate	50 packets/sec	50 packets/sec

In the experiments, the Smartbit generates multiple RTP streams identified with different source/destination IP address pairs, and injects them to the MS. The RTP packets are transmitted over UDP. Two kinds of voice codecs, G.711 [29] and G.729 [30], are used in generating the RTP streams. G.711 is a high bit-rate codec with a sample period of 20 ms. G.729 is a low bit-rate codec with a sample period of 10 ms. The codec attributes are summarized in Table 3.1.

When the MS receives the RTP packets from the Smartbit, it encapsulates these packets with IPsec ESP in the tunnel mode. The RTP packets are encrypted by the 3DES algorithm [31], and are authenticated by the HMAC-SHA-1-96 algorithm [32]. Then the encapsulated packets are sent to the WAG/PDG via the IPsec tunnel. Upon receipt of the IPsec packets, the WAG/PDG executes the IPsec decryption procedure. The decrypted RTP packets are then collected by the Smartbit. From the measured packets, the Smartbit produces the output statistics.

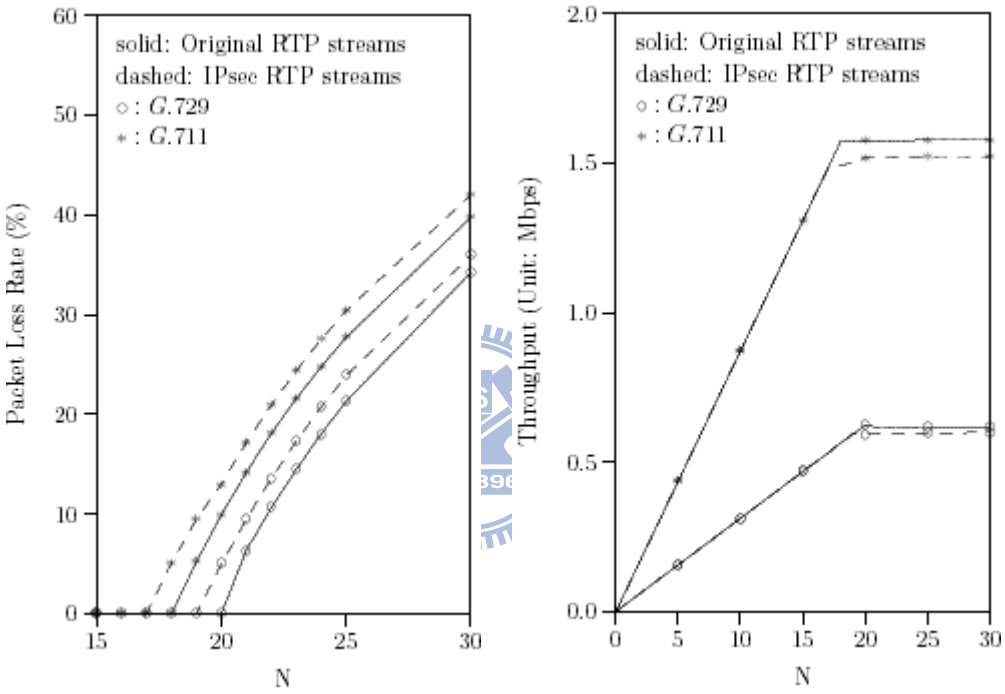
3.4 Performance Measurement

Based on the experimental environment described in Section 3.3, we measure the IPsec overhead in terms of throughput, packet loss rate, latency, and jitter.

3.4.1 Throughput and Packet Loss Rate

Figure 3.2 (a) illustrates the packet loss rate measured by the Smartbit. The study in [26, 27] derived an upper bound for the VoIP capacity (in terms of the number of the supported RTP streams) over IEEE 802.11b without packet loss. By using the equation derived in [26, 27], we compute the upper bound of network capacity for IPsec-based VoIP. Table 3.2 compares

the VoIP capacity without packet loss between the measured values in our experiments (Figure 3.2) with the upper bounds derived in [26, 27]. This table shows that the measured capacity without packet loss achieves about 85% of the calculated upper bound capacity. Table 3.2 also indicates that after IPsec encryption, the capacities without packet loss degrade by 4% (in terms of the number of RTP streams supported) for G.729 and 5% for G.711, respectively.



(a) Packet Loss Rate (b) Throughput

Figure 3.2 Packet Loss and Throughput (N: Number of RTP Streams)

Table 3.2 Measured Capacities without Packet Loss and Their Upper Bounds

Codec	IPsec Encrypted	Upper Bound Capacity	Measured Capacity
G.729	no	24.1 RTP Streams	20 RTP Streams
	yes	23.2 RTP Streams	19 RTP Streams
G.711	no	21.5 RTP Streams	18 RTP Streams
	yes	20.7 RTP Streams	17 RTP Streams

The study in [25] showed that the IEEE 802.11b AP can support 28 IPsec VoIP connections

for G.711 with packet loss rate less than 1%. This result is probably misleading because the reported number of VoIP connections (i.e., 28) exceeds the capacity upper bound derived in [26, 27]. Our experiments and the studies in [26, 27] show that when the packet loss rate is less than 1%, the IEEE 802.11b AP can only support 20 and 17 IPsec VoIP connections for G.711, respectively.

Figure 3.2 (a) indicates that to maintain the same packet loss rate, the system can support one less IPsec RTP streams than original RTP streams. For examples, when the packet loss rate is 10%, the system can support 21.86 original RTP streams or 21.13 IPsec RTP streams for G.729 (i.e., the IPsec overhead is 3.34%), and can support 20.02 original RTP streams or 19.15 IPsec RTP streams for G.711 (i.e., the IPsec overhead is 4.35%).

Figure 3.2 (a) also shows that as the system attempts to support more RTP streams (and therefore the packet loss rate increases), the IPsec overhead becomes less significant. For example, when the packet loss rate increases from 5% to 20%, the IPsec overhead decreases from 3.80% to 3.29% for G.729, and from 5.12% to 3.55% for G.711.

Based on the mathematical analysis in [26, 27], we further derive the capacity of an IEEE 802.11b AP for IPsec VoIP. Figure 3.2 (a) indicates that the packet loss rate increases to 5% – 6.3% when the VoIP traffic is one RTP stream larger than the network capacity without packet loss. This result is consistent with the simulation results in [26, 27].

Figure 3.2 (b) illustrates the throughput performance. We note that the following relationship holds:

$$\text{Packet Loss Rate} = \frac{\text{Arrival Rate} \times \text{Packet Size} - \text{Throughput}}{\text{Arrival Rate} \times \text{Packet Size}}$$

where the arrival rate is 50 packets/sec times the number of RTP streams. This figure indicates

that the system saturates if we pump more than 20 original RTP streams or 19 IPsec RTP streams for G.729 (or 18 original RTP streams or 17 IPsec RTP streams for G.711). By exercising IPsec, the maximum throughput for the system degrades by 5% for G.729, and 5.56% for G.711. When the system is not saturated, the throughput for supporting both original and IPsec RTP streams are the same.

3.4.2 Latency

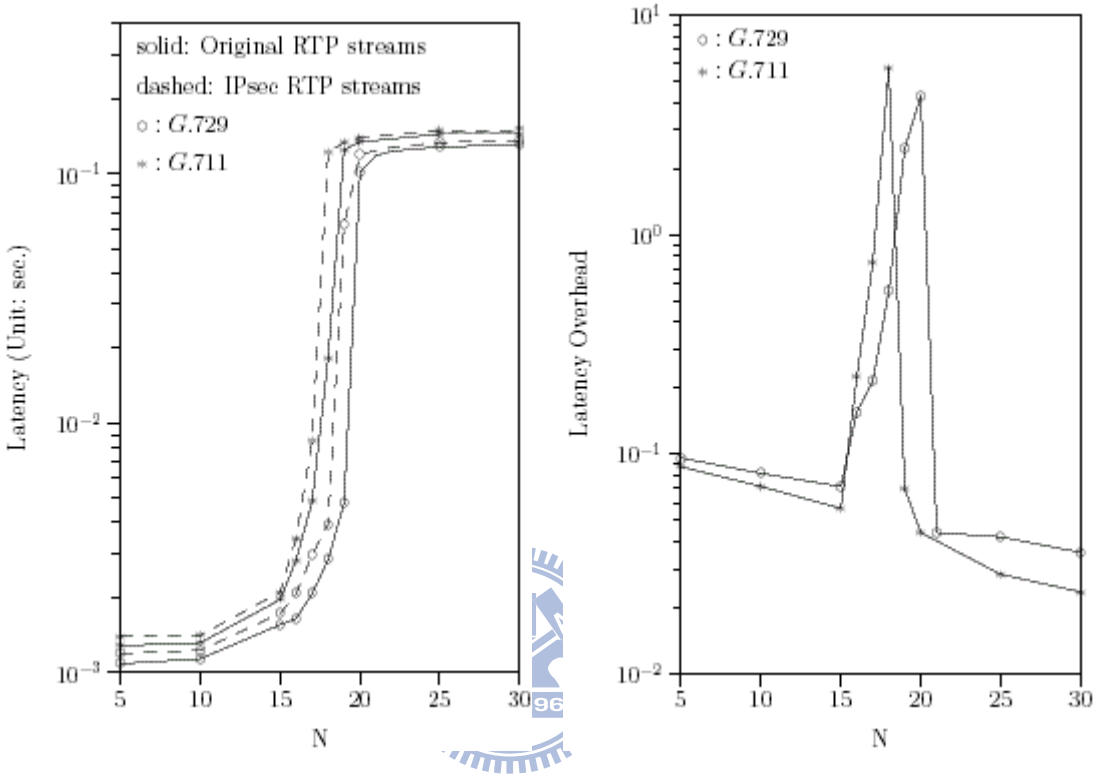
The latency performance is affected by the following factors: (1) packet processing, (2) packet delivery, and (3) packet loss. Both packet processing and packet delivery contribute to queueing, and therefore will increase the latency. During delivery, packets may be retransmitted due to transmission errors or collisions (i.e., congestion of the radio link). In IEEE 802.11b, a packet is transmitted after a backoff delay. For each retransmission, the average backoff delay is doubled. When the number of retransmissions for a packet reaches the retry threshold, it is discarded. Packet loss mitigates queueing effect, and therefore will "stop" increasing of the latency caused by packet retransmission.

Figure 3.3 (a) shows that the mean latency is an "S"-shape increasing function of the number N of RTP streams.

Case I. When $N < 10$, increasing the stream number insignificantly affects the latency. In this case, there is no packet loss, and packet retransmission seldom occurs. Therefore, the latency is caused by the queueing effect of packet processing. For example, when N increases from 5 to 10, the latency increases from 1.17 ms to 1.30 ms (i.e., increases 11.11%) for G.729, and from 1.39 ms to 1.41 ms (i.e., increases 1.44%) for G.711.

Case II. When $10 \leq N \leq 20$, the latency increases significantly as N increases due to both

packet processing and packet retransmission. For example, when N increases form 15 to 20, the latency increases from 1.72 ms to 101.22 ms for G.729, and from 2.10 ms to 139.38 ms for G.711.



