



Enabling Delivery of Mobile Services Over Heterogeneous Converged Infrastructures

Nikos Houssos, Vangelis Gazis and Athanassia Alonistioti*

Communication Networks Laboratory, Department of Informatics & Telecommunications, University of Athens, 157 84, Athens, Greece

E-mail: nhoussos@di.uoa.gr

E-mail: gazis@di.uoa.gr

E-mail: nancy@di.uoa.gr

Abstract. Traditionally, end customers have been offered different categories of communication, data and media services (e.g., fixed/mobile voice, fixed/mobile data, broadcasting) through vertically separated, rigidly integrated infrastructures. Major advances in a variety of technological fields, mainly in the area of mobile computing and networking, have created prospects for a fully converged environment, where ubiquitous access to an abundance of value-added services will be offered over a single, as perceived by the users, highly reconfigurable system. This vision can be enabled by the seamless “plugging” of diverse access networks to a high-speed IP backbone; however the path to its realization poses a variety of additional challenges. The required support of complicated business models and service delivery over highly diverse contexts introduces significant complexity to service management and provision. The present contribution presents object-oriented mediating service management platforms as a catalyst for making these demanding tasks feasible, identifies their desired functionality and provides an overview of such a distributed framework that we have designed and prototyped. Furthermore, we demonstrate that the latter is able to be a critical enabler of flexible, adaptable service provision over future heterogeneous networks, while supporting advanced business paradigms.

Key Words. mobile service provision, business models, 3G, 4G, reconfigurability, adaptability, converged networks

1. Introduction

The information technology and telecommunications fields witnessed in the 1990s two major revolutions that took place in parallel on a global scale: the widespread adoption of the Internet and the phenomenal growth of 2nd generation (2G) cellular telephony. These developments created the trend for a convergence

between the traditionally not closely related worlds of Internet and telecommunications, whose early results like Voice over IP (VoIP) services over IP networks and mobile Internet access have already made some impact and shown clear merits. However, the potential of this convergence, as envisioned mainly by the mobile computing and communications research community can go much further: In the longer term, systems commonly referred to as 4th generation (4G) Jamalipour and Tekinay, 2001; Pereira, 1999; Gazis et al., 2002; Weiser, 1991), are envisioned, offering “always-on” connectivity and ubiquitous, customizable value-added services (VAS)¹ over highly diverse, dynamically reconfigurable communication infrastructures, including wireline and wireless wide-area telecommunication networks, Digital Video Broadcasting (DVB) systems and Home Area Networks, to name a few. The first milestone on this route is the introduction of 3rd generation (3G) (UMTS Forum, 2003a, 2003b, 2003c, 2003d) networks, characterised by a shift of focus from plain voice and short messaging services to value-added, content-oriented multimedia applications, provisioned over open networks and contributed by various market players.

Naturally, putting the above into practice is anything else than trivial. Business partnerships and multi-party co-operation are nowadays still hampered by the lack of efficient, largely automated mechanisms for inter-domain interactions (e.g., service deployment across administrative boundaries, billing

*To whom correspondence should be addressed.

and accounting). Indeed, the fact that applications are developed and delivered by many co-operating entities and need to be deployed on multiple types of networks significantly complicates service delivery and management. To make things worse, the evolution to more dynamic, horizontally integrated environments requires reconfigurability and adaptability of systems and services (Katz, 1994; Dillinger, Alonistioti and Madani, 2003), leading to increased complexity in their development. In the present contribution, we discuss how advanced multimedia service provision in a converged environment can be achieved through the evolution of current mobile networks and services. We argue that a catalyst for the resolution of these issues in a more advanced 4G context, can be found in the form of service provision platforms that mediate between service developers, infrastructure operators and end-users for facilitating and optimizing personalized, flexible service delivery. The administering entity of such a platform, termed the *service platform operator (provider)*, is introduced as a new role in the mobile services value chain that can form the basis for the application of novel business models in service provision over heterogeneous networks. Furthermore, we identify the functionality that such a platform should incorporate, we elaborate on the architecture and high-level functionality of a platform prototype that we have designed and implemented and illustrate with an example how these models can be applied with the support of the proposed platform for customized delivery of mobile services in diverse environments.

The rest of the present paper is structured as follows: First, we discuss certain issues regarding the evolution of service provision to 3G and beyond and in particular briefly describe the current status worldwide and indicate the potential of mediating service management platforms. Subsequently, we introduce the new *service platform operator* business role and elaborate on certain business models of mobile service provision, including identification of business actors and revenue flows. We then present the architecture and functionality of an object oriented service provision platform that supports the application of such business models and acts as a catalyst in the delivery of services, by intelligently mediating between involved parties. Next, a scenario demonstrating context-aware mobile service delivery through the platform is presented. Then, a detailed evaluation of the proposed approach and its applicability in real-world deployments is presented. The last sections

of this document are dedicated to conclusions and acknowledgements.

2. Mobile Value-Added Service Provision in 3G and Beyond

2.1. Mobile services—From plain voice telephony to 4G

Service provision in 2G mobile systems is performed according to the paradigms that have been applied in traditional fixed telephony networks (Lazar, 1997): service logic is vertically placed in telecommunication network nodes and tightly coupled with the underlying infrastructure; application development is performed by network operators or equipment vendors using specialized tools; access to services is exclusive to the subscribers of a particular operator and typically forms part of subscription-based offerings. Important negative effects of this approach include tedious and time-consuming (in the order of months or years) procedures for the introduction of a new service as well as increased complexity of the value-added applications development task, resulting in a limited range of available services (only a few dozens of services are available from a typical 2G mobile operator).

3G cellular networks are expected to enable the provision of a much wider repertoire of non-standardized multimedia value-added applications that will be developed through the contribution of many different business players including mobile operators, equipment vendors, application developers (commonly termed Value-Added Service Providers, VASPs), content providers, financial institutions (UMTS Forum, 2003a, 2003b, 2003c, 2003d). These objectives are to a large extent reflected in fundamental design principles of 3G networks that have been adopted by the corresponding standardization bodies and fora (3GPP, 2003; 3GPP2, 2003) and are listed below:

- Traditional circuit-switching is (to some extent) being replaced by IP-based transport, which enables higher degrees of network programmability and flexibility in service deployment and provision.
- Services are decoupled from the functionality of the underlying network, which is accessible through open APIs, like OSA/Parlay (URL, 2003; 3GPP TS 23.127); Service logic is not exclusively hosted inside the network; parts of it can be distributed to

enterprise servers outside the mobile operator's domain or downloaded to terminals (3GPP TS 23.057).

- The access and core network are to a significant extent decoupled, so that service provision over different access technologies is possible (3GPP TS 23.002).

