Available online at www.sciencedirect.com

ScienceDirect

Computer Networks 52 (2008) 1521–1544

Computer Networks

www.elsevier.com/locate/comnet

Group registration with distributed databases for location tracking in 3G wireless networks

Dimitrios D. Vergados [a,b,*], Alexandros Panoutsakopoulos [a], Christos Douligeris [b]

[a] University of the Aegean, Department of Information and Communication Systems Engineering, Karlovassi, Samos GR-832 00, Greece
[b] University of Piraeus, Department of Informatics, 80 Karaoli and Dimitriou St., GR-185 34, Piraeus, Greece

Received 10 December 2005; received in revised form 15 January 2008; accepted 15 January 2008
Available online 1 February 2008

Responsible Editor: W. Wang

Abstract

The increase of subscribers in wireless networks has led to the need for efficient location tracking strategies. Location tracking is used to keep track of a Mobile Terminal (MT). The network retains the Registration Area (RA), where the MT last updated its location, so when an incoming call arrives for the MT, the network with the help of a location tracking strategy can find the area where the MT resides and then deliver the call. In this paper, we introduce a 2-level distributed database architecture combined with the *Group Registration* (GR) location tracking strategy to be used in 3G wireless networks. The GR strategy reduces the location management total cost, by updating the location of MTs in an RA with a single route response message to the HSS (Home Subscriber Server). More specifically, the IDs of the MTs newly moving into an RA are buffered and sent to the HSS for location updating in the route response message of the next incoming call to any MT in the RA. An analytical model is developed and numerical results are presented. It is shown that the GR strategy integrated with a 2-level distributed databases architecture in 3G networks can achieve location management cost reduction compared to costs of the distributed databases without the GR strategy and the GR strategy without distributed databases. Moreover, the proposed strategy results in small call delivery latency.
© 2008 Elsevier B.V. All rights reserved.

Keywords: Local anchor; Location area; Location management; Location tracking; Location update

1. Introduction

In order to overcome the demands of mobile users in mobile networks, a procedure has been created that helps mobile networks know where exactly (in the sense of an exact cell) a Mobile Terminal (MT) resides. This procedure is called Location Tracking (LT). When an incoming call arrives from a caller to a callee, the location tracking procedure takes place and the network is informed about the residence area of the caller and the callee subscribers. Thus, the network can create the routing path between them and deliver the call. In 3G wireless networks [1] there is only one database that records

* Corresponding author. Address: University of Piraeus, Department of Informatics, 80 Karaoli and Dimitriou St., GR-185 34, Piraeus, Greece. Tel.: +30 210 4142479.

E-mail addresses: vergados@aegean.gr, vergados@unipi.gr (D.D. Vergados), alexandros.panoutsakopouloss@googlemail.com (A. Panoutsakopoulos), cdoulig@unipi.gr (C. Douligeris).

1389-1286/$ - see front matter © 2008 Elsevier B.V. All rights reserved.
doi:10.1016/j.comnet.2008.01.010

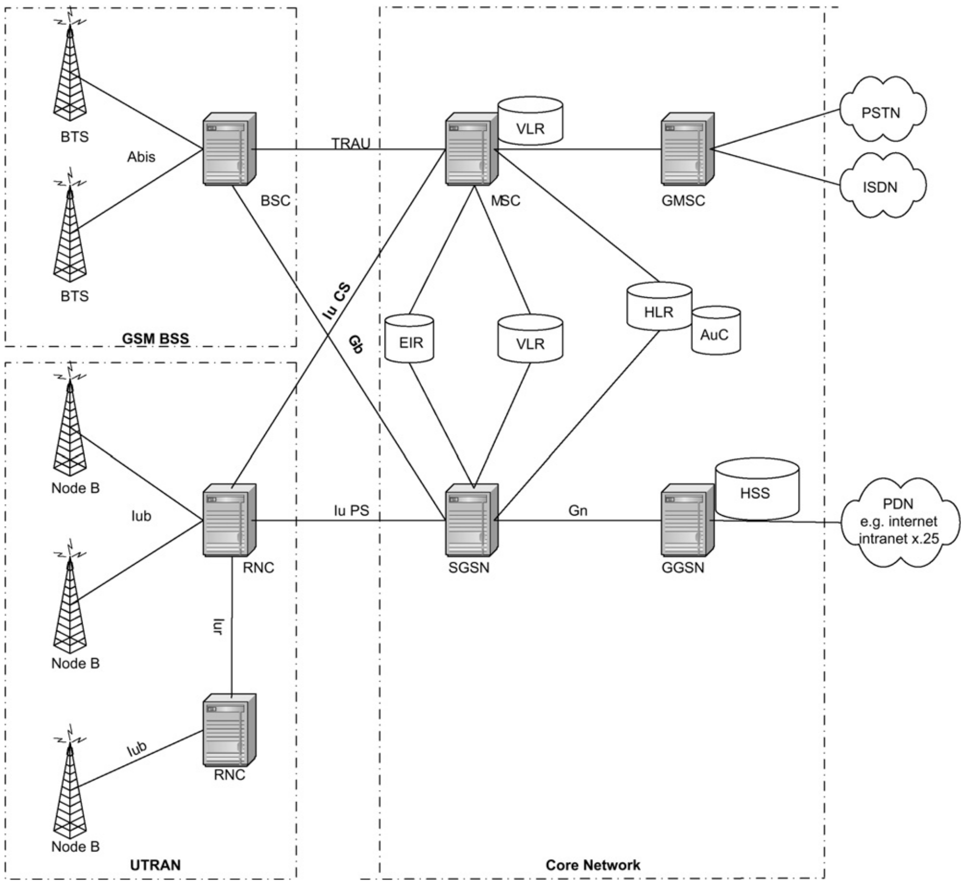

Fig. 1. The architecture of a 3G network.

the transactions between the Registration Areas (RAs) of the MTs (Fig. 1). This database is called Home Subscriber Server (HSS) and communicates directly with the GGSN (Gateway GPRS Support Node) and with the SGSN (Service GPRS Support Node). HSS, GGSN and SGSN exchange signaling messages only. Moreover, every GGSN can serve multiple SGSNs and every SGSN can serve multiple Radio Network Controllers (RNCs). The RNC provides the air interface (Iub) between the MTs and the core network and also the air interface (Iur) between RNCs.

In 3G wireless networks, the following strategy is employed for location tracking [1]. When an MT changes an RA which is controlled by an SGSN, a location registration request is sent to the HSS, which in turn updates the MTs service profile. Then the HSS records the new RA that now serves the MT and deletes the old RA that the MT used to reside in from its database. When an incoming call for the MT arrives, the HSS is queried to find out the current RA serving the MT. The HSS, after having found in which RA the callee is residing, creates the routing path between the caller and the callee. A call connection between the RA of the caller and the RA of the callee is then set up and the calling procedure ends.

There are two main differences between 3G/UMTS and GSM networks: (a) 3G networks use a strategy with only one database (HSS) to record the movements of the MTs, in contradiction to the GSM networks which use a strategy of more than one database communicating to each other (VLRs and HLR). As one may notice in a 3G network the HSS is engaged in every movement of the MTs. Due to the locality behavior of the MTs, in this paper (Section 4) a 2-level distributed databases architecture is introduced in order to reduce the processing that is necessary in the HSS database. Also, this architecture has been chosen because the 2-level distributed databases architecture suits better to the 3G networks architecture (Section 4) than the 3-level or 4-level ones, since 3G UMTS networks are most often divided in two levels: SGSNs–GGSN and RNCs,

(b) The neighboring RNCs have the ability to communicate directly with each other. On the other hand in GSM networks, there is no direct communication between neighboring VLRs and the communication is taking place only through the HLR.

The basic idea of our proposal is based on the scheme proposed by Mao and Douligeris [19], where the GR strategy is applied in GSM networks. The basic idea behind the GR strategy is to send location update requests to the root DB for multiple MTs in a single message to reduce the location update cost. Specifically, the IDs of the MTs newly moving into an RA are buffered and sent to the root DB for location update in the route response message of the next incoming call to any MT in the RA. According to this scheme the root DB is the DB attached to the HLR of a GSM network.

From the performance evaluation, it is shown that the proposed strategy has a remarkable good behavior compared with distributed databases (2-level architecture) without the GR strategy and the GR strategy without distributed databases in 3G networks, reducing significantly the total location management cost [2].

Apart from the remarkable good behavior that the proposed strategy shows, compared to the other strategies, the application of the GR strategy with distributed databases in 3G networks scales also well, as it has also been observed in GSM networks.

This paper is organized as follows: A description of signaling and tracking strategies in 3G/UMTS is presented in Section 2. In Section 3, a description of the proposed strategy is given, while a location management procedure with distributed databases is provided in Section 4. In Sections 5 and 6, the walk model and the performance analysis of the proposed strategy are presented. Finally, Section 7 concludes the paper.

2. Signaling and tracking strategies in UMTS

A UMTS network consists of three interacting domains (Fig. 1). The Core Network (CN), the UMTS Terrestrial Radio Access Network (UTRAN) and the User Equipment (UE) or mobile terminals.

2.1. Core Network

The Core Network is divided in two distinct separate domains, the circuit switched and the packet switched domains. Some of the circuit switched elements are: the Mobile services Switching Centre (MSC), the Visitor Location Register (VLR) and the Gateway MSC. Packet switched elements are the Serving GPRS Support Node (SGSN) and the Gateway GPRS Support Node (GGSN). Some network elements, like the EIR, the HLR, the VLR and the AUC are shared by both domains. The Asynchronous Transfer Mode (ATM) is defined for the UMTS core transmission. The ATM Adaptation Layer type 2 (AAL2) handles the circuit switched connection and the packet connection protocol, while AAL5 is used for data delivery. The architecture of the Core Network may change when new services and features are introduced. A Gateway Location Register (GLR) may be used to optimize the subscriber handling between network boundaries. The MSC, the VLR and the SGSN can merge to become a UMTS MSC.

2.2. Radio Access

The wide band CDMA technology has been selected for the UTRAN air interface. The UMTS WCDMA (Wireless Code Division Multiple Access) is a Direct Sequence (DS) CDMA system, where user data is multiplied with quasi-random bits derived from WCDMA spreading codes. In UMTS, in addition to canalization, codes are used for synchronization and scrambling.

2.3. User equipment

The UMTS standard does not restrict the functionality of the user equipment. Terminals work as an air interface counter part for Node-B and have many different types of identities. Most of these UMTS identity types are taken directly from the GSM specifications (see Table 1).

2.4. General Tracking Strategies

As mentioned before, the location tracking strategy imposes that the HSS must be accessed for each

Table 1
UMTS identity types

Types of IDs	Abbreviation
International Mobile Subscriber Identity	IMSI
Temporary Mobile Subscriber Identity	TMSI
Packet Temporary Mobile Subscriber Identity	P-TMSI
Temporary Logical Link Identity	TLLI
Mobile Station ISDN	MSISDN
International Mobile Station Equipment Identity	IMSEI
International Mobile Station Equipment Identity and Software Number	IMSEISV

RA change of every MT as well as for each incoming call for every MT. The HSS may become the bottleneck of a wireless network as the number of mobile users is increasing. In addition, a high volume of signaling traffic is generated by these procedures. To reduce the impact of location registration on mobile networks, a number of location tracking strategies have been proposed to reduce the network signaling load and the database burden. In this paper, we focus on location management strategies for location registration. These location strategies can be classified into:

