An overview of online charging in 3GPP networks: new ways of utilizing user, network, and service related information

Tomislav Grgic1* and Maja Matijasevic1

1University of Zagreb, Faculty of Electrical Engineering and Computing, Unska 3, Zagreb, Croatia

SUMMARY

Modern charging systems use user, network, and service related information when performing online charging. Compared, however, to the overall information available and used in network management processes as a whole, charging systems only use a limited subset. This work is motivated by the challenge to identify which information is used, and how it is used in online charging related processes, and also to explore whether it could be utilized “better” or “smarter” to improve future online charging systems functionality. We do not attempt to predict which information will be utilized in such systems and for what purpose, but instead we summarize open issues in view of the emerging trend of exploiting the user, network and service related information in service provisioning. We focus on the latest 3GPP standards and review relevant research papers, and propose three key aspects of online charging, with respect to information utilization: a) signaling aspect, b) inter-domain aspect, and c) service- and component-based aspect. We present a state of the art review by grouping the works found in the literature based on the aspects they are associated with, and compare them based on proposed comparison criteria. The discussion presented at the end of the paper indicates three common open issues, namely: 1) lack of common charging information specification and structure, 2) lack of mechanisms for information sharing between stakeholders in the service delivery process, and 3) lack of a common framework for sharing information while protecting user privacy.

Copyright © 2012 John Wiley & Sons, Ltd.

1. INTRODUCTION

Modern telecommunication systems continuously track and regularly access information regarding users and services, such as user’s location or subscription data. This paper is motivated by the challenge to determine to which extent is this information used for online charging (i.e., a real time service cost calculation) [1] in the Next Generation Network as defined by the Third Generation

*Correspondence to: University of Zagreb, Faculty of Electrical Engineering and Computing, Unska 3, HR-10000 Zagreb, Croatia. E-mail: tomislav.grgic@fer.hr

Copyright © 2012 John Wiley & Sons, Ltd.

Prepared using nemauth.cls [Version: 2010/05/13 v2.00]
Partnership Project (3GPP) and what are the benefits thereof. We examine the latest 3GPP online charging standards and the most recent research works in the area of online charging in mobile networks. By analyzing the input information they use for charging, compared to the overall available information, we identify open issues in information utilization and discuss potential development courses for charging systems. A subset of this information is already used by online charging systems (e.g., available credit, user subscriptions), but there is also a part that is not necessarily used (e.g., user location, preferences, information about other users nearby). We believe that by fully utilizing the information, online charging may be functionally improved. For example, assume that a mobile user has a credit of 50 minutes for watching the television program provided by a third party streaming service provider. At some time, the user requests streaming of a previously recorded soccer game (which normally lasts 90 minutes). “Traditional” online charging systems would allow initiation of the stream, but would stop the stream after 50 minutes that the available credit would allow. However, if the knowledge about the allowed 50 minutes of stream is shared between the online charging system and the service provider (if disclosure of this information is permitted by the user), the streaming service provider may offer a shorter version of the game to the user, containing only the most interesting parts.

The focus of the paper is the 3GPP Online Charging System (Releases 10 and 11) and other related 3GPP architectures. Additionally, recent research efforts in online charging are elaborated, that are - or may be - relevant for the 3GPP charging architecture. The overview addresses three aspects of online charging: a) charging signaling aspect, b) inter-domain charging aspect, and c) the service- and component- based charging aspect. To the best of our knowledge, none of the charging related overview papers includes discussion about the input information used for charging as their central point. Furthermore, most of the online charging related overviews are somewhat dated [2, 3, 4, 5, 6, 7] and do not represent current achievements in charging. In a very recent and thorough paper by Kuhne et al. [8], general requirements for future charging systems are identified, but information utilization is not particularly discussed. This paper complements that work in this respect.

The rest of the paper is structured as follows: in Section 2, charging-related terminology is explained in more detail, including related processes, business relationships, and charging data used. Section 3 elaborates on 3GPP Releases 10 and 11 Online Charging System, including charging requirements and charging information used, and provides a background of charging-related architectures and protocols specified by the Internet Engineering Task Force (IETF) and used in 3GPP. Section 4 explains contributions in the area of online charging of recent research projects, grouped into three aspects listed above. Finally, in Section 5 we discuss the information used for charging in all examined charging systems, and identify open issues and potential improvements in information utilization.

2. CHARGING-RELATED TERMINOLOGY AND PROCESS INTERACTIONS

Charging related terminology found in the literature and standards is unfortunately not harmonized. For the benefit of the reader and to facilitate understanding of terms to newcomers in the field, we state the terms and definitions adopted for use in this paper and note the differences between
terms as used in other sources where applicable. (An experienced reader may skip this section and move on to Section 3.) There are four terms that refer to the most important charging-related processes (Fig. 1). *Metering* refers to a process that collects information about resource usage at a particular network element. *Accounting* uses metering logs to aggregate information about resource usage from different network elements. *Charging* refers to a process of calculating a cost, expressed in units acceptable for network management processes, of a given service consumption, by using the accounting information. Within the *billing* process, the charges are collected, and the service payment procedures are managed towards the party that consumed the service(s).

Depending on the payment type agreed, *prepaid* (also known as *prepay*) and *postpaid* (also known as *postpay*) billing method may be used. These methods are defined as charging options by Kurtansky and Stiller [4]. In *prepaid billing*, a certain amount of money must be deposited in advance to an account, which will then be spent in accordance with a service usage. In *postpaid billing*, an account is debited as the services are being used, but the payment is performed only after a certain time interval has expired (e.g., a month) by issuing a bill that aggregates costs of all services that have been used in the given interval.