(a) The Mean Latency (b) Latency Overhead

Figure 3.3 Latency Performance

Case III. When N > 20, the latency only slightly increases as N increases due to packet loss. As shown in Figure 3.3 (a), packet loss significantly increases as N increases for N ≥ 20. When N increases form 25 to 30, the latency increases from 132.99 ms to 134.95 ms (i.e., increases 1.47%) for G.729, and from 147.82 ms to 149.23 ms (i.e., increases 0.95%) for G.711.

The latency performance reported in [25] showed the same trend as our results for Case I and II. However, Case III was not investigated in [25].

Since the packet size for G.711 is larger than that for G.729, Figure 3.3 (a) shows that the packet processing time for G.711 is longer than that for G.729. Similarly, the latencies for IPsec RTP streams are longer than that for original streams. Specifically, the latency overhead for IPsec can be calculated as

$$\text{Latency Overhead} = \frac{\text{IPsec RTP Latency} - \text{Original RTP Latency}}{\text{Original RTP Latency}} .$$

When N < 15, the IPsec overhead is less than 9.26% due to insignificant queueing effect contributed by packet processing. When 15 ≤ N ≤ 20, the latency overhead for IPsec significantly increases (can be up to 570.97%). In this case, for the same N value, IPsec streams experience heavy packet retransmission. On the other hand, packet retransmission is not significant for original streams. For N > 20, the system saturates for both IPsec and original streams. That is, many packets will reach the retransmission threshold and are dropped in both IPsec and original streams. Therefore, the latency overhead for IPsec drops to less than 4.38%.

3.4.3 Jitters

Jitter or the variation of packet inter-arrival time may create unexpected pauses between utterances, and therefore affects the intelligibility of the VoIP speech. It was reported that an average jitter exceeding 35 ms results in unacceptable QoS for VoIP [28]. To reduce jitter, a buffer is used to store the incoming packets before they are played. If the jitter buffer size is too small, network jitter will result in packet loss and therefore degrade the intelligibility of the voice. If the jitter buffer size is too large, long packet delay will be experienced, which results in QoS degradation (e.g., echo level may be more easily perceptible). NTT Communications specifies that the average jitter should be no more than 0.5 ms for VoIP [33].

Figure 3.4 shows that without jitter buffer, the jitter is an "S"-shape increasing function of the number N of RTP streams in our experiments. When N < 5, the RTP packets experience less link congestions and the average jitters are less than 0.5 ms. When 5 ≤ N ≤ 25, the average jitter significantly increases as N increases. When N > 25, the system saturates and the average jitter increases to about 10 ms. To maintain the same average jitter, the system supports about one less IPsec RTP streams than original RTP streams. For example, to limit the average jitter to 1 ms, the system cannot support more than 17.77 original RTP streams or 16.75 IPsec RTP streams for G.729 (i.e., the IPsec overhead is 5.74%), and cannot supports more than 15.88 original RTP streams or 15.39 IPsec RTP streams for G.711 (i.e., the IPsec overhead is 9.79%).

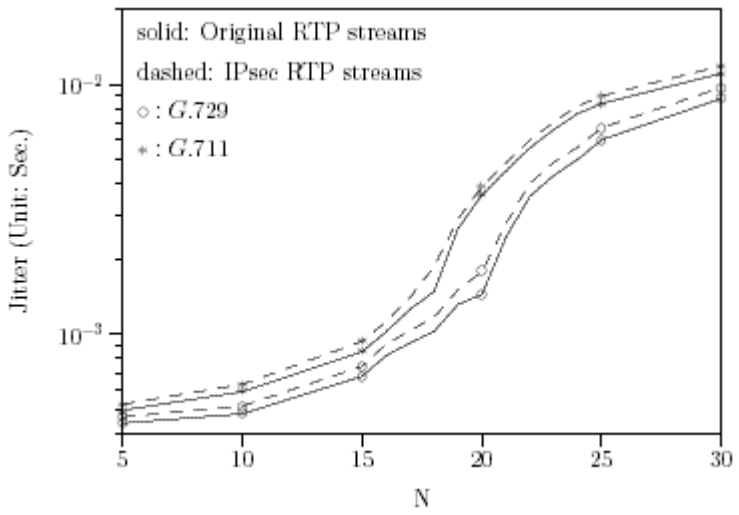


Figure 3.4 Jitter Performance (Without Jitter Buffer)

Figure 3.4 also indicates that the jitters for the G.711 RTP streams are larger than that for G.729 RTP streams. Since packet size for G.711 is larger than that for G.729, G.711 causes more link congestion than G.729 does.

In our experiments without jitter buffer, the average network jitters (between the MS and the WAG/PDG) range from 0.44 ms to 14.55 ms. Thus, to eliminate the jitter effect caused by

WLAN, at least one G.711 or G.729 RTP packet (i.e. 20 ms) should be buffered to achieve the jitter performance specified by [33]. Also, the jitter buffer size is not affected by IPsec in this case.

3.5 Conclusions

This chapter investigated the performance of IPsec-based VoIP service measured in the WLAN environment. The overheads caused by IPsec encapsulation were discussed in several aspects.

After IPsec encapsulation, the performance of IEEE 802.11b without packet loss degrades by 4% (in terms of the number N of RTP streams supported) for G.729 and 5% for G.711, respectively.

The IPsec encapsulation causes more packet processing, packet retransmisstions, and packet loss, and therefore results in extra latency. When $N < 15$, the IPsec overhead is less than 9.26%. When $15 \leq N \leq 20$, the latency overhead for IPsec significantly increases (can be up to 570.97%). For $N > 20$, the latency overhead for IPsec drops to less than 4.38%.

Similarly, IPsec encapsulation results in extra jitter overhead. When $N < 15$, the IPsec overhead is about 10%. When $15 \leq N \leq 25$, the jitter overhead for IPsec significantly increases (can be up to 26.3%). For $N > 25$, the jitter overhead for IPsec drops to less than 13.47% for G.711 and 8.24% for G.729.

Our study provides guidelines to select appropriate system parameters setups for the VoIP service over WLAN environment. Specifically, for an IEEE 802.11b AP, the system saturates if we pump more than 20 original RTP streams or 19 IPsec RTP streams for G.729 (or 18 original RTP streams or 17 IPsec RTP streams for G.711). Also, for transmission in IEEE 802.11b, at least one G.711 or G.729 RTP packet (i.e. 20 ms) should be queued in the jitter

buffer to achieve acceptable VoIP performance specified by [33]. We also observed that the jitter buffer size is not affected by IPsec encapsulation in the IEEE 802.11b configuration in our study.

Chapter 4
M-Taiwan Experience in VoIP-WiMAX Trial

Considering voice as a dominant telecommunication service, the performance of *Voice over IP* (VoIP) plays a critical role in deployment of WiMAX technology providing All-IP network services. To that effect, in this chapter we investigate the performance of a WiMAX-based VoIP established under the *Mobile Taiwan* (M-Taiwan) field-trial funded program. To achieve the objectives of the trial the measurement results expressed in the form of *Mean Opinion Score* (MOS), packet loss, packet delay and jitters. For the worst-case-scenario, the tests were conducted under a stringent condition of both communicating devices, wirelessly connected to the same WiMAX base station under a heavy background traffic and interference, were experiencing simultaneous handovers during the communication.

Upon our analysis the field measurements confirm an excellent performance when both communicating devices kept stationary and show an acceptable quality for the service when both communicating devices are on the move at a speed of 50 Km/h.

4.1 Introduction

Taiwan's Wi-Fi industry accounts for more than 90% of the global market share. In its quest to identify the next generation products, the Taiwan government has chosen *Worldwide Interoperability for Microwave Access* (WiMAX) [5] as one of the major directions for Taiwan's wireless industry, and has established the *Mobile Taiwan* (M-Taiwan) Program as the blueprint for an island-wide WiMAX environment. M-Taiwan aims at developing chip

sets and *base stations* (BS). For example, WiMAX chip sets have been developed by Mediatek, and the BSs have been developed by T-Com and ZyXEL. Furthermore, by creating its own WiMAX ecosystem, Taiwan offers not only manufacturing capabilities, but also an entire service and application test-bed for mobile services, mobile learning and mobile life. Since 2006, 18 large-scale WiMAX service trials have been deployed in Taiwan [34, 35].

In M-Taiwan, the *Voice over IP* (VoIP) service is considered as an enabling technology integrating broadband data applications with the voice. In particular, *IP Multimedia Core Network Subsystem* (IMS) [15] is utilized for voice and data integration. Under the support of the M-Taiwan Program, this chapter investigates the VoIP performance for a WiMAX deployment in Taipei City. In this chapter, we elaborate on the VoIP experimental environment, describe the output measures, and demonstrate the VoIP performance with and without mobility. The remainder of this chapter is organized as follows. Sections 4.2 and 4.3 provide a brief overview of VoIP service and WiMAX system. The general configuration set-up for the experimental field tests explained in Section 4.4 followed by the service performance measurement system in Section 4.5 and detailed results in Section 4.6.

4.2 VoIP Overview

With the explosive growth of the Internet subscriber population, VoIP has become the most promising low-cost option for voice communication over the IP network. In the M-Taiwan program, VoIP is implemented by using the *Session Initiation Protocol* (SIP) [22] and *Real-Time Transport Protocol* (RTP) [23].

4.2.1 Session Initiation Protocol and Real-Time Transport Protocol

IETF RFC 3261 defines SIP for Internet telephony [22]. As an application-layer signaling protocol over the IP network, SIP is designed for creating, modifying, and terminating multimedia sessions or calls. A SIP *customer premise equipment* (CPE) is installed with a user agent. The user agent contains both a *User Agent Client* (UAC) and a *User Agent Server* (UAS). The UAC (or calling user agent) is responsible for issuing SIP requests, and the UAS (or called user agent) receives the SIP requests and responds.

The SIP message specifies the *Real-Time Transport Protocol* (RTP) [23], which delivers the data in the multimedia sessions. Implemented on top of UDP, RTP detects packet loss and ensures an ordered delivery. The RTP packet also indicates the packet sampling time from the source media stream. The destination application can use this time stamp to calculate delay and jitter to provide the QoS feedback.

SIP conjuncts with protocols such as *Session Description Protocol* (SDP) [36] to describe the multimedia information. It conveys sufficient information to enable applications to join a session. During the session initiation, SDP describes the media type, media protocol, and codec number supported by the session endpoints to announce the endpoints capabilities. SDP provides the RTP information such as the network address and the transport port number of the RTP connection. Details of SIP and RTP can be found in [15].

4.2.2 E-Model

The quality of a communication service is traditionally based on subjective perception and typically measured by the *Mean Opinion Score* (MOS), which considers the effects of

equipment and impairment factors to subjectively quantify the perceived quality of a transmission such as voice based on typical users' perceptions]. The MOS values range are quantized to 5 levels, from 1 to 5, where 1 is unacceptably bad, 2 is poor, 3 is fair, 4 is good, and 5 is excellent. The ITU-T G.107, however, defines an E-Model which provides a computational model for rating the end-to-end transmission performance for the VoIP service [37]. The E-Model considers different kinds of transmission impairments add on linearly to the scale of the rating factor R. The model then converts the value of R into a MOS scale that quantifies an overall conversational quality.

The rating factor R is then expressed as follows,

$$R = R_o - I_s - I_d - I_{e\text{-}eff} + A. \quad (4.1)$$

In the right-hand side of (4.1), these factors are described as follows:

R_o: The basic signal-to-noise ratio includes the noise sources such as circuit noise and room background noise.

