Credit Allocation Schemes for Quality-class-oriented Services in Next Generation Networks

Hsiu-Fang Ho
Institute of Computer and Communication Engineering
National Cheng Kung University
Tainan, Taiwan 701, R.O.C
q36984202@mail.ncku.edu.tw

Sok-Ian Sou
Department of Electrical Engineering
Institute of Computer and Communication Engineering
National Cheng Kung University
Tainan, Taiwan 701, R.O.C
sisou@mail.ncku.edu.tw

Jeu-Yih Jeng
Customer Service Information Technology Laboratory
Telecommunication Laboratories
Chunghwa Telecom Co.
Taipei, Taiwan 106, R.O.C
jyjeng@cht.com.tw

Abstract—This paper studies credit allocation schemes for quality-class oriented services based on the 3GPP policy and charging control (PCC) architecture. According to users' preference for the service quality, three credit allocation schemes, Minimum Credit First (MCF), Average Quality First (AQF) and Best Quality First (BQF), are proposed and investigated. Specifically, we study the expected number of sessions (n_c) supported and the expected lifetime (T_c) of an online charging account for each scheme. The above performance metrics provide useful information to operators for online account management.

I. INTRODUCTION

Next Generation Network (NGN) supports real-time IP multimedia services through IP Multimedia Subsystem (IMS) over heterogeneous IP networks [1], [2], [3]. In recent years, the NGN architecture requires a convergent charging solution that allows both prepaid and post-paid accounts handled in one billing platform for different kinds of services [4], [5], [6]. Before a session with online charging starts, the *Packet Data Gateway* (PDNGW) needs to reserve a certain amount of *online credit* from the charging system for this session. The online credit is maintained in a central node called *Online Charging System* (OCS) [7]. The flexibility in real-time online credit allocation attracts investments from content and service providers in NGN. With OCS, an operator can reduce the bad debt risk; a subscriber does not have a bill shock [8].

In NGN, the IP-based multimedia services specify critical charging requirements. In traditional charging plan, the services are charged by time-based, volume-based or content-based [9]. For example, a user spends NT.7 dollars to make 1-minute outgoing call time; a user spends NT.8 dollars to download a 200-KB data; a user downloads a ringing tone for NT.30 dollars [10]. However, the billing plan for content-based services is hard to design. Many IMS services are served with different charging requirements. The value for an IMS service is hard to measure only by time-based, volume-based or content-based method, since the bandwidth requirement among different multimedia services greatly differ from traditional telecom services. Also, the competition in telecom markets is very tough and the price reduction pressure from government and subscribers is high. The charging plan for IMS services should satisfy the expectations from the customers, the *Internet Service Provider* (ISP) and the *Content Provider* (CP).

Besides of the billing plan, a mobile operator requires an efficient way to manage network resources for bandwidth allocation and packet filtering. Therefore, combining policy control with online charging is a new trend for mobile operators to carry out an advanced billing platform. In NGN, the *Policy and Charging Control* (PCC) is standardized by 3GPP to realize dynamic network resource control and charging management [11], [12]. Through the PCC architecture, operators can support more advanced billing plans for mobile services. This paper studies how online credit allocation can be effectively applied to charging plan that considering which quality class provided to the session. Fig. 1 shows the PCC architecture, where a main component *Policy and Charging Rules Function* (PCRF; Fig. 1 (a)) is used to provide PCC rules (see also Table I) for a service flow such that policy enforcement and charging management can be performed in NGN. The *Policy and Charging Enforcement Function* (PCEF) is implemented at the PDNGW (Fig. 1 (b)). The *Subscriber Profile Repository* (SPR; Fig. 1 (e)) stores the user PCC-related information such as resource requirement and service personalization. According to the billing class of a subscriber, the type of the application to be accessed and the local control policy defined by the telecom operator, the PCRF makes policy decision and provides PCC rules to the PDNGW/PCEF through the Gx interface (see Chapter 9 in [13]). The OCS (see Fig. 1 (f)) is responsible for online charging credit and billing plan management.