The discussions for the definition of 4G networks and the relevant research activities are still in their early stages. However, there is a common vision of those systems forming an ambient, "invisible" computing environment (Weiser, 1991) where a vast space of services will impact almost every aspect of users' daily activities. The following basic characteristics of next generation mobile systems and services could be identified:

- **User centrality:** The fact that current communication systems are technology- instead of user-driven is widely recognized as one of their major limitations. Nowadays communication service consumers experience the simultaneous existence of various types of systems (e.g., wireline or wireless telecommunication networks, wired Internet, DVB), each providing a particular set of communication services. In an ideal fully converged environment, an end-user should be given the view of a single communication infrastructure supporting the entirety of services that are being offered and consequently be totally agnostic of the underlying technologies. For example, he/she does not need to be aware of whether the video clip that he/she is watching is downloaded over a cellular network, a wireless LAN or a DVB system. Moreover, user intervention should be minimal (to the extent possible) (Garlan et al., 2002). Intelligent software agents could undertake the task of discovering services that can be useful in a particular situation as well as of customizing them and rendering them to the user in the best possible way (e.g., in terms of the wireless communication technology, the user terminal or the human-machine interaction paradigm employed).
- **Pervasiveness:** Service delivery in the medium to long term will be performed over a multiplicity of devices that will be spread all over the physical environment and enable the creation of ubiquitous distributed services not imaginable today.
- **Co-existence of multiple access technologies.** 4G systems will be characterized by the seamless co-existence of competing and/or complementary connectivity technologies (Jamalipour and Tekinay, 2001; Mohr, 2000; Annoni et al., 2001) (see Table 1

for examples of such technologies) that will provide end-users with access to the global core IP backbone infrastructure. Capabilities for dynamic switches between different access technologies (i.e., vertical handovers), based on intelligent, context-aware decisions should be commonplace.

- **Adaptability/reconfigurability and context-awareness.** Adaptation has been widely recognized as a crucial factor for next generation mobile services (Katz, 1994; Dillinger, Alonistiotti and Madani, 2003), which will be delivered in a variety of circumstances that will not be a-priori predictable. This unpredictability can be appreciated by examining Table 1, where only one aspect of context is presented, namely access technologies and the resulting network parameters (e.g., achievable data rates). The constantly changing context will considerably impact service access and provision functionality across all functional layers and planes. The corresponding components should therefore be dynamically reconfigurable and adaptable to rapidly varying conditions.

2.2. Next generation mobile services—towards realization

A major issue regarding mobile services is the effective commercial exploitation of the huge research and standardization efforts related to 3G and 4G systems, to the benefit of end users as well as industry stakeholders. As of early June 2003, initial 3G systems were available (e.g., in Japan, South Korea, as well as five European countries). However, outstanding success has been only achieved so far by service provision models that are based on so-called 2.5G systems, which have introduced the use of packet-based networks (packet-switched PDC or GPRS), namely the i-mode system (NTT DoCoMo's I-mode, 2003) and its local competitors (EzWeb and J-Phone) in Japan. In those cases, mobile operators offer mobile users a wide range of third-party content/applications through a portal-like interface incorporated in the cellular phone menu. Similar offerings have subsequently become available worldwide.

Such paradigms base their success potential mainly on appropriate business models and marketing strategy, while not introducing significant technical innovations. They are designed to operate without sophisticated support for inter-domain interactions like automated service registration/deployment and advanced revenue sharing; for instance, services are offered to subscribers

Table 1. Main parameters of different access technologies

System	Data rate	Technology	Range	Mobility	Frequency range	Original application area
GSM including • GPRS • HSCSD • EDGE	9.6 Kb/s up to 384 Kb/s	TDMA FDD	Up to 35 Km GSM Lower for data	High	900 MHz 1800 MHz 1900 MHz	Public environment Private environment
IMT-2000 UMTS (UTRA)	Max. 2 Mb/s	IMT-2000 Family WCDMA (FDD) TD-CDMA (TDD)	30 m–20 Km	High	2 GHz	Public environment Private environment
DECT DECTLink	Max. 2 Mb/s	TDMA/TDD	Up to 50 m	Low	1880–1900 MHz	Office environment Residential environment
Bluetooth	Max. 721 kb/s		0.1–10 m	Very low	2.4 GHz ISM band	Cable replacement SoHo environment
HiperLAN/2	25 Mb/s	OFDM, TDD	50–300 m	Low	5 GHz	Corporate environment Public hot spots
IEEE 802.11a	~20 Mb/s	OFDM, TDD	50–300 m	Low	5 GHz	Corporate environment Public hot spots
HIPERACCESS	~25 Mb/s	Not yet specified	2...10 km	No	5–40 GHz	Business access, feeder
DAB	1.5 Mb/s	OFDM	100 km	High	176–230 MHz 1452–1467.5 MHz	Audio broadcasting
DVB-T	5–31 Mb/s per 8 MHz channel	OFDM	100 km	Medium to High	TV bands below 860 MHz	Video broadcasting
Cable modem	Down <40 Mb/s Up <10 Mb/s	FDD QAM/QPSK	5 to ~20 km	No	Down ~60 to 860 MHz Up 10 to ~40 MHz	Residential environment
ADSL	Down 6.144 (8) Mb/s Up (0.640 Mb/s)	DMT Carrierless AM/PM CAP	2–6 Km	No	Baseband	SoHo SME Residential environment

of a single operator that are typically charged a monthly subscription fee for each application. Income apportioning is usually based on a fixed, content/application provider-specific split of this fee between content provider and network operator. Thus, there is no support for functions like complex charging and network monitoring schemes.

These approaches constitute a considerable improvement compared to the 2G era (third parties had the chance to enter the market and widen the range of available content/applications, while service time-to-market has decreased) and provide a good basis for the wider adoption of mobile data services. However, they do not seem to be sufficient in the medium to long term. To achieve a new level of offerings and

significantly enhance the user experience, applications need to fully exploit the constantly growing potential of the underlying infrastructure (e.g., reconfigurability, accessibility of ambient information) and in the same time be adaptable and portable to a large heterogeneity of environments. Thus, significant functionality of particular complexity is required to enhance and augment service delivery and management as known until today. This functionality essentially should accomplish the appropriate orchestration of system and application resources for the sake of optimal and consistent service provision in highly diverse environments. A non-exhaustive list of tasks of particular importance include management of reconfigurability of system functions (e.g., traffic monitoring, charging/billing) in

Table 2. Platform/business model benefits for VASPs

Service development/deployment/commercial exploitation tasks	Effort required	
	Without platform	With platform
Acquisition/development of client software for open network API gateway frameworks	Major	Minor (only requirement: agreement with service platform provider)
Business agreements with open network API providers (for a single network)	Medium	Minor (only requirement: agreement with service platform provider)
Business agreements with open network API providers (for multiple networks/network types)	Proportional to number of networks	Fixed—the same as in the case of a single network
Development of context collection logic	Major	None or Minor
Development of adaptation logic	Major	None or Minor
Development of infrastructure reconfiguration logic	Major	Minor (typically only requires specification of service profile metadata)
Marketing the service	Major	None or Minor

terminals and networks, support of application adaptation, intelligent collection of contextual information as well as handling of inter-domain management interactions.