1. Distance-based location update: A distance-based location update scheme is considered in [3,4]. When an incoming call arrives, cells are paged in a shortest-distance-first order such that cells closest to the cell where the last location update occurred are polled first. The delay in locating an MT is, therefore, proportional to the distance traveled since the location update. The authors have proposed an iterative algorithm that can generate the optimal threshold distance that results in the minimum cost.
2. Movement-based location update: The location update is performed whenever the mobile user completes d movements between cells. The value d is called the location update *movement threshold*. When an incoming call arrives, the system pages a location area including all cells within a distance d from the last registered location of the called mobile user [3,5–8].
3. Time-based location update scheme: The location update is performed every t units of time. The size of the location area can then be calculated as a function of the mobility of the mobile user in the scheme [3,9].
   Evaluation of the performance of the above schemes, has demonstrated that the distance-based scheme produces the best performance but its implementation incurs the highest overhead. For the time-based and the movement-based schemes, the MT has to keep track of the time elapsed and the number of movements performed, respectively since the last location update. This can be achieved by implementing a timer or a movement counter at the MT. The distance-based scheme, however, assumes that the MTs have knowledge of the distance relationship among cells. The network must be able to provide this information to each MT in an efficient manner.
4. Novel Tracking Strategy: The basic idea of this strategy is to allow the MT to transmit update messages only at specific cells (the number of which is small compared to the total number of cells in the network), while restricting every search for a mobile user to a small subset of cells. In this approach, a subset of the base stations called the reporting centers is selected among all base stations. The other cells are called non-reporting [10].
5. Hierarchical Architecture Strategy: This strategy is basically applied in MANs (Metropolitan Area Networks), in combination with a 3-level distributed database architecture. The three levels are: the access network MAN, the core network MAN and the MAN. Every mobile user has a home access MAN and a home MAN address. The tracking procedure is firstly performed at the local MA address of the caller and then at the home MA address of the caller [11].
6. Alternative Tracking Strategy: This strategy is based on the fact that the behavior of the MTs can be predicted. It is observed that an MT's user stays within certain geographic areas such as home, workplace, etc. for considerable time periods. The basic VLR/MSCs associated with these places are selected as the user's Local Anchors (LAs). In this strategy, at most two LAs, one within the office environment and another within the home environment of a certain user, are assumed [12–14].
7. Forwarding Pointer Strategy: Instead of reporting a location change to the HLR every time the MT moves to an area belonging to a different VLR, the report can be eliminated by simply setting up a forwarding pointer from the old VLR to the new VLR [15–17]. When a call for the MT is initiated, the network locates the MT by first determining the VLR at the beginning of the pointer chain and then follows the pointers to the current serving VLR of the MT. To minimize the delay in locating an MT, the length of the pointer chain is limited to a predefined maximum value K. When the length of the pointer chain reaches K, additional forwarding is not allowed. A location change must be reported to the HLR when the next movement occurs. Fig. 2 demonstrates the operation of the pointer forwarding. Pointers are set up from VLR1 to VLR2, from VLR2 to VLR3 and from VLR3 to VLR4 as the MT moves from MSC1 to MSC2 from MSC2 to MSC3 and from MSC3 to MSC4, respectively. For $K = 3$, the pointer chain

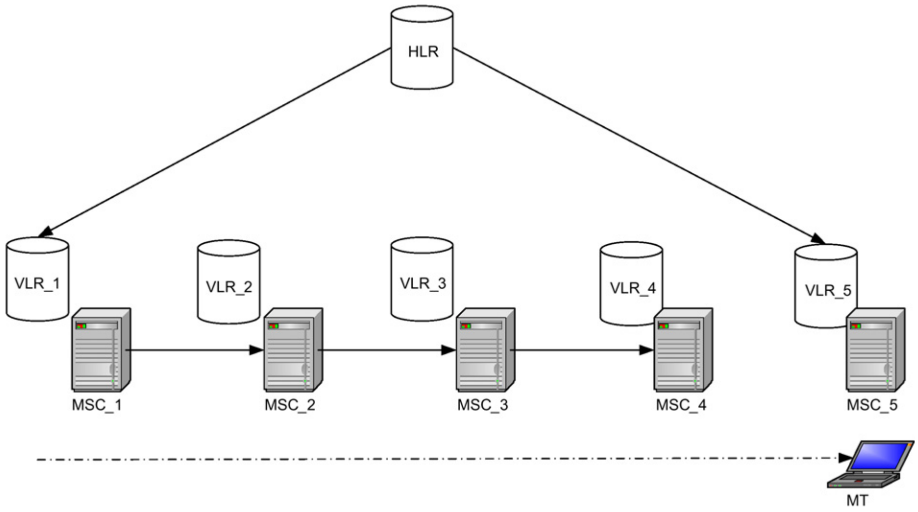

Fig. 2. Forwarding Pointer Strategy.

cannot be extended any further. An additional movement from MSC4 to MSC5 will result in a location registration at the HLR. The original pointers are deleted and the HLR records the ID of the current serving VLR of the MT. It is demonstrated that depending on the mobility and call arrival parameters and the value of *K*, this scheme may not always result in a reduction in cost as compared with the original IS-41 scheme [15].

8. Eager Cashing Strategy: The basic idea of this strategy is the buffering of the location of the callee into the database of the caller. So after the first call has taken place, when another call is created from the same caller to the same callee, the call is delivered immediately. A pair of bypass pointers *forward* and *reverse* are used. These pointers are created at databases *s*, *t*, respectively [11,18].
9. Local Anchoring proposal. In this scheme the signaling load is reduced by eliminating the need to report location changes to the HLR [22]. The VLR where the MT is roaming most of the time is the local anchor. The HLR stores the pointer into this VLR. The location of the MT is reported to the local anchor instead of the HLR. So, the registration cost may decrease because the location of the MT is not reported to the HLR. When an incoming call arrives, the HLR is queried. The output is a pointer to the local anchor. If the MT is located in the local anchor, the paging procedure starts. If the local anchor stores a pointer into a VLR, the MT is reached in the pointed VLR. Fig. 3 summarizes the local anchoring scheme. There are two different ways to use the anchor. In the static scheme, the VLR where the last incoming call is received becomes the local anchor for the called MT. In the dynamic scheme, the local anchor changes when an incoming call arrives, but the system can also decide to change the local anchor. This decision is based on the movements of the MT. Depending on the mobility and incoming call rates, this proposal may offer a better performance than the classical GSM-MAP or IS-41 scheme.

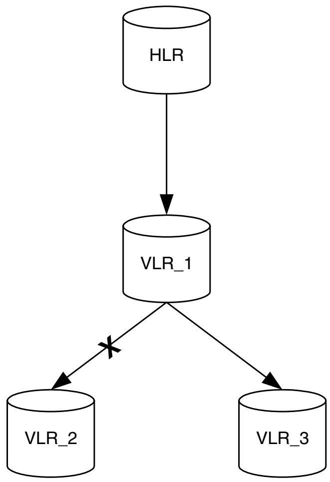

Fig. 3. Local Anchoring scheme.

2.5. General Categories of Tracking Strategies in 2.5G–3G networks

The strategies for location update in 2.5G–3G networks are described in the following subsections:

2.5.1. Static Strategies

In order to minimize the location management cost, location areas must be carefully designed. Analytical studies assumed that all cells have the same size and shape, all LAs have the same structure, and the movements of all MTs are homogenous [18]. Two mobility models are commonly used: the Markovian mobility model and the fluid flow mobility model. If a simplified fluid flow mobility model is used, it can be shown that

$$N_{\text{opt}} = \sqrt{\frac{vC_{\text{PG}}}{\pi RC_{\text{LU}}}}, \tag{1}$$

where N is the optimum number of cells per LA, R is the cell radius and v is the MT speed [20]. The PG (paging) cost per incoming call is denoted as C_{PG} and the LU (location update) cost as C_{LU}. An algorithm is proposed which takes into account the kind of cell and cluster pattern to evaluate its perimeter (the LA perimeter) [21].

The GSM-MAP standard uses a method that combines a time-based scheme and an LA-based policy. As we already know, when an MT moves to a new LA, it triggers an LU message. There is also a timer in the MT, which is reset by incoming calls and LA-based LU messages. When this timer expires, the MT sends an LU message. Time-based LU messages are refreshing packets for the system DB, used to avoid unnecessary paging.

2.5.2. Dynamic Strategies

Dynamic Management of the Registration Area: This proposal considers a mesh cell configuration allowing the evaluation of the optimal size of the registration area in order to reduce LU and PG costs [18]. The MT is located in a $k \times k$ cells registration area, where k is evaluated for each user according to its mobility pattern and incoming call rate. This method provides a better performance than a static LA-based scheme. However, it is not easy to implement this method. In order to page the MT, for example, the system DB has to store the centre cell of the registration area and the radius k for each user. Furthermore, for each user LU message received, the MT has to evaluate and send to the MT the new set of cells of its registration area for the MT to store them. In short, in order to evaluate MTs registration areas, the system has to store the cell layout in its DB.

Combining Movement-based and Distance-based schemes: The MT is located in a set of cells within a distance d of the centre cell [18]. The MT stores in its memory only the set of cells within a distance h to the centre cell. This set can be an empty set. It can include all the cells within a distance d to the centre cell, or it can be an intermediate option between the last two alternatives. It can be shown that $h \leqslant d$, and the movement threshold is $m = d - h + 1$. When the MT moves out of the set, the movement counter is reset. The location management signaling cost diminishes as the size of the local memory in the MT increases.

3. Description of the proposed strategy

A GSM network is designed according to a 2-level distributed database architecture, where in the upper level (level 1) a database (root database) is located that is attached to the HLR node and in the lower level (level 2) there are the databases that are attached to VLR nodes. An HLR may serve multiple VLRs and a VLR may serve multiple BSCs. Moreover a BSC node may serve one, two or even more geographical areas which are called RAs. The GSM architecture is shown in Fig. 4.

In the Group Registration strategy (GSM), an MT upon an RA change updates its location at a local anchor. In addition, the MTs ID is copied in a list called *Registration Waiting List* (RWL), which is maintained by the DB of the BSC serving the new RA. This list holds all newly moved-in MTs' IDs for group registration. As the next incoming or outgoing call associated with the RA occurs, the RWL is sent to the VLR for updating the location of all MTs whose IDs are in the RWL. As a result, the current RAs VLR becomes the new local anchor of those MTs. As an incoming call arrives, the VLR forwards the route request to the called MT's local anchor for call delivery. The same procedure is applied while an MT moves between RAs under different VLRs.

In the scenario mentioned above we can see that the whole registration procedure of an MT takes place under a registration area which is served by the same HLR, where only level 2 of the GSM architecture is used. In the case of a transaction of an MT between RAs served by different HLRs, the whole registration procedure is executed using both levels of the GSM architecture (level 1 and level 2). In that case the only difference with the previous scenario (registration procedure under a registration area which is served by the same HLR) is that the RWL is sent to the DB attached to the

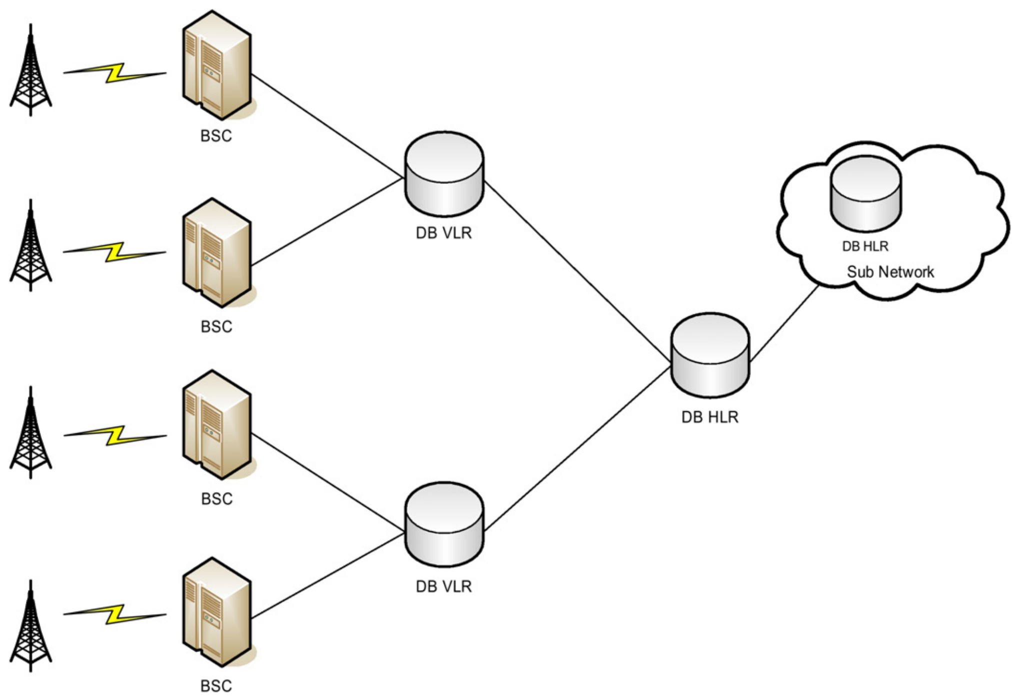

Fig. 4. The GSM architecture.

HLR, loading the network with extra registration messages.

In this paper, we propose a 2-level distributed database architecture integrated with the *Group Registration* (GR) strategy for 3G networks (Fig. 5), in order to combine location update requests for multiple MTs in one single message to the HSS. In such a way, the signaling cost of location registration is expected to reduce significantly.

