國 立 交 通 大 學

資訊工程學系

博士論文

網路電話與公眾電話/行動電話之智慧型服務

Intelligent Services for VoIP and PSTN/PLMN Networks

研究生：呂芳森

指導教授：張明峰 博士

鄭枸澮 博士

中華民國九十六年十月

網路電話與公眾電話/行動電話之智慧型服務

Intelligent Services for VoIP and PSTN/PLMN Networks

研 究 生：呂芳森　　Student： Fang-Sun Lu

指導教授：張明峰 博士　　Advisor： Dr. Ming-Feng Chang

鄭枸澺 博士　　Dr. Jeu-Yih Jeng

國 立 交 通 大 學

資 訊 工 程 學 系

博 士 論 文

A Dissertation

Submitted to Department of Computer Science

College of Computer Science

National Chiao Tung University

in partial Fulfillment of the Requirements

for the Degree of

Doctor

in

Computer Science

October 2007

Hsinchu, Taiwan, Republic of China

中華民國九十六年十月

網路電話與公眾電話/行動電話之智慧型服務

學生：呂芳森　　指導教授：張明峰 博士

鄭枸澐 博士

國立交通大學資訊工程學系

摘要

在越來越競爭的電信市場上，如何提供創新的服務以滿足客戶的需求越形重要。論文中針對智慧型服務相關的整合型預付式扣款服務、安全的行動電子付款模式、回呼叫機制與電信客戶服務帳務系統做了廣泛且深入之研究。

首先整合型預付式扣款服務探討在一個預付式帳戶下如何同時提供行動通信語音與數據服務，我們研究 CAMEL 協定提出通信及扣款程序；而如何降低扣款過程中二種服務同時被中斷的機率是複雜的問題，我們提出一演算法，建立其分析模式並完成數據的模擬以驗證結果。

當越來越多行動加值服務的推出以吸引客戶使用時，就需要一個安全的電子付款交易模式，我們探討提出基本需求及建立這個模式的架構，並因應行動終端設備性能的限制研究以低複雜度的 ID-based 驗證技術來滿足交易的安全性。

在私有電信網路的配號限制下，對外通信常造成被叫者一些識別不便，我們提出回呼叫表格處理機制來克服問題，並研究通信行為提出避免回呼叫碰撞之演算法，及建立分析模式。

最後於附錄中加入電信客戶服務帳務系統，這是電信業者必備之營業支援系統，藉此管理客戶資訊以提供客戶滿意的服務。

關鍵字：行動網路增強邏輯的訂製應用(CAMEL)、General Packet Radio Service (GPRS)、預付式扣款服務、行動應用、安全、小額付款、Bilinear Pairing、ID-based 加密法、帳務、私有電信網路、Voice over IP (VoIP)、回呼叫。

Intelligent Services for VoIP and PSTN/PLMN Networks

Student: Fang-Sun Lu Advisor: Dr. Ming-Feng Chang

Dr. Jeu-Yih Jeng

Department of Computer Science

National Chiao-Tung University

Abstract

In today's highly competitive telecommunications environment, the emphasis has shifted to delivering innovative service while to satisfy increasingly sophisticated customers' need. This dissertation first introduces an Integrated Mobile Prepaid Services. Prepaid personal communication service (PCS) users have outnumbered postpaid users. We study the charging issues of an integrated GSM and GPRS prepaid service, where a single prepaid account provides a user both voice and data services. The call setup and charging procedures for GSM and GPRS are presented using the CAMEL network architecture. To reduce the probability of terminating both on-going voice and data calls, we suggest that no new call be admitted when the user credit is below a threshold. An analytic model has been developed to evaluate the performance of the approach. Computer simulations have also been used to verify the results. The numeric results indicate that the forced termination probability can be significantly reduced by choosing an appropriate threshold of the user credit.

When the basic functionalities of a wireless mobile network have been achieved, customers are then more interested in value-added mobile applications. In order to attract more customers to such mobile applications, a solid, secure and robust trading model is a

must. We propose such a secure trading model named Mobile Electronic Payment (MEP) for wireless mobile networks, which applies the emerging ID-based cryptography for key agreement and authentication. Our MEP attempts to alleviate the computational cost, reduce the memory space requirement in mobile devices, and meet the requirements for secure trading: avoidance of overspending and double spending, fairness, user anonymity and privacy. Our design is transparent to the bearer networks and is of low deployment cost. We expect that our MEP provides a viable trading architecture model for the future mobile applications.

Private Telecommunications Networks (PTNs), such as enterprise telephony and Voice over IP (VoIP), have been widely deployed in the recent years. A common limitation of PTNs is that the PTN users are usually not assigned with E.164 numbers that are routable in Public Switched Telephone Network (PSTN). When a PTN user initiates a call to a PSTN user, it is not possible to use the PTN calling party's identifier as the caller ID; instead, an E.164 number of the PBX or VoIP gateway is used. Therefore, the PSTN called party cannot use the caller ID to call back to the PTN calling party directly. We propose a callback table approach to resolve this issue. We describe how the PTN call-out information is stored into and retrieved from the callback table in the PBXs or VoIP gateways to provide callback service. Numeric analysis has been performed to evaluate the feasibility of this approach. The numeric results indicate that the callback method performs better as there is more voice traffic served by a VoIP gateway.

These research results presented in this dissertation can be viewed as a useful foundation for further study in intelligent services for VoIP and PSTN/PLMN networks.

In appendix we introduce customer care and billing system for telecommunication, which is of the Business Support System (BSS). We present how customer service and billing system are managed in the mobile market.

Key Words: Customized Applications for Mobile network Enhanced Logic (CAMEL), General Packet Radio Service (GPRS), Prepaid services, Mobile Application, Security, Micropayment, Bilinear Pairing, Identity-Based Cryptography,
Billing, Private Telecommunications Network (PTN), Voice over IP (VoIP), Callback.

Acknowledgement

I would like to express my sincere thanks to my advisors, Prof. Ming-Feng Chang and Dr. Jeu-Yih Jeng. Without their supervision and perspicacious advice, I can not complete this dissertation. Special thanks to my committee members, Prof. Chung-Ta King, Prof. Yi-Bing Lin, Dr. Sheng-Lin Chou, Prof. Wen-Nung Tsai, Dr. Kuang-Yao Chang, and Prof. Ai-Chun Pang for their valuable comments. Thanks also to the colleagues in Internet Communication Laboratory.

I also express my appreciation to all the faculty, staff and colleagues in the Department of Computer Science and Information Engineering, NCTU and Customer Service Systems Laboratory, Telecommunication Laboratories of CHT. In particular, I would like to thank Prof. Phone Lin, Prof. Wei-Zu Yang, Dr. Miin-Luen Day, Dr. Yuan-Kai Chen, Dr. Meng-Ta Hsu, Mr. Chung-Yung Chia, Mr. Moen Song and Mr. Ming-Shew Wu for their friendship and support in various ways.

Finally, I am grateful to my family, my wife, Hui-Ling, my children, and friends for their encouragement and support during these years.

Contents

List of Figures

List of Tables

Abbreviation

The abbreviations used in this dissertation are listed below.

2G: The Second Generation Mobile System
3G: The Third Generation Mobile System
3GPP: 3rd Generation Partnership Project
APN: Access Point Name
ARPU: Average Revenue Per User
BSS: Business Support System
CAMEL: Customized Applications for Mobile network Enhanced Logic
DID: Direct Inward Dialing
PCS: Personal Communication Service
GGSN: Gateway GPRS Support Node
GPRS: General Packet Radio Service
GSM: Global System for Mobile Communication
IAM: Initial Address Message
IN: Intelligent Network
IVR: Interactive Voice Response
LEC: Llocal Exchange Carrier
MEP: Mobile Electronic Payment
MGCP: Media Gateway Control Protocol
MS: Mobile Station
NGN: Next Generation Network
OCS: Online Charging System
PBX: Private Branch Exchange
PLMN: Public Land Mobile Network
PSTN: Public Switched Telephone Network
PTN: Private Telecommunications Network
SBCF: Session Based Charging Function
SCP: Service Control Point
SGSN: Serving GPRS Support Node
SIP: Session Initiation Protocol
SS7: Signaling System Number 7
SSF: Service Switching Function
UMTS: Universal Mobile Telecommunications System
URI: Universal Resource Identifier

VoIP: Voice over IP
VPG: VoIP Gateway
WLAN: Wireless Local Area Network

CHAPTER 1 Introduction

In recent years, telecommunication industry is growing fast especially in mobile market. Many technologies have been developed and deployed, such as 2G/GPRS/3G, VoIP, and NGN (Next Generation Network). They provide not only the traditional voice communication service but also many advanced data and information services. However, technical advances no longer drive the market trend. In today's highly competitive environment, the emphasis has shifted to delivering innovative service to satisfy increasingly sophisticated customers' need. Customers have changed from being the passive role to active, and operators need to focus on customers' feeling. The traditional belief of targeting at the general public has been rapidly modified to become targeting at differentiated products and services. It is important to satisfy all kinds of customers and make them feel that the service is tailored for them, for their benefits and interests. No matter who they are, themselves are actually aware of being the most important and can get the simplest, the most direct and personalized services. The customers' real requests will dominate the new business model, depicted in Fig. 1.1. As the ARPU (Average Revenue Per User) falls off with increased telecom penetration, operators must offer more value-added services to stimulate the user demand. Further more, operators need effective value-added service suppliers for their entire customer base. Nowadays' innovative services are deployed not just for the simple relation between operators and customers. Operators and service/content providers need to work together to make a win-win new telecom business market.

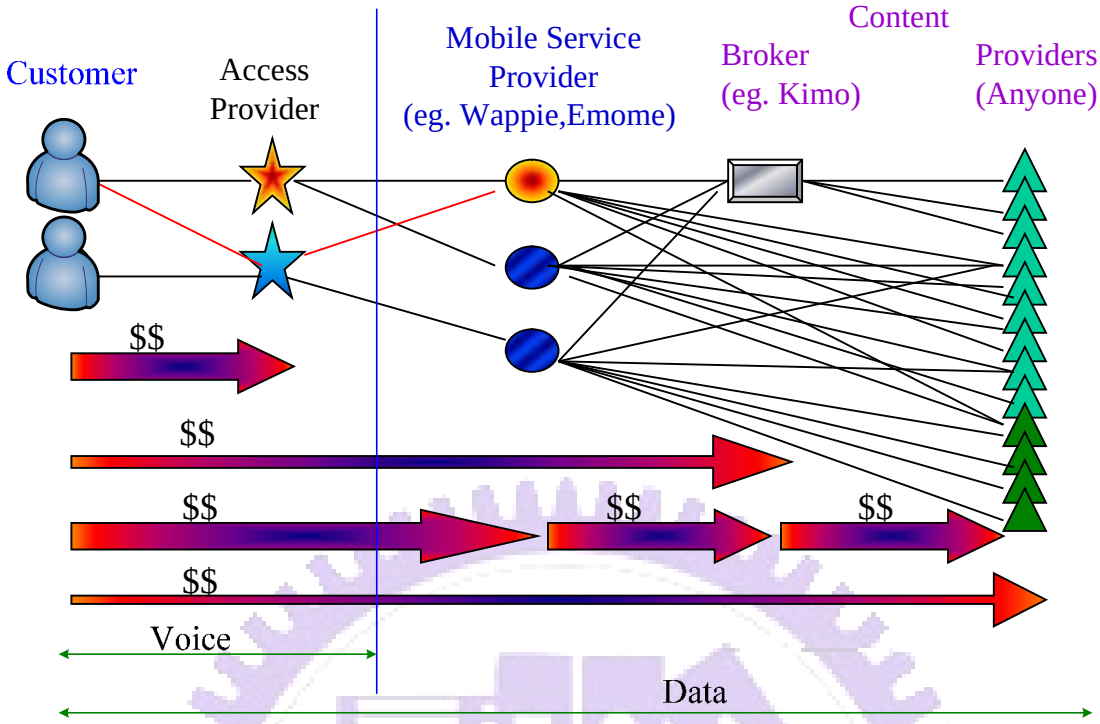


Fig. 1.1: New Telecom Business Model

1.1 An Integrated Mobile Prepaid Services

Due to the advances in wireless communication technologies, the personal communication service (PCS) market has grown exponentially over the past ten years. Prepaid service is a telecommunication service that requires subscribers to pay before they make calls. This service was offered in Europe and Asia in 1982 and became popular in U.S. in 1992 [5]. In 1997, there were about 60 million GSM subscribers across the world and 8% of them subscribed to prepaid service. At the end of Q3 in 2001, 50% of the subscribers world-wide were prepaid users [22]. In USA, the prepaid calling market grew 56% to about two billion US dollars in 1998 and expected to maintain a high growth rate to 2005 [51]. Asia countries such as Philippine, Australia, Hong Kong, Singapore and Taiwan have already shown successful examples for prepaid services.

Four billing technologies have been used in mobile prepaid service: hot billing approach, service node approach, intelligent network approach and handset-based approach. We have

studied performance of the hot billing approach in [16] and the service node approach in [17]. However, these studies focused only on voice services. As the ARPU falls off, operators must offer more value-added services to stimulate the user demand. By providing the data services and multimedia services, the ARPU is expected to increase [20].

GPRS (General Packet Radio Service) reuses the GSM infrastructure and provides ”always on” connection, flexible radio resource allocation and fast data transmission to the customers [52]. Although GPRS is just being launched, it is estimated to attract 110M GPRS users across western Europe by 2006, which may generate €23 billion in mobile data revenue [21]. By providing the integrated prepaid service that includes the voice service and data service in the GSM and GPRS network, the operator can offset the decline of the voice revenue.

In GSM, the billing records are generated by the MSC when a voice call terminates. The information in the record includes type of service, date/time of usage, user identification and location information [31]. In GPRS, the charging information of data service is collected by the SGSN (Serving GPRS Support Node) and GGSN (Gateway GPRS Support Node). The charging information collected by the SGSN includes the charging ID, APN (Access Point Name) ID, location of the mobile station and the amount of data transmitted through the radio interface with QoS profiles [2], [39]. The GGSN collects the information of external data network usage including the source address and the destination address of the packets and the amount of transmitted packets to the external network. CAMEL (Customized Applications for Mobile network Enhanced Logic) integrates the intelligent network (IN) techniques into the mobile telecommunication network, enabling the operator to provide service to its subscribers inside and outside the home network. In Phase 3, CAMEL specified the capability to control of GPRS sessions and PDP contexts [40]. Very few studies have been done on the analysis of the prepaid services with concurrent sessions. We have included two related papers recently published. Cai, et. al., presented secure authorization mechanisms for prepaid services [12].

Lin, et. al., investigated the credit allocation problem of concurrent service sessions [36]. In this topic, we studied how to provide GSM and GPRS prepaid services under the CAMEL architecture. We design CAMEL message flows for GPRS prepaid service based on the attach/detach state model and PDP state control mode. In addition, we study the performance of GSM/GPRS network where the charges (i.e., voice charge and data charge) of a subscriber are debited in one account. The prepaid credit may deplete either when the subscriber is in a voice conversation, a data transport session, or both. It is preferable that the billing system should avoid terminating both the ongoing voice call and data sessions simultaneously when the prepaid credit runs out. We proposed a simple algorithm to reduce this probability, and an analytic model to evaluate the improvement.

1.2 A Secure Mobile Electronic Payment Architecture Platform

With the vast development and deployment of wireless mobile networks such as 3G UMTS [26,39], WiMAX [30] and Wi-Fi [29], mobile networking applications enabling customers to gain network access anywhere and anytime have attracted more and more attention in our daily lives. When the basic functionalities of a wireless network have been in place, customers are now more interested in value-added mobile applications over this network. Most mobile applications come with the emergence of electronic trading (mobile commerce or m-commerce), hence good secure mobile trading model must be designed to attract more mobile users for doing business wirelessly. Thus, how to integrate the mobile applications with a secure trading model becomes an important design issue, which will significantly affect the success of any value-added mobile application.

Mobile applications can be categorized into session-based applications and event-based applications. In event-based applications, user's payment is reected by one-time events. Examples include sending a message, querying traffic information, or purchasing a song. A

session-based application consists of three phases: the session-setup phase, the communication phase and the session release phase. A customer is charged for a session-based application based on either time spent or data volume transferred, e.g., VoIP-calling, video-streaming, audio-streaming, or video-conferencing.

There are a few payment models proposed in the literature [6,35], which can be classified into two categories: the traditional payment model and the micropayment model. The examples of traditional payment models include the credit card platforms [9,3,42,41] and the electronic cash platforms [10,43,13]. The traditional payment models allow only one payment in a payment transaction, which has been widely adopted for the event-based applications. Since a session-based application usually requires multiple payments during the execution of this application, with the traditional payment model, it requires multiple payment transactions to complete a session-based application. This is inefficient because heavy signaling and computational overheads are introduced into the network. On the other hand, the micropayment models allow multiple payments in a payment transaction, which is considered more efficient than the traditional payment model. Thus, the micropayment models [55,27,54,46] are often adopted for most of mobile applications. To secure transactions, in [55,46], the public-key cryptography (e.g., RSA [33]) is adopted. Unfortunately, the public-key cryptography requires heavy computation and long execution time, which may not be a good solution in wireless mobile networks. Yang et al. [54] applied the symmetric-key cryptography such as Advanced Encryption Standard (AES) [19] that is more efficient than the public-key cryptography in terms of computational cost and is more suitable for mobile devices. Unfortunately, the symmetric-key cryptography requires more frequent key establishments and updates to prevent the shared key from being compromised, and hence induces more communication cost due to key establishment and key updates. Moreover, how to establish the shared key in wireless mobile networks for the symmetric-key cryptography is very challenging.

Compared with fixed networks, mobile networks have lower bandwidth, longer transmission latency, and more unreliable connections, and mobile devices are restricted by limited memory size and low CPU computational capability [34]. The installation of mobile applications on a mobile network should be quick and of low cost. To summarize, the following requirements should be addressed when designing a suitable trading mechanism on a mobile network. First, customers expect a robust, secure, and fair trading mechanism which can be applied in different mobile networks. Second, the trading mechanism should be light-weight (i.e., with low computational complexity and low communication overhead) so that it can be easily run on mobile devices. Third, user anonymity should be achieved, that is, users' purchasing behavior or preference should not be traceable by others. Finally, a trading mechanism should be of low implementation cost.

In this topic, we design an application-level secure payment model, named Mobile Electronic Payment (MEP), for wireless mobile networks, which attempts to meet these requirements. It is based on a more general trading architecture model, which combines both public-key cryptography and symmetric-key cryptography to overcome the disadvantages of both technologies. Specifically, we apply the emerging ID-Based Cryptography (IBC [11]; cf. Section 3.2.2) in the MEP to generate the public-private key pairs so that the certificate overheads among the network operator (denoted as **O**), the user (denoted as **U**), and the mobile application developer or content provider (denoted as **P**) commonly required in the traditional public-key cryptography can be eliminated. Then, from these public-private key pairs, we generate three symmetric keys $k_{u\text{-}o}$ (held by **O** and **U**), $k_{o\text{-}p}$ (held by **O** and **P**), and $k_{u\text{-}p}$ (held by **U** and **P**) to encrypt and decrypt the signaling messages exchanged among **O**, **U**, and **P**. An important observation is that these three symmetric keys are established without actually exchanging them among the concerned parties, a unique feature of ID-based cryptography. To prevent the symmetric keys from being compromised, in each payment transaction, the three public-private key pairs ($k_{pub,o}$, $k_{pri,o}$) held by **O**, ($k_{pub,p}$, $k_{pri,p}$) held by **P**,

and ($k_{pub,u}$, $k_{pri,u}$) held by **U** are used to generate the new symmetric keys.

Our design keeps the key freshness and thus provides more robust security protection. Moreover, MEP supports both event-based and session-based applications and is suitable for the resource-constrained mobile devices because MEP attempts to alleviate the computational cost and reduce the memory space requirement in mobile devices. We expect that our MEP provides a viable trading model for the future mobile applications.

1.3 A Callback Mechanism for Private Telecommunications Network

A Private Telecommunications Network (PTN) is a telecommunications network that has its own number plan other than the public E.164 numbering used in the PSTN [14]. Examples of PTNs are enterprise telephone systems and Voice over IP (VoIP) networks without being assigned PSTN E.164 numbers. User of enterprise telephone systems are identified by extension numbers. VoIP users are often identified by URIs (universal resource Identifiers), such as sip:tom@nctu.edu.tw, not by public telephone numbers. PTNs have been widely deployed in companies and the Internet/Intranet. Fig. 1.2 illustrates an abstract PTN architecture that interconnects to the PSTN. The core component in this architecture is the Private Branch Exchange (PBX) in a telephony-based PTN or the VoIP Gateway (VPG) in an Internet-based PTN. Without loss of generality, we consider the Internet-based PTNs. Example Internet-based PTNs include H.323/SIP telephony networks [1, 32, 45], MGCP/MEGACO-based networks [4, 15], and Skype [48]. A phone call between a VoIP user and a PSTN user consists of two connections: a VoIP connection between the VoIP user and the VPG, and a PSTN connection between the VPG and the PSTN user. This type of calls will be referred to as IP-PSTN calls. At present, a VoIP connection is free of charge or much cheaper than a PSTN connection. On the other hand, the cost of a PSTN connection depends on its distance. For example, an international/long-distance call costs more than a local one.

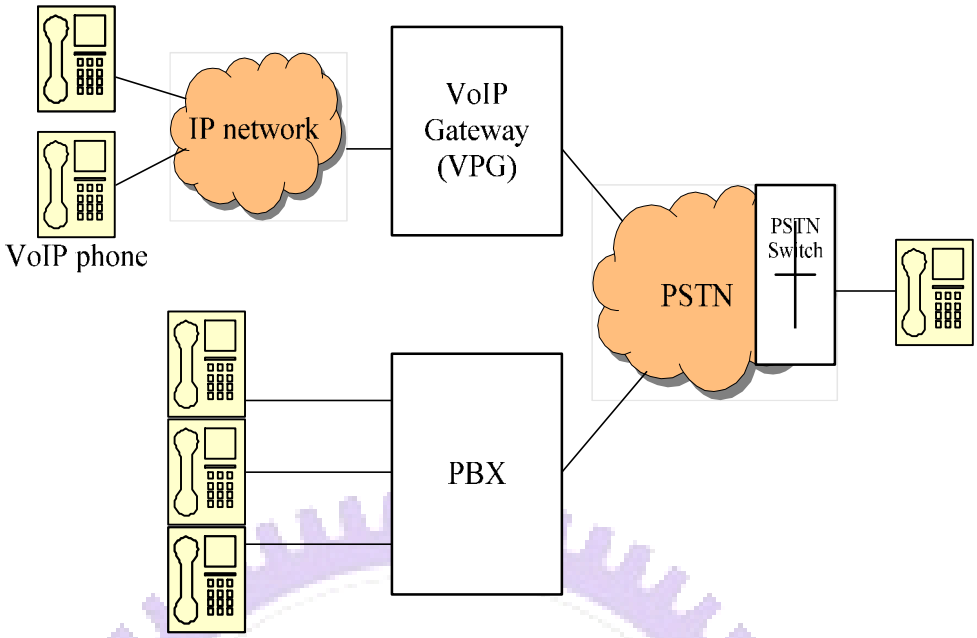


Fig. 1.2: Private Telecommunications Network Architecture

A VoIP service provider may have a large number of VPGs, each of which connects to a tariff zone of a local exchange carrier (LEC). Fig. 1.3 depicts an example of three VPGs located in three different tariff zones of Chunghwa Telecom (CHT), Taiwan. To minimize the cost of an IP-PSTN call, the PSTN user should be connected to the VPG at the same tariff zone, i.e. the shortest PSTN connection is chosen. In this way, when serving IP-PSTN calls, the PSTN is partitioned into small tariff zones, each of which is served by a VPG. The number of leased lines of each VPG can be determined by the voice traffic served by the VPG. A general rule is to keep the call blocking probability at busy hours below 1%.

A common limitation of VoIP is that the VoIP users, such Skype users, are usually not assigned with E.164 numbers. When a PTN user initiates a call to a PSTN user, it is not possible to use the PTN calling party's identifier as the caller ID; instead, an E.164 number of the serving VPG is used. Therefore, the PSTN called party cannot use the caller ID to call back to the PTN calling party later. As a result, although it is easy for a Skype user to initiate a call to a PSTN user, it is not possible for a PSTN user to call a Skype user if the Skype user

does not subscribe SkypeIn service (i.e., the user is not assigned an E.164 number). In this topic, we propose a callback table approach to resolve this problem.

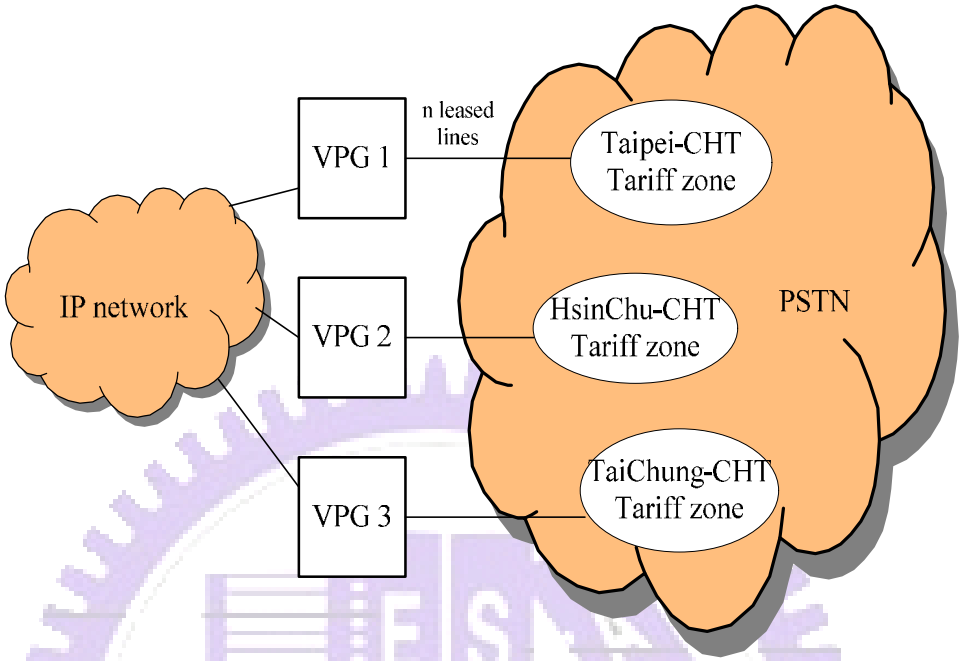


Fig. 1.3: The PSTN is partitioned by the VPGs

1.4 Synopsis of This Dissertation

This dissertation is organized as follows. Chapter 2 presents An Integrated Mobile Prepaid Services. Chapter 3 presents A Secure Mobile Electronic Payment Architecture Platform for Wireless Mobile Networks. In Chapter 4, we describe A Callback Mechanism for Private Telecommunications Network. Chapter 5 concludes this dissertation and describes the future work. Appendix A describes Customer Care and Billing System for Telecommunication.

