Location Management in Wireless Data Networks

Abstract

Location Management (LM) has become a diverse and broad field for research, and Location Based Services (LBS) are the "Application Layer" of LM. Research areas for LM include LM in Wireless Networks, Ad-hoc networks, WiFi, 802.15, WiMax, and mostly; cellular networks. The integration of IP (especially IPv6) into LM is the hottest topic of today. This survey paper tends to introduce the reader to the latest research on LM in Wireless Data Networks and explore the various technologies under research and construction.

Table of Contents

1. Introduction

Mobile wireless devices with wireless connection facilities are changing the way people think about the use of computing and communication. These wireless devices can communicate with one another even though the user is mobile. People carrying a mobile computer will be able to access information regardless of the time and their current position. Over 100 million wireless Internet users were recorded as of September 2003 with the majority in Japan and Korea, while fast growth rates were recorded in Europe. Significant growth is expected in specialized mobile services such as driving directions, traffic report, tour guides, and commerce services such as mobile shopping. However, Location Management (LM) will be an important issue in these situations because wireless devices can change location while connected to a wireless network. New strategies must be introduced to deal with the dynamic changes of a mobile device's network address.

The ability to change locations while connected to the network creates a dynamic environment. This means that data, which is static for stationary computing, becomes dynamic for mobile computing. There are a few questions that must be answered when looking at a LM scheme. What happens when a mobile user changes location? Who should know about the change? How can you contact a mobile host? Should you search the whole network or does anyone know about the mobile users moves?

LM schemes are essentially based on users' mobility and incoming call rate characteristics. The main task of LM is to keep track of a users' location all the time while operating and on the move so that incoming messages (calls) can be routed to the intended recipient.

LM consists mainly of:

1. **Location Tracking and Updating (Registration):** A process in which an end-point initiates a change in the *Location Database* according to its new location. This procedure allows the main system to keep track of a user's location so that for example an incoming call could be forwarded to the intended mobile user when a call exists or maybe bring a user's profile near to its current location so that it could provide a user with his/her subscribed services.
2. **Location Finding (Paging):** The process of which the network initiates a query for an end-point's location. This process is implemented by the system sending beacons to all cells so that one of the cells could locate the user. This might also result in an update to the location register.

As we can see, the main difference between location tracking and paging is in who initiates the change. While location tracking is initiated by a mobile host, paging is initiated by the base system. Most LM techniques use a combination of location tracking and location finding to select the best trade-off between the update overhead and the paging delay.

LM methods are classified into two groups:

1. Group one includes methods based on network architecture and algorithms, mainly on processing capabilities of the system.
2. Group two includes methods based on learning processes (i.e. which require the collection of statistics on subscribers' mobility behavior). This method emphasizes the information capabilities of the network.

For LM purposes, a wireless network usually consists of *Location Areas* (LAs) and *Paging Areas* (PAs). While LAs are a set of areas over which location updates take place, PAs are a set of areas over which paging updates take place. Usually, LAs and PAs are contiguous, but that's not the case always. In addition, a LA usually contains several PAs, see Figure 1.1.
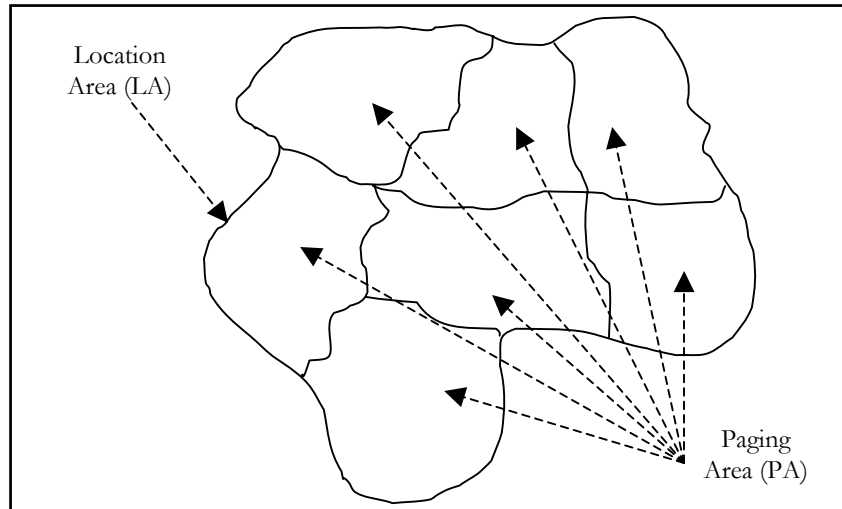
Figure 1.1 Location Areas vs. Paging Areas