![Figure 1. Charging and charging-related processes](image)

The literature differentiates between *charging models* and *charging mechanisms*. A *charging model* defines a list of criteria that will be applied in order to calculate a monetary cost of a used service, as well as a (list of) price(s), also known as *tariffs*, of a defined service unit that will be used in cost calculation, by applying the given criteria. *Volume-based* charging model uses the amount of data transferred as a criterion, for example, defines a price of 1$ for every 10 kB of transferred data. Other commonly used charging models are, the *time-based* model, which uses the time spent for service usage as a criterion, and the *content-based* model, which uses the service content of the provided service as a criterion [9]. A survey of the state of the art charging models (also known as *pricing schemes*) is given by Gizelis and Vergados [10].

A *charging mechanism*, which may be marked as offline or online [2], identifies whether the service is charged while it is being provisioned or thereafter.

*Offline charging mechanism* separates charging and service provisioning in time: charging is requested for the particular service as the service is started, but only accounting and metering processes are initiated. After the service is terminated, charging processes the accounting data, calculates the final service cost, and forwards it to the billing domain.

*Online charging mechanism* is performed in real-time in accordance with service provisioning, requiring accounting and metering to be performed in real-time as well. The main advantage of this approach is the ability to control the service cost at each point of the service session. Additionally, this enables introduction of service authorization mechanisms, i.e., granting or denying particular service components. Finally, online charging process can make decisions regarding service termination if certain conditions are met. From now on, if not otherwise noted, the term *charging* refers to online charging.
2.1. Differences in terminology in standards

There exist differences in how metering, accounting, charging, and billing are defined by the two relevant standardization bodies: the IETF and the 3GPP. For what is in 3GPP and in this paper termed charging, IETF uses the term rating [11]. IETF RFC 2975 [12] defines other processes in the same way as in this paper. 3GPP defines billing mainly as explained here, but refers to accounting as the process of “apportioning charges between the parties involved in service delivery”, TS 21.905 [13]. Furthermore, charging is in 3GPP used in a broader sense, as an umbrella term for all involved sub-processes [8]. The difference in terminology is explained in more detail in [4, 5, 6].

2.2. Roles and stakeholders

There are two key roles that are played by entities who participate in charging and in charging-related processes: 1) a user, and 2) a provider, as shown in Fig. 2. A user is an entity that uses/receives the given service and is being charged for using it. A provider is an entity that provides the given service. In a general case, a provider role may be played by different stakeholders, e.g., a Mobile Network Operator, an Internet Service Provider, or a third-party service provider. From now on, the term service provider will be used as a general reference to any of these stakeholders.

A user role may be played either by a natural person that receives a service (in that case, it is usually an end user), or by a service provider. This enables any service provider to provide a service, but also to be a user of other service(s) provided by any other service provider(s). Primary service provider (PSP) is a service provider that has a business relationship established with an end user, usually by means of a Service Level Agreement (SLA) [14], and it is responsible for overall provision and control of services towards him (Fig. 3). In 3GPP, a network under the jurisdiction of the primary service provider is also known as a Home Environment [13]. From the point of view of the end user, his PSP is the only stakeholder that is responsible for charging [15]. In a case when a service is provisioned by using resources or services of additional sub-providers, inter-domain charging procedures are needed. Different aspects of the charging responsibility reassignment
between the service providers are discussed in [16, 17, 18, 19], addressing the problems of charging data correlation and charging process coordination.

Recent standardization and research efforts have been turning towards policy-based charging systems [20, 21, 22]. As defined by Westerinen et al. [23], a policy is “1) A definite goal, course or method of action to guide and determine present and future decisions; 2) a set of rules to administer, manage, and control access to network resources”. By using policy-based systems, it is possible to exchange messages containing aggregated management information (including the charging information), and thus to reduce signaling.

2.3. Information used in charging

Depending on the information the charging data contain, we aggregate the information into three groups: User profile, Charging model, and Accounting record, as shown in Fig. 4.

Figure 4. Input and output data used in charging

*User profile* is a collection of information related to an end user, maintained by the PSP. In this paper we are focused only on a fraction of the user profile, the one that contains the information used in charging. Such parameters include, for example, the amount of credit that is available for spending, a list of services the end user is allowed to use, potential credit limits for certain services, and references to the used charging models. The user profile is updated during and/or after the charging process, by, e.g., modifying counters or deducting the credit spent during the charging process from the overall available credit.

*Charging model* is a collection of information containing 1) rules that determine how to calculate a service cost, and 2) tariffs used.

*Accounting record* (generated by the accounting process) is a collection of usage data records (generated by the metering process) aggregated from different network elements. Accounting records are mostly used in offline charging for post-processing purposes, and are in that case called Charging Data Records. In a broader sense, they are also used by online charging as input parameters for real time service cost calculation.

2.4. Interaction between processes, stakeholders, and charging information

Fig. 5 aggregates and puts into a mutual relation the overall terminology. For simplicity, it is assumed that no sub-providers are included in service provisioning.

Since the billing, charging, accounting, and metering processes are carried out within the PSP’s home environment, the PSP has full control over them. The end user, situated in the home environment, has an SLA established with his PSP, thus agreeing the terms of service usage,
including the content of the user profile, as well as the charging mechanism, charging model, and the billing method used.

Once the service is initiated, the charging process retrieves the user profile and the agreed charging model, uses both of them as input charging data, and starts charging. Accordingly, accounting and metering are used for generating the accounting records and the usage data records, respectively. Accounting record is then used to calculate the cost of the service at a certain point of the service session, which will enable charging process to allow or deny further service usage. During or after the charging process, the user profile is updated accordingly.

Figure 5. Charging in a perspective of related processes, stakeholders, and data

3. ONLINE CHARGING IN 3GPP STANDARDS

Online charging is covered by a number of 3GPP documents. The 3GPP Technical Specification (TS) 32.240 [25] serves as an umbrella document for charging in 3GPP (both offline and online), and contains a listed structure of other 3GPP charging specifications. The specifications are categorized into two sets: the first set covers specifications for different charging-related 3GPP architectures, and the second set covers charging parameters, syntax descriptions, and interactions within the network.

