International Journal of Electrical, Electronic and Communication Sciences
ISSN: 2517-9438
Vol:1, No:1, 2007

Location Management in Cellular Networks

Bhavneet Sidhu, and Hardeep Singh

Abstract**—**Cellular networks provide voice and data services to the users with mobility. To deliver services to the mobile users, the cellular network is capable of tracking the locations of the users, and allowing user movement during the conversations. These capabilities are achieved by the location management. Location management in mobile communication systems is concerned with those network functions necessary to allow the users to be reached wherever they are in the network coverage area. In a cellular network, a service coverage area is divided into smaller areas of hexagonal shape, referred to as cells. The cellular concept was introduced to reuse the radio frequency. Continued expansion of cellular networks, coupled with an increasingly restricted mobile spectrum, has established the reduction of communication overhead as a highly important issue. Much of this traffic is used in determining the precise location of individual users when relaying calls, with the field of location management aiming to reduce this overhead through prediction of user location. This paper describes and compares various location management schemes in the cellular networks.

Keywords**—**Cellular Networks, Location Area, Mobility Management, Paging.

I. INTRODUCTION

IN the past decade, cellular communications have experienced an explosive growth due to recent technological advances in cellular networks and cellular phone manufacturing. It is anticipated that they will experience even more growth in the next decade. In order to accommodate more subscribers, the size of cells must be reduced to make more efficient use of the limited frequency spectrum allocation. A cellular communication system must track the location of its users in order to forward calls to the relevant cell within a network. This will add to the challenge of some fundamental issues in cellular networks. Location management is one of the fundamental issues in cellular networks. It deals with how to track subscribers on the move. The purpose of this paper is to survey recent research on location management in cellular networks. The study of location management aims to reduce the overhead required in locating mobile devices in a cellular network. The paper is organized as follows: The next section will briefly describe some fundamental concepts about cellular networks and location management. The section 3 will address recent researches on location management schemes. Finally the paper is concluded by summarizing the main points.

Manuscript received November 30, 2006.

Authors are with Department of Computer Science & Engineering, Guru Nanak Dev University, Amritsar, India (e-mails: bhavneet_sidhu@yahoo.co.in, hardeep_gndu@rediffmail.com).

II. CELLULAR NETWORKS

In a cellular network, a service coverage area is divided into smaller areas of hexagonal shape, referred to as cells. Each cell is served by a base station. The base station is fixed. It is able to communicate with mobile stations such as cellular phones using its radio transceiver. The base station is connected to the mobile switching center (MSC), which is, in turn, connected to the public switched telephone network (PSTN). Fig. 1 illustrates a typical cellular network. (*A base station is marked with a triangle.)*

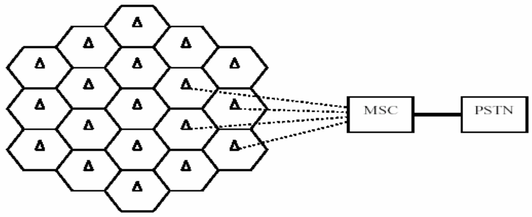

Fig. 1 A typical cellular network

The frequency spectrum allocated to wireless communications is very limited. The cellular concept has been introduced to reuse the frequency. Each cell is assigned a certain number of channels. To avoid radio interference, the channels assigned to one cell must be different from the channels assigned to its neighboring cells. However, the same channels can be reused by two cells, which are far apart such that the radio interference between them is tolerable. By reducing the size of cells, the cellular network is able to increase its capacity, and therefore to serve more subscribers.

For the channels assigned to a cell, some are forward (or downlink) channels which are used to carry traffic from the base station to mobile stations, and the other are reverse (or uplink) channels which are used to carry traffic from mobile stations to the base station. Both forward and reverse channels are further divided into control and voice (or data) channels. The voice channels are for actual conversations while the control channels are used to help set up conversations.

A mobile station communicates with another station, either mobile or land, via a base station. A mobile station cannot communicate with another mobile station directly. To make a call from a mobile station, the mobile station first needs to make a request using a reverse control channel of the current cell. If the request is granted by MSC, a pair of voice channels will be assigned for the call. To route a call to a mobile station is more complicated. The network first needs to know the MSC and the cell in which the mobile station is currently located. How to find out the current residing cell of a mobile station is an issue of location management. Once the MSC knows the cell of the mobile station, the MSC can assign a pair of voice channels in that cell for the call. If a call is in progress when the mobile station moves into a neighboring cell, the mobile station needs to get a new pair of voice channels in the neighboring cell from the MSC so the call can continue. This process is called handoff (or handover). The MSC usually adopts a channel assignment strategy, which prioritize handoff calls over new calls.

