WIRELESS COMMUNICATIONS AND MOBILE COMPUTING
Wirel. Commun. Mob. Comput. 2008; **8**:463–481
DOI: 10.1002/wcm.464

Multicast in UMTS: Evaluation and recommendations

Antonios Alexiou and Christos Bouras*,†
Computer Engineering and Informatics Department, University of Patras, Research Academic Computer Technology Institute, Patras, Greece

Summary

It is known that multicasting is an efficient method of supporting group communication as it allows the transmission of packets to multiple destinations using fewer network resources. Along with the widespread deployment of the third generation cellular networks, the fast-improving capabilities of the mobile devices, content and service providers are increasingly interested in supporting multicast communications over wireless networks and in particular over Universal Mobile Telecommunications System (UMTS). Multicasting is a more efficient method of supporting group communication than unicasting or broadcasting, as it allows transmission and routing of packets to multiple destinations using fewer network resources. In this paper, the three above mentioned methods of supporting group communication are analyzed in terms of their performance. The critical parameters of primary interest for the evaluation of any method are the packet delivery cost and the scalability of the method. Copyright © 2006 John Wiley & Sons, Ltd.

KEY WORDS: UMTS; multicast in UMTS; MBMS; multiple unicast; broadcast

1. Introduction

Universal Mobile Telecommunications System (UMTS) constitutes the third generation of cellular wireless networks which aims to provide high-speed data access along with real time voice calls. Wireless data is one of the major boosters of wireless communications and one of the main motivations of the next generation standards [1].

Multicast communications for wireline users has been deployed in the Internet for at least the past 10 years. The multicast transmission of real time multimedia data is an important component of many current and future emerging Internet applications, such as videoconference, distance learning, and video distribution. The multicast mechanism offers efficient multidestination delivery, since data is transmitted in an optimal manner with minimal packet duplication [2].

Although UMTS networks offer high capacity, the expected demand will certainly overcome the available resources. Thus, the multicast transmission over the UMTS networks constitutes a challenge and an area of research. Actually, the adoption of multicast routing over mobile networks poses a different set of challenges in comparison with multicasting over the Internet. First of all, multicast receivers are non-stationary, and consequently, they may change their access point at any time. Second, mobile networks are generally based on a well-defined tree topology with the non-stationary multicast receivers being located at the leaves of the network tree.

---

*Correspondence to: Christos Bouras, Computer Engineering and Informatics Department, University of Patras, Research Academic Computer Technology Institute, N. Kazantzaki str, GR 26500, Patras, Greece.
†E-mail: bouras@cti.gr

The construction of a source-rooted shortest-path tree over such a topology is trivial and may be achieved by transmitting only a single packet over the paths that are shared by several multicast recipients. However, as a result of user mobility, there are several cases where this simplified view of the mobile network is violated [3]. It is, therefore, not appropriate to apply IP multicast routing mechanisms in UMTS, since they are not designed to take into account the need for mobility management that mobile networks require.

In this paper, we present an overview of three different one-to-many packet delivery schemes for UMTS. These schemes include the Broadcast, the Multiple Unicast, and the Multicast scheme. We analytically present these schemes and analyze their performance in terms of the packet delivery cost and the scalability of each scheme. Furthermore, for the evaluation of the schemes, we consider different transport channels for the transmission of the data over the Iub and Uu interfaces. Since the performance of these schemes depends mainly on the configuration of the UMTS network that is under investigation, we consider different network topologies and user distributions.

This paper is structured as follows. In Section 2, we provide an overview of the UMTS in packet switched domain. Section 3 is dedicated to describing the related work, while Section 4 presents the Multimedia Broadcast/Multicast Service framework of UMTS. In Section 5, we present a number of alternative one-to-many packet delivery schemes for UMTS. Following this, Section 6 analyzes the different delivery schemes in terms of telecommunication costs, while Section 7 presents some numerical results that characterize the above schemes. Finally, some concluding remarks and planned next steps are briefly described.

2. Overview of the UMTS in the Packet Switched Domain

UMTS network is split in two main domains: the User Equipment (UE) domain and the Public Land Mobile Network (PLMN) domain. The UE domain consists of the equipment employed by the user to access the UMTS services. The PLMN domain consists of two land-based infrastructures: the Core Network (CN) and the UMTS Terrestrial Radio-Access Network (UTRAN) (Figure 1). The CN is responsible for switching/routing voice and data connections, while the UTRAN handles all radio-related functionalities. The CN is logically divided into two service domains: the Circuit-Switched (CS) service domain and the Packet-Switched (PS) service domain [1,4]. The PS portion of the CN in UMTS consists of two kinds of General Packet Radio Service (GPRS) Support Nodes (GSNs), namely Gateway GSN (GGSN) and Serving GSN (SGSN) (Figure 1). SGSN is the centerpiece of the PS domain. It provides routing functionality, interacts with databases (like Home Location Register (HLR)) and manages many Radio Network Controllers (RNCs). SGSN is connected to GGSN via the Gn interface and to RNCs via the Iu interface. GGSN provides the interconnection of UMTS network (through the Broadcast Multicast–Service Center) with other Packet Data Networks (PDNs) like the Internet [1].

UTRAN consists of two kinds of nodes: the first is the RNC and the second is the Node B. Node B constitutes the base station and provides radio coverage to one or more cells (Figure 1). Node B is connected to the UE via the Uu interface (based on the Wideband Code Division Multiple Access, W-CDMA technology) and

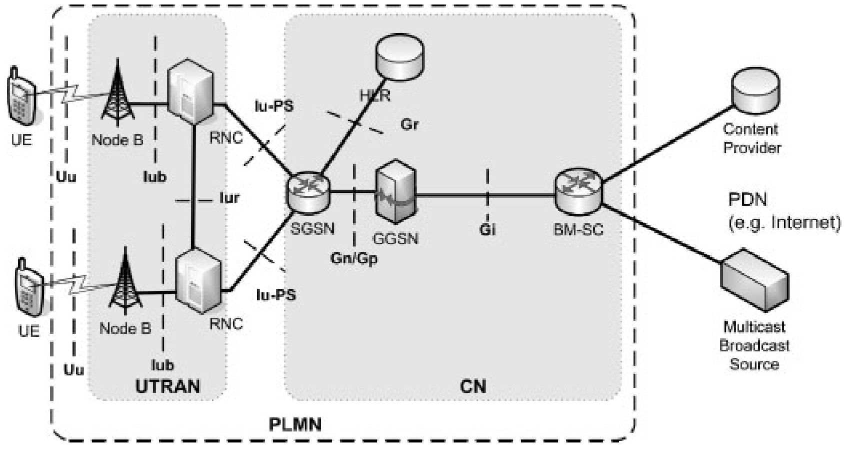


Fig. 1. UMTS and MBMS Architecture.

to the RNC via the Iub interface. One RNC with all the connected to it Node Bs is called Radio Network Subsystem (RNS).

Before a UE can exchange data with an external PDN, it must firstly establish a virtual connection with this PDN. Once the UE is known to the network, packets are transferred between it and the network, based on the Packet Data Protocol (PDP), the network-layer protocol carried by UMTS. An instance of a PDP type is called a PDP context and contains all the parameters describing the characteristics of the connection to an external network by means of end-point addresses and Quality of Service (QoS). A PDP context is established for all application traffic sourced from and destined for one IP address. A PDP context activation is a request–reply procedure between a UE and the GGSN. A successful context activation leads to the creation of two GPRS Tunneling Protocol (GTP) sessions, specific to the subscriber: between the GGSN and SGSN over the Gn interface and between the SGSN and RNC over the Iu interface. IP packets destined for an application using a particular PDP context are augmented with UE- and PDP-specific fields and are tunneled using GTP to the appropriate SGSN. The SGSN recovers the IP packets, queries the appropriate PDP context based on the UE- and PDP-specific fields, and forwards the packets to the appropriate RNC. The RNC maintains Radio-Access Bearer (RAB) contexts. Equivalently to PDP contexts, a RAB context allows the RNC to resolve the subscriber identity associated with a GTP-tunneled network packet data unit. The RNC recovers the GTP-tunneled packet and forwards the packet to the appropriate Node B. Finally, a Tunnel Endpoint IDentifier is used across the Gn and Iu interfaces to identify a tunnel endpoint at the receiving network node [4].

In the UMTS PS domain, the cells are grouped into Routing Areas (RAs), while the cells in a RA are further grouped into UTRAN Registration Areas (URAs). The Mobility-Management activities for a UE are characterized by two finite state machines: the Mobility Management (MM) and the Radio Resource Control (RRC). The Packet MM (PMM) state machine for the UMTS PS domain is executed between the SGSN and the UE for CN-level tracking, while the RRC state machine is executed between the UTRAN and the UE for UTRAN-level tracking. After the UE is attached to the PS service domain, the PMM state machine is in one of the two states: PMM idle and PMM connected. In the RRC state machine, there are three states: RRC idle mode, RRC cell-connected mode, and RRC URA connected mode [5].

3. Related Work

Several multicast mechanisms for UMTS have been proposed in the literature. In Reference [6], the authors discuss the use of commonly deployed IP multicast protocols in UMTS networks. Three potential Internet multicast architectures are analyzed. The first is the existing multicast architecture that is standardized as an optional feature in the UMTS networks. In this architecture, the IP multicast routing protocol is terminated in the gateway between the Internet and the UMTS network. This solution requires few multicast aware UMTS nodes. However, this architecture does not provide any bandwidth savings in the UMTS network. The two other designs are Internet multicast architectures where the multicast functionality is pushed successively further out towards the UMTS terminal. These two architectures require multicast awareness from an increased number of UMTS network nodes. Higher complexity is introduced to achieve network resource savings. The presented multicast mechanism employs the Internet Group Management Protocol (IGMP) for group management and relies on the standard hierarchical tunneling of UMTS for distributing multicast packets to the group. The hierarchical tunneling mechanism of UMTS, however, does not lend itself to efficient multicast packet delivery, since each tunnel may only be established for a single subscriber. Considering a group of N multicast users, a single multicast packet must be duplicated and transmitted N times throughout the network in order to reach all the destinations. Depending on the distribution of the multicast users within the coverage area, this may lead to an inefficient usage of resources within the network.

A solution to the above described problem is presented in Reference [3]. The authors, in order to overcome the one-to-one relationship between a single subscriber and a GPRS Tunneling Protocol (GTP) tunnel that is inherent to the hierarchical routing in UMTS, implement a Multicast-Packet Data Protocol (M-PDP) context for each multicast group in the GGSN and SGSN. In this approach, the authors do not adopt the use of IP multicast protocols for multicast routing in UMTS and present an alternative solution. For multicast group management, the authors propose the introduction of a number of new tables in GGSN, SGSN, and RNC, while for multicast packet forwarding some trivial changes in the GTP are required.