Based on the standardized OCS and PCC architecture, we can achieve flexible credit allocation in advanced mobile services (such as IMS calls with different quality requirement). However, how to efficiently allocate online credit to advanced mobile services according to the subscriber preference is not discussed in 3GPP specifications. To fill this gap, we study new kinds of credit allocation schemes for quality-class-oriented services in NGN.

Denote a user refresh cycle as the time between when a user refreshes his/her online account and when all credit in the account is consumed. A user refresh cycle is refreshed to as the lifetime of an online charging account. In each user refresh cycle, consider there is a fixed amount of credit in an online charging account. Before new online credit allocation schemes for quality-class-oriented services are brought to the telecom market, operates need to evaluate the following performance metrics:

- How long a newly refresh online charging account can be used before all the credit is consumed for quality-class-oriented services?
- How many sessions can be supported in each user refresh cycle?

In order to answer the above questions, we study three credit allocation schemes according to the quality-class the user preferred and the online credit charged by the service. We study the expected number of sessions (n_c) supported and the expected lifetime (T_c) of an online charging account for each scheme.

II. The OCS/PCC Management For Advance Mobile Service

This section explains the cooperation between the OCS and the PCC in NGN. When a User Equipment (UE; Fig. 1 (c)) initiates a new online service session, the PDNGW requests a PCC rule from the PCRF, which includes the details about the end-to-end services that need to be transferred, such as the service session filters (source/destination IP address and port number), the related QoS description (QoS class, maximum and guaranteed bit rate for uplink/downlink traffic), and the charging information (the measurement method and the charging key) as listed in Table I. The Subscriber Profile Repository (SPR; Fig. 1 (e)) stores the policy requirement and the online account rule settings for the subscribers. When the PDNGW successfully requests the PCC rules from the PCRF, the PDNGW needs to request online credit from the OCS. The OCS determines the rating method based on the PCC rule and allocates the granted credit to the PDNGW.

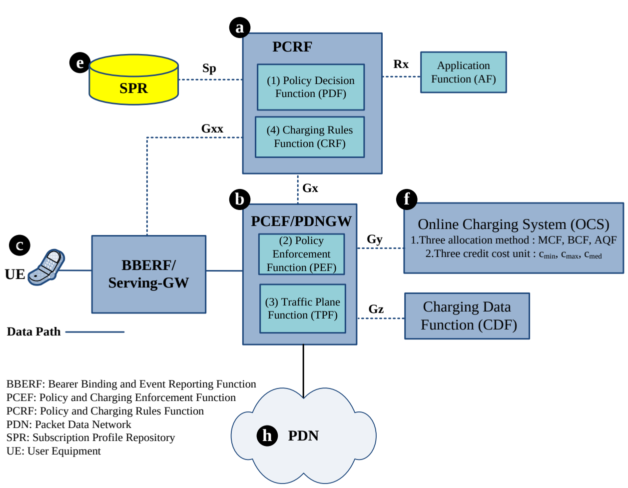


Figure 1. The 3GPP-based Policy Control and Charging Architecture.

Table I. The PCC rule information

Information Name	Description
Rule identifier	It is used to uniquely identify the PCC rule within an EPS session.
Service data flow template	A list of service data flow filters within an EPS session.
Precedence	It is used to determine the order in which the service data flow templates are applied.
Charging key	It is used to determine the tariff for the service data flow in the OCS.
Service identifier	The identity of the service data flow.
Charging method	It is used to indicate the required charging method for the PCC rule.
Measurement method	It indicates whether the service data flow data volume, duration, or event information is measured.
Gate status	It indicates whether the service data flow may pass or be discarded at the PCEF.
QoS class identifier	The identifier for the authorized QoS parameters.
UL/DL maximum bit rate	The uplink/downlink maximum bit rate authorized for an EPS session.
UL/DL guaranteed bit rate	The uplink/downlink guaranteed bit rate authorized for an EPS session.