I_s: The simultaneous impairment factor combines the impairments that occur simultaneously with the voice signal. These impairments include the quality degradation caused by the overall loudness, non-optimum sidetone and quantizing distortion.

I_d: The delay impairment factor represents the impairments due to delay in arrival of the voice signal.

$I_{e\text{-}eff}$: The effective equipment impairment factor represents impairments caused by low bit-rate CODEC and the impairments due to random packet loss.

A: The advantage factor allows for compensation of impairment factors when there are other advantages of access to the user. ITU-T G.107 suggests the default value 0 for *A*.

The rating factor *R* is then converted into an estimated MOS value as follows,

$$\begin{cases} \text{For } R < 0\text{: MOS} = 1 \\ \text{For } 0 < R < 100\text{: MOS} = 1 + 0.035R + R \times (R - 60) \times (100 - R) \times 7 \times 10^{-6} \\ \text{For } R > 100\text{: MOS} = 4.5 \end{cases} \quad (4.2)$$

Therefore, the estimated MOS values range from 1 to 4.5

The relation (4.2) between the estimated MOS value and the rating factor *R* is illustrated in Figure 1.

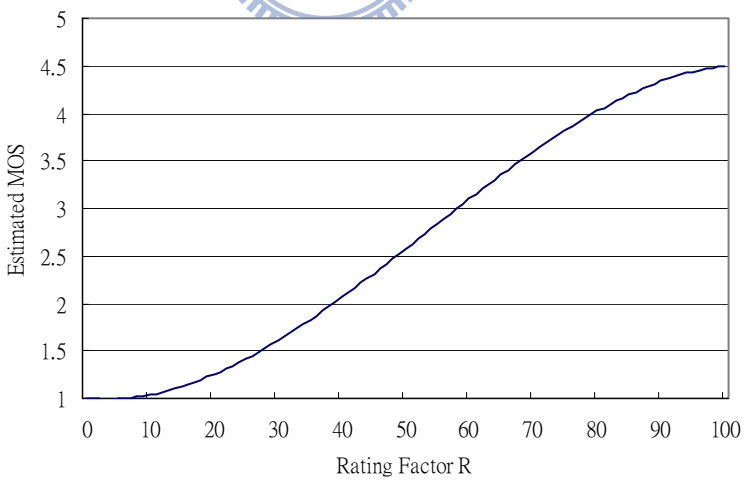


Figure 4.1 Estimated MOS Value as a Function of Rating Factor *R*

4.3 WiMAX Overview

Following the success of the Internet technology, broadband data communication services have been provisioned to the expert communities for decades, which for the wired and fiber connections have been achieved with the turn of the century. For wireless it is due any time within the next decade where superior mass production of quality wireless components to extend the frequency range and overcome shadowing and multipath fading issues using super sensitive receivers [38]. Now, with the industry capable of providing the WiMAX technology for superiority of virtually nil infra-structure costs, we are able to offer a data-enabled very low cost *wireless metropolitan area network* (WMAN) style *wireless broadband access* (WBA) solutions which in long run may overshadow competitive solutions [39] due to the fact that WiMAX is able to provide broadband wireless access with wide service coverage, high data throughput, high mobility and greater service flexibility [40, 41]. Figure 4.2 shows a simplified WiMAX network architecture, which consists of the *access service networks* (ASNs; see Figure 4.2 (a)) and the *connectivity service networks* (CSNs; see Figure 4.2 (b)). An ASN provides radio access (such as radio resource management, paging and location management) to the WiMAX *mobile station* (MS; Figure 4.2 (e)). The ASN comprises *ASN gateways* (ASN-GWs; see Figure 4.2 (c)) and WiMAX BSs (see Figure 4.2 (d)). Every ASN-GW connects to several BSs. The ASN-GWs are also connected to each other to coordinate MS mobility. A CSN consists of network nodes such as the *mobile IP* (MIP) *home agent* (HA; see Figure 4.2 (f)) [3], the *authentication authorization, and accounting* (AAA) server (see Figure 4.2 (g)) and the *dynamic host configuration protocol* (DHCP) server (see Figure 4.2 (h)). The CSN provides IP connectivity (such as Internet access and IP address allocation) to a WiMAX MS and interworks with the ASNs to support capabilities such as AAA and mobility management. Before an MS is allowed to access WiMAX services, it must

be authenticated by the ASN-GW (which serves as the authenticator) and the AAA server in the CSN.

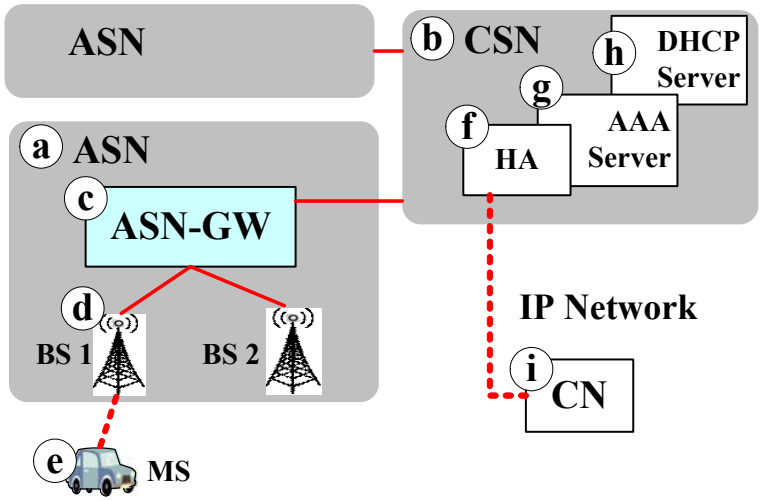


Figure 4.2 Simplifiied WiMAX Network Architecture

The WiMAX *Physical* (PHY) and *Media Access Control* (MAC) layers are defined in IEEE 802.16 standard to support multiple services with point-to-multipoint and mesh broadband wireless access [35]. The point-to-multipoint mode defines one-hop communication between a BS and an MS, while the mesh mode allows traffic to be directly exchanged and forwarded among neighboring BSs. IEEE 802.16 is initially designed as an access technology for WMAN. The first specification IEEE 802.16-2004 targets on fixed and nomadic accesses. In IEEE 802.16e-2005 amendment, the IEEE 802.16e system (Mobile WiMAX) further provides functions to facilitate mobile accesses. We introduce the functions of MAC and PHY layers in the following subsections. Details of WiMAX technology can be found in [38]. Figure 4.3 illustrates the IEEE 802.16 protocol stack. The functions of the WiMAX PHY and MAC layers are described in the following subsections.

4.3.1 The Media Access Control Layer

There are three sublayers in IEEE 802.16 MAC layer: service-specific *convergence sublayer* (CS; see Figure 4.3 (a)), the MAC common part sublayer (see Figure 4.3 (b)), and the security sublayer (see Figure 4.3 (c)).

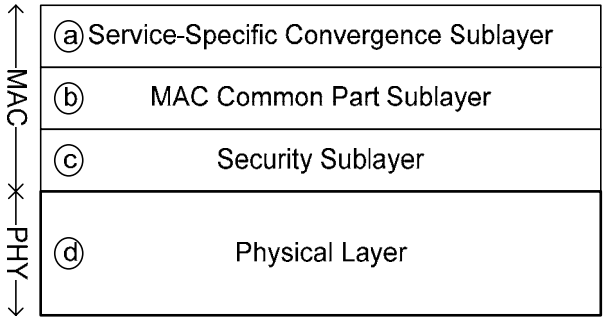


Figure 4.3 IEEE 802.16 Protocol Stack

The service-specific CS performs packet classification, header suppression, and converts packets between the upper layer and the MAC layer. The IEEE 802.16 currently supports packet CS and ATM CS to interface with IP and ATM protocol layers, respectively. In IEEE 802.16, the connections between the MSs and the BSs can be identified with unique *connection identifications* (CIDs). The packet CS may check the IP or TCP/UDP header of a packet to determine its CID. Besides the CID mapping, the CS may perform the optional payload header suppression to eliminate the redundant parts of the packets during the transmission over the air interface.

The MAC common part sublayer provides the medium access, connection management, and QoS functions that are independent of specific CSs. After the packets are processed by the CS, the MAC common part may perform *automatic repeat request* (ARQ) for retransmitting lost packets. ARQ is optional in IEEE 802.16 but is mandatory for IEEE 802.16e.

In IEEE 802.16, QoS functions are implemented in the MAC common part sublayer. Several service classes are defined to satisfy various QoS requirements. For example, a VoIP connection is often associated with *unsolicited grant service* (UGS) to support *constant bit-rate* (CBR) or CBR-like flows with constant bandwidth allocation. According to the QoS associated, the BS schedules radio resources with various scheduling disciplines, such as round-robin and first-in-first-out.

The security sublayer provides privacy and protections through encryption, decryption, and authentication. In IEEE 802.16, an MS is requested to perform the authentication and authorization before attaching to a WiMAX network. During the authorization procedure, the MS negotiates with the BS to generate the session key. To perform packet encryption and decryption, each connection is linked with a *security association* (SA), which contains the security information and settings such as encryption keys. Packet encryption and decryption are exercised based on the information in the SA.

Before accessing the WiMAX network, an MS should perform a complete spectrum search and synchronize the time and frequency with a BS through the ranging procedure. Then the MS starts the network entry procedure to negotiate the capabilities with the BS and performs authorization process to generate the keys used between the MS and the BS. Finally, the MS obtains an IP address from the BS, and establishes data connections with the BS.

4.3.2 The Physical Layer

In the PHY layer (see Figure 4.3 (d)) IEEE 802.16 defines several specifications for different frequency ranges and applications. For example, *orthogonal frequency division modulation* (OFDM) is used for non-line-of-sight operations in the frequency bands below 11 GHz. By

extending the OFDM technology, *orthogonal frequency division multiple access* (OFDMA) allows one channel to be shared by multiple users. The IEEE 802.16 standard defines a set of adaptive modulation and coding rate configurations that can be used to trade off data rate against system robustness under various wireless propagation and interference conditions. The allowed modulation types are *binary phase shift keying* (BPSK), *quadrature phase shift keying* (QPSK), *16-quadrature amplitude modulation* (16QAM), and 64QAM [35].

Several duplexing technologies are provided in IEEE 802.16. In *time division duplex* (TDD), a WiMAX frame consists of a *downlink* subframe and an *uplink* subframe and a short transition gap is placed between the downlink and uplink subframes for receive and transmission transitions in the radio. The gap between the downlink burst and the subsequent uplink burst is called *transmit/receive transition gap* (TTG). The gap between the uplink and the subsequent downlink is called *receive/transmit transition gap* (RTG).

The duration of an OFDM symbol includes the useful symbol time and a prefix. In OFDM, all users within the same cell or sector use orthogonal subcarriers to carry the OFDM symbols. The OFDM symbol uses a fixed-length *cyclic prefix* (CP) to counteract the intersymbol interference. The ratio of the CP length to the useful symbol time is defined as the guard interval, which is used by the receiver to collect signals from multiple paths and improve system performance.