In Reference [7], a multicast mechanism for CS GSM networks is outlined that only sends multicast messages to Location Areas (LAs) in which multicast users reside. This mechanism uses the existed

DOI: 10.1002/wcm

UMTS/GSM short message architecture in order to perform multicast routing. In particular, two new tables are considered in the Home Location Register (HLR) and in the Visitor Location Register (VLR). The multicast table at the HLR records the Mobile Switching Centers (MSCs) that serve multicast users, while the VLR keeps track of the LAs that have multicast users. However, the multicast messages are delivered to all the cells of an LA, independently of whether or not multicast users are located in all cells. This is inefficient if an LA is large or only sparingly populated with multicast users.

Finally, the Multimedia Broadcast/Multicast Service (MBMS) framework of UMTS is currently being standardized by the third Generation Partnership Project (3GPP) [8]. MBMS relies on the definition of service areas for delivering multimedia traffic to subscribers. The MBMS framework is presented in detail in the following section.

4. Multimedia Broadcast/Multicast Service in UMTS

3GPP is currently standardizing the MBMS. Actually, the MBMS is an IP datacast type of service, which can be offered via existing GSM and UMTS cellular networks. It consists of an MBMS bearer service and a MBMS user service. The latter represents applications, which offer for example multimedia content to the users, while the MBMS bearer service provides methods for user authorization, charging and QoS improvement to prevent unauthorized reception [8].

The major modification in the existing GPRS platform is the addition of a new entity called Broadcast Multicast-Service Center (BM-SC). Figure 1 presents the architecture of the MBMS. The BM-SC communicates with the existing UMTS-GSM networks and the external PDNs [8].

As the term MBMS indicates, there are two types of service mode: the broadcast and the multicast mode. In broadcast mode, data is delivered to a specified area without knowing the receivers and whether there is any receiver at all in this area. However, in the multicast mode, the receivers have to signal their interest for the data reception to the network and then the network decides whether the user may receive the data or not.

Since the multicast mode is more complicated than the broadcast mode, it is more useful to present the operation of the MBMS multicast mode and the way that the mobile user receives the multicast data of a service. Actually, the reception of an MBMS multicast service is enabled by certain procedures. These are: Subscription, Service Announcement, Joining, Session Start, MBMS Notification, Data Transfer, Session Stop, and Leaving. The phases Subscription, Joining, and Leaving are performed individually per user. The other phases are performed for a service, that is, for all users interested in the related service. The sequence of the phases may be repeated, for example, depending on the need to transfer data. Also Subscription, Joining, Leaving, Service Announcement, as well as MBMS Notification may run in parallel to other phases.

5. One-to-Many Packet Delivery Schemes for UMTS

In this section, we present an overview of three different one-to-many packet delivery schemes for UMTS. These schemes include the multiple unicast scheme, the broadcast scheme, and the multicast scheme. The above schemes are presented in detail in the following paragraphs. Additionally, we analyze the number of the GTP tunnels established in every edge of the network.

Figure 2 shows a subset of a UMTS network. In this architecture, there are 2 SGSNs connected to a GGSN, 4 RNCs, and 12 Node Bs. Furthermore, 11 members of a multicast group are located in 6 cells. The BM-SC acts as the interface towards external sources of traffic [9]. In the presented analysis, we assume that a data stream that comes from an external PDN through BM-SC, must be delivered to the 11 UEs as illustrated in Figure 2.

The analysis presented in the following paragraphs, covers the forwarding mechanism of the data packets between the BM-SC and the UEs (Figure 2). Regarding the transmission of the packets over the Iub and Uu interfaces, it may be performed on common (ex. Broadcast Channel—BCH, Forward Access Channel—FACH), dedicated (Dedicated Channel—DCH) or shared transport channels (ex. High Speed Downlink Shared Channel—HS-DSCH). As presented in Reference [10], the transport channel that the 3GPP decided to use as the main transport channel for point-to-multipoint MBMS data transmission is the FACH with turbo coding and QPSK modulation at a constant transmission power. Multiple services can be configured in a cell, either time multiplexed on one FACH or transmitted on separate channels. DCH is a point-to-point channel and hence, it suffers from the inefficiencies of requiring multiple DCHs to carry the data to a group of users. However, DCH can employ fast closed-loop power control and soft handover mechanisms and generally is a highly reliable channel.

DOI: 10.1002/wcm

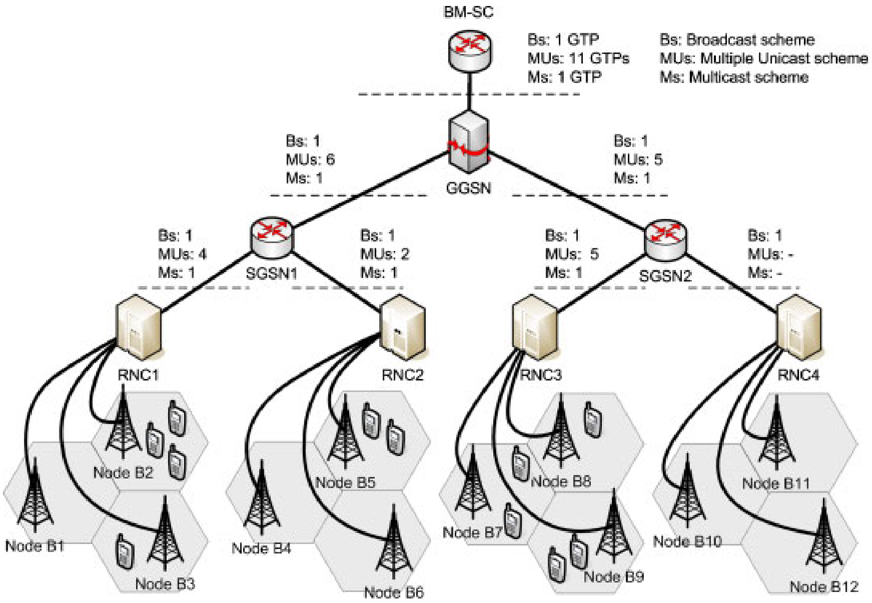


Fig. 2. Packet delivery in UMTS.

Additionally, DCH consists of an uplink channel. BCH is a point-to-multipoint channel, which has no uplink channel. Thus, even with a large number of multicast receivers, only one BCH is required for a multicast service in the cell. Furthermore, a new transport channel named HS-DSCH has been introduced as the primary radio bearer in Release 5 of UMTS. The HS-DSCH resources can be shared between all users in a particular sector. The primary channel multiplexing occurs in the time domain, where each Transmission Time Interval (TTI) consists of three slots (or 2 ms) [1].

It has to be mentioned that no decision has yet been made within 3GPP on how to optimize the MBMS data flow over the Iub. It has been proposed to avoid the duplication of Iub data flows and hence, only one Iub data flow per MBMS service should be used [11].

5.1. Description of the Broadcast Scheme (Bs)

With broadcast, the network simply floods multicast packets to all the nodes within the network. In this approach, the single packet is initially transmitted from the BM-SC to the GGSN. This procedure implies that the first GTP session is created between the BM-SC and the GGSN. Then, when the GGSN receives the packet, it forwards it to every SGSN that exists, which in turn forward the packet to the corresponding RNCs. As we described above, this leads to the creation of two GTP sessions for every path between the GGSN and RNCs. The first GTP session is created in the edge between GGSN and each SGSN, while the second one between the SGSN and each RNC. Then, each RNC forwards the packet to the corresponding Node Bs and the Node Bs ought to transmit the packet to the mobile users in their range.

Having analyzed the way that the packets are transmitted in the broadcast method, we can calculate the number of the GTP sessions created in the example that is shown in Figure 2. First, we have one GTP session in the edge between BM-SC and GGSN. Second, given the fact that we have two SGSN nodes, the corresponding GTP sessions that are created are two-one for each edge between the GGSN and SGSN. Furthermore, we observe four additional GTP sessions, because we assume that there are four RNCs and hence, four edges between the SGSNs and RNCs. The RNCs forward the packets to all Node Bs that have established the appropriate radio bearers. Figure 2 shows the number of the GTP sessions that are created in the edges of the network.

5.2. Description of the Multiple Unicast Scheme (MUs)

With multiple unicast, each packet is forwarded once to each member of the group separately. This means that when the BM-SC receives a packet, the packet is duplicated and each copy of the packet is transferred to a single mobile user. Consequently, in the edge between the BM-SC and the GGSN, the number of the GTP sessions that are created is equal to the number

DOI: 10.1002/wcm

of the multicast mobile users. Thus, in our example, the number of the GTP sessions in this edge is 11. Additionally, when the GGSN receives the packets, it forwards them to the corresponding SGSNs that serve at least one mobile user of the multicast group. In our example, the number of the GTP sessions in the edge between the GGSN and the SGSN1 is six, while in the edge between the GGSN and the SGSN2 is five.

Then, each of the two SGSNs forwards the received packets to the corresponding RNCs that serve mobile users. Additionally, the number of the GTP sessions created in the edges SGSN1–RNC1 and SGSN1–RNC2 are four and two, respectively. Likewise, the number of the GTP sessions for the edges SGSN2–RNC3 is five. Then, the RNCs forward the packets to the Node Bs that serve multicast users and have already established the appropriate radio bearers. The multicast users receive the packets on the appropriate radio bearers by FACHs or DCHs or HS-DSCHs. The number of the GTP sessions established in the edges of the network is shown in Figure 2.

However, this scheme, especially when the multicast group consists of a great number of UEs, produces excessive redundancies in the radio transmissions. These redundancies add up as interference to the network and limit its potential capacity.

5.3. Description of the Multicast Scheme (Ms)

With multicast, the packets are forwarded to those Node Bs that have multicast users. Therefore, in Figure 2, the Node Bs 2, 3, 5, 7, 8, 9 receive the multicast packets issued by the BM-SC. We briefly summarize the five steps occurred for the delivery of the multicast packets. First, the BM-SC receives a multicast packet and forwards it to the GGSN that has registered to receive the multicast traffic. Then, the GGSN receives the multicast packet and by querying its multicast routing tables, it determines which downstream SGSCs have multicast users residing in their respective service areas. The term 'downstream' refers to the topological position of one node with respect to another and relative to the distribution of the multicast data flow. In Figure 2, the GGSN duplicates the multicast packet and forwards it to the SGSN1 and the SGSN2.

Then, both destination SGSNs receive the multicast packets and, having queried their multicast routing tables, determine which RNCs are to receive the multicast packets. The destination RNCs receive the multicast packet and send it to the Node Bs that have established the appropriate radio bearers for the multicast application. In Figure 2, these are Node B2, B3, B5, B7, B8, B9. The multicast users receive the multicast packets on the appropriate radio bearers, either by point-to-point channels transmitted to individual users separately or by point-to-multipoint channels transmitted to all group members in the cell. In our analysis, we consider transport channels such as FACH, HS-DSCH, and DCH.

