The multimedia broadcast/multicast service

George Xylomenos, Vasilis Vogkas and George Thanos

{xgeorge,vvogkas,gthanos}@aueb.gr

Mobile Multimedia Laboratory, Department of Informatics

Athens University of Economics and Business

Patision 76, Athens 104 34, Greece

Abstract— This article provides an overview of the multimedia broadcast/multicast service (MBMS) for universal mobile telecommunications system (UMTS) networks. We first outline the features of UMTS networks as defined by the 3rd generation partnership project (3GPP) in order to provide a background for the discussion to follow. We then present the overall MBMS architecture, the services that it provides to the users and the differences between the broadcast and multicast options. The implementation details of MBMS are explained in terms of the modifications needed to the network, the new signaling procedures required and the impact of MBMS on the radio part of the network. We then describe how the security architecture of UMTS is used to support content protection and key distribution for MBMS groups. We evaluate the prospects of MBMS by comparing it first with IP multicast and then with DVB-H, considered by many to be the closest competitor of MBMS. Finally, we conclude with a discussion of the technical and business challenges still faced by MBMS.

Index Terms— MBMS, UMTS, Multimedia, Broadcast, Multicast.

I. INTRODUCTION

One of the most important emerging aspects of *3rd Generation* (3G) networks, and in particular of the *Universal Mobile Telecommunications System* (UMTS) defined by the *3rd Generation Partnership Project* (3GPP), is the introduction of the *Multimedia Broadcast/Multicast Service* (MBMS). The MBMS provides native support for broadcasting and multicasting IP packets in UMTS networks, thus allowing high bandwidth services to be offered to multiple users in an economical manner. Examples include video streaming and software distribution via multicast, as well as transmission of local news and weather reports via broadcast or multicast.

This article provides a comprehensive introduction to MBMS from both the services and the implementation viewpoints, as well as an evaluation of its potential for success. The introduction to MBMS aims to present the advantages and disadvantages of MBMS from the viewpoint of UMTS network operators and users; while MBMS can potentially provide considerable savings in transmission bandwidth over the air, it incurs considerable signaling overhead and may not even make sense for small numbers of users. The evaluation of MBMS discusses whether MBMS can succeed where IP multicast has failed, especially in the face of competing systems like DVB-H, and presents the technical and business obstacles that MBMS still needs to overcome.

Corresponding Author: George Xylomenos, Athens University of Economics and Business, Patision 76, Athens 104 34, Greece.

The outline of the remainder of this article is as follows. In Section II we present a short description of UMTS networks. Section III introduces the MBMS architecture and its services. Section IV discusses MBMS implementation issues covering both the extensions required to UMTS networks and the new signaling procedures needed. Section V covers the security aspects of MBMS. In Section VI we discuss the differences between MBMS and IP multicast and their potential effects on the commercial success of MBMS, while in Section VII we compare MBMS with DVB-H in an attempt to determine whether there is room for both systems in the market. Finally, in Section VIII we discuss the challenges that MBMS still needs to address in order to succeed.

II. UMTS NETWORKS

During the 1980s, the desire for the full digitization of the *Public Switched Telephone Network* (PSTN) led to the introduction of the *Integrated Services Digital Network* (ISDN). Similarly, during the 1990s, *2nd Generation* (2G) digital cellular systems were introduced to replace the earlier analog ones. The most successful 2G system turned out to be the *Global System for Mobile Communications* (GSM), created by a consortium of European manufacturers and operators. While GSM is circuit switched and optimized for voice, the increased importance of data led to its extension with the *General Packet Radio Service* (GPRS). GPRS allows multiple packet calls to dynamically share the radio link, using a packet switched data backbone that is separate from the GSM circuit switched voice backbone.

The incompatibilities and technical limitations of 2G systems led the *International Telecommunications Union* (ITU) to initiate work on a worldwide 3G standard. Eventually however, two separate standardization groups were formed. The *3G Partnership Project* (3GPP) is developing a system based on *Wideband CDMA* (W-CDMA) for the radio part and GSM/GPRS for the network backbone, while the *3G Partnership Project 2* (3GPP2) is developing a CDMA2000 based system [1]. In this article we focus on the 3GPP work related to the MBMS. A UMTS network as defined by 3GPP consists of two parts, the *Radio Access Network* (RAN) and the *Core Network* (CN), with the *User Equipment* (UE) connecting to the CN via the RAN, as shown in Figure 1. Since the RAN provides wireless access to the network, it is inherently tied to a specific wireless technology; in contrast, since the CN provides generic UMTS services to the users, it is independent of the underlying wireless technology. This separation

allows different RAN and CN options to be combined, thus supporting multiple evolutionary paths from 2G to 3G.

The CN consists of the *Circuit Switched* (CS) and *Packet Switched* (PS) domains, as shown on the right side of Figure 1. The CS domain is an evolution of GSM, while the PS domain is an evolution of GPRS. The *Home Subscriber Server* (HSS) holds user information for both domains. Calls in the CS domain are handled by two *Mobile services Switching Centers* (MSC): the *Gateway MSC* (GMSC), located at the user's home network, and the *Visitor MSC* (VMSC), located at the network that the user is currently visiting. When the user enters a new network, its VMSC informs the local *Visitor Location Register* (VLR) about the user, and the VLR informs the HSS about the user's new location. Incoming calls are first directed to the GMSC, which passes the call to the proper VMSC based on information from the HSS. For ISDN or PSTN originated or terminated calls, a *Media GateWay* (MGW) is employed for media and signaling conversions.

In the PS domain, calls are also handled by two *GPRS Support Nodes* (GSN): the *Gateway GSN* (GGSN), located at the home network, and the *Serving GSN* (GGSN), located at the visited network. However, in the PS domain the GGSN always knows the SGSN currently handling a user, therefore, the VLR is not needed. For Internet originated or terminated calls, the GGSN acts as an IP gateway router. For a UE to gain IP connectivity, it must first *attach* to the GPRS, whereby the user is authenticated by the HSS, and the SGSN creates a mobility management context for the UE. The UE then creates a *Packet Data Protocol* (PDP) context at the GGSN, containing the IP address assigned to the user. At this point, the user can send and receive IP packets via the GGSN [2].