CHAPTER 2 An Integrated Mobile Prepaid Services

2.1 Introduction

Due to the advances in wireless communication technologies, the personal communication service (PCS) market has grown exponentially over the past ten years. Prepaid service is a telecommunication service that requires subscribers to pay before they make calls. Four billing technologies have been used in mobile prepaid service: hot billing approach, service node approach, intelligent network approach and handset-based approach. However, these studies focused only on voice services. As the ARPU falls off, operators must offer more value-added services to stimulate the user demand. GPRS (General Packet Radio Service) reuses the GSM infrastructure and provides ”always on” connection, flexible radio resource allocation and fast data transmission to the customers [52]. By providing the integrated prepaid service that includes the voice service and data service in the GSM and GPRS network, the operator can offset the decline of the voice revenue.

CAMEL (Customized Applications for Mobile network Enhanced Logic) integrates the intelligent network (IN) techniques into the mobile telecommunication network, enabling the operator to provide service to its subscribers inside and outside the home network. In Phase 3, CAMEL specified the capability to control of GPRS sessions and PDP contexts [40]. Very few studies have been done on the analysis of the prepaid services with concurrent sessions. We studied how to provide GSM and GPRS prepaid services under the CAMEL architecture. We design CAMEL message flows for GPRS prepaid service based on the attach/detach state model and PDP state control mode. In addition, we study the performance of GSM/GPRS network where the charges (i.e., voice charge and data charge) of a subscriber are debited in one account. The prepaid credit may deplete either when the subscriber is in a voice conversation, a data transport session, or both. It is preferable that the billing system should

avoid terminating both the ongoing voice call and data sessions simultaneously when the prepaid credit runs out. We proposed a simple algorithm to reduce this probability, and an analytic model to evaluate the improvement.

2.2 Prepaid Services in CAMEL

Fig. 2.1. depicts the CAMEL network architecture for the integrated GSM and GPRS prepaid service. In this architecture, we add three components in the GSM/GPRS network for prepaid service: P-SCP (prepaid service control point), gsmSSF (GSM Service Switching Function) and gprsSSF (GPRS service switching function). The SGSN (MSC) interfaces with the PSCP via the gprsSSF (gsmSSF) for CAMEL session control. The P-SCP integrates the billing of GSM and GPRS services in one account. When a subscriber originates a voice call, the gsmSSF issues a CAP (CAMEL Application Part) message, ”InitialDP” (Trigger Detection Point Request), to the P-SCP. The P-SCP checks the account status and starts a count-down timer if the call is allowed. When a subscriber originates a data session (activate a PDP context), the gprsSSF issues a CAP message ”InitialDPGPRS”, (Trigger Detection Point Request), to the P-SCP. The P-SCP checks the account status and gives predefined units of quota to the gprsSSF for the charging of the data session.

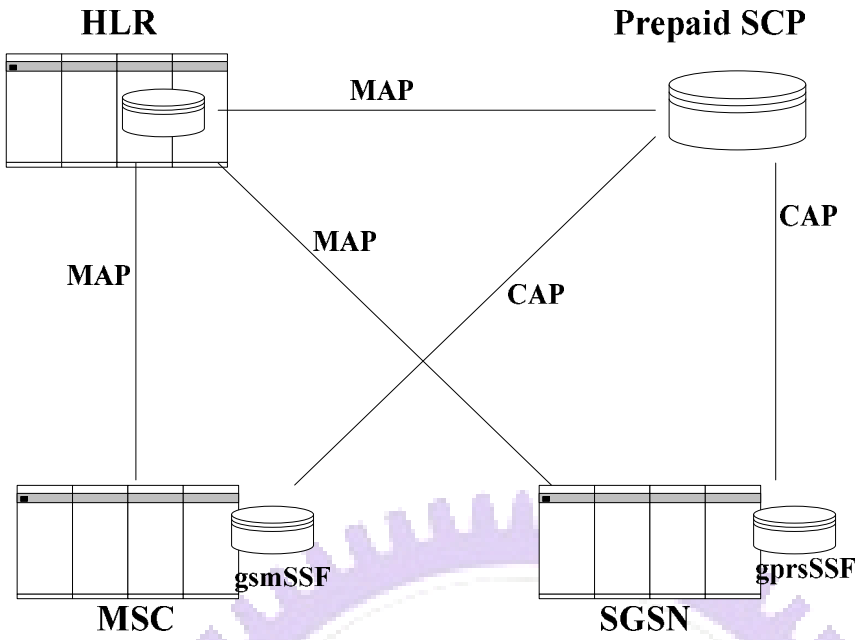


Fig. 2.1: CAMEL architecture for integrated GSM/GPRS prepaid services

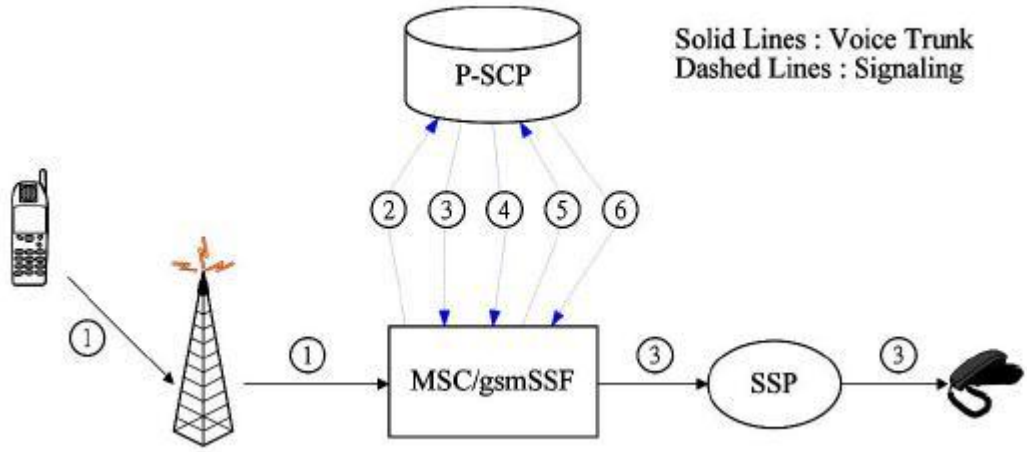


Fig. 2.2: IN prepaid call origination procedure

Fig. 2.2. illustrates the voice call origination procedure, which comprises the following steps.

Step1) The prepaid customer initiates a call by dialing the called party's telephone number.

Step2) The MSC/gsmSSF encounters an IN call setup trigger. The call setup process is

suspended, and a CAP message,”InitialDP”, is sent to the P-SCP. The message includes the MSISDN, location information of the MS, and the called party telephone number. The P-SCP determines whether or not the customer can make the call by querying its database. Based on threshold processing parameters defined in the prepaid billing system, the P-SCP may deny or accept the call. (In this example, assume that the call is accepted.)

Step3) The P-SCP asks the MSC to resume the call setup procedure with a CAP message,”ApplyCharging”, and the call is eventually connected. The P-SCP starts a countdown timer. The rate of credit decrement (from the current balance) is derived from carrier-defined threshold parameters, the rate plan, the destination, and time/date dependency.

Step4) The call terminates when either the balance depletes or the call completes. If the countdown timer ends before the customer terminates the call, the P-SCP instructs the MSC with message,”ReleaseCall”, to terminate the call. For normal call completion, this step does not exist.

Step5) Once the call is completed, the MSC/gsmSSF encounters an IN call-release trigger, which sends a disconnect message,”EventReportBCSM”, to P-SCP indicating the completion time of the call.

Step6) The P-SCP rates the completed call and updates the customer’s prepaid balance accordingly. Then it sends the current balance and cost of the call to the MSC. The MSC release the call.

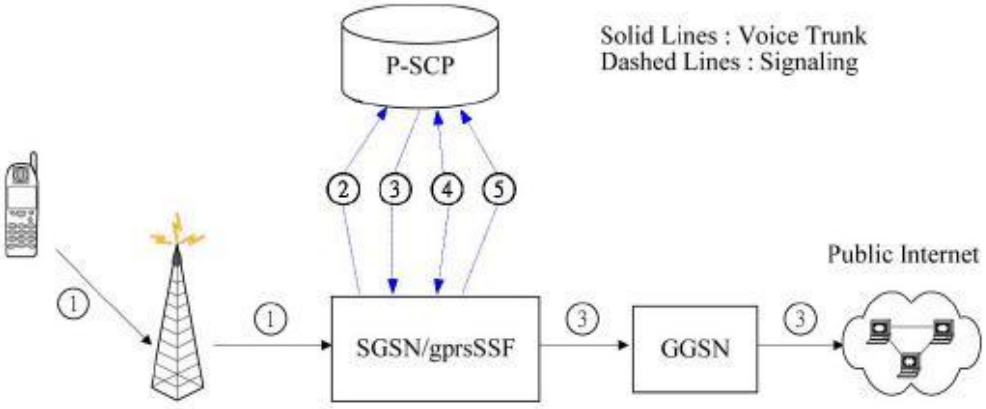


Fig. 2.3: Basics of prepaid GPRS call procedure

According to the popular view on GPRS, data volume is the natural measure for GPRS resource usage and should be the crucial parameter to drive the charge. This proposed charging principle reflects the GPRS vision that the user may stay always on-line, and may send/receive data as the need arises. Fig. 2.3. illustrates the data call procedure and charging by data volume when PDP context activated, which comprises the following steps.

Step1) A mobile station performs a GPRS PDP context activation.

Step2) The SGSN/gprsSSF sends an ”InitialDPGPRS” message to the P-SCP, containing a service indicator as a key to GPRS prepaid service. The P-SCP invokes the operator-defined GPRS prepaid data service. Tariff and authorization features are applied to the request for service. The P-SCP sends the result with predefined units of quota in an ”ApplyChargingGPRS” message to the gprsSSF.

Step3) The SGSN receives the response that either authorizes or rejects the call. If authorized, the SGSN sends a ”Create PDP Context Request” message to the GGSN, which contains the MS identifiers, location, QoS(quality of service) and APN. The session of data access is continued.

Step4) The SGSN also starts to decrement the quota of accessible data volume. When the quota is depleted, the SGSN suspends the session (data queued in buffer) and

sends an ”ApplyChargingReportGPRS” message to the P-SCP asking additional quota. The P-SCP starts a new check to user account and sends the result with predefined units of quota in an ”ApplyChargingGPRS” message to the gprsSSF again. The SGSN receives the response, and let the session continued. Checking quota process is repeated during the session.

Step5) On call termination, the SGSN sends an ”ApplyChargingReportGPRS” message to the P-SCP indicating the charge of the service. The SGSN sends a ”PDP Context Disconnection” message to the mobile station.

Step6) The account status notification feature can be enabled at the P-SCP. The P-SCP sends a status notification message to the SMS-SC, which is forwarded to the prepaid subscriber.

2.3 The Analytic Model

In this section, we present an analytic model for GSM and GPRS prepaid services. For a GSM/GPRS network providing prepaid services, there are two types of calls, namely 1) voice calls and 2) data calls. The voice calls and data calls that originates from a MS are assumed to be Poisson with rates λ_1 and λ_2, respectively. The call holding time of voice calls and data calls are assumed to be exponentially distributed with mean $1/\mu_1$ and $1/\mu_2$, respectively.

Let B be the prepaid credit. We assume that the prepaid credit of voice service and data service shares the same account and is decremented by the P-SCP in a real-time fashion. When the credit of the subscriber depletes, the P-SCP terminates the on-going calls. Let $P_{f,v}$, $P_{f,d}$ and $P_{f,vd}$ be the probabilities that a voice call, data call, and both types of calls are forced to terminate when the credit depletes, respectively. Note that

$$P_{f,v} + P_{f,d} + P_{f,d} = 1 \qquad (2.1)$$

Our previous study on prepaid voice calls shows that the decrement of a subscriber’s

credit by the successive calls of the subscriber can be modeled as a renewal process, since the charge of each call is independent and has an identical distribution. Our study also indicated that the time when the prepaid credit depletes can be treated as a random observer of a renewal process if the initial credit is much larger than the charge of one call [4]. Hence, $P_{f,v}$, $P_{f,d}$ and $P_{f,vd}$ can be derived by using the theory of Markov chain. In this process, state 0 represents that the subscriber has no on-going call. State 1 (state 2) represents that the subscriber is going a voice call (data call). State 3 represents that the subscriber is going a voice call and a data call simultaneously. Fig. 2.4. illustrates the state transition diagram of the prepaid GSM/GPRS services.

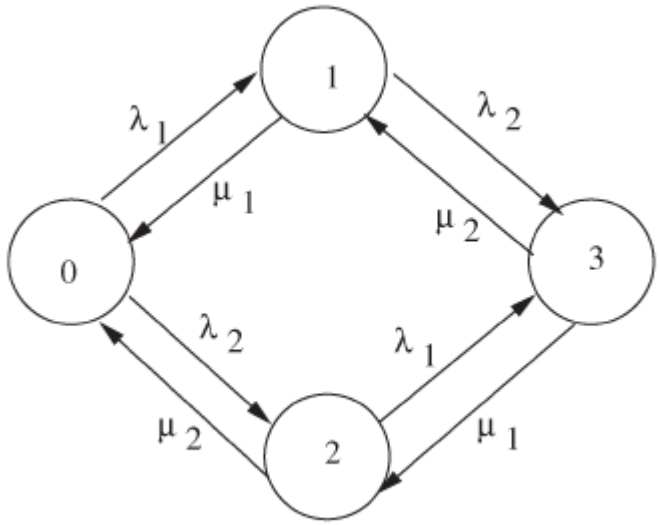


Fig. 2.4: State transition diagram of prepaid GSM/GPRS services

Let π_i be the steady state probability that the system is at state i. Let r_1 be the charge rate of voice calls, and r_2 be that of data calls. If the initial credit is sufficiently large, the probability that the credit depletes at state i is proportional to the fraction of credit depleted at state i, i.e., $P_{f,v} : P_{f,d} : P_{f,vd} = \pi_1 r_1 : \pi_2 r_2 : \pi_3 (r_1 + r_2)$. From Fig. 2.4., we have the balance equations as follow.

$$\begin{aligned}
(l_1 + l_2)p_0 &= m_1 p_1 + m_2 p_2 \\
(l_2 + m_1)p_1 &= l_1 p_0 + m_2 p_3 \\
(l_1 + m_2)p_2 &= l_2 p_0 + m_1 p_3 \\
(m_1 + m_2)p_3 &= l_2 p_1 + l_1 p_2
\end{aligned} \tag{2.2}$$

Note that

$$\sum_{i=0}^{3} p_i = 1 \tag{2.3}$$

From Eqs. (2.2) and (2.3), p_0, p_1, p_2, and p_3 can be expressed as follows.

$$\begin{aligned}
p_0 &= \frac{m_1 m_2}{l_1 m_2 + l_2 m_1 + l_1 l_2 + m_1 m_2} \\
p_1 &= \frac{l_1 m_2}{l_1 m_2 + l_2 m_1 + l_1 l_2 + m_1 m_2} \\
p_2 &= \frac{l_2 m_1}{l_1 m_2 + l_2 m_1 + l_1 l_2 + m_1 m_2} \\
p_3 &= \frac{l_1 l_2}{l_1 m_2 + l_2 m_1 + l_1 l_2 + m_1 m_2}
\end{aligned} \tag{2.4}$$

From Eqs. (2.1)-(2.4), we have

$$\begin{aligned}
P_{f,v} &= \frac{r_1 l_1 m_2}{r_1 l_1 m_2 + r_2 l_2 m_1 + (r_1 + r_2) l_1 l_2} \\
P_{f,d} &= \frac{r_2 l_2 m_1}{r_1 l_1 m_2 + r_2 l_2 m_1 + (r_1 + r_2) l_1 l_2} \\
P_{f,vd} &= \frac{(r_1 + r_2) l_1 l_2}{r_1 l_1 m_2 + r_2 l_2 m_1 + (r_1 + r_2) l_1 l_2}
\end{aligned} \tag{2.5}$$

For prepaid GSM/GPRS services, a subscriber would be annoyed if both on-going voice call and data call are forced to terminate simultaneously when the credit depletes. To enhance the user's satisfaction, we propose an algorithm, namely, no-more-call (NMC), to reduce the forced termination probability. It is described as follows.

Algorithm NMC: When the user credit is less than *b*, no more new call, except the ongoing calls, can be served.

Note that *b* can be sufficiently small such that the remaining credit accommodates only the on-going calls.

2.3.1 Analytic Model for NMC

In this section, we present an analytic model to derive the forced-termination probability of NMC algorithm. The credit depletes in one of the following cases:

Case I) The subscriber is only in a voice conversation when $B = b$.

Case Ia) The voice call depletes the rest of credit and is forced to terminate.

Case Ib) The voice call completes, and there is remaining credit.

Case II) The subscriber is only using the data service when $B = b$.

Case IIa) The data call depletes the rest of credit, and is forced to terminate.

Case IIb) The data call completes, and there is remaining credit.

Case III) The subscriber has both an on-going voice call and an on-going data call when $B = b$.

Case IIIa) Both calls deplete the rest of the credit, and are forced to terminate.

Case IIIb) The data call terminates before the credit is depleted. The voice call depletes the remaining credit and is forced to terminate.

Case IIIc) The voice call terminates before the credit is depleted. The data call depletes the remaining credit and is forced to terminate.

Case IIId) Both calls complete, and there is remaining credit.

Let $P_{f,v}^{*}$ be the probability that a voice call is forced to terminate in NMC. If the subscriber's initial credit is sufficiently larger than b, Prob(Case I)= $P_{f,v}$, Prob(Case II)= $P_{f,d}$ and Prob(Case III)= $P_{f,vd}$. Let x and y be the amount of credit depleted by the voice call and data call after the remaining credit equals to b, but the calls are allowed to complete normally. Then, $P_{f,v}^{*}$ can be expressed as equation 2.6.

$$
\begin{aligned}
P^*_{f,v} &= \Pr ob(CaseI) \times \Pr ob(CaseIa \mid CaseI) \\
&+ \Pr ob(CaseIII) \times \Pr ob(CaseIIIb \mid CaseIII) \\
&= P_{f,v} \int_{x=b/r_1}^{\infty} u_1 e^{-m_1 x} dx + P_{f,vd} \int_{y=0}^{\frac{b}{r_1+r_2}} \int_{x=\frac{b-r_2 y}{r_1}}^{\infty} m_1 e^{-m_1 x} m_2 e^{-m_2 y} dxdy \\
&= \frac{r_1 l_1 m_2 e^{-\left(\frac{b}{r_1}\right) m_1}}{r_1 l_1 m_2 + r_2 l_2 m_1 + (r_1 + r_2) l_1 l_2} \\
&+ \left[\frac{(r_1 + r_2) l_1 l_2 e^{-\left(\frac{b}{r_1}\right) m_1}}{r_1 l_1 m_2 + r_2 l_2 m_1 + (r_1 + r_2) l_1 l_2} \right] \\
&\times \left[\frac{m_2}{m_2 - \left(\frac{r_2}{r_1}\right) m_1} \right] \left\{ 1 - e^{-\left(\frac{b}{r_1+r_2}\right)\left[m_2 - \left(\frac{r_2}{r_1}\right) m_1 \right]} \right\}
\end{aligned}
\tag{2.6}
$$

Let $P^*_{f,d}$ be the probability that only a data call is forced to terminate in NMC and $P^*_{f,vd}$ be the probability that both the voice and data calls are forced to terminate in NMC. Using approaches similar to those previously mentioned, $P^*_{f,d}$ and $P^*_{f,vd}$ can be expressed as

$$
\begin{aligned}
P^*_{f,d} &= \Pr ob(CaseII) \times \Pr ob(CaseIIa \mid CaseII) \\
&+ \Pr ob(CaseIII) \times \Pr ob(CaseIIIc \mid CaseIII) \\
&= P_{f,d} \int_{y=b/r_2}^{\infty} u_2 e^{-m_2 y} dy + P_{f,vd} \int_{x=0}^{\frac{b}{r_1+r_2}} \int_{y=\frac{b-r_1 x}{r_2}}^{\infty} m_1 e^{-m_1 x} m_2 e^{-m_2 y} dydx \\
&= \frac{r_2 l_2 m_1 e^{-\left(\frac{b}{r_2}\right) m_2}}{r_1 l_1 m_2 + r_2 l_2 m_1 + (r_1 + r_2) l_1 l_2} \\
&+ \left[\frac{(r_1 + r_2) l_1 l_2 e^{-\left(\frac{b}{r_2}\right) m_2}}{r_1 l_1 m_2 + r_2 l_2 m_1 + (r_1 + r_2) l_1 l_2} \right] \\
&\times \left[\frac{m_1}{m_1 - \left(\frac{r_1}{r_2}\right) m_2} \right] \left\{ 1 - e^{-\left(\frac{b}{r_1+r_2}\right)\left[m_1 - \left(\frac{r_1}{r_2}\right) m_2 \right]} \right\}
\end{aligned}
\tag{2.7}
$$

$$P_{f,vd}^{*} = \Pr ob(CaseIII) \times \Pr ob(CaseIIIa \mid CaseIII)$$

$$= P_{f,vd} \int_{x=\frac{b}{r_1+r_2}}^{\infty} \int_{y=\frac{b}{r_1+r_2}}^{\infty} m_1 e^{-m_1 x} m_2 e^{-m_2 y} dydx \qquad (2.8)$$

$$= \left[\frac{(r_1+r_2) l_1 l_2}{r_1 l_1 m_2 + r_2 l_2 m_1 + (r_1+r_2) l_1 l_2} \right] e^{-(m_1+m_2)\left(\frac{b}{r_1+r_2}\right)}$$

Note that $P_{f,v}^{*}$, $P_{f,d}^{*}$ and $P_{f,vd}^{*}$ are equal to zero when $b \to \infty$. Fig. 2.5 also shows that the call/data termination probabilities decrease as b increases. However if b is too large, it may not be good for the user experience because calls cannot be made, even though there is still much remaining credit b. It is desirable to choose an appropriate value of "b" so that the call/data termination probabilities are lower than a threshold.

Let P_c^{*} be the probability that the calls terminate normally before the credit depletes. Note that

$$P_c^{*} = 1 - P_{f,v}^{*} - P_{f,d}^{*} - P_{f,vd}^{*} \qquad (2.9)$$

Assume that b^{*} is the remaining credit when the last call terminates. The expected value of b^{*} can be expressed as equation 2.10, shown at the next page.

$$E[b^*]=\left(\frac{1}{P_c^*}\right)\Big\{P_{f,v}\int_{x=0}^{\frac{b}{r_1}}(b-r_1x)m_1e^{-m_1x}dx+P_{f,d}\int_{y=0}^{\frac{b}{r_2}}(b-r_2y)m_2e^{-m_2y}dy$$

$$+P_{f,vd}\int_{x=0}^{\frac{b}{r_1}}\int_{y=0}^{\frac{b-r_1x}{r_2}}(b-r_1x-r_2y)m_1e^{-m_1x}m_2e^{-m_2y}dydx\ \Big\}$$

$$=\left(\frac{1}{P_c^*}\right)\Big\{P_{f,v}\left[b+\left(\frac{r_1}{m_1}\right)e^{-m_1\left(\frac{b}{r_1}\right)}-\frac{r_1}{m_1}\right]+P_{f,d}\left[b+\left(\frac{r_2}{m_2}\right)e^{-m_2\left(\frac{b}{r_2}\right)}-\frac{r_2}{m_2}\right]$$

$$+P_{f,vd}\left\{b\left[1-e^{-\left(\frac{b}{r_1}\right)m_1}\right]-\left[\frac{bm_1e^{-m_2\left(\frac{b}{r_2}\right)}}{m_1-\left(\frac{r_1}{r_2}\right)m_2}\right]\left\{1-e^{-\left[m_1-\left(\frac{r_1}{r_2}\right)m_2\right]\left(\frac{b}{r_1}\right)}\right\}\right.$$

$$+be^{-m_1\left(\frac{b}{r_1}\right)}-\left(\frac{r_1}{m_1}\right)\left[1-e^{-m_1\left(\frac{b}{r_1}\right)}\right]+\left[\frac{r_1m_1e^{-m_2\left(\frac{b}{r_2}\right)}}{m_1-m_2\left(\frac{r_1}{r_2}\right)}\right]$$

$$\times\left\{-\left(\frac{b}{r_1}\right)e^{-\left[m_1-\left(\frac{r_1}{r_2}\right)m_2\right]\left(\frac{b}{r_1}\right)}+\left[\frac{1}{m_1-\left(\frac{r_1}{r_2}\right)m_2}\right]\left[1-e^{-\left[m_1-\left(\frac{r_1}{r_2}\right)m_2\right]\left(\frac{b}{r_1}\right)}\right]\right\}$$

$$+\left[\frac{bm_1e^{-\left(\frac{b}{r_2}\right)m_2}}{m_1-\left(\frac{r_1}{r_2}\right)m_2}\right]\left\{1-e^{-\left[m_1-\left(\frac{r_1}{r_2}\right)m_2\right]\left(\frac{b}{r_1}\right)}\right\}-\left[\frac{r_1m_1e^{-\left(\frac{b}{r_2}\right)m_2}}{m_1-\left(\frac{r_1}{r_2}\right)m_2}\right]$$

$$\times\left\{-\left(\frac{b}{r_1}\right)e^{-\left[m_1-\left(\frac{r_1}{r_2}\right)m_2\right]\left(\frac{b}{r_1}\right)}+\left[\frac{1}{m_1-\left(\frac{r_1}{r_2}\right)m_2}\right]\left\{1-e^{-\left[m_1-\left(\frac{r_1}{r_2}\right)m_2\right]\left(\frac{b}{r_1}\right)}\right\}\right\}$$

$$-\left(\frac{r_2m_1}{m_2}\right)\left\{\left(\frac{1}{m_1}\right)\left[1-e^{-\left(\frac{b}{r_1}\right)m_1}\right]-\left[\frac{e^{-\left(\frac{b}{r_2}\right)m_2}}{m_1-\left(\frac{r_1}{r_2}\right)m_2}\right]\left\{1-e^{-\left[m_1-\left(\frac{r_1}{r_2}\right)m_2\right]\left(\frac{b}{r_1}\right)}\right\}\right\}\ \Big\}\Big\}$$

(2.10)

2.4 Numeric Results

This section investigates the performance of prepaid GSM and GPRS services based on the analytic model developed in the previous section. Simulation experiments have been conducted to validate the analytic results. Each simulation experiment was repeated 500,000 times to ensure stable results. To reflect the situation of prepaid service and GPRS service in Taiwan, R.O.C., the prepaid credit B is NT$500, the charge rate of a voice call is NT$0.2/sec and the charge rate of a data call ranges from NT$0.08/sec (for high usage) to NT$0.8/sec (for low usage). Table 2.1 compares the results of analytic and simulation models. The table shows that the analytic results are consistent with the simulation results for various charge rates r_1 and r_2.