The specifications are grouped in stable releases. This paper is mostly based on 3GPP Release 10, but also explains selected features covered by Release 11. Release 11 is scheduled to be completed in September 2012.

End user’s home network operator (i.e., the PSP) is defined by the 3GPP as a central entity that is responsible for charging the end user. If the PSP has roaming agreements with other network operators, the end user is able to use their network infrastructure while roaming. In this case, the PSP’s charging system is also responsible for charging the end user, but additional inter-domain charging procedures must take place, too. Additionally, the end user may be able to consume services provided by third-party content and application providers.
General charging requirements and principles [31] for the 3GPP architecture are listed in TS 22.115 [32]. Important requirements for end users are here summarized as follows:

- End users must be aware of all charges that are related to them. This includes the ability of informing end users of the charges that are about to happen, and enabling end users to accept or reject the service, regarding the calculated amount of charge.
- There must exist an ability to charge separately each media component within a single session, and to perform charging according to the network resources used.
- There must exist an ability to charge end users depending on their location, presence, i.e., the context the service is being consumed in.
- An end user must be able to use the same charging model when roaming, the same way as if he were in the home environment.

A high-level view of 3GPP online charging is given in Fig 6. Online charging procedures may be used at three different architectural levels: a bearer level, a subsystem level, and a service level. Different charging functionalities are provided at different levels. At the bearer level, charging is able to control network resources used for service delivery (in both circuit switched and packet switched domain). At the subsystem level, that is, the level where IMS is situated [24], charging controls service sessions and is able to allow/deny session initiation as a whole. At the service level, specific services may be charged (e.g., a Multimedia Service or MMS) as well as particular service content (e.g., movies or music).

![Figure 6. Operating levels of online charging in 3GPP](image)

An Online Charging System (OCS), defined in TS 32.296 [26] is a functional architecture that provides support for all three levels of online charging in 3GPP. The OCS also serves the Policy and Charging Control (PCC) architecture [21], a framework that logically connects processes at the subsystem level with the processes at the bearer level. The OCS and the charging-related functionalities of the PCC architecture are elaborated in more detail later in the paper.

### 3.1. Background in IETF standards

The generic AAA architecture given in RFC 2903 [27] specifies a framework that incorporates user authentication, service authorization, and accounting procedures, aiming to be used in an Internet environment. Published in 2000, the architecture is still used (partially or entirely) as a basis in building many state of the art charging architectures, including the 3GPP OCS. Telecom-specific requirements, important for the 3GPP OCS, adopt two key ideas from the AAA architecture: (1)
accounting (and consequently charging) is carried out at a central point in the network, and (2) accounting (and consequently charging) is incorporated within the service authorization procedures.

A more detailed view on accounting procedures in IETF can be found in RFC 3334 [20], where policy-based accounting is specified for the generic AAA architecture. The document defines a three-layer reference model incorporating the layers of metering, accounting, and charging policies. However, only accounting policies are considered in more detail, and are in this case used to instruct a remote network entity on how to collect accounting data. The concept of policies is accepted in 3GPP for network resource management (carried out by the PCC architecture) and for service authorization procedures (carried out by the OCS).

The main signaling protocol used in 3GPP OCS for transmission of AAA information is the Diameter protocol, specified in RFC 3688 [28], a successor of the RADIUS protocol, specified in RFC 2865 [29]. Diameter Base Protocol [28] contains basic Diameter messages for generic AAA functionality as well as message content specification, stored in data structures called Attribute Value Pairs (AVP). All other AAA procedures, which are specific for a particular service or a network function that uses Diameter, can be defined by Diameter applications. Such applications are, the Diameter Credit-Control Application, RFC 4006 [11], used by the 3GPP OCS for cost control during the service session, and 3GPP-specific interface applications used for signaling purposes between 3GPP functions.

Credit control is a mechanism used for real time interaction between an OCS and a service provider, to control and/or monitor all charges related to the service usage. The mechanism can be used for both event-based and session-based services in online charging scenarios. Fig. 7 depicts an example credit control scenario between a service provider and a charging system for a single service session. Initial Diameter credit control message Start is sent by the service provider at the same time as the service session is started. The message is sent to the OCS, containing, among other
parameters, the number of requested service units $r_1$, the service identifier $s_1$, and the user identifier $u_1$. OCS then grants a certain number of service units $g_1$, by sending the Diameter authentication message $Auth$. After receiving the $Auth$ message, service provider can deliver the granted number of service units to the user. Once the granted units are spent, Diameter $Interim$ message is sent, requesting new units $r_2$ and containing the number of used units $b_1$, where $b_1 \leq g_1$. A new number of service units $g_2$ is granted, allowing the service provider to continue service provisioning. At the end of the service session, a final Diameter message $Stop$ is sent containing the number of used units $b_2$, where $b_2 \leq g_2$, and no further service units are requested.

From the point of view of an end user, credit control mechanism takes place between his/her PSP, i.e., the function in the PSP’s home environment that provides the given service (usually a gateway node providing a network access service), and the PSP’s OCS. Considering the charging-related information stored in the end user’s User Profile (particularly the available credit information), the OCS grants or denies service usage by using the credit-control mechanism. Future Diameter extensions are being developed within the IETF Diameter Maintenance and Extensions (DIME) work group. The overview of their work is available on the DIME WG website [30].

3.2. 3GPP Online Charging System

The OCS [26], illustrated in Fig. 8, stands as a central function for online charging within the 3GPP. It consists of four functional elements: an Online Charging Function (OCF), an Account Balance Management Function (ABMF), a Rating Function (RF), and a Charging Gateway Function (CGF). By using the standardized interfaces, it is connected with other functions/nodes across the three levels of charging, e.g., the Mobile Switching Center (MSC), the Serving GPRS Support Node (SGSN), IMS Application Servers (AS), and PCC functions. Next, the OCS is also connected with the Recharging Server. A more detailed description of OCS functions and the interaction between the OCS and the PCC functions is given next. (A comprehensive tutorial about online charging at the subsystem level is given by Kuhrie et al. [2].)

