EFFICIENT LOCATION FINDING UTILISING ROUTING FUNCTIONALITY IN 3RD GENERATION WIRELESS/MOBILE SYSTEMS

E. D. Sykas*, J. A. Korinthios, P. I. Philippopoulos***

* National Technical University of Athens (NTUA)
Department of Electrical Engineering and Computer Science
9 Heroon Polytechniou Str., Zographou
157 73 Athens, GREECE
PH: +30-1-772-2528, FAX: +30-1-772-2534
e-mail: sykas@cs.ntua.gr
** Panafon S.A., GREECE.

Abstract - *One of the core functionalities in wireless/mobile networks, is tracking of Mobile Terminals (MTs) location. Massive user roaming requires a location management mechanism that provides for speed and signalling efficiency. This paper proposes an efficient Location Finding mechanism, based on flexible routing functionality in the Radio Access Network (RAN) of 3G mobile systems. To assess its applicability, two operational scenarios are identified: A "traditional" one, featuring transaction-based routing updates and paging in the RAN, and an innovative one, that minimises Paging and enhances accuracy of the MTs location information, at the cost of more frequent updates. A trade-off analysis between the two approaches is provided, focusing on the incurred air-interface signalling load and identifying the important traffic parameters and thresholds.*

I. INTRODUCTION

Third Generation (3G) wireless/mobile systems are emerging. UMTS, viewed as the European member of the family of 3G mobile systems developed within the ITU IMT-2000 framework, is forseen to provide for a unified user access to a multiplicity of services, through advanced low-power terminals and will guarantee high capacity, in order to support the envisaged traffic from users in a variety of mobile environments.

Tracking of roaming MTs is a core functionality in mobile networks. Considering the huge number of subscribers and the variety of mobile environments employing different cellular structures, mobility tracking should be characterised by two main merits: speed of execution and signalling efficiency. Two basic mobility functions are currently utilised for the deployment of mobility tracking: Location Updating (LU) and Paging. Through LU, location information kept in the network databases, is refreshed via LU messages emitted from the MTs. Location information is organised in logical sets named Location Areas (LAs). The formation of LAs is subject mainly to performance and signalling preservation constraints, but also relates to geographical and administrative issues. Thus, the network knows the position of a MT with the accuracy of a LA. *Paging* has to be performed within a Paging Area (PA), to reveal the exact position of the MT. Since the location information held by the system is consistent, it is reasonable to expect that the PA is equal to the LA (like in current 2G cellular systems, e.g. GSM). When designing large LAs/PAs the overall LU rate for the mobile system is minimized, thus providing a signalling traffic gain, in both the air-interface and the fixed network. However, signalling due to paging is increased, as a larger number of cells is polled. Therefore, the design of PAs/LAs is currently a trade-off between the volume of LUs and the bandwidth required to perform paging.

In 3G wireless/mobile systems (like UMTS), a separation between the concepts of PA and LA may be introduced. Studies (e.g. [1]) indicate, that signalling efficiency can be improved if the PAs are dynamically determined (Intelligent Paging), based on user specific information (e.g. roaming history, user preferences/habits, subscription information). The formation of PAs should thus be determined by the envisaged paging traffic and be a pure matter of signalling bandwidth preservation, while LA design should be a matter of network DB dimensioning (LU/Interrogation dependent) and administrative considerations. Location management strategies proposed in current literature (e.g. [2], [3]), are in principal dynamic LU schemes combined with appropriate Paging strategies, implemented with a certain degree of supplementary LU/paging-specific signalling to keep the network updated. This paper aims to provide a solution for the decoupling of the formation of PAs and LAs,

0-7803-5435-4/99/$10.00

utilising a connectionless routing protocol, originally designed for the transport of signalling in the UMTS Access Network ([4]). The proposed Location Finding (LF) mechanism provides an extra accuracy of location information, without imposing any additional traffic on the core network databases. A trade-off study is presented, between a "traditional" approach for LF and a novel approach, which maintains more accurate MT positional routing information within the RAN. The two approaches can be concurrently applied on different MTs, or interchangeably used for the same MT (e.g. according to power consumption requirements). While this paper focuses on signaling traffic, the routing mechanism is also applicable for packet data user services.