The RAN provides two connectivity options, as shown on the left side of Figure 1. The first option is the *GSM EDGE RAN* (GERAN), which is based on the *Enhanced Data rates for GSM Evolution* (EDGE) technology. While EDGE reuses the GSM frequency allocations, it provides higher bandwidth via advanced modulation and coding techniques. In the GERAN, each radio cell is served by a *Base Transmitter Station* (BTS), multiple BTSs are controlled by a *Base Station Controller* (BSC) and multiple BSCs form a *Base Station Subsystem* (BSS). The GERAN allocates each frequency band to either voice or packet data calls, and then employs *Time Division Multiple Access* (TDMA) to multiplex calls within a band. In TDMA, time is divided into slots; a voice call gets to use a single slot periodically, while a packet data call is dynamically allocated one or more slots whenever it needs to transmit or receive packets. The GERAN provides data rates of at least 384 kbps for urban environments and 144 kbps for rural environments.

The second RAN option is the *Universal Terrestrial RAN* (UTRAN), where one or more cells are served by a Node-B, multiple Node-Bs are controlled by a *Radio Network Controller* (RNC), and multiple RNCs form a *Radio Network Subsystem* (RNS). The UTRAN uses *Code Division Multiple Access* (CDMA) to share the frequency range between multiple simultaneous calls. In CDMA, all calls employ the entire frequency range all the time, but each call uses a different code to scramble its data at a very high speed. By choosing

appropriate codes and coding/decoding schemes, each call distinguishes its own data, treating all other transmissions as random noise. Higher data rates are supported by assigning multiple codes to a call. The UTRAN supports speeds of more than 2048 kbps in small cells or indoor areas. To achieve these rates, the UTRAN employs *power control*, where the transmissions between a UE and a Node-B use the minimum power required for acceptable connectivity. This reduces the noise perceived by each user due to the activity of other users, thus allowing higher data rates to be achieved.

Originally, the UMTS specifications were released on a yearly basis by the *European Telecommunications Standards Institute* (ETSI) and named accordingly, for example, Release 97. Starting from Release 99, however, the 3GPP assumed responsibility for UMTS and changed the numbering scheme. The next step, Release 4, introduced modifications to the CN, such as the use of IP to transport voice and data. Release 5 also allowed IP to be used inside the UTRAN to transport voice and data. Release 6 introduced, among other things, the MBMS [3]; we use Release 6 as the basis for the following discussion.

III. MBMS ARCHITECTURE

The main attraction of MBMS is that it allows many receivers in the same radio cell to be served by a common signal transmission facility, or *bearer*, thus conserving radio resources. This is not a new idea; since Release 4, UMTS networks provide a *Cell Broadcast Service* (CBS) that transmits unacknowledged short messages to all users in a particular area. An information provider passes these messages to a *Cell Broadcast Center* (CBC) for transmission, and the CBC broadcasts each message periodically, at a frequency and duration arranged with the information provider. For example, road traffic reports will probably be retransmitted more frequently than weather reports [4]. The CBS, however, is targeted to text messaging, therefore, it is unsuitable for high bandwidth multimedia services.

Since Release 99, UMTS networks added support for IP multicast, where IP packets are forwarded to all users belonging to a multicast *group* identified by a class D IP address. Any UE may join or leave a multicast group without restrictions. As shown on the left side of Figure 2 though, Release 99 IP multicast is implemented by separately sending each packet from the GGSN to each UE, therefore, no sharing gains are achieved, and high bandwidth multimedia services remain expensive.

The Release 6 specifications in a sense combined ideas from IP multicast and the CBS to produce the MBMS, a new service that supports native multicast and broadcast over UMTS networks. MBMS can transport any type of data, unlike the text oriented CBS. MBMS uses IP multicast packets, that is, packets sent to a class D IP address, but, as shown on the right side of Figure 2, the GGSN and SGSN send multicast packets only once to each downstream node. More importantly, packets are transmitted only once in each cell, at least when a sufficient number of intended receivers is present there.