The 2-level distributed databases architecture consists of two sets of databases (Fig. 5). The architecture has the form of a tree, where the root of the tree (level 0) is called DB_0 (HSS) and the databases of level 1 are called DB_1. Neighboring DB_1s communicate directly with each other, as well as with the DB_0 (HSS) because of the nature of 3G wireless networks (3G UMTS). For example in Fig. 5, only RNCs under the same SGSN are neighboring RNCs and the corresponding DB_1s have entries of the neighboring RNCs, as well. Every DB_1 is attached to an RNC and neighboring RNCs communicate with each other by means of the air interface (Iur). Thus, we observe that a number of DB_1s are grouped under one and unique DB_0. The routing protocol that is useful to exchange messages between the databases is the PDP (Packet Data Protocol) [23,24]. DB_0 maintains the service profile of all the mobile users residing in its region, and also the service profiles of MTs residing in other database subsystems (DS). Moreover, DB_0 maintains an import for each user it serves. This import in the DB_0 contains a pointer to another DB_0 that serves at that particular moment the user or a pointer to a DB_1 which is connected to the user that particular time moment.

Each DB_1 has a copy of the service profiles of the users, which reside provisionally in the region that covers DB_1, as well as in the two neighboring DB_1s of the neighboring RNCs. With this architecture, we avoid the frequent queries to DB_0 (HSS), because of the locality that is presented in user mobility and user call patterns. Of course, when a call or a location update is not local, then more databases, including the DB_0 will also be visited. If we apply, for example, the proposed 2-level distributed databases architecture (Fig. 5) along a motorway, there is no need for RNC_4 to communicate directly with RNC_1, because it belongs to a different SGSN and it communicates with a different RWL (Fig. 6). In that case, in order to cover the entire motorway, we may need a different network architecture, more GGSNs, SGSNs, and RNCs under the same SGSN and so on. The numbers of the nodes, which implement a network, vary due

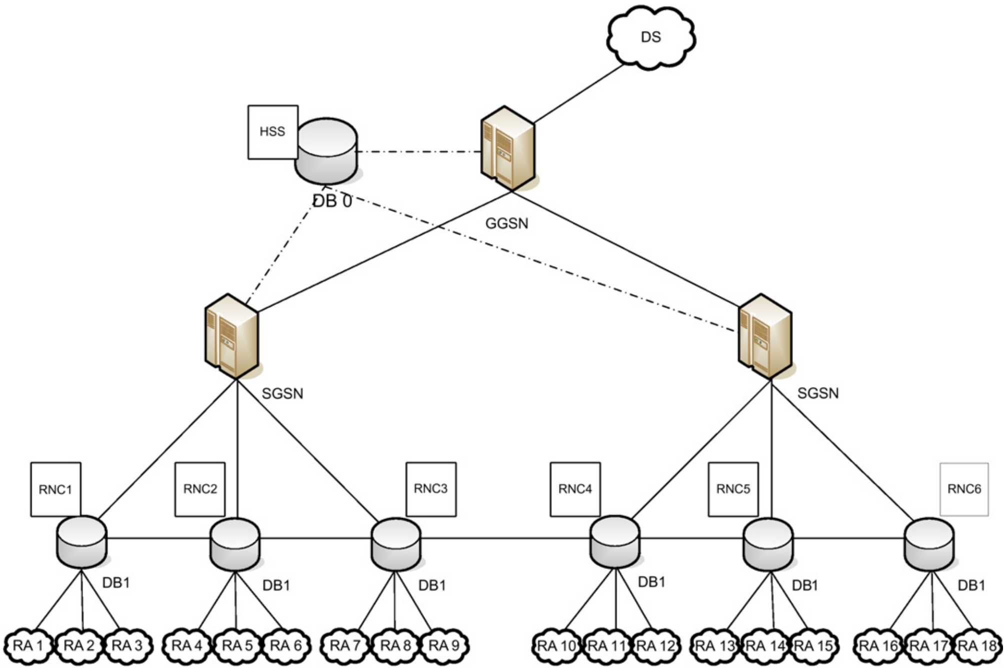

Fig. 5. The proposed 2-level distributed databases architecture in 3G wireless networks.

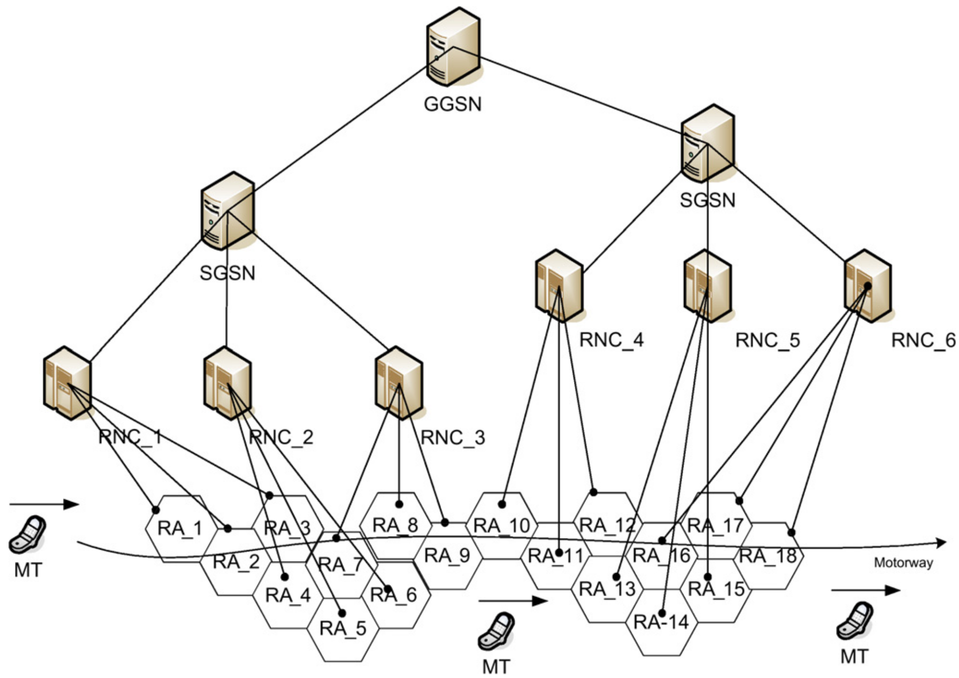

Fig. 6. The application of the proposed 2-level distributed databases architecture to a motorway.

to the needs of the users (bandwidth), the characteristics of the users (velocity, etc.), the geographic environment and the number of the users.

Also, every DB_1, in general, ought to have different capabilities (i.e. capacity, CPU speed, etc.), since it supports, different RAs. In our study, we assume that all DB_1s are the same, but without being a bottleneck to the performance of the entire network. In this paper, we assume a flat geographical environment, where the range of an RNC varies from 10 to 60 km. If we want to gain in capacity, but not in coverage area, a short RNC range may be used. If in the coverage area, the mean number of subscribers is small, then the use of long range RNCs is preferable. Therefore, the numbers of RNCs under the same SGSN may vary. A 3G network may support 60 RNCs under the same SGSN. In our study, a typical network architecture is used, which is applicable in a small country and it is applied to all strategies.

In the GR strategy (3G/UMTS), upon an RA change, an MT updates its location at a local anchor. In addition, the MTs ID is copied in a list called Registration Waiting List (RWL), which is maintained by the DB_1 of the RNC serving the new RA to hold all newly moved-in MTs IDs for group registration. As the next incoming or outgoing call associated with the RA occurs, the RWL is sent to the DB_0 (HSS) for updating the location of all MTs whose IDs are in the RWL. As a result, the current RA's RNC becomes the new local anchor of those MTs. As an incoming call arrives, the DB_0 (HSS) forwards the route request to the called MT's local anchor for call delivery.

The proposed GR strategy (3G/UMTS) has the following two characteristics:

- 3G networks apply only one database (HSS) to record the movements of the MTs. Thus, the HSS is engaged for every movement of MTs in a 3G network. In order to reduce the processing applied in the HSS, plus the locality of MTs movements (they used to move through RAs which are served from the same RNC), we introduce a 2-level distributed database architecture with the intention to reduce the tracking processing costs of the HSS by transferring them to the distributed databases.
- Neighboring RNCs have the ability to communicate directly with each other. So, during a transaction of an MT from one RA to another RA of the same RNC, the DB attached to the RNC is not engaged in the location tracking process. The DB has a record of the MT independently of the fact that the MT has changed RA. Thus, the number of location registration messages is reduced, the location registration cost is reduced and of course the total cost of the proposed strategy is also reduced.

4. Location update – call delivery procedures, and analytical analysis of the proposed strategy

In this section, the application of Group Registration in 3G networks is studied (Fig. 7). Also, the location update and the call delivery procedure are presented here. The location management and call delivery costs of the proposed strategy are calculated through an analytical model.

The basic idea behind the GR strategy (3G/UMTS) is to send location update requests to DB_0 (HSS) for multiple MTs in a single message to reduce the location update cost. Under each SGSN there is a RWL to keep the newly arrived MTs IDs. This RWL is located at the DB_1 of RNC_1 or RNC_2 or RNC_3 for the one SGSN or at the DB_1 of RNC_4 or RNC_5 or RNC_6 of the other SGSN (Fig. 5). However, before the location of the newly moved-in MTs is updated at DB_0 (HSS), a mechanism should be in place so that any incoming calls for these MTs can be successfully delivered to their current RA. For this purpose, either forwarding or local anchoring can be used to setup a forwarding pointer from an MTs old RA or local anchor to its new RA as the MT changes its RA.

Since local anchoring has lower call delivery latency than the location forwarding technique [22], we adopt local anchoring for call delivery in the proposed strategy.

A crucial factor that determines the performance of the proposed strategy is the time instance and the way to send the RWL to the DB_0 (HSS) for location updating. There could be different ways to send the RWL to the DB_0. The RWL could be sent to the DB_0 when the next incoming call into any MT in the RAs (under the same RNC) arrives and the RWL would then be piggybacked to the route request acknowledgement message. The RWL could also be sent to DB_0 upon the arrival of either the next incoming or outgoing call for any MT in the RAs (under the same RNC). For illustrative purposes, the first method is used for analysis in this paper. In Section 4.1, the location update and call delivery procedure of the proposed strategy are described.

4.1. Location update procedure

We consider that the DB_1s under the same SGSN have one RWL that is located at the DB_1 of RNC_1 or RNC_2 or RNC_3. The RWL has an entry of all the IDs of the MTs residing in the

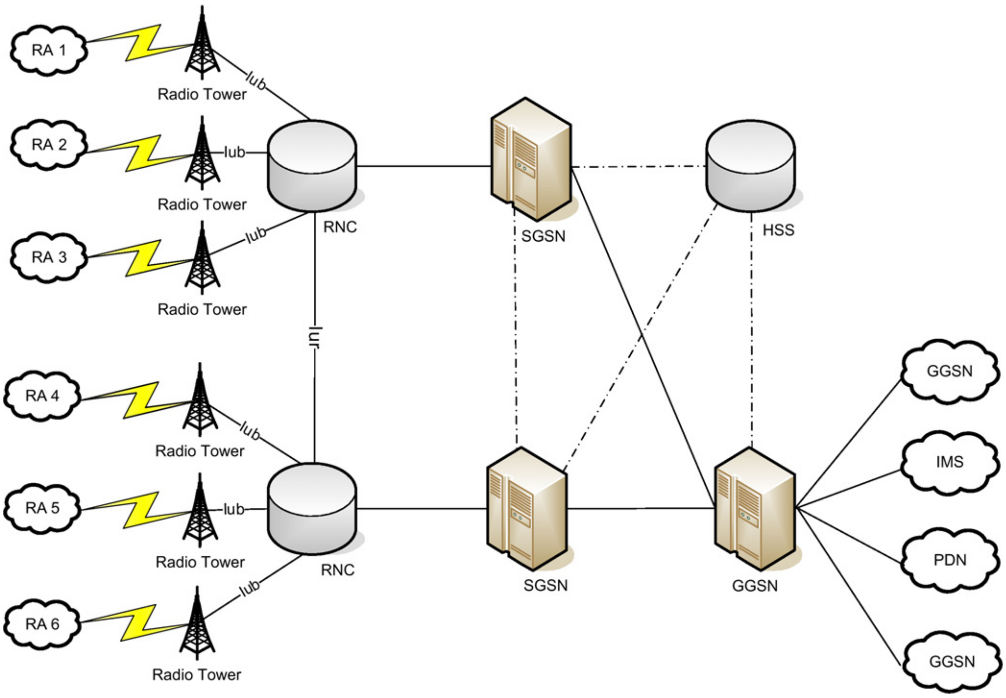

Fig. 7. The 3G UMTS architecture.

RAs of these DB_1s. Thus, in the proposed architecture (Fig. 5) there will be two RWLs, each one under an SGSN.