II. THE LF MECHANISM

The proposed LF scheme is based on a network layer (NL) routing protocol ([4]), applicable on any RAN topology that can be reduced to a logical tree. It comprises of a Data Transfer, a Routing Update (RU) and a Route Finding (RF) mechanism. The Data Transfer mechanism is connectionless, thus avoids connection release and re-establishment. The RU mechanism maintains a path towards the MT, while the MT is roaming and changes its point-of-attachment (PoA). For this purpose, the MT emitts towards the network, appropriate RU messages that update the routing tables in the network nodes, thus maintaining routing information along only one active path towards the MT. If the exact position of the MT is unknown, the RF mechanism is initiated. *Table 1,* outlines the basic message types used and *Figure 1* depicts the creation of a routing path towards the MT.

Messages	Usage
Unitdata (UDT)	Connectionless data packet carrying user (upper layer) data.
Routing Update (RU)	Packet updating the route towards a MT within the RAN. It might contain upper layer data.
Find (FND)	Broadcasted over the air-interface within an RAN. It is used by the RF funtion, to alert and effectively locate the appropriate MT.

Table 1: RAN Routing Protocol messages

Three types of Nodes are identified, i.e. *MT*, *Edge* and *Fixed* nodes. MT1 with address MTA_1 issues a RU packet which reaches the current PoA (Edge_Node1) with address BSA_1. Each network entity (NE) receiving a RU, creates a routing entry with the MT address pointing to the previous NE in the path and, if necessary, forwards an RU to the NE above it in the tree hierarchy. Thus, routing in the direction from the RAN root node towards the MT becomes possible. The routing tables that are maintained in this case, and the relevant entries, are shown in the figure. "⊙" denotes the current NE.

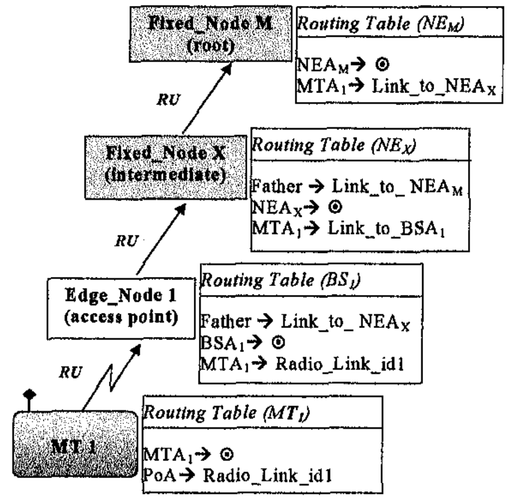

Figure 1: Complete route update mechanism.

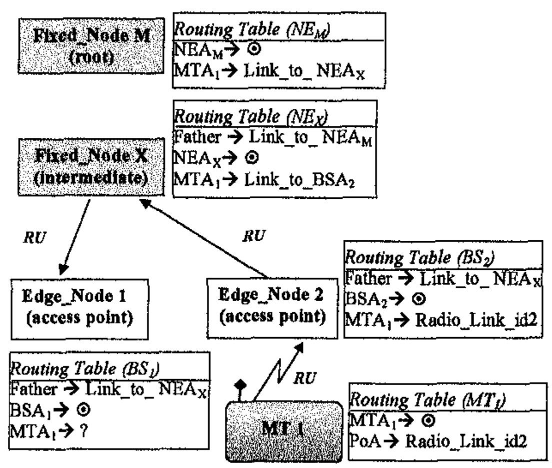

Figure 2: Routing information maintenance.

Figure 2, depicts the maintenance of the routing path towards MT1, while changing its PoA from Edge_Node1 to Edge_Node2. An RU packet updates the routing table in Edge_Node2 and a subsequent RU updates the routing table in Fixed_NodeX. Fixed_NodeM has the correct entry, so an RU is not forwarded towards it, i.e. the path is updated as high as required (up to the first common node in the old and new path). A RU may also be forwarded towards Edge_Node 1, to remove the old path. If no routing information is available in the RAN for a requested MT, then the RF mechanism is initiated and a paging message is appropriately multicasted in the RAN. However, the edge nodes may not broadcast the paging message over the air-interface, but temporarily store it and broadcast a short FND message (possibly addressing multiple MTs) instead, to save bandwidth. The MT reacts with an RU, thus establishing a complete path up to the root. Two operational scenarios are identified regarding the triggering of RUs via which LUs are effected:

- The *"Update on a transaction basis"* scenario: RUs are only emitted when the MT has an on-going transaction with the network. Routing information is therefore only used if available. Otherwise, a RF procedure is performed within the LA. However, if the MT has already an on-going transaction, its exact position (PoA) is known. This information is also used in 2G mobile systems. The RU mechanism in used in a "traditional" way and paging is performed similarly to the 2G systems.
- The *"Update on a regular basis"* scenario: Routing information is regularly updated when the MT changes Atomic areas (AA) while roaming. *Atomic* is the area, that when crossed, a routing update is effected. As AA we may consider the area covered by an Edge node, which (for simplicity) we will assume equivalent to a single cell. Consequently, route finding (i.e. paging) is not performed, since the exact location of the MT is known. This results in the minimisation of the paging traffic. On the other hand, additional signalling traffic is incurred due to the regular RUs. An certain increase in MT power consumption should also be expected. This scenario results in the complete decoupling of the LA and PA concepts. The PA is limited to the AA.