Table 2.1: Comparison of analytic and simulation models

	$P^*_{f,v}$(%)		$P^*_{f.d}$(%)	
(r_1,r_2)	Simulation	Analytic	Simulation	Analytic
(0.2, 0.08)	75.66	75.65	8.63	8.59
(0.2, 0.3)	54.79	54.75	29.87	29.91
(0.2, 0.5)	43.64	43.76	41.45	41.33
(0.2, 0.8)	33.60	33.63	51.82	51.88
(r_1,r_2)	$P^*_{f,vd}$(%)		$E[b^*]$(NT$)	
	Simulation	Analytic	Simulation	Analytic
(0.2, 0.08)	5.33	5.37	1.54	1.54
(0.2, 0.3)	7.46	7.47	1.52	1.52
(0.2, 0.5)	8.56	8.58	1.51	1.52
(0.2, 0.8)	9.68	9.61	1.52	1.51

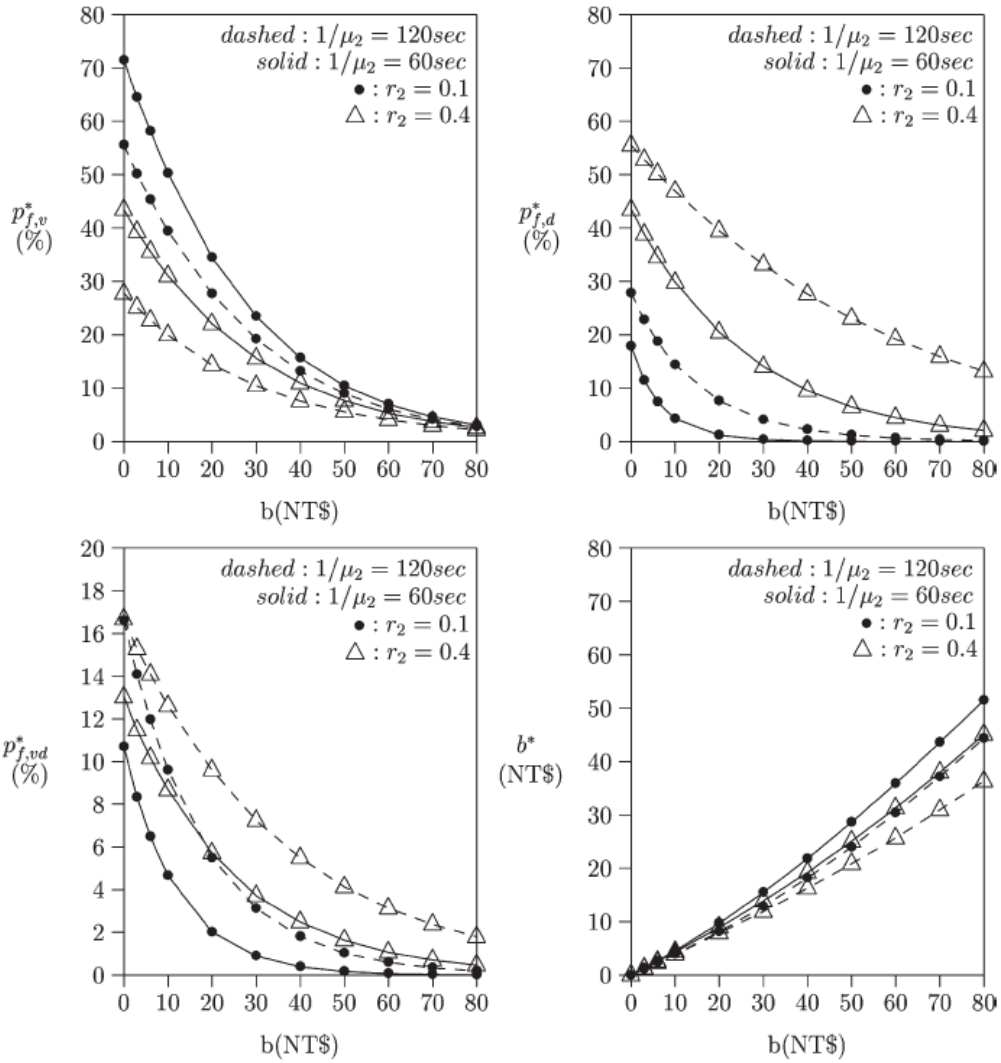


Fig. 2.5: Effect b on $P^*_{f,v}$, $P^*_{f,d}$ and $P^*_{f,vd}$

2.4.1 Effect of Threshold

Fig. 2.5 depicts the effect of b on $P^*_{f,v}$, $P^*_{f,d}$ and $P^*_{f,vd}$. The arrival rates of voice calls and data calls are both Poisson distributed with $\lambda_1=\lambda_2=1/600$ sec^{-1}. The arrival rates were chosen to emulate the behavior of a user with high usage. The voice call holding times are exponentially distributed with mean value of 2 minutes (i.e., $1/\mu_1=120$ sec) and the charge rate of a voice call is NT$0.2/sec. In Fig. 2.5, two cases are considered for the mean of data call holding time: $1/\mu_2=60$ sec and 120 sec. In a GPRS network, the maximum transmission speed

over the air is in the range of 40-53 kb/s[49].Hence, the mean size of data transmission in our experiments corresponds to 300 and 600kB.

The results in Fig. 2.5 indicates that $P^*_{f,v}$, $P^*_{f,d}$ and $P^*_{f,vd}$ decrease as the NMC threshold b increases. The decreasing rates are rapid when b is small, but they slow down as b increases. Smaller forced-termination probabilities can provide better quality of service to the subscriber. This improvement is obtained at the cost of increasing the NMC threshold that the subscriber is not allowed to use. Fig. 2.5 also shows that when b is less than NT$30, b^* equals to about b/2. When b is chosen to be the total cost of a voice call and a data call, $P^*_{f,vd}$ can be reduced to 2% in our experiment.

2.4.2 Effect of the Variance of Data Call Holding Times

We have also studied the effects of the variance of data call holding time on $P^*_{f,v}$, $P^*_{f,d}$ and $P^*_{f,vd}$. The data call holding time is assumed to have a Gamma distribution. A Gamma distribution has the following density function :

$$f(t) = \frac{m^a}{\Gamma(a)} t^{a-1} e^{-mt} \quad \text{for } t > 0 \qquad (2.11)$$

where α (>0) is the shape parameter, μ (>0) is the scale parameter and $\Gamma(q) = \int_{z=0}^{\infty} z^{q-1} e^{-z} dz$.

Let C_{x^*} be the coefficient of the variation of data call holding time. For Gamma distribution, $C_{x^*} = (1/\sqrt{a})$. A small C_{x^*} means that the variation is small and the lengths of most calls are close to the mean length. When C_{x^*} is large, there are a few long calls and many short calls. In our experiment, C_{x^*} varies in the range from 10^{-2} - 10^2. The arrival rates, call holding times and the charge rates are assumed to be the same as those in the previous experiment. The NMC threshold b is chosen to be NT$24.

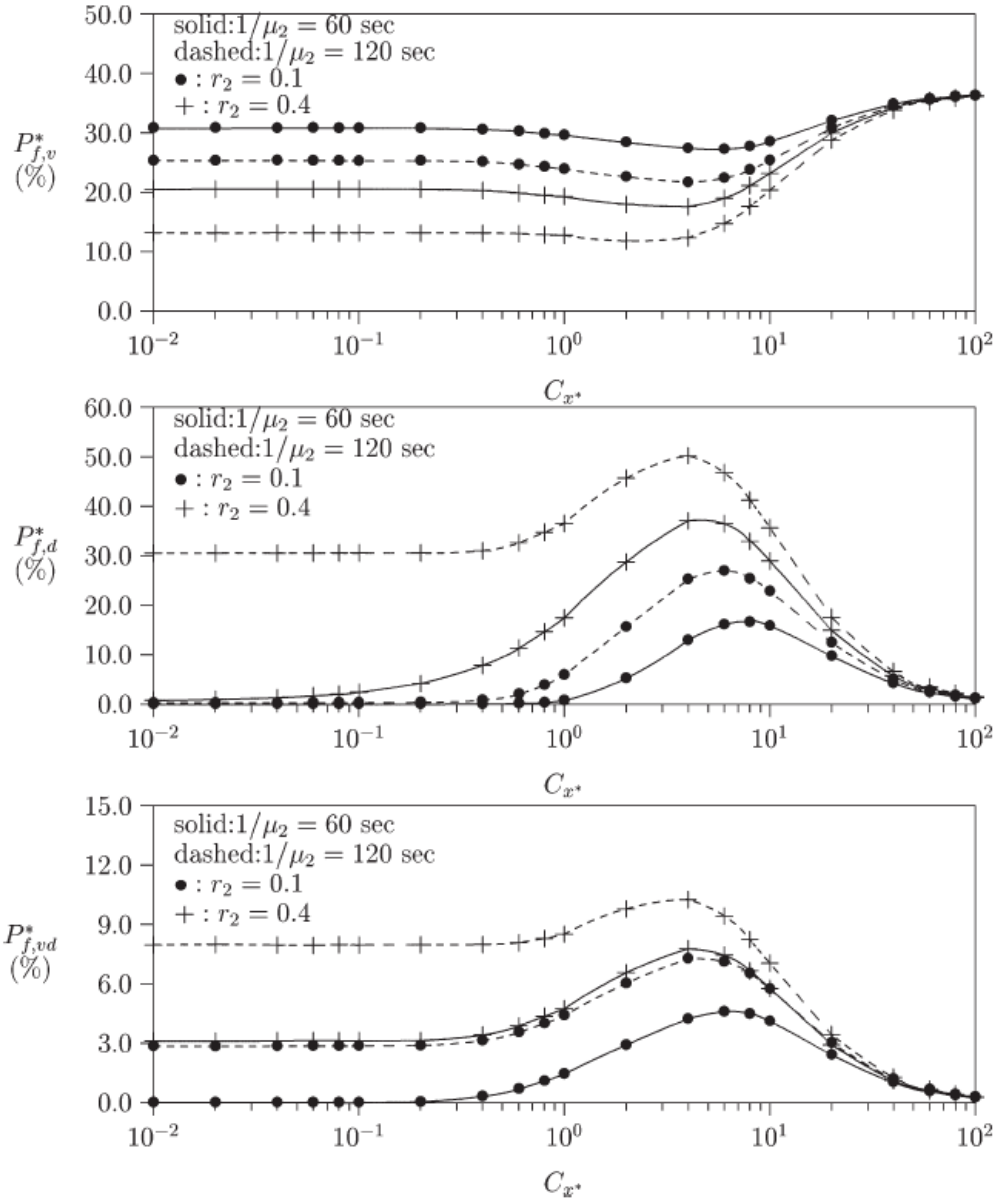


Fig. 2.6: Effect of C_{x^*} on $P^*_{f,v}$, $P^*_{f,d}$ and $P^*_{f,vd}$

Fig. 2.6 shows that when $C_{x^*} < 0.5$, the forced termination probabilities $P^*_{f,v}$, $P^*_{f,d}$ and $P^*_{f,vd}$ are insensitive to C_{x^*}, but are sensitive to the charge of a data call. This is because when C_{x^*} is very small, most of the length of data calls are close to the mean value. The data calls with higher charge rate are more likely to be forced to terminate. Therefore, $P^*_{f,d}$

increases and $P_{f,v}^{*}$ decreases as r_2 increases.

As C_{x^*} increases (i.e., $C_{x^*} > 0.5$), the number of long data calls whose charge is larger than the NMC threshold increases. As a result, more data calls are forced to terminate; $P_{f,d}^{*}$ and $P_{f,vd}^{*}$ increase as C_{x^*} increases. When C_{x^*} increases further (i.e., $C_{x^*} > 5$), the number of short data calls increase. More data calls complete normally and $P_{f,d}^{*}$ and $P_{f,vd}^{*}$ decrease as C_{x^*} increases. For $C_{x^*} > 10^2$, $P_{f,d}^{*}$ and $P_{f,vd}^{*}$ are near zero. This is because most of the data calls are very short and can complete normally. Only the very long calls, whose number is small, may be forced to terminate. We come to the conclusion that NMC effect is more significant when C_{x^*} is large. Since the variance of voice call length remain the same, the probability that a voice call is forced to terminate [i.e., $(P_{f,v}^{*} + P_{f,vd}^{*})$] slightly increases as C_{x^*} increases. When $P_{f,vd}^{*}$ increases, $P_{f,v}^{*}$ decreases and vise versa.

2.5 Conclusions

This chapter investigates the charging issues of an integrated GSM and GPRS prepaid service, where a single prepaid account provides the user both voice and data services. Base on the CAMEL network architecture, the call setup and charging procedures for GSM and GPRS have been illustrated. We propose an NMC algorithm to reduce the probability of terminating both on-going voice and data calls, where no more new calls are admitted when the user credit is below a threshold. An analytic model has been developed to evaluate the performance of the NMC algorithm. Computer simulations have also been used to verify the analytical results. The numeric results indicate that the forced termination probability can be significantly reduced by choosing an appropriate threshold of the user credit. In addition, the forced termination probability of voice calls slightly increases as the call pattern of data calls

becomes irregular. As the need to rapidly roll out new services in mobile networks, our model can be easily extended to accommodate different real-time services. Our analytic method could provide guidelines helping the operators to generate higher revenues.

CHAPTER 3 A Secure Mobile Electronic Payment Architecture Platform for Wireless Mobile Networks

3.1 Introduction

With the vast development and deployment of wireless mobile networks such as 3G UMTS [26,39], WiMAX [30] and Wi-Fi [29], mobile networking applications enabling customers to gain network access anywhere and anytime have attracted more and more attention in our daily lives. When the basic functionalities of a wireless network have been in place, customers are now more interested in value-added mobile applications over this network. Most mobile applications come with the emergence of electronic trading (mobile commerce or m-commerce), hence good secure mobile trading model must be designed to attract more mobile users for doing business wirelessly. Thus, how to integrate the mobile applications with a secure trading model becomes an important design issue, which will significantly affect the success of any value-added mobile application.

Mobile applications can be categorized into session-based applications and event-based applications. In event-based applications, user's payment is reflected by one-time events. Examples include sending a message, querying traffic information, or purchasing a song. A session-based application consists of three phases: the session-setup phase, the communication phase and the session release phase. A customer is charged for a session-based application based on either time spent or data volume transferred, e.g., VoIP-calling, video-streaming, audio-streaming, or video-conferencing.

Compared with fixed networks, mobile networks have lower bandwidth, longer transmission latency, and more unreliable connections, and mobile devices are restricted by limited memory size and low CPU computational capability [34]. The installation of mobile

applications on a mobile network should be quick and of low cost. To summarize, the following requirements should be addressed when designing a suitable trading mechanism on a mobile network. First, customers expect a robust, secure, and fair trading mechanism which can be applied in different mobile networks. Second, the trading mechanism should be light-weight (i.e., with low computational complexity and low communication overhead) so that it can be easily run on mobile devices. Third, user anonymity should be achieved, that is, users' purchasing behavior or preference should not be traceable by others. Finally, a trading mechanism should be of low implementation cost.

We design an application-level secure payment model, named Mobile Electronic Payment (MEP), for wireless mobile networks, which attempts to meet these requirements. It is based on a more general trading architecture model, which combines both public-key cryptography and symmetric-key cryptography to overcome the disadvantages of both technologies. Our design keeps the key freshness and thus provides more robust security protection. Moreover, MEP supports both event-based and session-based applications and is suitable for the resource-constrained mobile devices because MEP attempts to alleviate the computational cost and reduce the memory space requirement in mobile devices. We expect that our MEP provides a viable trading model for the future mobile applications.

3.2 Preliminaries

3.2.1 General Conceptual Trading Model

Fig. 3.1. illustrates the general conceptual trading model for mobile applications [24, 55], which consists of three major components: the network operator **O** (Fig. 3.1 (1)), the user (customer) **U** (Fig. 3.1 (2)), and the mobile applications/content provider **P** (Fig. 3.1. (3)). The Ps supply mobile applications to Us. The **O** provides network bearer services (e.g., the UMTS

bearer services or the WLAN services) to **U**s, through which **U**s may use different kinds of mobile devices to access the applications. **P** and **O** may reside in different networks. For example, **O** is the operator of a cellular network, and **P** resides in the Internet.

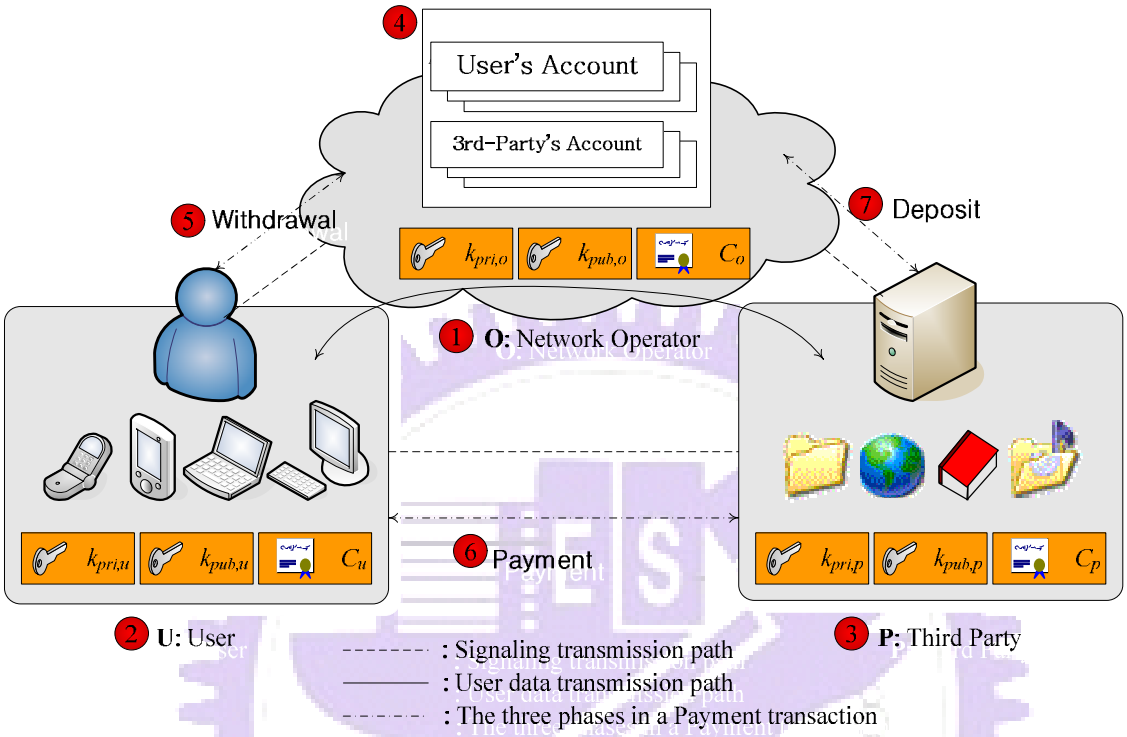


Fig. 3.1: The General Trading Model for Mobile Applications

In this trading model, **O** has to be trusted by **U** and **P**. Initially, **U** and **P** apply for accounts from **O**, and **O** maintains an account balance (Fig. 3.1 (4)) for each account. The public-private key pairs, ($k_{pub;o}$, $k_{pri;o}$), ($k_{pub;p}$, $k_{pri;p}$) and ($k_{pub;u}$, $k_{pri;u}$), and certificates, c_o, c_p, and c_u, which are held by **O**, **U**, and **P**, respectively, are used to address the security issues such as the confidentiality and authentication. The certificate is used to verify the owner of a public key. The certificate uses a digital signature to bind a public key with an individual's identity information (e.g., telephone number or email address). The public-private key pairs are used to encrypt and decrypt all the signaling messages exchanged among **O**, **U**, and **P**.

Before **U** purchases a mobile application from **P**, it initiates a Payment Transaction among **O**, **P**, and **U**. The creation process of a payment transaction consists of three phases

[24]: the Withdrawal phase (Fig. 3.1 (5)), the Payment phase (Fig. 3.1 (6)), and the Deposit phase (Fig. 3.1 (7)). The process begins at the Withdrawal phase where **U** obtains the electronic means (e.g., the electronic tokens [10, 43] or the value-added smart card [23]) from **O**. Then, the process enters the Payment phase. In the Payment phase, **U** issues the electronic means to **P**, which is known as “payment". Then **P** checks the validity of the electronic means. If it is valid, **U** is permitted to purchase a mobile application. The payment may be performed either once or many times, which depends on whether the application is event-based or session-based. For an event-based application, only one payment is made in this phase. For a session-based application, multiple payments may be executed. When the mobile application ends, the process gets into the Deposit phase. In this phase, **P** uses the electronic means obtained from **U** to exchange the payment with **O**, where **O** verifies the electronic means and deposits the payment into **P**'s account.

3.2.2 Basics of the ID-based Cryptography

This section briefly discusses the fundamentals of the ID-based cryptography (IBC) [11]. The general IBC concept was proposed in [47] in 1984. Only after 2001 when Boneh and Franklin [11] successfully implemented the IBC concept by using the bilinear pairing function does IBC gain more popularity and show many more useful applications of this technique [56, 57, 58, 59, 60]. In IBC, there is no binding between the user ID and public keys. With this future, the proposed MEP adopts IBC to save transactions and cost associated with either communications or computation.

In the IBC, each user owns a key pair (k_{pub}; k_{pri}). The k_{pub} is a public key, which is derived from a user's identity information (e.g., user's telephone number or email address). The derivation of k_{pub} can be done at the user's device or at the trusted authority (e.g., a network operator). The k_{pri} is a private key, which is generated by the trusted authority by

taking the k_{pub} into a function *f* and is passed to the user through a secure link. As mentioned in [11, 18, 60], the main advantage of the IBC is that there is no need to have certificate to bind user names with their public keys.

The IBC relies on a concept called pairing. A bilinear pairing function is a mapping defined as $\hat{e} : G_1 \times G_1 \rightarrow G_2$ where G_1 and G_2 are cyclic groups satisfying the following three properties:

- *Property 1*: (G_1; +) and (G_2;×) must be cyclic groups of the same order *K*, i.e., both G_1 and G_2 contain *K* elements [25].
- *Property 2*: Given $p,q \in G_1$, it is computationally infeasible to extract *X* (where $X \in Z_K^* = \{i \mid 1 \le i \le K-1\}$ from $q = X \cdot p$, where $X \cdot p$ is *p* added to itself *X* times.
- *Property 3*: Given $m,n \in G_2$, it is computationally infeasible to extract *Y* (where $Y \in Z_K^*$) from $n = m^Y$. Note that m^Y is m multiplied by itself *Y* times.

The details for G_1 and G_2 can be found in [11]. A bilinear pairing function $\hat{e} : G_1 \times G_1 \rightarrow G_2$ has the following four properties:

- *Bilinearity*: For all $p,q \in G_1$ and all $C,D \in Z_K^* = \{i \mid 1 \le i \le K-1\}$, $\hat{e}(C \cdot p, D \cdot q) = \hat{e}(C \cdot p, q)^D = \hat{e}(p, D \cdot q)^C = \hat{e}(p, q)^{C \cdot D}$
- *Non-degeneracy*: If g is a generator of G1, then $\hat{e}(g, g)$ is a generator of G_2.
- *Computability*: There is an *efficient* algorithm to compute $\hat{e}(p, q)$ for all $p,q \in G_1$.
- *Symmetry*: For all $p,q \in G_1$, $\hat{e}(p, q) = \hat{e}(q, p)$.

Modified Weil [11] and Tate [7] pairings are examples of such bilinear maps for which the Bilinear Diffie-Hellman Problem (BDHP) is believed to be hard. We refer to [11, 7] for a more comprehensive description of how the pairing parameters should be chosen in practice.

3.3 The Mobile Electronic Payment (MEP) Platform

In this section, we present the MEP platform which follows the general trading model. When a new user **U** or a mobile application/content provider **P** joins the MEP, the Key Distribution procedure (to be elaborated later) is executed to distribute **U** or **P** public-private key pairs denoted as ($k_{pub,u}$, $k_{pri,u}$) or ($k_{pub,p}$, $k_{pri;p}$), respectively. Then, **U** can purchase a mobile application from **P** by running a payment transaction. In a payment transaction, the signaling messages exchanged among **O**, **U**, and **P** are encrypted using three symmetric keys $k_{u\text{-}o}$ (held by **O** and **U**), $k_{o\text{-}p}$ (held by **O** and **P**), and $k_{u\text{-}p}$ (held by **U** and **P**). The three symmetric keys are updated (by utilizing the public-private key pairs) at the beginning of every payment transaction. A payment transaction consists of three phases, the Withdrawal phase (where **U** obtains the tokens from **O**), the Payment phase (where **U** uses the tokes to purchase a mobile application from **P**), and the Deposit phase (where **P** redeems the obtained tokens from **O**).

In the following subsections, we first illustrate the key distribution procedure and then describe how a payment transaction is executed in MEP.

3.3.1 The Key Distribution Procedure

The key distribution procedure generates public-private key pairs for **O**, **U** and **P**. The design of this procedure utilizes the IBC to eliminate the certificate overhead from binding one's ID with its public key. Fig. 3.2 illustrates the message flow for this procedure with the following steps:

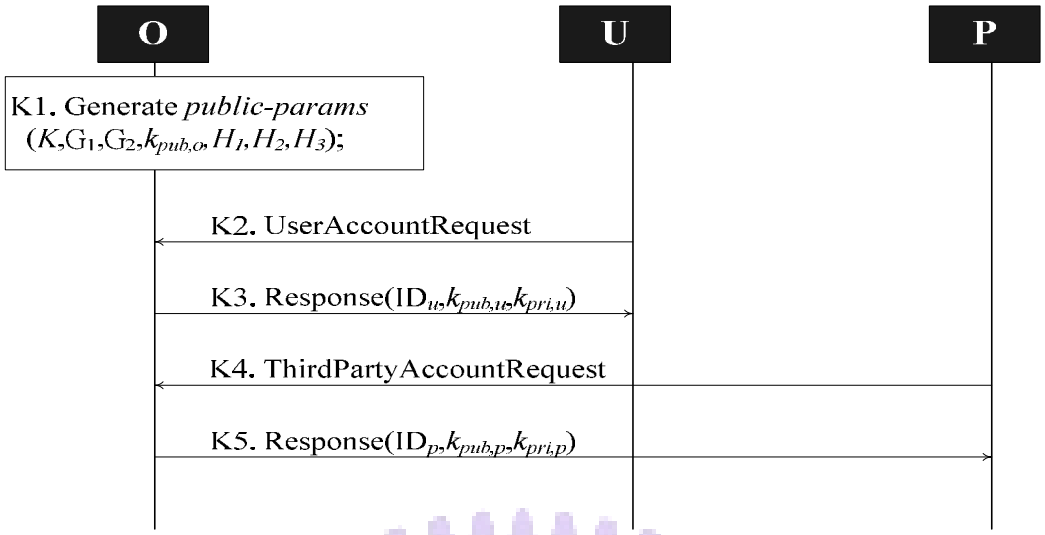


Fig. 3.2: Message flow for The Key Distribution Procedure

Algorithm 1 GENERATE-PARAMS

1: Generate the pairing parameters (K, G_1, G_2, $\hat{e}$);
2: Select an arbitrary generator for G_1 as the public key $k_{pub,o}$;
3: Choose a hash function H_1:$\{0,1\}^* \rightarrow G_1$;
4: Choose a hash function H_2:$G_2 \rightarrow \{0,1\}^N$ for some integer N;
5: Choose a one-way hash function H_3:$\{0,1\}^* \rightarrow \{0,1\}^M$ for some integer M (e.g., H_3 can be SHA-1 and MD5 [50]);
6: return (K,G_1,G_2,$\hat{e}$, $k_{pub,o}$,H_1,H_2,H_3)

Fig. 3.3: The GENERATE-PARAMS algorithm

Step K1. **O** first generates a public-params set (K,G_1,G_2,$\hat{e}$, $k_{pub,o}$,H_1,H_2,H_3) by the generate-params algorithm as shown in Fig. 3.3 The public-params set contains all parameters required in MEP. The usage of the parameters is listed in Table 3.1. Then **O** publishes the generated public-params set in a public place (e.g., website). **O** selects a random number $S \in Z_K^*$, and derives its private key $k_{pri,o}$ by computing

$$k_{pri,o} = S \cdot k_{pub,o} \quad (3.1)$$

where " $\cdot$ " is defined in Property 2 in Section 3.2.2. **O** keeps S and $k_{pri,o}$ confidential.