Traditionally, there are three kinds of rating methods, namely, the time-based method (e.g., for voice call), the volume-based method (e.g., for data session), and the event-based method (e.g., for the short message service). With the diversity in IMS services, the requirement and expectation in charging method for NGN services greatly differ from that for traditionally telecom services. For example, when a user views a video clip, he/she can choose the quality (video size and quality class) that he/she can afford. Therefore, new kinds of time/volume/event –based services combined with specified quality class are expected. In this paper, by considering the quality class enforced on time-based services, we propose three credit allocation schemes to investigate the effects on credit allocation in quality oriented services for NGN. Note that these allocation schemes are not defined by 3GPP but are necessary for NGN. Based on the simulation framework for time and quality-class-based services, this work can be extended to support more complicated rating rates. The notations used in allocation schemes are described as Table II.

Table II. The Notations of allocation scheme

Notation	Description
C_r	Remaining credit of an online charging account.
M	The total number of quality class.
N	The total number of session type.
$c_{m,n}$	The credit charged for each time unit for a session with type n ($1 \leqq n \leqq N$) and quality class m ($1 \leqq m \leqq M$).
$t_{h,n}$	The session holding time for a service with type n.
$t_{a,k}$	The cumulative allocation time at the k-th CCR reservation ($t_{a,0}$=0).
β	The probability threshold of the expected cumulative holding time in a session.

A. *Best Quality First (BQF) Scheme*

In the Best Quality First (BQF) scheme, a customer requests a service session with the highest quality-class that the remaining credit in the OCS account can support. The idea behinds the BQF scheme is that some users want to enjoy IMS services with the best quality and do not mind how much to pay. In the BQF scheme, the OCS chooses quality class m according to the following rule:

To take maximum m, we subject to

$$(t_{a,k} - t_{a,k-1})c_{m,n} \leqq C_r$$

$$\Pr[t_{a,k} > t_{h,n} | t_{h,n} > t_{a,k-1}] \geqq \beta$$

B. *Minimum Credit First (MCF) Scheme*

In the Minimum Credit First (MCF) scheme, a customer requests a service session with the lowest quality-class. The idea behinds the MCF scheme is that some users want to enjoy IMS services with the cheapest price. In the MCF scheme, the OCS chooses the quality class m according to rule of MCF scheme:

To take minimum m, we subject to

$$(t_{a,k} - t_{a,k-1})c_{m,n} \leqq C_r$$

$$\Pr[t_{a,k} > t_{h,n} | t_{h,n} > t_{a,k-1}] \geqq \beta$$

C. *Average Quality First (AQF) Scheme*

The Average Quality First (AQF) scheme, a customer requests a service session with a medium quality-class. The idea behinds the AQF scheme is that some users want to enjoy IMS services with a common price, which is not the cheapest or the expensive one. In the AQF scheme, the OCS chooses the quality class m according to the following rule:

To take average credit cost unit with AQF scheme, we subject to

$$(t_{a,k} - t_{a,k-1})c_{m,n} \leqq C_r$$

$$\Pr[t_{a,k} > t_{h,n} | t_{h,n} > t_{a,k-1}] \geqq \beta$$

$$c_{m,n} \leqq \frac{1}{M}\sum_{m=1}^{M} c_{m.n}$$

Based on the above three allocation schemes, we investigate how long a new refresh online charging account can be used before all the credit is consumed when multiple quality classes are provided; and how many sessions can be supported in each refresh cycle. In the next section, we establish a simulation model to model credit allocation in OCS with quality-oriented services.

III. SIMULATION MODEL

In this paper, we develop a C++ discrete-event simulation to test the performance for the above three credit allocation schemes. For this study, each data point on the plots shown in this section is an average of 1,000,000 samples of such cases. We simulate three kinds of event sessions, ARRIVAL, UPDATE and DEPARTURE. An ARRIVAL event represents a new session event (which may be a circuit-switched voice call session, an IMS VoIP session or an IMS data session).