A UMTS network may provide many simultaneous MBMS *user* services, each consisting of one or more MBMS *bearer*
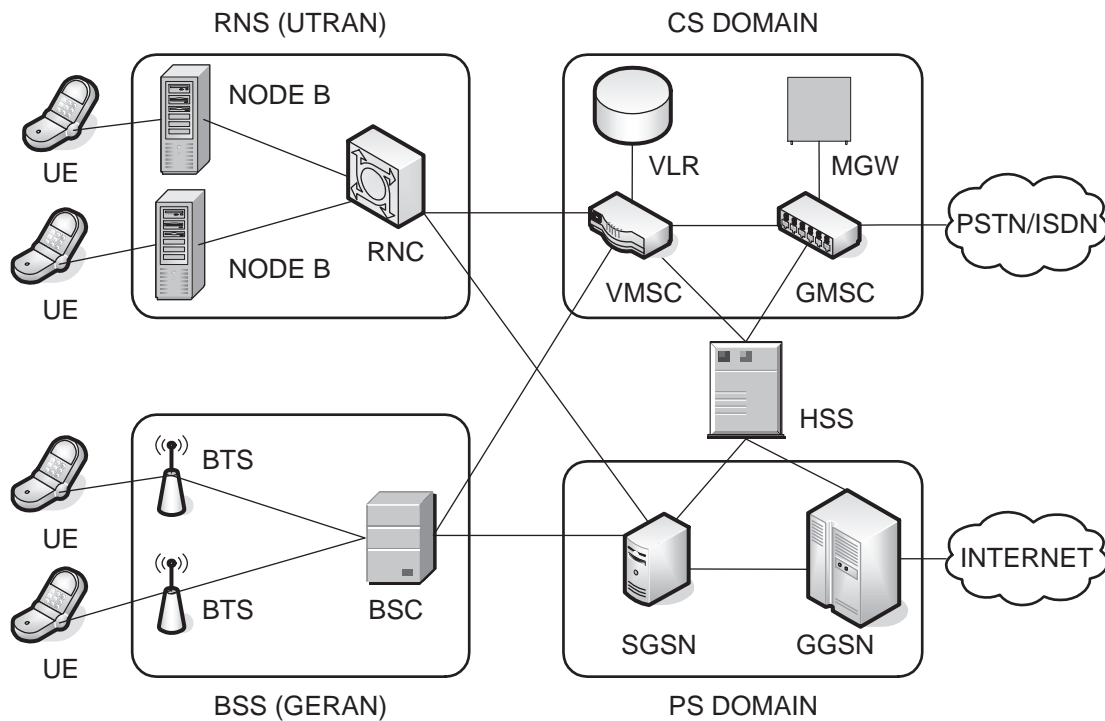
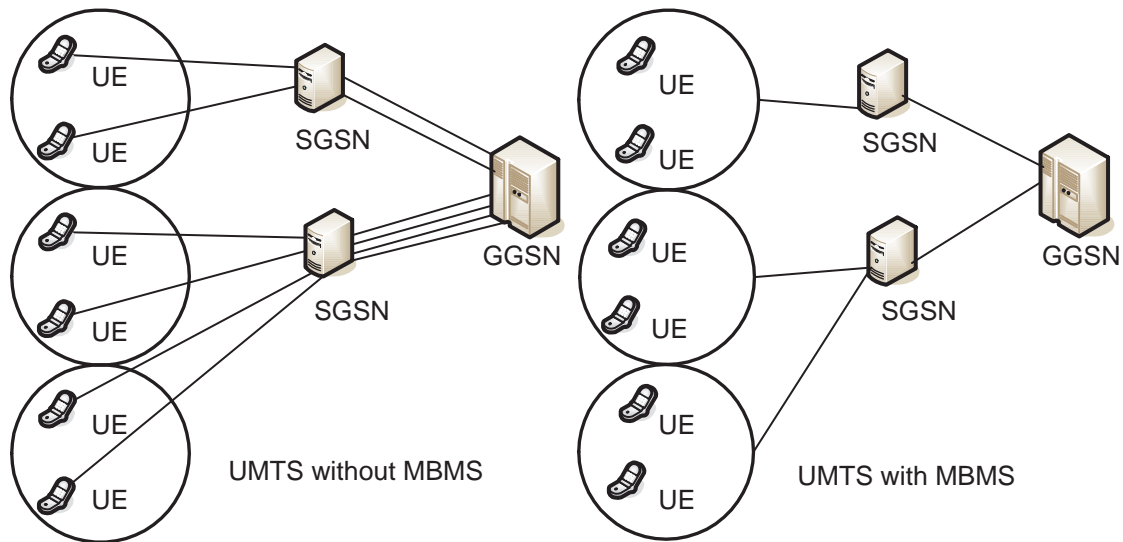Fig. 1.   Core network and radio access network options.

Fig. 2.   IP multicast with and without MBMS.

services; for example, a television user service may consist of separate audio and video bearer services. We will concentrate below on bearer services. Each service may target a specific set of cells, or *service area*. The actual transmission of data within a service is called a *session*. A service can only have one active session at a time, but it may use multiple sessions throughout its lifetime, transmitting the same or different content. The service types targeted by MBMS are classified as follows [5]:

- *Streaming:* continuous media such as audio and video, plus supplementary text and images, similar to TV chan-

nels but enhanced with multimedia content.
- *Download:* reliable binary data transfers without stringent delay constraints, similar to conventional file transfers but with multiple receivers.
- *Carousel:* a combination of streaming and download, combining static media with synchronization constraints, similar to stock quote ticker tapes.

To maximize the number of terminals that are able to receive MBMS services, the 3GPP recommends the *Real-time Transport Protocol* (RTP) for use with streaming services, and the

File Delivery over Unidirectional Transport (FLUTE) protocol for use with download services. The 3GPP has defined MBMS specific profiles for these protocols, as well as procedures for reception reporting and file repairs [6].

In order to support MBMS, a new functional entity, the *Broadcast/Multicast Service Center* (BM-SC) must be added to the PS domain of the CN, as shown in Figure 3, mediating between the content providers and the UMTS network. The BM-SC has the overall responsibility for both the control and user planes of an MBMS service. MBMS services may employ one of two modes: the *broadcast mode* sends data to all users in the service area, while the *multicast mode* sends data only to users in the service area that have explicitly requested a particular service. Data transfer for both multicast and broadcast mode services is strictly unidirectional, from the BM-SC towards the UEs [7].

The life cycle of a broadcast session is shown on the left side of Figure 4. In the *service announcement* phase the users are informed about the service's availability in their area. The BM-SC may transmit service announcements using an MBMS service of the download type, or interested UEs may directly query the BM-SC to discover what services are offered; other options are the use of CBS, SMS or web page listings. In all cases, the *Session Description Protocol* (SDP) is used to describe the contents of a session, such as media types and bandwidth requirements. Users may choose to either activate or deactivate reception for each broadcast service locally, that is, in their UE.

When data are about to be delivered, the BM-SC initiates the *session start* phase, during which the appropriate UMTS network nodes are instructed to establish bearers. In broadcast mode, bearers are established towards all cells in the service area. In the *MBMS notification* phase the users are notified that a data transmission is about to begin so that they may start listening at the appropriate broadcast radio channel. In the *data transfer* phase the actual data are transmitted to all cells in the service area and are received by those UEs that have not locally deactivated the service. Finally, when the data transfer is finished, the BM-SC initiates the *session stop* phase during which the established bearers are released.

The life cycle of a multicast session, shown on the right side of Figure 4, adds three more phases which are performed separately by each UE desiring to receive a multicast session; in contrast, the phases presented above are performed on a per session basis. In the *subscription* phase, interested users subscribe to multicast services via external means, for example, via a web page. This informs the BM-SC that the user has agreed to receive, and possible be charged for, a multicast service. The service announcement phase is similar as in broadcast mode, but it also includes the multicast address to be used for the session. Either before or after session start, users interested in the session initiate the *joining* phase, that is, they send a message to the GGSN indicating the multicast group that they wish to join. These messages are screened by the BM-SC which verifies whether the user has indeed subscribed to the service.