Step K2. **U** sends **O** the UserAccountRequest message to apply for a user account.

Table 3.1: The usage of the parameters in public-params set

Parameter	Usage	Parameter	Usage
K	The order of G_1 and G_2	$k_{pub,o}$	The **O**'s public key
G_1	The cyclic group with operation "+"	H_1	The hash function used to derive one's ID to its public key
G_2	The cyclic group with operation "×"	H_2	The hash function used to derive the output of Bilinear Pairing function ê to a symmetric key
ê	The Bilinear Pairing function	H_3	The hash function used to generate the electronic means

Step K3. Upon receiving **U**'s request, **O** selects an ID, ID_u, for **U** and creates an account for **U**.

Then **O** generates U's public key $k_{pub,u}$ and private key $k_{pri;u}$ by

$$k_{pub,u} = H_1(ID_u), \tag{3.2}$$

and

$$k_{pri,u} = S \cdot k_{pub,u}. \tag{3.3}$$

O sends $k_{pub,u}$, $k_{pri,u}$, and ID_u to **U** through the bearer network link. Since **U** is the customer of **O**, the bearer network is considered secure.

Steps K4 and K5. The two steps are similar to Steps K2 and K3, respectively. **P** applies a third-party account by sending the ThirdPartyAccountRequest message to **O**. **O** selects an ID, ID_p, creates an account for **P**, and generates **P**'s public key $k_{pub,p}$ and private key $k_{pri,p}$ by

$$k_{pub,p} = H_1(ID_p) \tag{3.4}$$

and

$$k_{pri,p} = S \cdot k_{pub,p} \tag{3.5}$$

O sends $k_{pub,p}$, $k_{pri,p}$, and ID_p to **P** through the secure link between **O** and **P**.

3.3.2 Payment Transaction in MEP

In this section, we describe the execution of a payment transaction in MEP for **U** to purchase a mobile application from **P**. Following the general trading model, a payment transaction in MEP consists of three phases: the Withdrawal phase, the Payment phase and the Deposit phases, which are described below.

3.3.2.1 Withdrawal Phase

In this phase, **U** obtains the electronic means (i.e., the tokens) from **O**. Fig. 3.4 illustrates the message flow for this phase with the following steps. To simplify our description, we use k(D) to denote that the data D is encrypted by the symmetric key k with an efficient symmetric-key algorithm.

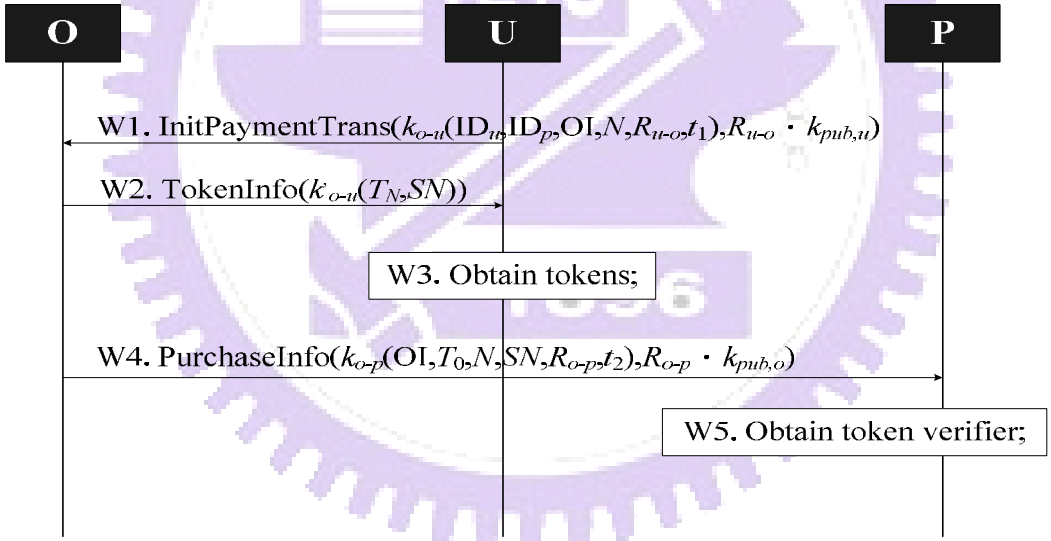


Fig. 3.4: Message flow for the Withdrawal phase of a payment transaction

Step W1. By browsing **P**'s website, **U** selects a mobile application, gets **P**'s ID, ID$_p$, and obtains the Order Information (OI) containing the ID and the data unit price of the mobile application. Then, **U** randomly selects an integer $R_{u\text{-}o}$ from Z_K^* and generates the symmetric key $k_{u\text{-}o}$ by computing

$$k_{u\text{-}o} = H_2(\hat{e}\ (R_{u\text{-}o} \cdot k_{pub,o}, k_{pri,u})). \quad (3.6)$$

Then **U** sends an InitPaymentTrans message to **O** to initiate a payment transaction,

where $k_{u\text{-}o}$ (ID_u, ID_p, OI,N, $R_{u\text{-}o}$, t_1) and $R_{u\text{-}o} \cdot k_{pub,u}$ are carried in the message. The first parameter $k_{u\text{-}o}$ (ID_u, ID_p, OI,N, $R_{u\text{-}o}$, t_1) contains the necessary information for **O** to generate the tokens for **U**. N is the amount of data units **U** will purchase, and t_1 is the current system time, which is used to prevent message replay and impersonation attacks [44]. The second parameter $R_{u\text{-}o} \cdot k_{pub,u}$ will be used by **O** to derive the symmetric key $k'_{u\text{-}o}$ (see Step W2) and authenticate **U**. Note that $k'_{u\text{-}o}$ is the same as $k_{u\text{-}o}$ (*Proposition 1*), so that **O** can decrypt the $k_{u\text{-}o}$ (ID_u, ID_p, OI,N, $R_{u\text{-}o}$, t_1) parameter.

Step W2. Upon receipt of the InitPaymentTrans message, **O** will perform the following tasks:

(i) **O** extracts the second parameter $R_{u\text{-}o} \cdot k_{pub,u}$ from the InitPaymentTrans message, and uses this parameter and **O**'s private key $k_{pri,o}$ to derive the symmetric key $k'_{u\text{-}o}$ as

$$k'_{u\text{-}o} = H_2(\hat{e}(R_{u\text{-}o} \cdot k_{pub,u}, k_{pri,o})). \quad (3.7)$$

Then **O** uses $k'_{u\text{-}o}$ to decrypt $k_{u\text{-}o}$ (ID_u, ID_p, OI,N, $R_{u\text{-}o}$, t_1), and **O** obtains the ID_u, ID_p, OI, N, $R_{u\text{-}o}$, and t_1.

Algorithm 2 GENERATE-TOKEN(T_N,N)

1: for $i \leftarrow N$**-1 downto 0 do**

2: $T_i \leftarrow H_3(T_{i+1})$

3: return $\langle T_{N-1}, T_{N-2}, \ldots, T_0 \rangle$

Fig. 3.5: The GENERATE-TOKEN algorithm

(ii) To autenticate the sender of the InitPaymentTrans message, **O** verifies whether $R_{u\text{-}o} \cdot H_1(ID_u)$ (where $R_{u\text{-}o}$ and ID_u are obtained in (i)) is equal to the second parameter $R_{u\text{-}o} \cdot k_{pub,u}$. If they are not equal (i.e., $R_{u\text{-}o} \cdot H_1(ID_u) \neq R_{u\text{-}o} \cdot k_{pub,u}$), the sender is illegal, and the phase quits without sending extra messages. If they

are equal (i.e., $H_1(\mathrm{ID}_u) = k_{pub,u}$), the sender is authenticated and then **O** checks whether the difference between t_1 and the local clock time is within an acceptance window to prevent from message replay and impersonation [44].

(iii) If the authentication is successful, **O** will then generate the tokens for **U**. Suppose that each data unit consumes one token, and N tokens are required for **U**. Let $\langle T_N, T_{N-1}, T_{N-2},\ldots, T_1 \rangle$ denote the N tokens. Initially, **O** selects a random number as the token root T_N. Then **O** executes the Generate-Tokens algorithm (see Fig. 3.5) with arguments T_N and N to generate N tokens. After executing the algorithm, **O** obtains the tokens T_{N-1}, $T_{N-2},\ldots,$ T_1 and a token verifier T_0. The token verifier T_0 will be used by **P** to make sure that the tokens are sent from **U** in the Payment phase. Each token indicates the data unit price of the mobile application. Then **O** deducts the cost for N tokens from **U**'s account.

(iv) The payment transaction is assigned an unique serial number SN by **O**. Then **O** sends **U** the TokenInfo message carrying $k'_{u\text{-}o}(T_N, SN)$.

Step W3. Upon receipt of the TokenInfo message, **U** uses $k_{u\text{-}o}$ to decrypt the message and obtains T_N and SN. Then, **U** uses the token root T_N to generate N tokens by executing the Generate-Token algorithm. Note that due to the lightweight feature of the hash function H_3 [46], the tokens are generated efficiently.

Step W4. O selects a random integer $R_{o\text{-}p}$ from Z_K^* and generates the symmetric key $k_{o\text{-}p}$ as

$$k_{o\text{-}p} = H_2(\hat{e}(R_{o\text{-}p} \cdot k_{pub,p}, k_{pri,o})). \quad (3.8)$$

where **P**'s public key $k_{pub,p}$ is obtained by $k_{pub,p} = H_1(\mathrm{ID}_p)$. Then, **O** sends **P** a PurchaseInfo message to notify that **U** wants to purchase the mobile application. The parameters carried in the message contains $k_{o\text{-}p}(\mathrm{OI}, T0, N, SN, R_{o\text{-}p}, t_2)$ and $R_{o\text{-}p} \cdot k_{pub,o}$, where t_2 is the current system time used to prevent from message replay

and impersonation and the second parameter $R_{o\text{-}p} \cdot k_{pub,o}$ will be used by P to derive the symmetric key $k'_{o\text{-}p}$ (to be elaborated in next step). Then, using *SN* as the index, **O** stores the information (containing ID_u, ID_p, N, T_N,and $k_{o\text{-}p}$) required in Deposit phase into its database.

Step W5. Upon receipt of the PurchaseInfo message, **P** extracts the second parameter $R_{o\text{-}p} \cdot k_{pub,o}$ from the message, and uses this parameter and **P**'s private key $k_{pri,p}$ to derive the symmetric key $k'_{o\text{-}p}$ as

$$k'_{o\text{-}p} = H_2(\hat{e}(R_{o\text{-}p} \cdot k_{pub,o}, k_{pri,p})). \quad (3.9)$$

Note that from Proposition 1, we know $k'_{o\text{-}p} = k_{o\text{-}p}$. **P** uses $k'_{o\text{-}p}$ to decrypt the first parameter $k_{o\text{-}p}$ (OI, T_0, N, *SN*, $R_{o\text{-}p}$, t_2).

Similar to Step W2.(ii), **P** calculates $R_{o\text{-}p} \cdot k_{pub,o}$ ($R_{o\text{-}p}$ is obtained from the first parameter and $k_{pub,o}$ is obtained from the public-params set) and checks whether the result is equal to the second parameter $R_{o\text{-}p} \cdot k_{pub,o}$ carried in the PurchaseInfo message. If they are not equal, the sender is illegal, and the phase quits without sending extra message. If they are equal, the sender is ensured to be **O**. **P** checks whether the difference between t_2 and the local clock time is within an acceptance window to prevent from message replay and impersonation. Then, using *SN* as the index, **P** stores the information (containing OI, T_0, N, and $k_{o\text{-}p}$) required in the payment phase and the deposit phase into its database.

The following result ensures the correctness of our MEP.

Proposition 1 $k_{u\text{-}o} = k'_{u\text{-}o}$ and $k_{o\text{-}p} = k'_{o\text{-}p}$.

Proof:

$$k_{u\text{-}o} = H_2(\hat{e}(R_{u\text{-}o} \cdot k_{pub,o}, k_{pri,u})) \quad \text{(from (3.6))}$$

$$= H_2(\hat{e}(R_{u\text{-}o} \cdot k_{pub,o}, S \cdot k_{pub,u})) \quad \text{(from (3.3))}$$

$$= H_2(\hat{e}\left(k_{pub,o}, k_{pub,u}\right)^{R_{u\text{-}o} \cdot S}) \quad \text{(Bilinearity of } \hat{e}\text{)}$$

$= H_2(\hat{e}(k_{pub,u}, k_{pub,o})^{R_{u-o} \cdot S})$ (Symmetry of $\hat{e}$)

$= H_2(\hat{e}(R_{u-o} \cdot k_{pub,u}, S \cdot k_{pub,o}))$ (Bilinearity of $\hat{e}$)

$= H_2(\hat{e}(R_{u-o} \cdot k_{pub,u}, k_{pri,o}))$ (from (3.1))

$= k'_{u-o}$ (from (3.7))

Similarly, we can prove the other identity.

3.3.2.2 Payment Phase

In the Payment phase, **U** uses the tokens to purchase a mobile application from **P**. This phase may consist of one or more payments. We assume that **U** pays J_i tokens in the ith payment. Fig. 3.6 illustrates the message flow for the Payment phase with the following steps.

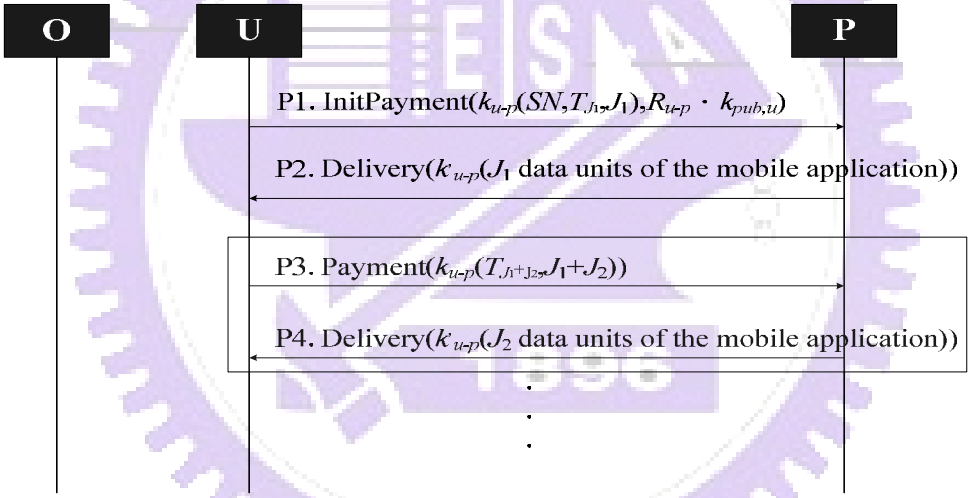


Fig. 3.6: Message flow for the Payment phase of a payment transaction

Step P1. **U** randomly selects an integer R_{u-p} from Z_K^* and generates the symmetric key k_{u-p} by

$$k_{u-p} = H_2(\hat{e}(R_{u-p} \cdot k_{pub,p}, k_{pri,u})). \quad (3.10)$$

where **P**'s public key $k_{pub,p}$ is obtained from $k_{pub,p} = H_1(\mathrm{ID}_p)$. Then **U** initiates the first payment by sending a InitialPayment message, where J_1 tokens $\langle T_1, T_2, \ldots T_{J_1} \rangle$ are carried in this message. The parameters of the InitialPayment message

include $k_{u\text{-}p}$ (SN, T_{J_1}, J_1) and $R_{u\text{-}p} \cdot k_{pub,u}$. The second parameter $R_{u\text{-}p} \cdot k_{pub,u}$ will be used by **P** to derive the symmetric key $k'_{u\text{-}p}$ (see (3.11), Step P2).

Note that in this step, **P** cannot directly extract $k_{pub,u}$ easily from the second parameter $R_{u\text{-}p} \cdot k_{pub,u}$, and the InitialPayment message does not contain any information that may leak out **U**'s identity. Therefore, "user anonymity" is well protected.

Step P2. Upon receipt of the InitialPayment message, **P** uses the second parameter $R_{u\text{-}p} \cdot k_{pub,u}$ and **P**'s private key $k_{pri,p}$ to generate the symmetric key $k'_{u\text{-}p}$ by

$$k'_{u\text{-}p} = H_2(\hat{e}(R_{u\text{-}p} \cdot k_{pub,u}, k_{pri,p})). \qquad (3.11)$$

From Proposition 1, $k'_{u\text{-}p}$ is the same as $k_{u\text{-}p}$. Using the symmetric key $k'_{u\text{-}p}$, **P** decrypts the first parameter $k_{u\text{-}p}$ (SN, T_{J_1}, J_1) and obtains SN, T_{J_1} and J_1. **P** uses SN as the index to query its database for the token verifier T_0, N, and OI. According to the mobile application ID contained in OI, **P** identifies the mobile application that **U** wants to purchase, and prepares N data units of the mobile application (e.g., streaming data for N seconds). To verify the token T_{J_1}, **P** checks whether the equation $\underbrace{H_3(H_3 \cdots (H_3}_{J_1}(T_{J_1})\cdots)) \stackrel{?}{=} T_0$ holds. If it holds, **P** ascertains that the token T_{J_1} is legal. P stores T_{J_1} for verifying the token carried in the next message and discards T_0 to release the memory space. Then **P** encrypts the first unit to the J_1th unit of the mobile application using the symmetric key $k'_{u\text{-}p}$ and responds with the J_1 data units carried in the Delivery message, to **U**.

Step P3. Upon receipt of the Delivery message, **U** decrypts the message using the symmetric key $k_{u\text{-}p}$ and obtains the J_1 data units. Then, **U** starts the 2nd Payment to purchase the next J_2 data units by sending **P** the Payment($T_{J_1+J_2}$, $J_1 + J_2$) message.

Step P4. Upon receipt of the Payment message, **P** decrypts the message using the symmetric key $k'_{u\text{-}p}$ and obtains $T_{J_1+J_2}$ and $J_1 + J_2$. **P** gets J_2 by subtracting ($J_1 + J_2$) $-$ J_1. Then **P** checks whether the equation $\underbrace{H_3(H_3 \cdots (H_3}_{J_2}(T_{J_1+J_2})\cdots)) \stackrel{?}{=} T_{J_1}$ holds. If the equality holds, **P** ascertains that the token $T_{J_1+J_2}$ is legal. **P** stores $T_{J_1+J_2}$ for verifying the token carried in next message and discards the token T_{J_1}. Then, **P** encrypts the next J_2 data units of the mobile application using the symmetric key $k'_{u\text{-}p}$ and delivers the J_2 data units to **U**.

Repeating Steps P3 and P4, **U** sends the succeeding tokens to **P**, and **P** delivers the succeeding data units to **U**. This phase may be terminated if **U** stops paying the token or **P** stops delivering the mobile application.

3.3.2.3 Deposit Phase

Assume that **P** receives J ($J \leq N$) tokens after the Payment phase. In the Deposit phase, **P** redeems the J tokens from **O**. This phase consists of the following two steps.

Step D1. **P** sends **O** the deposit message carrying the parameters SN and $k_{o\text{-}p}(T_J, J)$.

Step D2. Upon receipt of the deposit message, **O** uses the first parameter SN and the index to query its database for ID_u, ID_p, N, T_N, and $k_{o\text{-}p}$. Using the symmetric key $k_{o\text{-}p}$, **O** decrypts the second parameter $k_{o\text{-}p}(T_J, J)$ and obtains T_J and J, and then checks whether the equation $\underbrace{H_3(H_3 \cdots (H_3}_{N-J}(T_N)\cdots)) \stackrel{?}{=} T_J$ holds to verify the token T_J. If the equation holds, **O** deposits the credit for J tokens into **P**'s account and takes the cost for J tokens from **U**'s account. The payment transaction is completed. Otherwise (i.e., $\underbrace{H_3(H_3 \cdots (H_3}_{N-J}(T_N)\cdots)) \neq T_J$), **O** treats the sender of the deposit message as an adversary, and the deposit phase will not carried through.

Note that if **P** does not exercise Step D1 after the payment phase in a predefined time period (e.g., one day), the payment transaction is considered incomplete, and **O** gives the credit for all tokens into **U**'s account, and terminates the payment transaction.

3.4 Features and Overhead Analysis of MEP

3.4.1 Features of MEP

There are a few useful features of MEP including the avoidance of overspending and double spending, the fairness, the user anonymity, and the privacy, which are discussed next.

3.4.1.1 Avoidance of Overspending and Double Spending

Overspending means that **U**'s account does not have enough credit to purchase a mobile application. Double spending is that **U** uses the same tokens to purchase mobile applications from different **P**s. Both overspending and double spending cause financial loss to **O** and **P**. MEP adopts the “prepaid" approach, that is, **U**'s account is deducted (see Step W2 of the withdrawal phase, Section 3.3.2.1) before **U** purchases a mobile application. If **U**'s account does not have enough credit to purchase a mobile application, **O** will not issue tokens to **U**. Hence, MEP can avoid the loss due to a user's overspending. Moreover, when **U** withdraws tokens from **O**, it has to inform **O** of **P**'s ID (see Step W1 of the withdrawal phase, Section 3.3.2.1). Then, **O** sends tokens and the corresponding token verifiers to **U** and **P** (see Steps W2 and W4 of the withdrawal phase, Section 3.3.2.1), respectively. If **U** applies the same tokens to another **P**’, **P**’ will not accept the tokens because it does not own the corresponding token verifiers. Therefore, the risk of double spending is avoided.

3.4.1.2 Fairness

After a payment transaction, **U** can get the data units of a mobile application, whose

value is equivalent to the credits **U** pays for, and **P** can get the credits equivalent to the value of data units of the mobile application **P** provides [53], which is referred to as the “fairness".

In MEP, during the execution of the payment phase (see Section 3.3.2.2), **P** provides **U** the data units of the mobile application after **U** has paid the tokens (i.e., credits). Therefore, there is no risk for **P** to provide data units. Furthermore, **U** can terminate the token payment immediately if **P** does not send the requested data units. In this case, at most one token is lost, which is considered insignificant. The fairness feature can be accommodated in MEP.

3.4.1.3 User Anonymity

The user anonymity is defined in two levels: untraceability and unlinkability [6]. Untraceability means that **P** is not allowed to know **U**'s identity during the execution of a payment transaction. Unlinkability means that two different payment transactions (which involve the same **U**) cannot be linked by **P**, i.e., **P** is not allowed to identify the two payment transactions initiated by the same **U** so that any user profiling attempt fails.

In MEP, **U** and **P** negotiate only in the payment phase (see Section 3.3.2.2) for purchasing mobile applications. The information of **U** sent to **P** in this phase contains tokens, the serial number of a payment transaction, the total number of tokens paid to **P**, and the parameter $R_{u\text{-}p} \cdot k_{pub,u}$, which does not include any **U**-related information. Through the payment phase, **P** cannot identify **U**. Consequently, MEP ensures untraceability. Furthermore, the public key $k_{pub,u}$ of **U** cannot be extracted from the parameter $R_{u\text{-}p} \cdot k_{pub,u}$, which is shown in Property 2 of a bilinear pairing function in Section 3.2.2. Since **P** is not able to obtain either **U**'s ID or **U**'s public key, the unlinkability between any two different payment transactions can be achieved in MEP.

3.4.1.4 Privacy

Privacy means that the data of a mobile application exchanged between **U** and **P** cannot

be revealed by any unauthorized third party except **O** who distributes the keys. In MEP, the data of a mobile application is encrypted by the symmetric key $k_{u\text{-}p}$, where the $k_{u\text{-}p}$ is only known by **U** and **P** (see Steps P1 and P2 of the payment phase, Section 3.3.2.2). Hence, privacy is well protected in MEP.

3.4.2 Computational Overhead of MEP

This section analyzes the computational overhead for a payment transaction in MEP. The computational cost for a payment transaction can be evaluated in the following three aspects.

3.4.2.1 Token Generation and Verification

Let C_{h3} be the computational cost for executing the H_3 hash function for token generation and verification. Assume that **U** obtains N tokens from **O** and pays J ($J \leq N$) tokens to **P** for a mobile application (i.e., in one payment transaction). **O** and **U** generate N tokens by executing the Generate-Token algorithm in Steps W2 and W3, respectively, where the hash function H_3 is executed N times with computational cost $2C_{h3}N$. During a payment transaction, **P** verifies J tokens sent from **U** in Steps P2 and P4 of the payment phase by executing the H_3 function J times with the computational cost $C_{h3}J$. In the deposit phase, **P** sends the last token of the received J tokens to redeem token from **O** in Step D1, and **O** runs the hash function H_3 $N-J$ times to verify the last token (see Step D2), that is, the computational cost is $C_{h3}(N-J)$. The total computational cost for token verification in a payment transaction is $C_{h3}N$. Thus, the total computational cost for token generation and verification is $3C_{h3}N$.

As mentioned in [46], H_3 is a light-weight function with low computational cost, which is about 100 times faster than the RSA signature verification and about 10,000 times faster than the RSA signature generation. Usually, the number N of tokens required in a payment transaction is 50 to 50,000, and the computational cost is 150 C_{h3} to 150,000 C_{h3}, which is considered to be reasonably smaller than what RSA needs.

3.4.2.2 Message Encryption and Decryption

Let C_m be the computational cost of the symmetric key algorithm AES (applied in MEP) for message encryption and decryption. There are three messages (including the InitPaymentTrans, TokenInfo, and PurchaseInfo messages), encrypted in the payment phase. Suppose that there are P ($1 \leq P \leq N$) payments processed in the payment phase, where P messages (including InitPayment and payment messages) are encrypted. The deposit message is encrypted in the deposit phase. Each message requires two operations (encryption and decryption). Thus, the total computational cost for message encryption and decryption in MEP is $2(4 + P)\ C_m$.