III. Location Management Schemes

The continued growth of wireless communication systems, and expansion of network subscription rates, signals increased demand for the efficient location management. A cellular communication system must track the location of its users in order to forward calls to the relevant cell within a network. Cells within a network are grouped into *Location Areas* (LAs).

Users are free to move with a given location area without updating their location, informing the network only when transitioning to a new LA. If a call is to be forwarded to a user, the network must now page every cell within the location area to determine their precise location. Network cost is incurred on location updates and paging, the balance of these defining the field of *Location Management* (LM). So there are two basic operations involved with location management:

- **Location Updating:** Informing the network of a devices location.
- **Paging:** Polling a group of cells to determine the precise location of a device [3].

The frequency of updates and paging messages relates closely to user movement and call arrival rates, along with network characteristics such as cell size. As mobile devices move between cells in a network they must register their new location to allow the correct forwarding of data. Continual location updates can be a very expensive operation, particularly for users with comparatively low call arrival rates. This update overhead not only puts load on the core (wired) network but also reduces available bandwidth in the mobile spectrum. Importantly, unnecessary location updating incurs heavy costs in power consumption for the mobile device. The study of location management aims to reduce the overhead required in locating mobile devices in a cellular network.

An additional area of consideration in location management is the mobility model used to estimate user movement in the network, aiding in the optimization of location updating and paging schemes.

A. Location Update

A location update is used to inform the network of a mobile device's location. This requires the device to register its new location with the current base station, to allow the forwarding of incoming calls. Each location update is a costly exercise, involving the use of cellular network bandwidth and core network communication; including the modification of location databases. A wide variety of schemes have hence been proposed to reduce the number of location update messages required by a device in a cellular network. Location update schemes are often partitioned into the categories of **static** and **dynamic** [3].

A location update scheme is static if there is a predetermined set of cells at which location updates must be generated by a mobile station regardless of it mobility. A scheme is dynamic if a location update can be generated by a mobile station in any cell depending on its mobility. Static schemes offer a lower level of cost reduction but reduced computational complexity. Dynamic schemes adjust location update frequency per user and are hence able to achieve better results, while requiring a higher degree of computational overhead.

1) Static Location Update Schemes

Static schemes define the frequency and occurrence of location updates independently from any user characteristics. Such static mechanisms allow efficient implementation and low computational requirements due to the lack of independent user tracking and parameterization. The various static location update schemes are discussed below:

a) Always-Update vs. Never-Update

The always-update strategy is the simplest location update scheme; performing a location update whenever the user moves into a new cell. The network always has complete knowledge of the user's location and requires no paging to locate the user when an incoming call arrives. The always-update scheme performs well for users with a low mobility rate or high call arrival rate. It performs quite poorly for users with high mobility however, requiring many location updates and excessive use of resources. While not used in practice, this scheme forms the basis of many more complex location management mechanisms, such as location area [12] and profile-based update [10].

The never-update scheme is the logical counterpart to always-update, never requiring the mobile device to update its location with the network. This entails no location update overhead but may result in excessive paging for a large network or for high call arrival rates. These two strategies represent the extremes of location management - always-update minimizing paging cost with never-update minimizing location update cost.

b) Location Areas

The location area topology is widely used to control the frequency of location updates, both in the current GSM [14] and IS-41 [6] telecommunications architectures. Here the network is partitioned via the amalgamation of groups of cells into larger meta-cells, or *location areas*. The scheme then functions very similarly to the always-update mechanism – mobile devices only update their location when leaving their current location area [12]. This partitioning is shown in Fig. 2, with four separate location areas.

