Overhead Reduction for Real-Time Charging in UMTS Networks

Bálint Dávid ARY, Gábor DEBREI and Sándor IMRE*

Budapest University of Technology and Economics, Department of Telecommunications
H-1117. Budapest, Magyar tudósok körútja 2. Hungary
E-mails: balint.ary@isolation.hu, debg@freemail.hu, imre@hit.bme.hu

Abstract—**With the advent of UMTS and with the fulfillment of the new generation network concept content and services will be offered by 3rd party providers, thus a proper and real-time charging mechanism is needed. The current real-time charging solutions are generating huge overhead, which spoils the network utilization, and deflates the incomes of the network operator. In our paper we present a novel charging model which is compliant with the related standards, and reduce the network overhead needed for real-time charging.**

Index Terms—**Charging, UMTS, Telecommunication Management, Network Overhead**

I. Introduction

At the beginning of the 21st century a new, packet based mobile telecommunication network was (or being) introduced in many countries in Europe. The Universal Mobile Telecommunications System (UMTS) contains major changes to the well known GSM (Global System for Mobile Communications) network. The huge number of available functions and the standardized interfaces give the possibilities to 3rd party providers to use the infrastructure of the UMTS system, and offer value-added services to the subscribers. Because these providers are financially isolated from the network operator, and because of the continuous growth of the number of pre-paid services and subscribers, an accurate and real-time charging mechanism is indispensable. The corresponding standards are defining the non-real-time (offline), and real-time (online) charging methods, functions and mechanisms, which solve the above mentioned problems. However, offline charging is not suitable for pre-paid users, and the real-time approach invokes significant network overhead.

In this paper we present our simple, but novel concept, which ensures real-time charging with a relatively small administrative traffic. As mentioned above, charging is crucial in the telecommunication world, but its realization is (more or less) differs in every network operator domain. Our concept is a high level model, thus can be adopted in every mobile telecommunication network. This paper organized as follows: Section 2 summarizes the current state of UMTS charging. We introduce our new concept in Section 3. Finally Section 4 concludes the work and presents future directions. Due to the page limitations, the short survey of business models, and the state of the art technical challenges section was removed. In this paper the terms "subscriber" and "user" are used as synonyms.

II. Charging in UMTS Networks

A proper charging model must be constructed according to the related standards. 3GPP specifies guidelines and recommendations for UMTS charging systems; these guidelines include the architecture, the main role and functionality of the network elements.

Standards define two different payment types, and two different charging methods. The payment type indicates the time of the settlement of the bill, and so, pre-paid and post-paid modes are distinguished. The charging method (offline and online) indicates how the subscribers' bills are handled. These methods are described in the following subsections at large.

A. *Offline charging*

In data measurement the gateway (GGSN – Gateway GPRS Support Node) and the inner-nodes (SGSN – Serving GPRS Support Node) are sending charging information to the Billing System (BS). The Charging Trigger Functions (CTFs) of the nodes are generating charging events based on the observation of network resource usage. The Charging Data Function (CDF) receives charging events from the Charging Trigger Functions, and then uses the information contained in the charging events to construct the standardized CDRs (Charging Data Records). These records are sent to the Charging Gateway Function (CGF), which acts like a storage buffer, cleans, and preprocesses the CDRs, and then sends the processed CDRs to the Billing System (Fig. 1). Since this information is collected after the event / service, and sent trough a widespread network, real-time charging is not possible. According to the standards, post-paid users can limit their account for a specific service, and so, a real-time charging method should be used for pre-paid users and for post-paid users with credit limit. To ensure this, online, real-time charging should be applied. Standards also define the trigger events to send these Charging Data Records [1]. The events include:

– determinate data amount or time-interval,
– the change of charging conditions,
– the change of QoS, tariff, position or cell
– and the closure of voice, data or multimedia sessions.

Because these charging records carry information about the services required, the functionality of the CDRs extends beyond charging. With CDRs it's possible to analyze

* The research is supported by PHOENIX project and OTKA F042590.