The ABMF contains information about end user’s credit as well as counters associated with the end user. A counter is an aggregation of units of service usage (or monetary units), which may be in relation to the end user contractual terms with his PSP (e.g., a number of minutes for voice calls that are free of charge per month). As a service is being used, the value of counters associated with the service is updated accordingly. By using counters, it is possible to establish an end-user specific loyalty program such as service price discounts or bonus programs. The ABMF is additionally connected with the Recharging Server, used to buy more credits if available credits are (nearly) depleted.

The RF contains prices of all available services. It is able to provide the OCF either with information about a price of a certain service unit considering the charging model used, or with information about price of a given service session, considering the price of the service unit, charging model used, and the number of service units consumed. Additionally, the RF maintains the end user context information, which is in this architecture defined as a list of currently active services per end user. By using the context information, the RF is able to perform a correlation process, i.e., a process in which service prices applied to the end user may be modified depending on other active services in progress.
The OCF is a central OCS function that interacts with other functions in the network and performs online charging. It is possible to perform charging in an event-based (i.e., only once), or in a session based manner (i.e., continuously during the session), depending on the service used. The OCF is connected with the ABMF and the RF by using Diameter-based Rc and Re reference points, respectively. These interfaces are used for retrieval and/or update of end user’s available credit, service prices, possible counters, etc.

Figure 8. 3GPP Online Charging System and selected functions that use it (Release 11). Sy interface is not included in Release 10.

Bearer level procedures are responsible for establishment, modification, and termination of IP flows used for service delivery, TS 23.125 [33]. At this level, it is possible not only to allow/deny service usage by using the OCS functionalities, but also to reserve network resources, in order to achieve a certain level of QoS for each IP flow within the service session, according to end user’s subscription profile.

To support these mechanisms, 3GPP defines the PCC architecture, TS 23.203 [21], which provides a framework for mapping session-related data in the signaling layer to the network-related data in the connectivity layer, i.e., QoS data and charging data. Session-related data are received from an Application Function, situated at the subsystem level. Each media flow is given a policy for QoS assurance and charging, called a PCC rule. Policy decision making and PCC rules creation is performed at the Policy and Charging Rules Function (PCRF). When creating PCC rules, a decision is made on how a certain media flow is treated in the packet switched network, depending on the user-specific information (e.g. allowed QoS), retrieved from the Subscription Profile Repository (SPR). Policy and Charging Enforcement Function (PCEF) enforces the policies by creating, modifying or deleting IP flows (used for transmission of media flows the given service consists of), and assuring the required QoS for each flow. Starting with Release 8, additional functions are introduced in the PCC architecture to support mobility and roaming, well explained by Balbas et al. [34]. A need for such support was previously discussed by Kueh and Wilson [35].
Table I. List of selected charging information (3GPP TS 22.115)

<table>
<thead>
<tr>
<th>Charging information provided by an end user</th>
<th>End user identity</th>
<th>Home environment identity</th>
<th>Terminal Identity</th>
<th>Resource requested</th>
<th>QoS parameters</th>
<th>Identifier for service requested</th>
</tr>
</thead>
<tbody>
<tr>
<td>Charging information provided by a home environment or a roaming network</td>
<td>Serving network identity</td>
<td>Universal Time for specific events during the charging session</td>
<td>Quantity of data transferred both to and from the user</td>
<td>Resource allocated to the user</td>
<td>QoS provided to the user</td>
<td>Location of an end user</td>
</tr>
<tr>
<td>Charging information provided by a third party service provider</td>
<td>Third party identity</td>
<td>Universal Time for specific events during the charging session</td>
<td>Type of service</td>
<td>Type of content</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

PCC architecture is connected with the OCS by using the Gy reference point (between the PCEF and the OCF), and, as of Release 11, Sy reference point (between the PCRF and the OCS).

Gy reference point implements the IETF’s Credit-Control application [11] to allow online charging of each IP flow used during the service session, as explained in Section 3.1.

Sy reference point is used to utilize the counters that are maintained at the OCS, whose status may influence the policy decision making process at the PCRF. By using Sy, PCRF is able to access the information about the end user’s spending limits for a certain service, stored at the OCS. If a limit threshold is reached, the PCRF is able to, e.g., adjust QoS of the service. Note that a need for such an interface was also previously expressed by Grgic et al. [36].

A possible evolution of the PCC architecture in an aspect of fixed - mobile convergence is further discussed by Ouellette et al. [37].

3.3. Charging information used in 3GPP OCS

Information used in 3GPP online charging may be grouped in two groups: (1) information provided by network functions to the OCS when requesting online charging, including service level, subsystem level, and bearer level functions; and (2) information already stored at the OCS and used in online charging.

The first group of information is shown in Table I. Instead of listing what information is provided by which network function, each row in the table aggregates the given information, depending on the stakeholder the information is related to. In a general case, the following stakeholders may be included in service delivery (and accordingly in charging): an end user, an end user’s home environment (i.e., the PSP), a visited network (if an end user is in roaming), and a third party service provider.
provider (if an end user has access to its service(s)). The data mostly contain information regarding identities (for identifying, e.g., a network, a user terminal, and a requested service), requested service units (measured as network resources or a content provided), requested QoS, location, and presence information.

The second group of information is stored at functions that belong to the OCS, and has mostly been mentioned in the previous section. The most important information includes:

- End user’s available credit;
- Counters related to end user’s service usage history, allowed number of service units, etc.;
- A list of active services per end user; and
- A list of service prices.