The MT position is known in the core network DB with the accuracy of a LA, in both approaches. Location accuracy could be improved in the 2nd scenario via the AA concept, without incurring any additional signalling traffic on the DB, as (Network Layer) routing functionality in the RAN, transparently refreshes location information. This results in the reduction of the paging related signalling over the air-interface. Moreover, the implementation of paging in one polling cycle, minimises paging delay and enhances location tracking speed. However, additional signalling traffic is incurred in the RAN, due to frequent updates. Nevertheless, this traffic will be evenly distributed throughout the RAN.

III. LF MECHANISM EVALUATION

The performance of the indicated scenarios in terms of signalling load imposed on the air-interface will be evaluated and compared.

The MT crossing rate through a random curve of perimeter L_a , assuming an average density of MTs σ_{MT}, and an average speed of MTs v_{MT} in the neighbourhood of the curve is ([5]):

$$\lambda_{in,a} = \lambda_{out,a} = 1/\pi \cdot L_a \sigma_{MT} v_{MT}$$

where, $\lambda_{in,a}$ and $\lambda_{out,a}$ are the crossing rates in and out of the area respectively and $1/\pi$ represents the random movement of the MTs (assumed here for simplicity - see [6] for alternative factors).

"Update on a Regular basis" scenario

The total air-interface signalling load (bytes/sec) in an AA ($\lambda_{reg,s,a}$) comprises the following components (expressed in messages/sec): the update rate ($\lambda_{reg,u,a}$) due to the RU and LU messages, the handover rate ($\lambda_{h,a}$), and call rate $\lambda_{c,a}$ (incoming and outgoing). Thus:

$$\lambda_{reg,s,a} = \lambda_{reg,u,a} l_u + \lambda_{h,a} l_h + \lambda_{c,a} l_c$$

Where, l_u is the average length of update messages, l_h the length of handover messages and l_c is the length of call set-up messages.

The handover rate is proportional to the rate of MTs that cross into the atomic area while having an active call. Therefore, it can be estimated as:

$$\lambda_{h,a} = A \lambda_{in,a}$$

The probability of having an active call (A), is given by the following equation:

$$A = \lambda/(\lambda + \mu)$$

Where, $1/\lambda$ is the average time between calls and $1/\mu$ the average duration of the call. We assume that when a MT with an active call crosses an AA, the HO message (AL-message) is encapsulated within the RU (NL-message). Thus, the update rate is given by:

$$\lambda_{reg,u,a} = \lambda_{in,a} - \lambda_{h,a} = (1 - A)\lambda_{in,a}$$

In order to estimate the update load, note that LU messages differ in length from RUs. When the AA is at the border of the LA, the RU message carries a LU message. We observe that for each AA, only a fraction of its perimeter is also part of the LA perimeter. This fraction will be called x_a , where $0 \le x_a \le 1$. If $x_a = 0$, then the AA is in the middle of the LA and has zero LU rate. If $x_a = 1$, the complete AA border is part of the LA border. The average length of update messages is thus:

$$l_u = [x_a l_{lu} + (1 - x_a) l_{ru}]$$

where, l_{lu} is the length of LU message and l_{ru} is the length of RU message. The call rate (incoming and outgoing) depends on the number of the MTs ($\sigma_{MT}\ S_a$), where S_a is the surface of the AA. Therefore:

$$\lambda_{c,a} = \sigma_{MT} S_a A\mu$$

"*Update on a transaction basis*" scenario

In this scenario, it is assumed that the MT, in parallel to having an active transaction, also performs RU. If routing information is available, the RAN can use it in order to locate the MT. Note, that in this scenario RU happens only in parallel to other procedures (LU or handover). The signalling load ($\lambda_{trans,s,a}$) in an AA is by:

$$\lambda_{trans,s,a} = \lambda_{trans,u,a} l_{lu} + \lambda_{h,a} l_h + \lambda_{trans,p,a} l_{pm} + \lambda_{c,a} l_c$$