4.4 VoIP Experimental Environment

The main bulk of this trial service performance measurement has been conducted during 2007-08 in the Taipai area under various communication conditions. Based on the abstract network in Figure 4.2, Figure 4.4 illustrates the network architecture for one of the WiMAX

deployments in the M-Taiwan Program. Based on mobile WiMAX (IEEE 802.16e-2005) technology [5], more than 52 WiMAX BSs have been deployed. The WiMAX ASN-GW (a Foundry's Netlron XMR400 plus Motorola's CAP Controller) is located in Taipei County. The distances between the BSs to be tested in our study and the ASN-GW range from 18.5 Km to 21 Km. Every BS is connected to the ASN-GW through a 50 Mbps optical fiber link. The ASN-GW connects to the *Foreign Agent* (FA; which is a Redback's SmartEdge 400) through *Gigabit Ethernet* (GE). The FA connects to a core router (Juniper's M120) through another GE. The core router connects to an L2 switch (Cisco's Catalyst 3560E) through GE. The L2 switch connects to the HA (a Starent's ST-16 Intelligent Mobile GW) through GE, and connects to an FTP server through a 10/100M *Fast Ethernet* (FE). In the above configuration, backup for ASN-GW controller, FA, and core router are also deployed to support reliability and availability.

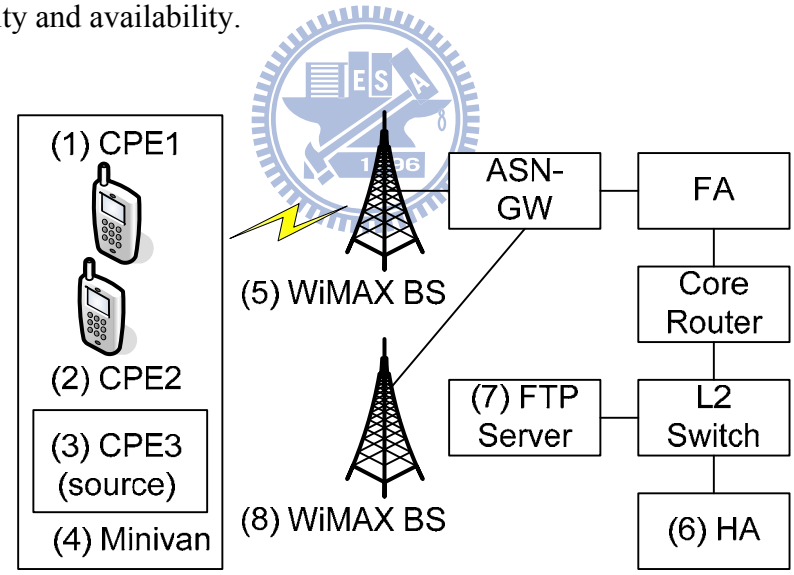
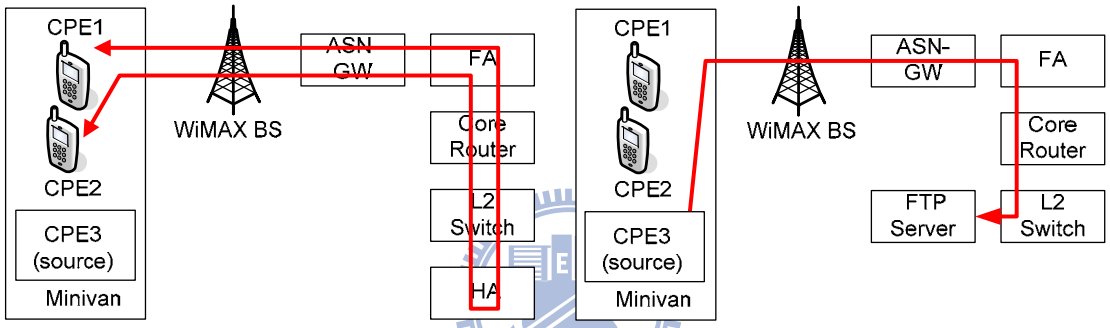


Figure 4.4 M-Taiwan VoIP Experimental Environment

In this experimental environment, the WiMAX MSs are installed SIP call agents, and serve as SIP CPEs. The VoIP calls are generated and measured between WiMAX CPE1 (Figure 4.4 (1)) and WiMAX CPE2 (Figure 4.4 (2)). Our experiments also include the background data traffic, which is generated from WiMAX CPE3 (Figure 4.4 (3)) to the FTP server Figure 4.4

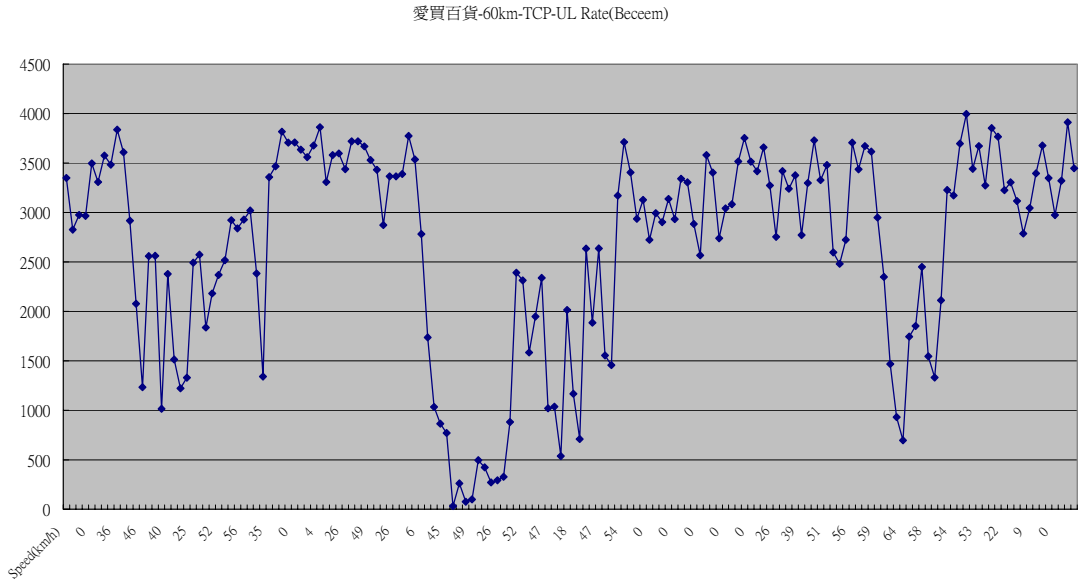
(7)). These three CPEs are notebooks connected to Quanta/Beceem's WiMAX (wave 2) USB dongles, and are all located in a minivan (see Figure 4.4 (4) and Figure 4.6 (left)). As illustrated in Figure 4.5 (a), it is clear that a one-way VoIP link between CPE1 and CPE2 consists of 12 hops (CPE1 ←→ BS ←→ ASN-GW ←→ FA ←→ Core Router ←→ L2 Switch ←→ HA ←→ L2 Switch ←→ Core Router ←→ FA ←→ ASN-GW ←→ BS ←→ CPE2). In Figure 4.5 (b), the data path between CPE3 and the FTP server includes 6 hops (CPE3 ←→ BS ←→ ASN-GW ←→ FA ←→ Core Router ←→ L2 Switch ←→ FTP Server).



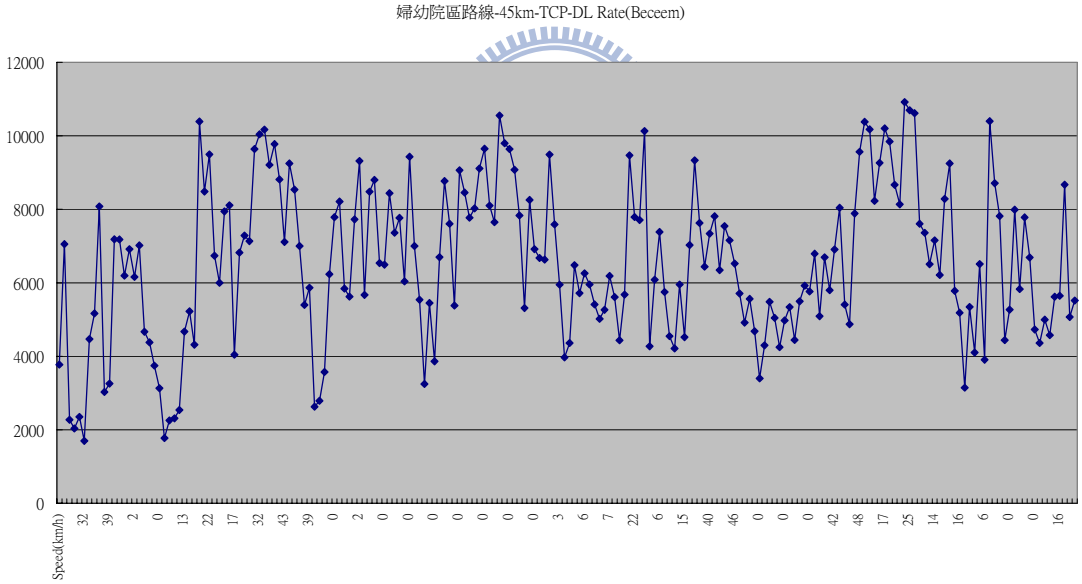
(a) VoIP Path between CPE1 and CPE2 (b) FTP Path between CPE3 and FTP Server

Figure 4.5 Data Paths in the Experiments

Figure 4.6 WiMAX CPEs in the Minivan (left) and the WiMAX Antenna (right)



(a) TCP uplink transmission rate (Kbps) vs CPE speed (Km/h)



(b) TCP downlink transmission rate (Kbps) vs CPE speed (Km/h)

Figure 4.7 Real-Time Measures of TCP Transmission Rate at Various CPE Speeds

In our study, a 3-sector WiMAX BS (Figure 4.6 (right)) is typically installed at the roof of a building with the coverage of 1.5 Km in diameter. To fully utilize existing cellular infrastructure, the WiMAX antenna may be collocated with the WCDMA antenna. The

WiMAX antenna is an adaptive system with beamforming. In this WiMAX network deployment, the TDD ratio for downlink and uplink can be 3 to 1 or 3 to 2. In our experiments, 3-to-1 ratio is considered. The modulation schemes are 16QAM 3/4,16QAM 1/2, QPSK 3/4, QPSK 1/2 for uplink, and 64QAM 5/6, 64QAM 3/4, 64QAM 2/3, 64QAM 1/2, 16QAM 3/4,16QAM 1/2, QPSK 3/4, QPSK 1/2 for downlink. We observed that the bandwidth performance is significantly improved by up to 100% in our measurements when the modulation scheme is enhanced from 64QAM 1/2 to 64QAM 5/6 for downlink, and from 16QAM 1/2 to 16QAM 3/4 for uplink. Through measurements of 14 experiments, the average TCP uplink transmission rate is 3.668 Mbps. Figure 4.7 (a) plots a typical experiment of TCP uplink transmission rate as the CPE speed changes. The sample points are measured for every 2-3 seconds. The figure indicates that the transmission rate drops significantly as the CPE suddenly accelerates (e.g., when the speed increases from 6 Km/h to 59 Km/h).

The average downlink TCP transmission rate of the BSs is 10.01 Mbps (per sector). Figure 4.7 (b) plots a typical experiment of TCP downlink transmission rate as the CPE speed changes.