4.1.1. Under the same SGSN

If we assume that an MT moves from RA_x to RA_y located under the same SGSN, the registration steps are as follows (Fig. 8):

1. The MT moves from RA_x to RA_y. A registration request message, which is initiated by the MT, is sent to the DB_1 through RNC_y, of RA_y.
2. DB_1 is queried as well as the neighboring DB_1s. This DB_1 realizes that there is a registration of the MT in it, so the registration procedure stops.
3. DB_1 of RNC_x deletes the record of the MT and creates a pointer pointing to the DB_1 of RNC_y.

4.1.2. Under different SGSNs

We consider that an MT moves from RNC_x (starting point) to RNC_y (ending point) (Fig. 9).

1. A registration request message, which is initiated by the MT, is sent to the DB_1 through the RNC_y of RA_y. RNC_y detects an entrance of an MT to the RA that serves. Then it sends a message to the RNC_x, in order to inform it that the MT has left RA_x.
2. RNC_x sends an acknowledgement message to RNC_y, including the service profile of the MT. Then, it deletes the service profile of the MT from the DB_1 that RNC_x serves.
3. After RNC_y has received the acknowledgement message from RNC_x, it creates the service profile of the MT and imports the ID of that MT to the RWL of RA_y.
4. The DB_1 of RNC_x deletes the ID of the MT from the RWL. Then, RNC_x sends a location update message to DB_0, the local anchor of the MT. DB_0 is updated to point now to the DB_1 of the RNC_y.
5. DB_0 (local anchor) sends an acknowledgement message, which is initiated by the GGSN, to the old RNC of the MT and the location update procedure is completed.

4.1.3. Under different DB_0s

When a transaction of an MT takes place under different DB_0s, the application of the proposed strategy is not sufficient with the use of two RWLs under each SGSN. To resolve this issue, a separate RWL list for each DB_0 to store the IDs of those

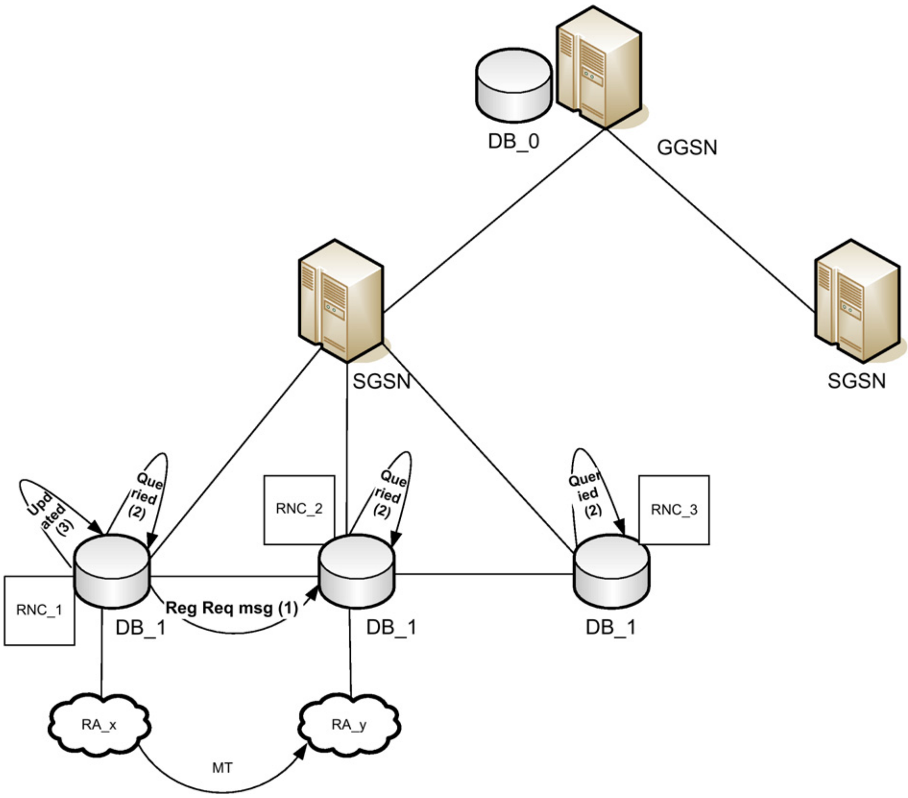

Fig. 8. Location update procedure under the same SGSN.

MTs that move between different DB_0s (different Database Subsystems) may be used.

While an MT moves from one Database Subsystem (DS) to another, the same GR strategy is applied and exactly the same location update and call delivery procedures follow. A DS consists of a GGSN, SGSNs and RNCs connected to each other as shown in Fig. 5. Two DSs communicate through the Gateway GPRS Support Nodes (GGSNs). For example, if an MT moves from a DS (the old one) to another (the new one), a registration request message, which is initiated from the MT, will be sent from the DB_0 of the old DS to the DB_0 of the new DS. The DB_0 of the new DS will be queried. Since there is no entry of the newly coming MT, the new DB_0 will be updated by recording the entrance of the MT, and it will send an acknowledgement message through the GGSN to the old DB_0. Moreover, a pointer will be created in the old DB_0 pointing to the new DB_0. If a call originates at an MT residing in the old DS to the MT newly moved in the new DS, the DB_0 of the old DS will be queried and it will inform the network that the called MT is residing in another DS. According to this procedure the network will be able to deliver the call. The architecture of Fig. 5 covers a very large geographical area. The users do not move from one DS to another DS frequently (i.e. when users roam from one country operator to another). So, because of the locality that the mobile users present (they usually follow a specific route, for example from their home to their office and backwards) the possibility to move to another DS is very small. For this reason, in this paper, we assume that an MT will always stay inside the Database Subsystem, as depicted in Fig. 5.

4.2. Call delivery procedure

Local anchor changes can also occur during the call delivery procedure. In other words, when an incoming call arrives at an RA, if the RAs RWL is not empty, the RWL is sent to DB_0 (HSS) for local anchor changes during the call delivery procedure. The call delivery procedure is exactly the same when a call originates between MTs located under the same or under different SGSN.

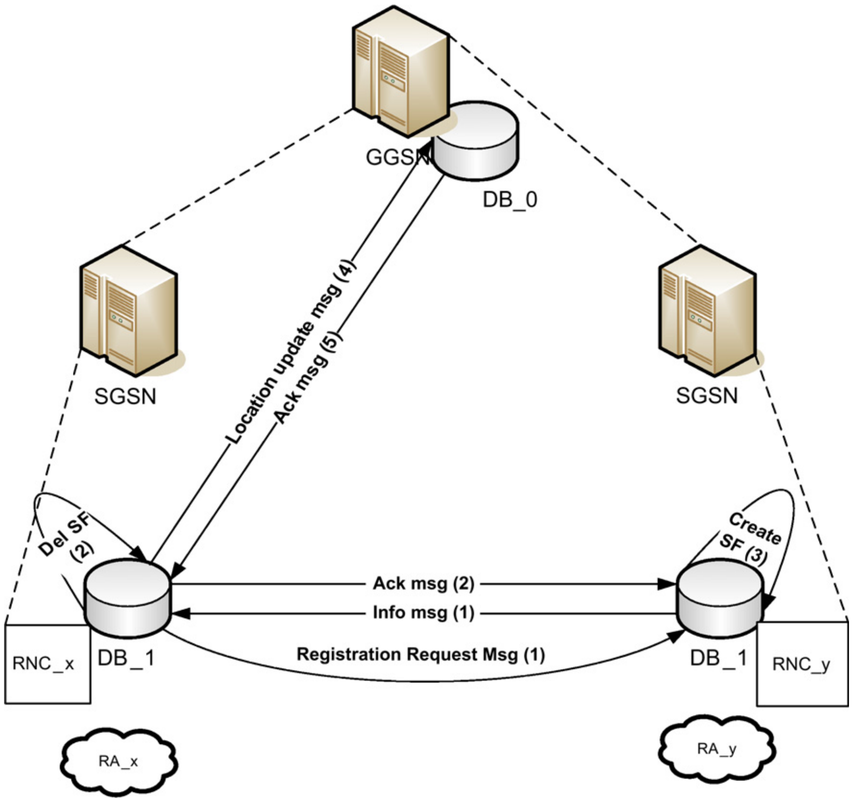

Fig. 9. Location update procedure under different SGSN.

The call delivery procedure proceeds as follows:

1. When a call for an MT originates (the origin of the call can be either from a wire line or from a mobile network), a location request message is sent to the MT's DB_0.
2. DB_0 queries itself in order to find which RA serves the MT. After having found the ID of the MT, the DB_0, through the GGSN, sends a route request message to the DB_1 of the serving RA.
3. If that RA is the local anchor of the callee, then the algorithm proceeds to the next step (step 4, Fig. 10). In case two, the route request message is forwarded to the DB_1 of the RNC that serves the MT at this time (local anchor, Fig. 11). So, the RA that forwarded the route request message to the new RA deletes the pointer that pointed to the RA that now serves the MT.
4. The callee's RNC searches for the MT. If the MT is found, a Temporary Local Directory Number (TLDN) is allocated to the MT.
5. If the RWL of the called MT's current RA is not empty, all MT IDs in the RWL are sent to the DB_0 in the route response message along with the TLDN. The route response message is initiated by RNC_x to the GGSN.
6. Meanwhile, except from the called MT, for every other MT in the RWL, the DB_1, through RNC_x, of the called MT sends a deregistration message to its local anchor, which removes the MT's forwarding pointer entry. Afterwards the RWL is emptied. So, the DB_1 of the called MT becomes the local anchor of all MT's in the RWL.
7. GGSN of DB_0 after having received the route response message from the DB_1 of the called MT, it forwards the TLDN to the calling RNC. If the route response message contains any of the MTs' IDs, DB_0 changes these MTs' local anchors to the current RNC.
8. After having received the TLDN, the calling RNC can set up the connection between the caller and the callee.

4.3. Location update cost of the proposed strategy

Consider that the MTs arrive in an RA following a Poisson process, that the incoming calls arriving at

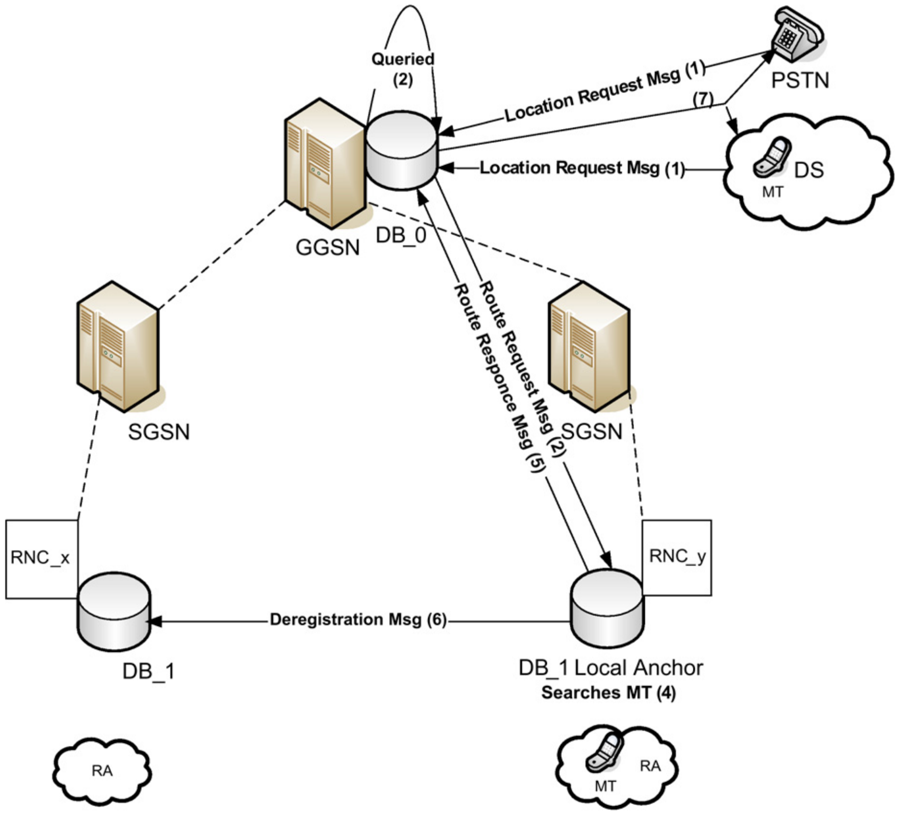

Fig. 10. The call delivery procedure (case 1).

an MT also follow a Poisson process and that the MTs' residence time in an RA follows an exponential distribution. In order to evaluate the total cost of the proposed strategy, the following cost notations are introduced:

- C_u cost for a query or an update to DB_1. (RNCs).
- C_h cost for a query or an update to DB_0.
- C_{uu} cost for transmitting a signaling message between two DB_1s (RNCs).
- C_{hu} cost for transmitting a signaling message between DB_0 and DB_1 (RNCs and GGSN).
- p_1 the probability that an MT that moves from an RA to another, to be under the same SGSN.
- p_2 the probability that a DB_1 that serves a certain m_0 (the name of the MT), not to be its local anchor.