In this context, the potential of intelligent service management platforms that act as mediators between VASPs, end-users and mobile network operators can be clearly identified (Houssos et al., 2002; Grech et al., 2000). The above mentioned functionality, which is crucial for mobile service provision in 3G and beyond can be more efficiently handled in a unified way by a mediating platform residing between the networks and applications than by entities like network operators and VASPs, which cannot by themselves carry this excessive burden (cf. Section 6 and Table 2). In particular, we propose the introduction of a new business actor, called the *service platform operator/provider* who owns and administers such a mediating mobile service management platform. This approach has the important benefit of simplifying and lowering the overhead of certain important tasks, by assigning them to a specialized entity (service provision platform) instead of distributing the corresponding responsibility to a multitude of actors (e.g., connectivity providers, VASPs). Ultimately this could lead to drastically decreasing the overall cost of mobile service provision for the involved stakeholders (e.g., economies of scale when deploying an application over multiple networks in a single operation). That also lowers entry barriers to the mobile applications market, making it more appealing for new third-party application and content providers, a factor that is recognized as one of the key aspects for the success of any service provision model (Natsuno, 2002, 2003). Moreover, the administration of the platform by an

entity other than a network operator enables more flexibility in service deployment over heterogeneous infrastructures, inline with the requirements of 4G.

3. Advanced Business Model for Mobile Service Provision

The main business actors that we identify in the mobile service provision value chain are depicted in the UML class diagram of Fig. 1 and elaborated in the following. In the identification of business model actors we intentionally did not take into account players acting in “supporting” roles (e.g., terminal and network equipment manufacturers, terminal execution platform providers); only actors that directly participate in business transactions are included (Amit and Zott, 2001). A transaction is defined as a compound operation of some monetary value (e.g., service delivery, service deployment), which consists of a sequence of interactions between business actors.

In the proposed paradigm of the following roles are defined:

(Mobile) User: An entity that is the actual consumer of value-added services.

Connectivity provider/Network operator: The connectivity provider/network operator is an entity that provides the transport medium for mobile user access to basic communication services, which support the provision of value-added applications. A specialization of this role is the *Open API provider*, an entity that enables third-party applications to directly access selected network functionality. This is typically

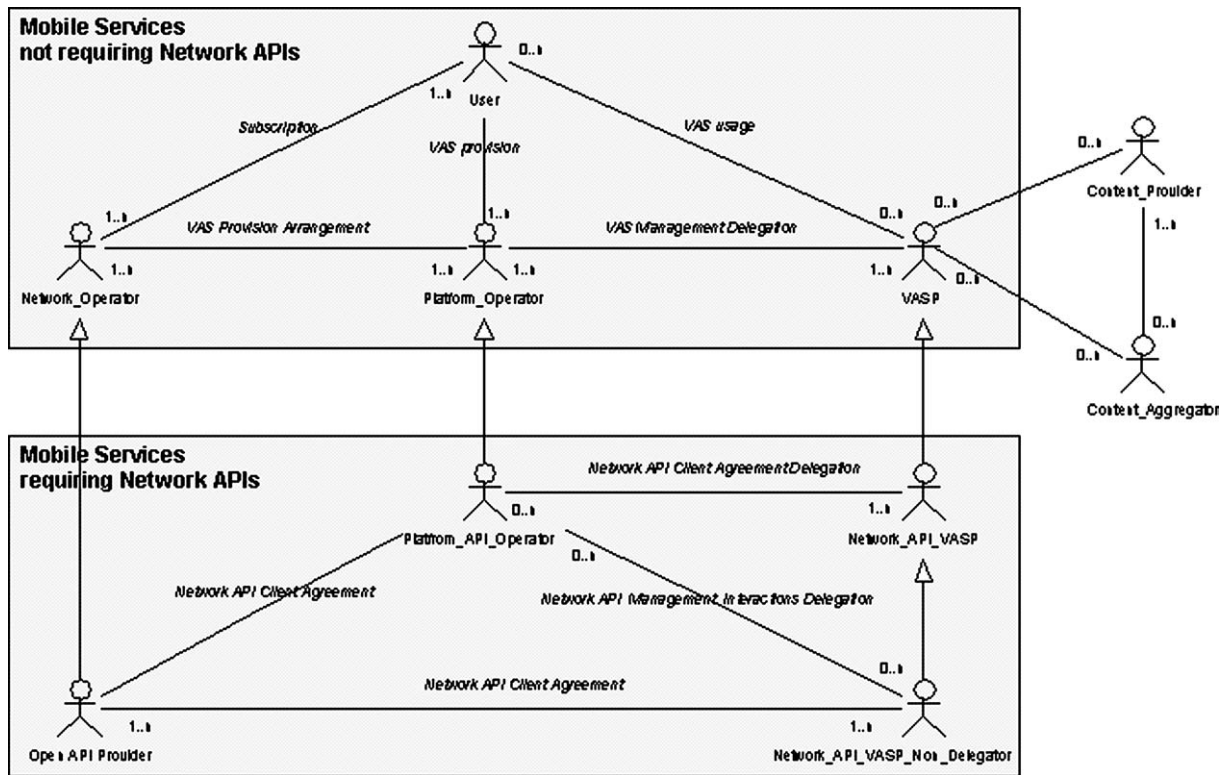


Fig. 1. Business actors in mobile service provision.

accomplished through open, standardized APIs (e.g., OSA/Parlay). In our model, the open API provider is a specialization of the connectivity provider (in other words, the open API provider is itself a connectivity provider). From a technical viewpoint, this is not mandatory; however it is expected to be the dominant paradigm in 3G networks and also to survive the advent of 4G systems, since the undertaking of the open API provider role from a third party introduces complex security issues.

Service platform operator/provider: A business entity that mediates between service developers, network providers and end-users by operating a software platform for service management and provision. The platform operator co-operates with network operators and VASPs to deliver applications to end-users. Relevant important tasks that are addressed by the platform provider are service deployment, management, discovery, adaptation and downloading. A particular category of platform operators

is modeled by a separate role: the *Platform API Operator*, which is able to handle the delivery and management of services that make use of network APIs.

Value-Added Service Provider (VASP): A business entity that owns and administers the implementation (e.g., software infrastructure) of a value-added service. The concept of *value-added services* encompasses all application-oriented services, whose value resides mostly in their functionality and content, instead of transport or connectivity. Such services can range from advanced messaging to mobile commerce and location-aware navigation guides. Service providers establish business level agreements with platform operators, outsourcing to them the deployment and provisioning of their services over various networks. Business-level agreements between VASPs and users are not required; however, they are not precluded. A specialized class of VASPs encompasses the providers of services that make use of the OSA/Parlay API

(*Network_API_VASP*). These providers may choose to undertake themselves the required business interactions with the open API operator (this corresponds to the *Network_API_VASP_Non_Delegator* role) or delegate them to the platform API operator.

Content Provider: An entity that is able to provide some form of content that can be of value to mobile users (e.g., stock prices, music, news). The content provider may establish business relationships with content aggregators or directly with VASPs, so that the content can be exploited in the context of mobile services.

Content Aggregator: An entity that collects content from various sources (content providers) and aggregates/packages it in a meaningful way so that it is more conveniently discoverable and accessible or more usable for interested parties (e.g., VASPs) than the data retrieved directly from content providers.