In this approach, every multicast packet is initially transmitted from the BM-SC to the GGSN. This procedure implies that the first GTP tunnel session is created between the BM-SC and the GGSN. The GGSN forwards exactly one copy of the multicast packet to each SGSN that serves multicast users. This leads to the creation of one GTP tunnel session between the GGSN and the SGSN1 and one GTP tunnel session between the GGSN and SGSN2 (Figure 2). Having received the multicast packets, the SGSN1 forwards exactly one copy of the multicast packet to the RNCs that serve multicast users, which are the RNC1 and the RNC2. In parallel, the SGSN2 forwards the multicast packets to the RNC3, which is the only RNC, covered by the SGSN2 that serves multicast users. If none of the SGSNs does not have valid routing information for any multicast user, the paging procedure is performed in order to determine the required information from the multicast users. Regarding the edges between the SGSNs and the RNCs in Figure 2, the first GTP tunnel is created between the SGSN1 and RNC1, the second between the SGSN1 and RNC2 session, and the third between the SGSN2 and RNC3. Finally, the RNCs forward the multicast packets to those Node Bs that multicast users reside in. Additionally, Figure 2 shows the exact number of the GTP sessions that are established in edges of the network for the Multicast scheme.

6. Evaluation of Different One-to-Many Packet Delivery Schemes

In this section, we present an evaluation, in terms of the telecommunication costs, of different one-to-many delivery schemes. We consider different UMTS network topologies and different transport channels for the transmission of the multicast data.

6.1. General Assumptions

We consider a subset of a UMTS network consisting of a single GGSN and N_{SGSN} nodes connected to the GGSN. Furthermore, each SGSN manages a number of N_{ra} RAs. Each RA consists of a number of N_{rnc} RNC

DOI: 10.1002/wcm

nodes, while each RNC node manages a number of N_{ura} URAs. Finally, each URA consists of N_{nodeb} cells. The total number of RAs, RNCs, URAs, and cells are:

$$N_{\mathrm{RA}} = N_{\mathrm{SGSN}} \cdot N_{\mathrm{ra}} \qquad (1)$$

$$N_{\mathrm{RNC}} = N_{\mathrm{SGSN}} \cdot N_{\mathrm{ra}} \cdot N_{\mathrm{rnc}} \qquad (2)$$

$$N_{\mathrm{URA}} = N_{\mathrm{SGSN}} \cdot N_{\mathrm{ra}} \cdot N_{\mathrm{rnc}} \cdot N_{\mathrm{ura}} \qquad (3)$$

$$N_{\mathrm{NODEB}} = N_{\mathrm{SGSN}} \cdot N_{\mathrm{ra}} \cdot N_{\mathrm{rnc}} \cdot N_{\mathrm{ura}} \cdot N_{\mathrm{nodeb}} \qquad (4)$$

The total transmission cost for packet deliveries including paging is considered as the performance metric. Furthermore, the cost for paging is differentiated from the cost for packet deliveries. We make a further distinction between the processing costs at nodes and the transmission costs on links, both for paging and packet deliveries. As presented in Reference [12] and analyzed in Reference [3], we assume that there is a cost associated with each link and each node of the network, both for paging and packet deliveries. For the analysis, we apply the following notations:

that a UE is in PMM idle/RRC idle state is P_{RA}, the probability that a UE is in PMM connected/RRC URA connected state is P_{URA}, and finally the probability that a UE is in PMM connected/RRC cell-connected state is P_{cell}.

Additionally, in the multicast scheme, we consider different values for the processing costs (P_{gM}, P_{sM}, P_{rM}) at the nodes of the UMTS network than the corresponding values (P_{g}, P_{s}, P_{r}) in the other two schemes since some overhead is needed in the UMTS nodes in order to maintain the routing tables required for the packet forwarding in the multicast scheme.

In the remainder of this section, we describe a method that models the multicast user distribution in the network. In particular, we present a probabilistic method that calculates the number of multicast users in the network (n_{UE}), the number of SGSNs that serve multicast users (n_{SGSN}), the number of RNCs that serve multicast users (n_{RNC}), and finally the number of Node Bs that serve multicast members (n_{NODEB}).

D_{gs}	Transmission cost of packet delivery between GGSN and SGSG
D_{sr}	Transmission cost of packet delivery between SGSN and RNC
D_{rb}	Transmission cost of packet delivery between RNC and Node B
D_{BCH}	Transmission cost of packet delivery over Iub and Uu with BCHs
D_{DCH}	Transmission cost of packet delivery over Iub and Uu with DCHs
$D_{\mathrm{HS-DSCH}}$	Transmission cost of packet delivery over Iub and Uu with HS-DSCHs
D_{FACH}	Transmission cost of packet delivery over Iub and Uu with FACHs
S_{sr}	Transmission cost of paging between SGSN and RNC
S_{rb}	Transmission cost of paging between RNC and Node B
S_{a}	Transmission cost of paging over the air
p_{g}	Processing cost of packet delivery at GGSN
p_{gM}	Processing cost of multicast packet delivery at GGSN
p_{s}	Processing cost of packet delivery at SGSN
p_{sM}	Processing cost of multicast packet delivery at SGSN
p_{r}	Processing cost of packet delivery at RNC
p_{rM}	Processing cost of multicast packet delivery at RNC
p_{b}	Processing cost of packet delivery at Node B
a_{s}	Processing cost of paging at SGSN
a_{r}	Processing cost of paging at RNC
a_{b}	Processing cost of paging at Node B

The total number of the multicast UEs in the network is denoted by N_{UE}. For the cost analysis, we define the total packets per multicast session as N_{p}. Since network operators will typically deploy an IP backbone network between the GGSN, SGSN, and RNC, the links between these nodes will consist of more than one hop. Additionally, the distance between the RNC and Node B consists of a single hop ($l_{\mathrm{rb}} = 1$). In the presented analysis we assume that the distance between GGSN and SGSN is l_{gs} hops, while the distance between the SGSN and RNC is l_{sr} hops.

Furthermore, we assume that the probability that a UE is in PMM detached state is P_{DET}, the probability

As introduced in Reference [7] and analyzed in Reference [3], we classify the RAs into L_{RA} categories. For $1 \leq i \leq L_{\mathrm{RA}}$ there are $N_i^{(\mathrm{RA})}$ RAs of class i. Therefore, the total number of RAs within the network is $N_{\mathrm{RA}} = \sum_{i=1}^{L_{\mathrm{RA}}} N_i^{(\mathrm{RA})}$.

Suppose that the distribution of the multicast users among the classes of RAs follows the Poisson distribution with $\lambda = \theta_i^{(\mathrm{RA})}$ where $1 \leq i \leq L_{\mathrm{RA}}$. In general, the probability that k exactly multicast users reside in the RAs of class i is calculated from the following equation:

$$p\left(k, \theta_i^{(\mathrm{RA})}\right) = \frac{e^{-\theta_i^{(\mathrm{RA})}} \cdot \left(\theta_i^{(\mathrm{RA})}\right)^k}{k!} \qquad (5)$$

DOI: 10.1002/wcm

Thus, the probability none of the RAs of class i serves multicast users is $p(0, \theta_i^{(\mathrm{RA})}) = e^{-\theta_i^{(\mathrm{RA})}}$, which in turn means that the probability at least one multicast user is served by the RAs of class i is $p = 1 - p(0, \theta_i^{(\mathrm{RA})}) = 1 - e^{-\theta_i^{(\mathrm{RA})}}$.

Since every class i consists of $N_i^{(\mathrm{RA})}$ RAs, the total number of the RAs in the class i, that serve multicast users is $N_i^{(\mathrm{RA})}(1 - e^{-\theta_i^{(\mathrm{RA})}})$. Thus, the total number of the RAs of every class that serve multicast users is:

$$n_{\mathrm{RA}} = \sum_{i=1}^{L_{\mathrm{RA}}} N_i^{(\mathrm{RA})} \left(1 - e^{-\theta_i^{(\mathrm{RA})}}\right) \tag{6}$$

where $\theta_i^{(\mathrm{RA})}$ represents the number of multicast users for the $N_i^{(\mathrm{RA})}$ RAs of class i.

If there are n_{RA} RAs that are serving multicast users, the probability that an SGSN does not have any such RA is:

$$p_{\mathrm{SGSN}} = \begin{cases} \dfrac{\dbinom{N_{\mathrm{RA}} - N_{\mathrm{ra}}}{n_{\mathrm{RA}}}}{\dbinom{N_{\mathrm{RA}}}{n_{\mathrm{RA}}}}, & if \quad n_{\mathrm{RA}} \le N_{\mathrm{RA}} - N_{\mathrm{ra}} \\ 0, & othewise \end{cases} \tag{7}$$

Based on Equation (7), the total number of SGSNs that are serving multicast users can be calculated as follows: $n_{\mathrm{SGSN}} = N_{\mathrm{SGSN}}\,(1 - p_{\mathrm{SGSN}})$.

The total number of multicast users in the network is:

$$N_{\mathrm{UE}} = \sum_{i=1}^{L_{\mathrm{RA}}} N_i^{(\mathrm{RA})} \theta_i \tag{8}$$

where θ_i is the number of multicast users in a RA of class i.

As in Reference [3], we assume that all RNCs within a service area of class i have the same multicast population distribution density as in the RA case. Based on a uniform density distribution within a single RA, the multicast population of an RNC within the service area of a class i RA is $\theta_i^{(\mathrm{RNC})} = \theta_i^{(\mathrm{RA})}/N_{rnc}$. The total number of RNCs of class i is $N_i^{(\mathrm{RNC})} = N_i^{(\mathrm{RA})} \cdot N_{\mathrm{rnc}}$.

Assuming that the number of RA categories is equal to the number of RNC categories ($L_{\mathrm{RNC}} = L_{\mathrm{RA}}$), the total number of RNCs that serve multicast users is:

$$n_{\mathrm{RNC}} = \sum_{i=1}^{L_{\mathrm{RNC}}} N_i^{(\mathrm{RNC})} \left(1 - e^{-\theta_i^{(\mathrm{RNC})}}\right) \tag{9}$$

The same are applied to the cells within the service area of an RNC. The average number of multicast users for a single cell of class i is $\theta_i^{(\mathrm{B})} = \theta_i^{(\mathrm{RNC})}/(N_{\mathrm{ura}} \cdot N_{\mathrm{nodeb}})$.

The number of Node Bs belonging to class i is $N_i^{(\mathrm{B})} = N_i^{(\mathrm{RNC})} \cdot N_{\mathrm{ura}} \cdot N_{\mathrm{nodeb}}$. Assuming that the number of the RNC categories is equal to the number of the Node B categories ($L_{\mathrm{RNC}} = L_{\mathrm{NODEB}}$), the total number of Node Bs that serve multicast users is:

$$n_{\mathrm{NODEB}} = \sum_{i=1}^{L_{\mathrm{NODEB}}} N_i^{(\mathrm{B})} \left(1 - e^{-\theta_i^{(\mathrm{B})}}\right) \tag{10}$$

6.2. Broadcast Scheme (Bs)

In this scheme, the packets are broadcasted to all the nodes of the network and no paging procedure is required. Regarding over the air transmission, the transport channel that is used in this scheme is the BCH. It is obvious that the total cost of the packet delivery is independent from the number of the multicast users (N_{UE}). The total cost of the packet delivery to the multicast users is computed as follows:

$$\begin{aligned} Bs = \big[&p_{\mathrm{g}} + N_{\mathrm{SGSN}}\left(D_{\mathrm{gs}} + p_{\mathrm{s}}\right) + N_{\mathrm{RNC}}(D_{\mathrm{sr}} + p_{\mathrm{r}}) \\ &+ N_{\mathrm{NODEB}}(D_{\mathrm{rb}} + p_{\mathrm{b}} + D_{\mathrm{BCH}})\big]\, N_{\mathrm{p}} \end{aligned} \tag{11}$$

6.3. Multiple Unicast Scheme (MUs)

With multiple unicast, each packet is forwarded once to each member of the group separately. In Figure 2, RNC1, for instance, would receive four duplicate copies of the same multicast packet from the SGSN1. It is obvious that the cost of a single packet delivery to a multicast user depends on its MM and RRC state.