In multicast mode, during session start, bearers are created only towards cells containing UEs that have joined the multicast group. During MBMS notification the RAN decides whether a single *Point-to-Multipoint* (PtM) or multiple individual *Point-to-Point* (PtP) bearers must be established in each cell and notifies the UEs accordingly; we discuss this in Section IV. During data transfer the actual data are transmitted to all UEs that have joined the service and during session stop, the bearers are released. Either before or after session stop, users that are no longer interested in receiving data initiate the *leaving* phase, that is, they send a message to the GGSN, indicating the multicast group that they wish to leave.

Each MBMS session supports a specific set of *Quality of Service* (QoS) parameters; these are common for the entire distribution tree so as to simplify tree creation and maintenance. The 3GPP defined QoS model for UMTS networks [8] is used for MBMS, with one limitation: only the (unidirectional) *streaming* and *background* traffic classes can be used for MBMS services depending on whether they are delay sensitive or not, respectively. The (bidirectional) *conversational* and *interactive* traffic classes are obviously unsuitable for MBMS. It is likely, but not mandatory that download and carousel services will use the background class, and media streaming services will use the streaming class.

For broadcast services, MBMS supports charging the content provider based on various parameters, including service transmission time and duration, data transferred and number of cells reached. Receivers cannot be charged directly for the reception of broadcast services, since the network does not know whether a UE has activated them. This sender pays charging model is suitable for advertising supported applications. Multicast services, on the other hand, support charging for both sending and receiving data as the network is aware of when each UE joins or leaves a service. This receiver and/or sender pays model is therefore also suitable for pay per view applications. When a service is encrypted, the receivers may also be charged for the decryption keys (see Section V), for both multicast and broadcast services.

IV. MBMS IMPLEMENTATION

Besides the addition of a new node in the UMTS network, MBMS also requires modifications to existing nodes. The BM-SC is the new node; it is responsible for authenticating and authorizing the content providers, receiving and possibly modifying their data, for example, by encrypting them and passing these data to the GGSN for transmission. The BM-SC may repeat an entire data transmission or specific parts of it for error recovery purposes. The BM-SC also provides information about its services for both service announcement and bearer setup purposes, screens the UEs wishing to join a multicast session to verify that they have subscribed to the service, and initiates the session start and stop phases.

The GGSN is responsible for transmitting data via tunnels to the appropriate SGSNs, that is, all SGSNs in the case of broadcast services or only those SGSNs serving group members for multicast services. The SGSN sends these data via tunnels to the appropriate RNCs or BSCs, depending on whether the UTRAN or the GERAN is used. The SGSN is also responsible for instructing the RAN to establish and release
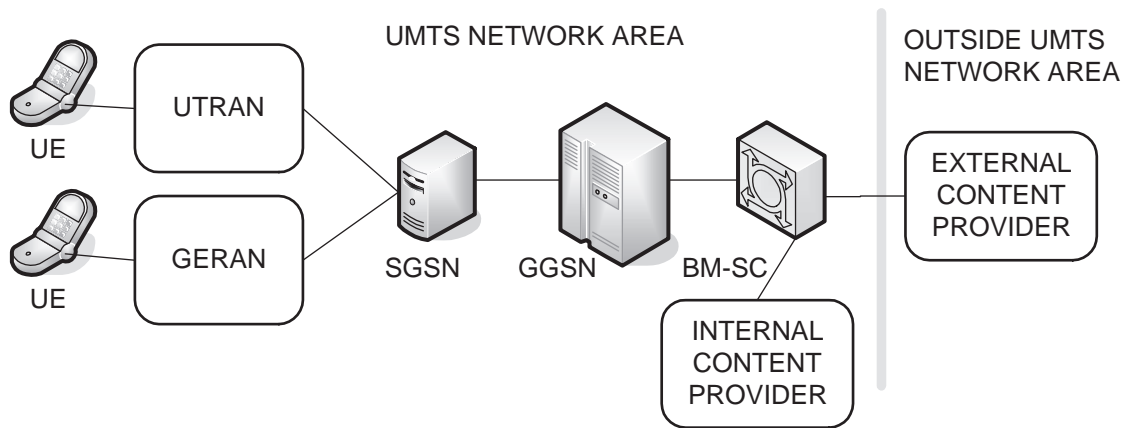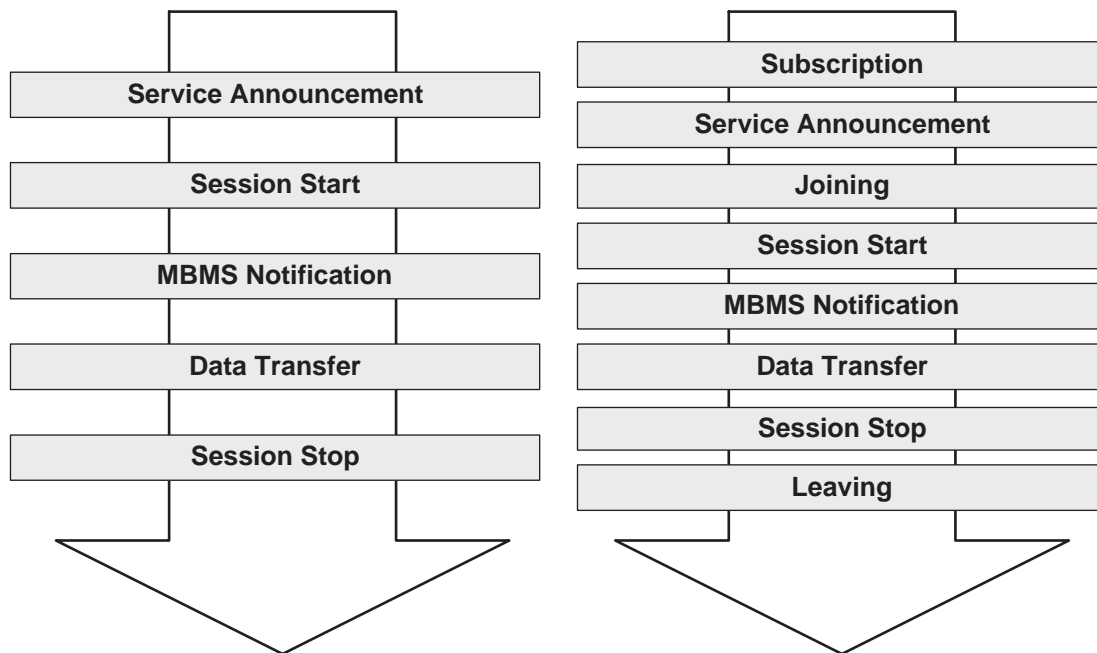
Fig. 3.   MBMS network architecture.