Table III. The Notations of allocation scheme fllow-chart

Notation	Description	
Event	We create three types of event to simulate each session	
	ARRIVAL	To generate a new session
	UPDATE	To handle an existing session that requests more credit
	DEPARTURE	To handle a session termination
Event's parameter	Detail of sessions record in each event	
	TimeStamp	The arrival time of an event
	ResudualTime	The residual time of a session
	HoldingTime	The holding time of a session
C	The amount of total credit when an online charging account is newly refreshed.	
C_B	The emergency credit threshold provided by OCS for the last session.	
C_r	The remaining credit of an online charging account in the usable duration.	
C_now	Credit remains now.	
C_extra	Extra pay for over time call.	
$1/\mu_s$	The expected session holding time for an IMS session. (minutes/ session)	
λ_a	The expected session arrival rate for an IMS session. (sessions/ minute)	
ts	The service time of a call event.	
ta	The arrival time of a call event.	
AllScheme	Allocation schemes: BQF, MCF and AQF.	
	BQF	Best Quality First
	MCF	Minimum Credit First
	AQF	Average Quality First
$c_{m,n}$	Credit charged per time unit for a session with type n and quality class m ($1 \leqq m \leqq M$). Here, class M represents the highest quality-class.	
T_avg	Reserve average service time.	
T_allocate	The time period system allocate for an *UPDATE/ARRIVAL* event.	
C_allocate	The credit cost when *T_allocate* allocate for an *UPDATE/ARRIVAL* event.	
NumDeparture	Departure session number before exhausting the credit of an online charging account.	
β	The probability threshold of the expected cumulative holding time in a session.	
n_c	Number of sessions supported.	
T_c	The expected lifetime of an online charging account.	
P_E	The probability of the expected cumulative holding time in the last session.	

The OCS reserves credit for a time period (*T_allocate*) to this session. When the PDNGW handling this session consumes all credit, the PDNGW requests more credit from the OCS. In our simulation, we generate an UPDATE event to simulate the operation of credit allocation from OCS to an existing session. When the user terminates a session, we generate a DEPARTURE event to simulate the session termination. The simulations flow-chart is shown in Fig. 2 and the notation used in the simulation is explained in Table III.

IV. NUMERICAL RESULTS

Based on the simulation model proposed in Section III, we evaluate the refresh cycle and the number of sessions served in an online charging account by considering three charging rates (c_{min}, c_{med}, c_{max}), which typically can be referred as three QoS classes (*Bronze*, *Silver*, *Gold*) in telecom market. Table IV list the credit charged for each QoS class per time unit.