If the multicast member is in PMM connected/RRC cell-connected state, then there is no need for any paging procedure neither from the SGSN nor from the serving RNC. In this case, the packet delivery cost is derived from Eqnuation (12). It has to be mentioned that this quantity does not include the cost for the transmission of the packets over the Iub and Uu interfaces since this cost depends first on the number of multicast users and second on the transport channel used for data transmission.

$$C_{\mathrm{cell}} = p_{\mathrm{g}} + D_{\mathrm{gs}} + p_{\mathrm{s}} + D_{\mathrm{sr}} + p_{\mathrm{r}} \tag{12}$$

If the multicast member is in PMM connected/RRC URA connected state, then the RNC must first page all the cells within the URA in which mobile users reside and then proceeds to the data transfer. After the

DOI: 10.1002/wcm

subscriber receives the paging message from the RNC, it returns to the RNC its cell ID. The cost for paging such a multicast member is:

$$C_{\mathrm{URA}} = N_{\mathrm{nodeb}}\,(S_{\mathrm{rb}} + a_{\mathrm{b}} + S_{\mathrm{a}}) + S_{\mathrm{a}} + a_{\mathrm{b}} + S_{\mathrm{rb}} + a_{\mathrm{r}} \tag{13}$$

If the multicast member is in PMM idle/RRC idle state, the SGSN only stores the identity of the RA in which the user is located. Therefore, all cells in the RA must be paged. The cost for paging such a multicast member is:

$$\begin{aligned} C_{\mathrm{RA}} = {} & N_{\mathrm{rnc}}\,(S_{\mathrm{sr}} + a_{\mathrm{r}}) + (N_{\mathrm{rnc}} \cdot N_{\mathrm{ura}} \cdot N_{\mathrm{nodeb}}) \\ & \times (S_{\mathrm{rb}} + a_{\mathrm{b}} + S_{\mathrm{a}}) + S_{\mathrm{a}} + a_{\mathrm{b}} + S_{\mathrm{rb}} \\ & + a_{\mathrm{r}} + S_{\mathrm{sr}} + a_{\mathrm{s}} \end{aligned} \tag{14}$$

Taking into consideration that the paging procedure is performed on the first packet of a data session, the total cost of the Multiple Unicast scheme is derived from the following equations for the three different transport channels (DCH, FACH, HS-DSCH), where n_{NODEB} represent the number of Node Bs that serve multicast users.

$$\mathrm{MUs} = \begin{cases} \big[P_{\mathrm{cell}} \cdot C_{\mathrm{cell}} \cdot N_{\mathrm{p}} + P_{\mathrm{URA}}\,(C_{\mathrm{URA}} + C_{\mathrm{cell}} \cdot N_{\mathrm{p}}) \\ \quad + P_{\mathrm{RA}}\,(C_{\mathrm{RA}} + C_{\mathrm{cell}} \cdot N_{\mathrm{p}}) + (P_{\mathrm{cell}} + P_{\mathrm{URA}} \\ \quad + P_{\mathrm{RA}}) \cdot (D_{\mathrm{DCH}} + D_{\mathrm{rb}} + p_{\mathrm{b}}) \cdot N_{\mathrm{p}}\big]\,N_{\mathrm{UE}} & (15) \\ \big[P_{\mathrm{cell}} \cdot C_{\mathrm{cell}} \cdot N_{\mathrm{p}} + P_{\mathrm{URA}}\,(C_{\mathrm{URA}} + C_{\mathrm{cell}} \cdot N_{\mathrm{p}}) \\ \quad + P_{\mathrm{RA}}\,(C_{\mathrm{RA}} + C_{\mathrm{cell}} \cdot N_{\mathrm{p}})\big]\,N_{\mathrm{UE}} \\ \quad + n_{\mathrm{NODEB}} \cdot (D_{\mathrm{FACH}} + D_{\mathrm{rb}} + p_{\mathrm{b}}) \cdot N_{\mathrm{p}} & (16) \\ \big[P_{\mathrm{cell}} \cdot C_{\mathrm{cell}} \cdot N_{\mathrm{p}} + P_{\mathrm{URA}}\,(C_{\mathrm{URA}} + C_{\mathrm{cell}} \cdot N_{\mathrm{p}}) \\ \quad + P_{\mathrm{RA}}\,(C_{\mathrm{RA}} + C_{\mathrm{cell}} \cdot N_{\mathrm{p}}) + (P_{\mathrm{cell}} + P_{\mathrm{URA}} \\ \quad + P_{\mathrm{RA}}) \cdot (D_{\mathrm{HS\text{-}DSCH}} + D_{\mathrm{rb}} + p_{\mathrm{b}}) \cdot N_{\mathrm{p}}\big]\,N_{\mathrm{UE}} & (17) \end{cases}$$

The last term in each of the above three equations represent the cost of the packet transmission in the Iub and Uu interfaces. In general, in case we use the FACH as transport channel, each multicast packet send once over the Iub interface and then the packet is transmitted to the UEs that served by the corresponding Node B. In case we use DCHs for the transmission of the multicast packets, each packet is replicated over the Iub as many times as the number of multicast users that the corresponding Node B serves. Finally, with HS-DSCH, a separate timeslot must be used to transport the multicast data to each multicast receiver. However, one could envision that all multicast receivers could receive the same timeslot that contains the multicast data, but in its current form the HS-DSCH has not been modified to allow this. Thus, the number of time slots required for the transmission of the multicast data to the multicast users is equal to the number of multicast users reside in the corresponding cell.

6.4. Multicast Scheme (Ms)

In the multicast scheme, the multicast group management is performed at the GGSN, SGSN, and RNC and multicast tunnels are established over the Gn and Iu interfaces. All multicast users that are in PMM idle/RRC idle or PMM connected/RRC URA connected state must be paged. After the paging procedure, the RNC stores the location of any UE at a cell level. The cost for that paging procedure is given by Equations (13) and (14), respectively. In multicast, the SGSN and the RNC forward a single copy of each multicast packet to those RNCs or Node Bs respectively that are serving multicast users. After the correct multicast packet reception at the Node Bs that serve multicast users, the Node Bs transmit the multicast packets to the multicast users via common, dedicated, or high speed shared transport channels. The total cost for the Multicast scheme is derived from the following equation where n_{SGSN}, n_{RNC}, n_{NODEB} represent the number of SGSNs, RNCs, Node Bs, respectively that serve multicast users.

$$\begin{aligned} \mathrm{Ms} = {} & \big[p_{\mathrm{gM}} + n_{\mathrm{SGSN}}\,(D_{\mathrm{gs}} + p_{\mathrm{sM}}) + n_{\mathrm{RNC}}\,(D_{\mathrm{sr}} + p_{\mathrm{rM}}) \\ & + Y\big]\,N_{\mathrm{p}} + (P_{\mathrm{RA}} \cdot C_{\mathrm{RA}} + P_{\mathrm{URA}} \cdot C_{\mathrm{URA}})\,N_{\mathrm{UE}} \\ = {} & D_{\mathrm{packet\_delivery}} + D_{\mathrm{paging}} \end{aligned} \tag{18}$$

where

$$Y = \begin{cases} n_{\mathrm{NODEB}} \cdot (D_{\mathrm{FACH}} + D_{\mathrm{rb}} + p_{\mathrm{b}}) & \text{if} \quad \text{channel} = \mathrm{FACH} \\ N_{\mathrm{UE}} \cdot (D_{\mathrm{DCH}} + D_{\mathrm{rb}} + p_{\mathrm{b}}) & \text{if} \quad \text{channel} = \mathrm{DCH} \\ N_{\mathrm{UE}} \cdot (D_{\mathrm{HS-DSCH}} + D_{\mathrm{rb}} + p_{\mathrm{b}}) & \text{if} \quad \text{channel} = \mathrm{HS-DSCH} \end{cases}$$

$$\begin{aligned} D_{\mathrm{packet\_delivery}} & = \big[p_{\mathrm{gM}} + n_{\mathrm{SGSN}}\,(D_{\mathrm{gs}} + p_{\mathrm{sM}}) \\ & \quad + n_{\mathrm{RNC}}\,(D_{\mathrm{sr}} + p_{\mathrm{rM}}) + Y\big]\,N_{\mathrm{p}} \\ D_{\mathrm{paging}} & = (P_{\mathrm{RA}} \cdot C_{\mathrm{RA}} + P_{\mathrm{URA}} \cdot C_{\mathrm{URA}})\,N_{\mathrm{UE}} \end{aligned}$$

The parameter Y represents the multicast cost for the transmission of the multicast data over the Iub and Uu

Table I. Chosen values for the calculation of transmission costs in the links.

Link	Link capacity factor (k)	Number of hops (l)	Transmission cost (D)
GGSN–SGSN	$k_{\text{gs}} = 0.5$	$l_{\text{gs}} = 6$	$D_{\text{gs}} = 12$
SGSN–RNC	$k_{\text{sr}} = 0.5$	$l_{\text{sr}} = 3$	$D_{\text{sr}} = 6$
RNC–Node B	$k_{\text{rb}} = 0.2$	$l_{\text{rb}} = 1$	$D_{\text{rb}} = 5$

interfaces. This cost of the multicast scheme depends mainly on the distribution of the multicast group within the UMTS network and secondly on the transport channel that is used. In case we use the FACH as transport channel, each multicast packet send once over the Iub interface and then the packet is transmitted to the UEs that served by the corresponding Node B. However, in case we use DCHs or HS-DSCHs for the transmission of the multicast packets over the Iub each packet is replicated over the Iub as many times as the number of multicast users that the corresponding Node B serves.

7. Results

Having analyzed the costs of the above presented one-to-many packet delivery methods, we try to evaluate the cost of each scheme assuming a general network topology. In general, we assume a network configuration, with $N_{\text{SGSN}} = 10$, $N_{\text{ra}} = 10$, $N_{\text{rnc}} = 10$, $N_{\text{ura}} = 5$ and $N_{\text{nodeb}} = 5$.