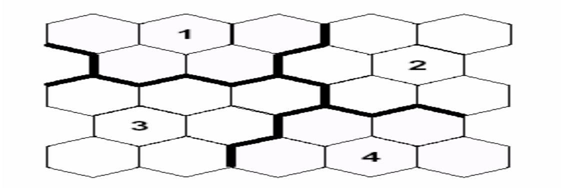

Fig. 2 Network partitioned into location areas

If the network knows the users current location area, the paging required to locate a user is confined to the meta-cell within which the user resides. The location update process may be instigated **periodically** or, more commonly, on location area **boundary crossings**. The periodic location-updating scheme is the simplest to implement, merely requiring the mobile device to send a location update message containing its current location area at regular time intervals. The methodology used here captures none of the details of user movement however, and enforces an update frequency that may be highly unsuitable given the users current rate of movement. It also offers no guarantees that the network will have knowledge of the exact location area the user resides in, requiring sophisticated paging across multiple location areas. The boundary crossing method is more precise; updating the user location only when crossing to another location area. This method has its own inherent weaknesses however, particularly when a user makes multiple crossings across a location area boundary [5].

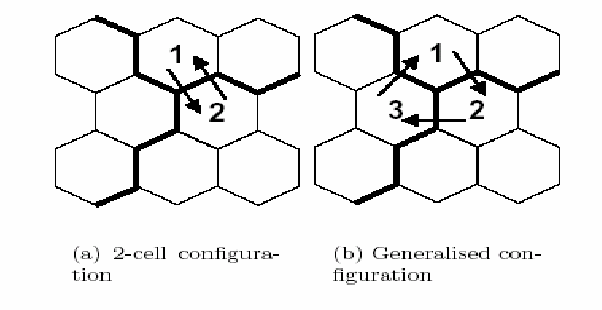

Fig. 3 Cell ping-pong effect

The ping-ponging effect, illustrated in Figures 3(a) and 3(b), is the major weakness of location area schemes. Here a user moves repeatedly between the boundaries of two or more location areas, inducing a high location update rate with comparatively low physical mobility.

A number of schemes have been proposed to address this problem,, such as the Two Location Area (TLA) [10] and Three Location Area (TrLA) [5] mechanisms. These assign multiple location areas to a mobile device, requiring location update only when one of the respective two or three location areas has been exited.

2) Dynamic Location Update Schemes

Dynamic location update schemes allow per-user parameterization of the location update frequency. These account for the dynamic behavior of users and may result in lower location management costs than static schemes. Unlike static location management strategies, a location update may be performed from any cell in the network, taking into consideration the call arrival and mobility patterns of the user.

a) Threshold-Based

In threshold-based schemes each mobile device maintains a particular parameter, updating its location when the parameter increases beyond a certain threshold. The update threshold may be optimized on a per-user basis, dependent on a user's call-arrival and mobility rate [1]. The most common thresholding schemes are time-based, movement-based and distance-based; evaluated in [3].

Time-Based Update: The time-based strategy requires that users update their location at constant time intervals. This time interval may then be optimized per-user, to minimize the number of redundant update messages sent. This only requires the mobile device to maintain a simple timer, allowing efficient implementation and low computational overhead. Fig 4 illustrates this scheme, with the updates (U1, U2 and U3) performed at each time interval Δt, regardless of individual movements. It was found in [11] that signalling load may be reduced using a time-based scheme, with the network additionally able to determine if a mobile device is detached if it does not receive an update at the expected time.

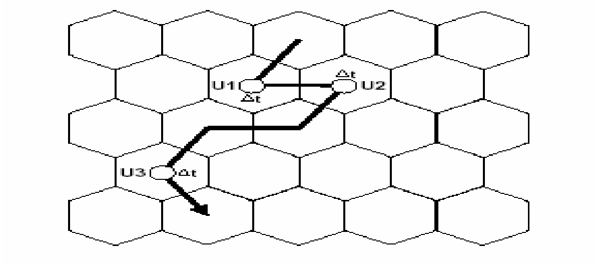

Fig. 4 Time-based location update

[13] analyses the performance of the time-based location update scheme, independent of user mobility constraints, and finds it to outperform the location area method used in current systems. Thus this scheme does however entail a high degree of overhead in a number of situations, such as when a user has only moved a very small distance or has not moved at all.

Movement-Based Update: This scheme requires mobile devices to update their location after a given number of boundary-crossings to other cells in the network. This boundary-crossing threshold may be assigned per-user, optimized for individual movement and call arrival rates [1].