As it was presented in Section 4.A, in the proposed strategy a location update is performed when an MT changes its RA. However, the location update procedure can be divided in two cases:

Case one: An MT moves from one RA to another under the same SGSN. Then the location update cost C_{R1}, is given by the following equation:

$$C_{R1} = 4C_u + C_{uu}. \tag{2}$$

Without losing the generality of our scheme, the exact expression of Eq. (2) for n nodes, depends on how the DBs are connected. Thus, if n nodes are connected in a continuous chain form, the location update cost is given by: $C_{R1} = (n+1)C_u + C_{uu}$. In our case, for $n = 3$, we obtain the Eq. (2).

Case two: An MT moves from one RA to another under different SGSNs. Then the location update cost C_{R2}, as well as the total location update cost C_R, are given by the following equations:

$$C_{R2} = 2(C_u + C_{hu}) + 3C_{uu} + C_h, \tag{3}$$

$$C_R = p_1 C_{R1} + (1 - p_1)C_{R2}. \tag{4}$$

Furthermore, we calculate the expected location update cost per call arrival for m_0. Let $f(y_0)$ be the density function of m_0's RA residence time with mean $1/\lambda_0$ and $g(x_0)$ be the density function of m_0's inter-call arrival time with mean $1/\mu_0$, i.e.

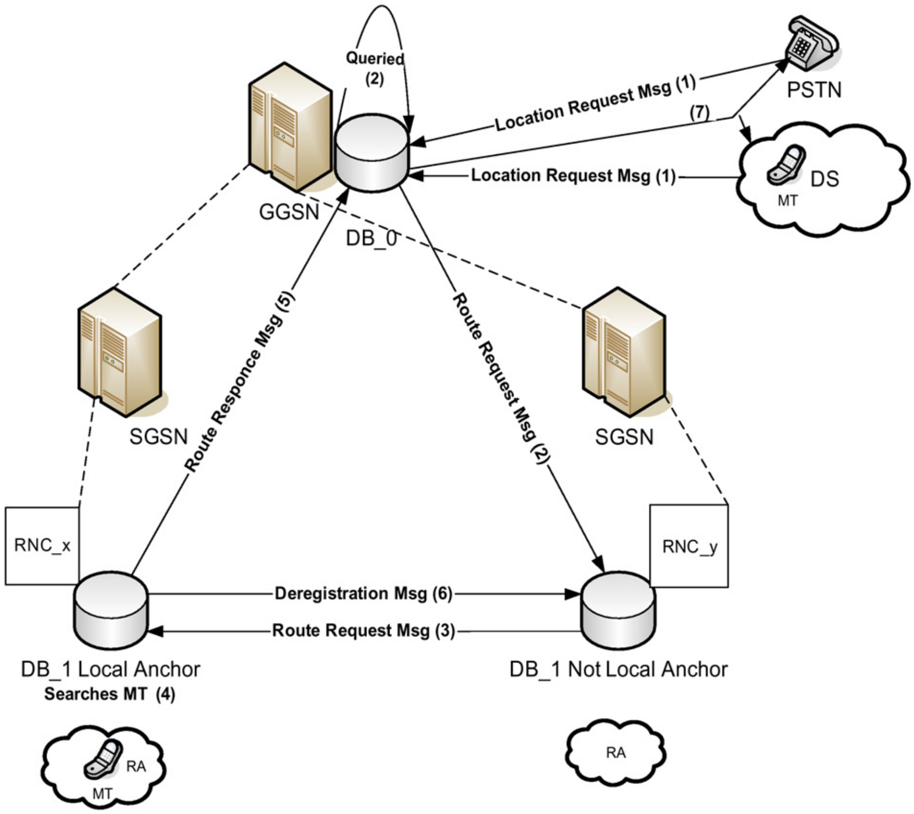

Fig. 11. The call delivery procedure (case 2).

$$f(y_0) = \lambda_0 e^{-\lambda_0 y_0} \tag{5}$$

and

$$g(x_0) = \mu_0 e^{-\mu_0 x_0}. \tag{6}$$

The probability of n boundary crossings that occur between two call arrivals is given by [19]

$$a(n) = \left\{ \begin{array}{ll} 1 - \frac{1}{p}[1 - f^*(\mu_0)], & n = 0 \\ \frac{1}{p}[1 - f^*(\mu_0)]^2 [f^*(\mu_0)]^{n-1}, & n > 0 \end{array} \right\}, \tag{7}$$

where $p = \frac{\mu_0}{\lambda_0}$ is the call-to-mobility ratio (CMR) of m_0 and $f^*(s)$ is the Laplace–Stieltjes transform of $f(y_0)$ i.e.

$$f^*(s) = \frac{\lambda_0}{s + \lambda_0}. \tag{8}$$

Then the expected location update cost $C_{R,G}$ per call arrival is given by

$$C_{R,G} = C_R \sum_{n=1}^{\infty} n a(n). \tag{9}$$

Furthermore, some additional notations should be introduced, in order to better describe the analysis model.

θ is the probability of at least one MT be in the RWL, when an incoming call arrives

$$\theta = \frac{\eta}{\eta + \sum_{i=1}^{M} \Phi_i}. \tag{10}$$

R_L is the expected number of MTs in the RWL:

$$R_L = \frac{\eta}{\sum_{i=1}^{M} \Phi_i}. \tag{11}$$

γ is the average number of MTs in the RWL (excluding the called MT)

$$\gamma = R_L - \frac{\eta}{\eta + \sum_{i=1}^{M} \Phi_i}, \tag{12}$$

where Φ_i is the arrival rate of the incoming calls to MTs. Φ_i is uniformly distributed over $(0.2, 3)$ [19].

M is the number of the MTs registered in the RAs under the same RNC.

η is the arrival rate of MTs in the RAs under the same RNC.

4.4. Call delivery cost of the proposed strategy

Case one: Since m_0 moves to an RA, in which the first incoming call is for another MT, then DB_0 will point to the DB_1 that serves m_0, when an incoming call arrives for m_0. In that case, the call delivery cost C_{C1} is given by

$$C_{C1} = (3+\theta)C_{hu} + C_h + C_u + R_L(C_u + C_{uu}). \qquad (13)$$

From Fig. 10 it can be seen that the factor $3^* C_{hu}$ is the sum of 2, 5, 7 messages and the factor $\theta^* C_{hu}$, is the RWL that is sent with the route response message when a call is originating to an MT which has an entry in the RWL (RWL is not empty).

Case two: When an MT receives the first incoming call in an RA after the MT moved into that RA, the MT's DB_0 is still pointing to the MT's local anchor, and the call delivery cost C_{C2} is given by:

$$C_{C2} = (3+\theta)C_{hu} + \gamma(C_u + C_{uu}) + C_h + C_u. \qquad (14)$$

Then, the total call delivery cost C_C will be the sum of C_{C1} and C_{C2}:

$$C_c = (1-p_2)C_{C1} + p_2 C_{C2}, \qquad (15)$$

where p_2 is the probability that a DB_1 that serves m_0, and DB_0 it is not the local anchor of m_0 (as calculated in [19])

$$p_2 = \prod_{n_5=1}^{N_5} \frac{\mu_0}{\mu_0 + \mu_{n_5,5} + \lambda_{n_5,5}} \prod_{n_6=1}^{N_6} \frac{\lambda_{n_6,6}}{\mu_0 + \mu_{n_6,6} + \lambda_{n_6,6}} \\ x \prod_{n_7=1}^{N_7} \frac{\mu_0}{\mu_0 + \mu_{n_7,7} + \lambda_{n_7,7}} \prod_{n_8=1}^{N_8} \frac{\lambda_{n_8,8}}{\mu_0 + \mu_{n_8,8} + \lambda_{n_8,8}} \\ x \frac{\mu_0}{\mu_0 + \lambda_0}, \qquad (16)$$

where

$\lambda_{n_i,i}$ $(i = 1,2,\ldots,8)$: are uniformly distributed over (0.1, 3)
$\mu_{n_i,i}(i = 1,2,\ldots,8)$: are uniformly distributed over (0.2, 3).

After the first incoming call, the cost for every other incoming call to m_0, since it stays in the same RA, is given by Eq. (13). Thus, the expected cost, $C_{C,G}$ per call arrival is

$$C_{C,G} = \begin{cases} C_C, & p \leqslant 1, \\ \frac{1}{p}(C_C + [p_2 - 1]C_{C1}), & p > 1, \end{cases} \qquad (17)$$

Therefore, the total cost per call arrival for the proposed strategy, denoted by C_{TOTAL} is

$$C_{\text{TOTAL}} = C_{R,G} + C_{C,G}. \qquad (18)$$

4.5. Total Cost of the GR Strategy without Distributed Databases in 3G UMTS networks

In this subsection, we compute the location update cost and the call delivery cost of the GR strategy without Distributed Databases in 3G/UMTS networks. The sum of these two costs gives the total cost of the proposed strategy.

4.5.1. Location update cost

The location update procedure is the same when the MT transacts under the same or under different SGSN (Fig. 12). In order to compute the equation of the location update cost, Fig. 12 must be used. In this figure there will be only one database, the DB_0. The DB_1s do not exist. Whenever an MT transacts from an RA to another, independently of whether these RAs are under the same or under different SGSNs, the DB_0 will always be queried. So, the location update cost is given by:

$$C_R = 2C_{uu} + 2C_{hu} + C_h. \qquad (19)$$

From Fig. 12 it can be seen that the location update cost (Eq. (19)) is the sum of messages 1, 5 ($2C_{uu}$), plus the sum of messages 2, 4 ($2C_{hu}$) plus the cost for updating the DB_0 (message number 3). In both cases, under the same and different SGSNs, the cost is the same.

4.5.2. Call delivery cost

For the computation of the call delivery cost the Local Anchor procedure is not applied, since there is only one Database which records the transactions of the MTs. The call delivery cost can be easily computed by

$$C_C = (4+\theta)C_{hu} + C_h + C_u. \qquad (20)$$

From Fig. 13 it can be seen that the call delivery cost (Eq. (20)) is the sum of the messages 3, 5, 6, 7 (messages exchanged between DB_0 and DB_1) plus $\theta^* C_{hu}$, which is the route response message that it is sent also with the RWL, when a call is originating to an MT, which has an entry in the RWL (RWL is not empty), plus message 2 (update of DB_0), plus message 4 (update of DB_1).

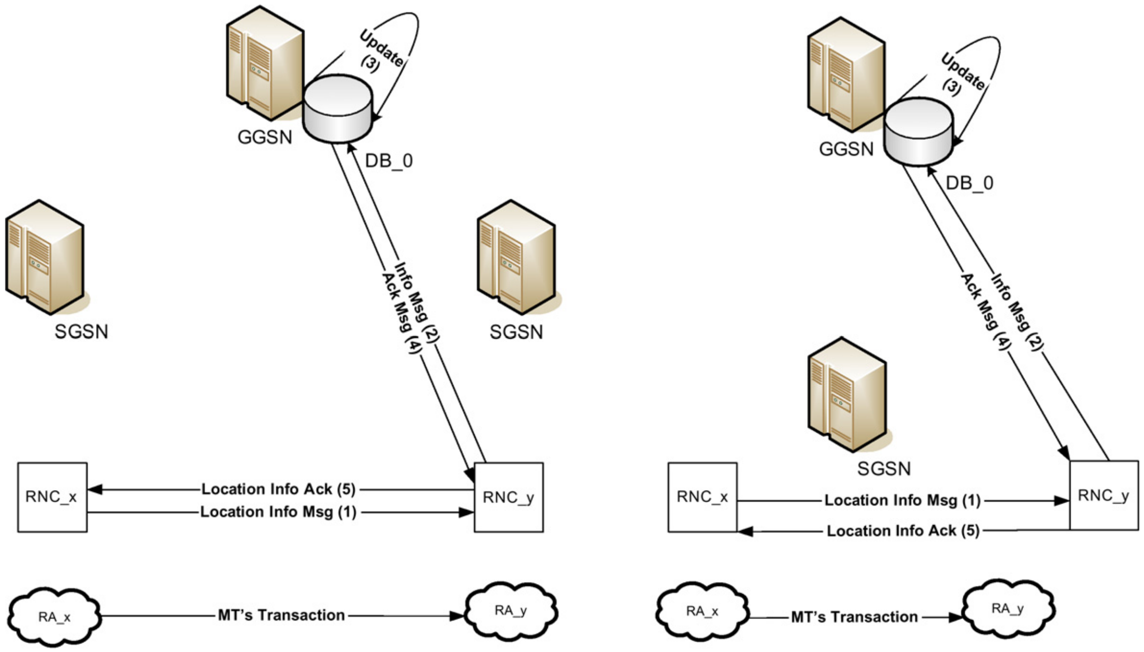

Fig. 12. Location update procedure of GR strategy without distributed databases.