In our model, particular emphasis is placed on a special (albeit very broad) category of end user applications. These services are characterized by the fact that their implementations make direct use of the underlying infrastructure functionality (e.g., for establishing calls or retrieving location data), typically through open APIs. Applications of this class hold the potential of pushing the way forward for mobile service provision, since they combine access to information available from the network (including user context and status data) as well as to application-specific content, while being portable to different infrastructures (as they use standardized, network-independent APIs). However, their deployment and management imposes considerably increased requirements in comparison with “plain” services; for example, application registration with open interface gateways is necessary, including the specification of parameters (e.g., maximum permitted request frequency) related to the secure access and usage of the APIs. This is why our model includes explicit actor specializations that handle these tasks (in Fig. 1 there is a clear grouping of these actors).

Notably, in the proposed paradigm, there is not necessarily a one-to-one mapping between actors and real-life entities. For example, a connectivity provider could undertake the role of platform API operator. Moreover, a platform operator or content provider could also itself be a VASP.

4. An Advanced Service Management Platform

4.1. Platform overview and architecture

In this section, we introduce the architecture of a distributed middleware platform, aiming to address the issues described in Section 2.2. The proposed software framework should be administered by an entity undertaking the service platform operator/provider role, as elaborated in the preceding section and is designed to support the advanced business models previously described in this paper. The entire framework comprises several distributed software components, as shown in Fig. 2, whose functionality is briefly described in the following:

- The *Reconfiguration Control and Service Provision Manager (RCSPM)* (Houssos et al., 2002) is the central component of the platform, catering for the automatic service registration and deployment by 3rd party VASPs including the necessary reconfiguration actions on the underlying network for optimal service provision and the management interactions between mobile network/open API providers and VASPs. Moreover, it handles service adaptability to highly variable environments by providing to the end user a mobile portal for personalized and context-aware service discovery and access (Houssos, Pantazis and Alonistioti, 2002).
- The *Charging, Accounting and Billing system (CAB)* (Koutsopoulou, Farmakis and Gazis, 2001; Koutsopoulou, Kaloxylou and Alonistioti, 2002), calculates the overall user charges induced by service consumption and apportions the aggregate revenue among the involved business parties (e.g., VASPs, service platform providers, network operators). The system has been designed to allow the application of arbitrarily complex billing models and tariffing schemes and to be dynamically reconfigurable by the RCSPM.
- The *VASP* component of the platform is located in the VASP domain (e.g., enterprise servers) and handles secure terminal access to a repository of application clients, while also providing web interfaces to RCSPM functionality, through which VASPs can carry out sophisticated service management operations (e.g., service deployment).
- The *End-User Terminal Software (EUT)* (Fouial, Fadel and Demeure, 2002), which resides in the

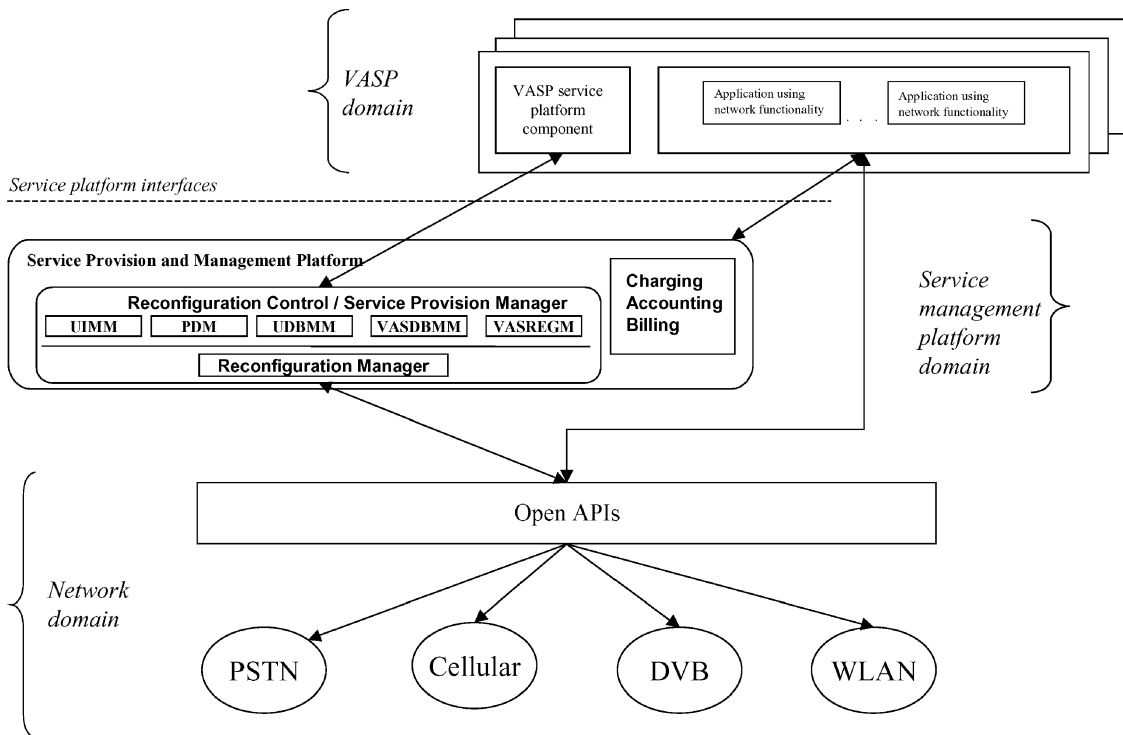


Fig. 2. Service management platform architecture.

end-user terminal (it is not depicted in Fig. 2), incorporates an execution environment for applications and functionality such as a graphical user interface for service discovery, secure downloading and lifecycle management (e.g., start/stop/abort/resume) of application clients, as well as initial user registration with the platform and subsequent user login/authentication. It is worth noting that the EUT needs not be built-in the terminal; it can itself be downloaded and dynamically installed, by any user that can become the appropriate authorization. The latter can take the form of a valid subscription with a network operator that has a contract with the service platform provider or a certificate issued by a trusted third party.

An important observation in Fig. 2 relates to the fact that the service platform is mainly orthogonal to the actual end-user applications and does not directly interact with them in most cases; that is, service code is not aware of the platform existence. A major advantage of this approach is that platform agnostic (e.g., legacy) applications can be seamlessly accommodated. However, there is an exception to this rule: open inter-

faces are exposed by the RCSPM and can be invoked by service logic for tasks like charging/billing and user context information retrieval.

4.2. Internal architecture of the RCSPM

As mentioned in the preceding paragraph, the Reconfiguration Control and Service Provision Manager (RCSPM) is the main component of the proposed architecture. It is responsible for a number of important tasks including mechanisms for application deployment and data management, service discovery by end-users, VASP and user authentication and adaptation of service provision to user preferences and terminal characteristics. The architecture of the RCSPM is depicted in Fig. 2.

In the following paragraphs, the functionality of the main modules of the RCSPM is presented.