where, $\lambda_{trans,u,a}$ is the update rate (LUs), $\lambda_{trans,p,a}$ the paging rate, l_{pm} the average length of RF (i.e. paging related) messages. The handover and call rates coincide with those of the first scenario. As LUs happen at LA boundaries, the update rate is:

$$\lambda_{trans,u,a} = x_a \lambda_{reg,u,a}$$

If routing information is not available in the RAN, RF is performed in the complete LA, which consists of N AAs. This results in a short FND message, broadcasted over the complete LA. The MT responds with an RU. Then, the paging message is sent over the air-interface to the appropriate AA only. The paging load is:

$$\lambda_{trans,p,a} l_{pm} = \lambda_{trans,p,a} (l_{ru} + N l_f)$$

where, l_f is the length of the FND message and N is the number of AAs in a LA. The above holds for $N > 1$. For $N = 1$, the position of the MT is accurately known and as a result, there is no reason for applying route finding. As only incoming calls are paged and assuming that incoming and outgoing call rate are equal, the paging rate is given by:

$$\lambda_{trans,p,a} = 1/2 \cdot \sigma_{MT} S_a A\mu$$

Performance Comparison

In an AA, the "regular basis" scenario is preferable over the transaction based, when:

$$\lambda_{reg,s,a} < \lambda_{trans,s,a} \Leftrightarrow$$
$$(1 - x_a)\lambda_{reg,u,a} l_{ru} < (N l_f + l_{ru})\lambda_{trans,p,a}$$

In order to minimise the signalling load in the complete LA, we sum all AAs signalling loads:

$$\sum_{a=1}^{N}[(1 - x_a)\lambda_{reg,u,a} l_{ru}] < \sum_{a=1}^{N}[(N l_f + l_{ru})\lambda_{trans,p,a}] \Leftrightarrow$$

$$\frac{(1-A)v_{MT} l_{ru}}{\pi}\sum_{a=1}^{N}[(1-x_a)L_a] < \frac{A\mu}{2}\sum_{a=1}^{N}[(N l_f + l_{ru})S_a] \quad (1)$$

where, α is the α^{th} AA in a LA ($\alpha = 1, \ldots, N$). Formula (1) is a generic result, corresponding to any geometrical area pattern. Assuming that L_α and S_α are constant for each AA:

$$\frac{(1-A)L_a v_{MT} l_{ru}}{\pi}[N - \sum_{a=1}^{N} x_a] < \frac{N(N l_f + l_{ru})S_a A\mu}{2} \quad (2)$$

From (2) it is evident that as N increases, i.e. the larger the LA becomes, the "regular basis" scenario becomes preferable. There will be a threshold N_0, such that (2) is satisfied for every $N > N_0$, since the right hand side of (2) is proportional to N^2, while the left hand side is proportional to N.

$\sum_{a=1}^{N}(x_a)$, L_α and S_α geometrical parameters in Appendix A, are approximated for a LA consisted of randomly shaped cells having equal surface. After some algebraic calculations and based on the definition of an equivalent radius R_α, we express the v_{MT}/λ ratio as follows:

$$\frac{v_{MT}}{\lambda} < \frac{\pi}{4}\frac{N(N l_f + l_{ru})}{N - \sum_{1}^{N}(x_a)}\frac{1}{l_{ru}}R_a \quad (3)$$

Equation (3) indicates that for given LA and AAs, there is a given v_{MT}/λ (average distance crossed by a MT between successive calls) threshold, below which the regular update scenario is always preferable. By increasing v_{MT}/λ, the update on a transaction basis scenario becomes preferable. The preferred scenario also depends on R_α. By increasing R_α, the update on a regular basis scenario becomes preferable. *Table 2* depicts some typical v_{MT}/λ values for different user environments:

Categories of Users	v_{MT} *(km/h)*	λ *(calls/h)*	v_{MT}/λ *(km/call)*
Office User	**0,5**	**3**	**0,17**
Pedestrian User	**5**	**3**	**1,67**
Car User	**50**	**3**	**16,67**

Table 2: Typical values of the v_{MT}/λ

Further calculations of v_{MT}/λ ratio, based on the results of Appendix A lead to:

$$\frac{v_{MT}}{\lambda} < \frac{\pi^2}{4}\frac{Nl_f + l_{ru}}{N}\frac{1}{l_{ru}}R \qquad (4)$$

Considering (A.1) and solving for N we obtain:

$$N > \frac{l_{ru}}{l_f}\left(\frac{4}{\pi}\frac{v_{MT}}{\lambda}\frac{R}{S_a} - 1\right) \qquad N \neq 0 \qquad (5)$$