As the size of the LA increases, the cost of paging will also increase as more PAs are to be paged to find a called mobile host. On the other hand, reducing the size of a LA will increase the number of crossings per unit time. Hence, the cost of location update or registration will rise. Both paging and location updates consume scarce resources like wireless network bandwidth and power of mobile hosts. Each has a significant cost associated with it. So, LA planning is to be based on a criterion that guarantees the total signaling load, which comprises paging and registration, is kept under tolerable limits. Therefore, it is characterized by the trade-off between the number of location updates and the amount of paging signaling that the wireless network has to deal with.

Having discussed briefly about LM, updating, and paging, sections 3 and 4 go more deep into discussing updating and paging in more detail. The next section talks briefly about Location Based Services (LBS) and their impact on LM. The aim of the next section is to provide the reader with an overview as to why LM is necessary.

2. Location Based Services

Although Location Based Services (LBS) have been an issue in the field of mobile communications for many years, no common definition has been devised for it. The terms *location-based service*, *location-aware service*, *location-related service*, and *location service* have been used interchangeably. These various terms have lead to several definitions of LBS. [Schiller04] defines LBS as "services that integrate a mobile device's location or position with other information so as to provide added value to a user". The GSM association has defined LBS as "services that use the location of the target for adding value to the service, where the target is the entity to be located (this entity isn't necessarily the user of the service)". The 3rd *Generation Partnership Project (3GPP)* defines LBS as "a service provided by a service provider that utilizes the available information of the terminal". One of the most traditional examples of LBS is Global Positioning System (GPS). GPS was used by the US Department of Defense since the 1970s. However, in the 1980s, GPS became available freely for the industry worldwide.

Not only does user location allow companies to conceive completely new service concepts (e.g. tracking applications), but it also has the potential to make many messaging and mobile Internet services more relevant to customers as information is adjusted to context (e.g. weather information adjusted to the region one is in). Finally, location information can considerably improve service

usability. The next few subsections classify LBS and give out a few examples. Finally, an overview on research on LBS is provided.

2.1 Classification of LBSs

LBSs can be classified as *Reactive LBSs* and *Proactive LBSs*:

1. **Reactive LBSs:** Reactive LBSs are always explicitly activated by the user. The interaction between LBS and users is roughly as follows: the user first invokes the service and establishes a service session, either via a mobile device or a desktop PC. The user then requests for certain functions or information whereupon the service gathers location data (either of the user or of another target person), processes it, and returns the location-dependent result to the user. This request/response cycle may be repeated several times before the session is finally terminated. Thus, a reactive LBS is characterized by a synchronous interaction pattern between user and service.

2. **Proactive LBSs:** Proactive LBSs are automatically initialized as soon as a predefined location event occurs, for example, if a user enters, approaches, or leaves a certain point of interest or if he/she approaches, meets, or leaves another target. Thus, proactive services are not explicitly requested by the user, but the interaction between them happens asynchronously. In contrast to reactive LBSs where the user is only located once, proactive LBSs require to permanently track a user in order to detect location events.

2.2 Applications of LBSs

LBSs are often used via web browsers and are considered a particular type of web services. A representative application example of LBSs is that of "Personalized Web Services for the Olympic Games in Beijing in 2008". LBSs are mainly used in three areas: military and government industries, emergency services, and the commercial sector.

As was mentioned earlier, the first location system in use was the satellite based GPS that allows for precise localization of people and objects of up to 3 meters accuracy. GPS is an example of the first area of LBSs, i.e. military and government industries.