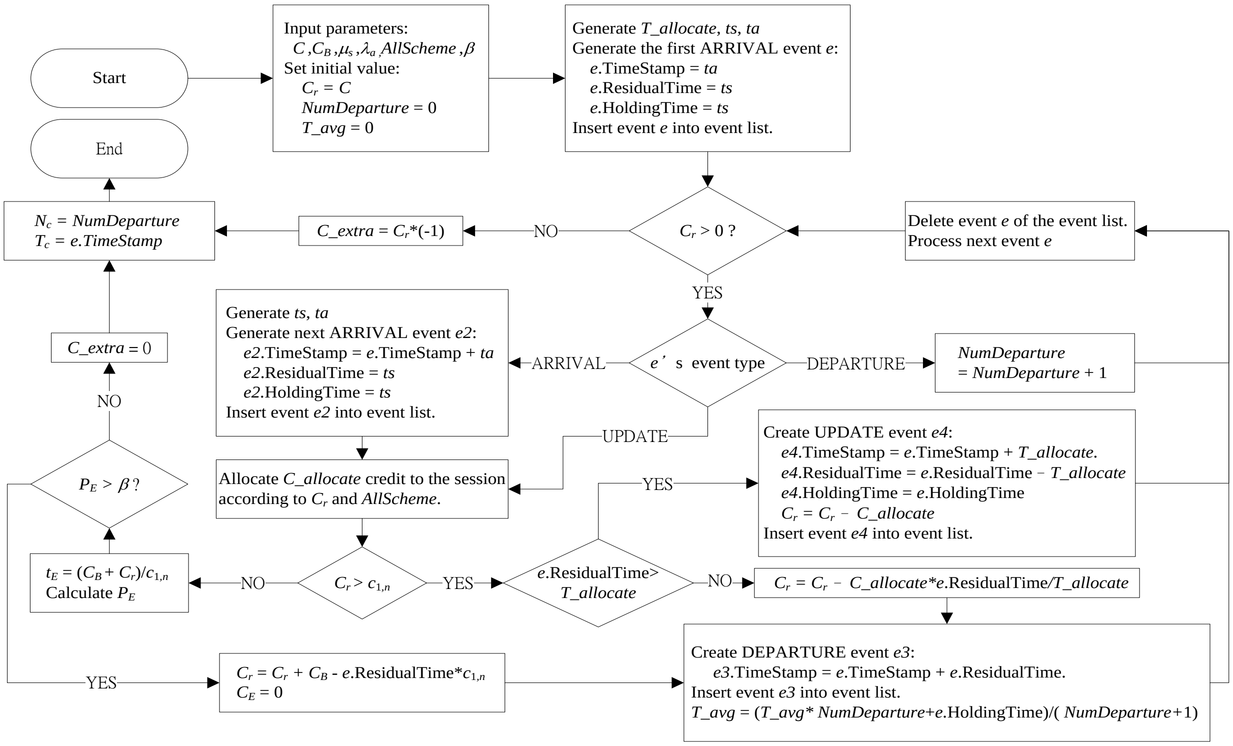


Figure 2. The simulations flow-chart.

Table IV. Credit charged for three QoS classes in simulation

QoS Class (*Billing plan*)	**Notation**	**Credit Charged**
Gold	c_{max}	6 unit/ per minute
Silver	c_{med}	4 unit/ per minute
Bronze	c_{min}	2 unit/ per minute

Based on the amount of credit C in a newly refresh account, the charging unit $c_{m,n}$, session completion rate μ_s (sessions/minute) and arrival rate λ_a (sessions/minute), we calculate the number of departure sessions by Eq. (1) and the analytic lifetime by Eq. (2). Let T_a and N_a be the upper bounds of the lifetime and the number of completed session in an online account.

First, we calculate the lifetime of an online charging account with credit C. For example, when C=NT500, we want to know the expected lifetime when a user consumes all the credit and when he/she needs to refresh the account. Sometimes, a user does not notice that his/her account is going to deplete before making a new call (session), in this case, he/she wants to complete the call first and performs an account refresh later. In Taiwan, we notice that there is a setup fee when a new (prepaid) account is setup, or a contract is signed between a user and the operator, and a user will not shift to another operator easily. However, the last call can be a very important (emergency) call to a user and the user will like to borrow some emergency credit (C_B) before the account refresh. C_B=0 implies that no emergency credit will be provided to the user. Usually, a C_B setting that less than the account setup fee is reasonable. Providing emergency credit to a user increases user satisfaction without taking a big risk in revenue loss. Hence in Eq. (1), we consider that a user can use $C + C_B$ credit in his/her account with the session completion rate (μ_s), charging unit $c_{m,n}$, based on different service types. By considering the session completion rate, the upper bound for the number of sessions (N_a) completed in an online account can be computed as (1).

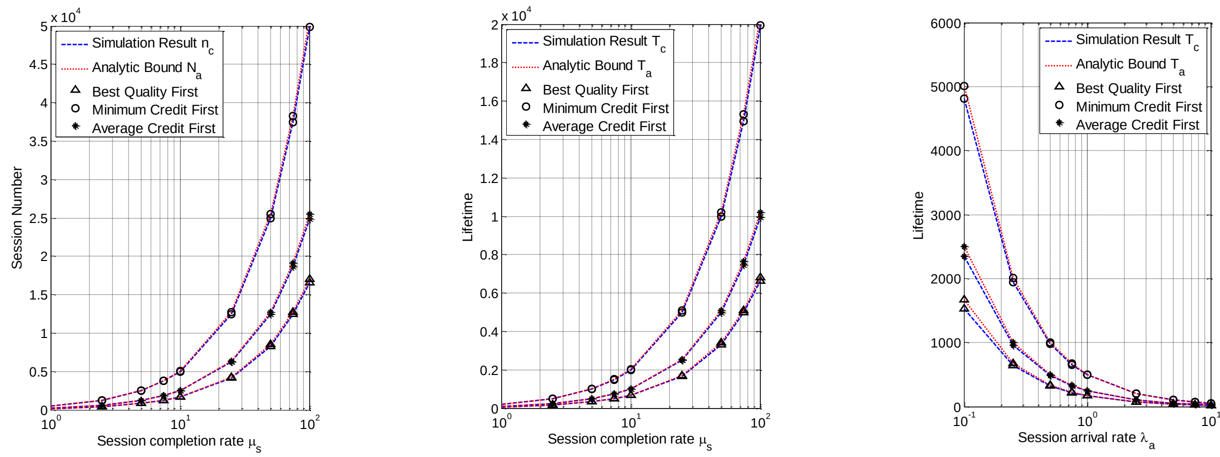
$$N_a = \frac{(C_B + C)}{c_{m,n}} \times \mu_s \quad (1) \qquad T_a = \frac{(C_B + C)}{c_{m,n}} \times \frac{\mu_s}{\lambda_a} \quad (2)$$