Fig. 4.   Broadcast and multicast mode phases.

actual radio bearers in the session start and stop phases, respectively. Finally, the UE is responsible for joining and leaving multicast services, a procedure that requires interaction with the UMTS network, and for enabling or disabling broadcast services, a purely local procedure. Since the UE consumes the data, it usually has additional application specific tasks, such as media playback.

In order for MBMS services to operate, a set of *signaling* procedures are defined as extensions to the existing control plane UMTS protocols [9]. These procedures can be divided into *user specific* and *session specific* ones. We will briefly outline below these procedures as they apply to the more complicated multicast mode. The first user-specific procedure is *MBMS Activation*, triggered when a UE sends an IGMP (for IPv4) or MLD (for IPv6) Join message indicating the IP address of a multicast group, as shown in Figure 5. While

this is *not* the normal mode of IGMP/MLD operation, it is compatible with its use for IP multicast [10]. The GGSN asks the BM-SC to verify whether the user has subscribed to the service; if so, the BM-SC, via the GGSN and the SGSN, directs the UE to the GGSN serving as the source for this service, which may be different than the one the UE originally communicated with.

The second user-specific procedure is *MBMS Context Creation*, initiated by the UE when the MBMS Activation completes (first dashed line). This procedure causes the CN nodes, that is, the BM-SC, GGSN, and SGSN to create a *MBMS UE Context* (MUEC) to store the information pertaining to a UE for a specific MBMS service. When the first MUEC for a service is created at a GGSN or SGSN, they initiate the first session-specific procedure, the *MBMS Registration*. In this procedure (dotted lines), the GGSN (SGSN) asks its

| UE | RAN | SGSN | GGSN | BM-SC |

IGMP Join

Authorization Request

Authorization Response

Notification Request

Notification Response

Request Context Activation

Activate Context Request

Create Context Request

Authorization Request

Authorization Response

Registration Request

Registration Response

Create Context Response

Registration Request

Registration Response

Provision of MBMS UE Context to RAN

Activate Context Accept

Session Start Request

Session Start Response

Session Start Request

Session Start Response

Session Start Request

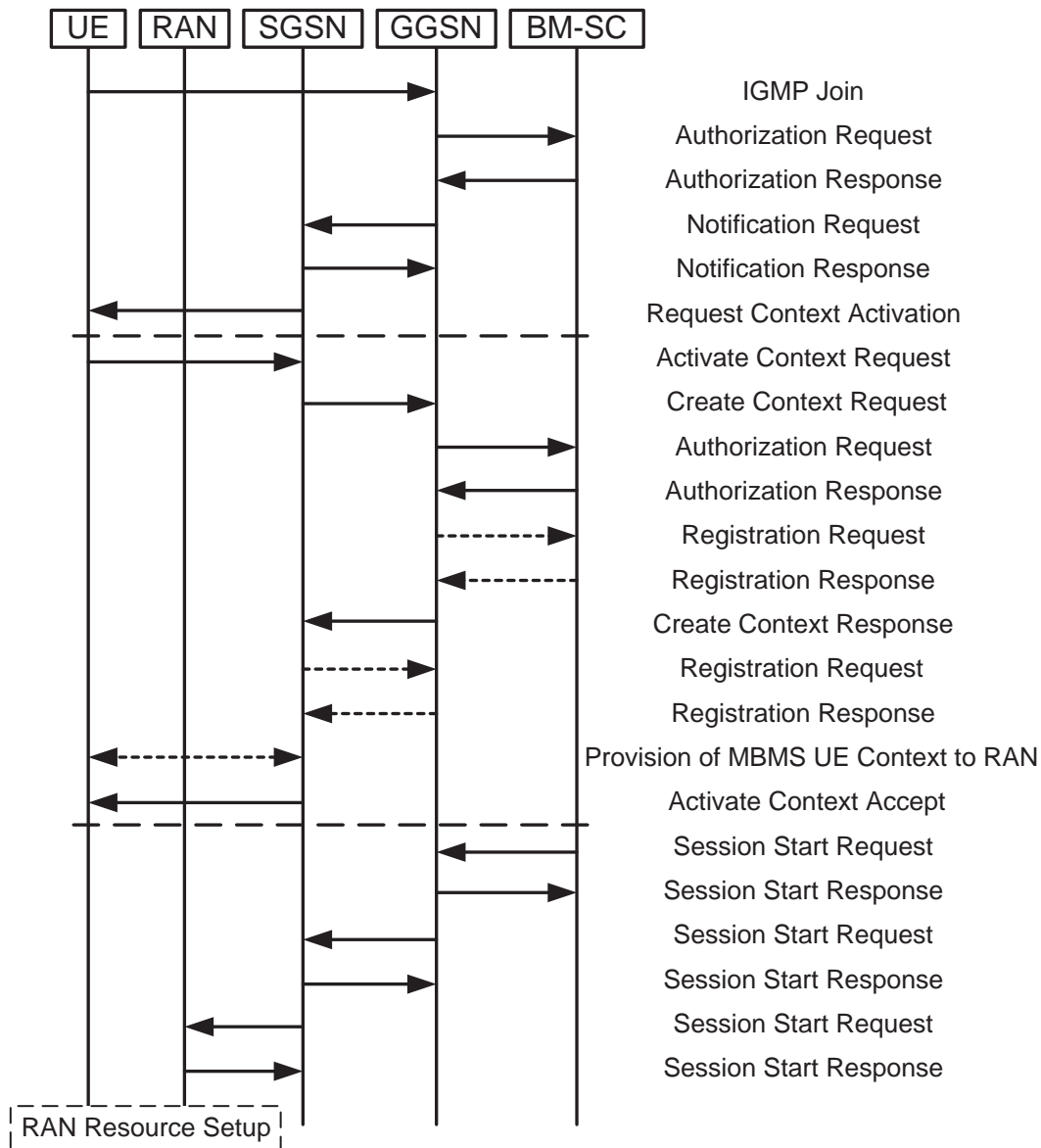Session Start Response

RAN Resource Setup

Fig. 5.   MBMS signaling procedures.