For a stationary CPE, the maximum and minimum uplink TCP bandwidths are 2.879 Mbps and 2.306 Mbps, respectively. The average uplink bandwidth is 2.492222 Mbps. The maximum and minimum downlink TCP bandwidths are 7.781 Mbps and 4.741 Mbps, respectively. The average downlink bandwidth is 6.881778 Mbps.

We also measure the handover delays. The average handover delays of 5 measurements are 67.78 ms for inter-BS handover at 30 Km/h, 68.125 ms for inter-BS handover at 50 Km/h, 63.5 ms intra-BS handover at 30 Km/h, and 65 ms for intra-BS handover at 50 Km/h. Therefore, as the CPE speed increases form 30 Km/h to 50 Km/h, the handover delay increases by 0.51%-2.3%. The inter-BS handover time is 4.8%-6.7% longer than the intra-BS

handover.

Figure 4.8 Moving Path for Mobility Tests (the solid path is covered by one base station, and the dashed path is covered by another base station)

In the stationary tests, the distance between the CPEs and the BS is about 210 meters, and the output data are measured at a base station located at the 7th floor of a building in Nei-Hu area of Taipei City. In the mobility tests, two WiMAX BSs are involved. These BSs are located near the Taipei City Hall. To produce the handover effect under the controlled condition, the minivan carrying the CPEs repeatedly drove on the roads around a square area adjacent to the City Hall (see Figure 4.8). This path is covered by two WiMAX BSs and the distances between the BSs and the CPEs range from 150 meters to more than 400 meters.

4.5 VoIP Experimental Setup for Output Measurement

As described in Section 4.2.1, in M-Taiwan, SIP is used to control a VoIP call and RTP is used to deliver the voice data. In this chapter, we focus on the RTP packet performance. The SIP call setup signaling can be found in [15], and will not be discussed in this chapter.

We use the NetIQ Chariot tool [42] to measure the MOS values following the E-model described in Section 2.2. To collect the measured data, the Chariot endpoints are installed in the VoIP CPEs running on Windows XP 6.2. The Chariot Console resides in one of the CPEs.

In VoIP, *CODEC* are used to convert analog voice to digital samples so that the voice information can be delivered in the IP network. The VoIP codec techniques determine the maximum MOS value. Our experiments utilize the high-quality G.711 codec that consumes larger bandwidth as compared with other CODEC. The maximum MOS value for G.711 is 4.4 (which is lower than the theoretical value 5). Although other CODEC such as G.729 (with maximum achievable MOS value 4.07) are also supported in M-Taiwan deployments, G.711 is selected for presentation in this chapter because WiMAX support broadband applications, and therefore can comfortably accommodate G.711 CODEC. The default G.711 data rate is 64 Kbps.

There are several techniques to improve the performance of the codecs. With silence suppression, when no one is talking during a call, VoIP data are not delivered to save the network bandwidth. In our experiments, silence suppression is disabled to obtain a more intense assessment of the network. The G.711 *packet loss concealment* (PLC) option mitigates the VoIP data loss effects and therefore improves the MOS estimate. This option is turned off in our experiments for a stringent scenario investigation.

Besides MOS, the following output measures are also investigated in this chapter:

Packet loss can severely impair call quality because voice information cannot be received by the listeners. Data loss in bursts is more serious than uniform random loss because humans perceive bursts of loss as impairments to audio quality much more than uniform random loss. Bursts of loss are often observed in radio links, and are a major measure we would

like to investigate in this study. Packet loss is included in the MOS calculation (see factor $I_{e\text{-}eff}$ in Section 4.2.2). WiMAX Forum requires that packet loss be less than 1%.

One-way Packet Delay is the time period for a VoIP packet to travel from one CPE to another. Typically, the voice quality is acceptable when the voice delay is less than 150 ms. When it exceeds 200 ms, the listeners will experience the walkie-talkie effect with poor audio quality [3]. In our experiments, the one-way packet delay includes the *propagation delay* contributed by 10 IP hops (from the BS to the HA and back), the *transport delay* contributed by 2 WiMAX radio links, the G.711 *packetization delay***,** and *jitter buffer delay*. G.711 introduces packetization delay to convert a signal from analog to digital. In our experiments, the packetization delay is set to 20 ms. This delay is included in the MOS estimate (see factor I_d in Section 4.2.2). The WiMAX Forum specifies the preferred packet delay to be less than 150 ms, and the limited delay to be 200 ms.

Jitters or the variation of packet inter-arrival time may create unexpected pauses between utterances, and therefore affect the intelligibility of the VoIP speech. It was reported that an average jitter exceeding 35 ms [28] or 50 ms [42] results in unacceptable QoS for VoIP. The WiMAX Forum requires that jitter is less than 25 ms. In order to reduce the jitters, a buffer is used to store the incoming packets before they are played. If the jitter buffer size is too small, network jitter will result in packet loss and therefore degrade the intelligibility of the voice. If the jitter buffer size is too large, long packet delay will be experienced, which results in QoS degradation (e.g., echo level may be more easily perceptible). Our previous study indicated that buffering one packet is sufficient for WLAN if the core network delay (transport delay) is not considered [40]. Default G.711 jitter buffer delay is 40 ms (2 packets) in our experiments, which is also included in the MOS estimate.

In some wireless VoIP experiments, only one call party resides in wireless network and the

other call party is directly connected to Internet through wired network [40, 41]. We consider a stringent scenario where both CPEs (CPE1 and CPE2 in Figure 4.4) are wirelessly connected to the same WiMAX BS, and will handover at the same time. To our knowledge, the behavior of simultaneous handovers for both call parties is seldom reported in the literature. We also use a third CPE (CPE3 in Figure 4.4) to generate the uplink background traffic. Downlink background traffic is not considered because the WiMAX uplink is the bottleneck (due to the 1 to 3 uplink-to-downlink bandwidth ratio), and our past experience indicated that the impact of WiMAX uplink background traffic is more significant than that of downlink background traffic.

Based on the configuration illustrated in Figure 4.4, there are two VoIP links and one background traffic link in every experiment. The background traffic with 512 Kbps, 1 Mbps, 2 Mbps and 3 Mbps are considered. We note that the 3 Mbps background traffic consumes most of WiMAX uplink bandwidth. In terms of CPE mobility, we consider three cases: stationary (no mobility), 30 Km/h, and 50 Km/h.

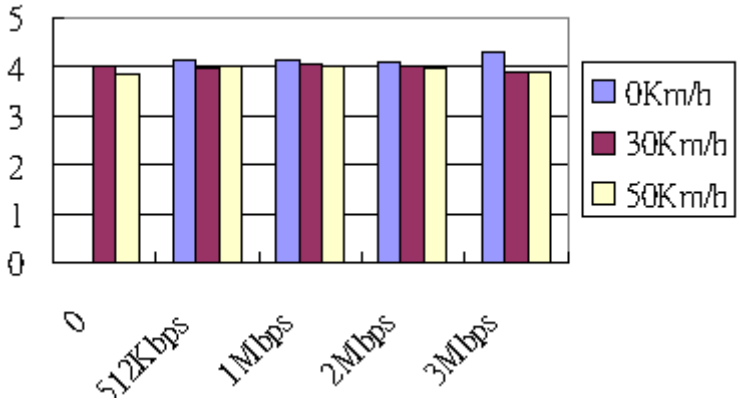
4.6 Wireless-to-Wireless VoIP Measurement Results

Our study is conducted in Taipei City, where the RF environment is affected by tall buildings and heavy vehicle traffics, and more serious interference is observed as compared with the line-of-sight environment. Every stationary test is conducted for 5 minutes, and every mobility test is conducted for 2 minutes. During a stationary test, roughly 2,400,000 bytes were sent in one-way VoIP link. Similarly, in a mobility test, we measured roughly 960,000 bytes for one-way VoIP link. Therefore, the equivalent bandwidth consumed is about 80 Kbps, which is higher than the default G.711 data rate (i.e., 64 Kbps) due to extra RTP header overhead.

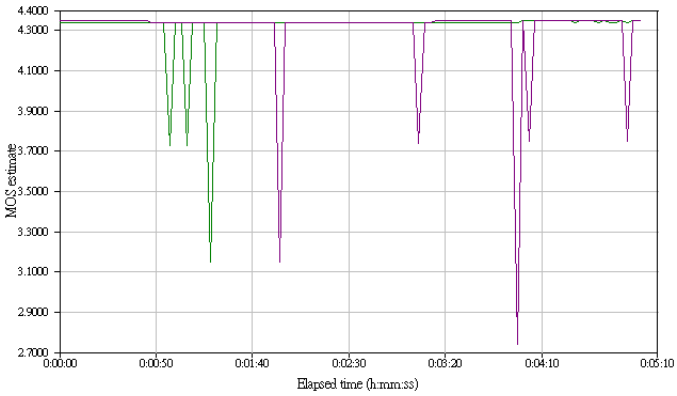
This section shows the effects of CPE mobility and background traffic on MOS, packet loss, packet delay and jitter. Figure 4.9-4.12 (a) show the expected MOS values. Figures 4.9-4.12 (b)-(d) give the real-time measures of an example experiment.

4.6.1 Mean Opinion Score (MOS)

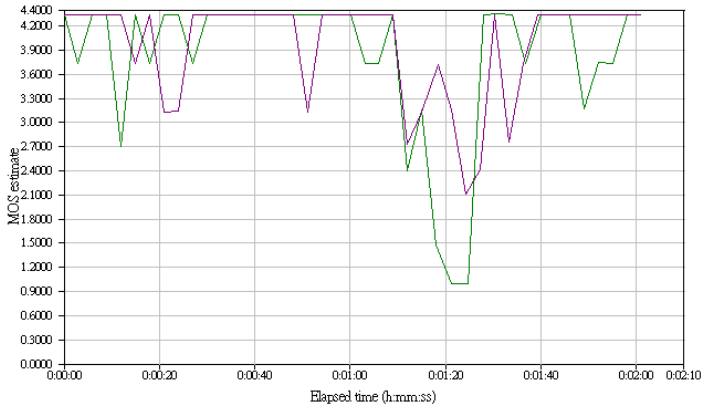
Figure 4.9 shows the MOS performance. Figure 4.9 (a) indicates that the average MOS values for stationary CPEs are above 4.0, and are larger than 3.9 for moving CPEs. The MOS slightly decreases as the CPE speed increases. The MOS values are insignificantly affected by the background traffic.