Based on (1), by considering the inter-arrival rate (λ_a) of the session, the upper bound of the lifetime (refresh cycle) of an online account with initially credit (C) and emergency credit (C_B) is computed as

A. *Performance for the credit allocation schemes*

In this subsection, the effect of session completion rate (μ_s) is illustrated in Fig. 3(a) and Fig. 3(b), where C=1000, C_B=2, β=0.5, λ_a=2.5 (sessions/minute) and the session completion rate (μ_s) varies from 0 to 100. Fig. 3(a) shows that as the session completion rate increases, the number of session completed in an online charging account also increases. We also observe that MCF scheme can serve more sessions than other schemes while AQF scheme serves more sessions than BQF scheme. Because MCF scheme chooses the quality that the least credit charged per minute as its top priority. To validate the accuracy in simulation model, we compute the analytic upper bounds for n_c in the MCF, BQF, and AQF schemes based on Eq. (2). The analytic results are very close to the simulation bounds as shown in Fig. 3 (a).

Fig. 3(b) shows that as μ_s increases, the refresh cycle (lifetime) also increases. We also observe that the lifetime in MCF scheme is the longest among three schemes;



(a) The session number of increasing μ_s. (C =1000, C_B =2, β =0.5 and λ_a =2.5 minutes/ session)

(b) The lifetime of increasing μ_s. (C =1000, C_B =2, β =0.5 and λ_a =2.5 minutes/ session)

(c) The lifetime of λ_a. (C =1000, C_B =2, β =0.5 and μ_s =1session/ minutes)

Figure 3. The session number/ lifetime of increasing session arrival/completion rate.

the lifetime in AQF scheme is longer than that in BQF scheme. Based on Eq. (1), we compute analytic upper bound in lifetime of an online charging account in the MCF, BQF, and AQF schemes. Clearly, the analytic results are very close to the simulation results as shown in Fig. 3(b).

Based on the above analytic model and the simulation assumptions, Fig. 3(c) plots the refresh cycle (T_c) against different session arrival rate (λ_a), where the initial credit amount C =1000, C_B =2, β =0.5 and μ_s =1. First, Fig. 3(c) illustrates that the online charging account lifetime decreases as the session arrival rate increases. We also observe that among three kinds of credit allocation schemes, the account lifetime of MCF scheme has the largest value while that of BQF scheme has the lowest value. This observation is consistent with what the users expect when they select their credit allocation scheme. To validate the accuracy in simulation model, we further compute the analytic upper bound in lifetime of an online charging account in MCF, BQF, and AQF schemes based on Eq. (1). Clearly, the simulation results are close to the analytic upper bounds.
We observe that the session arrival rate has no effect on session number. Here, in BQF, AQF and MAC schemes, the session number is around 167, 250 and 500 sessions where the session arrival rate varies from 0 to 100. The initial credit amount C =1000, C_B =2, β =0.5 and μ_s =1.

B. Performance for the emergency credit

In this subsection, we investigate the last session continuity in each refresh cycle. By providing an extra amount of emergency credit to the user, we can increase the service continuity in the last session before an account is refreshed. Fig. 4 investigates the effect of the extra credit threshold with two different session completion rates (μ_s=1, μ_s=0.5). As the emergency credit threshold (C_B) increases, the extra cost increases until it reaches a peak value equal to c_{min}/μ_s, which is the minimum credit cost of the average call session completion time.