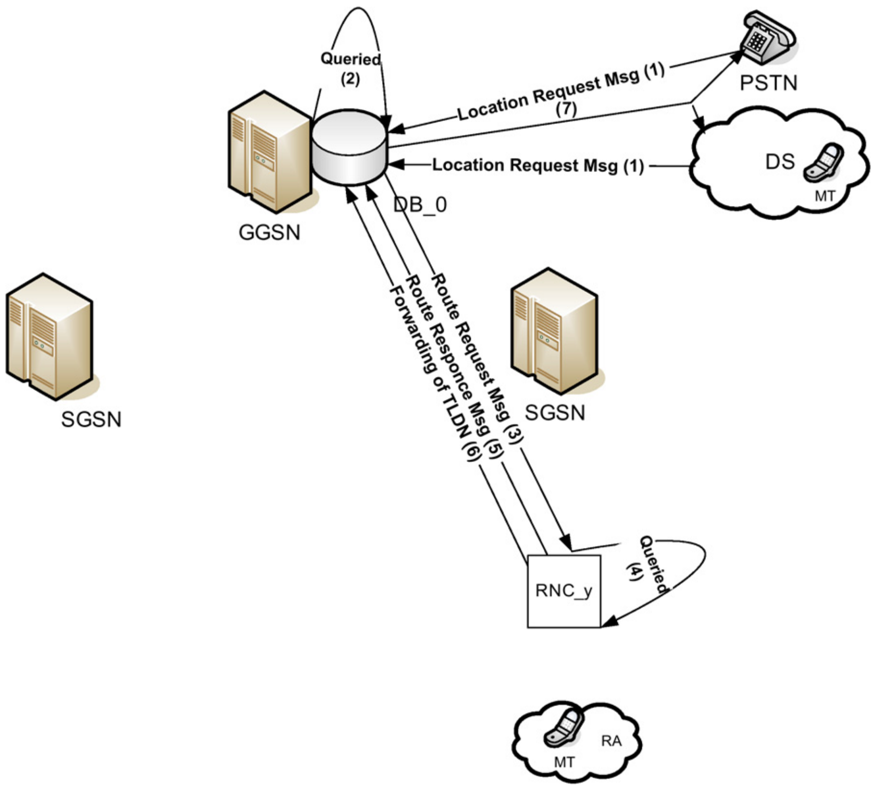

Fig. 13. Call delivery procedure of GR strategy without distributed Databases.

The total cost of our scenario can be computed as

$$C_{\mathrm{TOTAL}} = C_{R,G} + C_C, \tag{21}$$

where $C_{R,G}$ is computed by 9 substituting the C_R value from Eq. (19) and C_C is given by 20.

4.6. Total cost of Distributed Databases (2-level architecture) without the GR strategy in 3G UMTS networks

In this subsection, we will compute the total cost of the proposed strategy mentioned above. The total cost equation derives from the sum of the location update cost and the call delivery cost.

4.6.1. Location update cost

Since we apply a 2-level distributed databases architecture the equations that compute the total location update cost are the same with Eqs. (2) and (3) (derived from Figs. 9, 10). Moreover, the GR factors (θ, R_L) of Section 4.3 are not taken into account for the computation of the Eqs. (2) and (3). Thus, the total location update cost of distributed databases scheme without the GR strategy, is computed by Eq. (4), which is the weighted sum of Eqs. (2) and (3).

4.6.2. Call delivery cost

The only difference between the call delivery cost computed in this subsection with the one in Section 4.4 is that in Section 4.4 the GR strategy is applied. The computation of the call delivery cost is divided into two cases:

Case one: Since m_0 moves to an RA, in which the first incoming call is for another MT, then DB_0 will point to the DB_1 that serves m_0, when an incoming call for m_0 arrives. In that case, the call delivery cost C_{C1} is given by

$$C_{C1} = 4C_{hu} + 2C_h + C_u + C_{uu}. \tag{22}$$

Case two: When an MT receives the first incoming call in an RA after the MT has moved into that RA, the MT's DB_0 is still pointing to the MT's local anchor, and the call delivery cost C_{C2} is given by

$$C_{C2} = 4C_{hu} + 2C_h + 3C_{uu} + C_u. \tag{23}$$

The total call delivery cost will be the sum of C_{C1} and C_{C2}

$$C_C = (1 - p_2)C_{C1} + p_2 C_{C2}. \tag{24}$$

The total cost per call arrival for the proposed strategy, denoted by C_{TOTAL} is given by

$$C_{\mathrm{TOTAL}} = C_{R,G} + C_{C,G}, \tag{25}$$

where $C_{R,G}$ and $C_{C,G}$ are defined in Eqs. (9) and (17).

5. The random walk model

In the proposed walk model, the architecture of our scenario consists of four SGSNs. Under every SGSN three DB_1s and every DB_1 that serves three RAs are located. While an MT is within an RA, it chooses one of the eight directions with the same probability (Fig. 14). That probability is equal to 1/8. Moreover, we assume that the MT keeps moving in the same direction until it reaches a distance equal to the given moving step. Then it changes its direction randomly. We also assume that the probability that an MT that moves from an RA to another, to be under the same SGSN is $p_1 = 3/4$, while the probability to change SGSN is $1 - p_1 = 1/4$.

In the proposed walk model, the RA is a square with side $a = 50$ km. Thus the radius of the RA is $R = 25\sqrt{2}$ km. The residence time of an MT in an RA is exponentially distributed with mean value $1/\mu$.

5.1. Schematic description of the RAs under SGSNs

The white SGSNs in Fig. 15 are of a different Database Subsystem (DS) than the one under study. The only reason that we have added them in Fig. 15 is to show the borders of our DS. The SGSNs enclose the RAs shown below:

SGSN_1 = {RA1, RA2, RA3, RA7, RA8, RA9, RA13, RA14, RA15}.
SGSN_2 = {RA4, RA5, RA6, RA10, RA11, RA12, RA16, RA17, RA18}.

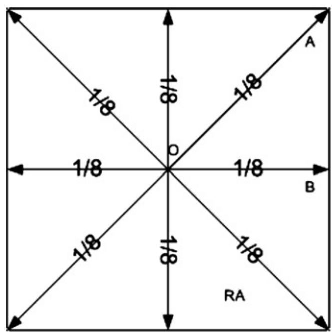

Fig. 14. The probability that an MT chooses one of the eight directions in an RA.

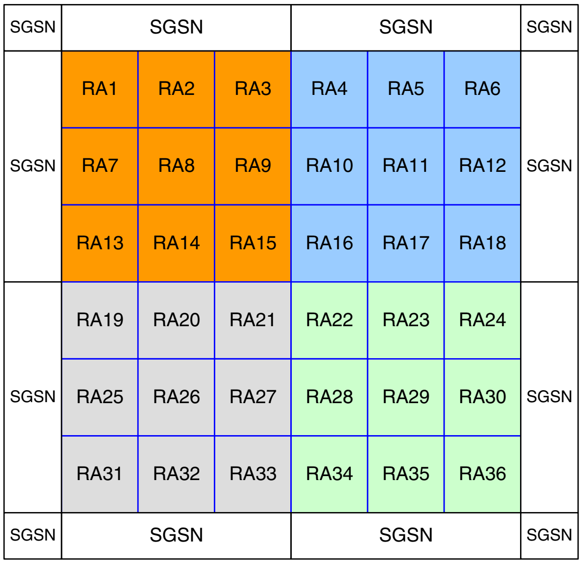

Fig. 15. Schematic description of the proposed DS.

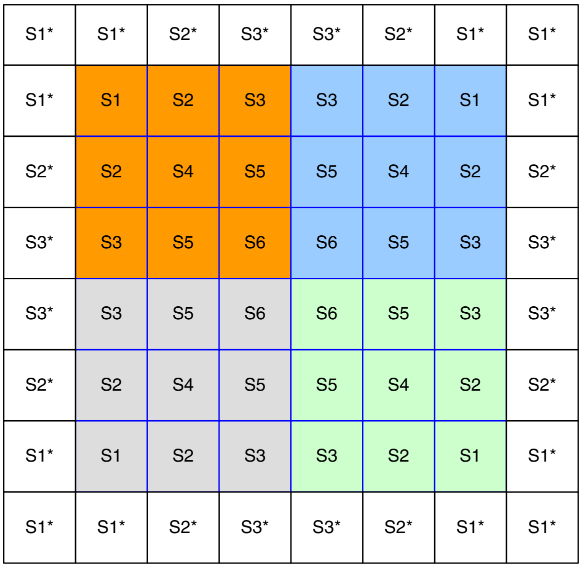

Fig. 16. Schematic description of aggregate states.

SGSN_3 = {RA19, RA20, RA21, RA25, RA25, RA27, RA31, RA32, RA33}.
SGSN_4 = {RA22, RA23, RA24, RA28, RA29, RA30, RA34, RA35, RA36}.

5.2. States definition

As we mentioned in Section 4.1, an MT moves from an RA to another with probability 1/8. In some of the RAs, with reference to the DS, the mobile user will have identical movement patterns to other RAs within the same DS and to RAs in adjacent DSs. For example, from RAs numbered 1, 6, 31, and 36 in Fig. 15, the mobile user movement pattern and transition probabilities will be identical but rotated by 90^0. The rotation of the movement pattern will not affect the studies to be conducted using Markov chains. Hence, these four RAs can be grouped into one state in the Markov chain model. This property of lumping the states was carried out to reduce the total number of computational states in the mathematical analysis. After the aggregation process, the following state aggregation is obtained [23] (Fig. 16):

$\mathbf{S}_1 = \{$RA1, RA6, RA31, RA36$\}$
$\mathbf{S}_2 = \{$RA2, RA5, RA7, RA12, RA25, RA32, RA30, RA35$\}$
$\mathbf{S}_3 = \{$RA3, RA4, RA33, RA34$\}$
$\mathbf{S}_4 = \{$RA8, RA11, RA26, RA29$\}$
$\mathbf{S}_5 = \{$RA9, RA10, RA14, RA17, RA20, RA23, RA27, RA28$\}$
$\mathbf{S}_6 = \{$RA15, RA16, RA21, RA22$\}$

In Fig. 16, the states $\mathbf{S}_1^*$, $\mathbf{S}_2^*$, $\mathbf{S}_3^*$, are the states of SGSNs of other DSs.

5.3. States-transition diagram and transitions matrix

After having defined the states where an MT can be located and the probability of an MT to move to a different RA, the next step is to define the state-transition diagram (Fig. 17).

Furthermore, in the simulation we assumed that the MT was moving in the schematic description of the proposed DS (Fig. 15), for three hours, since we want our results to be as reliable as possible. Also, we want the conditions of the moving pattern to be as close as it can be to the conditions of a real moving pattern. During that time frame the MT committed fifteen RA transactions and four SGSN transactions.

For illustrative purposes, we assume that every time an MT reaches a distance equal to $0.6R$ (moving step), it changes direction by choosing randomly an angle φ over $[0, 2\pi]$.

The coordinates of the current location of a user are given by the equations

$$x = 0.6R\cos\phi,$$
$$y = 0.6R\sin\phi. \tag{26}$$

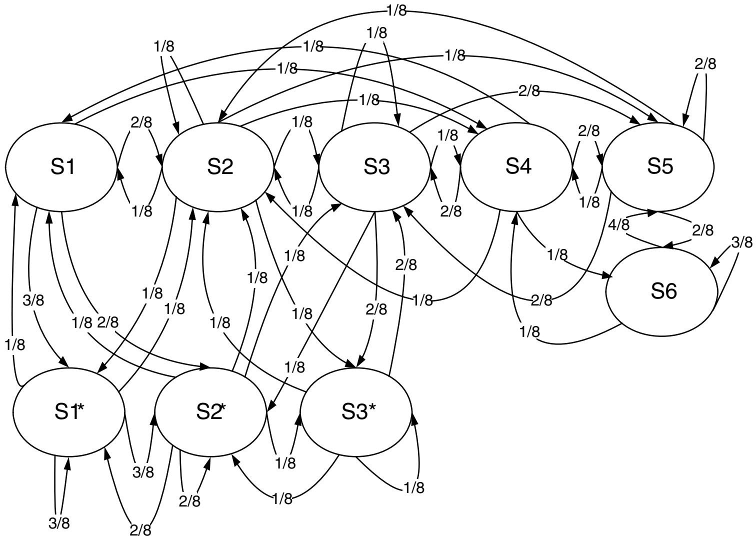

Fig. 17. States-transition diagram.