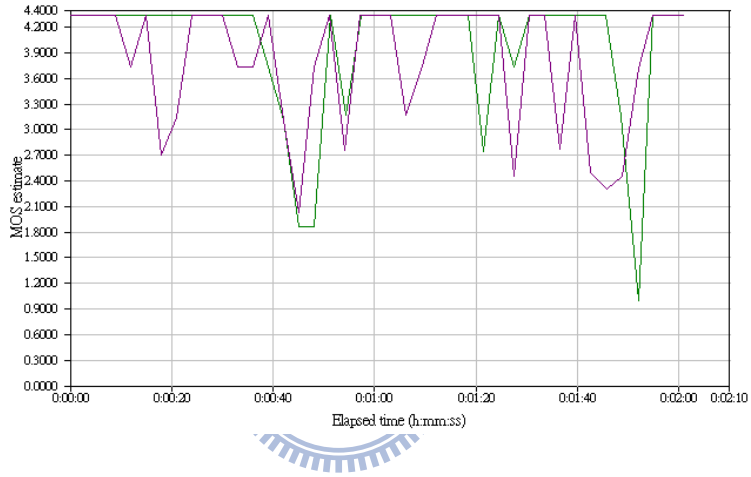
` (a) Average MOS



(b) CPE Speed: 0 Km/h (Uplink Background Traffic: 3 Mbps)



(c) CPE Speed: 30 Km/h (Uplink Background Traffic: 3 Mbps)



(d) CPE Speed: 50 Km/h (Uplink Background Traffic: 3 Mbps)

Figure 4.9 MOS Measurements

Figures 4.9 (b)-(d) illustrate real-time MOS measurements of CPE1 (the green curve) and CPE2 (the purple curve) in a typical experiment, where the uplink background traffic of CPE3 is 3 Mbps. When both CPEs are stationary, the real-time MOS values are measured for 5 minutes. In this case, the MOS is typically maintained higher than 3.8, and the values of most MOS drops are still above 3.0. These MOS drops are partly due to tall buildings and the street traffic surrounding the minivan of the CPEs. The 5-minute average MOS values are 4.32 at CPE1, and 4.3 at CPE2.

In every mobility test, the real-time MOS values are measured in 2 minutes. At the speed of 30 Km/h, a handover occurs roughly at the 80th second. The real-time MOS may drop significantly from 4.35 to 1.0 (for CPE1) and 2.1 (for CPE2) as illustrated in Figure 4.9 (c). The average MOS values are 3.83 (for CPE1) and 3.96 (for CPE2). At the speed of 50 Km/h (Figure 4.9 (d)), after the first handover occurring at the 50th second, the MOS values become very unstable, and the MOS is not recovered back to 4.34. The 2-minute average MOS values are 3.97 (for CPE1) and 3.79 (for CPE2), respectively.

Our study indicates that the CPE mobility does not degrade the MOS performance except when the handovers occur.

4.6.2 Packet Loss

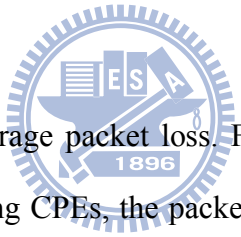


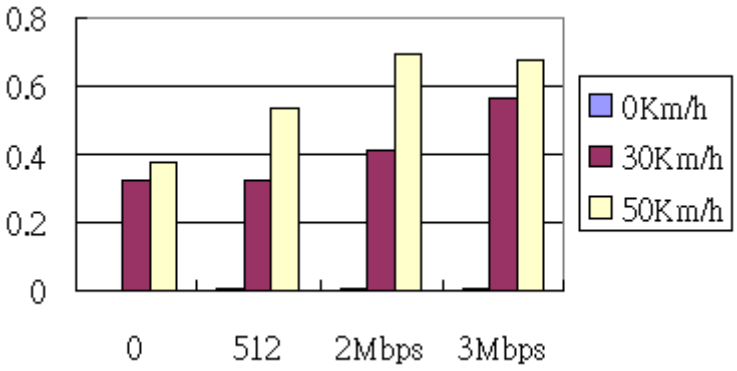
Figure 4.10 (a) illustrates the average packet loss. For stationary CPEs, the average packet loss is less than 0.01%. For moving CPEs, the packet loss is less than 0.7%. The packet loss increases as the CPE speed increases.

The packet loss of stationary CPE is not affected by the background traffic. On the other hand, the background traffic significantly affects the moving CPEs. At CPE speed of 30 Km/h (50 Km/h), the packet loss increases from 0.325% (0.375%) to 0.566% (0.675%) when the background traffic increases from 0 to 3 Mbps.

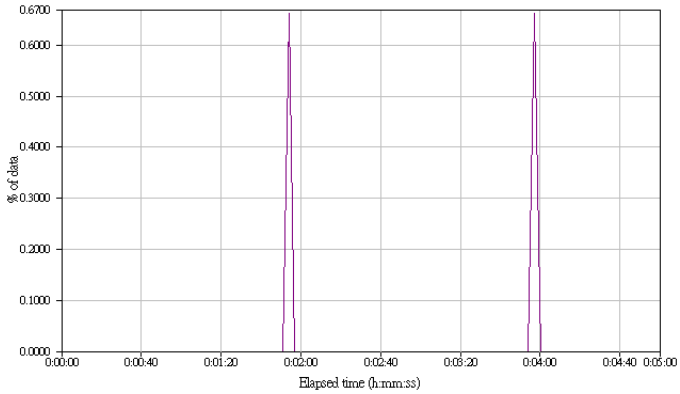
Like Figures 4.9 (b)-(d), Figures 4.10 (b)-(d) illustrate an example of real-time packet loss measurements with various CPE speeds. When both CPEs are stationary, most packet loss values are less than 0.064%. The 5-minute average packet loss value is 0.013%.

At the speed of 30 Km/h (Figure 4.10 (c)), the 2-minute average packet loss values are

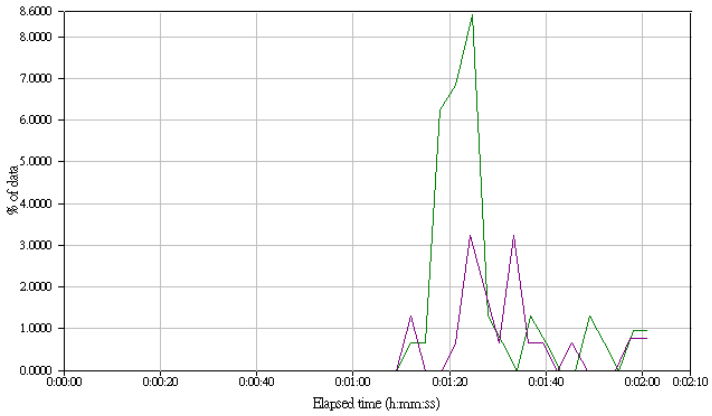
0.783% (for CPE1) and 0.35% (for CPE2). At the speed of 50 Km/h (Figure 4.10 (d)), the 2-minute average packet loss is 0.7% (for CPE1) and 0.65% (for CPE2), respectively.



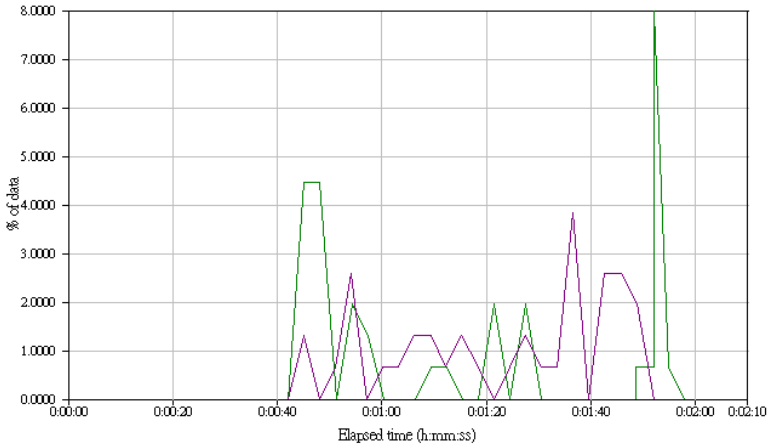
(a) Average Packet Loss (%)



(b) CPE Speed: 0 Km/h (Uplink Background Traffic: 3 Mbps)



(c) CPE Speed: 30 Km/h (Uplink Background Traffic: 3 Mbps)



(d) CPE Speed: 50 Km/h (Uplink Background Traffic: 3 Mbps)

Figure 4.10 Packet Loss Measurements

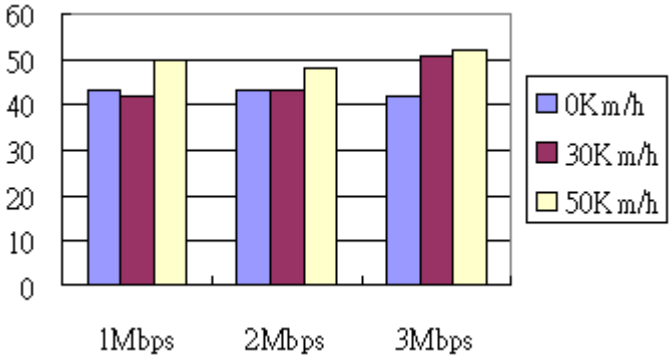
We note that packet loss in bursts is more damaging than uniform random packet loss. The voice quality is affected when consecutive five or more packets are lost at a time. Figure 4.10 (b) shows that for stationary CPEs, the maximum number of consecutive lost packets is 1. For CPEs moving at 30 Km/h (Figure 4.10 (c)), packet loss in bursts are observed when handovers occur, and the maximum number of consecutive lost packets is 3 occurring at handover. For CPEs moving at 50 Km/h (Figure 4.10 (d)), lost packets in bursts are more serious. The maximum number of consecutive lost packets is 3. In our experiments, the packet loss measures satisfy the requirement of WiMAX Forum (i.e., less than 1%).

4.6.3 One-way Packet Delay

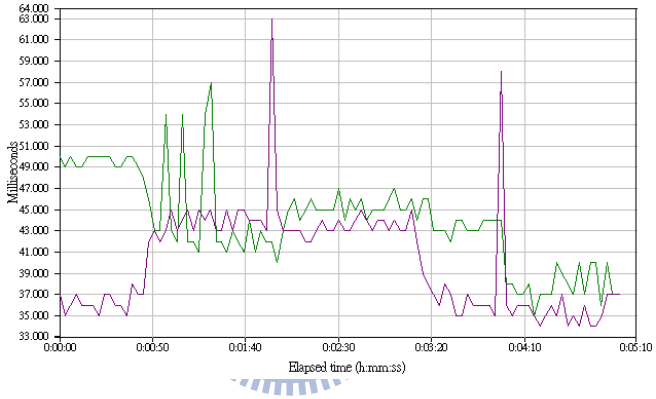
In Figure 4.11 (a), the average one-way packet delay (including 10 IP hops and two WiMAX radio links) is less than 45 ms for stationary CPEs, and is less than 52 ms for moving CPEs. The delay increases as the CPE speed increases (because of the handover impact).

When the background traffic increases, the one-way packet delay tends to increase for moving

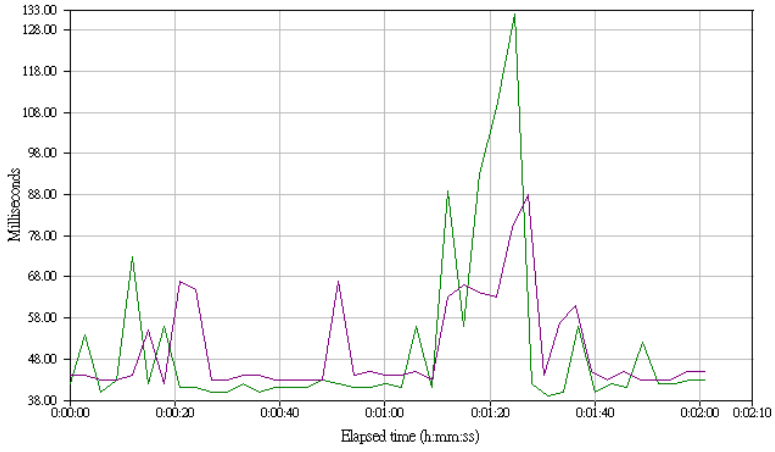
CPEs. The background traffic effect on stationary CPEs is negligible.