Besides the military use of location data, emergency services have turned out to be an important field. Every day, 170000 emergency calls are made in the USA, and 1/3 of these originate from mobile phones, and in most cases, from people who don't know their exact location so as to guide the emergency provider with directions. As a result, the US Federal Communications Commission (FCC) set an October 2001 deadline for commercial wireless carriers to provide the caller's location information in a 911 emergency call. This means that when placing an emergency call from a mobile device, a caller's device position is automatically transmitted to the closest emergency station. Consequently, people in such situations don't have to explain at length where they are but rather are located in seconds. However, none of the carriers were able to meet the deadline of the FCC, and the deadline date was left open. It is expected for a few more years before the entire system is implemented with full coverage.

The third and final application area of LBSs is commercial. For some time, marketers have been unsure whether lower levels of accuracy as they are obtained from well Cell-ID (a mobile positioning system) would be sufficient to launch compelling consumer and business services. Yet, early service examples show that the accuracy level required depends very much on the service. Even with Cell-ID, location information can successfully be integrated by operators into many existing and new applications that enhance current value propositions and usability.

2.3 LBSs in Research

In research, LBSs are often considered to be a special subset of *context-aware services* (from where the term *location-aware service* has its origin). Generally, context-aware services are defined to be services that automatically adapt their behavior to one or several parameters reflecting the context of a target. These parameters are termed *context information*. The set of potential context information is broadly categorized and, as depicted in Figure 2.1, may be subdivided into personal, technical, spatial, social, and physical contexts. It can be further classified as *primary* and *secondary contexts*. *Primary contexts* comprises any kind of raw data that can be selected from sensors, microphones, accelerometers, location sensors … This raw data may be refined by combination, deduction, or filtering in order to derive high-level context information, which is termed *secondary context* and is more appropriate for processing by a given *context-aware* service.
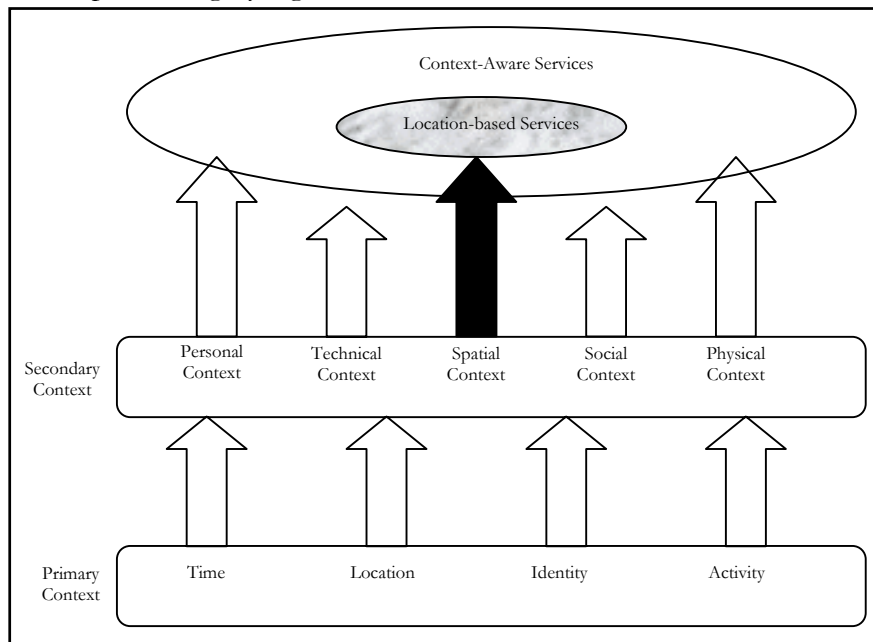

Figure 2.1 Context-Aware and Location-Based

As can be derived from Figure 2.1, LBSs are always context-aware services because location is one special case of context information. In many cases, the concept of primary and secondary contexts can also be applied to LBSs, for example, when location data from different targets are related or the history of location data is analyzed to obtain high-level information such as the distance between targets or their velocity and direction of motion. Therefore, there is no sharp distinction between LBSs and context-aware services. In many cases, context information that is relevant to a service, for example, information such as temperature, pollution, or audibility are closely related to the location of the target to be considered. Hence, its location must be obtained first before gathering other context information.