The packet transmission cost (D_{xx}) in any segment of the UMTS network depends on two parameters: the number of hops between the edge nodes of this network segment and the capacity of the link of the network segment. This means that $D_{\text{gs}} = l_{\text{gs}}/k_{\text{gs}}$, $D_{\text{sr}} = l_{\text{sr}}/k_{\text{sr}}$ and $D_{\text{rb}} = l_{\text{rb}}/k_{\text{rb}}$. Parameter k_{xx} represents the profile of the corresponding link between two UMTS network nodes. More specifically, in the high capacity links at the CN, the values of k_{xx} are greater than the corresponding values in the low capacity links at UTRAN. For the cost analysis and without loss of generality, we assume that the distance between the GGSN and SGSN is 6 hops, the distance between SGSN and RNC is 3 hops, and the distance between RNC and Node B is 1 hop. The above parameters as well as the values of the k_{xx} are presented in detail in Table I. Regarding the transmission cost of paging (S_{xx}) in the segments of the UMTS network, it is calculated in a similar way as the packet transmission cost (D_{xx}). More specifically, S_{xx} is a fraction of the calculated transmission cost (D_{xx}) and in our case, we assume that it is three times smaller than D_{xx}.

As we can observe from the equations in the previous section, the costs of the schemes depend on a number of other parameters. Thus, we have to estimate the value of these parameters. This procedure implies that we choose the values appropriately, taking into consideration the relations between them. The chosen values of the parameters are presented in Table II.

As it is shown in Table II, the values for the transmission costs of the packet delivery over the air with each of the four transport channels are different. More specifically, the transmission cost over the air with BCH (D_{BCH}) or FACH (D_{FACH}) is bigger than the D_{DCH}, which in turn is bigger than the transmission cost of the packet delivery over the air with HS-DSCH ($D_{\text{HS-DSCH}}$). This occurs because BCH and FACH as common channels require high transmission power in order to reach all the users within the coverage area even if they are not members of the multicast group. Additionally, regarding the relation of the costs of the DCH and the HS-DSCH, the DCH is a power controlled channel while the HS-DSCH is rate controlled. In Reference [1], it is shown that the HS-DSCH throughput is by far bigger than the DCH throughput for a giver fraction of Node B power allocated for these channels. For example, if we consider a macro cell environment

Table II. Chosen parameters' values.

D_{gs}	D_{sr}	D_{rb}	S_{sr}	S_{rb}	S_{a}	p_{g}	p_{s}	p_{r}	p_{gM}	p_{sM}	p_{rM}	p_{b}	a_{s}	a_{r}	a_{b}	p_{RA}	p_{URA}	p_{cell}
12	6	5	2	5/3	4/3	1	1	1	2	2	2	1	1	1	1	0.6	0.2	0.1

D_{BCH}	D_{DCH}	D_{FACH}	$D_{\text{HS-DSCH}}$
16	14	16	12

DOI: 10.1002/wcm

configuration with a ITU Pedestrian-A delay profile and 7W allocated to HSDPA from the available Node B power, the average cell throughput on the HS-DSCH is 1.4 Mbps while the value of the throughput with only non HSDPA terminals active (64 kbps R99 DCH users only) in the cell is 1 Mbps. The average throughput that the HSDPA users experience depends on the number of simultaneous users that are sharing the HS-DSCH, as well as their relative experienced signal quality (E_b/N_0). On the contrary, the total downlink transmission power allocated for DCH is variable and increasing exponentially while the UE distance from the node B is increasing. Also, the more the UEs in the cell thus the higher the interference, the more exponential the increase in the total power required. Furthermore, the power required for a reference DCH user depends on the experienced signal quality (E_b/N_0). The following equation calculates the total required transmission power at the Node B for a reference DCH user [13].

$$P_{\mathrm{T}i} = L_{p,i} \frac{P_{\mathrm{N}} + \chi_i + p\frac{P_T}{L_{p,i}}}{\frac{W}{(E_b/N_0)_i R_{\mathrm{b},i}} + p} \tag{19}$$

where P_{T} is the base station transmitted power, $P_{\mathrm{T}i}$ is the power devoted to the ith user, $L_{\mathrm{p},i}$ is the path loss, $R_{\mathrm{b},i}$ the ith user transmission rate, W the bandwidth, P_{N} the background noise, p is the orthogonality factor, and χ_i is the intercell interference observed by the ith user.

Regarding the parameters P_{gM}, P_{sM}, and P_{rM} in the multicast scheme and their relation with the parameters P_{g}, P_{s}, and P_{r} of the other two schemes, the latter parameters have lower values that the values of the former parameters since some overhead is needed in the UMTS nodes in order to maintain the routing tables required for the multicast packet forwarding in the multicast scheme. At this point, we have to mention that since the nodes that are responsible for the establishment of the radio bearers are the RNCs we assume that there is no need for maintaining any routing tables in the Node Bs and thus there is no any additional cost for the processing at these nodes.

Furthermore, we have chosen appropriately the probabilities P_{RA}, P_{URA}, and P_{cell} More specifically, the probability that a UE is in PMM idle/RRC idle state is $P_{\mathrm{RA}} = 0.6$. The probability that a UE is in PMM-connected/RRC URA-connected state is $P_{\mathrm{URA}} = 0.2$ and the probability that a UE is in PMM connected/RRC cell-connected state is $P_{\mathrm{cell}} = 0.1$. Additionally, there is a probability that the UE is not reachable by the network and we consider it to be 0.1.

It is true that the performance of the three schemes depends mainly, on the configuration of the UMTS network that is under investigation. Therefore, we consider a general network configuration, with $N_{\mathrm{SGSN}} = 10$, $N_{\mathrm{ra}} = 10$, $N_{\mathrm{rnc}} = 10$, $N_{\mathrm{ura}} = 5$ and $N_{\mathrm{nodeb}} = 5$. In our analysis, we assume that we have two classes of RAs. A class $i = 1$ RA has multicast user population of $\theta_1 = 1/\delta$ and a class $i = 2$ RA has a multicast user population of $\theta_2 = \delta$. If $\delta \gg 1$, the class $i = 1$ RA has a small multicast user population and the class $i = 2$ RA has a large multicast user population. Let α be the proportion of the class $i = 1$ RAs and $(1 - \alpha)$ be the proportion of the class $i = 2$ RAs [7]. Thus, the number of class $i = 1$ RAs is $N_1^{(\mathrm{RA})} = \alpha N_{\mathrm{RA}}$ and the number of class $i = 2$ RAs is$N_2^{(\mathrm{RA})} = (1 - \alpha)N_{\mathrm{RA}}$. Each RA of class $i \in \{1, 2\}$ is in turn subdivided into N_{rnc} RNCs of the same class i and similarly, each RNC of class $i \in \{1, 2\}$ is subdivided into $N_{\mathrm{ura}} \cdot N_{\mathrm{nodeb}}$ Node Bs of the same class i. Take into consideration the above mentioned parameters, Equation (8) can be rewritten as follows:

$$\begin{aligned} N_{\mathrm{UE}} &= \sum_{i=1}^{2} N_i^{(\mathrm{RA})} \cdot \theta_i = N_1^{(\mathrm{RA})} \cdot \theta_1 + N_2^{(\mathrm{RA})} \cdot \theta_2 \\ &= N_{\mathrm{RA}} \left(\frac{\alpha}{\delta} + \delta - \alpha\delta \right) \end{aligned} \tag{20}$$

It is obvious from Equation (20) that as α decreases and δ increases the number of multicast users increases rapidly.

In Figures 3 and 4, we plot the cost of the three schemes in function of α, for different values of δ and N_{p} respectively. Since in our model we consider three different transport channels over the air for the MUs and the Ms, in the following figures only the channel with the lowest cost of each scheme is presented. Additionally, we provide the comparison of the costs of every scheme using different transport channels in separate figures. Generally, we observe from Figure 3 and Figure 4, that the cost of the broadcast scheme is constant, while the costs of the other two schemes decrease as α increases. With $\delta \gg 1$, it has to be mentioned that there are no multicast users in a RA of class $i = 1$ and there are many multicast members in a RA of class $i = 2$. Furthermore, as α increases, the number of class $i = 1$ RAs with no multicast users increases and hence, the costs of the MUs and the Ms decrease as it is shown in Figures 3 and 4.

More specifically, in Figure 3, we present the costs of the three schemes in function of α for $\delta = 100$ (Figure 3a) and $\delta = 1000$ (Figure 3b). For $\delta = 100$, the

DOI: 10.1002/wcm

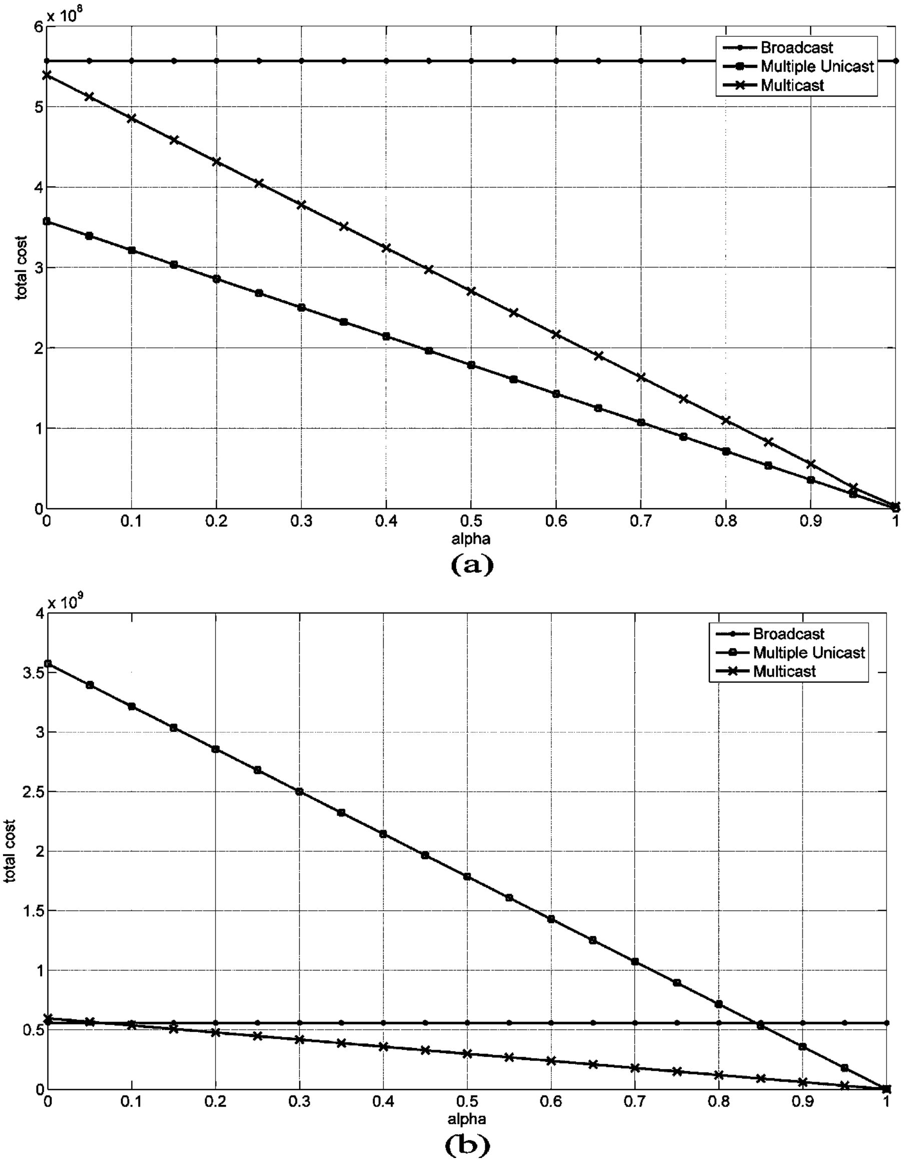


Fig. 3. Total cost for the three schemes in function of α with $N_p = 1000$. (a) $\delta = 100$, (b) $\delta = 1000$.