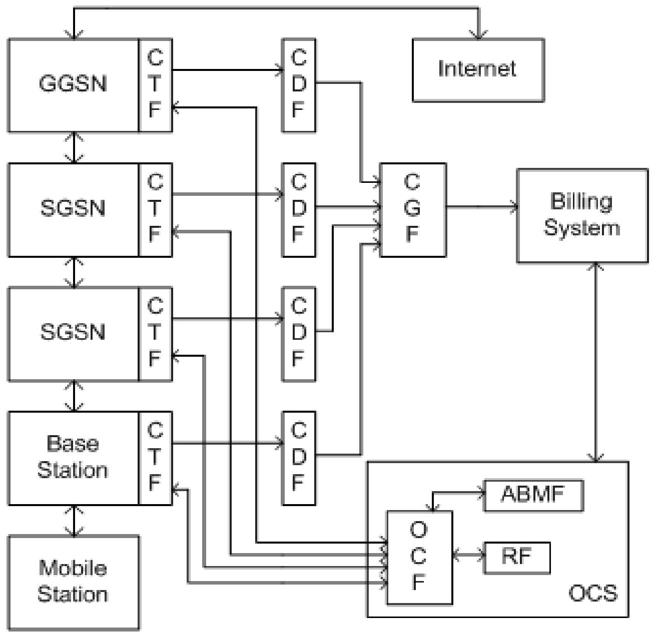

Fig. 1. UMTS charging architecture

service-utilization, and gain statistical information about the services and content [1]. By archiving the CDRs, the user-complaints can also be easily settled.

B. *Online charging*

In online mode (real-time charging mode), the Billing System allocates proper amount of money to the serving nodes, and these nodes are deflating the delegated units simultaneously with the service providing. The Online Charging System (OCS) is responsible for the proper and continuous unit allocation / reservation. The Online Charging Function within the OCS communicates with the CTFs of the serving nodes, and granting credits for the service; these credits are deducted from the user's account. This communication is bidirectional, as the CTFs also generating charging events, and credit demands for the OCF (Fig. 1). If the service terminates before all credits consumed, the network elements are retransferring the remaining credits to the OCF. To assure continuous service delivery, if the users do not terminate the service, a new amount of granted credit should be sent to the serving network element before the previous one runs out [3] [4].

The OCS also includes the Account Balance Management Function (ABMF) and the Rating Function (RF). The Account Balance Management Function is the location of the subscriber's account balance within the OCS. The Rating Function is used to determinate the value of the network resource usage and responsible for the

- rating of data volume,
- rating of session / connection time,
- and for rating of service events.

III. Mode-Switching Model

For correct modeling it is obligatory to suit the related standards. The optimal model can be developed using the proper determination of the free parameters. Such variable parameters are the amount of data and / or time that triggers the CDR generation, the amount of granted credit during unit reservation, and the exact operations of the network elements.

Our idea was not to glue the charging mode to the type of the payment (pre-paid, post-paid), but to dynamically switch between offline and online charging (if online charging is required) considering the demanded service, and user's account as well. Moreover, the overhead of the continuous unit reservation can also be reduced, by granting units only once. The quality of service should also be supervised, in order to charge services properly.

In our model, we assign a service specific limit to every service offered. If the user's account is above this limit, then charging is done in offline mode. If the subscriber's account drops below this limit, the online charging mechanism is applied (if required), and we grant all the consumable credit to the serving network element. In multi-task systems it is possible to access more than one service at a time. In such cases, when the account drops below the limit, we shall delegate the credits to multiple network elements. A good solution is to distribute the account among the services with statistical methods, considering the money-consumption and properties of the services, and the behavior of the user.

A. *Estimation of Network Overhead*

Let us denote d_{data} for the amount of useful data transmitted. If the size of a CDR is d_{cdr} and the amount of data that triggers the CDR generation is referred as t_{cdr}, then the quotient of the charging messages and the useful data ($O_{offline}$) is

$$O_{offline} = [\ (\ d_{data} / t_{cdr}\ ) \cdot d_{cdr}\ ] / d_{data} = d_{cdr} / t_{cdr} \quad (1)$$

in offline mode (if we use the same unit for d_{data}, d_{cdr} and t_{cdr}). Similar to this, the quotient of the charging messages and the useful data in online mode (O_{online}) is