parent BM-SC (GGSN) to send it information about an MBMS service. The child node stores this information in an *MBMS Bearer Context* (MBC) for the group, while its parent notes in its own MBC that this child has asked to receive data for the service. Using these procedures, each CN node knows which of the UEs that it serves are interested in a service and which of its children nodes should receive the data transmitted to this service. However, no bearers are established yet. When data transmission is about to begin (second dashed line), the second session-specific procedure, the *MBMS Session Start*, causes each CN node to establish bearers towards each of its children that has registered for this service.

The remaining signaling procedures (not shown) reverse the results of the ones above. The *MBMS Session Stop* procedure causes bearers to be released when the session ends. The *MBMS Deactivation* procedure is triggered by a UE sending an IGMP/MLD leave message to the GGSN whenever it

desires to stop receiving data for a service. The *MBMS Context Deletion* procedure that follows it leads to the deletion of the MUEC for this UE from the CN nodes serving the UE. Finally, when a MUEC is deleted, if a CN node does not serve any more UEs, it initiates the *MBMS Deregistration* procedure to indicate to its parent that it is no longer interested in this service, and then deletes the corresponding MBC.

We have so far concentrated on CN signaling procedures, as the RAN procedures are rather involved and depend on whether the GERAN or the UTRAN is used. We will now outline the most interesting aspects of UTRAN-specific MBMS signaling. We must first point out that the UTRAN does *not* track the cell where each and every UE resides, as this would require excessive amounts of signaling. While every UE always notifies the CN when it moves to a new routing area, that is, a set of cells served by a different SGSN, only *PS connected* UEs, that is, UEs that have established PS signaling

connections with the UMTS network in order to exchange data with it, inform the RAN as they move from cell to cell. When a PS connected UE enters the area served by an RNC or when it joins a group, the *MBMS UE Linking* procedure is performed, whereby the SGSN serving the UE sends to the RNC information about the groups that the UE is a member of. When the UE exits the area served by an RNC or when it leaves a group, the *MBMS UE De-Linking* procedure is performed to remove this information from the RNC.

Therefore, when an RNC is notified about a session start, it does not know how many UEs are interested in this service in each cell; it is only aware of PS connected UEs that were *linked*. This information is critical, however, due to the power control used in UTRAN: PtP bearers require only small amounts of transmission power for UEs close to the center of the cell, while PtM bearers always require large amounts of power so as to reach UEs at the edge of the cell, so for a few UEs it may be better to use individual PtP bearers rather than a common PtM one. Thus, the RNC must estimate the number of UEs interested in a service in each cell, even if they are not PS connected in order to make a choice. Rather than requesting *all* UEs to become PS connected, the RNC can ask all UEs in a cell to establish a PS signaling connection to the network with a specific probability. Depending on the number of UEs connecting, the RNC estimates the total number of UEs interested in the service. For example, if three UEs connect with an announced probability of 5%, it may be estimated that 60 UEs are interested. The procedure may be repeated with increasing probabilities in order for the RNC to increase its confidence before deciding whether to establish a PtP or a PtM bearer. This mechanism is commonly called *UE Counting* [11].

V. MBMS SECURITY

Prior to the introduction of MBMS, all UMTS security procedures were performed separately between each UE and the network. After mutual authentication, a shared secret was established between the UE and the network; this secret was then used to ensure the confidentiality and integrity of transmitted data by encrypting and digitally signing them, respectively. With MBMS, however, the same data may be delivered to many UEs over PtM bearers, therefore, confidentiality and integrity have to be achieved on a one to many basis. This means that the data source must share the same secret with many UEs. To prevent a UE from using this secret after leaving a service, the shared secret must be periodically updated.

Fortunately, the UMTS security architecture provides support for application level security by allowing a UE and an application server to establish a shared secret for use by any protocol they desire. For MBMS, the BM-SC is the entity responsible for generating and distributing shared keys to the UEs participating in a service [12]. Since the BM-SC is also responsible for content distribution, it can use the appropriate keys to encrypt and/or digitally sign the content before transmission. As a result, integrity and confidentiality are achieved end-to-end between the BM-SC and the UEs, transparently to the UMTS network.
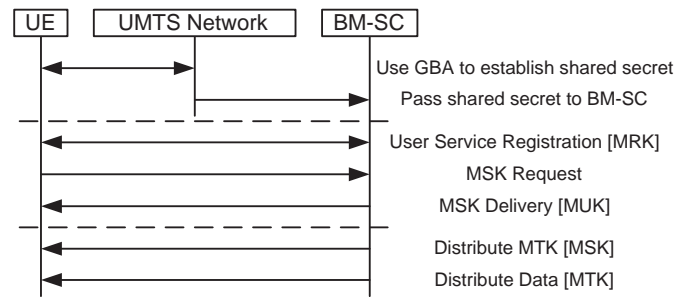

Fig. 6.   MBMS security procedures.

If a service employs MBMS security, this is indicated in its service announcement. As shown in Figure 6, a UE that desires to receive such a service must first use the *Generic Bootstrapping Architecture* (GBA) [13] to establish a shared secret with the UMTS network for application specific purposes; this secret is then passed by the network to the BM-SC. At this point (first dashed line) the shared secret is used by both the UE and the BM-SC to derive two UE specific keys, the *MBMS Request Key* (MRK) and the *MBMS User Key* (MUK). The UE then initiates the *User Service Registration* procedure with the BM-SC, during which the two parties authenticate each other by using the MRK. If the UE has subscribed to the service, it is registered by the BM-SC as a key recipient. The UE then performs the *MSK Request* procedure, asking the BM-SC for the key of a specific MBMS service. The BM-SC sends this *MBMS Service Key* (MSK) to the UE with the *MSK Delivery* procedure, using the MUK to protect it (second dashed line). The BM-SC may periodically send a new MSK to the UE so as to invalidate older keys.