MUs has the lowest cost while the Bs has the higher cost independently of the value of α. This occurs because in small values of α, there are many RAs with large multicast users' population (class $i = 2$ RAs). However, the value of $\delta = 100$ results to a small number of multicast users within the network and hence, the costs of MUs and Ms are kept in lower values than the value of the cost of the Bs. In addition, if $\delta = 1000$ (Figure 3b), the number of multicast users within the network is increased and this results to an increased cost for the MUs. The costs of the other two schemes behave as follows: for $\alpha < 0.06$ the Bs has the lowest cost and for $\alpha > 0.06$ the Ms has the lowest cost. The latter occurs because for small values of α and increased number of δ, the number of multicast users within the network is increased and furthermore, there are many class $i = 2$ RAs in the network with large multicast users' population. This means that the multicast users within the network are spreaded to many RAs and hence, the cost of the paging which is required in the Ms is increased, making the Ms inefficient for this network topology. On the other hand, when the value of α is increased, the number of class $i = 1$ RAs with no multicast users is increased and all the multicast users are located in a small number of class $i = 2$ RAs. Thus, the Ms is more efficient than the Bs as it is shown in Figure 3b.

DOI: 10.1002/wcm

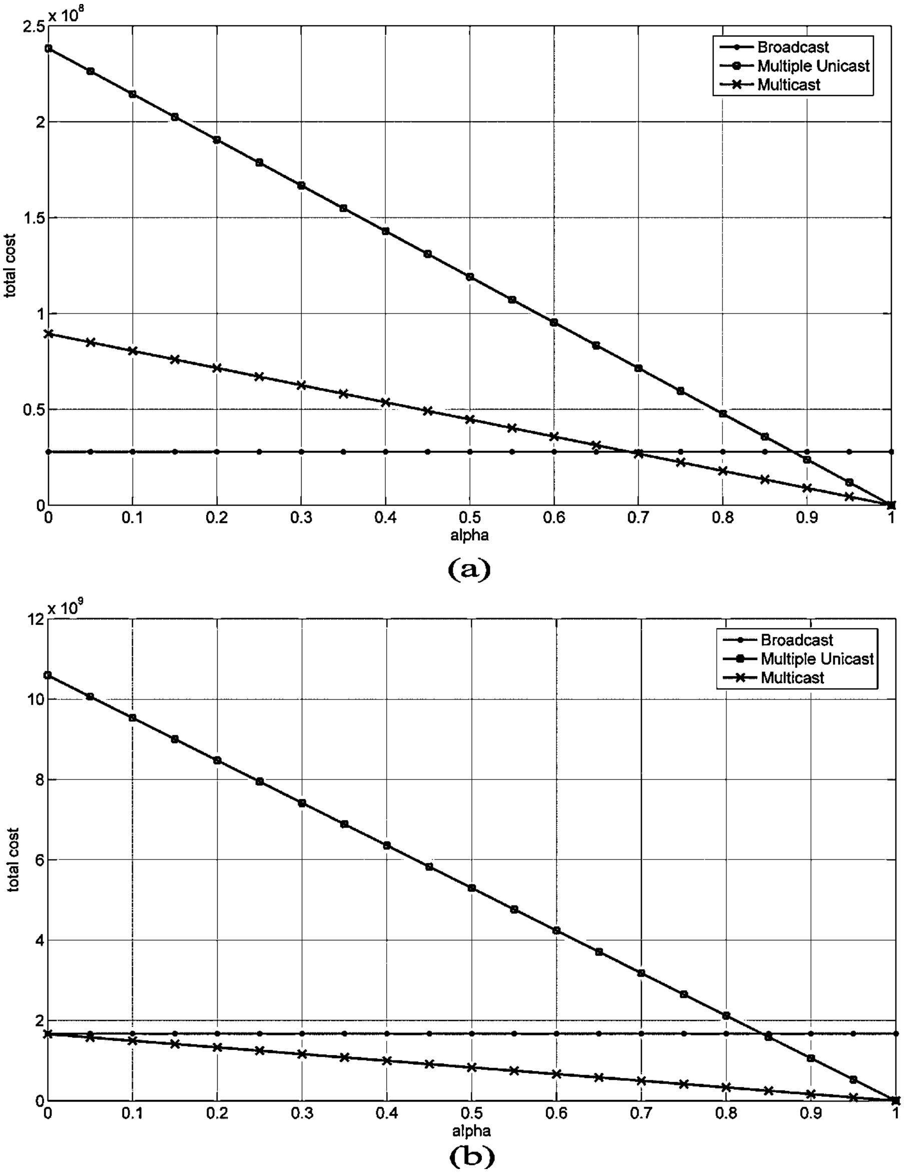


Fig. 4. Total cost for the three schemes in function of α with $\delta = 1000$. (a) $N_\mathrm{p} = 50$, (b) $N_\mathrm{p} = 3000$.

In Figure 4, we present the costs of the three schemes in function of α for $N_\mathrm{p} = 50$ (Figure 4a) and $N_\mathrm{p} = 3000$ (Figure 4b). The value of $N_\mathrm{p} = 50$ corresponds to a multicast session with few transmitted packets while the value of $N_\mathrm{p} = 3000$ corresponds to a multicast session consisting of a large number of packets. For $N_\mathrm{p} = 50$ (Figure 4a), if $\alpha < 0.70$ the Bs outperforms both the Ms and the MUs. The later has the highest value for these values of α. This occurs because a small value of α results to a topology with large number of high populated RAs (class $i = 2$ RAs). As the parameter α increases the number of RAs with no multicast users increases and the Ms becomes more favorable than the other two schemes. Obviously, for $\alpha > 0.70$ the Ms has the lowest cost compared to Bs and MUs. Additionally, for $\alpha > 0.90$, the MUs outperforms the Bs. This occurs because a value of α close to 1 in conjunction with a big value of δ ($\delta = 1000$) results to small number of multicast users according to Equation (21). The same observation regarding the MUs and the Bs occurs also in Figure 4b. Furthermore, for $N_\mathrm{p} = 3000$ (Figure 4b), the multicast session is consisted of a large number of packets and hence, the Ms has the lowest cost independently of the parameter α.

DOI: 10.1002/wcm

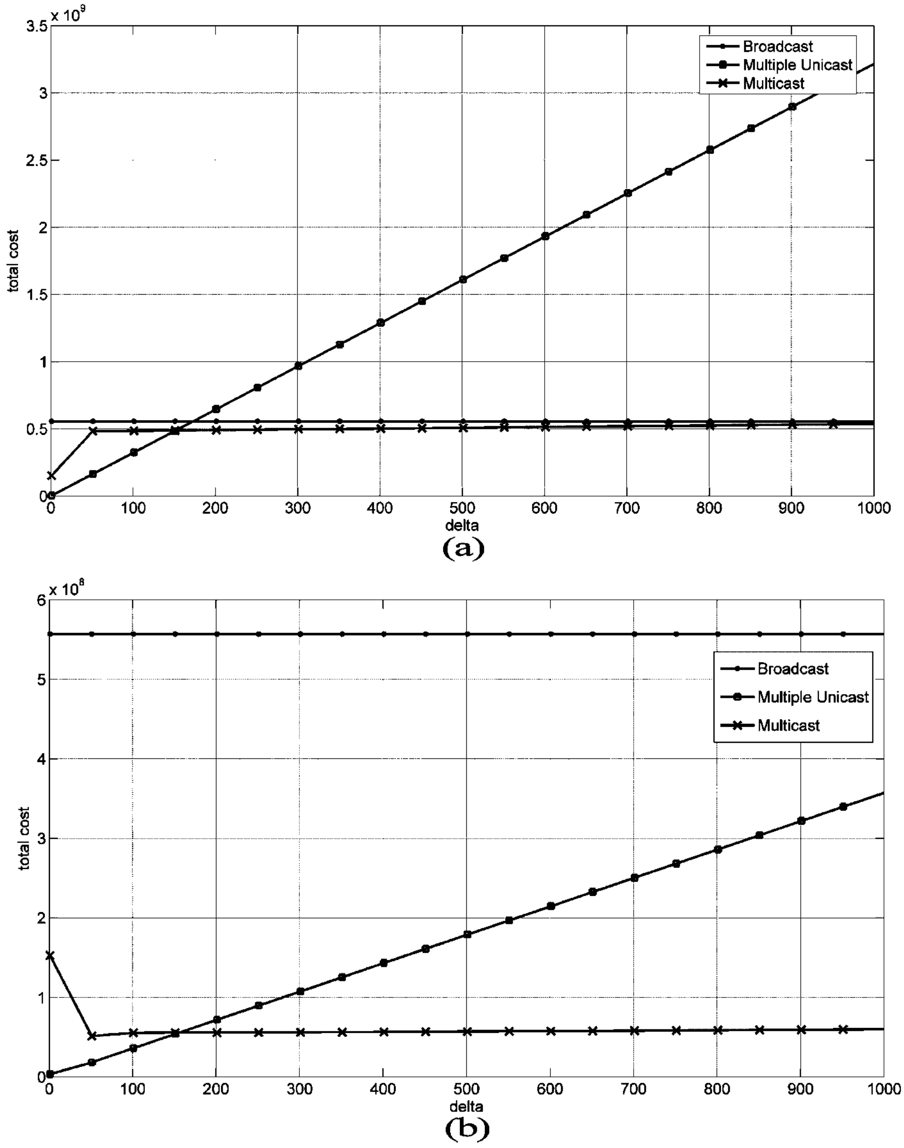


Fig. 5. Total cost for the three schemes in function of δ with $N_\mathrm{p} = 1000$. (a) $\alpha = 0.1$, (b) $\alpha = 0.9$.

Figure 5 presents the costs of the three schemes in function of δ for $\alpha = 0.1$ (Figure 5a) and $\alpha = 0.9$ (Figure 5b). Our first observation is that the cost of the Bs is constant and it is not depending on the parameter δ while the cost of the MUs increases as δ increases. As it has already been mentioned, a class $i = 1$ RA has $1/\delta$ multicast users while a class $i = 2$ RA has δ multicast users. Furthermore, if α converges to 1 (Figure 5b), the number of class $i = 2$ RAs in the network is small. Thus, the multicast users are located in small number of RAs and as a result, the cost of the Ms is kept in lower values (independently of the parameter δ) than the cost of the MUs and the Bs (Figure 5b). The later, obviously, has the higher cost as it is shown in Figure 5b. Additionally, if δ is small, the MUs outperforms Ms since the number of multicast users is limited.