$$O_{online} = [\ (d_{data} / t_{ur}) \cdot d_{ur}\ ] / d_{data} = d_{ur} / t_{ur}\ , \quad (2)$$

where d_{ur} is the size of the unit reservation message and t_{ur} is the amount of granted data. Since the unit reservation message should contain more or less the same information as the CDRs, we assume that d_{ur} is equal to d_{cdr}. In online charging, if the service reserves a large amount of credit from the user's account, access to additional, parallel resources could be denied, because there is no credit left on the account for another resource usage request; even if some service terminates afterwards, and the unused credits are returned to the users. In light of this, a more frequent unit reservation, with a smaller amount of credit should be applied. Because CDRs indicate the used services/data, this problem doesn't occur during offline charging. As follows, t_{ur} is smaller than t_{cdr}, and thus online charging causes bigger network overhead, than offline charging:

$$O_{\mathrm{offline}} = d_{cdr} / t_{cdr} < d_{ur} / t_{ur} = O_{online}\ . \quad (3)$$

Our idea is to apply offline charging for pre-paid users if their account is far above zero and for post-paid users with credit threshold, if their account is far from the specified limit. If the user's account is close to zero (or to the specified credit limit) online charging should be applied. A crucial question is to determine the threshold limit to switch between offline and online charging.

B. Mode-Switching Limit

Let us define a function called unit consumption speed

$$C(T)\ , \tag{4}$$

having the measure of [*unit/sec*], which represents the consumed units in one second. The consumption rate depends on time to give the possibility to the operators to assign different prices to different time of the day and week for traffic shaping reasons. The consumed unit and money can be calculated from the consumption rate by means of the following equations

$$unit = C(T)\cdot t\ , \tag{5}$$
$$money = unit\cdot R(T)\ , \tag{6}$$

where $R(T)$ represents the relation [2] between unit and money. The time-dependence of this function can be used to change the price of the units in case of inflation or discounts, or to apply different prices for different groups of users. Although the time dependence of the price can be divided into consumption speed and rating, it is not necessary, and it depends on the needs of the network operator.

Let T_c represent the time needed to query the user's proper account. The network elements are sending the CDRs usually in bigger time-intervals and the Billing System debit the user's account periodically. T_c represents these intervals. With these notations and definitions the limit for mode-switch can be calculated. In ideal case it is

$$L = C(T)\cdot T_c\ . \tag{7}$$

If we own more units on our account than L, the charging is done offline with small network overhead; otherwise accounting is done online, with unit reservation. If we require more than one service at a time, the limit can be calculated by the sum of the limits of the services:

$$L = \sum L_i\ . \tag{8}$$

To reduce the network overhead, all credits below this threshold can be reserved. In case of multiple service demands, the units can be distributed to the serving network elements with the rate of the service's consumption speed. A re-sharing should be done every time a service ends, a new service started, or when an event based service occurs (Short Messaging Service for example). In order to ensure this, new functionality is required. The online charging function (OCF) should be able to force the network elements to retransfer the currently unused credits. After the transfer, the online charging function could re-share the credits among the services considering the new circumstances.

When a fix consumption speed can not be assigned to the service (browsing, or interactive content), the average consumption speed should be determined using various statistical models.

C. Propagation Delay

The events occurring in a distributed, wide network (signaling, queries) have propagation delay, which is not constant in general. If we want to determine the mode-switching threshold properly, we have to consider the time needed the query the account (T_c) and to switch between modes (T_d), together with the variation of these values (T_{cj} and T_{dj}):

$$L = C(T)\cdot(T_c+T_{cj}+T_d+T_{dj})\ . \tag{9}$$

To ensure accurate charging, we should count with the maximum values of the jitters (T_{ci} and T_{di}). If we want to reduce the values of the mode-switching limits (in order to reduce the network overhead), we shall count with smaller values (with the expected value for example). In this case the possibility of users gaining more service than they paid for can be calculated from the distributions of the jitters.

In case of re-sharing the control messages should be labeled with proper time-stamps to be able to charge the services gained during the retransfer and mode switching process.

The mode-switching thresholds can be calculated offline for every service offered, and the system can use these pre-calculated values to switch between the charging modes. However, the actual limit can be dependant on the time of the day and on the user profile (discounts for group of users, statistical behavior for interactive content).