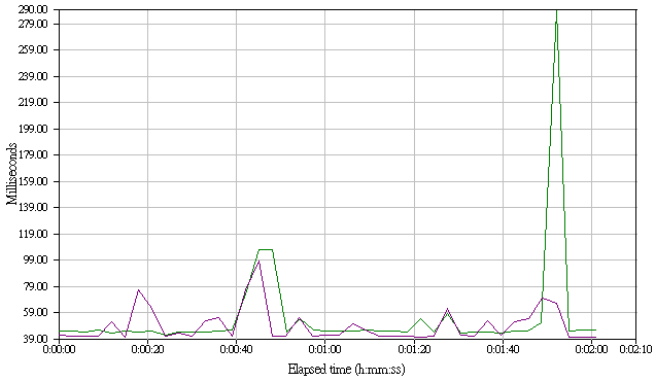
(a) Average One-way Packet Delay (ms)



(b) CPE Speed: 0 Km/h (Uplink Background Traffic: 3 Mbps)



(c) CPE Speed: 30 Km/h (Uplink Background Traffic: 3 Mbps)



(d) CPE Speed: 50 Km/h (Uplink Background Traffic: 3 Mbps)

Figure 4.11 One-way Packet Delay Measurements

Figures 4.11 (b)-(d) illustrate an example of real-time packet delay measurements with different CPE speeds. When both CPEs are stationary, packet delays are always less than 63 ms. The 5-minute average packet delay is 44 ms at CPE1, and 40 ms at CPE2.

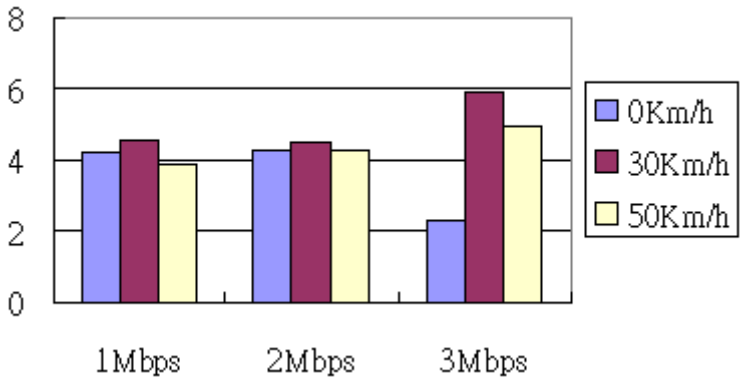
At the speed of 30 Km/h (Figure 4.11 (c)), most packet delays are less than 88 ms, and the maximum packet delay is 132 ms occurring at the handover. The 2-minute average packet delay is 51 ms (for CPE1) and 50 ms (for CPE2). At the speed of 50 Km/h (Figure 4.11 (d)), most packet delays are less than 100 ms, and the maximum packet delay is 289 ms. The 2-minute average packet delay is 55 ms (for CPE1) and 49 ms (for CPE2).

In our experiments, most packet delays are much less than the acceptable upper limit of packet delay (i.e., 150 ms).

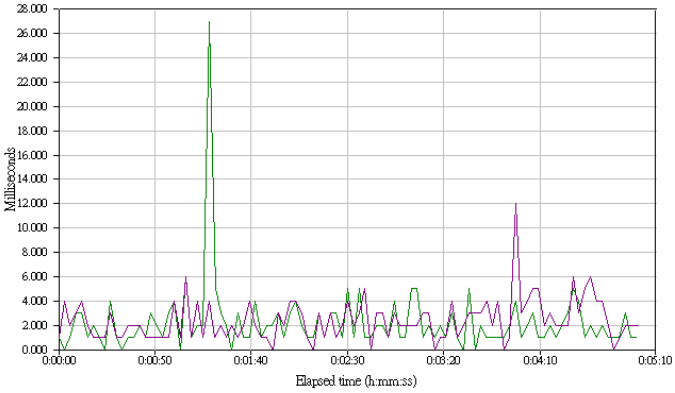
4.6.4 Jitters

Figure 4.12 (a) shows that the average jitter is less than 4.3 ms for stationary CPEs, and is less than 6 ms for moving CPEs. Our experiments indicate that for stationary CPEs, the background traffic seems not affect jitter. For moving CPEs, jitter increases as the background

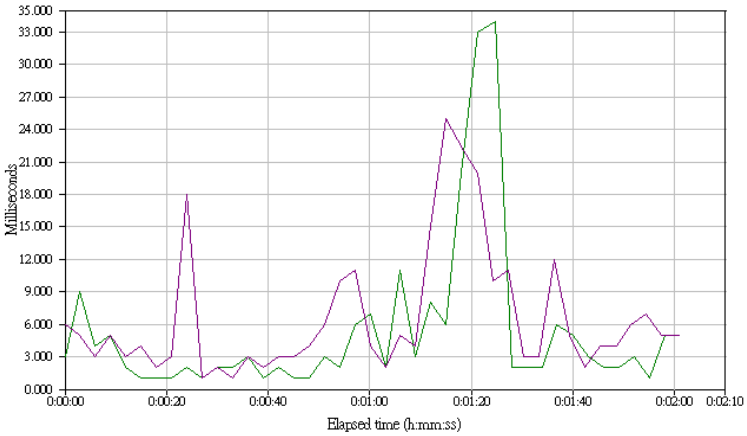
traffic increases. However, it is not clear why CPE speed at 30 Km/h tends to have the worst jitter performance (such phenomenon was also observed in other experiments).



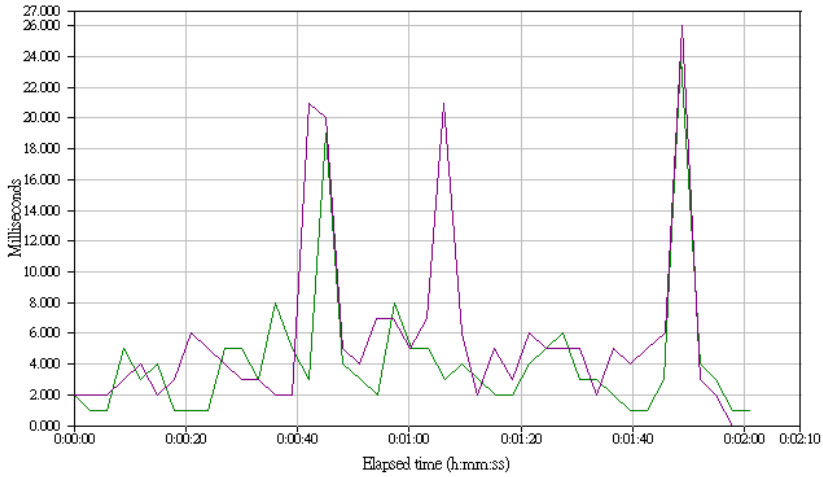
(a) Average Jitter (ms)



(b) CPE Speed: 0 Km/h (Uplink Background Traffic: 3 Mbps)



(c) CPE Speed: 30 Km/h (Uplink Background Traffic: 3 Mbps)



(d) CPE Speed: 50 Km/h (Uplink Background Traffic: 3 Mbps)

Figure 4.12 Jitter Measurements

Figures 4.12 (b)-(d) give examples of real-time jitters measurements. When both CPEs are stationary, all jitters values are less than 27 ms. The 5-minute average jitter is 2.23 ms at CPE1, and 2.42 ms at CPE2.

At the speed of 30 Km/h (Figure 4.12 (c)), the maximum jitter is 34 ms when the handover occurs. The 2-minute average jitter values are 5.225 ms (for CPE1) and 6.6 ms (for CPE2). At the speed of 50 Km/h (Figure 4.12 (d)), maximum jitter is 27 ms. The average jitters values are 4.2 ms (for CPE1) and 5.75 ms (for CPE2). After the handover, jitters occur in bursts. Figure 12 (c) and (d) show that jitter is more seriously affected by handover at 30 Km/h than that at 50 Km/h.

The study also indicates the following observations:

- The impact of background traffic on VoIP is mostly insignificant.

- The MOS values are slightly decreases as the CPE speed increases. The MOS values are not affected by the background traffic.

- The packet loss increases as the CPE speed increases. The packet loss of stationary CPE is insignificant, and is not affected by the background traffic. On the other hand, the background traffic significantly affects the moving CPEs.

- The one-way packet delay increases as the CPE speed increases. The background traffic slightly affects the packet delays for moving CPEs. The background traffic effect on stationary CPEs is negligible.

- Impacts of CPE speed and background traffic on the jitters is not clear in our study. However, all experiments indicate resilience against jitters.

- The values of all jitter-samples observed in our study are much lower than the unacceptable jitter value (i.e., 25 ms).

4.7 Conclusions

Our investigation upon the experimental results indicates the performance of a VoIP service using the WiMAX-based infrastructure of the M-Taiwan Program conforms very well to the standard requirements of G.107 under the worse-condition and stringent scenario where both VoIP CPEs are wirelessly connected to the same WiMAX base station with both moving CPEs at the speeds up to 50 Km/h while both going under handovers at the same time.

Chapter 5
Conclusions and Future Work

Voice over Internet Protocol (VoIP) is a promising low-cost voice communication over the wired or wireless Internet network. In the mobile/wireless environment, the radio resource is restricted and the reliability of the wireless transmission is much poor than that of the wired environment. To provide satisfactory VoIP services in the mobile/wireless network, the *Quality of Service* (QoS) of the mobile/wireless network should be guaranteed. This dissertation investigated the VoIP performance in the mobile/wireless network environment. This chapter concludes our work presented in this dissertation, and briefly discusses future directions of our work.

5.1 Concluding Remarks

In Chapter 2, we conducted a modeling study to tune the IEEE 802.1X parameters to yield better performance. In the IEEE 802.1X standard, several timeout timers are defined for message exchanges in the authentication mechanism, where the same fixed value is suggested for these timeout timers. We observed that the delays for the *Extensible Authentication Protocol over LAN* (EAPOL) message exchanges may significantly vary. In this WLAN-3G integrated security approach, the access of *Home Location Register/Authentication Center* (HLR/AuC) in the 3G network may incur long delay. Therefore the setup of timeout periods is very critical for WLAN VoIP call setup. To decrease the false failure detection probability and significantly improve the expected response time of the IEEE 802.1X authentication procedure, we provided guidelines to select appropriate timeout values for IEEE 802.1X operation.

In Chapter 3, we provided guidelines to select appropriate system parameter values for VoIP in the environment (i.e., WLAN network). In 3GPP specifications, the authenticated WLAN *Mobile Station* (MS) is allowed to access the 3G network through a *WLAN Access Gateway/Packet Data Gateway* (WAG/PDG). However, to ensure telecom grade security, the VoIP traffic between the MS and the WAG/PDG must be protected with IPsec. We analyzed the performance of IPsec-based VoIP service in a IEEE 802.11b WLAN environment. Specifically, an IEEE 802.11b AP can support 15 IPsec VoIP connections with acceptable latency, small jitter, and no packet loss. We also indicated that the IPsec overhead is not serious. To maintain the same packet loss rate and jitter, the system will support one less IPsec VoIP connection than original VoIP connection.