Therefore, in the case where the complete LA is of circular shape with radius R and consists of N AAs of equal surface S_α, through (5) we can calculate the positive integer N_0, such that for $N > N_0$, the innovative "regular update" scenario becomes preferable. N_0 depends on the characteristics of the user traffic (v_{MT}/λ) and on geographical considerations (the ratio between the radius of the LA and the surface of each individual AA). Since $N \neq 0$, N_0 is formally expressed by:

$$N_0 = Max[Trunc(\frac{l_{ru}}{l_f}(\frac{4}{\pi}\frac{v_{MT}}{\lambda}\frac{R}{S_a} - 1) + 1), 1]$$

Figure 3, depicts some values for N_0, calculated for various area patterns (R/S_a), based on the anticipated user traffic characteristics (v_{MT}/λ) of some environments, as presented in *Table 2*. A typical message length ratio (l_{ru}/l_f =1) is used.

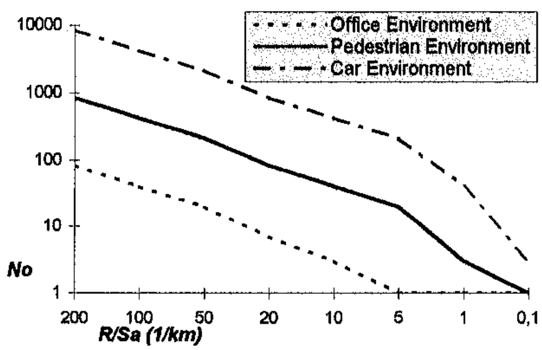

Figure 3: N_0 values for randomly shaped area consisted of equally sized cells.

IV. CONCLUSIONS

The study compared the air-interface signalling load for the innovative and the traditional approach regarding location finding within the RAN. As the number of cells per LA (N) increases, the "regular update" scenario becomes preferable. There is always a threshold N_0, such that for every $N > N_0$, the innovative scenario minimises signaling over the traditional one. The analytical solution given to the above trade-off problem holds for any geometrical area pattern. The choice between the scenarios depends on the average distance crossed by a MT between successive calls (v_{MT}/λ) and on geographical considerations (the ratio between the radius of LA and the surface of each individual AA). On the other hand, the choice does not depend on the user density (σ_{MT}), the LU message length (l_{lu}) and the average duration of a call ($1/\mu$). The scenarios are not mutually exclusive and can be used interchangeably.

V. APPENDIX A

Regarding the geometrical aspects of a randomly shaped AA, we define the equivalent radius (R_a) of a randomly shaped AA such that: $S_a = \frac{1}{2}R_a L_a$

In order to estimate R_a , we view the AA as sector of a circle (radius R) approximating the LA. The following equations hold:

$$S_a = \frac{S_{LA}}{N} = \frac{\pi R^2}{N} = \frac{1}{2}R_a L_a \qquad (A.1)$$

$$L_a = 2R + 2\pi R/N \qquad (A.2)$$

$$x_a = \frac{2\pi R/N}{L_a} = \frac{\pi}{\pi + N} \qquad (A.3)$$

$$\sum_{a=1}^{N}(x_a) = X = \frac{\pi N}{\pi + N} \qquad (A.4)$$

VI. REFERENCES

[1] G.L.Lyberopoulos, et al., "Intelligent Paging strategies for 3rd Generation mobile telecommunication systems", *IEEE Transactions on Vehicular Technologies,* Vol. 44, No. 3, August 1995.

[2] A.Bar-Noy, I.Kessler and M.Sidi, "Mobile Users: To Update or not to Update?", *ACM-Baltzer Journal of Wireless Networks*, Vol. 1, No. 2, July 1995.

[3] V.Casares-Giner, J.Mataix-Oltra, "On movement-based mobility tracking strategy. An enhanced version", *IEEE Communications Letters,* Vol.2, Feb. 1998.

[4] H.Mitts, G.Luijten, J.A.Korinthios and J.Nelson, "Connectionless Signalling Network Layer in UMTS", *IEEE Personal Communications Magazine*, June 1996.

[5] G. morales, M. Vilen, "An approach to modelling subscribers mobility in cellular radio network", *Forum Telecom '87,* Geneva, 1987.

[6] J.G.Markoulidakis, E.D.Sykas, "Model for Location Updating and Handover Rate Estimation in Mobile Telecom.", *IEEE Electr. Letters*, Vol. 29, No. 17, Aug. 1993.