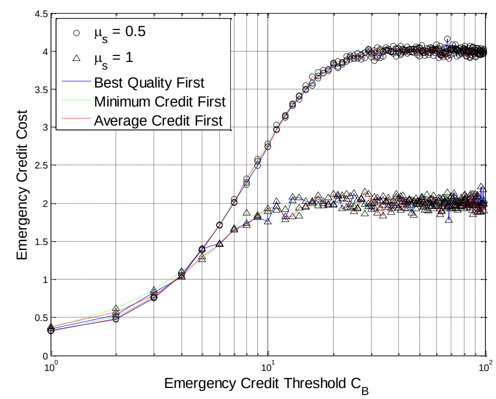


Figure 4. The online charging account emergency credit cost of increasing extra credit threshold.

C. A case study based on session statistics in Taiwan

In this subsection, we use simulation experiments to investigate the effects on the service session distribution for voice and data applications. Studies on non-VoIP mobile phone calls indicated that the mean call holding time is 40.6 s during working hours and 63.3 s during non-working hours, respectively [14]. Measured data from Taiwan's mobile operators indicate that the mean call holding time is 45 s. The mean VoIP call holding time distribution of Taiwan-mobile is 110 s [15]. Study on data applications indicated that the WWW network or data service network can be modeled by Pareto distribution. Table V lists the charging rate for three quality classes and three allocation schemes based on Chunghwa Telecom data [10].

In this subsection, we simulate the VoIP call with average session holding time 110 seconds, the non-VoIP call with average session holding time 45 seconds and the data session holding time that follows a Pareto distribution (location=1, scale=0.8). The credit charged for each time unit (i.e., 6 seconds) in each quality class and service type is listed in Table V.

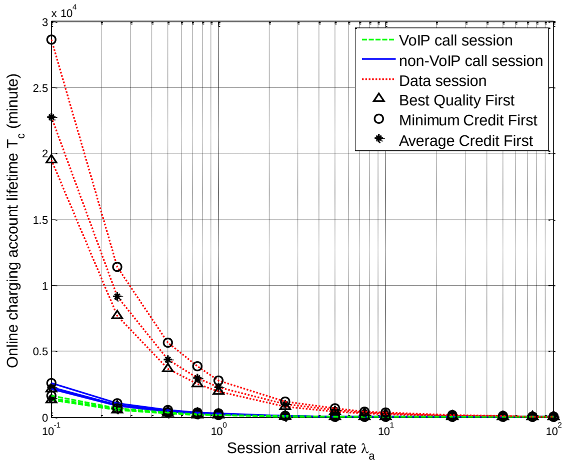


Figure 5. The online charging account lifetime of increasing session arrival rate

Table V. Charing rate for three quality classes and session types

Service type (N=3)	Quality class (M=3)			Session distribution
	1	2	3	
(n=1) non-VoIP call session	0.59/6s	0.56/6s	0.5/6s	Exponential (mean:45s)
(n=2) VoIP call session	0.4/6s	0.36/6s	0.32/6s	Exponential (mean:110s)
(n=3) data session	0.3/6s	0.25/6s	0.2/6s	Pareto (ON): location 1, scale 0.8

Fig. 5 shows an online credit account lifetime varies a lot within different kinds of sessions. The lifetime in VoIP call sessions has the lowest value among all. It is clear that the account lifetime of pure VoIP call environment is shorter than that in pure non-VoIP call environment, since subscribers tend to make a long VoIP call session due to an attractive cheaper rate for the lower equipment cost in IP-based platform. Surprisingly, we observe that the lifetime in pure data session environment has the highest value among all. It is because the charging rate for per data session is very low in Taiwan. Because the telecom operation needs to complete data service with other ISP, the data rate charging in Chunghwa Telecom is very low compared with making VoIP or non-VoIP calls.

V. CONCLUTION

Based on the PCC architecture, three credit allocation schemes, Minimum Credit First (MCF), Average Quality First (AQF) and Best Quality First (BQF) are proposed and investigated in this paper. Specifically, we study the expected number of sessions supported and the expected lifetime of an online charging account for each scheme. Through extensive simulation, we observe that among three kinds of credit allocation schemes, the account lifetime of MCF scheme has the largest value while that of BQF scheme has the lowest value. As the session completion rate increases, the number of session completed in an online charging account also increases. We also observe that MCF scheme can serve more sessions than other schemes while AQF scheme serves more sessions than BQF scheme.