In Chapter 4, we investigated the VoIP performance in the vehicle environment. We conducted trials in the real *Worldwide Interoperability for Microwave Access* (WiMAX) network which supports high-speed mobile broadband services and investigated the WiMAX-based VoIP of a *Mobile Taiwan* (M-Taiwan) funded program conducted during 2007-08 in the Taipei area. We observed that the performance of a VoIP service using the WiMAX-based infrastructure of the M-Taiwan Program conforms very well to the standard requirements of G.107 under the worse-condition and stringent scenario where both VoIP CPEs are wirelessly connected to the same WiMAX base station with both moving CPEs at the speeds up to 50 Km/h while both going under handovers at the same time.

5.2 Future Work

Based on the research results of this dissertation, the following issues can be further investigated:

Videophone Service: In this dissertation, we evaluated the VoIP performance according to the packet loss, latency, and jitter of the voice streams. Base on the previous IP voice studies, we will further analyze the video performance in the *IP Multimedia Core Network Subsystem* (IMS) network. ITU-T Recommendation G.1070 proposed an algorithm to estimate the videophone quality [29]. However, the proposed computation model estimates the speech quality and video quality individually, which is not practical in videophone performance measurement. We will further study the video quality affected by speech quality and vice versa to estimate the performance of the videophone service.

Call Transmission Performance: In this dissertation, we focused on the VoIP services involving the endpoints in the same wireless environment. In the further study, we will study the call transmission performance that the endpoints are located in different wireless access networks; for example, one is in the WLAN and the other is in the WiMAX network. Furthermore, we are interested in the next generation 3GPP radio access network called *Evolved UMTS Terrestrial Radio Access Network* (E-UTRAN) [30] and we will evaluate the call transmission performance of the VoIP services involving the endpoints in the next generation mobile network and the current wireless network.

Internet Call Server: In this dissertation, we discussed the telecom-grade call control and the security performance of the mobile wireless network. In the further study, we will investigate the service performance in a managed IP network, where the call application server is developed on the IBM WsT platform described in Chapter 1 and the mobile user connects to the *Chunghwa Telecom* (CHT) IMS network. We will show a telecom-grade call server implementation example using the IBM WsT

platform and evaluate the performance of the Internet call services. This study will provide a guideline for the third party service provider.

Bibliography

[1] 3GPP. 3rd Generation Partnership Project; Technical Specification Group Services and System Aspects; Feasibility study on 3GPP system to Wireless Local Area Network (WLAN) interworking. Technical Specification 3GPP TR 22.934 version 9.0.0 (2009-12), 2009.

[2] 3GPP. 3rd Generation Partnership Project; Technical Specification Group Core Network and Terminals; Technical realization of Short Message Service (SMS). Technical Specification 3GPP TR 23.040 version 9.1.0 (2009-9), 2009.

[3] 3GPP. 3rd Generation Partnership Project; Technical Specification Group Core Network and Terminals; Multimedia Messaging Service (MMS); Functional description; Stage 2. Technical Specification 3GPP TR 23.140 version 6.16.0 (2009-3), 2009.

[4] 3GPP. 3rd Generation Partnership Project; Technical Specification Group Services and System Aspects; IP Multimedia Subsystem (IMS); Stage 2. Technical Specification 3GPP TR 23.228 version 9.2.0 (2009-12), 2009.

[5] WiMAX Forum. WiMAX Forum Network Architecture (Stage 3: Detailed Protocols and Procedures), Release 1, V1.2, Jan. 2008.

[6] IBM. http://www-01.ibm.com/software/webservers/appserv/was/; WebSphere Application Server.

[7] LAN/MAN Standards Committee of the IEEE Computer Societ, IEEE Standard for Local and Metropolitan Area Networks – Port-Based Network Access Control. IEEE Std 802.1X-2001, 2001.

[8] Y.-B. Lin and I. Chlamtac, *Wireless and Mobile Network Architectures*. John Wiley & Sons, Inc., 2001.

[9] Y.-B. Lin and Y.-K. Chen, Reducing Authentication Signaling Traffic in Third Generation Mobile Network. *IEEE Transactions on Wireless Communications*, 2002.

[10] IETF. Extensible Authentication Protocol (EAP). IETF RFC 3748, 2004.

[11] IETF. Remote Authentication Dial In User Service (RADIUS). IETF RFC 2865, 2000.

[12] A. K. Salkintzis, C. Fors, and R. Pazhyannur, WLAN-GPRS Integration for Next-generation Mobile Data Networks. *IEEE Wireless Communications*, 2002.

[13] K. Ahmavaara, H. Haverinen, R. Pichna, Interworking architecture between 3GPP and WLAN systems. *IEEE Communications Magazine*, 2003.

[14] A. K. Salkintzis, Interworking techniques and Architectures for WLAN/3G Integration toward 4G Mobile Data Networks. *IEEE Wireless Communications*, 2004.

[15] Y.-B. Lin and A.-C. Pang, *Wireless and Mobile All-IP Networks*. John Wiley & Sons, Inc., 2005.

[16] IETF. Extensible Authentication Protocol Method for GSM Subscriber Identity Modules (EAP-SIM). IETF Internet Draft draft-haverinen-pppext-eap-sim-05, June 2002.

[17] 3GPP. 3rd Generation Partnership Project; Technical Specification Group Services and System Aspects; 3G Security; Security Architecture. 3G TS 33.102, v6.2.0, 2004.

[18] L. Klenrock, *Queueing Systems; Volume I: Theory*. John Wiley & Sons, Inc., 1976.

[19] J. Banks, J. S. Carson, B. L. Nelson, and D. M. Nicol. *Discrete-Event System Simulation*. Prentice Hall, 2001.

[20] *3GPP TS 23.234, 3GPP System to Wireless Local Area Network (WLAN) Interworking; System Description* (release 7), 3rd Generation Partnership Project (3GPP), 2006; http://www.3gpp.org/ftp/Specs/archive/23_series/23.234/.

[21] S. Kent and R. Atkinson, "*IP Encapsulating Security Payload (ESP)*", RFC 2406, Nov 1998.

[22] J. Rosenberg, H. Schulzrinne, G. Camarillo, A. Johnston, J. Peterson, R. Sparks, M. Handley, and E. Schooler, "*SIP: Session Initiation Protocol*", RFC 3261, Jun. 2002.

[23] S. Casner, R. Frederick, V. Jacobson, and H. Schulzrinne, "*RTP: A Transport Protocol for Real-Time Applications*", RFC 3550, Jul. 2003.

[24] A. Nascimento, A. Passito, E. Mota, E. Nascimento, and L. Carvalho, "*Can I Add a Secure VoIP Call?*", Proceedings of the XIII IEEE International Conference on Networks, Volume 1, Page 151– 155, Nov. 2005.

[25] R. Rajavelsamy, V. Jeedigunta, B. Holur, M. Choudhary, and O. Song, "*Performance Evaluation of VoIP over 3G-WLAN Interworking System*", IEEE Wireless Communications and Networking Conference, Volume 4, Page 2312– 2317, Mar. 2005.

[26] W. Wang, S.-C. Liew, and V.O.K. Li, "*Solutions to Performance Problems in VoIP over a 802.11 Wireless LAN*", IEEE Transactions on Vehicular Technology, Volume 54, Issue 1, Page 366– 384, Jan. 2005.

[27] D.P. Hole and F.A. Tobagi, "*Capacity of an IEEE 802.11b Wireless LAN Supporting VoIP*", IEEE International Conference on Communications, Volume 1, Page 196– 201, Jun. 2004.

[28] SmartVoIPQoS User Guide, http://www.spirentcom.com/documents/438.pdf

[29] ITU-T G.711 Recommendation, "*Pulse Code Modulation (PCM) of Voice Frequencies*", Nov. 1988.

[30] ITU-T G.729 Recommendation, "*Coding of Speech at 8 kbit/s Using Conjugate-Structure Algebraic-Code-Excited Linear-Prediction (CS-ACELP)*", Mar. 1996.

[31] ANSI X9.52-1998, "*Triple Data Encryption Algorithm Modes of Operation*", American National Standard Institute, 1998.

[32] FIPS 180-1, "*Secure Hash Standard*", National Institute of Standards and Technology, US Department of Commerce, Washington DC: Springer-Verlag, April 1995.

[33] NTT Communications, http://www.ntt.com/index-e.html

[34] Y.-B. Lin, "Keynote Speech on M-Taiwan: A WiMAX Experience", Mobility 2008 Conference, I-Lan, Taiwan, 9-12 September, 2008.

[35] MOEA, 2008 WiMAX Expo, Taipei & WiMAX Operator Summit, June 2-6, 2008. See also http://wimaxtaipei.tw/.

[36] IETF. "*SDP: Session Description Protocol*", RFC 4566, Jul. 2006.

[37] ITU-T. The E-model, a Computational Model for Use in Transmission Planning, ITU-T Recommendation G.107, 2003.

[38] H. F. Rashvand, Lecture Notes for ES433 & ES9R8, *Wireless Communication Systems*, University of Warwick.

[39] H. F. Rashvand, Special Guest Speaker, *WiMAX Cybercity & Next Generation Networks*, Mobility 2008, Ilan, Taiwan, 9-12 September, 2008.

[40] Y.-C. Sung and Y.-B. Lin, IPsec-based VoIP Performance in the WLAN Environment, *IEEE Internet Computing*, 2008.

[41] D. P. Hol and F. A. Tobagi, Capacity of an IEEE 802.11b Wireless LAN Supporting VoIP, *IEEE International Conference on Communications*, Volume 1, pps.196– 201, Jun. 2004.

[42] NetIQ, Performing a VoIP Assessment with Vivinet Assessor, White Paper, NetIQ, 2007.

Curriculum Vitae

Ya-Chin Sung received the B.S. and the M.S. degrees from National Chiao Tung University (NCTU), Hsinchu, Taiwan, R.O.C., in 2002 and 2003, respectively. Her current research interests include design and analysis of personal communications services networks, mobile computing and performance modeling.

Publication List

● International journal papers

1. Ya-Chin Sung and Yi-Bing Lin; "Effects of the EAPOL Timers in IEEE 802.1X Authentication". IEEE Transactions on Wireless Communications 6(6): 2276-2281, 2007

2. Ya-Chin Sung and Yi-Bing Lin; "IPsec-Based VoIP Performance in WLAN". IEEE Internet Computing 12 (6): 77-82, 2008

3. Yi-Bing Lin, Ya-Chin Sung, Habib F. Rashvand, Chia-Lung Liu and Yang-Jang Liao; "M-Taiwan Experience in VoIP-WiMAX Trial". Accepted and to appear in IET Communications

4. Ya-Chin Sung, Yi-Bing Lin, and Ren-Huang Liou, Lon-Fon Shieh; "NCTU-VT: A Freeware for Wireless VoIP Performance Measurement". Accepted and to appear in Wireless Communications and Mobile Computing

● Conference paper

1. Shiang-Ming Huang, Ya-Chin Sung, Shie-Yuan Wang, and Yi-Bing Lin; "NCTUns Simulation Tool for WiMAX Modeling". 3rd International Wireless Internet Conference (WICON), Austin, Texas, USA, October 2007