The actual content is protected by a service specific *MBMS Traffic Key* (MTK). This key is distributed as a part of the content itself, using the MSK to protect it; therefore, the MTK is transmitted over either a PtP or a PtM bearer to all UEs, unlike previous keys that were sent over PtP bearers to individual UEs. The MTK may also be periodically refreshed so as to invalidate older keys. Note that while the MRK and MUK are UE specific, the MSK and MTK are service specific and common to all UEs receiving a service; after the MSK is separately received by each UE protected by a MUK, it is used to recover the MTK received by all UEs. Each UE then uses the common MTK to decrypt the protected content.

VI. MBMS AND IP MULTICAST

While IP multicast has been available for more than a decade, it has not taken off. There are many reasons for this, which we will now review in order to assess whether MBMS also suffers from them. A technical shortcoming is that IP multicast routing does not scale well. Unlike unicast addresses that can be aggregated for each network area, allowing a router to use a single entry for many nearby networks, multicast addresses are logical, so nearly identical addresses may be referring to completely different member sets. As a result, each router on the distribution path of an IP multicast group must allocate memory and perform separate

signaling procedures for each group [14]. Another issue is that there is no mechanism for global multicast address allocation, therefore, applications may face address collisions.

The most important issue, however, is a business rather than a technical one. The main selling point of multicast to network providers is the cost savings when distributing multimedia sessions. IP multicast groups are *open* though, that is, anyone can receive data sent to a group without authorization from the sender and anyone can send data to a group without authorization from the receivers. Content providers are obviously reluctant to use a service in which there is no guarantee that the receivers will compensate them, while receivers are reluctant to pay for a service that may be abused by random senders [15].

MBMS services are defined with respect to a particular UMTS network, therefore, their scope is limited to that network. As a result, while MBMS may seem to be no better than IP multicast in terms of routing scalability, as it also requires a separate entry per multicast group, since MBMS groups are limited to a single UMTS network, the number of groups per network should be smaller. In addition, while IP multicast routing entries are needed for each incoming interface to a router and, for some routing protocols, different entries are required for different sources, with MBMS only a single incoming interface exists, from the node's parent, and only a single source exists, the BM-SC, therefore, only a single entry per group is needed at each node. Finally, MBMS avoids address collisions as each group is defined with respect to a specific UMTS network and all addresses are allocated by the BM-SC.

More importantly, MBMS groups are *closed*, that is, UEs must subscribe to a service before being allowed to join it and only the BM-SC is allowed to send data to a group. Receivers must be authorized before joining a service and the network may charge them in a very flexible manner. The network also controls who sends to a service, thus preventing third parties from sending data to a multicast service for which the UEs are paying. Therefore, an attractive business model is offered to both content providers and users. And while MBMS is clearly less flexible than IP multicast, it was never intended as a substitute for IP multicast. For example, a multiparty conferencing application is unsuitable for MBMS; MBMS is best suited to content distribution from a single, operator controlled, source to large groups of receivers.

VII. MBMS AND DVB-H

The predicted expansion of the mobile multimedia services market, especially for TV distribution to mobile handheld devices, has resulted in the creation of a number of solutions for broadcast and multicast content distribution. Some have a mobile cellular background, including MBMS and its counterparts from 3GPP2, while others have a broadcasting background. Among the latter, the most likely competitor to MBMS is DVB-H, an extension of the DVB-T system used for terrestrial TV broadcasting in Europe. DVB-H, an ETSI specification, provides additional features over DVB-T to support reception for users moving at high speeds,

power savings for battery constrained terminals, single antenna reception for compact devices, and seamless handover in single frequency networks [16]. As DVB-H transmits IP packets, encapsulated inside the MPEG-2 transport stream of DVB-T, it is *not* limited to TV.

DVB-H can be considered as both a competing and as a complementary technology to MBMS. Both systems support mobile handheld devices, offering IP based services that include, but are not limited to, mobile TV. DVB-H not only supports handovers, it can also operate within the 5 MHz bands traditionally used for cellular networks. DVB-H also provides higher capacity than MBMS: DVB-H can support up to 30 video streams at a resolution of 352×288 using MPEG-4 in an 8 MHz channel, while MBMS can only support up to 15 video streams at a reduced resolution of 176×144 in a 5 MHz channel. Therefore, DVB-H seems to be a better option for transmitting TV channels to mobile handheld devices.

On the other hand, MBMS is coupled with uplink services, unlike DVB-H which is a pure downlink service. MBMS can use the uplink to provide interactive facilities without an additional transmitter, as well as to request repairs for damaged data streams, something quite useful for software downloads. Maybe more importantly, the MBMS multicast mode can be used to transmit different data streams to each cell, based on user demands. For example, MBMS may offer hundreds of video streams in an area, even though it can only transmit a small number of them in each cell. Therefore, MBMS seems to be a more appropriate option than DVB-H for transmitting specialized content to smaller user communities.

VIII. MBMS CHALLENGES

While MBMS successfully addresses the shortcomings of IP multicast (see Section VI) and has a potential market even when competing with DVB-H (see Section VII), in order to succeed, it must overcome additional technical and business challenges. As MBMS introduces considerable complexity into a UMTS network (see Section IV), it must provide sufficient benefits to outweigh its costs. Since the most critical resource in a UMTS network is transmission power, MBMS services should ideally use a single PtM bearer to serve a large number of users at a fixed cost. As PtM bearers must serve UEs even at the edges of a cell, however, they significantly reduce the power available for other UMTS services.

In order to reduce the power requirements of PtM bearers, MBMS can take advantage of *macro-diversity* techniques, where a Node-B transmits on PtM bearers with less power than what is required to reliably reach the edge of the cell, assuming that UEs in this area will combine the transmissions from neighboring Node-Bs in order to successfully reconstruct the transmitted data. On the network side, this requires sufficient time synchronization of identical MBMS transmissions in different cells to allow combining to occur; on the UE side, this requires the capability to receive and decode the same content from multiple transmitters simultaneously, either by *selective combining*, that is, independently decoding the bits from each transmission, or by *soft combining*, that is, jointly decoding the bits from each transmission. Macro-diversity techniques

show great promise for reducing the power requirements of PtM bearers [17].