TABLE I
PARAMETERS FOR VIDEO-STREAMING

Parameter name	*Value*
Data rate of the streaming-video	24576 bit/sec
Price of the service ($C(t)$)	0.2 €/min
Size of a CDR (d_{cdr})	102400 bit
Size of an Unit Reservation message (d_{ur})	81920 bit
Amount of data that triggers the CDR generation (t_{cdr})	7372800 bit
Amount of granted data in one Unit Reservation message (t_{ur})	1474560 bit
Time needed to query the user's account and switch between charging modes ($T_c+T_{cj}+T_d+T_{dj}$)	7.5 sec

D. Calculations for Streaming-Video

To demonstrate the profits of our model let's assume, that a user requires a streaming video service with 4 € on his/her account. The used parameters of the network and the streaming-service are listed in Table I. With these parameters, the network overhead in offline and in online mode can be calculated, such as the mode-switching limit using (1), (2) and (9):

$$O_{offline} = 102400\ /\ 7372800\ =\ 0.01388\ , \tag{10}$$
$$O_{online} = 81920\ /\ 1474560 = 0.05555\ , \tag{11}$$
$$L = 0.2\cdot 7.5 = 1.5\ . \tag{12}$$

With 4 € the user can stream twenty minutes of video. During that period four CDRs or twenty Unit Reservation messages are produced. In offline mode, the user's account drops below 1.5 € after 900 seconds (as the system receives the third CDR). With mode-switching three CDRs and one Unit Reservation message are sent, and so the network overhead is:

$$O_{mode\text{-}swith} = 0.01319\ . \tag{13}$$

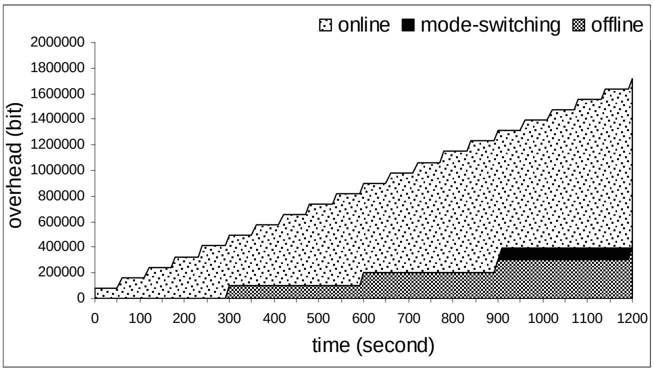

Fig. 2. Size of the charging overhead during the service

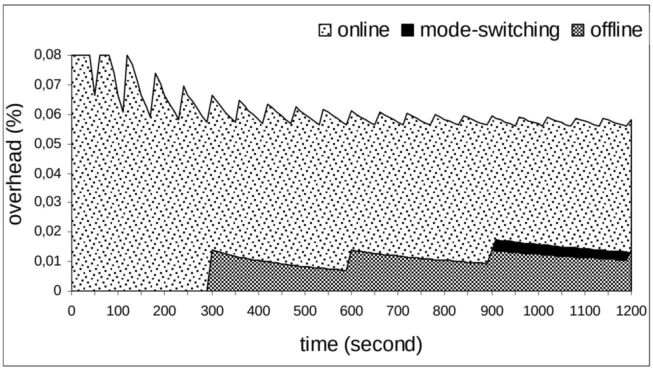

Fig. 3. Percentage of the charging overhead during the service

Fig. 2 and Fig. 3 demonstrate the graphical representation of the network overhead during the streaming-service. The overhead of the mode-switching model is equal to the overhead of the offline mode till the account drops below 1.5 €. At the end of the service (1200 second) the overhead of the mode-switching model is smaller than as it is with the other two methods.

IV. Conclusions and Future Plans

In our study, we gave a small summary about the current state of the 3GPP standards – according to the latest documents. We also gave a new charging model to increase the network utilization. The model operates in such a way, that charging is made in the network offline, without a need for a real-time approach, to a large volume of users (who have more money on their account than the critical amount). This method invokes low CDR transfer, and low network overhead. Charging critical users is more complicated, but supported by 3GPP standards. With this idea the necessary network overhead can be decreased. Moreover, with a small function extension and statistical estimation, the overhead can be further reduced.