The first three items represent information that is related to the end user: available credit is a part of the User Profile; certain counters may have values depending on the end user’s subscription; and the list of active services represents a part of the context in which the end user consumes the given service. However, information about the services that are charged is reduced to the information about the service prices. Additionally, the OCS itself does not provide any specifications about the inter-domain charging procedures, which come out of the information stored in SLA agreements established between providers. Finally, due to numerous network functions that are able to connect with the OCS by using standardized reference points, to provide the OCS with the necessary charging information and to request charging (particularly when a per-flow charging is required by the PCC architecture), there emerges a problem of the amount of signaling generated by charging.

Therefore, in the next section we elaborate the state of the art research related to the 3GPP online charging open issues, grouped into aspects of charging signaling, inter-domain charging, and the service aspect of charging.

4. STATE OF THE ART IN ONLINE CHARGING

This section describes and compares the state of the art research work in online charging. Continuing from the 3GPP three levels of charging (Fig. 6), we categorize the relevant research work into three aspects (Fig. 9): 1) Signaling aspect, 2) Inter-domain aspect, and 3) Service- and component-based aspect. Each aspect is presented as an aggregation of work related to one of the 3GPP charging levels. First, the signaling aspect aggregates research work that may be used as an extension to existing bearer level signaling procedures. Next, the inter-domain aspect broadens the view of the service session charging at the subsystem level in respect to inter-domain problems and known limitations. Finally, the service- and component based aspect considers charging of composed services, as a (possible) extension of service level 3GPP charging procedures. In each aspect we define different set of comparison parameters and use them to compare the research work. Note that charging functionalities elaborated in this section, especially their functions and interfaces, do not necessarily match the 3GPP OCS architecture nor are designed for it. However, key ideas used in those systems are discussed in light of their possible use in 3GPP.
4.1. Charging signaling aspect

Charging signaling aspect addresses research work related to charging signaling procedures, which are used to authorize network resources and hereby allow or deny service usage. An example of such procedure is the Credit-Control application [11].

The work is compared based on the following parameters:

- **Amount of signaling exchanged** – Depending if research is focused on adding new features to signaling interfaces or optimizing the existing features, amount of necessary signaling may be increased or decreased respectively. The parameter is normalized to the standard amount of signaling;

- **Premature service session termination probability** – This parameter compares existing procedures with respect to their ability to react on time to charging events that may result in terminating the service session, e.g., if the end user’s credit is depleted and no additional service provisioning is allowed, and to prevent such termination if possible. The parameter is normalized given the standard session termination probability;

- **Multi-service support** – This parameter assesses how a certain signaling procedure enhances (if at all) charging of multiple services used by a single end user in parallel. Parameter values are low, medium, and high;

- **3GPP compatibility** – This parameter assesses to which extent the proposed signaling procedure is compatible with the existing 3GPP charging architecture. Parameter values are low, medium, and high.

The problem of charging end user’s multiple parallel services is addressed by Kurtansky and Stiller [41], proposing a concept of service bundles, defined as a list of services that may be used by an end user. Instead of performing credit control procedure separately for every media flow in each service session that is in progress, it is done per bundle by using a Time Interval Calculation Algorithm (TICA). Depending on the input parameters, TICA calculates a time interval wherein an end user can consume any combination of services in the bundle. When the time interval elapses, another interval is calculated. In [42] algorithm improvements are proposed, including statistical prediction of resource consumption, learning from the past methods, and service classification. The evaluation of the algorithm in [43] showed the reduction of credit check messages, compared to traditional credit check procedures, with minimal charging fraud detected. As a next step of this work, Oumina and Ranc [44] deal with incorporating service bundles into the 3GPP OCS. They discuss when to initiate time interval adaptation, e.g., in case of a service bundle change, or in case of a change in a price of a service used within the bundle. Finally, in [45] they propose a modification...
of an OCS’s Rating Function in order to support applications using multiple (third party) services. The idea is to be able to calculate a service price depending on a combination of other services used throughout the application. To do so, they propose a service identifier (used when requesting the Rating Function to determine service price) to consist of two parts: a fixed part (which defines in an unique way the service in the OCS), and a variable part (which refers to the application that uses the service for which charging has been requested).

Lin et al. [39] deal with multiple parallel services consumed by an end user, in a situation when additional new service is initiated. A Prepaid Credit Reclaim (PCR) mechanism is proposed, enabling the credit re-distribution between the services that are already in progress and the ones that are about to be established. This approach also generates an additional amount of signaling, but enables all services to be equally allocated with the credit, instead to (possibly) reject new services due to insufficient credit left, while at the same time already established services are allocated with (potentially) too much credit. This approach is opposite from the one later proposed by Yang et al. [40], where the probability of terminating ongoing service sessions is reduced by introducing a mechanism that rejects new session requests, when the end user’s credit is below a certain threshold.

Sou et al. [38] propose the Recharge Threshold-based Credit Reservation (RTCR) mechanism, enabling the charging system to recharge end user’s account by sending a message to the credit recharge server. An analytical model calculates the optimal credit threshold, which, when reached, triggers the charging system to request credit recharge. If the threshold is set too small, credit units may be depleted before the credit has been recharged, resulting in forced service termination. If the threshold is too large, the end user will receive recharge messages too frequently, resulting in heavy network traffic.

In [46], the same authors deal with charging signaling optimization for mid-session charging events, e.g., when a change of QoS occurs. Differing from the traditional systems where previously allocated but unused credits will be freed and credits for the modified session will be requested, the authors propose a Threshold-based Scheme for Credit Re-authorization (TSCR) procedure, enabling the charging system to omit the re-authorization procedure if the remaining credit units are sufficient to accommodate the new reservation and thereby reduce amount of signaling.

Fig. 10 shows comparison results of the explained charging signaling enhancements given the identified criteria. RTCR and PCR introduce charging functionality that results in a reduced
probability of a premature session termination. However, the amount of signaling they produce is increased compared to the standard signaling. A particular load on the network is being put by the PCR mechanism, due to the credit reclaiming procedure for every parallel session in progress. Both TICA and TSCR reduce the signaling, but due to the increased inaccuracy of credit information [46], TSCR poses a potential threat to session termination, increasing the termination probability.