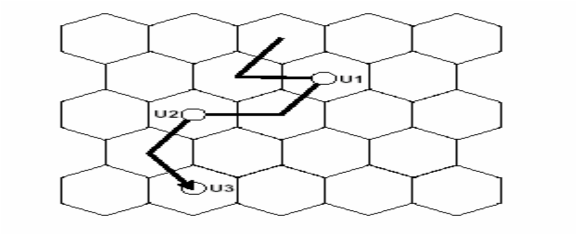

Fig. 5 Movement-Based location update

Fig. 5 shows a movement-based scheme, with a movement threshold of two. Here the device updates its location every two crossings between cells. The required paging area is restricted to a neighborhood of radius equal to the distance threshold around the last updated location. The paging area requirement is reduced through this scheme, although unnecessary updates may still be performed as a result of repeated crossings over the same cell boundary. [3] finds the movement-based scheme to perform better than the time-based scheme under a memory-less (random) movement model. Under Markovian analysis, the time-based scheme is found to perform better for special cases of extremely low update rates, with the movement-based scheme performing slightly better for the majority of user classes.

Distance-Based Update: In a distance-based scheme the mobile device performs a location update when it has moved a certain distance from the cell where it last updated its location. Again, this distance threshold may be optimized per-user according to movement and call arrival rates.

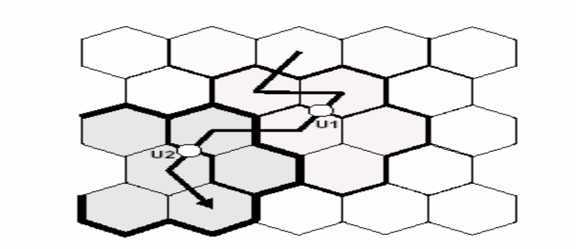

Fig. 6 Distance-based location update

A distance-based update system is shown in Fig. 6. Here the location is only updated when the user travels beyond a certain radii from the previous update location. This scheme has the benefit of not requiring an update when a user repeatedly moves between a small subset of cells, provided these cells reside within the distance-threshold radius. [3] find the distance-based scheme to significantly outperform the time- and movement-based schemes under both simulation and theoretical analysis.

Distance-based update strategies are quite difficult to implement in a real-world network [7]. Each mobile node is required to track its location and determine the distance from the previously updated cell. This not only requires the device to retain information about its starting position, but also to possess some concept of a coordinate system, allowing the calculation of distance between the two points. This coordinate system is a non-trivial requirement in a heterogeneous network, where cell adjacencies and distances may not be clearly defined.

b) Profile-Based

Under a profile-based scheme the network maintains a profile for each user in the network, based on previous movements, containing a list of the most probable cells for the user to reside within [10]. On a location update the network sends this list to the mobile device, forming what may be considered a complex location area. The mobile device updates its location only when entering a cell not contained in the list. [16] found that when movement exhibits a medium to high predictability, the location management cost is lower than that provided by schemes with a more general representation of location area. When low predictability is encountered, the high overhead of sending a large cell list to users outweighs the cost reduction provided by the profile-based scheme.

B. Paging

While mobile devices perform updates according to their location update scheme, the network needs to be able to precisely determine the current cell location of a user to be able to route an incoming call. This requires the network to send a paging query to all cells where the mobile device may be located, to inform it of the incoming transmission. It is desirable to minimize the size of this paging area, to reduce the cost incurred on the network with each successive paging message [7]. Ideally the paging area will be restricted to a known group of cells, such as with the currently implemented location area scheme [12]. An optimum paging area size calculation involves a trade-off between location update cost and paging cost. This technique is used in many location management schemes to reduce the location management costs incurred. The most commonly used paging schemes are summarized below. These have seen extensive use in real-world telecommunications networks [7].

1) Simultaneous Paging

The simultaneous paging scheme, also known as blanket paging, is the mechanism used in current GSM network implementations. Here all cells in the users' location area are paged simultaneously, to determine the location of the mobile device. This requires no additional knowledge of user location but may generate excessive amounts of paging traffic. Implementations of simultaneous paging favor networks with large cells and low user population and call rates. This scheme does not scale well to large networks with high numbers of users, necessitating the development of more advanced paging techniques.