The computational cost of symmetric key cryptography is less complex than public-key cryptography [50]. Our design has the lower computational cost then what is needed than traditional public-key cryptography.

3.4.2.3 Symmetric-Key Update

In MEP, we update three symmetric-keys, $k_{u\text{-}o}$, $k_{o\text{-}p}$, in the withdrawal phase (see Equations (3.6), (3.7), (3.8) and (3.9)), and $k_{u\text{-}p}$ in the payment phase (see Equations (3.10) and (3.11)) by running the $H_2(\hat{e}(a \cdot b, c))$ function, where $a \cdot b$ and c are the input parameters, and " $\cdot$ " is the scalar multiplication operation. Let C_k be the computational cost of $H_2(\hat{e}(a \cdot b, c))$. Computing the $H_2(\hat{e}(a \cdot b, c))$ function requires to execute a hash function H_2, a bilinear pairing function $\hat{e}$ and a scalar multiplication. Let C_{h2} be the computational cost for the hash function H_2, $C_{\hat{e}}$ be the computational cost for the bilinear pairing function $\hat{e}$, C_{G_1} be the computational cost for the scalar multiplication. Then C_k can be expressed as $C_k = C_{h2} + C_{\hat{e}} + C_{G_1}$. As noted in [28], the $C_{\hat{e}}$ is much larger than the C_{h2} and the C_{G_1}, and we have $C_k = C_{h2} + C_{\hat{e}} + C_{G_1} < 3C_{\hat{e}}$. The study [8] has shown that the computation of $\hat{e}$ can be completed within 20 milliseconds on a Pentium III 1 GHz machine, and the computation of $H_2(\hat{e}(a \cdot b, c))$ can be completed within 60 milliseconds, which is reasonably low and suitable for a

resource-restrained mobile device.

3.5 Conclusions

This chapter proposed a secure Mobile Electronic Payment (MEP) platform for the mobile commerce (m-commerce) over wireless mobile networks. In this platform, we take advantage of the emerging the ID-Based Cryptography which eliminates the necessity of certificates commonly required by other public key cryptography. Moreover, since ID-based cryptography can establish the shared key between two parties without additional message exchanges, symmetric key cryptography can be still used effectively, leading to significant computational cost. Our study shows that our MEP platform satisfies the requirements of secure trading (such as avoidance of overspending and double spending, fairness, user anonymity, and privacy) and has low computational cost. We expect that our MEP will provide a viable trading model for the future mobile applications and play an important role in the emerging m-commerce industries.

CHAPTER 4 A Callback Mechanism for Private Telecommunications Network

4.1 Introduction

A Private Telecommunications Network (PTN) is a telecommunications network that has its own number plan other than the public E.164 numbering used in the PSTN [14]. Examples of PTN are enterprise telephone systems and Voice over IP (VoIP) networks. PTNs have been widely deployed in companies and the Internet/Intrant. And they connect to the PSTN through the Private Branch Exchange (PBX) in a telephony-based PTN or the VoIP Gateway (VPG) in an internet-based PTN. Without loss of generality, we consider the Internet-based PTN. A phone call between a VoIP user and PSTN user consists of two connections: a VoIP connection between the VoIP user and the VPG, and a PSTN connection between the VPG and the PSTN user. A common limitation of VoIP is that the VoIP users, such Skype users, are usually not assigned with E.164 numbers. When a PTN user initiates a call to a PSTN user, it is not possible to use the PTN calling party's identity as the caller ID; instead, an E.164 number of the serving VPG is used. Therefore, the PSTN called party cannot retrieve the caller ID and call back to the PTN calling party later. As a result, although it is easy for a Skype user to initiate a call to a PSTN user, it is not possible for a PSTN user to call a Skype user if the Skype user does not subscribe SkypeIn service (i.e, is not assigned an E.164 number). We proposes a callback table approach to resolve this problem and presents a mechanism to allow a calling party outside a PTN to call back a user within the PTN without knowing the PTN user's extension number. We describe how the call-out information is stored into and retrieved from the callback table in the PBXs or VoIP gateways to provide callback

service.

4.2 The traditional PTN call process

When a PTN user attempts to make a call to a PSTN user (this procedure is referred to as PTN *call origination*), the call is first set up to the target VPG, which is located at the same tariff zone as the called PSTN user (Step 1 in Fig. 4.1) using VoIP signaling, such as *Session Initiation Protocol* (SIP), H.323, or Media Gateway Control Protocol (MGCP). The VPG then chooses an available leased line (among the *n*-leased-line pool; Step 2 in Fig. 4.1), and connects the call to the PSTN switch (Step 3 in Fig. 4.1) using *Signaling System Number 7* (SS7) [39].

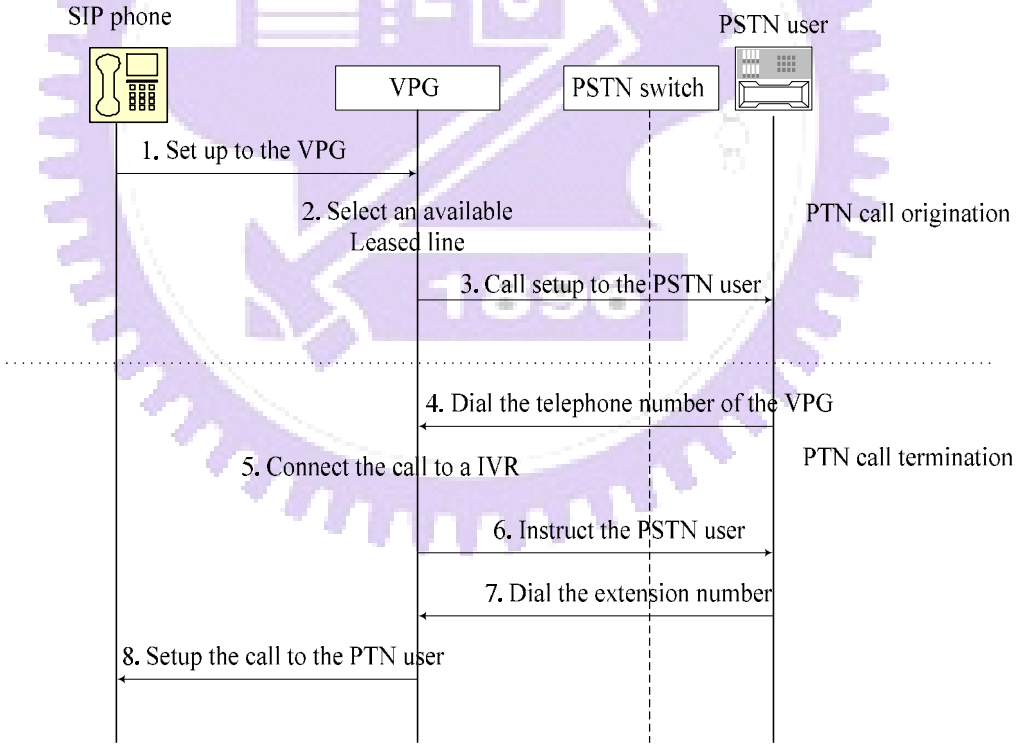


Fig. 4.1: A call origination from an IP user to a PSTN user.

One major problem with a PTN is that if a user in the PTN does not have an E.164 number, the user cannot be reached directly from the PSTN. When a PSTN user attempts to

call a PTN user (this procedure will be referred to as *PTN call termination*), he or she first dials the telephone number of the VPG (i.e., the number is one of the ***n*** leased lines owned by the VPG). The PSTN switch serving the PSTN caller sets up the call to the VPG using SS7 (Step 4 in Fig. 4.1). Then the VPG requests the PSTN caller (e.g., through an *Interactive Voice Response* (IVR) system) to input the extension number of the called PTN user (Steps 5~7 in Fig. 4.1). After the PSTN caller has input the extension number (e.g., through DTMF dialing [37]), the VPG sets up the call to the called PTN user through VoIP protocols (Step 8 in Fig. 4.1). In this call termination procedure to a PTN user, the PTN user is accessed indirectly through two-stage dialing. That is, the PSTN user must first dials the number of the VPG of the PTN, and then dials the extension number of the called PTN user.

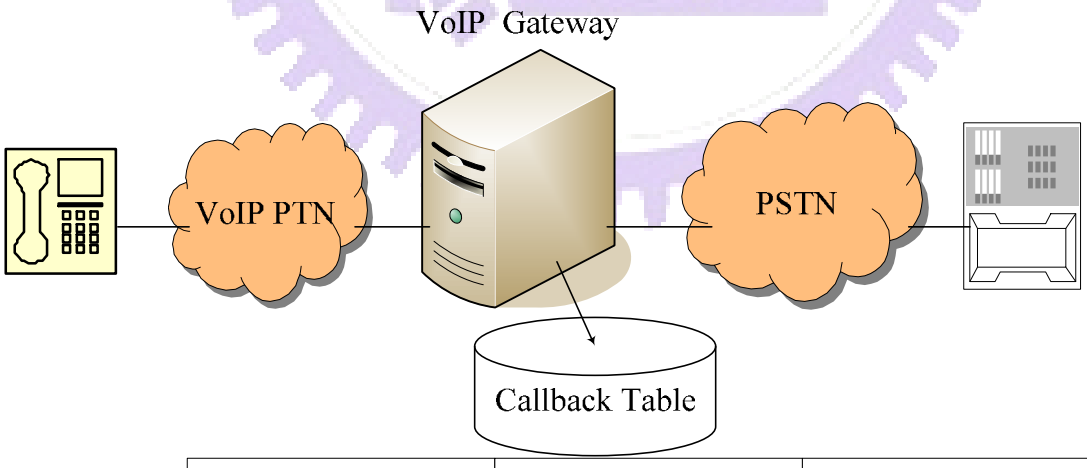
When setting up a call, the SS7 Initial Address Message (IAM)[39] can deliver the caller's telephone number (i.e., the ***caller ID)*** to the called party. The caller ID can be automatically stored in the called party's telephone device (e.g., a mobile phone with address book). The called party can then call back without dialing the caller's telephone number. Another limitation of PTNs is that from the viewpoint of the PSTN, a PTN user without an E.164 number does not have a caller ID. When a PTN user reaches a PSTN user, the PTN's identity cannot be carried by the SS7 IAM. Instead, the PBX or VPG's telephone number is used as the caller ID. Consequently, the PSTN user cannot utilize the call back service to reach the PTN user later. This limitation may cause frustration to the PSTN user, especially if the PSTN user is not available when the PTN user calls this PSTN user. The PSTN user will not be able to call back based on the caller ID because he or she does not even know who made the call previously. When the PSTN user does call back the caller ID of the missed call, the user will be connected to the PBX/VPG, but he or she may be unable to reach the PTN user because the PTN user's name or number is unknown. Therefore, when a PBX/VPG dials out for a PTN call origination, the caller ID is usually blocked out in order not to causing confusion to the called party. In this case, unanswered calls cannot be returned, and important

calls may be lost.

If a PTN user does have a public E.164 number, one-stage dialing is possible. A well-know example is the Direct Inward Dialing (DID) service provided by PSTN operators. When a PSTN user makes a call to a PTN user by dialing the PTN user's DID number, which is an E.164 number. The call will be routed to a designated PBX or VPG along with a signaling indicating the DID number dialed. The signaling can be an ISUP IAM, a Q.931 setup, or Caller Line Identifier supported by CLASS [38]. The PBX or VPG retrieves the DID number dialed without picking up the call, and forwards the call the PTN user called. However, if a PTN user has no public E.164 number, no one-stage dialing scheme has ever been presented.

To resolve the above issue, this paper presents a mechanism to allow a calling party outside a PTN to call back a user within the PTN without knowing the PTN user's extension number.

4.3 The PTN Callback Mechanism



PTN Caller's ID	Called party's tel no.	Leased line no.
11	5752121	5720250
12	5755476	5720251

Fig. 4.2: A VPG with a callback table

The PTN callback mechanism is established by introducing a *callback table* in the VPG. Fig. 4.2 illustrates an abstract VPG architecture with a callback table. The callback table stores callback tags. Each callback tag represents a record of PTN call origination; it consists of at least three fields: the PTN caller's identifier (including the user's private phone number, SIP URI, or any ID representation), the PSTN called party's telephone number, and the leased line used for the call.

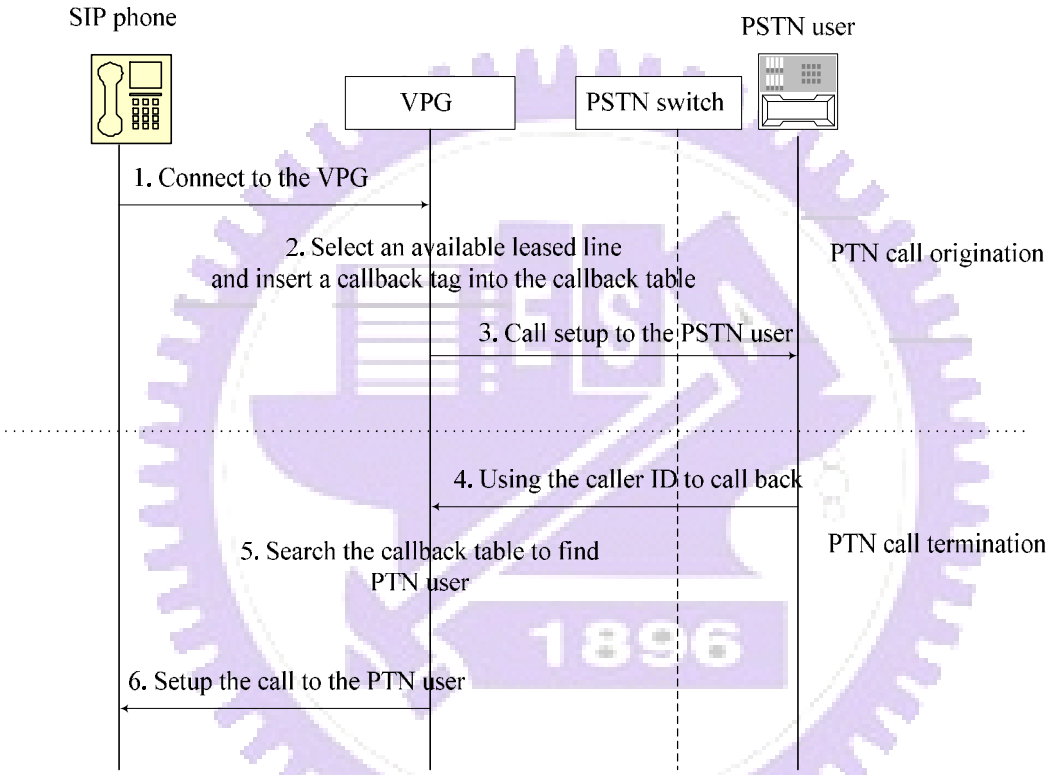


Fig. 4.3: The call flow of our callback mechanism

For a PTN call origination, the VPG uses the data pair (the PTN user identifier, the PSTN called party's number) as the key to search the callback table (Step 2 in Fig. 4.3). If there is a matched tag, meaning the PTN user has called the PSTN user before, the leased line in the matched record is used to set up the call (Step 3 in Fig. 4.3). If no matched tag is found, the VPG tries to find an unused leased line to set up the call and insert a callback tag (the PTN user identifier, the called party's number, leased line) in the callback table. An unused leased

line is a line that has not been used by other PTN users to call the same PSTN user. If the PSTN user has been called by *n* different PTN users, no unused leased line can be found. This problem will be referred to as the leased line collision problem. When this happens, the call cannot be served by this VPG without deleting an old callback tag. In this situation, we route the call to a nearby VPG, and the new VPG tries to connect the call. Since a VPG in a different tariff zone is used, the call may be charged at a higher rate.

Note that in Step 2 the leased line selected may be busy, and thus the line cannot be used to call out to the PSTN if analog signals are used between the VPG and PSTN. In our design, we assume that SS7 or Q.931 is used between the VPG and PSTN. In this case, there is no fixed binding between a circuit and a telephone number. Even an ongoing VPG originating call is using a certain caller ID, the caller ID can still be carried in another VPG originating call request (IAM for SS7 or CONNECT for Q.931).

In a PTN call termination, when the VPG receives a call from the PSTN (Step 4 in Fig. 4.3), the VPG received the PSTN calling party's number and the called leased line number, and use the data pair (the PSTN calling party's number, the leased line called) as the key to search the callback table (Step 5 in Fig. 4.3). If a matched tag is found, the VPG sets up the call to the PTN user (Step 6 in Fig. 4.3). If no matched callback tag is found, meaning no PTN user has used the leased line to call the PSTN user, the call can be routed to an IVR for further assistance. In this way, a PSTN user can call back a PTN user by one-stage dialing, and the call is charged as a local phone call since the VPG and the PSTN caller are located at the same tariff zone. Note that when a PSTN user dials a SkypeIN number, the call is often a long-distance or even an international call.

4.4 The Analytic Method

The feasibility of this callback method supported by the VPGs depends on how often leased line collisions would occur. In our analysis, we assume a VPG only has *L* leased lines and the number of PSTN users served by the VPG is *G*. In addition, the number of PTN users calling a PSTN user, *k*, is normally distributed with mean $\overline{K}$ and standard deviation σ. The calls between a PTN user and a PSTN user form a Poisson process with rate l. The service time of a voice call is assumed to be exponentially distributed with mean $1/m$. Using Erlang's B formula in Equation 4.1, we can obtain the fraction of time that all leased lines on the VPG are busy, P_L, where E_b is the Erlang B value and can be expressed in Equation 4.2. To provide an acceptable call blocking probability (P_L < 1%), *L* has to be large enough, and its value can be determined from Equation 4.1.

$$P_L = \frac{(E_b)^L / L!}{\sum_{i=0}^{L} (E_b)^i / i!} \tag{4.1}$$

$$E_b = \frac{(\overline{K} G l) / m}{3600} \tag{4.2}$$

When a PSTN user has received calls from *L* PTN users, and receives a call from a new PTN user, the VPG encounter a leased line collision problem because all *L* leased lines have been used. In this case, the call is re-routed to a nearby VPG. For each PSTN user, the percentage of his or her correspondent PTN users who may experience leased line collision problems is denoted by P_r, and it can be obtained in Equation 4.3.

$$P\_r = P(k > L) \tag{4.3}$$

Since *k* is normally distributed, P_r can be expressed as

$$P\_r = 1 - P(k \le L) = 1 - F\left(\frac{L - \overline{K}}{s}; 0,1\right) \tag{4.4}$$

where $F\left(x;\overline{K},s\right)=\frac{1}{s\sqrt{2p}}\int_{-\infty}^{x}e^{-\left(\frac{\left(u-\overline{K}\right)^2}{2s^2}\right)}du$

Assuming the calls to a PSTN user are uniformly distributed among his or her correspondent PTN users. Let the probability that a call from a PTN user to a PSTN user encounter a leased line collision problem and needs to be re-routed be denoted by T_r; it can be expressed in Equation 4.5.

$$T\_r = \frac{\sum_{i=1}^{\infty}\left\{i\times\left[P\_r(L+i-\frac{1}{2})-P\_r(L+i+\frac{1}{2})\right]\right\}}{\overline{K}} \qquad (4.5)$$

4.5 The Numeric Results

In general, *G*, the number of users in a PSTN tariff zone, can be as large as millions. When *G* is large, the number of leased line required on the VPG is also large, and as a result, the leased line collision problems may not occur. In our experiments studying the leased line collision problems, *G* varies in the range of 1000-8000. For each PSTN user, the number of correspondent PTN users, *k*, is assumed to be normally distributed. It is clear that as *k* increases the leased line collisions occurs more often. However, its mean value, $\overline{K}$, and its standard deviation, σ, are difficult to determined because no such data have been collected. In our experiments, $\overline{K}$ varies in the range of 50-300 and the standard deviation equals to $\frac{1}{4}\overline{K}$. Considering the number of different PTN users calling the same PSTN user, we believe 300 is large enough to represent the worst case realistic situation. The mean service time of a call $(1/m)$ is assumed to be 60 seconds and the arrival rate of voice calls (l) between two corresponding PTN and PSTN users varies in the range of 0.005-0.012 calls/hour. The number leased required on the VPG can be determined by Equation 4.1; *L* is chosen to be large enough so that $P_L \leq$ 1%.

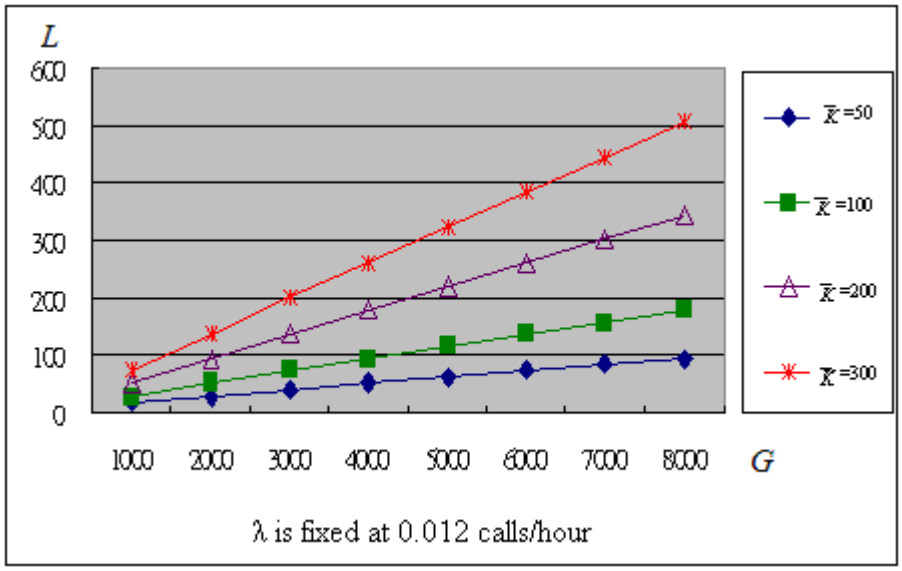


Fig. 4.4: The number of leased lines (L) required on a VPG.

Fig. 4.4 depicts the number of leased lines (L) required on the VPG. The arrival rate of voice calls (l) is fixed at 0.012 calls/hour. The results indicate that L increases linearly as $\overline{K}$ and G increase, i.e., a larger $\overline{K}$ or G indeed needs more leased lines. Note that $L=\overline{K}$ when G is in the range of 4000-5000. When k of a PSTN user is less than L, the PSTN user's corresponding PTN users experience no leased line collision. This indicates the leased line collisions are very common when G is less than 4000.

Fig. 4.5 depicts the effects of $\overline{K}$ and G on P_r, the percentage of PTN users who call the same PSTN user and experience leased line collision problem. The arrival rate of voice calls (l) is fixed at 0.012 calls/hour. The results indicate that P_r increases as $\overline{K}$ increases, i.e., a larger $\overline{K}$ causes more leased line collisions. However, the difference between the cases $\overline{K}$=200 and 300 is very small. For all different values of $\overline{K}$, P_r decreases as G increases, and drops to zero as G is larger than 8000. The results are also consistent with the results in Fig. 4.4 in that P_r decrease rapidly when G reaches around 4000.

Fig. 4.6 depicts the effects of $\overline{K}$ and G on T_r, the probabilities that a PTN originating call needs to be re-routed to a nearby VPG. The results indicate that similar to P_r, T_r decreases as G increases and as $\overline{K}$ decreases. It is interesting to note that T_r drops even more rapidly than P_r does as G increases. T_r drops to nearly 0 when G is larger than 7000. Since the number of PSTN user in a tariff zone, G, is in general much larger than 8000, we conclude that leased line collision problems will not occurred if l is more than 0.012 calls/hour.

.

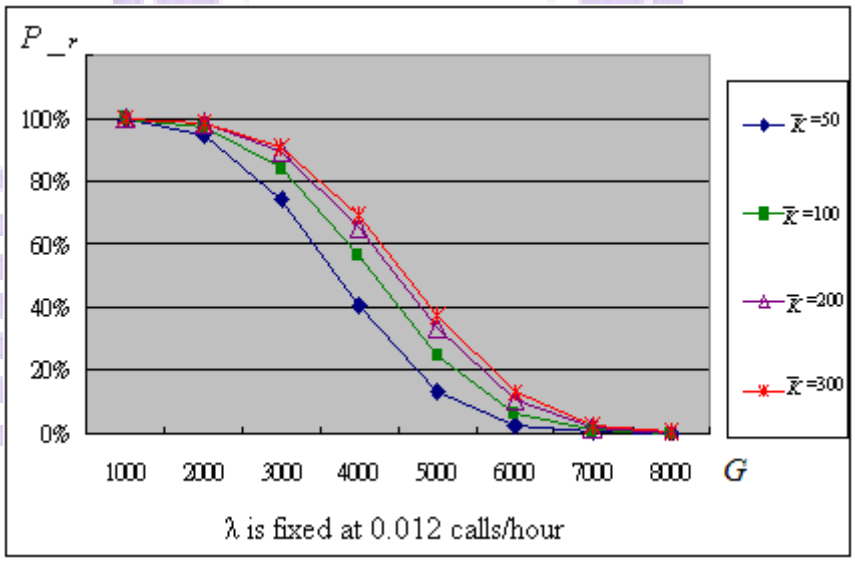


Fig. 4.5: P_r at a high call arrival rate

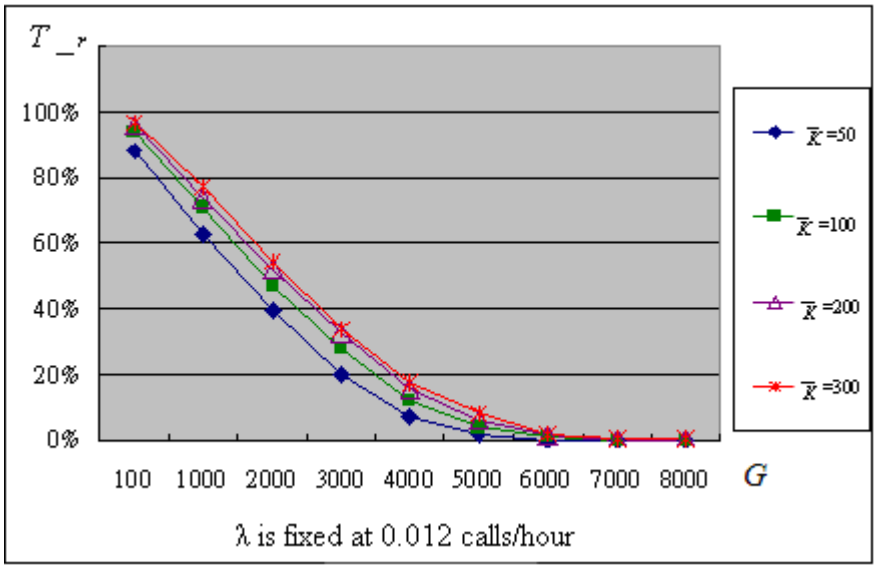


Fig. 4.6: T_r at a high call arrival rate

However, at the initial stage after the VPGs are installed, the users may call less often. There we studied the effects of call arrival rates on P_r, and T_r. The average number of correspondent PTN users of each PSTN user, $\overline{K}$, is assumed to be 100. Fig. 4.7 depicts the number of leased lines, *L*, required on a VPG when $\overline{K}$ is fixed at 100. The results indicate that the leased line *L* increases as l and *G* increase. Note that when the call arrival rate is as low as 0.005 calls/hour, *G* has to be as large as 11000 so that *L* equals to $\overline{K}$ and thus P_r starts to drops rapidly as *G* increases. However, when l is high, say 0.025 calls/hour, the transition position is when *G* equals to 2500.