We assess the 3GPP compatibility for TSCR and RTCR as high, since the proposed functionalities are able to use existing 3GPP infrastructure without any changes in signaling interfaces. TICA’s and PCR’s compatibility are assessed as medium, due to the proposed modifications that are needed in the network. When comparing the multi-service support, TSCR and RTCR deal only with the single service scenario, without any signaling optimization proposals when multiple services are initiated, and are consequently assessed as medium. On the other hand, PCR and TICA base their solutions on the signaling optimization techniques in case of multiple services are concurrently charged, and are marked as high.

The following work deals with functional improvements of the 3GPP OCS functions, which are not necessarily related to service authorization, therefore not included in the above comparison. For integrity, they are described next.

Grgic et al. [47] extend standard Rx and Gx interfaces within the PCC architecture with additional Diameter messages, in order to support advanced service adaptation mechanisms for complex multimedia services. The adaptation mechanisms, introduced by Skorin-Kapov et al. [48], specify and utilize a data structure called Media Degradation Path (MDP), which enables agreement of several service configurations for a single service session, any of which may be used during the service session, enabling the service, by selecting the adequate configuration, to adapt to different network conditions. In [36], modifications of the standard OCS are proposed and implemented in a laboratory prototype, to support the model presented in [47]. Finally in [49], the knowledge about end user’s service subscriptions (and the agreed service prices) is included in a process of finding optimal service configuration, by using the proposed interface situated between the PCRF and the OCS (as of 3GPP Release 11, also known as the Sy reference point).

Albaladejo et al. [50] notice that the 3GPP specifications (deliberately) omit the mapping procedure between the session data and the PCC rules at the PCRF, and propose two possible strategies of filling the PCC rules with the session information, followed by performance measurements and comparison of the two approaches. Shengyao [51] presents a design of the IMS Gateway Function, enabling the call session control functions to initiate/modify/terminate session-based online charging with the OCS.

4.2. Inter-domain aspect

This section studies work related to inter-domain charging, explaining the issues that emerge from limitations posed by business agreements among stakeholders. The research works are mutually independent and presented in a chronological order. The following parameters are used for comparison of works:

- Ability to dynamically change business agreements – assesses possibilities to create and/or modify business agreements among stakeholders depending on the current situation and interests, including change of charging model, service tariff, etc.;
• Privacy assurance mechanisms between stakeholders – analyzes available mechanisms in each research approach that assure end user's privacy while disseminating his data in the network;
• 3GPP compatibility – assesses to which extent a certain work is compatible with the 3GPP subsystem-level and inter-domain charging procedures. Assessment values are the same as in section 4.1.

Koutsopulou et al. [17, 18, 19], propose an architecture in which Charging, Accounting, and Billing (CAB) functions are deployed within a network of a third trusted party. By using a CAB gateway, charging data are collected from the entities that participate in service delivery, but belong to different administrative domains. By using open application programming interfaces, service providers can define their accounting and charging policies. This approach is not compatible with the 3GPP network operator-centric model, where all charging-related functions (from the point of view of an end user) are performed by the PSP. (Other models, e.g., content provider-centric model and content aggregator-centric model, may be found in the literature [16].)

Bormann et al. [57, 58] propose context-aware charging and billing mechanism, as a part of the project entitled Local Mobile Services (LOMS). They developed a system that enables providers of local mobile services to offer their services to end users by using a larger network operator’s infrastructure. These services are often context-aware (e.g., aware of a user’s location), thus enabling service providers to offer and utilize charging models that take the given context in consideration.

Within the Ambient Networks Project [52], Huitema et al. [53] created the architecture that supports negotiation mechanisms in accounting, charging, and billing, called Compensation architecture. The key idea is to allow for automated and dynamic negotiation about relationships between parties known or unknown to each other, and to realize the negotiated agreement in near real-time. Without a need to establish business agreements, the architecture enables negotiation of, e.g., tariffs and time of payment between the parties. A charging system to be used in this environment is defined, too.

In a novel approach proposed by Murata et al. [55], an open network is described as one in which network resources can be deployed not only by existing operators, but also by other interested parties: companies, universities, etc. The idea is to create an open-type mobile business environment as similarly as it has already been achieved in the Internet. In their work they propose an Open Heterogeneous Mobile Network (OHMN) consisting of five functional layers. The layers are also mapped to the existing layers standardized by the TISPAN NGN. Although it is believed the OHMN will open the mobile market, since the proposed architecture is still in an experimental phase, charging models to be used in such network are yet to be determined.

Simultaneously with the growing number of stakeholders and new business relationships, the market is flooded with numerous charging models and available tariffs, and offered to end users. Cheboldaeff [54] explains known approaches in organizing the charging models, such as service buckets and discounts usage, and highlights charging related issues when using these approaches.

In the latest research, Tran and Tuffin [56] list the key requirements for pricing schemes (i.e., charging models) when used in inter-domain charging, and discuss the important characteristics an ideal inter-domain pricing scheme must have.

When comparing the elaborated approaches given their ability to support dynamic change of business agreements, only the OHMN and the Compensation architecture (which is designed for that purpose) would support such scenarios. The motivation for introducing a five-layer environment
in OHMN is to ease introduction and promote new business models for the benefit of users. Other approaches assume the existence of a fixed business model (e.g., a third party charging provider in CAB, small provider – large provider relationship in LOMS, or offering groups of services by using predefined buckets in [54]). Next, if considering the aspect of privacy assurance, CAB enables exchanging CDRs between the network operator and the CAB gateway without specifying any mechanism that controls end user’s (potentially) private information contained within the CDR. In LOMS, it is assumed that users freely offer information about their context (e.g., preferences or location). However, Compensation allows end users to negotiate about the information they provide to the service provider. Finally, no architecture provides full compatibility with the 3GPP specifications. CAB enables connecting with the IMS via Open Service Architecture, but does not provide any signaling scenarios. LOMS uses Parlay X API for interconnecting with other networks. Finally, Compensation does not consider mapping to 3GPP at the current stage of development.