2) Sequential Paging

Sequential paging avoids paging every cell within a location area by segmenting it into a number of *paging areas*, to be polled one-by-one. It is found in [15] that the optimal paging mechanism, in terms of network utilization, is a sequential poll of every cell in the location area individually, in decreasing probability of user residence. The individual delays incurred in this scheme may be unacceptable however, and hence it is suggested that paging areas are formed from a larger number of cells. The number of cells per paging area is a factor which needs to be optimized and may lead to excessive call delays, particularly in large networks. The order by which each area is paged is central to the performance of the sequential paging scheme. [15] suggests several methods to determine the ordering of paging areas in a sequential scheme. The simplest ordering constraint is a purely random assignment, where each paging area is polled in a random order. While this reduces the total number of polling messages over a blanket scheme, it is far from optimal. Schemes favoring paging areas located geographically closer to the previously updated location are found to further reduce the total number of paging messages required. These schemes necessitate knowledge of the geographical structure of the network however, and may perform poorly for high movement rates.

3) Intelligent Paging

The intelligent paging scheme is a variation of sequential paging, where the paging order is calculated probabilistically based on pre-established probability metrics [15]. Intelligent paging aims to poll the correct paging area on the first pass, with a high probability of success. This efficient ordering of paging areas requires a comprehensive knowledge of user residence probabilities. [3] discusses that the success of an intelligent paging scheme hinges on the ability of an algorithm to calculate the probability of a user residing in each cell of the current location area.

An algorithm to calculate the paging area ordering based on a probability transition matrix is presented in [8], claiming to result in the optimal poll ordering. The computational overhead involved in this scheme is quite high however, requiring the computation of an $n \times n$ matrix for a system with n cells, and hence is infeasible for a large cellular network.

C. Comparison of Location Update and Paging

Location update involves reverse control channels while paging involves forward control channels. The total location management cost is the sum of the location update cost and the paging cost. There is a trade-off between the location update cost and the paging cost. If a mobile station updates its location more frequently (incurring higher location update cost), the network knows the location of the mobile station better. Then the paging cost will be lower when an incoming call arrives for the mobile station. Therefore both location update and paging costs cannot be minimized at the same time. However, the total cost can be minimized or one cost can be minimized by putting a bound on the other cost. For example, many researchers try to minimize the location update cost subject to a constraint on the paging cost.

The cost of paging a mobile station over a set of cells or location areas has been studied against the paging delay [17]. There is a trade-off between the paging cost and the paging delay. If there is no delay constraint, the cells can paged sequentially in order of decreasing probability, which will result in the minimal paging cost. If all cells are paged simultaneously, the paging cost reaches the maximum while the paging delay is the minimum. Many researchers try to minimize the paging cost under delay constraints.

D. Mobility Models

The mobility pattern also plays an important role when evaluating the performance of a location management scheme. A mobility model is usually used to describe the mobility of an individual subscriber. Sometimes it is used to describe the aggregate pattern of all subscribers. The following are several commonly used mobility models.

- **Fluid Flow**

The fluid flow model has been used in [18] to model the mobility of vehicular mobile stations. It requires a continuous movement with infrequent speed and direction changes. The fluid flow model is suitable for vehicle traffic in highways, but not suitable for pedestrian movements with stop-and-go interruption.

- **Random Walk**

The random walk mobility model is regularly used to model the movements of users in a cellular network. It assumes that the direction of each user-movement is completely random, and hence each neighboring cell may be visited with equal probability. This model is easily implemented, as it requires no state information to predict the next cell occupied by a user. [3] refers to random walk as the memory-less movement model owing to this property. [4] presents a two-dimensional random walk model, capable of representing both cell crossing rate and cell residence time, for either square or hexagonal cells. This model represents a large amount of movement information while remaining highly tractable, due to the random approach. While the accuracy of a random model is higher than may be expected, given the complexity and irregularity of real-world networks the simplistic scheme fails to capture sophisticated movement patterns exhibited by real users.

- **Markovian**

The Markovian mobility model defines distinct probabilities for movement from a given cell to each of its neighbors. These probabilities are dependent on individual user movement histories. This scheme is based on the assumption that a user moving in a given direction will continue in a similar direction with greater probability than a divergence from course. Markovian models are able to capture movement patterns well but need an established concept of user movement from which to base probabilities [2].