The mean residence time λ_0 of an MT in an RA is

$$\lambda_0 = \frac{180(\mathrm{min})}{15(\mathrm{RAs})} = 12(\mathrm{min})/(\mathrm{RA}). \tag{27}$$

The mean velocity value $\overline{U}$ of an MT is given by

$$\overline{U} = \frac{0.6R}{\lambda_0} = \frac{15\sqrt{2}\ (\mathrm{km})}{0.2\ (\mathrm{h})} \cong 106\ (\mathrm{km/h}). \tag{28}$$

At this point, it is worthwhile to mention that in the transition probability matrix (Table 2), we observe that the sum of the elements of the ninth line is not equal to 1. This happens because we study a particular Database Subsystem. As we can understand, there will be transactions from **S3*** states to other states (i.e. **S4***, **S5*** or **S6***). These transactions are taking place in different database subsystems and thus they are not inscribed in the transition probability matrix.

Table 2
Transitions probability matrix

$P_{[s_i,s_j]} =$

	S1	S2	S3	S4	S5	S6	S1 *	S2 *	S3 *
S1	0	$\frac{2}{8}$	0	$\frac{1}{8}$	0	0	$\frac{3}{8}$	$\frac{2}{8}$	0
S2	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{1}{8}$	0	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{1}{8}$
S3	0	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{2}{8}$	0	0	$\frac{1}{8}$	$\frac{2}{8}$
S4	$\frac{1}{8}$	$\frac{2}{8}$	$\frac{2}{8}$	0	$\frac{2}{8}$	$\frac{1}{8}$	0	0	0
S5	0	$\frac{1}{8}$	$\frac{2}{8}$	$\frac{1}{8}$	$\frac{2}{8}$	$\frac{2}{8}$	0	0	0
S6	0	0	0	$\frac{1}{8}$	$\frac{4}{8}$	$\frac{3}{8}$	0	0	0
S1 *	$\frac{1}{8}$	$\frac{1}{8}$	0	0	0	0	$\frac{3}{8}$	$\frac{3}{8}$	0
S2 *	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{1}{8}$	0	0	0	$\frac{2}{8}$	$\frac{2}{8}$	$\frac{1}{8}$
S3 *	0	$\frac{1}{8}$	$\frac{2}{8}$	0	0	0	0	$\frac{1}{8}$	$\frac{1}{8}$

6. Performance analysis and evaluation

The scope of this paper is to estimate the total location management cost of the MTs according to the proposed strategy and to ascertain the variation between the total cost and the Call-to-Mobility Ratio (CMR). In this section, we study the performance of the proposed strategy based on numerical examples. First, several critical parameters of the proposed strategy, i.e. p_2 and θ, are evaluated and their impact on the proposed strategy is discussed. Then, the proposed strategy is compared with the GR strategy without distributed databases and with the distributed databases scheme (2-level architecture) without the GR strategy. Some issues about the comparison of these strategies are discussed, in order to realize how efficient the proposed strategy is. The simulations in this section are run using the walk model of Section 5. Please, also note that in Figs. 16–20, it is assumed that the $\lambda_{n_i,i}$ $(i = 1,2,\ldots,8)$ are uniformly distributed over (0.1, 3) and the incoming call arrival rates Φ_i $(i = 1,2,\ldots,\mathrm{M})$ to all MTs in the RAs under the same RNC are uniformly distributed over $(0.2, 3)$, which implies that the $\mu_{n_i,i}$ $(i = 1,2,\ldots,8)$ are also uniformly distributed over $(0.2, 3)$.

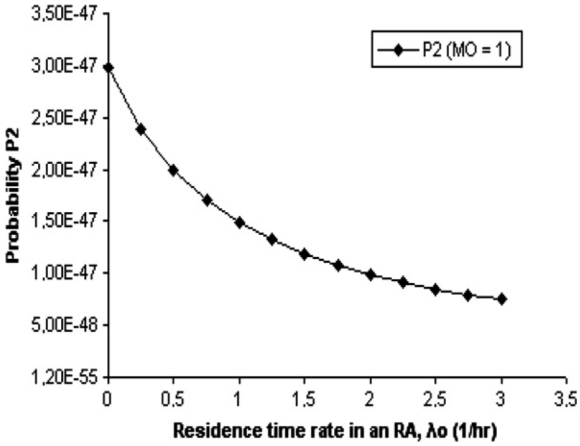

Fig. 18. Probability p_2 (proposed strategy).

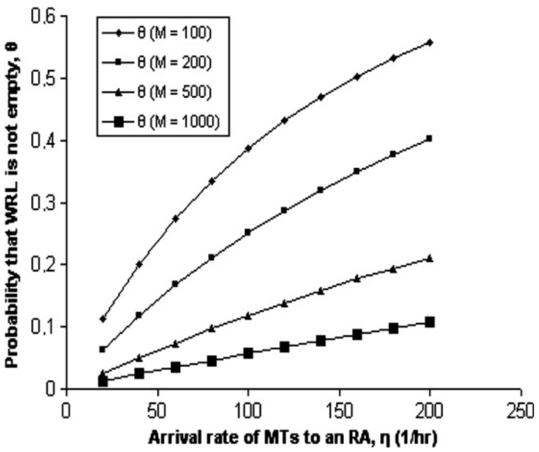

Fig. 19. Probability that the RWL is not empty when an incoming call arrives (proposed strategy).

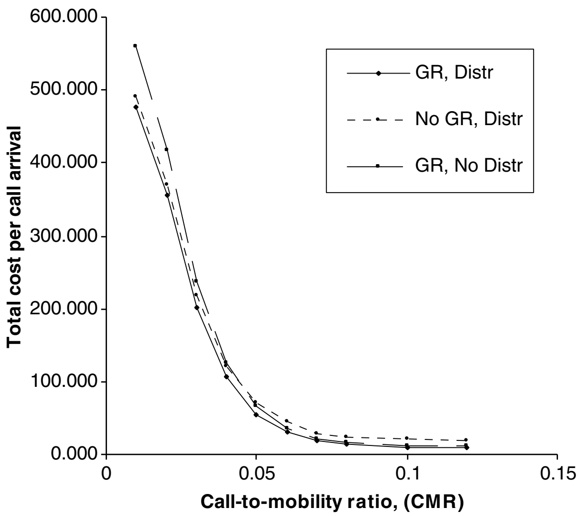

Fig. 20. Total cost per call arrival versus CMR (proposed strategy, GR strategy without distributed databases, no GR strategy with distributed databases).

The performance observed in Fig. 18 for the proposed strategy is similar to the performance observed in [19] (one has to be careful though that in [19] the graph of $1/p_2$ is plotted).

6.1. Probability P_2 versus λ_0

Fig. 18 shows the probability $p2$ versus λ_0, under a significant value of μ_0 (μ_0=1), where $N_i = 20$ ($i = 1, 2, \ldots, 8$) is used for demonstration purposes. In Fig. 18 we can see that p_2 is very small and that it decreases as λ_0 increases. This is true since as the RA residence time rate λ_0 increases the probability of an RA not to be the local anchor of an MT decreases. Thus, the longer an MT stays in an RA, the smaller the probability of the RA not to be the local anchor of the MT becomes. Also, it should be noted that the value of μ_0 has an impact only on the scale of p_2 but not on the general form of the graph of Fig. 18 (see Eq. (16)). For a significant value of μ_0 ($\mu_0 = 1$) we can see how p_2 varies with λ_0 (Fig. 18).

6.2. Probability θ versus MT the arrival rate η

Fig. 19 studies the impact of the MT arrival rate η at RAs under the same RNC on the probability of the RWL not to be empty, when an incoming call arrives at the RAs, under different number M of MTs registered at RAs. From Fig. 19, it can be seen that θ decreases as η decreases or M increases. This is true because a smaller MT arrival rate at the RAs (under the same RNC) increases the probability of the RWL to be empty upon the arrival of a call. On the other hand, more MTs in the RAs (under the same RNC) will make RAs receive calls more often and the RWL will be emptied more frequently, thus the probability of the RWL not to be empty becomes smaller, resulting in a smaller cost for the proposed strategy. The graph of Fig. 19 is similar to that of the GR strategy [19] (θ versus η).

6.3. Total cost versus CMR

Fig. 20 shows the total cost per call arrival under value λ_0 ($\lambda_0 = 12\,\text{min}$), where (M, η) is set to $(200, 100)$ for the proposed strategy, the GR strategy without distributed databases and the scheme of distributed databases without the GR strategy. In Figs. 20 and 21, the following cost values are used: $C_u = 1$, $C_h = 1.5$, $C_{uu} = 1$ and $C_{hu} = 2$. It is assumed that a query or an update at the DB_0 (HSS) incurs a larger cost than that at DB_1, and the messages exchanges between DB_0 and a DB_1 incur a larger cost than those between two DB_1s. Additionally, we can see that under a specific CMR value (CMR $\approx$ 2.4) the value of total cost is stabilized ($C_{\mathrm{TOTAL}} \approx 8$).

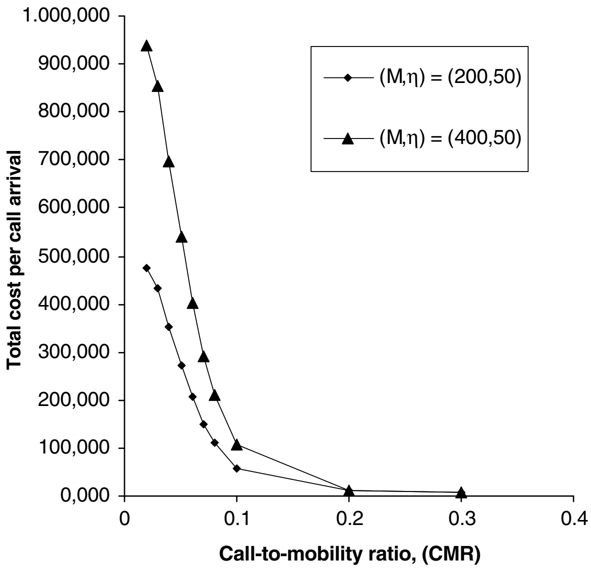

Fig. 21. Total cost per call arrival versus CMR (proposed strategy) (a) $M = 200$ and $\eta = 50$ (b) $M = 400$ and $\eta = 50$.

As it is shown in Fig. 20, the total cost per call arrival versus CMR in 3G networks shows similar behavior as the one in GSM Networks (Fig. 7, Ref. [19]). The only difference is that the maximum value of the total cost of our proposal is higher ($\approx$476,247) than that of the GR for GSM networks ($\approx$350). Moreover, the total cost of the GR strategy without distributed databases and the total cost of the scheme, with distributed databases without GR strategy, is greater than our proposed strategy. Thus, the total cost (491.247) of the scheme with distributed databases without GR strategy is 3% greater than the proposed strategy and the total cost (560.033) of the GR strategy without distributed databases is 15% greater than the total cost of GR strategy with distributed databases.

Fig. 21 shows the total cost per call arrival versus CMR, under two different sets of (M, η), $(200, 50)$ and $(400, 50)$ and also $\lambda_0 = 12$ min. We observe that as M increases while (η) is stable, the total cost decreases as CMR increases. Furthermore, a larger M results in smaller p_1 and p_2 reducing both the expected location update cost and the call delivery cost in the proposed strategy.

Additionally, in both sets of (M, η), under a certain value of CMR (CMR $\approx$ 2.4), the value of the

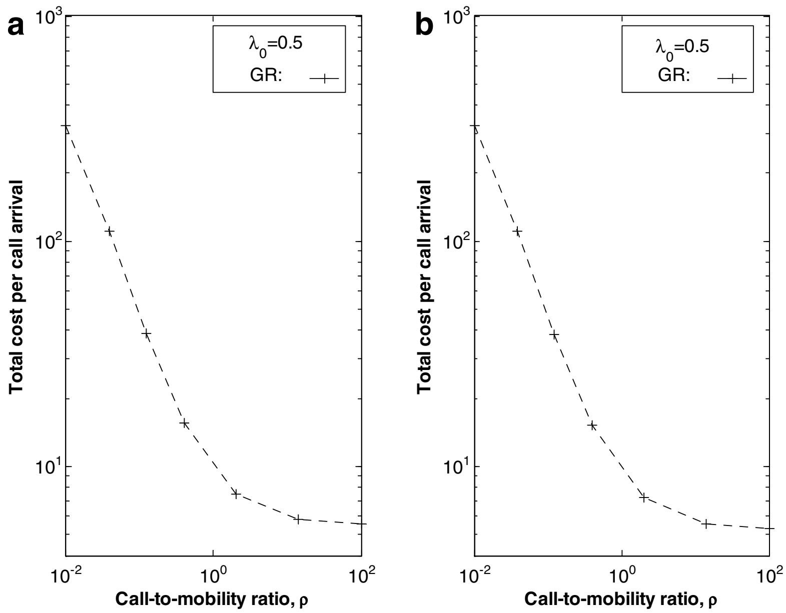

Fig. 22. Total cost per call arrival versus CMR (GR strategy in GSM networks) (a) $M = 200$ and $\eta = 50$ (b) $M = 400$ and $\eta = 50$, as depicted in [19].

total location update cost is stabilized ($M = 200$, $C_{\mathrm{TOTAL}} \approx 8$ and $M = 400$, $C_{\mathrm{TOTAL}} \approx 9$).