The *User Interaction Management Module (UIMM)* is responsible for providing the user with a highly personalizable, context-aware mobile portal. To accomplish the appropriate functions, such as service discovery, the UIMM maintains context information, which is bound to a specific user session. This information is collected from appropriate context sources

(e.g., the underlying network, service and user profiles databases) in a variety of ways (e.g., “pull” by the UIMM or “push” by context sources using the corresponding generic interfaces exposed by the UIMM) and provided as an input to the RCM that performs the required intelligent matching operations and returns the corresponding results to the UIMM, which further exploits them to co-ordinate the user experience of customised service provision. Notably, arbitrary, not a priori known, types of context parameters can be dynamically retrieved and processed (Houssos et al., 2003).

The *VAS Registrar Module (VASREGM)* is responsible for interacting with 3rd party service providers. Through the VASREGM the platform operator provides VASPs with a way to automatically deploy their services. The VASP compiles a descriptor, encoded in XML, of service attributes. Based on these attributes, the VASREGM co-ordinates service deployment, including any required management (e.g., interactions with OSA/Parlay gateways) and infrastructure reconfiguration actions. This way the underlying network/transport technologies are hidden from the applications, which therefore do not need to be aware of them. The service provider is able to manage (add/delete/update) its services via a convenient web interface.

The *Reconfiguration Manager (RCM)* undertakes network, platform and service reconfigurability. The RCM is responsible for executing the appropriate reconfiguration actions on the underlying network, during VASP-initiated service management procedures (registration/de-registration/update). Moreover, it supports the context-aware adaptation of service delivery (e.g., service discovery, service configuration).

A part of the RCM that plays a crucial role in the above procedures is the generic adaptation module (Houssos et al., 2003) that is used for intelligent profile matching. An important characteristic of the adaptation module is its independence of the types of profiles that need to be matched and, consequently, its ability to accommodate profiles of any kind without modifications in its code.

The *Packaging and Downloading Module (PDM)* is responsible for creating a single downloadable bundle containing all the software components and other supporting resources (e.g., images, etc.) required for executing a service. The bundling procedure takes into account the current service provision context, so that the allocation of application components between client

and server is optimal. The major advantage of generating a single file containing all the data necessary for service execution is that a client is able to download the archive over a single network connection, thus avoiding superfluous network traffic.

The *Value-Added Service Database Module (VASDBMM)* is essentially a front-end to a database where service profile information is stored. Service data can be managed (inserted/deleted/updated) by the corresponding VASP.

The *User Database Management Module (UDBMM)* handles user profile information required for customized service provision. The data included in the profiles concerns user interface preferences, service-independent user preferences (language, pricing level preferences) as well as the list of the user’s favorite services. The stored profiles can be viewed/modified by the user at any time.

4.3. Platform prototype

To provide proof of concept, a prototype of the proposed service management platform has been developed and successfully tested in a laboratory environment, comprising PCs and laptops in various configurations. The radio technologies employed for wireless access were IEEE 802.11 and Bluetooth. For QoS provisioning as well as generation of metering records required for billing, we used an implementation of reconfigurable L4+ systems by NEC Europe, Network Laboratories (Harbaum et al., 1999), which are essentially DiffServ-enabled IP routers with additional capabilities for flow monitoring and record generation based on transport and application layer information.

The platform was designed, using object-oriented techniques, with UML and SDL and developed mainly in Java (JDK 1.3). The RCSPM was implemented using mainly server-side enterprise Java technologies (e.g., Servlets, Java RMI, JDBC). Communication across interfaces that span administrative boundaries (e.g., VASP-RCSPM, end user terminal—RCSPM) is based on standard, highly interoperable technologies like HTTP and XML. The RCSPM has been designed and implemented in a modular fashion, so that it can operate in a variety of topologies (only IP connectivity is assumed between modules), the two extreme cases being the various components to be integrated in one physical element or completely distributed.

The ability of the platform prototype to handle service provision was validated using two example

applications, namely a streaming multimedia player and a location-aware instant messenger. The choice of the example services was such that they have different requirements in terms of critical delivery aspects like QoS, billing and adaptation. Different versions/configurations of these applications were delivered to mobile users over our wireless testbed, with the support of the platform prototype.

5. *Mobile Service Provision Scenario*

In this section, we present an example of service provision (for the specific case of a 3G network), where the previously mentioned business models can be applied with the support of the proposed platform. The case presented also illustrates the disparity in possible context-dependent situations in next generation service provision and the potential of service management platforms for addressing such environment unpredictability.

Before elaborating on our scenario, we should briefly refer to our perception of the “value added service” concept, which was taken into consideration in designing and implementing the platform. A service will typically consist of code that runs in the end-user terminal (service client) and optionally a server part. The former includes, besides other functionality, the service user interface. It can be pre-installed in the terminal, or downloaded and executed on demand. The latter case is aligned with the dynamic nature of service provision desired for 4G systems. The client part comes in multiple predefined versions or it can be even dynamically composed (in a context dependent manner) by application components on demand. For example, two distinct editions of a multimedia player application may support streaming of audio/video and audio-only files, respectively. Service execution normally, but not necessarily, includes interaction with the server part. If available, the application server part is located in the premises of the service provider (e.g., web server). Multi-tier architectures will be also commonplace, where the server part comprises many interacting components. These components may be located in the administrative domains of entities other than the service provider, like content providers marketing access to their databases, or network operators providing interfaces to network functionality (e.g., OSA, Parlay, JAIN).

Coming back to our example, we consider a hypothetical multimedia application for viewing news. The real-time content, regarding news in a wide geographical area spanning many different countries and coming in various forms (e.g., text, voice, video and a range of combinations of the above), is available through a large media company in the role of content provider/aggregator. A VASP develops an advanced application, which enables viewing of the news information, including news retrieval based on various criteria (topic, time, locality), and comes into agreement with the media company for access to the news repository. The service uses information from the underlying infrastructure regarding user location and also employs the Charging OSA API of 3G mobile operators to apply content-based charging (besides transport and service level charging that is available in co-operation with the platform), that is it is able to directly charge the user account upon the completion of certain “transactions” (e.g., downloading of a video file). The configuration (i.e., constituent components) of the application client used in the terminal depends on the media type(s) employed for news content presentation. The VASP also establishes a business-level agreement with a service platform provider, which concerns the deployment of the service over various 3G mobile networks (the latter includes arrangements for providing the required QoS to the application, as well as composition, packaging and transfer to the client of the appropriate application versions by the platform) and the delegation of OSA/Parlay related management interactions to the service platform operator. For each user request, several context-dependent factors determine the optimal service provision configuration. Let us consider the following hypothetical scenarios, for a user that requests textual representation of news information, accompanied with appropriate audio and video files (e.g., politician statements in audio form or videos with sports highlights):

1. The mobile terminal that the user is currently employing does not have persistent storage capabilities suitable for storing audio and video files, thus leaving streaming mode as the only feasible option of immediate content delivery. Since persistent storage at the user’s mobile terminal is not possible, the charging model applicable by the platform should reflect the cost resulting from the use of wireless and wireline network resources for one-time-delivery with a high-level of QoS appropriate for streaming.