The results in Fig. 4.8 also indicate when the call arrival rate is low, a VPG has less leased lines, we need more PSTN users to bring P_r down. For example, when the call arrival rate is 0.005 calls/hour, *G* has to be larger than 16000. Similar results can be found for T_r, as shown in Fig. 4.9. When the call arrival rate is 0.005 calls/hour, *G* has to be as large as 15000 so that T_r is close to 0, i.e., almost no calls are re-routed. However, at the initial stage after the VPGs are installed, we expect neither l nor *G* is large. Therefore, the system has to be built-ahead with capacity buffer for future expansion and for reducing the leased line

collision problem.

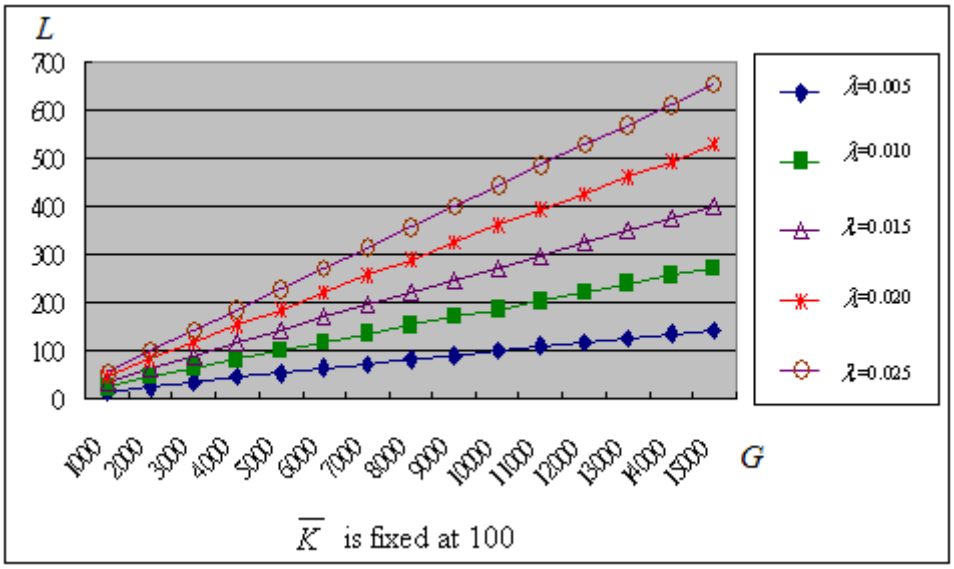


Fig. 4.7: The number of leased lines required on the VPG when $\overline{K}$ is fixed

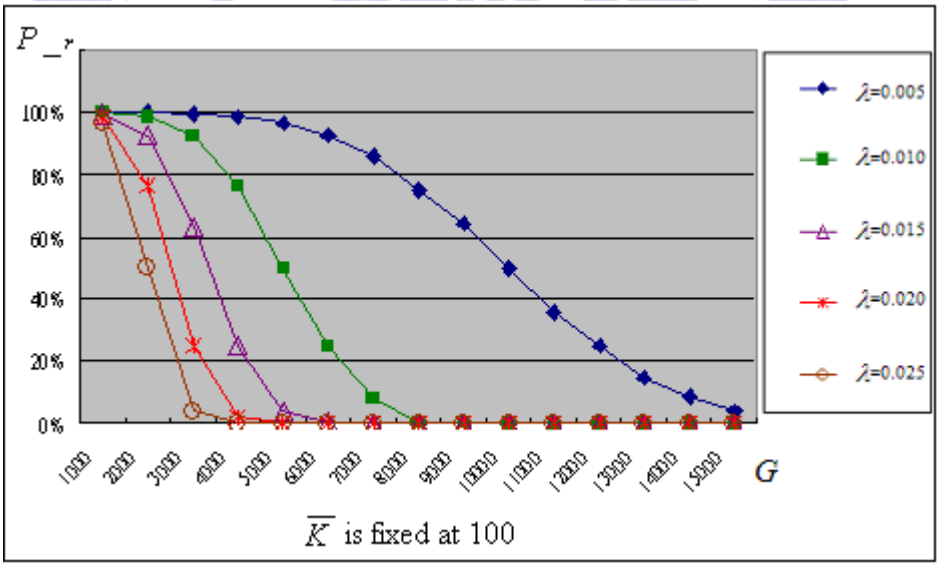


Fig. 4.8: The effects of call arrival rates on P_r

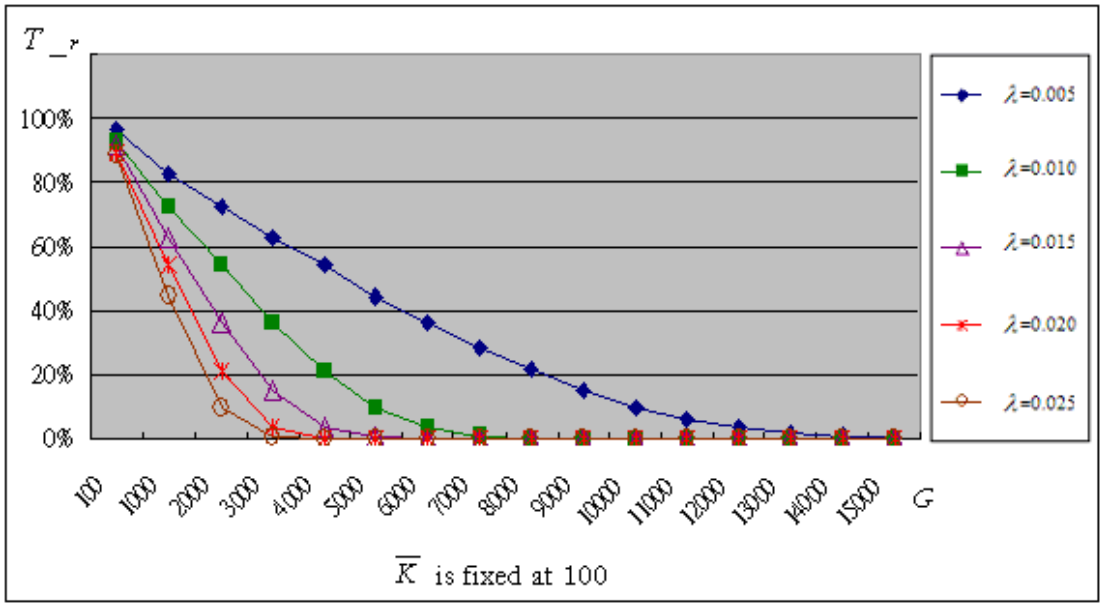


Fig. 4.9: The effects of call arrival rates on T_r

4.6 Conclusions

This chapter investigates the difficulty of PTNs in providing callback service. We present a callback table approach to resolve this issue. We describe how the call-out information is stored into and retrieved from the callback table in the PBXs or VoIP gateways to provide callback service. This is a plug-in solution that does not affect the existing call setup message flow. An analytic model has been developed to evaluate the lease line collisions issue. The numeric results are consistent with real world experience. Our results show that the callback mechanism performs better when more traffic and more PSTN users are served.

CHAPTER 5 Conclusions and Future Work

In this dissertation, we investigated design issues on the intelligent services for VoIP and PSTN/PLMN networks. This chapter summaries our study and contributions, and briefly discusses the future directions.

5.1 Summary

In this dissertation, we discussed intelligent services for VoIP and PSTN/PLMN networks design issues. In Chapter 2, an integrated mobile prepaid service was presented. We described the charging issues of an integrated GSM and GPRS prepaid service, where a single prepaid account provides the user both voice and data services. Based on the CAMEL network architecture, the call setup and charging procedures for GSM and GPRS have been illustrated. We propose an NMC algorithm to reduce the probability of terminating both on-going voice and data calls, where no more new calls are admitted when the user credit is below a threshold. An analytic model has been developed to evaluate the performance of the NMC algorithm. Computer simulations have also been used to verify the analytical results. The numeric results indicate that the forced termination probability can be significantly reduced by choosing an appropriate threshold of the user credit. In addition, the forced termination probability of voice calls slightly increases as the call pattern of data calls becomes irregular. As the need to rapidly roll out new services in mobile networks, our model can be easily extended to accommodate different real-time services. Our analytic method could provide guidelines helping the operators to generate higher revenues.

In Chapter 3, we proposed a secure Mobile Electronic Payment (MEP) platform for the mobile commerce (m-commerce) over wireless mobile networks. In this platform, we take advantage of the emerging the ID-Based cryptography which eliminates the necessity of certificates commonly required by other public key cryptography. Moreover, since ID-based

cryptography can establish the shared key between two parties without additional message exchanges, symmetric key cryptography can be still used effectively, leading to significant computational cost. Our study shows that our MEP platform satisfies the requirements of secure trading (such as avoidance of overspending and double spending, fairness, user anonymity, and privacy) and has low computational cost. We expect that our MEP will provide a viable trading model for the future mobile applications and play an important role in the emerging m-commerce industries.

In Chapter 4, we described a limitation of PTNs -- lack of the support for callback service. Based on VoIP SIP or MGCP protocols, we propose a callback table approach to resolve this issue. We describe how the call-out information is stored into and retrieved from the callback table at the PBXs or VoIP gateways to provide callback service. This is a plug-in solution that does not affect the existing call setup message flow. An analytic model has been developed to evaluate the leased line collision problem. The numeric results are consistent with the real world experience. The numeric results indicate that the callback method performs better as there is more voice traffic served by a VoIP gateway. Therefore at the initial stage, a VPG has to be built-ahead with capacity buffer for future expansion and for reducing the leased line collision problem.

5.2 Future Works

Based on the research results in this dissertation, the following design issues on intelligent services for VoIP and PSTN/PLMN networks can be investigated further.

Integrated prepaid services with priority algorithm

In this dissertation, we addressed the NMC algorithm to reduce the probability of terminating both on-going voice and data calls. To provide a better service to the users, the method making voice call (or data call exclusive) with higher priority to be served while

credit is not much left, can be applied to make the remaining credit more reasonable. This can be extended to integrated prepaid services.

Integration with 3GPP online charging system

3GPP has defined a common charging architecture and framework for GSM or UMTS PLMN. In particular, they define the Online Charging System (OCS) to support online charging for bear, value-added service and IMS services. How to move the IN prepaid process to use the Session Based Charging Function (SBCF) of OCS is an important issue that needs to be investigated. Furthermore, there are strong demands from the customers to request operator that they need support comprehensive charging mechanism to provide prepaid/postpaid convergent business model. The convergent charging system issue can be further discussed.

Appendix A Customer Care and Billing System for Telecommunication

To achieve the rapid service rollouts that the market demands, well integrated, easy to modify BSS(Business support system) systems are an absolute must. BSS systems are a key contributor to operators that can manage their subscribers and provide the potential function for daily operation's work. The management of any telecommunication business requires many elements, which can be generally put in the following three main categories:

- The first is the entity of management, which is the telecommunication company. The operator needs to obtain the business license, and then formulate the company's direction, goals, and marketing strategies. For example, some companies focus on providing the best customers care, some employ the most professional perspective to guide customers into using telecommunication services, whereas some just attract customers with low price.
- The second is the actual network devices (mobile phones, base stations, switches, routers, and gateways) handled by the service provider that form the network management service. This is critical to the quality of communication. Services can only be enhanced with excellent base facilities, and this is where most funding goes.
- The third is easily ignored by investors but is the pivotal element that connects the above two: it is CCBS (Customer Care & Billing System) that is customer-oriented. CCBS is in the frontline handling customers, and its quality is an important indicator to suggest how strong the customer loyalty might be.

Today, telecommunication operators do not face customers who only receive primitive service. Customers have changed from being the passive role to active, and operators need to

focus on customers' feeling. The traditional belief of targeting at the public view has been rapidly modified to become targeting at differentiated products and services. It is important to satisfy all kinds of customers and make them feel that the service is tailored for them. It is based on their benefits and interests. No matter who they are, themselves are actually aware of being the most important and can get the simplest, the most direct and personalized services. Thus, the importance of a frontline customer service system is gradually increasing.

According to the marketing strategies and methods, a billing system contains at least 4 characteristics:

- From the perspective of operator strategies, a billing system is the basis for determination of the prices, packaging, customer demands, and new business development.
- From the perspective of business management, a billing system is the basis for new products and services, increasing interactions with customers, and developing and evaluating niche markets.
- From the perspective of market responses, a billing system allows customers to determine and seek their own needs, provides instant interactions with customers, and responds to customers who speak different languages and require different products and services.
- From the perspective of marketing management, a billing system is able to create different billing combinations and processing models, accurately calculate and provide customers their billing statements, and offer businesses correct financial reports to corporations.

Thus, customer care and billing system (CCBS) have a major influence on telecom providers' operations and is not weaker than the technology of wireless communication.

A.1 The Scope of CCBS

CCBS's scope is rather broad, and it includes the life cycle from applying for a service to the termination of a service. Fig. A.1 shows the framework of CCBS' functions:

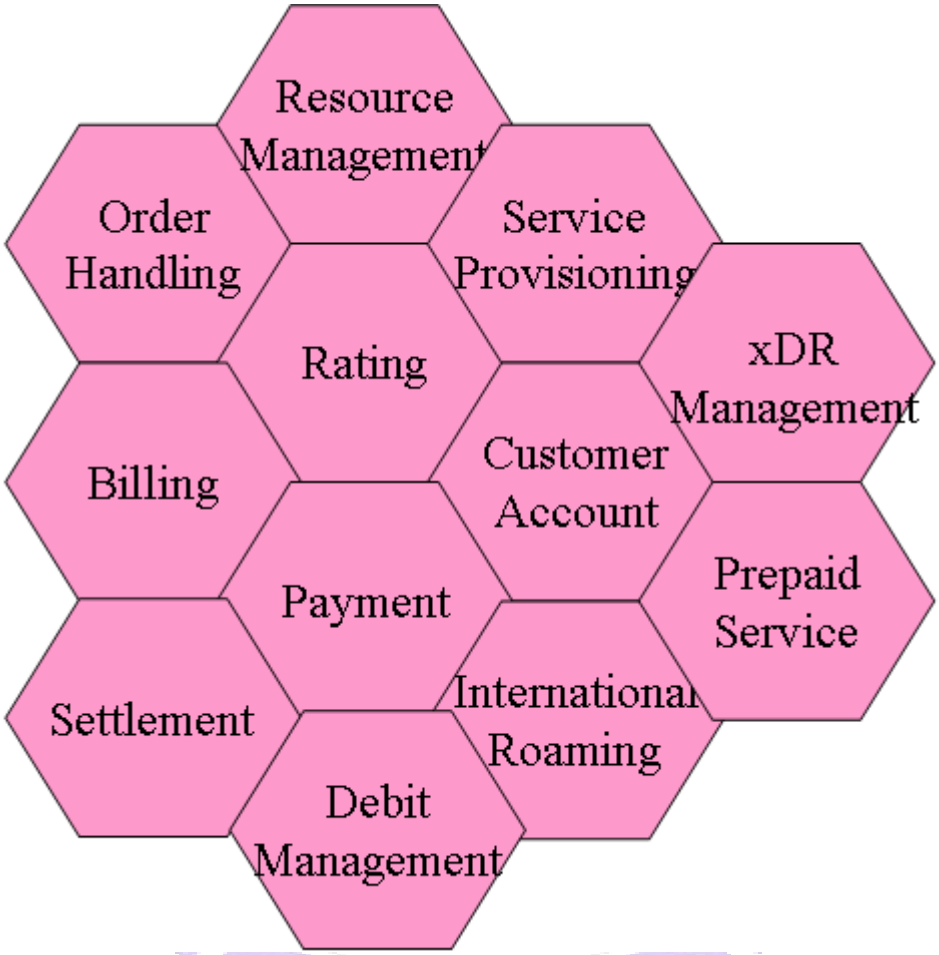


Fig. A.1: Framework of CCBS' functions

(1) Resource Management: The management of the resources that operators provide to their customers, including SIM cards, phone numbers, and cell phones.

(2) Order Handling Management: The management of the interactions between operators and their customers such as changing services and terminating or restoring halted services due to non-payment. The process of service-providing can also be controlled.

(3) Service Activation: Handles orders and configures switching centers or network

elements in order to provide the needed services.

(4) xDR Management: Collects the call detail records in switching centers or network components. After decoding, filtering, and processing through business logic, output files are prepared based on the follow-up modules such as rating and settlement.

(5) Rating Process: Calculates monthly charges based on customers' plans or calculates charges based on call detail records.

(6) Billing Process: Collects all the charges and prepares a statement before making a bill for the client.

(7) Payment Process: Inputs the information of the payments made by customers via different methods. Conducts appropriate measures for late payments, insufficient payment, excessive payment, or redundant payment.

(8) Customer Account Management: Conducts adjustments, changes, or striking a balance for abnormal accounts.

(9) Debit Management: Conducts tasks such as debt collection, service halting, or service termination for clients who fail to make complete payments.

(10) Settlement Management: Provides settlement services for Internet service providers, content providers, distributors, and other parties.

(11) Roaming Management: Handles the billing information of customers using roaming services in or out of Taiwan.

(12) Prepaid System: Provides instant billing for customers who use prepaid services instead of monthly bills.

A.2 Service Order and Activation System

Service order and activation system is the entrance from which operators serve their customers and their relationship begins. The processing and activation of different customer

applications such as getting a new phone, canceling an account, halting or restoring the service, additional services, changing SIM card/phone/number, additional numbers, changing the service area, adding/changing services (roaming, data transfer, fax, voice mail, SMS, and WAP services), changes of customers' basic information (name, ID, address), changes of services (e.g., billing rates and discounts), GPRS services, and paid content services. Fig. A.2 shows the scope of Service Order and Activation System, and its important modules (1. resource management, 2. customer management, 3. order management) and five other functions including customers' self services, analysis, and statistics report.

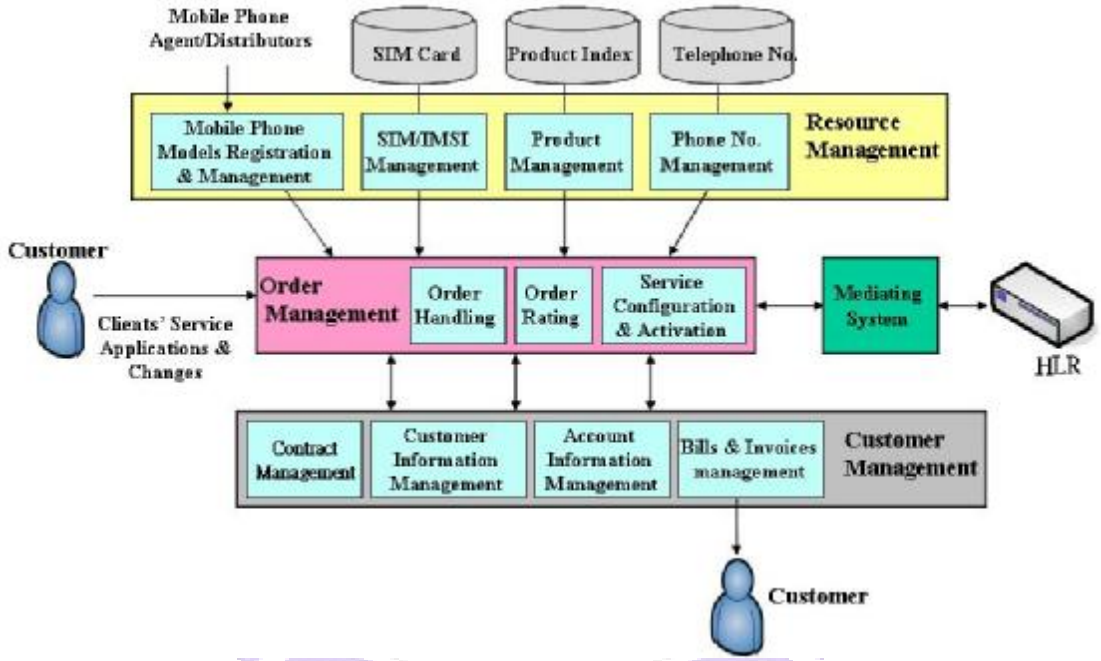


Fig. A.2: Scope of service order and activation system

A.2.1 Resource Management

To meet the managerial needs, telecomm operators need to make the following preparations before providing services:

Allocation of phone numbers: Each HLR stores a fixed set of phone numbers, which are allocated to different operation stations. The numbers can be put in the following categories: 1. Automatically provided number: the number that is randomly provided for a new client; 2.

Selected number: made available to clients at a service counter. The numbers include free or paid ones; 3. Reserved numbers: reserved for special purposes. However, since different service counters enjoy different amount of sales volumes, the numbers often need to be redistributed. Therefore, service providers are working on the planning of the "common pool," in which the numbers are allocated based on the order of applications instead of being put in fixed service counters. This allows better efficiency on the usage of numbers, but the management also becomes more complicated.

Manufacturing and delivering of SIM cards: Produced by SIM card manufacturers, brand new SIM cards are raw cards that cannot be used due to insufficient information. They must first be "personalized," which is writing data such as ISMI, Ki, PIN and PUK into its memory. At this time, the SIM card does not contain MSISDN. The personalized SIM cards will be sent to different service locations, and to prevent them from being illegally copied, each SIM card has a unique card number, and the special card numbers in each HLR range are not interchangeable. Therefore, if each service counter has different ranges, they need to be given the corresponding card numbers. In order to solve the problem of preparing a huge amount of SIM cards with different ranges, service providers are working on "Free Numbering," which only requires the matching between dynamic numbers and SIM cards. This, however, requires the management of additional HLR databases.

Providing mobile phones: In the competitive market, in order to increase customers' willingness to make the purchase, vendors will often sell the bundle that contains both a phone number and a mobile phone at a discounted rate instead of just selling phone numbers. The cost of mobile phones are covered by both vendors and distributors based on their strategic alliance.

Product packaging: In order to distinguish the market, appropriate services are provided to different clients, and operators would package their products differently, including high monthly account with low charges per call or low monthly account with high charges per call,

or selling a prepaid bundle that includes a cell phone, a phone number, and a SIM card all together.

In short, the system needs to provide the functions such as the management of card manufacturing, the management of the inventory of cards and mobile phones, the management of the sales of phone numbers, the management of the price of services, the management of mobile phones, and the management of prepaid cards.

A.2 2 Customer Management

In this section, we will use the process of applying for a mobile phone as an example to describe the functions that need to be included in Customer Management. Usually, a customer needs to first go to a telecom operator's service point to fill out an application, and the Service Order and Activation System needs to be capable of processing a request. First of all, a service agent needs to verify the client's history of applications and payments. If the client is on the "black list" or shows a bad history in making payments, the representative will reject the application. If the application is approved for a customer who is a VIP that has many phone numbers or has had a long history with the service provider, this customer will get to enjoy discounts. Thus, the system needs to be able to manage customers' information and keep each customer's transaction record. Afterwards, a customer begins to select the services provided. In terms of providing phone numbers, they can be allocated automatically or chosen by the customer out of the available numbers. Operators will also sell phone numbers with mobile phones together at a discounted rate. In terms of the types of services, different options will be given such as GSM voice mail, GPRS data service, GSM+GPRS combinations, or 3G services. In terms of the combination of rates, a customer can choose a suitable rate based on his or her amount of usage, such as the 88/188 provided by Chunghwa Telecom and the 199/399 by Far EastOne. Also, the customer can also choose supplementary services based on

his or her own needs, such as 3-way calls or restriction of roaming. This may also include value added services such as Multimedia Message Service (MMS) or Personal Ring Back Tone (PRBT), etc.

After a customer has confirmed the services he or she wishes to access, generally, an order will be generated that calculates, manages, and tracks this application. Please refer to Section A.2.3 for the details on Order Handling Management. When relevant charges are generated on an order, the system must decide which charges can be combined to the next bill using the method of "account keeping," and which charges need to be paid instantly (thus the statement will be printed out for a customer who needs to pay at the cashier). After the application is processed, the contract between the customer and the service provider is confirmed, and the customer can choose how he or she wishes to pay for the account in the future such as receiving a monthly statement or making automatic transactions through a financial institution. Thus, the system needs to provide the function of account information management that keeps track of a customer's method of payment, payment address, and billing cycle, etc.

Since a service agent allocates a SIM card to a customer based on the range in which the account is located, the customer can leave with the phone first without staying further. At this time, the service still will not be activated until the service provider configures the relevant network elements via provisioning.

A.2.3 Order Handling Management

As dispicted in A.2.2 regarding customers' different applications that are processed by the system, when an application is confirmed, the system produces an order, and then entries the scope of Order Handling Management. It manages and keeps track of all sorts of applications, verifies whether resources need to be allocated in its life cycle, whether billing is

applicable, or whether provisioning is needed…etc. This is generally done with the method of "finite state machine" If the service needs to allocate resources, the system will change the status to "awaiting corresponding resources" such as the allocation of SIM cards. After allocation is done, the system enters the next status. If the service requires billing, the status of the order changes to "awaiting the execution of price-quotation" and recalls the quotation module. If the quotation system requires cash payment, the status of the order will change to "awaiting payment." It enters the next status only after the customer has made the payment. Similarly, if the service requires the setting of network elements, the system then needs to execute provisioning or configure relevant network elements. Take data service or special service for example, provisioning only requires the configuration of HLR. When applying for value-added services, provisioning requires the configuration of corresponding network elements according to the type of value-added services. For example, MMS service requires the configuration of MMSC, and ring back tone requires the configuration of intelligent network component such as SCP. Relevant services are only activated after provisioning is done, which usually requires only 5 minutes.

A.2.4 Customer Self Service

In order to enhance the quality of service, most operators provide customers the function of self service that replaces going to a service counter. This function allows customers to change their monthly plans, inquire charges, halt or restore services, inquire the numbers on the black list or abnormal billing cycle, and activate or cancel services. Also, the system must provide interfaces for different terminals so that customers can access it via customer service centers, website, SMS, WAP cell phones, or PDAs.

A.2.5 Analysis and Statistics report

The statistical charts of the above mentioned managements show date, month, year, region, types of customers, types of services, types of products, sources of revenues, usage of the network, and other items for analysis.