Also, the total costs per call arrival versus CMR in 3G and GSM networks respectively (Figs. 21 and 22) show the same behavior. The total cost of the proposed strategy has a greater value in both cases than that of the GR strategy in GSM networks, but still very satisfying.

6.4. Call delivery latency versus CMR

Fig. 23 studies the expected call delivery latency of an incoming call of the proposed strategy, the GR strategy without distributed databases and the distributed databases without the GR strategy. The cost of a network element represents the delay incurred by its operation. The following parameter values are used in Fig. 23 for illustrative purposes: $C_u = 10$ ms, $C_{uu} = 10$ ms and $C_{hu} = 20$ ms. (M, η) is set to $(200, 50)$. Note that the delay incurred by the message exchanges between the caller and the called MTs DB_0 is not included in the comparison, since this delay is the same for the three strategies (the proposed strategy, the GR one without distributed databases and non-GR one with distributed databases).

From Fig. 23, we observe that the average call delivery latency remains constant (77 ms approximately for the proposed strategy), as the CMR increases, so it is independent from the call-to-mobility ratio factor.

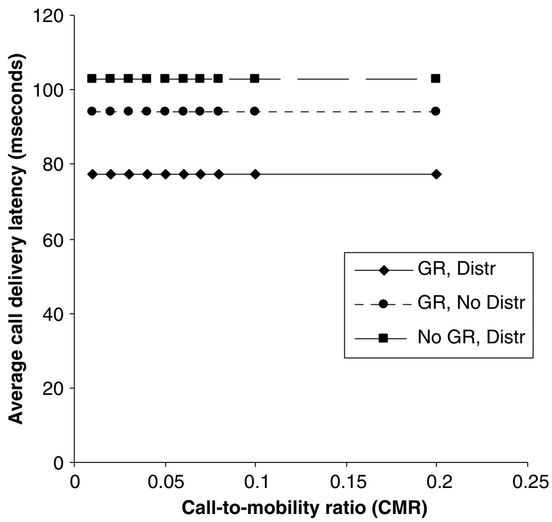

Fig. 23. Call delivery latency of an incoming call versus CMR (proposed strategy, GR strategy without distributed databases, no GR strategy with distributed databases), as depicted in [19].

Moreover, Fig. 24 of the call delivery latency versus CMR for GSM networks ([19]) is almost the same as the graph of Fig. 23. The only difference is the value of average call delivery latency, which is 50 ms. Additionally, the expected call delivery latency of the GR strategy without distributed dat-

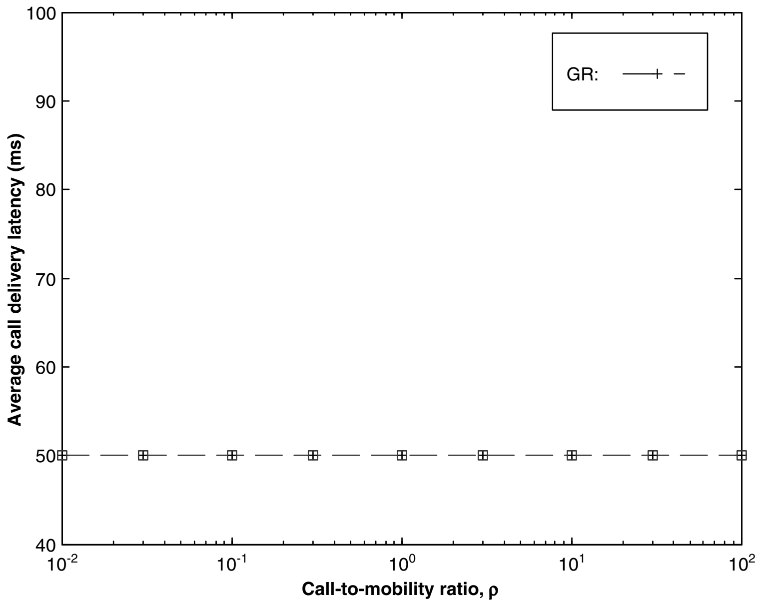

Fig. 24. Call delivery latency of an incoming call versus CMR (GR strategy in GSM networks).

abases and the distributed databases without the GR strategy is higher than that of the proposed scheme.

7. Conclusions

In order to ascertain the behavior of 3G wireless networks, we proposed the Group Registration strategy (GR strategy) for 3G networks integrated with a 2-level distributed databases architecture. The application of the GR strategy for 3G networks with distributed databases architecture shows a much better behavior than the GR strategy without the distributed databases architecture and the scheme of the distributed databases without the GR strategy.

Therefore, by applying the GR strategy in combination with a 2-level distributed databases architecture, we achieved a significant decrease of the total location management cost comparing with the total costs of these two strategies. Also, the results for the GR strategy in 3G networks show familiar behavior compared with the ones in GSM networks. The GR strategy in combination with the 2-level distributed databases architecture applied in 3G networks, leads to a significant reduction not only of the location management cost, but also of the mean call delivery time.

Thus, the resources of the network are allocated in a much more sufficient way and the quality of service of the network is improved. Since, the increase of subscribers in wireless networks has led to the need for efficient location tracking strategies, the development and application of efficient location management schemes in 3G Wireless Networks, minimizing the location management costs, are more than a necessity. Another emerging area in wireless networks where the above strategies may be applicable is that of Ad-hoc and Sensor networks. However, the new Location Management strategies that will be introduced should take into account the limitations arising by these networks.

References

[1] Cisco Mobile Exchange (CMX), Solution guide 0L-2947-01. 3GPP 23060.
[2] D.D. Vergados, A. Panoutsakopoulos, C. Douligeris, Location management in 3G networks using a 2-level distributed database architecture, in: Proceedings of the 18th Annual IEEE International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC'07), Athens, Greece, 2007.
[3] A. Bar-Noy, I. Kessler, M. Sidi, Mobile Users: to update or not to update? ACM-Baltzer J. Wireless Networks 1 (2) (1995) 175–186.
[4] V. Wong, V. Leung, An adaptive distance based location update algorithm for next generation PCS networks, IEEE J. Sel. Areas Commun. 19 (10) (2001) 1942–1952.
[5] I.F. Akyildiz, J. Mcnair, J.S.M. Ho, H. Uzunalioglou, W. Wang, Mobility management in next generation wireless systems, Proc. IEEE 87 (8) (1999) 1347–1348.
[6] J. Li, Y. Pan, X. Jia, Analysis of dynamic location management for PCS networks, IEEE Trans. Network. 5 (1) (1997) 25–33.
[7] Z. Mao, C. Douligeris, A location based mobility tracking scheme for PCS networks, Comput. Commun. 23 (18) (2000) 1729–1739.
[8] Y. Xiao, Optional fractional movement based scheme for PCS location management, IEEE Commun. Lett. 7 (2) (2003) 67–69.
[9] C. Rose, Minimizing the average cost of paging and registration: a timer based method, ACM-Baltzer J. Wireless Networks 2 (2) (1996) 109–116.
[10] A. Hac, X. Zhou, Location strategies for personal communication networks: a novel tracking strategy, IEEE J. Sel. Commun. 15 (8) (1997).
[11] J.-K. Wey, L.-F. Sun, W.-P. Yang, Using multilevel hierarchical registration strategy for mobility management, Int. J. Inform. Sci. 89 (1–2) (1996) 63–76.
[12] S. Tabbane, An alternative strategy for location tracking, IEEE J. Sel. Areas Commun. 13 (5) (1995) 880–892.
[13] Quintero O. Garcia, S. Pierre, An alternative strategy for location update and paging in mobile networks, Comput. Commun. 27 (15) (2004) 1509–1523.
[14] Z. Mao, C. Douligeris, Location tracking in mobile networks: a scheme and an analysis framework, Electron. Lett. 39 (19) (2003) 1406–1408.
[15] R. Jain, Y.B. Lin, An auxiliary user location strategy employing forwarding pointers to reduce network impacts of PCS, ACM-Baltzer J. Wireless Networks 1 (2) (1995) 197–210.
[16] E. Pitoura, I. Fudos, Distributed location databases for tracking highly mobile objects, Comput. J. 44 (2) (2001) 75–91.
[17] U. Madhow, M.L. Honig, K. Steiglitz, Optimization of wireless resources for personal communications mobility tracking, in: Proc. IEEE INFOCOM '94, June 1994, pp. 577–584.
[18] P.G.-Escale, V.C. Giner, A hybrid movement-distance-based location update strategy for mobility tracking, in: Proceedings of the 5th European Wireless Conference. Mobile and Wireless Systems beyond 3G, Barcelona (Spain), February 2004.
[19] Z. Mao, C. Douligeris, Group Registration with local anchor for location tracking in mobile networks, IEEE Trans. Mobile Comput. 5 (5) (2006) 583–595.
[20] G. Morales-Andres, M. Villen-Altamirano, An approach to modeling subscriber mobility in cellular radio networks, in: Proceedings of the Telecom Forum'87, Geneva, Switzerland, 1987, pp. 185–189.
[21] E. Alonso, K.S. Meier Helsten, G.P. Pollini, Influence of cell geometry on handover and registration rates in cellular and universal personal telecommunications networks, in: Proceedings of the 8th International Telecommunications Congress Special Seminar on Universal Personal Telecommunications, Genova, Italy, October 1992, pp. 261–270.

[22] J.S.M. Ho, I.F. Akyildiz, Local anchor scheme for reducing signaling costs in personal communication networks, IEEE/ACM Trans. Network. 4 (5) (1996) 709–725.
[23] A. Mishra, Performance characterization of signaling traffic in UMTS core networks, in: IEEE Global Telecommunications Conference, vol. 2, December 2003, pp. 1141–1146.
[24] Cisco Network Planning Solution – SPM Specialized Models User Guide, 8-UMTS Model User Guide, Release 11.5. <http://www.cisco.com/application/pdf/en/us/guest/products/ps6757/c1626/ccmigration_09186a008059b897.pdf>.

Dimitrios D. Vergados was born in Athens, Greece in 1973. He is a Lecturer in the Department of Informatics, University of Piraeus. He was a Lecturer in the University of the Aegean, Department of Information and Communication Systems Engineering. He received his BSc from the University of Ioannina and his PhD from the National Technical University of Athens, Department of Electrical and Computer Engineering. His research interests are in the area of Communication Networks (Wireless Broadband Networks, Sensor – Ad-hoc Networks, Military Networks, WLANs, IP, MIP, SONET Networks), Neural Networks, GRID Technologies, and Computer Vision. He has participated in several projects funded by EU and National Agencies and has several publications in journals, books and conference proceedings. He is, also, Guest Editor and Reviewer in several Journals and he has served in technical program committees of several international conferences.

Alexandros Panoutsakopoulos was born in Athens, Greece in 1978. He received his first degree from the University of the Aegean, Department of Mathematics and his MSc from the Department of Information and Communication Systems Engineering of Aegean University. His research interests are in the area of Communications Networks (Wired Networks, Wireless Networks, Ad-hoc Networks, Sensor Networks) and especially in mobility management techniques for wireless-mobile Networks.

Christos Douligeris held positions with the Department of Electrical and Computer Engineering at the University of Miami, where he reached the rank of associate professor. He is currently an associate professor at the department of Informatics, University of Piraeus, Greece. He is also an associate member of the Hellenic Authority for Information and Communication Assurance and Privacy, member of the board of trustees of the Greek Computer Society and member of the Executive Committee of the National Council on Electronic Commerce. Dr. Douligeris has published extensively in the networking scientific literature and he has participated in many research and development projects. He was the guest editor of a special issue of the IEEE Communications Magazine on Security for Telecommunication Networks. He is the co-editor of a book on "Network Security" published by IEEE Press/John Wiley. He is an editor of the IEEE Communications Letters, a technical editor of IEEE Network, Computer Networks (Elsevier), International Journal of Wireless and Mobile Computing (IJWMC), Journal of Internet Engineering, International Journal of Sensor Networks and the Euro Mediterranean Journal of Business (EMJB).