Based on the above observations, when there are more multimedia contents which need higher quality to support or need to occupy a longer service time, the initial amount of an online account should be raised to a higher level so that the user will not need to refresh his/her account so frequent. On the other hand, the operator should provide more promotion or rebate to users to increase their motivation for account refresh.

VI. ACKNOWLEDGEMENT

The work was supported by the National Science Council, Taiwan, R.O.C., under Grant NSC 98-2221-E-006-095020 and NSC 99-2221-E-006 -122, and in part by the Chung-Hwa Telecom project.

REFERENCES

[1] Y.-B. Lin and A.-C. Pang, Wireless and Mobile All-IP Networks, *Wiley*, 2005.

[2] ETSI, Telecommunications and Internet converged Services and Protocols for Advanced Networking (TISPAN); NGN Functional Architecture, ETSI ES 282 001, Release 3.4.1, 2009.

[3] 3GPP. 3rd Generation Partnership Project; Technical Specification Group Service and System Aspects; Architectural Requirements for Release 1999 (Release 1999). Technical Specification 3G TS 23.121 Version 3.6.0 (2002-06), 2002.

[4] S.-I. Sou, Y.-B. Lin and J.-Y. Jeng., "Reducing Credit Re-authorization Cost in UMTS Online Charging System," *IEEE Trans. Wireless Commun*., vol. 7, no. 9, pp. 3629-3635, Nov. 2008.

[5] S.-I. Sou, H.-N. Hung, Y.-B. Lin, N.-F. Peng, and J.-Y. Jeng. " Modeling Credit Reservation Procedure for UMTS Online Charging System," *IEEE Trans.Wireless Commun*., vol. 6, no. 11, pp. 4129-4135, Nov. 2007.

[6] S.-I. Sou, Y.-B. Lin, Q. Wu and J.-Y. Jeng, "Modeling of Prepaid Mechanism of VoIP and Messaging Services," *IEEE Trans. Veh. Technol*., vol. 56, no. 3, pp. 1434-1441, May 2007.

[7] 3GPP. 3rd Generation Partnership Project; Technical Specification Group Service and System Aspects; Telecommunication management; Charging management; Online Charging System (OCS): Applications and interfaces. Technical Specification 3G TS 32.296 Version 9.1.0 (2009-12), 2009.

[8] H. Waterman. "Bill Shock: Verizon, Family Spar over $18,00", http://www.von.com/news/

[9] 3GPP. 3rd Generation Partnership Project; Technical Specification Group Service and System Aspects; Telecommunication management; Charging Management; IP Multimedia Subsystem (IMS) charging. Technical Specification 3G TS 32.260 Version 9.4.0 (2010-06), 2010.

[10] Chunghwa Telecom Rate System, http://www.cht.com.tw/PersonalCat.php?Page=innerFrame&CatID=523

[11] 3GPP. Technical Specification Group Services and System Aspects; Policy and charging control architecture. 3G TS 23.203 Version 9.5.0 (2010-06), 2010.

[12] S. Zaghloul, W. Bziuk, and A. Jukan, "Signaling and Handoff Rates at the Policy Control Function (PCF) in IMS," *IEEE Commun. Lett.*, vol. 12, no. 7, pp. 526–528, 2008.

[13] Y.-B. Lin, and S.-I. Sou, Charging for Mobile All-IP Telecommunications, *John Wiley & Sons*, 2008.

[14] F. Barcelo and J. Jordan, "Channel holding time distribution in public telephony systems (PAMR and PCS)," *IEEE Trans. on Veh.Technol.*, vol. 49, no. 5, pp. 1615-1625, Sep. 2000.

[15] W.-E. Chen, H.-N. Hung, and Y.-B. Lin, "Modelling VOIP Call Holding Times for Telecommunications," *IEEE Network*, vol. 21, no. 6, pp. 22-28, Nov./Dec. 2007.