2. The mobile terminal that the user is currently employing has persistent storage capabilities suitable for audio and video files, thus implying the application of a different charging model, since now the value of the transaction lies in the permanent availability of (probably) copyrighted content. Thus, the charging model applied by the platform is purely related to the (probably best-effort) transport of the appropriate files, however, after the successful completion of the downloading the user is directly charged by the application via the OSA/Parlay interface.
3. The mobile terminal that the user is currently employing has persistent storage capabilities, however, the user's account balance is insufficient for the download-and-store delivery mode described in case 2 above, leaving the option of the download-and-view-once service offered by the value-added service provider. In this case, the applicable charges will include a minor direct charge by the application related to copyright and the best-effort level use of network resources.

The important advantage of such a scenario is that the delivery of the applications in all these different cases is attractive and beneficial for users, VASPs as well as network operators, since it is always performed through the best possible choice of alternatives regarding critical functions (e.g., charging). Evidently, delegating the complexity of coping with such multifaceted alternatives in an efficient manner to the application developer introduces unnecessary complexity in his/her service development process and bloats up the application code. We should note that it is not as much the functionality required by each aforementioned scenario (i.e., streaming functionality, download functionality, etc.) as the accommodation of the decision logic that introduces the major complexity. For instance, retrieving context information from various sources and selecting the most appropriate alternative leads to significant overhead for the service developer. The proposed service provision platforms address such multi-option situations by providing functionality that can efficiently support context-dependent decision-making processes on behalf of both mobile users and value-added service providers. Moreover, the delegation of interaction with open API providers to the service platform operator significantly reduces the service deployment overhead for the VASP.

6. *Applicability in Real-Life—Discussion*

In the present section, we discuss whether the paradigms and solutions presented in the previous sections can be applicable for real-life, commercial mobile service provision. At first we discuss business aspects and subsequently issues related to the technical feasibility of our approach.

An obvious question is whether the model of Section 3 makes sense in business terms. Its main goal is to attract VASPs into developing and deploying mobile services and tries to achieve this by making the corresponding processes substantially more efficient, which is one of the principal value-creating factors in a business model (Amit and Zott, 2001). In particular, the following benefits in terms of efficiency can be identified:

1. Automation of service installation and deployment drastically reduces service-to-market costs for network operators as well as VASPs and leads to the availability of a larger number of services and thus to more traffic and revenue for operators. Moreover, the reduced time between the development and the commercial exploitation of a service leads to a smaller payback period for VASPs that, cumulatively, facilitates production of a wider gamut of (mobile) services.
2. The platform support for service adaptation to context parameters (e.g., terminal and network capabilities, user location and preferences) and transparent reconfiguration of the underlying infrastructure during service deployment off-loads from the VASPs the cost of implementing on their own the logic for collecting environmental information as well for making and putting into force intelligent adaptation decisions.
3. The platform creates a one-stop shop on two sides: for the mobile user, since the available services can be discovered, downloaded and accessed in a unified and customizable way through a portal-like interface; and for the VASPs with the one-step deployment of a single service over multiple networks (and even different network types, e.g., UMTS and WLAN) and the delegation of interaction with the mobile operator for access to open network APIs. Furthermore, the automatic inclusion of services into a user portal with the advanced service discovery capabilities reduces the cost of marketing a new application, a benefit, which can be particularly

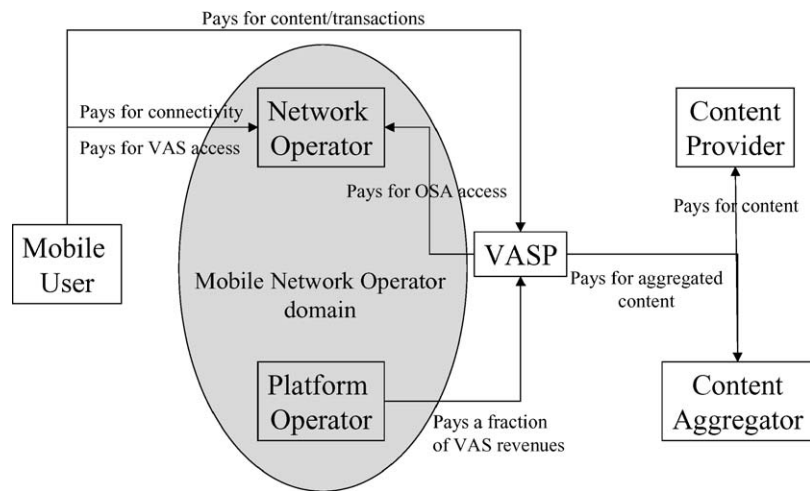


Fig. 3. Revenue flows for the case that the service management platform is administered by the mobile network operator.

helpful for small and medium sized VASPs without a strong brand.

The benefits for VASPs regarding service development, deployment and commercial exploitation are summarized in Table 2.

Figs. 3 and 4 depict revenue flows for two different projections of the business model in the real world, involving the business actors defined in this section, for mobile service provision in 3G networks and beyond. It is worth noting that revenue flows (“who gets money from whom”) ultimately depend on the relative strengths of the involved entities in a particular

context and cannot be accurately determined or predicted. Thus, revenue generation patterns other than those illustrated in the figures below could occur between the same business actors in a given real-world situation, although we believe that the given examples represent the most likely cases.

In the model shown in Fig. 3, the 3G mobile operator undertakes the role of the service platform provider and is directly engaged in business agreements with the VASPs for both VAS deployment and the management of VAS access to network functionality via the OSA/Parlay interface. The user receives a single bill

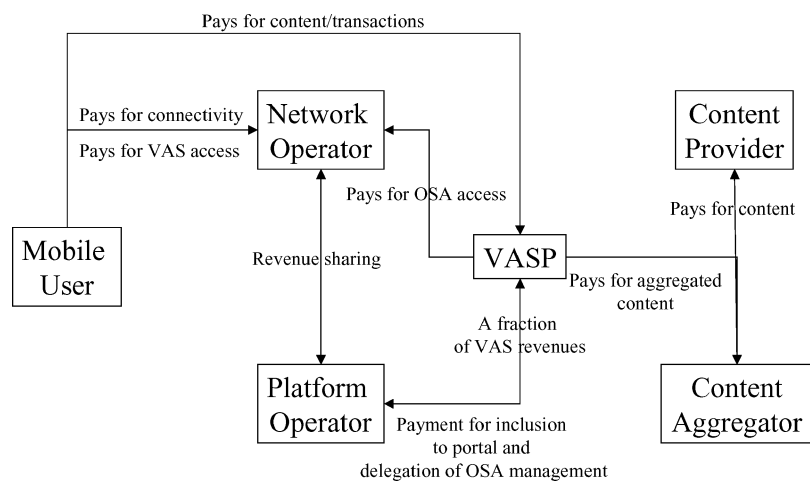


Fig. 4. Revenue flows for the case that the service management platform is administered by an entity other than the mobile network operator.

for basic and value added services. Part of the revenue from the latter goes to the VASP. This arrangement could be a natural evolution of the pre-3G paradigms (Tsalgatidou and Pitoura, 2001). The typical mobile network operator possesses significant strategic advantages, like an existing large subscriber base, direct customer relationship as well as easier access to information provided by the network, like user location. The trust relationship with its subscribers can be an important factor for attracting end users to 3G services that could benefit the overall development of the 3G market in the short to medium term. However, it is doubtful whether operators are able or willing to significantly extend their expertise in the IT domain so that they can cope with such a role. Moreover, a network operator running a service provision platform would probably offer deployment of services solely over its own network.