4.3. Service- and component-based aspect

This section addresses service level charging aspects of complex multimedia services. Such services, called composed services [60], usually consist of several elementary services, i.e., service components, often provided by different service providers. Charging of such services requires coordinated charging procedures for each service component included, and therefore creates problems of service price determination, charging data coordination/correlation, etc. According to [60], service component types are defined as follows:

- Media components — audio, video, data, etc.;
- Value-added service components, such as, a call forwarding service is an upgrade to standard call service and can be treated as a service component;
- Business-model-based components — includes different service providers in service delivery, each providing a different aspect of a service (e.g., access service or network service);
- Network-based components — in the case of roaming users, a service is divided into network components depending on the borders between the operators’ administrative domains.
- Content-based components, such as, music songs or traffic information messages.

In order to compare existing research work in this field, we define the following comparison criteria:

- Charging support for different component types – analyzes which component type(s) is/are supported by each work;
- Tariff adaptation ability – analyzes the ability of the charging system to adapt the tariff to the combination of included components; and
- 3GPP compatibility – analyzes whether the given work is compatible with service-level charging mechanisms in 3GPP.

One of the first research works in the field has been written by Van Le et al. [65], who propose mechanisms of outsourcing the accounting and charging processes, that is, having a specialized service provider in the network responsible only for these functions. In such environment, the proposed accounting and charging architecture supports service session composition across multiple domains.
Next, service tariff determination in QoS-enabled networks (e.g., a Differentiated Services-based network) is addressed by Wang and Schulzrinne [70]. They propose a model for dynamic definition of tariffs, which includes different parameters in tariff calculation, such as, a service priority, a service usage rate, or a current network load. The idea is to enable service providers to perform service adaptation to current network conditions by raising prices while congestion is in place.

Jennings and Malone [61] specify a two-phase rating process (i.e., a service price determination process) for composed services. In addition to summarizing prices of each included component (phase 1), the approach takes into consideration a context in which each component is used. That is, a tariff of each component is modified depending on the combination of other components that are used within the service (phase 2). Next, they develop an Accounting Logic Generator that automates the generation, deployment and application of charging schemes for composed IMS services [62].

Huang et al. [68, 69] address the problem of specifying the scheduling and load balancing algorithms for processing online charging requests depending on the service type, service priority, etc. The algorithm application results in response times to charging requests for different services.

Van Le et al. [66, 67] expand the standard OCS functionality to support charging for composed services. The included support consists of three elements: (1) the service composition information model used by the OCS; (2) functional components of the OCS to support composed services; and (3) the interaction between the functional components.

Faria and Nogueira [63, 64] take into consideration the context information present in collaborative service agreements between the involved service providers when charging composite services. This information enables, e.g., allowing certain discounts in tariffs or promotive periods. A mechanism to support inter-domain contextual information dissemination towards charging systems is also proposed.

When comparing the presented works regarding charging support for different component types, only OCS extension work [66, 67] supports media and content-based components. Works [65, 63, 64] include support for business-model-based components, while the two-phase rating [61] supports a combination of value-added and network-based components.

Next, in the DiffServ approach [70] tariff adaptation is performed based on network congestion status, while the outsourcing mechanism [65] uses charging records, but the tariff definition remains unclear. Only the two-phase rating [61] and the context utilization approach [63, 64] achieve a high level of tariff adaptation abilities based on the components included.

Currently, there is no support in standards for charging of composed services at the 3GPP OCS. Other than [66, 67], which propose an approach that enables the support, only the two-phase rating process could be compatible (although certain adaptations on the Rating Function are needed) with the 3GPP. Other approaches do not consider 3GPP-specific procedures.

5. CHARGING-RELATED INFORMATION: DISCUSSION ON OPEN ISSUES IN INFORMATION UTILIZATION

This section analyzes the charging-related information used in the presented research work, in addition to the information used in 3GPP standards and presented in Section 3.3. The open issues
in information utilization are identified and discussed, given the requirements of the future charging systems.

5.1. Charging information used in the signaling aspect

In the signaling aspect, research focus is oriented towards end user’s credit management, e.g., the ability to know when to recharge the user’s credit to avoid premature service termination, to optimize credit (re)distribution among different parallel services, or to reject initiation of a new service considering the available credit. Table II lists the information used in each approach. RTCR sets and modifies the level of the credit depletion threshold, enabling a dynamic mechanism of credit recharge. PCR and TSCR re-distribute the allocated credit for each service, depending on the occurrence of new service(s). The difference is that PCR uses information about credit of all end user’s active services, instead of TSCR’s single service. TICA maps available credit to a minimal time interval, within which all services in the bundle may be used without performing credit re-check. Online charging that uses MDP considers potential user’s service subscriptions while tariffing the service and deducing the available credit.

Table II. Comparison of the approaches using end user credit information in the signaling aspect of charging

<table>
<thead>
<tr>
<th>Approach</th>
<th>Key information used</th>
</tr>
</thead>
<tbody>
<tr>
<td>RTCR</td>
<td>Credit depletion threshold</td>
</tr>
<tr>
<td>PCR</td>
<td>Allocated credit per service, for all services used</td>
</tr>
<tr>
<td>TICA</td>
<td>Minimal time interval</td>
</tr>
<tr>
<td>TSCR</td>
<td>Allocated credit per service</td>
</tr>
<tr>
<td>MDP</td>
<td>Potential user subscription(s) that may determine how to use credit</td>
</tr>
</tbody>
</table>