A.3 Telecommunication Mediation Device

The main purpose of Mediation Device (MD) is to serve as the bridge between CCBS and the components of telecommunication network, allowing CCBS to access network elements thoroughly despite the complexity of the network and greatly reducing the cost of the development and maintenance of the system. MD can be divided into two categories, one being the Provisioning Mediation Device, and the other being the xDR Mediation Device which is used to provide the functions of (3) service activation and (4) management of call detail records in the framework. Fig. A.3 shows the scope of MD.

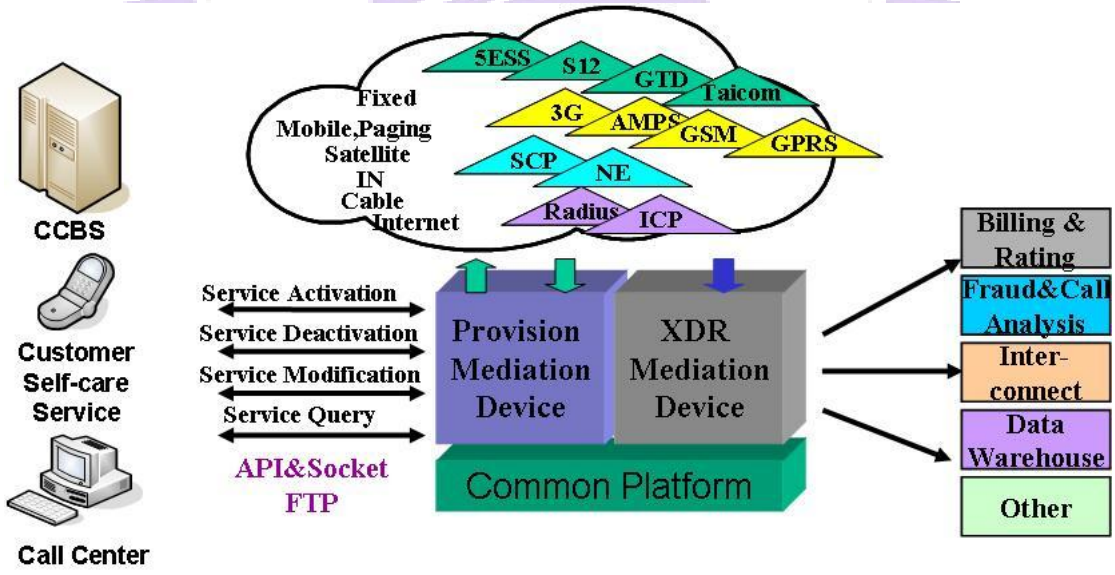


Fig. A.3: Scope of MD

A.3.1 Provisioning Mediation Device

Whether a telecom service can operate normally depends on whether the network elements are correctly configured. For example, in order to enable provisioning in mobile phones, HLR needs to be configured first, and the provisioning of intelligent services require SCP. The main task of Provisioning Mediation Device is to handle the provisioning requests in the Order Handling Management that was mentioned in section A.2.3, which converts the requests into network component instructions in order for network to be configured. Fig. A.4 describes the main modules of Provisioning Mediation Device.

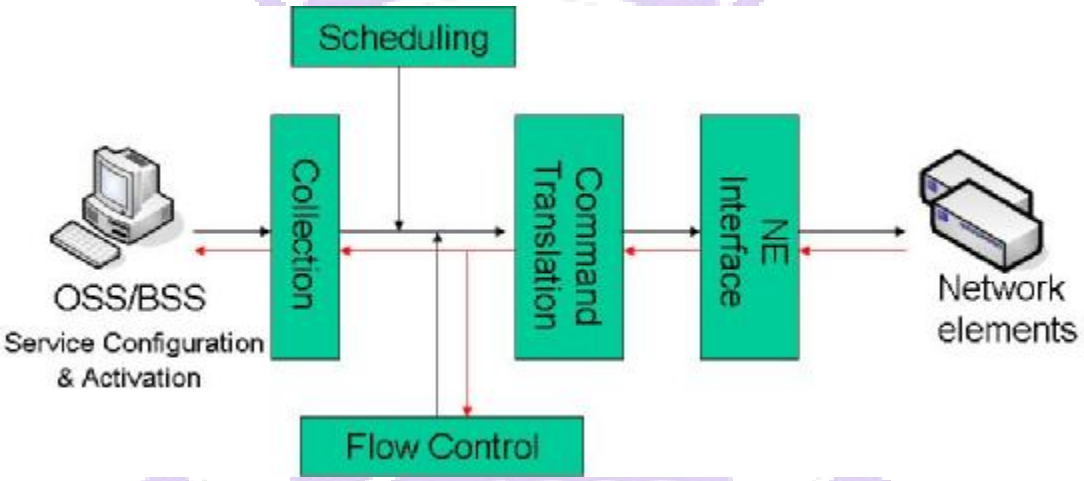


Fig. A.4: Main modules of provisioning mediation device

- Collection: Collects customer services (such as activating or terminating services) or their demands for network configuration that are listed on the order. These demands can be sent via the method of "process-to-process" to be temporarily stored in the related databases.
- Scheduling: Provisioning can be done instantly or booked in advance. Moreover, a phone number maybe configured differently within a short amount of time. A common example is the halting of service due to non-payments, and the customer makes the payment instantly and requests the service to be restored. If there is a

problem with the network, the execution of these two configurations could result in a major error that seriously undermines the customer's rights. Thus, scheduling allows the provisioning to be appropriate and timely.

- Flow Control: The activation of a service may involve the configurations of multiple network elements. For example, the provisioning of prepaid cards requires the configuration of HLR, intelligent network elements, and SCP. However, the order of this process determines the normal operation of the network; thus, a correct flow control is needed. In actual practices, the mechanism of finite state machine is used to control the flow.
- Command Translation: The configurations of the instructions for provisioning and network elements are drastically different. They require conversion and synchronization, and the result of the configuration of network elements needs to be given to the person who demands it after conversion.
- NE Interface: Different Network Elements (NE) have different communication interfaces. The most common method of communication is TCP/IP socket programming interface, which is still used in the earlier X.25 protocol. The demands of TCP/IP socket interfaces are divided into synchronization and a-synchronization based on their degree of real-time. Considering the reusability, these interfaces can be made into driver programs.

Besides the above functions, the following functions are also developed in order to allow easy maintenance for the system:

- Error Handling Mechanism: Configurations cannot be completed if there is a problem with the content of a request or a malfunction in the network component. The convenient Error Handling Mechanism (such as resend or reply) will reduce the probability of manual interventions, thus reducing the cost of maintenance.
- Central Monitoring System: The provisioning in mobile services needs to be

completed within a few minutes. CMS allows maintenance personnel to monitor the system at any time.

A.3.2 xDR Mediation Device

Besides the monthly plan, communication charges are also included in the monthly bill. The billing of the services in the earlier days that focused on voice calls was simpler since the CDR (Call Detail Record) is recorded by the switching center, and charges were calculated according to the duration of each call. As data services and value-added services develop quickly, the billing models become increasingly complex. The sources of this billing information increase rapidly, such as the CDR of GPRS Charging Gateway and the EDR (Event Detail Record) of different network events, and IPDR (Internet Protocol Detail Record). We call these records "xDR." The main task of xDR Mediation Device is to gather the records generated in different network elements, which are then processed and provided to end application systems. This makes the network very visible in the billing system and greatly reduces the complexity of the billing system. Fig. A.5 describes the main modules of xDR Mediation Device.

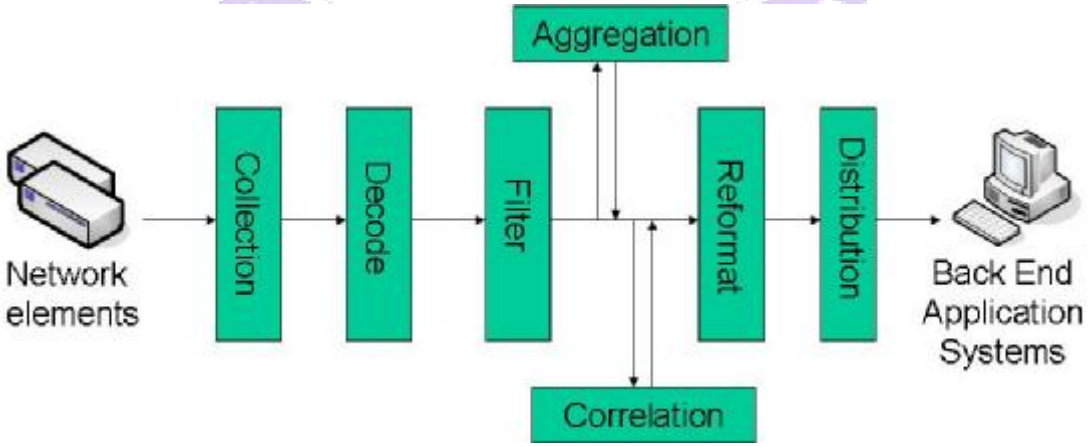


Fig. A.5: Main modules xDR mediation device

- Collection: Collects the xDR from all the network elements. xDR is usually

temporarily stored in the network elements as files. The mediation system is able to actively pull out (such as the “pulling” in ftp) or passively receive (such as the “pushing” in ftp) these files.s.

- Decode: There are no standards for the format and content of xDR. Also being a mobile switcing center (MSC), files produced by different companies are in different formats such as Binary or ASCII, the length maybe fixed or flexible, the fields also differ with different network elements. The languages that are more commonly used to define new network elements are ASN.1 and XML. The main function of decoding is to decipher files of different formats and analyze the data index.
- Filter: Filtrated data are not used by end billing system. For example, in Taiwan, “Mobile Origination Call” needs to pay for the call, and the “Mobile Termination Call” does not. Also making a call from a mobile phone, the charges for the call are already deducted from a prepaid card. Thus, the function of filter is to filter the needed data index according to business logic.
- Reformat: The content needed by the billing system is often less then that in the data index of network elements. For example, for a Mobile Origination Call, a MSC may involve hundreds of fields, whereas billing only involves less than 20 fields. The function of reformat is to convert the raw data into the format needed by the end application system.
- Distribution: xDR provides other functions other than billing such as preventing call-hijacking, data storage, communicating with police and official departments, and online bill settlement. The function of distribution is to send the reformatted files to end application systems through the most commonly used channel such as ftp.
- Aggregation: The actions in network elements often do not match with the needed

data in the end application system. For example, in WAP mobile Internet services, each transaction is recorded whenever a button is pressed in WAP Gateway. From the angle of billing, it is sufficient to charge for each session. The function of aggregation is to compile the relevant call detail records that can be used by the simpler data index required by the end application system. In practice, correlated databases are used for temporary storage of compiled data that will be sent to the end application system.

- Correlation: A transaction may generate different records on different network elements. Take MMS for example, a MMS transmission generates xDR in WAP Gateway, MSC, SMSC, GPRS Charging Gateway, MMSC, MMS Terminal Gateway, or MMS Email Gateway. Correlation analysis needs to be conducted in these different CDRs in order to come up with the data index needed by the end applications. The function of correlation is to determine the correlation between CDR and provide fuller data index for the end application system. In practice, correlated databases are used for temporary storage, the processed data are then sent to the end application system.

Besides these basic functions, the mediation system in service provisioning, Error Handling Mechanism, and Central Monitoring System are also necessary tools that allow easy maintenance of the system.

A.4 Billing Management System

Billing Management System includes calculating the charges based on the xDR, handling billing, handling refunds, inquiring into questionable charges, handling international roaming, managing payable accounts (collecting the amount or halting/terminating/restoring services), modifying accounts, switching accounts, making refunds, or charging against the

deposit. Fig. A.6 is the flow of the Billing Management System. Functions from (5) to (11) are all completed in this system.

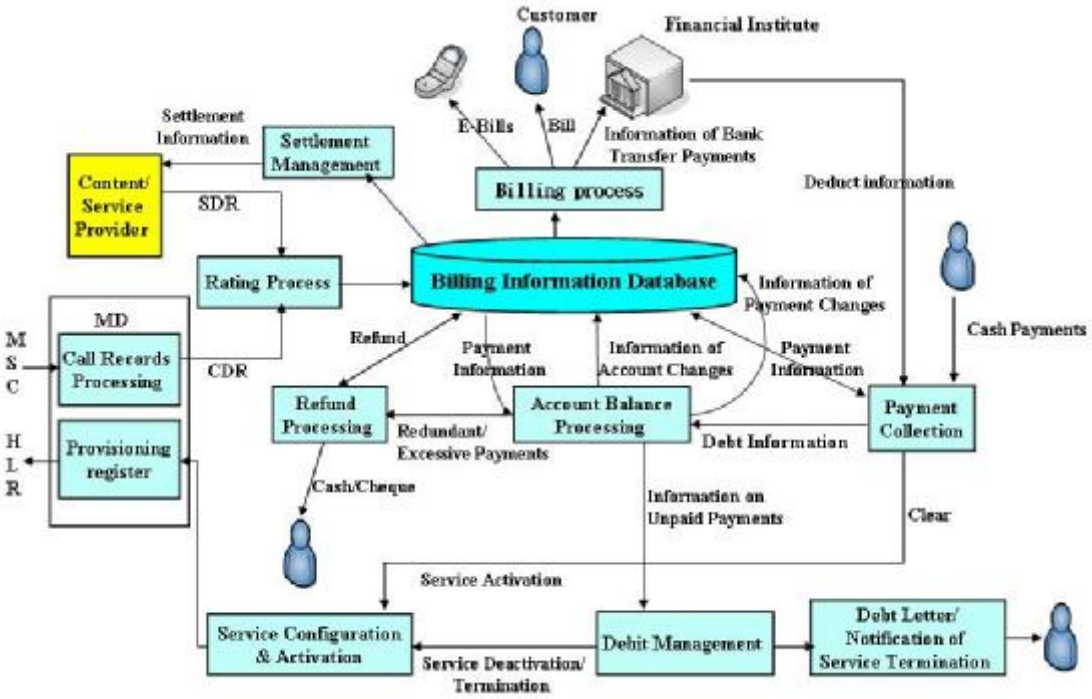


Fig. A.6: Procedures of billing

A.4.1 Rating Process

Whether a user makes voice calls, uses data transfer services, or valued added services, the most important part to service providers is collecting charges and making revenues and profits. Calculation of communication charges is based on the records in each switching center or network component. For example, records for voice services are from MSC CDR, SMS are from SMSC CDR, and data services are from Charging Gateway CDR. These records must be collected, filtered, and converted by the mediating device before being used to collect charges.

There are many factors that influence the rating process, and the main ones are the method of rating and the pricing strategies. The rating methods include “both callers and receivers paying for the calls” (practiced in the U.S.) or “callers paying for the calls only”

(practiced in Taiwan). Moreover, rating becomes more complicated as services diversify. Besides using the duration of each call to calculate the charges, charges can be also based on volume and number of events.

The pricing strategy often includes the following perspectives. After a pricing strategy is formulated, the method of billing is then planned accordingly in order to ensure the calculation is effective and accurate.

- Flat Rate
- Duration-Based
- Volume-Based
- QoS-Based
- Address-Based (APN)
- Location-Based
- Content-Based
- Transaction-Based
- Application-Based
- Prepaid or Postpaid

Besides determining the pricing strategies, there are other factors that influence the actual billing rates such as the calling number (area code), receiving number (area code, country code), type of call, type of customer (e.g., 188), date of call (workday or holiday), and time of call (rush hour or non-rush hour) all influence the charges.

A.4.2 Billing Process

The billing process collects all the charges in order to generate a billing statement for the client. To avoid having all the clients making payments at the same time and causing delays,

monthly billing is done in several cycles such as putting the phone numbers into several groups that have different billing dates. The charges include the monthly charges or communication charges. Some companies provide e-commerce services that involve smaller amounts, which can be combined into the monthly billing. Besides cash, clients can also make payments via financial institutions, which will receive the billing information and pay out the amount accordingly.

In order to match companies' management strategies and operational needs and provide the best services, today's customer information systems are focused on bundle services that integrate fixed network, data, and mobile services, allowing customers to receive complete coverage in less time. "Convergent billing" is a major revolution in this sense.

The basis of convergent billing depends on whether there is good account management. The configuration of the account (that represents the ID of an account) needs to be planned from the customer's perspective and should integrate the relationships between customers, accounts, contracts, and services (as shown in Fig. A.7). A customer may have multiple accounts, and an account may have multiple contracts. Contracts and telecom services (equipment and phone numbers) have multiple relationships. At any point in time, however, the relationship is one-to-one. Convergent billing can only become efficient when the relationships are accurate. Information systems relevant to billing should provide convenient functions that give customers enough information to determine whether one wishes to get the service of convergent billing. Except for the outward-displayed interface, account management (that identifies each account) needs to include flexible mechanisms for each payable service since this involves the calculation of the taxation for each charge (for different accounting categories). Without careful planning, the follow-up process of revenue-report will face problems. For example, the U.S. has federal taxes, state taxes, and local taxes, and their system is known as the most complicated one. Furthermore, the priority of each charge needs to be clarified so that there is a clear set of rules when payments made by the clients do not

match with the payment amounts. For example, charges acquired in business use should be paid first followed by agency receipt. Thus, the billing flow should consider all these issues.

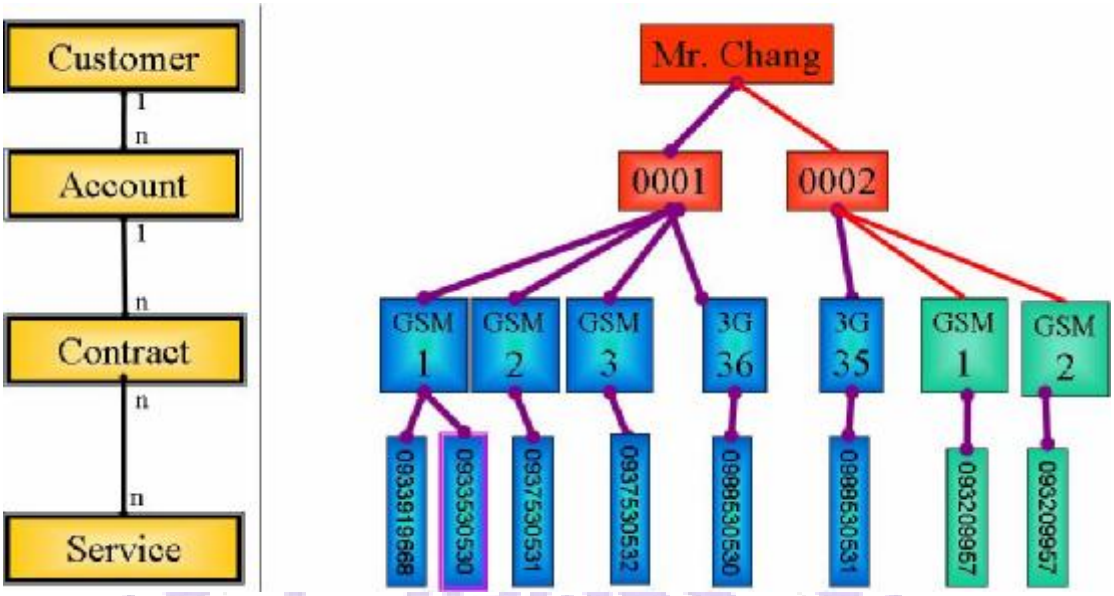


Fig. A.7: Relationships between customers, accounts, and contracts

Billing is at the pivotal point of the entire billing handling system. Its most fundamental function is to connect customer application information such as customer names and charges due to different billing plans in order for follow-up procedures to take place. The advanced functions also must match customers' choices, give them discounts based on their identifications (customers), services (accounts), and equipment numbers (contracts) so they are attracted by the services. For example, a service provider gives a senior customer a discount if the customer signs a contract for a certain duration, and this discount is calculated on the aspect of "equipment numbers" alone. However, if the discount is given on the aspect of "account," then it could be calculated based on all the monthly charges or monthly bills. Similarly, if discounts are given to all the phone numbers that one received from a single service provider, this will be a more attractive deal that in turn increases customer loyalty.

Although the charges are put on a single billing statement after integration, detailed information of each account's expenses should be retained in order to provide customers with

sufficient information. The format of the billing statement should consider that in Fig. A.8 in order to make it dependable.

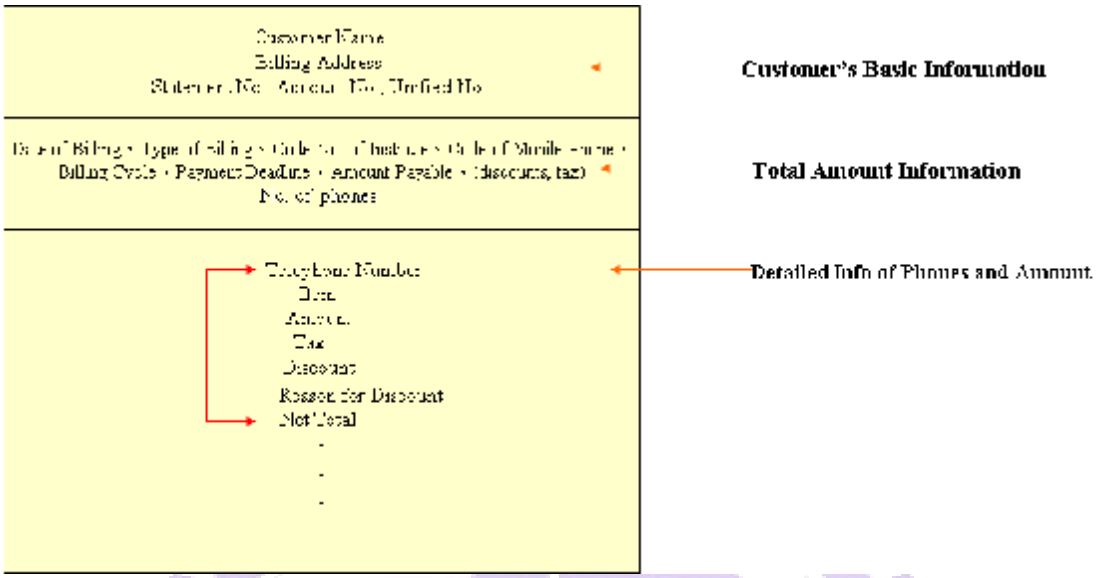


Fig. A.8: Sample of billing statement

A.4.3 Payment Process

After receiving the billing statement, clients go to the service centers, convenience stores, or financial institutes to make payments. The payment information will be sent to the operators. The payment process is comparing the amount of a bill with the amount paid by a customer in order for the bill to be checked off. In practice, service providers usually consider a 2-stage payment process: when a client makes a payment, a record is made. However, since different payment systems have different order of processing, it is possible that the billing system's record for a client remains unchanged even though the client has made the payment 2 or 3 days ago. The system often gets updated only after the payment has actually entered the billing system, and such delay often leaves customers very disappointed.

After the payment has been compared with the payable amount, situations such as insufficient, excessive, or redundant payments are often seen. Service providers should address this issue by having measures such as sending SMS to remind a client who has

delayed payments to make the payments timely, and the service of whom will be terminated or canceled if the customer ignores such request.

The importance of the payment process is to provide all sorts of financial reports that allow the business owners to stay on top of their cash flow since it is one of the matters that business owners cared about most as well as how investors evaluate the business owners.

A.4.4 Customer Account Management

To solve customers' disputes over their accounts, the account management system needs to provide functions such as account adjusting/modifying, refunds, and balance inquires so all the charges are explained clearly. Convenient, user-friendly interfaces also need to be provided for the customers.

A.4.4.1 Adjust Charge

If a customer has questions regarding his/her paid bill and asks the service agent to verify it, and there are indeed errors in the bill, adjustments would be needed to compensate the customer. In the next billing, the overpaid amount will be deducted from the payable amount. The phone number of the unit is counted as the unit in order to clearly calculate the charges. In order to clarify the phone numbers, the account number can also be used to list a customer's phone numbers before adjustments are made.

A.4.4.2 Modify Charge

If a customer has not yet paid a bill that has errors, the service agent can modify the charges after investigations so that the customer will pay for the correct amount. Therefore, the unit of "modifying charges" is the billing statement. The charges on a monthly bill needs to be listed and modified one by one, followed by adding them up and calculating the taxes.

A.4.4.3 Refund Process

Operators treat their bills as monthly, independent events, and this is different from the "roll-over" bills such as VISA credit cards." Therefore, "refund" is needed if there is additional money left after a bill has been paid or a single payment has been made twice. Refunds include cash, cheques, or used to deduct the charges in the next bill. Cash or cheques are used to return the money completely to the clients; if the money is used to deduct the next bill, the detailed charges need to be listed.

A.4.4.4 Charge Inquiry

Charge Inquiry is necessary to clarify each charge. The system should provide different methods of inquiries that use information such as customers' ID numbers, account numbers, or phone numbers.

A.4.4.5 Debit Management

Most billing systems' debit management systems include functions such as debt collection or service halting/termination. Debit management needs to be able to automatically list the phone numbers in an account that have unmade payments and collects the payment or halt/terminate services in an appropriate order. These procedures must be done in a correct sequence. The methods of collecting unmade payments include using automatic dialers to playback a voice mail for a client who has unmade payments or using SMS. Service halting/termination requires the work order uploaded to switching centers that integrates the Provisioning Mediation Device as discussed in Section A.3.1. The purpose of debit management is to remind the customers to make their payments; no service providers would like to see their customers leaving them. Thus, service halting includes the two-way (cannot make or receive calls) or one-way (can receive calls) method. This also allows quick restoration of the service after payment is completely made.

A.4.5 Settlement Management

Since the network service providers have the mechanisms of “billing system” and “handling system,” these two are used to guide the negotiations between collaborators and collect the applicable charges. Settlement Management allows network service providers, content providers, distributors, and other collaborators to settle the relevant charges. Whether the relationships between these parties are established and managed well is the key to the smooth process of Settlement Management; otherwise, disputes over account transactions maybe exist. Currently, the most common models of settlement include the “actual payment” and “payable payment.” In terms of actual payment, the settlement can only be done for the payment that has actually been made by a client. Thus, content providers and distributors would need to cover the loss. On the other hand, they pay less commission to the service providers. In terms of “payable payment,” the amount that is “payable” needs to be settled whether the clients make the payments or not. Thus, the network service providers cover the loss, but they receive more commissions from the distributors.

A.4.6 Inbound Roaming Billing Management

Compared to fixed networks, mobile phones do not only move freely but they also allow cross-operator roaming, which is divided into Inbound Roaming and Outbound Roaming. Inbound roaming means a foreign client whose account is registered in Taiwan uses the network resources in Taiwan to make calls. The network in Taiwan is in charge of the billing that includes the billing rates used by the overseas network that served this client. Fig. A.9 shows the procedures of Inbound Roaming.