In Fig. 4, a different setting is illustrated, where the service management platform is administered by an entity other than the mobile operator. Revenue sharing should then be applied also between network and platform operator. This paradigm enables third parties to become specialized in the service platform operator role, partner will various network operators and thus offer VASPs fast service deployment over more than one networks through only a single business-level agreement. Nevertheless, this would require significant effort, for gaining customer recognition (from both end-users and VASPs) as well as establishing non-exclusive business relationships with multiple operators. Entities that have the potential to overcome these difficulties and successfully undertake the platform operator role are enterprises that have an already established customer relationship and have built a strong brand (e.g., media companies like broadcasters and newspaper publishers or wired Internet portal providers).

Another major concern relates to whether the realization and operation of a platform like the one introduced in Section 4 is technically feasible in the short to medium term. There are clear signs that this is possible, as indicated by the already completed prototypical implementation of the platform, which has been performed using general-purpose software design and development tools that are not specific to telecommunication environments. Our experiments showed that the platform makes the service deployment and management process much easier for the VASPs (IST-10206-MOBIVAS, 2002) that do not have to alter their service implementations, but only to provide

to the RCSPM detailed XML-encoded application metadata, through a friendly web-based user interface. The realization of the complex, extensible logic that is necessary for service profile interpretation, decision making and reconfigurability management was greatly facilitated by the use of open, ubiquitous, interoperable technologies like HTTP, XML and Java and by the adoption of object-oriented techniques and design patterns. Through the platform, we were also able to validate the applicability of our service delivery paradigm through the demonstration of adaptable service discovery, downloading and execution by mobile users as well as of the corresponding billing and revenue sharing intelligence (IST-10206-MOBIVAS, 2002). All the required functionality proved to be implementable with reasonable effort using existing technologies.

An important feature that facilitates real-life platform utilization is that good levels of performance and scalability can be achieved, due to the fact that platform design is modular, results to the decoupling of the individual components and leads to infinite possibilities for performance-optimizing physical deployment and configuration of the platform. Moreover, platform signaling over the restricted capabilities air interface is minimal, moving the bulk of the processing and communication overhead to the resource-rich fixed infrastructure.

Another important aspect concerns support of third-party access to selected network functionality through open standardized APIs, so that the platform can seamlessly operate over various networks. There is a trend for overcoming the traditional reluctance of telecommunication operators to open-up their networks; this is indicated not only by the significant participation of telecommunication operators in international open API standardization organizations and fora (e.g., 3GPP/ETSI, Parlay Group) but also by initial open gateway commercial deployments that have already demonstrated the operator benefits for adopting a network openness policy (Caroll, Porteous and Taylor, 2002). These standardized open APIs cover a wide range of functionality, although extensions are definitely required to provide more functionality and in particular to enable more network reconfiguration options (Alonistioti, Houssos and Panagiotakis, 2001). Given the fast standardization processes employed by the aforementioned organizations and fora, currently ongoing, intensive activities for open API extension are expected to lead to specification and commercial adoption/implementation of a far richer set of

interfaces within a relatively short time period (in the range of 3–4 years). This was also approximately the time scale required for the first commercial operation of Parlay APIs since the initiation of the corresponding specification/standardization activities.

A significant enabler is also the availability of rich contextual information regarding each user session. Techniques for tracking and collecting certain types environment data (terminal capabilities, network characteristics, user location and preferences) are already there not only in research prototypes (like ours) but also in the industry. The potential of context-awareness can, however, be exploited to a substantially greater extent as soon as ambient information (e.g., from sensors spread all over the physical environment) and effective methods of identifying and deriving user status are widely available. Notably, our implementation is generic enough to accommodate such a development, since arbitrary (and not a priori known) types of context data can be handled on-the-fly without any code modifications or recompilation. However, this vision is unlikely to become commonplace in a timeframe shorter than a decade.

A critical issue also concerns user billing as well as automatic revenue sharing between stakeholders. Given that the necessary techniques and mechanisms are in place (Koutsopoulou, Kaloxylos and Alonistioti, 2002) and that similar ideas that have already found their way into commercial products (DSCP, 2003), the introduction of flexible schemes in mobile services charging seems quite feasible even in the short to medium term.

7. Conclusions

Future communication systems from 3G and beyond are expected to create a new, exciting era in mobile service provision. Their market opportunities are tremendous, but so are the technical and business challenges that must be confronted. To a great extent, success of next generation systems will depend on whether a sufficiently large mass of independent application developers will be engaged and committed to the production of innovative value-added services and feature-rich applications for diverse environments. With that in mind, technical solutions and business arrangements that lower the technical complexities and service deployment overhead imposed on VASPs and motivate

them to embark on the conception, specification and development of innovative service products should be given absolute priority.

Service provision platform approaches present an attractive solution that can contribute to the efficient resolution of relevant technical issues (e.g., inter-domain network and service management) whilst supporting the advanced and more complex business arrangements that can reap market success in the forthcoming 3G era and beyond. Particularly in combination with open, reconfigurable network environments and adaptable applications, service provision platforms can help realize important benefits for all players, users, mobile network operators and value-added service providers alike.

Acknowledgments

Preliminary work and prototypes related to the contributions presented in this paper have been developed within the frame of project IST-10206-MOBIVAS, <http://mobivas.cnl.di.uoa.gr>. The authors would like to acknowledge the fruitful co-operation and exchange of ideas between the MOBIVAS partners. The authors would also like to thank Spyros Pantazis, Zachos Boufidis, Kostas Saidis and Costas Sourbatis for their valuable contribution to the development of the service provision platform prototype.

References

- Alonistioti A, Houssos N, Panagiotakis S. A framework for reconfigurable provisioning of services in mobile networks. In: *Proceedings of the 12th International Symposium on Communication Technologies and Applications (ISCTA 2001)* Ambleside, UK, July 2001:21–26.
- Amit R, Zott C. Value creation in E-business. *Strategic Management Journal* 2001;22:493–520.
- Annoni M, Hancock R, Paila T, et al. Radio access networks beyond 3G: A first comparison of architectures from 4 IST projects. In: *Proceedings of the 2001 IST Mobile Communications Summit*, Barcelona, Spain, Sept. 2001.
- Carroll G, Porteous K, Taylor M. *The OSA/Parlay Report*. Moriana Group Report, 2002.
- Dillinger M, Alonistioti N, Madani K (eds.). *Software Defined Radio: Architectures, Systems and Functions*. John Wiley & Sons, 2003.
- Enabling UMTS/Third Generation Services and Applications. UMTS Forum, Report #11. URL <http://www.umts-forum.org>, June 2003d.
- Fouial O, Fadel KA, Demeure I. Adaptive service provision in mobile computing environments. In: *Proceedings of the 4th IEEE*