On the other hand, Figure 5a shows that the Ms performs efficiently if $\delta < 900$, but it is outperformed by the Bs for $\delta > 900$. This occurs because the value of α is small and the number of class $i = 2$ RAs with large multicast users' population is increased and hence, as the number of multicast users in a class $i = 2$ RA increases, there are many users among many cells in the network making the Bs more appropriate for the multicast packet delivery. Additionally, if δ converges to zero the MUs outperforms Ms since the number of multicast users in the network is limited.

DOI: 10.1002/wcm

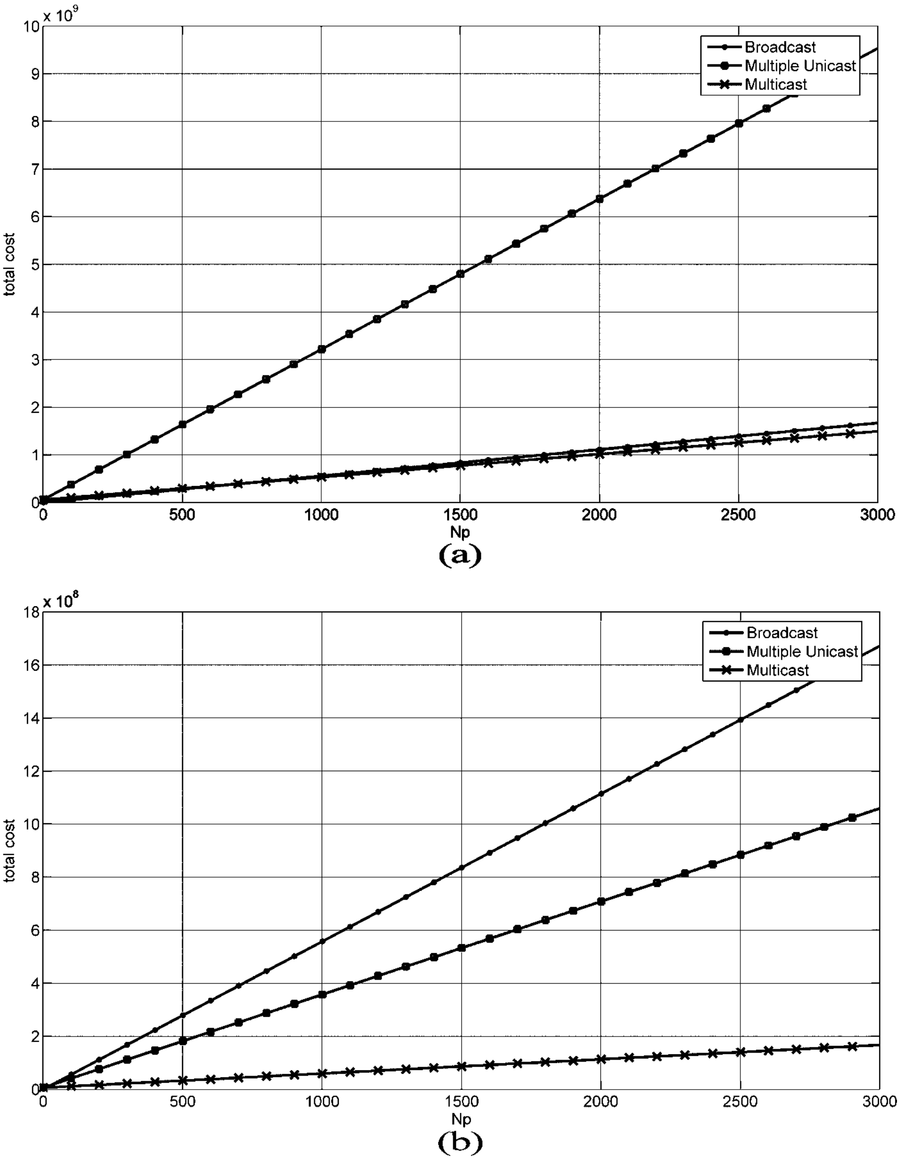
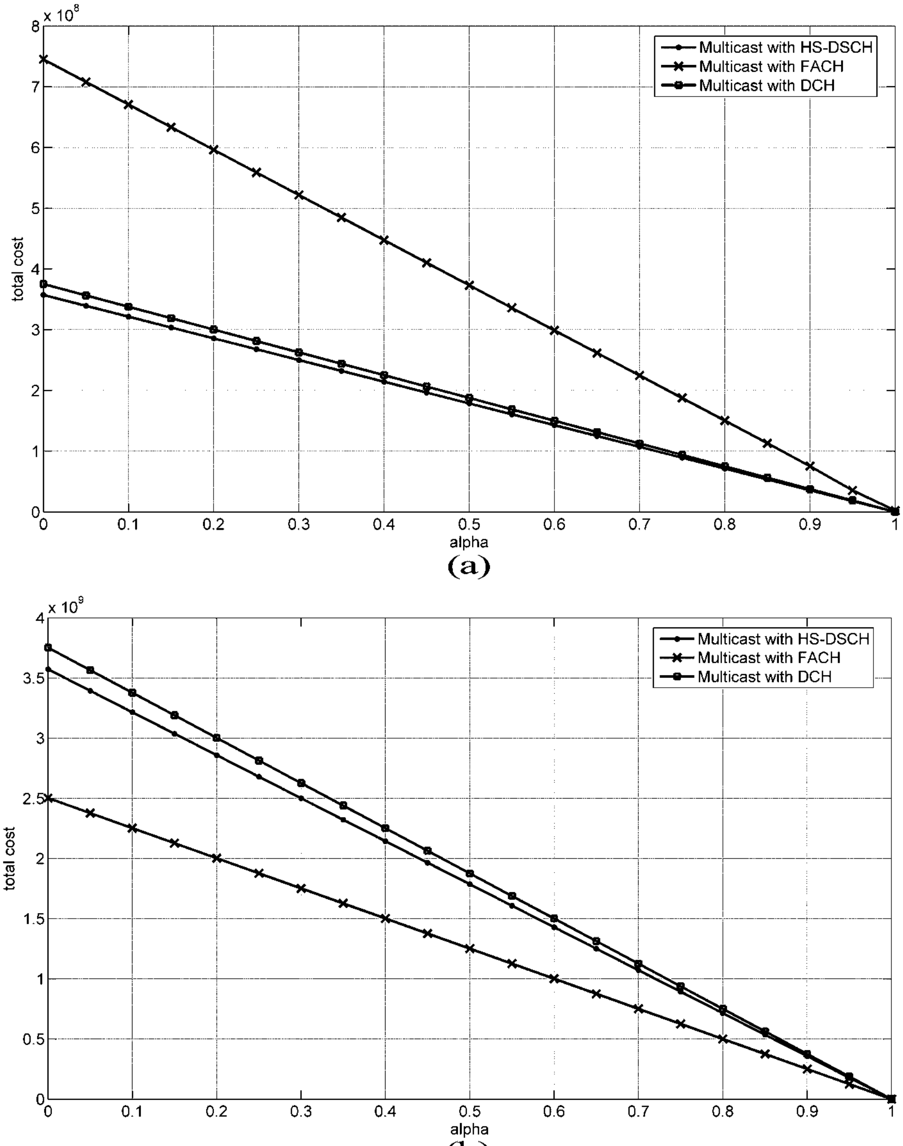


Fig. 6. Total cost for the three schemes in function of N_p with $\delta = 1000$. (a) $\alpha = 0.1$, (b) $\alpha = 0.9$.

The costs of the three schemes in function of the number of packets per multicast session are presented in Figure 6 for two different values of α. Our first observation is that as N_p increases, the total cost of each scheme increases. When α is small (Figure 6a), the number of RAs in the class $i = 2$ is increased and hence, the multicast users are spread to many RAs and cells within the network. Thus, the Ms and the Bs outperform the MUs for the given value of $\delta = 1000$. Regarding the Bs and the Ms they have similar costs in function of N_p. When $\alpha = 0.9$ (Figure 6b), the number of RAs in the class $i = 2$ is small and hence, the multicast users are residing in small number of RAs and cells within the network. In that case the Ms outperforms both MUs and Bs. The latter has obviously the highest cost for this network topology.

In Figure 7, the total costs for the Ms using different transport channels over the air in function of α, are presented. As we can observe, the cost decreases as α increases, independently of the type of the transport channel used for the transmission of the multicast data over the air. This occurs because the increment of α entails that the number of RAs that have small population increases. Thus, the total number of the multicast users decreases and the total cost for the transmission of the data decreases.



(a)

(b)

Fig. 7. Total costs for the Ms using different transport channels over the air in function of α with $N_p = 1000$. (a) $\delta = 100$, (b) $\delta = 1000$.

More specifically, in Figure 7a, we observe that the cost of the Ms using HS-DSCH is smaller than the cost of the Ms if we use DCHs or FACH. This occurs because the small value of δ, results to a reduced number of multicast users in the network and hence, the HS-DSCH (or DCH) is the more efficient transport channel in terms of the cost. On the other hand, the increased cost of the FACH (D_{FACH}) in conjunction with the small number of multicast users in the network makes FACH inefficient for this user distribution scheme. The opposite occurs in Figure 7b where the value of δ is increased, which means that the number of multicast users in the network is also increased. The increased number of multicast users in the network makes the use of DCHs and the HS-DSCHs inefficient for the transmission of the data over the Iub and Uu interfaces and the FACH the most appropriate transport channel as it is shown in Figure 7b. The same observations occur in the case of the MUs.