The current study is a part of a complex research, which is meant to develop a universal, integrated charging system, which could handle the existence of 3rd party providers. The high level functionality of the solution and the protocols should be also developed and revised. In the future, in order to develop the complete charging model, it is required to work out the exact method of measuring the amount and quality of the data flow. The amount and quality of the service should be derived from these parameters, but this solution needs further studies. The determination of the statistic parameters of the services and users are also necessary.

References

[1] ETSI TS 122 115 V5.3.0 (2003-06). “Digital cellular telecommunications system (Phase 2+); Universal Mobile Telecommunications System (UMTS); Service Aspects Charging and billing (3GPP TS 22.115 version 5.3.0 Release 5).” Technical report, 3GPP, 2003.

[2] ETSI TS 132 200 V5.7.0 (2004-06). “Digital cellular telecommunications system (Phase 2+); Universal Mobile Telecommunications System (UMTS); Telecommunication management; Charging management; Charging principles (3GPP TS 32.200 version 5.7.0 Release 5).” Technical report, 3GPP, 2004.

[3] Telecommunication management; Charging management; Charging architecture and principles (Release 6).” Technical report, 3GPP, 2004.

[4] 3GPP TS 32.260 V2.0.0 (2004-12). “3rd Generation Partnership Project; Technical Specification Group Services and System Aspects; Telecommunication management; Charging management; IP Multimedia Subsystem (IMS) charging; (Release 6).” Technical report, 3GPP, 2004.

[5] UMTS Forum Report No 11. “Enabling UMTS Third Generation Services and Applications.” Technical report, UMTS Forum, October 2000.

[6] UMTS Forum Report No 21. “Charging, Billing and Payment Views on 3G Business Models.” Technical report, UMTS Forum, 2002.

[7] Maria Koutsopoulou, Alexandros Kaloxylos, Athanassia Alonistioti, and Lazaros Merakos. “Charging, Accounting and Billing Management Schemes in Mobile Telecommunication Networks and the Internet.” *IEEE Communications Surveys, First Quarter 2004*, 6(1), 2004, pp 50-58.

[8] Susana Schwartz. “Next-Gen Rating: It Will Be Only As Good as the Network.” *Billing World & OSS Today*, February 2003, pp16-22.

[9] John Cushnie, David Hutchison, and Huw Oliver. Internet Services: Infrastructure for Charging and Billing. In *Advanced Internet Charging and QoS Technology (ICQT’01) Workshop WS3 of Infomatik 2001*, University of Vienna, Austria, September 2001.

[10] B. Stiller, G. Fankhauser, B. Plattner, and N. Weiler. Charging and accounting for integrated internet services - state of the art, problems, and trends. In *Proc. INET '98,* Geneva, Switzerland, July 1998.

[11] Harri Hakala, Leena Mattila, Juha-Pekka Koskinen, Marco Stura, and John Loughney. Diameter Credit-Control Application. Technical report, AAA Working Group, 2004.

[12] V. Gazis, M. Koutsopoulou, C. Farmakis, and A. Kaloxylos. A Flexible Charging and Billing Approach for the Emerging UMTS Network Operator Role. In *ATS 2001*, Seattle, Washington, USA, April 2001.

[13] Zsolt Butyka, Tamás Jursonovics, and Sándor Imre dr. A Real-Time Charging Model for UMTS Mobile Network. In *ETIK conference*, Budapest, Hungary, 2004.

[14] P. Pagliusi and C. J. Mitchell. PANA/GSM authentication for Internet access. In *SympoTIC ’03, Joint IST Workshop on Mobile Future and Symposium on Trends in Communications*, Bratislava, Slovakia, October 2003.

[15] Lucent Technologies. IP Multimedia Subsystem (IMS) Service Architecture, 2004. Whitepaper.

[16] Service Level Corporation. *Rating Support For Revenue Opportunities With GPRS & UMTS Services*. Whitepaper, http://www.servicelevel.net/ rating_matters/GPRS_and_UMTS_rating .htm, 2005-01-10, 00:25.