In order to determine, however, whether MBMS is worth its complexity, the main consideration is a business one: how popular will MBMS services be? For services that are not popular, PtP bearers will be used, therefore, rather than saving transmission power, the complicated MBMS signaling will actually lead to more overhead than with a simple unicast service. Even if a service is popular enough to warrant the use of a PtM bearer, for macro-diversity to be employed many cells will need to transmit the same content in PtM mode. As a result, only quite popular MBMS services will be able to amortize the cost of PtM bearers and MBMS signaling among many receiving UEs. One could indeed argue that the real challenge for MBMS is to offer services that are popular enough to provide transmission power savings, but inappropriate for DVB-H, either due to limited appeal or to an inherent need for a return channel. One such example is sending software updates to mobile devices: only specific terminals will be interested in any given update, and a return channel will be needed to repair damaged data blocks.

IX. SUMMARY

This article provided an overview of the 3GPP Multimedia Broadcast/Multicast Service. We first outlined the basic features of UMTS networks and then described the overall architecture and services provided by MBMS. We presented the implementation details of MBMS in terms of the additional node functionalities and signaling procedures required and then discussed the security architecture of MBMS. Finally, we discussed the issue of whether MBMS can succeed where IP multicast failed, compared MBMS with DVB-H in order to determine whether there is a market for both, and outlined the remaining technical and business challenges for MBMS.

ACKNOWLEDGEMENT

REFERENCES

[1] Zeng M, Annamalai A, Bhargava VK. Harmonization of global third-generation mobile systems. *IEEE Communications* 2000; **38**(12): 94–104.

[2] 3GPP. General Packet Radio Service (GPRS); Service description; Stage 2. *TS 23.060; V6.10.0* 2005.

[3] 3GPP. Evolution of 3GPP System. *TR 21.902; V6.0.0* 2003.

[4] 3GPP. Technical realization of Cell Broadcast Service (CBS). *TS 23.041; V6.2.0* 2003.

[5] 3GPP. Multimedia Broadcast/Multicast Service (MBMS) user services; Stage 1. *TS 22.246; V6.2.0* 2004.

[6] 3GPP. Multimedia Broadcast/Multicast Service (MBMS); Protocols and codecs. *TS 26.346; V6.2.0* 2005.

[7] 3GPP. Multimedia Broadcast/Multicast Service (MBMS); Stage 1. *TS 22.146; V6.6.0* 2004.

[8] Koodli R, Puuskari M. Supporting packet-data QoS in next-generation cellular networks. *IEEE Communications* 2001; **39**(2): 180-188.

[9] 3GPP. Multimedia Broadcast/Multicast Service (MBMS) user services; Architecture and functional description. *TS 23.246; V6.8.0* 2005.

[10] Xylomenos G. Group Management for the Multicast Broadcast/Multicast Service. In *Proc. of the 14th IST Mobile Summit*, 2005.

[11] 3GPP. Introduction of the Multimedia Broadcast Multicast Service (MBMS) in the Radio Access Network (RAN); Stage 2. *TS 25.346; V6.6.0* 2005.

[12] 3GPP. Generic Authentication Architecture (GAA); Generic Bootstrapping Architecture. *TS 33.220; V6.6.0* 2005.

[13] 3GPP. Security of Multimedia Broadcast/Multicast Service. *TS 33.246; V6.4.0* 2005.

[14] Zhang B, Mouftah HT. Forwarding state scalability for multicast provisioning in IP networks. *IEEE Communications* 2003; **41**(6): 46-51.

[15] Diot C, Levine BN, Lyles B, Kassem H, Balensiefen D. Deployment issues for the IP multicast service and architecture. *IEEE Network* 2000; **14**(1): 78-88.

[16] Faria G, Henriksson JA, Stare E, Talmola P. DVB-H: Digital broadcast services to handheld devices. *Proceedings of the IEEE* 2006; **94**(1): 194-209.

[17] Parkvall S, Englund E, Lundevall M, Torsner J. Evolving 3G mobile systems: broadband and broadcast services in WCDMA. *IEEE Communications* 2006; **44**(2): 68-74.

AUTHORS' BIOGRAPHIES

George Xylomenos is a member of the Mobile Multimedia Laboratory at the Athens University of Economics and Business, where he is also a faculty member in the Department of Informatics. He received his Diploma in Informatics from the Athens University of Economics and Business, Greece, and M.S. and Ph.D. degrees in Computer Science from the Department of Computer Science and Engineering at the University of California, San Diego. His current research focuses on improving the performance of Internet applications over wireless networks, designing link layer protocols for wireless links, supporting Quality of Service guarantees over wireless and mobile networks, and providing broadband services over next generation cellular and fixed wireless networks. His E-mail ID is xgeorge@aueb.gr.

Vasilis Vogkas is a member of the Mobile Multimedia Laboratory at the Athens University of Economics and Business in the Department of Informatics, where he is pursuing Ph.D. in the area of Multicast Distribution Protocols for Next Generation Wireless Networks. He received B.Eng. in Electrical and Electronic Engineering from the University of Glamorgan and M.Sc. in Computer Science (Networks and Distributed Systems) from Trinity College Dublin. His current research focuses on emerging technologies for multicasting and broadcasting of multimedia content over next generation fixed and mobile wireless telecommunication networks. His E-mail ID is vvogkas@aueb.gr.

George Thanos is a member of the Mobile Multimedia Laboratory at the Athens University of Economics and Business in the department of Informatics, where he is pursuing Ph.D. in the area of Interoperability between Next Generation Heterogeneous Networks. He received B.Eng. in Computer Systems Engineering from the University of Manchester (UMIST) and M.Sc. in Mobile and Satellite Communications from the University of Surrey. He was previously involved with industrial research and development of telecommunications systems for UMTS/3G networks. His current research focuses on investigating interoperability issues between different wireless multicasting and broadcasting systems and in the security of next generation pervasive/ubiquitous environments. His E-mail ID is gthanos@aueb.gr.