In Figure 8, the total costs for the Ms and MUs using different transport channels over the air in function of δ, are presented. In the case of the Ms, we chose a small value for the parameter α because the Ms becomes efficient when there is an increased density of multicast

DOI: 10.1002/wcm

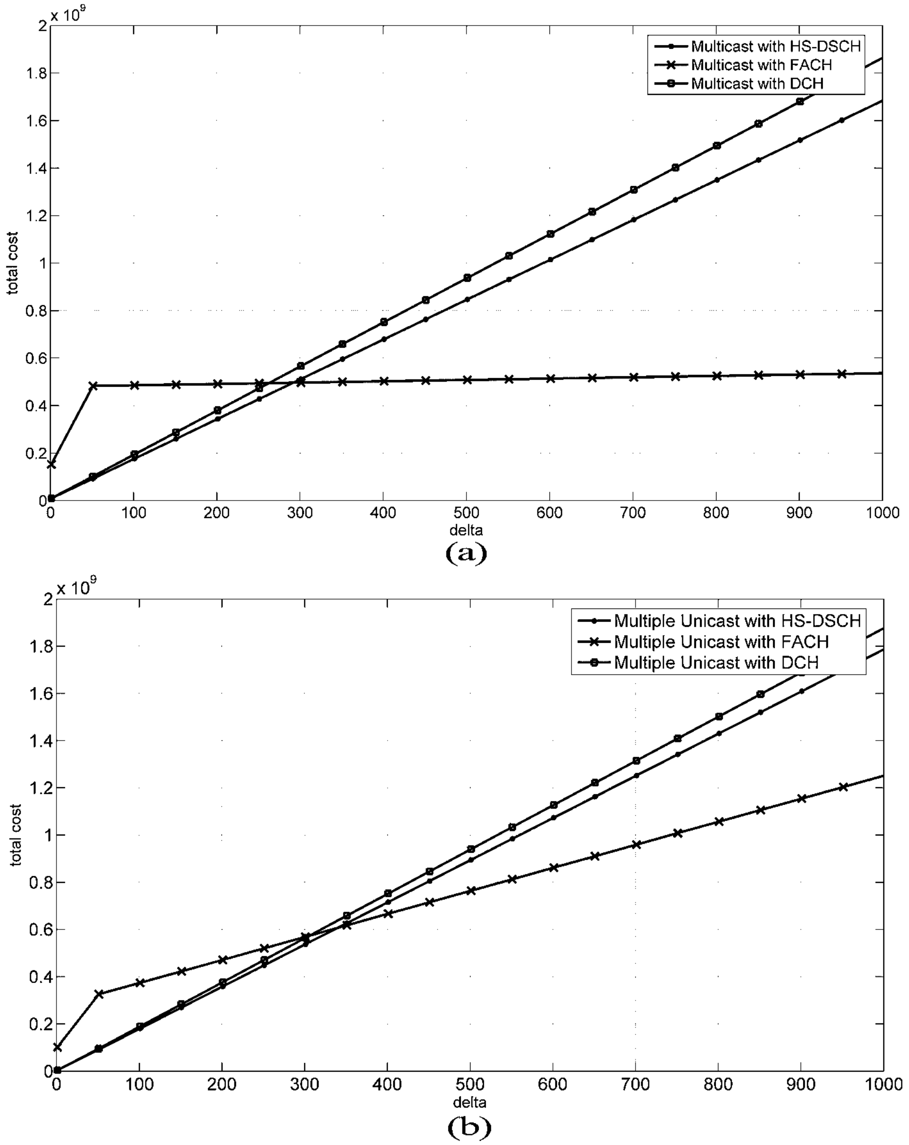


Fig. 8. Total costs for the Ms and MUs using different transport channels over the air in function of δ with $N_\mathrm{p} = 1000$. (a) $\alpha = 0.1$, (b) $\alpha = 0.5$.

users in the network. On the other hand, the value of α in the MUs must be big for the appropriate use of this scheme. Therefore, for the Ms (Figure 8a), we choose a value of $\alpha = 0.1$ which means that there are many RAs in the network with a great number of multicast users in these. Regarding the MUs (Figure 8b), we consider the same amount of RAs having large multicast users' population and RAs having small multicast users' population ($\alpha = 0.5$).

More specifically, in Figure 8a, we observe that for small values of δ, the cost of the Ms using DCHs or HS-DSCHs is small because there is a small number of multicast users in the network. Thus, the use of DCHs or HS-DSCHs for the data transmission over the air which reduces the cost of the scheme. On the other hand, bigger values of the parameter δ imply bigger number of multicast users in the network and hence, the use of FACH is more appropriate. The latter has obvious advantage in topologies with many high populated RAs since the multicast packets are send once over Iub and Uu interfaces. In Figure 8b, we present the cost of the MUs using the three transport channels. More specifically, as in the Ms, for small values of δ, the number of the multicast users is small and hence, the use of DCHs or HS-DSCHs for the data transmission over Uu is the most efficient. On the other hand, as

DOI: 10.1002/wcm

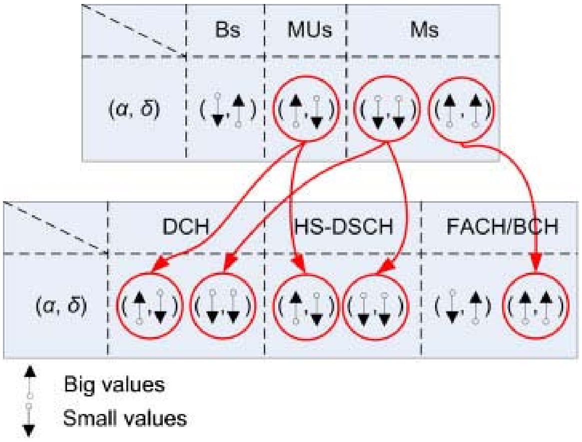


Fig. 9. Summary of the analysis.

δ increases, the use of FACH makes the scheme more efficient.

As it is shown in Figures 7 and 8, the cost of using HS-DSCHs for the transmission of the multicast data over the Iub and Uu interfaces in both the MUs and the Ms is always lower than the cost of using DCHs. Our model handles identical the two types of channels and its only difference is the lower cost of the HS-DSCH ($D_{\mathrm{HS-DSCH}}$) than the cost of DCH (D_{DCH}). Our decision of using lower cost for the HS-DSCH than the DCH is explained on the beginning of the results section. In general, the selection of the bearer type is operator-depended, typically based on the downlink radio resource situation so that the efficiency of the resource usage can be maximized.

A summary of our analysis is presented in Figure 9.

8. Future work

The step that follows this work is to carry out experiments using the NS-2 simulator. This means that we have to implement the above presented one-to-many packet delivery schemes and confirm the relation of the costs through the experiments. Further work could focus on evaluating the performance of MBMS, once the complete specifications are available, as well as studying the management of the user mobility. Furthermore, an MBMS handover mechanism in RNC should be realized in order to optimize the transmission of the multicast data in the Iub interface. Additionally, an innovative algorithm could be developed which would choose the most efficient transport channel for the transmission of the multicast packets depending on the distribution of the users and the network topology at any given time. In this way, interesting issues must be taken into account such as the user mobility, the max number of users that a channel can service as well as power issues.

9. Conclusions

In this paper, we presented an overview of three different packet delivery schemes for UMTS. These schemes include the Broadcast, the Multiple Unicast, and the Multicast scheme. We analytically presented these schemes and analyzed their performance in terms of the packet delivery cost and the scalability of each scheme. More specifically, we provided a formula to differentiate the number of the multicast users and their distribution. This formula is a function of the α and δ, where as α decreases and δ increases the number of multicast users increases rapidly and vice-versa. Thus, in the multicast scheme, which is useful when there are many multicast users gathered in some cells of the network, we chose a small value of α and a big value of δ. On the other hand, the multiple unicast is an efficient transmission scheme when there is a small number of multicast users and hence, we applied a big value of α and a small value of δ.

References

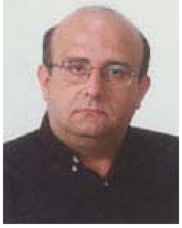
1. Holma H, Toskala A. *WCDMA for UMTS: Radio Access for Third Generation Mobile Communications*. John Wiley & Sons: Chichester, 2004. ISBN 0-470-87096-6.
2. Gossain H, Cordeiro C, Argawal D. Multicast: wired to wireless. *IEEE Communications Magazine* 2002; **40**(6): 116–123.
3. Rummler R, Chung Y, Aghvami H. Modeling and Analysis of an efficient multicast mechanism for UMTS. *IEEE Transactions on Vehicular Technology* 2005; **54**(1): 350–365.
4. 3GPP TS 23.060 V7.0.0. Technical Specification Group Services and System Aspects; General Packet Radio Service (GPRS); Service description; Stage 2 (Release 7). 2006.
5. Yang S, Lin Y. Performance evaluation of location management in UMTS. *IEEE Transactions on Vehicular Technology* 2003; **52**(6): 1603–1615.
6. Hauge M, Kure O. Multicast in 3G networks: employment of existing IP multicast protocols in UMTS. *5th ACM International Workshop on Wireless Mobile Multimedia* 2002, 96–103.
7. Lin Y. A multicast mechanism for mobile networks. *IEEE Communication Letters* 2001; **5**(11): 450–452.
8. 3GPP TS 22.146 V7.1.0. Technical Specification Group Services and System Aspects; Multimedia Broadcast/Multicast Service; Stage 1 (Release 7). 2006.
9. 3GPP TS 23.246 V6.9.0. Technical Specification Group Services and System Aspects; Multimedia Broadcast/Multicast Service (MBMS); Architecture and functional description (Release 6). 2005.
10. 3GPP, TR 23.846 v6.1.0. Technical Specification Group Services and System Aspects; Multimedia Broadcast/Multicast Service; Architecture and functional description (Release 6).

DOI: 10.1002/wcm

11. Boni A, Launay E, Mienville T, Stuckmann P. Multimedia broadcast multicast service—technology overview and service aspects. *Fifth IEEE International Conference on 3G Mobile Communication Technologies (3G 2004)*. 634–638.
12. Ho J, Akyildiz I. Local anchor scheme for reducing signaling costs in personal communications networks. *IEEE/ACM Transactions on Networking* 1996; **4**(5): 709–725.
13. Romero J, Sallent O, Agusti R, Diaz-Guerra M. *Radio Resource Management Strategies in UMTS*. John Wiley & Sons: Chichester, 2005. ISBN-10 0-470-02277-9.

Authors' Biographies

Antonios Alexiou obtained his Diploma from the Department of Electrical and Computer Engineering of the Aristotle University of Thessaloniki (Greece) and his Master Degree from the Computer Engineering and Informatics Department of Patras University. He is currently a PhD Candidate of the Department of Computer Engineer and Informatics of Patras University. Furthermore, he is working as R&D Computer Engineer at the Research Unit 6 of the Research Academic Computer Technology Institute in Patras (Greece). His research interests include Mobile Telecommunications Networks, Multicast Routing, User Mobility in Cellular Networks, Congestion Control and Quality of Service, Mobile and Wireless Ad-hoc Networks. He has published three papers in journals and 16 papers in well-known refereed conferences.



Christos Bouras obtained his Diploma and PhD from the Computer Science and Engineering Department of Patras University (Greece). He is currently an Associate Professor in the above department. Also he is a scientific advisor of Research Unit 6 in Research Academic Computer Technology Institute (CTI), Patras, Greece. His research interests include Analysis of Performance of Networking and Computer Systems, Computer Networks and Protocols, Telematics and New Services, QoS and Pricing for Networks and Services, e-learning, Networked Virtual Environments and WWW Issues. He has extended professional experience in Design and Analysis of Networks, Protocols, Telematics and New Services. He has published 200 papers in various well-known refereed conferences and journals. He is a co-author of 7 books in Greek. He has been a PC member and referee in various international journals and conferences. He has participated in R&D projects such as RACE, ESPRIT, TELEMATICS, EDUCATIONAL MULTIMEDIA, ISPO, EMPLOYMENT, ADAPT, STRIDE, EUROFORM, IST, GROWTH and others. Also he is member of, experts in the Greek Research and Technology Network (GRNET), Advisory Committee Member to the World Wide Web Consortium (W3C), IEEE Learning Technology Task Force, IEEE Technical Community for Services Computing WG 3.3 Research on Education Applications of Information Technologies and W 6.4 Internet Applications Engineering of IFIP, Task Force for Broadband Access in Greece, ACM, IEEE, EDEN, AACE and New York Academy of Sciences.

DOI: 10.1002/wcm