- **Activity-Based**

Activity-based models are the most recent and comprehensive in representing movement throughout a cellular network. They aim to reflect individual user behaviors based on parameters such as time of day, location and destination [9].

IV. Conclusion & Future Work

This paper has surveyed research on location management in cellular networks. Location management involves two operations: location update and paging. Paging is performed by the network to find out the cell in which a mobile station is located so the incoming call for the mobile station can be routed to the corresponding base station. Location update is done by the mobile station to let the network know its current location. There are three metrics involved with location management: location update cost, paging cost, and paging delay. The overall conclusion of this study is that static Location Management schemes are becoming increasingly out of date. While they are still used where cost or resource availability is an issue, upgrading to dynamic schemes is preferable. However, dynamic Location Management schemes have not yet become a panacea, as many schemes are still only theories, being insightful developments, but not useful under real-world conditions. Consequently, dynamic Location Management must continue to be researched and improved.

Given the increasing number of cellular users in an on-demand world, and transitions occurring from 2G to 3G and beyond, Location Management must and will continue to improve both Location Update and paging costs, while allocating appropriate Location Areas in a practical, implementable fashion.

References

[1] I.F. Akyildiz, J.S.M. Ho & Yi-Bing Lin. Movement-based location update & selective paging for pcs N/Ws. Networking, IEEE/ACM Transactions on, 4(4):629–638, 1996.

[2] I.F. Akyildiz, Yi-Bing Lin, Wei-Ru Lai, Rong-Jaye Chen. A new random walk model for pcs networks. Selected Areas in Communications, IEEE Journal on, 18(7):1254–1260, 2000.

[3] A. Bar-Noy, I.Kessler, M.Sidi. Mobile users: To update or not to update? Wireless Networks, 1(2):175–185, 1995.

[4] K. Chiang, N. Shenoy. A 2-d random-walk mobility model for LM studies in wireless networks. Vehicular Technology, IEEE Transactions on, 53(2):413–424, 2004.

[5] P.G. Escalle, V.C. Giner, and J.M. Oltra. Reducing location update and paging costs in a pcs network. Wireless Communications, IEEE Transactions on, 1(1):200–209, 2002.

[6] EIA/TIA. Cellular radio telecommunications intersystem operations. Technical report, July 1991.

[7] B. Furht and M. Ilyas. Wireless internet handbook - technologies, standards and applications. 2003.

[8] Daqing Gu and S.S. Rappaport. A dynamic location tracking strategy for mobile communication systems. In Vehicular Technology Conference, 1998. VTC 98. 48th IEEE, volume 1, pages 259–263 vol1, 1998.

[9] Thomas Kunz, Atif A. Siddiqi, and John Scourias. The peril of evaluating location management proposals through simulations. Wirel. Netw., 7(6):635–643, 2001.

[10] Yi-Bing Lin. Reducing location update cost in a pcs network. Networking, IEEE/ACM Transactions on, 5(1) 1997.

[11] Z. Naor, H. Levy. Minimizing the wireless cost of tracking mobile users: an adaptive threshold scheme. In INFOCOM ’98. Proceedings. IEEE, volume 2,1998.

[12] S. O., S. Onoe, S. Yasuda, A. Maebara. A new location updating method for digital cellular systems. In Vehicular Technology Conference, 1991. ’Gateway to the Future Technology in Motion’, 41st IEEE, pages 345–350, 1991.

[13] C. Rose. Minimizing the avg. cost of paging & regd: a timer-based method. Wirel. Netw., 2(2):109–116, 1996.

[14] M. Rahnema. Overview of the gsm system and protocol architecture. Communications Magazine, IEEE, 31(4), 1993.

[15] C. Rose, Roy Yates. Minimizing the avg. cost of paging under delay constraints. Wirel. Netw., 1(2):211–219, 1995.

[16] H. Xie, S. T, and D.J. Dynamic location area management & performance analysis. In Vehicular Tech. Conf., 1993 IEEE 43rd, pages 536–539, 1993.

[17] C. Rose & R. Yates, Minimizing the avg. cost of paging under delay constraints, Wireless Networks, 1 (1995) 211-219.

[18] H. Xie, S. Tabbane, and D. J., Dynamic Location Area Mgt. and Performance Analysis, Proc. 43rd IEEE Vehicular Tech. Conference, Secaucus, NJ, May 1993, pp. 536-539.