- International Conference on Mobile Wireless Communication Networks (MWCN 2002)*, Stockholm, Sweden, 9–11 Sept. 2002.
- Garlan D, Siewiorek D, Smailagic A, Steenkiste P. Project Aura: Towards distraction-free pervasive computing. *IEEE Pervasive Computing*, April–June 2002.
- Gazis V, Houssos N, Alonistioti A, Merakos L. Evolving perspectives of 4th generation mobile communication systems. In: *Proceedings of the 13th International Symposium on Personal, Indoor and Mobile Communications (PIMRC 2002)*, Coimbra, Portugal, 15–18 Sept. 2002:201–207.
- Grech M, McKinney R, Sharma S, Stanaway J, Varney D, Vemuri K. Delivering seamless services in open networks using intelligent service mediation. *Bell Labs Technical Journal*, July–Sept. 2000.
- 3rd Generation Partnership Project (3GPP), <http://www.3gpp.org>, June 2003.
- 3rd Generation Partnership Project 2 (3GPP2), <http://www.3gpp2.org>, June 2003.
- Harbaum T, Zitterbart M, Griffoul F, Röthig J, Schaller S, Stüttgen HJ. Layer 4+ switching with QoS support for RTP and HTTP. In: *Proceedings of the IEEE Globecom conference 1999*, Rio de Janeiro, Brazil, Dec. 1999.
- Houssos N, Alonistioti A, Merakos L, et al. Advanced adaptability and profile management framework for the support of flexible mobile service provision. *IEEE Wireless Communication Magazine*, Aug. 2003.
- Houssos N, Gazis V, Panagiotakis S, Gessler S, Schuelke A, Quesnel S. Value-added service management in 3G networks. In: *Proceedings of the 8th IEEE/IFIP Networks Operations and Management Symposium (NOMS 2002)*, Florence, Italy, 15–19 April 2002:529–545.
- Houssos N, Koutsopoulou M, Schaller S. A VHE architecture for advanced value-added service provision in 3rd generation mobile communication networks. In: *Proceedings of the IEEE GLOBECOM Conference 2000 Workshop on Service Portability and Customer Premises Environments*, San Francisco, USA, Dec. 2000:69–78.
- Houssos N, Pantazis S, Alonistioti A. Towards adaptability in 3G service provision. In: *Proceedings of the IST Mobile Communications Summit 2002*, Thessaloniki, 16–19 June 2002.
- IST-10206-MOBIVAS project Deliverable 6.1.1. *Evaluation Report*. September 2002.
- Jamalipour A, Tekinay S eds. Fourth generation wireless networks and interconnecting standards. Special Issue of *IEEE Personal Communications Magazine*, 2001.
- Katz RH. Adaptation and mobility in wireless information systems. *IEEE Personal Communications Magazine* First Quarter 1994.
- Koutsopoulou M, Farmakis C, Gazis V. Subscription management and charging for value added services in UMTS networks. In: *Proceedings of the IEEE VTC Spring 2001*, Rhodes, Greece, May 2001.
- Koutsopoulou M, Kaloxylos A, Alonistioti A. Charging, accounting and billing as a sophisticated and reconfigurable discrete service for next generation mobile networks. In: *Proceedings of the IEEE VTC Fall 2002*, Vancouver, Canada, Sept. 2002.
- Lazar A. Programming telecommunication networks. *IEEE Network Magazine* 1997.
- Mobile Station Application Execution Environment*. 3GPP TS 23.057.
- Mohr W. Access network evolution beyond third generation mobile communications. *IEEE Communications Magazine* Dec. 2000:122–133.
- Natsuno T. *Ten proven facts on Wireless Internet*. Keynote address at the *IST Mobile Communications Summit 2002*, Thessaloniki, Greece, 16–19 June 2002.
- Natsuno T. *i-mode strategy*. Halsted Press, 2003.
- Network architecture*. 3GPP TS 23.002.
- NTT DoCoMo's I-mode. URL <http://i-mode.nttdocomo.com>, June 2003.
- Parlay Specifications, URL <http://www.parlay.org/specs/index.asp>, June 2003.
- Pereira J. Beyond third generation. In: *Proceedings of the 2nd Wireless Personal Mobile Communications (WPMC) Conference*, Amsterdam, The Netherlands, 1999.
- Services and Service Capabilities*. 3GPP TS 22.105.
- Services Principles*. 3GPP TS 22.101.
- Shaping the Mobile Multimedia Future*. UMTS Forum, Report #10. URL <http://www.umts-forum.org>, June 2003c.
- Tsalgatidou A, Pitoura E. Business models and transactions in mobile electronic commerce: Requirements and properties. *Journal of Computer Networks*, Special Issue on Electronic Business Systems, 2001;37(2):221–236.
- The Future of Mobile Market*. UMTS Forum, Report #8. URL <http://www.umts-forum.org>, June 2003a.
- The UMTS Third Generation Market—Structuring the Service Revenues Opportunities*. UMTS Forum, Report #9. URL <http://www.umts-forum.org>, June 2003b.
- Virtual Home Environment*. 3GPP TS 23.127.
- Volubill Dialog Control and Charging Platform (DSCP), URL <http://www.volubill.com>, June 2003.
- Weiser M. The computer for the twenty-first century. *Scientific American* 1991;94–10.

Nikos Houssos has obtained a B.Sc. and a Ph.D. degree from the Department of Informatics at the University of Athens in 1998 and 2004, respectively, and a M.Sc. in Telematics (Communications & Software) from the department of Electronic and Electrical Engineering, University of Surrey in 1999. He is presently a senior researcher at the Communication Networks Laboratory of the University of Athens. He has participated in several national and European R&D projects. He has more than 30 publications in international journals, conferences and book chapters related to mobile computing and networking. His current research interests include middleware platforms and business models for service provision in 3G/4G networks, service adaptation, reconfigurability management and pervasive computing.

Vangelis Gazis received his B.Sc. and M.Sc. from the Department of Informatics and Telecommunications (University of Athens) in 1995, and 1998, respectively, and his MBA from the Athens University of Economics

and Business in 2001. Since 1995, he has been with the research staff of the Communication Networks Laboratory, at the aforementioned department, working in funded research and development projects. In parallel, he worked with a number of established companies in the Greek IT sector. He is currently pursuing a Ph.D. His research interests include adaptable service provision, reconfigurable systems, quality of service, billing, and business model issues in 3G/4G networks.

Athanassia Alonistioti has a B.Sc. degree and a Ph.D. degree in Informatics and Telecommunications

(University of Athens). She is senior researcher and project manager in Communication Networks Laboratory (CNL). She specializes in reconfigurable mobile systems and networks and adaptable services, pervasive computing and context awareness. She has participated in several national and European projects, and was Technical manager of the IST-MOBIVAS and ANWIRE projects. She is member of the PMT of the IST-E2R project. She is co-editor/author of "Software defined radio, Architectures, Systems and Functions," published by John Wiley in May 2003. She has over 60 publications in the area of mobile communications and reconfigurable systems and networks.