5.2. Charging information used in the inter-domain aspect

The key goal in the inter-domain aspect is to unify and simplify charging information that is exchanged between stakeholders. Information simplification would achieve having fewer compatibility issues, easier establishment of business agreements, and easier privacy management. Table III lists the information used in works described in section 4.2. CAB uses standard charging information, such as CDRs or tariff policies, but accesses them over the Application Programming Interface (API), thus having control over the information exchanged. LOMS uses user context for building context-aware services and defines service templates, which are then used by small service providers to build applications. For example, such a template may define tariff discount rules to be used in charging. Next, the most important information in the Compensation architecture is the result of negotiation between parties, which is then used as an input parameter to define all charging-related rules. Finally, the concept of service buckets uses rules that define which service belongs to which bucket (and thus how they should be charged).
Table III. Information used in key inter-domain charging approaches

<table>
<thead>
<tr>
<th>Approach</th>
<th>Key information used</th>
</tr>
</thead>
<tbody>
<tr>
<td>CAB</td>
<td>CDR, Tariff policies, CAB API</td>
</tr>
<tr>
<td>LOMS</td>
<td>Service templates, discount rules, context</td>
</tr>
<tr>
<td>Compensation</td>
<td>Service negotiation results</td>
</tr>
<tr>
<td>Service buckets</td>
<td>Service-to-bucket mapping rules</td>
</tr>
</tbody>
</table>

5.3. Charging information used in the service- and component-based aspect

Charging information used in this aspect mostly consists of a cross-referenced knowledge about service components and their prices in dependence of other components included in the service, or other contextual data. Table IV lists information used in the elaborated approaches. Tariff calculation model in Diffserv network uses three types of available charges (i.e., tariffs): holding charge (used when network resources are reserved for the service that does not currently generate any traffic), usage charge (used in normal network conditions), and congestion charge (used in case network is congested). Two-phase rating uses a charging scheme, i.e., a set of rules defining which tariff to apply in which phase and to which combination of components. The approach utilizing context stored in service agreements defines a new 3GPP function, called context broker, to collect and monitor the contextual data.

Table IV. Information used in key service- and component-based charging approaches

<table>
<thead>
<tr>
<th>Approach</th>
<th>Key information used</th>
</tr>
</thead>
<tbody>
<tr>
<td>Diffserv</td>
<td>Holding charge, usage charge, congestion charge</td>
</tr>
<tr>
<td>Two-phase rating</td>
<td>Charging scheme</td>
</tr>
<tr>
<td>OCS extensions</td>
<td>Standard 3GPP information</td>
</tr>
<tr>
<td>SLA context</td>
<td>Inter-domain context stored in context broker</td>
</tr>
</tbody>
</table>

5.4. Open issues in charging information utilization

Our analysis shows that the 3GPP OCS uses a limited set of information for online charging, the most important of which is: end user’s available credit, charging model and the corresponding service price, counters that can represent end users’ spending limits, and a list of active services per end user. The existing research work builds upon this information by introducing, e.g., different thresholds (credit, counter, etc.), knowledge stored in SLAs, or means to modify service prices according to other currently used services. However, along with the development of charging systems and more demanding requirements on charging, there emerges a need for accessing...
additional information that will become important, e.g., service usage statistics, device capabilities, or detailed information about active service configurations. We do not attempt to predict which of the information will be included in future charging configurations and for what purpose, but instead we summarize open issues that (will) occur when utilization of such information becomes inevitable. We identify the three most important open issues, as follows.

First, this information is already available in the network, but it is stored at different functional entities, and it is not conveniently structured for use in charging systems. Each of the existing research approaches explained in this paper uses only a segment of the overall available information, assuming it can be reached from the charging system. Future research works will have to tackle the problem of overall charging information specification and structuring.

Second, even if the information were structured, there is still unclear how to deliver the information to participants in service delivery and consequently in charging (PSP, third party providers, etc.) and how it would interfere with advanced service-related mechanisms, such as service adaptation. Therefore, the existing research work lacks the specification of signaling scenarios for charging information exchange.

Third, the user privacy protection rules (other than the generic ones provided in SLA agreements) while exchanging the information between stakeholders are currently not specified in an open and/or standardized way. It is likely that the future context-based services will require access to (some of) user’s personal information, and charging systems will have to find how to exchange such information (or a modified version of it) without compromising user’s privacy.

6. CONCLUSION

Our analysis shows that there is a trend of including additional information in the state of the art online charging systems. This information includes richer descriptions of users, services and network properties. Taking “user context” into account (e.g., list of active services or a credit depletion threshold) may become a crucial input for making advanced charging-related decisions. Knowledge about the service (e.g., service composition or its adaptation capabilities) is also important when calculating tariffs and/or performing service authorization. There is also a need for charging information simplification and unification in dynamically changing business environments.

The trend noted above creates several research challenges with respect to accessing and utilizing the relevant information. The most important issues, as identified in this paper, are the lack of information specification and structuring, information sharing issues, and user privacy issues. To make further advances in information utilization in future online charging systems, researchers will have to address them.

ACKNOWLEDGEMENT

This work was carried out within the research project 036-0362027-1639 "Content Delivery and Mobility of Users and Services in New Generation Networks“, supported by the Ministry of Science, Education and Sports of the Republic of Croatia.
REFERENCES


15. IP multimedia subsystem (IMS): Stage 2. 3GPP TS 23.228, Rev. 11.3.0, 2011.

16. Telecommunication management; charging management; charging architecture and principles. 3GPP TS 32.240, Rev. 11.2.0, 2011.

17. Telecommunication management; charging management; online charging system (OCS): Applications and interfaces. 3GPP TS 32.296, Rev. 11.2.0, 2011.


32. Service aspects; charging and billing. 3GPP TS 22.115, Rev. 11.5.0, 2011.
33. Overall high level functionality and architecture impacts of flow based charging; stage 2. 3GPP TS 23.125, Rev. 6.8.0, 2006.
51. C Shengyao PW. Design and implementation of an extensible online charging architecture for the open IMS playground. Proc. 3rd International Conference on Testbeds and Research Infrastructure for the Development of Networks and Communities, 2007; 1–5.