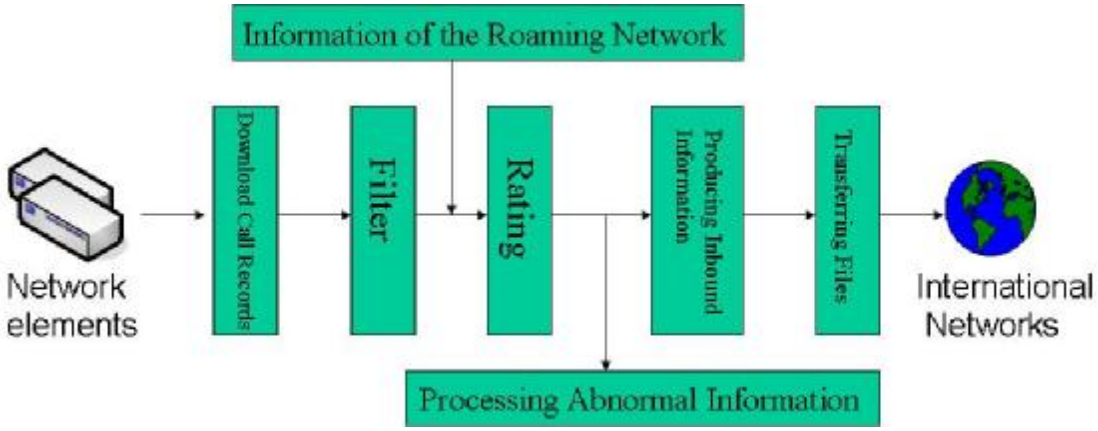


Fig. A.9: Procedures for Inbound Billing

A.4.7 Outbound Roaming Billing Management

Outbound Roaming refers to when a Taiwanese client uses an oversea network that has an agreement with a network in Taiwan. Since the wireless network in a foreign country is accessed, billing is based on the call records in the switching centers in this country. After the charges are calculated, the files are then sent to the Taiwanese operator which sends a bill to the client. Usually, the Taiwanese network would charge service charges for outbound roaming, and the current practice is to multiple the original charges by 15%. What needs to be noted is that a call made from a person in Taiwan to the mobile phone of the person who uses the roaming service in a foreign country is also counted as an international call, and the owner of the mobile phone needs to pay for the roaming charges. The same also applies to SMS and voice mails. Fig. A.10 shows the billing procedures of Outbound Roaming.

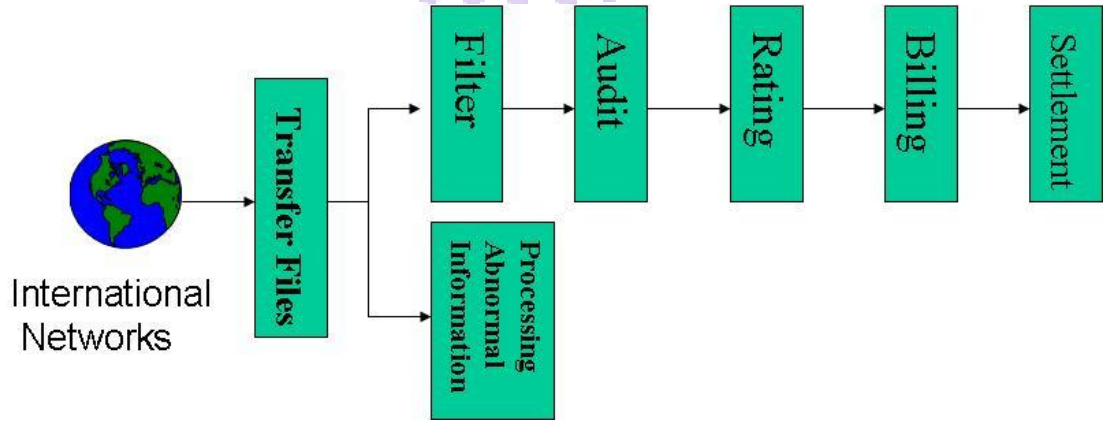


Fig. A.10: Outbound processing

A.4.8 Electric Bill Presentation and Payment

The communication between operators and their customers is often through monthly bills. The service providers send bills to their clients through out a month or collect payments from financial institutions. If a client chooses to make payments in a financial institute, he or she would need to take the time to go and make the payments, and the service providers need to spend money on printing the bills and postage. As a result, Electric Bill Presentation and Payment (EBPP) is proposed as an electronic method that replaces manual billing and payment mechanisms. Service providers can use the Internet, e-mails, or SMS to give the bills to their customers. On the other hand, their customers can use financial tools such as credit cards to make online payments via the Internet or mobile Internet. Using bank cards (an ATM-styled payment mechanism provided by financial institutes) to make payments is also getting popular. Fig. A.11 briefly describes the feasible software for EBPP.

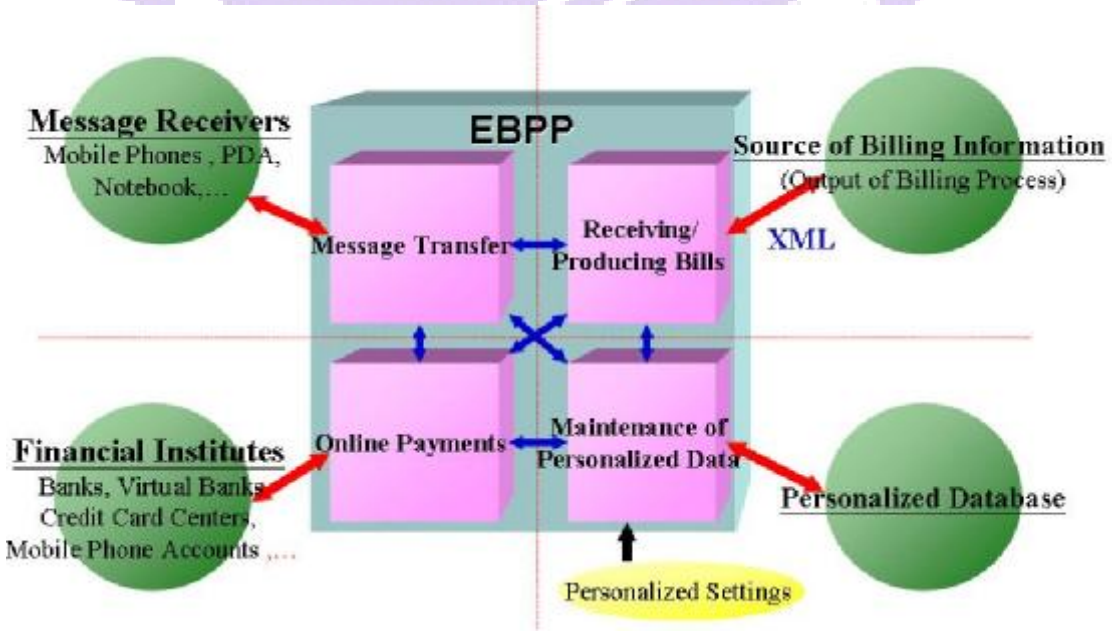


Fig. A.11: Framework of software utilized by EBPP

A.5 Prepaid System

The above sections focus on customers who use monthly services who make payments at the end of each month. After the first prepaid mobile service plan that was introduced in 1995, this service has grown rapidly, and now more than 250 wireless providers in more than 100 nations provide prepaid services. More than 600 million users use prepaid services, which is 60% of the total mobile phone users in the entire world. The percentage even exceeds 90% in some of the European nations. The so-called “prepaid” basically means a user pays a certain amount of money for the service when applying for it. Prepaid services have many advantages for the service providers:

- Prepaid services help reduce cost and increase the number of clients quickly.
- Payments are already fully made by the customers, eliminating bad debts.
- Eliminates the cost of printing and mailing bills.
- Eliminates the bad debts caused by illegal copies of SIM cards.
- Expands the potential pool of customers, including business travelers, corporations, or clients whose services were terminated due to bad credits. This also allows phone rental services.

To the customers, since they have paid for a certain length of duration for their services, they can effectively manage the amount of usage. For example, this helps parents to restrict their children’s cell phone bills. Moreover, since a receiver of a call does not need to pay for the call in Taiwan, clients who have low call volumes can save money on monthly service fees.

The prepaid system basically includes three actual types: the Billing System, the Intelligent Network, and Smart Card models. Since the most widely used system is the Intelligent Network system, we will describe its framework in Fig. A.12.

In this figure, HLR identifies whether a client is indeed a user of prepaid services since

prepaid clients often have special fields such as Service Key, DP2 trigger, and SCP address in their HLR. Moreover, SCP is used to store customers' basic information such as the phone number of a prepaid service, its expiry date, applicable billing plans, suitable languages, and account balance. When a prepaid user makes a call, since HLR/VLR contains DP2 trigger, MSC/SSP will look for the corresponding SCP based on the SCP address in order to inquire into the customer's remaining balance. If the balance is insufficient, SCP will prompt the IP (Intelligent Peripheral) to send a signal of "not enough balance" back to MSC/SSP, and MSC/SSP will playback a voice recording in the IP for the client. If there is enough balance remaining, SCP will convert it into the usable minutes based on the client's billing rate and notify MSC/SSP to let the call go through. During the call, MSC/SSP must constantly monitor to see whether the duration of the call exceeds the allowed time. When it does, IP will be activated to ask the client whether he/she wishes to add in more money or terminate the call. When the call is ended, MSC/SSP must let SCP know the client's remaining minutes so SCP can update the information on the remaining time. As for the source of customer information in HLR and SCP (please refer to Section A.2.2), the configuration is done by a service provider's Service Order Provisioning after a client has applied for the prepaid service.

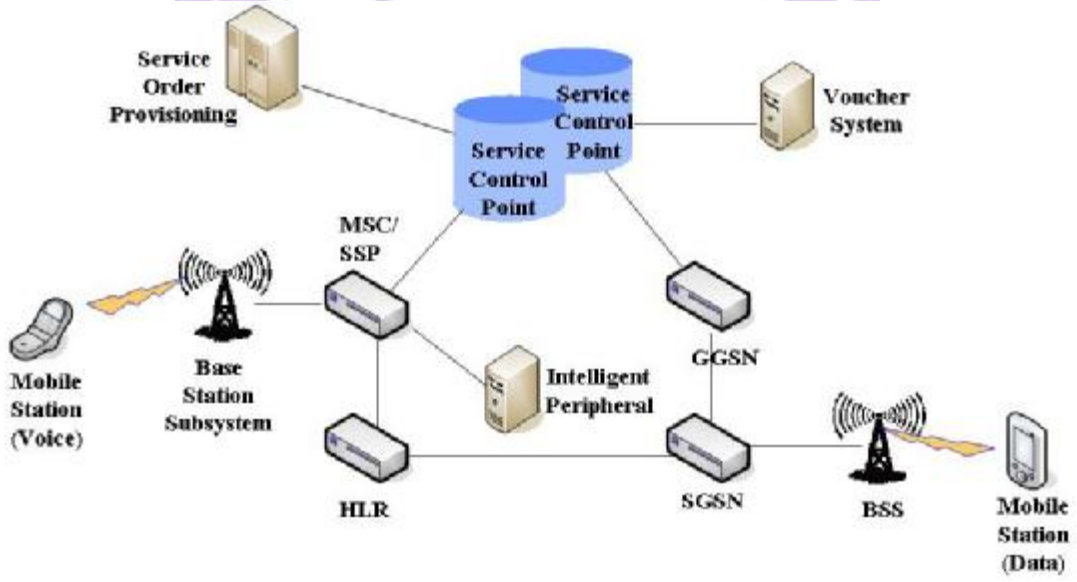


Fig. A.12: Framework of intelligent network prepaid system

Prepaid system's Voucher System is used to allow a user to add in more money. The user can go to the service window of the service provider or a convenience store to buy more "tokens" and use the code number provided by the service provider (Chunghwa's code is 928, for example) to add in the tokens.

A.6 Summary

A good CCBS must be constantly updated, and the only thing that remains unchanged is "continued changes." The activation and operation of a system is not the end but rather the beginning, and workers in this field must realize that the relevant functions and technologies must continue to be improved in order to satisfy their customers' demands. This is especially true for the integration of a large information system since it is extremely complicated. Without proper planning, appropriate tools, actual implementations, and careful collaborations, it would not be done within a limited amount of time. This is what software engineering is constantly working on.

Although information technologies can be acquired at schools, the key to successfully developing CCBS is "domain knowledge" – especially for telecommunication services that have a very long history. Therefore, the companies of many well-known brands of this type of software in the market mostly started as telecom operators or their long-term partners. In this era of knowledge-based economy, domain knowledge has become the most cherished and valuable asset.

Reference

[1] 3GPP. 3rd Generation Partnership Project; Technical Specification Group Core Network; Signalling flows for the IP multimedia call control based on SIP and SDP; Stage 3. version 5.15.0(2006-10). 3GPP TS 24.228, 2006.

[2] 3rd Generation Partnership Project, Technical Specification Group Services and System Aspects. Telecommunication Management, Charging and Billing, 3G call and event data for the Packet Switched (PS) domain. Technical Report 3GPP TS 32.015, version 3.9.0, 3GPP, Mar. 2002.

[3] Anderson, M. M. Financial Service Markup Language (FSML) Version 1.17.1. Technical Report Financial Services Technology Consortium, October 1998.

[4] Andreasen, F. and Foster, B. Media Gateway Control Protocol (MGCP) Version 1.0 IETF RFC3435, 2003.

[5] Arteta, A. Prepaid Billing Technologies - Which One is for you? *Billing World*, pages 54–61, Feb. 1998.

[6] Asokan, N., Janson, P. A., Steiner, M., and Waidner, M. The State of the Art in Electronic Payment Systems. IEEE Computer, 30(9):28-35, September 1997.

[7] Barreto, P., Kim, H., Bynn, B., and Scott, M. Efficient Algorithms for Pairing-Based Cryptosystems. Proceedings of 22nd Annual International Cryptology Conference Santa Barbara, 2442:354-368, Springer-Verlag 2002.

[8] Barreto, P. S. L. M., Ben, L., and Michael, S. Efficient Implementation of Pairing-Based Cryptosystems. Journal of Cryptology, 17(4):321-334, 2004.

[9] Bellare, M., Garay, J., Hauser, R., Herzberg, A., Krawczyk, H., Steiner, M., Tsudik, G., Herreweghen, E., and Waidner, M. Design, Implementation and Deployment of a Secure Account-Based Electronic Payment System. IEEE Journal on Selected Areas in Communications, 18:611-627, April 2000.

[10] Boly, J. P., Bosselaers, A., Cramer, R., Michelsen, R., Mjolsnes, S. F., Muller, F., Pedersen, T. P., P_tzmann, B., Rooij, P. de, Schoenmakers, B., Schunter, M., Vallee, L., and Waidner, M. The ESPRIT Project CAFE-High Security Digital Payment Systems. Proceedings of ESORICS, pages 217-230, 1994.

[11] Boneh, D. and Franklin, M. Identity-Based Encryption from Weil Pairing. Proceedings

of Crypto 2001, 2139:213-229, 2001.

[12] Cai, Y. et.al. Authorization mechanisms for mobile commerce implementations in enhanced prepaid solutions. *Bell Labs Technical Journal*, 8(4):121–131, Feb. 2004.

[13] Camenisch, J., Maurer, U., and Stadler, M. Digital Payment Systems with Passive Anonymity-Revoking Trustees. Proceedings of ESORICS, pages 33-43, 1996.

[14] CCITT. Numbering Plan for the ISDN Era. Technical Report Recommendation E. 164 (COM II-45-E), ITU-T, 1991.

[15] Chang, M.-F., Lin, Y.-B., and Pang, A.-C. vGPRS: A Mechanism for Voice over GPRS. *ACM Wireless Networks*, 2001.

[16] Chang, M.-F., Lin, Y.-B. and Yang, W.-Z. Performance of Hot Billing Mobile Prepaid Service. *Computer Networks Journal*, 36(2):269–290, Jul. 2001.

[17] Chang, M.-F., Yang, W.-Z. and Lin, Y.-B. Performance of Service-Node Based Mobile Prepaid Service. *IEEE Transactions on Vehicular Technology*, 51(3):597–612, May 2002.

[18] Chen R.-J. Hwu, J.-S. and Y.-B. Lin. An identity-based cryptosystem for end-to-end mobile security. IEEE Transactions on Wireless Communications, 2006. Accepted for publication.

[19] Daemen, J., Borg, S., and Rijmen, V. The Design of Rijndael: AES - The Advanced Encryption Standard. Springer-Verlag, 2002.

[20] Davies, W. Moving Wireless Office. *Taiwan Telecommunication Magazine*, pages 93–97, Dec. 2001.

[21] Davies, W. Who's Ready for GPRS. *Telecommunications*, Feb. 2002.

[22] D.L. Lu. presented in Chunghwa Telecom Co., Ltd. Apr. 2002.

[23] Fancher, C. H. In Your Pocket: Smartcards. IEEE Spectrum, 34:47-53, February 1997.

[24] Ferreira, L. and Dahab, R. A Scheme for Analyzing Electronic Payment Systems. Proceedings of ACSAC, pages 137-146, 1998.

[25] Grimaldi, R. Discrete and Combinational Mathematics, Fifth Edition. Addison-Wesley, 1999.

[26] Heikki, K., Ari, A., Lauri, L., Siamak, N., and Valtteri, N. UMTS Networks-Architecture, Mobility & Services. John Wiley & Sons, Inc., 2002.

[27] Herzberg, A. and Yochai, H. Mini-Pay: Charging Per Click on the Web. Proceedings of

the Sixth International World Wide Web Conference, Santa Clara, California, pages 301-307, April 1997.

[28] Hess, F. E_cient Identity Based Signature Schemes Based on Pairings. Proceeding of Selected Areas in Cryptography: 9th Annual International Workshop, 2595:310-324, August 2002.

[29] IEEE. IEEE Std 802.11-1997 Information Technology-Telecommunications and Information Exchange between Systems-Local and Metropolitan Area Networks-Specific Requirements-Part 11: Wireless Lan Medium Access Control (MAC) and Physical Layer (PHY) Specifications. Technical Report IEEE Std 802.11-1997, 1997.

[30] IEEE. IEEE Standard for Local and Metropolitan Area Networks-Part 16: Air Interface for Fixed Broadband Wireless Access Systems. Technical Report IEEE Std 802.16-2004, 2004.

[31] INFOCOMM. Intelligent Prepaid System (IPS) Functional Description. Technical Report PDIPS-00-00001-FD, INFOCOMM, 1998.

[32] ITU. Packet-based Multimedial Communications Systems. Technical Report ITU-T H.323, Version 5, International Telecommunication Union, 2003.

[33] Jonsson, J. and Kaliski, B. Public-Key Cryptography Standards (PKCS) #1: RSA Cryptography Specifications Version 2.1. Technical Report RFC 3447, 2003.

[34] Lai, Y.-C., Lin, P., and Huang, Y.-T. Design and Implementation of a Wireless Internet Remote Access Platform. Journal of Wireless Communications and Mobile Computing, 6(4):413-429, January 2006.

[35] Lee, Z.-Y., Yu, H.-C., and Ku, P.-J. An Analysis and Comparison of Di_erent Types of Electronic Payment Systems. Proceedings of Portland International Conference on Management of Engineering and Technology, 2001. PICMET'01, 2:38-45, 2001.

[36] Lin, P. et.al. Credit Allocation for UMTS Prepaid Service. *Accepted and to appear in IEEE Trans. on Veh. Technol.*

[37] Lin, Y.-B. Introduction to Telephony. Vekeg, 1997.

[38] Lin, Y-B. Telephone Network and PBX Software. Wei-Keg, 1999.

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

[40] Liu, C.-H. et. al. CAMEL Evolution and PPS Evaluation. *IEEE intelligent network workshop*, pages 9–13, May 2001.

[41] Low, S., and Maxemchuk, N. Anonymous Credit Cards. Proceedings of 2nd ACM Conference on Computer and Communications Security, pages 108-117, 1994.

[42] Mastercard and Visa. SET Secure Electronic Transactions Protocol, version 1.0 edition, Book One: Business Speci_cations, Book Two: Technical Speci_cation, Book Three: Formal Protocol Definition. May 1997.

[43] Medvinsky, F. and Neuman, B. NetCash: A Design for Practical Electronic Currency on the Internet. Proceedings of the First ACM Conference on Computer and Communications Security, pages 102-106, 1993.

[44] Menezes, A., Oorschot, P. V., and Vanstone, S. Handbook of Applied Cryptography; [Online] Available: http://citeseer.ist.psu.edu/428600.html, 1999.

[45] Rao, H.C.-H., Lin, Y.-B., and Chou, S.-L. iGSM: VoIP Service for Mobil Networks. *IEEE Communications Magazine*, 4(38):62-69, 2000.

[46] Rivest, R. L. and Shamir, A. PayWord and MicroMint: Two Simple Micropayment Schemes. IEEE Transactions on Vehicular Technology, 52(1):132-141, 2003.

[47] Shamir, A. Identity-Based Cryptosystems and Signature Schemes. Proceedings of CRYPTO 84 on Advances in cryptology, pages 47-53, 1985.

[48] Skype. http://www.skkype.com, 2007.

[49] Smith, C. and Collins, D. *3G Wireless Networks*. McGraw-Hill, 2002.

[50] Stallings, W. Cryptography and Network Security: Principles and Practice, Second Edition. Prentice-Hall, 1999.

[51] Tracey, L. V. Prepaid Calling: Market Update. *Telecommunications*, Apr. 1999.

[52] Vriendt, J. D. et.al. Mobile Network Evolution: A Revolution on the Move. *IEEE Communications Magazine*, pages 104–111, Apr. 2002.

[53] Wang, H. and Guo, H. Fair Payment Protocols for E-Commerce. Proceedings of the 22th Annual Symposium on Principles of Distributed Computing, pages 227-245, 2004.

[54] Yang, Z., Lang, W., and Tan, Y. A New Fair Micropayment System Based on Hash Chain. Proceedings of IEEE International Conference on e-Technology, e-Commerce and e-Service. EEE'04, pages 139-145, 2004.

[55] Yen, S.-M. PayFair: A Prepaid Internet Micropayment Scheme Ensuring Customer Fairness. Proceedings of IEE Computers and Digital Techniques, 148(6):207-213, November 2001.

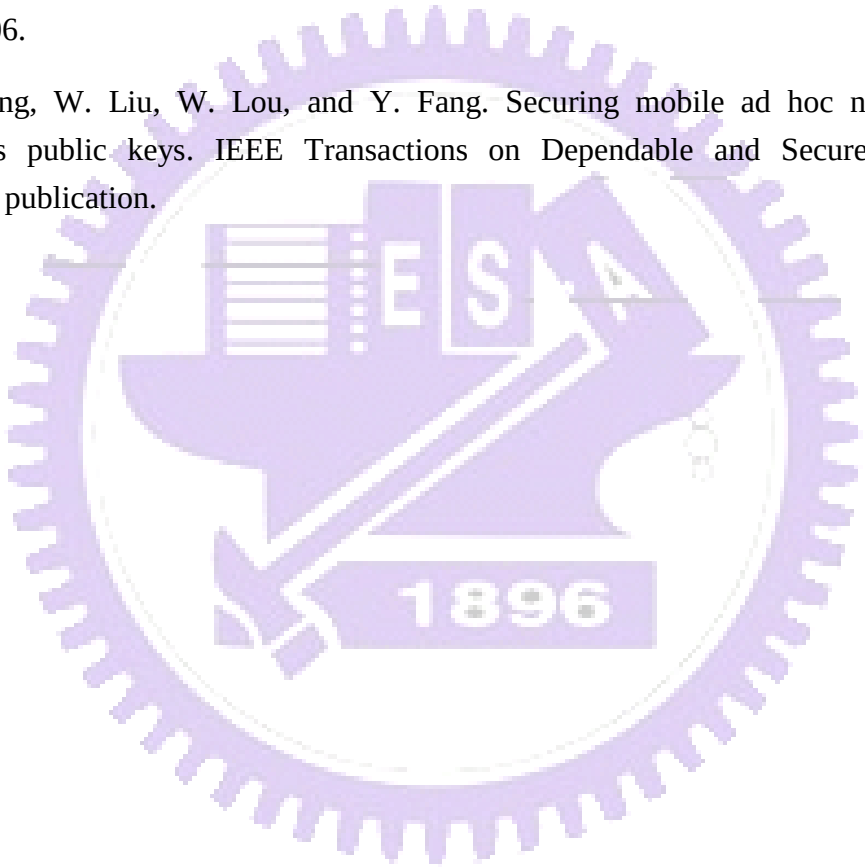
[56] Y. Zhang and Y. Fang. ARSA: an attack-resilient security architecture for multi-hop wireless mesh networks. IEEE Journal on Selected Areas in Communications, 2006. Accepted for publication.

[57] Y. Zhang and Y. Fang. A secure authentication and billing architecture for wireless mesh networks. ACM Wireless Networks, 2006. Accepted for publication.

[58] Y. Zhang, W. Liu, W. Lou, and Y. Fang. Anonymous handshakes in mobile ad hoc networks. IEEE Transactions on Wireless Communications, 2006. Accepted for publication.

[59] Y. Zhang, W. Liu, W. Lou, and Y. Fang. Location-based security mechanisms in wireless sensor networks. IEEE Journal on Selected Areas in Communications, 24(2):247-260, February 2006.

[60] Y. Zhang, W. Liu, W. Lou, and Y. Fang. Securing mobile ad hoc networks with certificateless public keys. IEEE Transactions on Dependable and Secure Computing, Accepted for publication.



Curriculum Vitae

Lu, Fang-Sun was born in Taipei, Taiwan, R.O.C., in 1960. He received the B.S. degree in applied mathematics, the M.S. degree in mathematics from Fu-Jen University in 1983 and 1985, respectively. He is currently a PhD candidate in the Department of Computer Science and Information Engineering at National Chiao Tong University. Since 1985, he has been with the Customer Service Systems Laboratory of Telecommunication Laboratories, Chunghwa Telcom Co., Ltd, where he is currently a Distinguished Researcher and a project manager. His research interests include design and analysis of personal communications services network, development of telecommunication operation support systems, and performance modeling.

Publication List

◆ Journal Publications

1. Wei-Zu Yang, Fang-Sun Lu*, Ming-Feng Chang. “Performance Modeling of an Integrated Mobile Prepaid Services.” IEEE Transaction on Vehicular Technology, 56(2):899-906, 2007.
2. Phone Lin, Hong-Yueh Chen, Yuguang Fang, Jeu-Yih Jeng, Fang-Sun Lu*, “A Secure Mobile Electronic Payment Architecture Platform for Wireless Mobile Networks.” accepted and to appear in IEEE Transaction on Wireless Communications.

◆ Conference Papers

1. Wei-Zu Yang, Fang-Sun Lu*, Ming-Feng Chang. “Performance Modeling of an Integrated Mobile Prepaid Services.” The 10th Mobile Computing Workshop, pp.128-133, March, 2004.

◆ Book

1. 鄭枸澮, 呂芳森*, “第十三章 電信客戶服務與帳務系統 (Customer Care & Billing System for Telecommunication).” 個人通訊服務網路. 教育部顧問室「通訊科技人才培育先導型計畫」 維科圖書有限公司 2006

◆ In Preparation

1. Meng-Fang Chang, Fang-Sun Lu* and Chung-Yung Chia , " A Callback Mechanism for Private Telecommunications Network."