In recent years, many location service protocols have been developed for Ad-hoc networks, including the Grid Location Service (GLS), the Simple Location Service (SLS), and the Legend Exchange and Augmentation Protocol (LEAP). In all of the existing location services, when a mobile node's location is needed, the previously saved information in the location table is used. [Luo05] proposes the Prediction Location Service (PLS), a service in which a mobile node uses information about its previous state to predict its future state. Results show that PLS has lower overhead and lower location error than GLS, SLS, and LEAP.

Much of the current location prediction research is focused on generalized location models where the geographic extent is divided into regular-shape cells. These models are not suitable for certain LBSs whose objective is to compute and present on-road services because a cell may contain several roads while the computation and delivery of a service may require the exact road on which the user is driving. [Karimi03] proposes a new model called *Predictive Location Model* (PLM) to predict locations in LBSs with road-level granularities. The premise of PLM is geometrical and topological techniques allowing users to receive timely and desired services. However, the proposed model has been analyzed only with synthetic data.

The topic of LBSs is very vast and diverse. In this section, key issues of LBS were pointed out. However, the reader is highly encouraged to see [Schiller04] and [Kupper05] for more information on this topic. Next, we talk about location tracking and updating and get to know more about the various static and dynamic updating strategies.

3. Location Tracking and Updating (Registration)

In updating a user's location, a key issue is how often the update process should occur. If the updating process is less than required, the main system would get into the paging phase and try to search for the intended user by sending beacons to all cells. This results in significant delivery delays. On the other hand, if the updating process is done more than required, uplink radio capacity and battery power would be exhausted for both the system and mobile hosts.

As mentioned in the introduction part of this paper, a wireless network usually consists of *Location Areas* (LAs) and *Paging Areas* (PAs). LAs are a set of areas over which location updates take place. There exists several location updating methods based on the LA structuring. The two most commonly used LA management schemes are:

1. Periodic location updating (it has the inherent drawback of having excessive resource consumption which is unnecessary at times)
2. Location updating on LA crossing.

Figure 3.1 shows a classification of possible update strategies used. As can bee seen in Figure 3.1, updating strategies can be classified as Static Strategies and Dynamic Strategies. Sections 3.1 and 3.2 explain Figure 3.1 in more details.
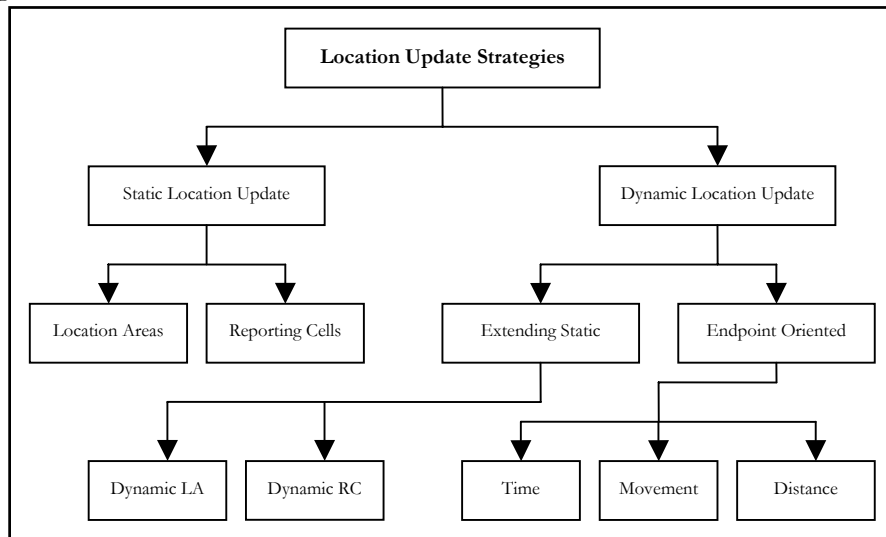

Figure 3.1 Classifications of Location Update Strategies

3.1 Static Update Strategies

In this approach, there are specific areas in which an update could take place. If a mobile host enters any one of these areas, an update takes place (though there might be instances in which an update doesn't happen every time).

Two approaches of static updating are as follows:

1. **Location Areas (LAs):** Also referred to as *Paging Areas* or *Registration Areas*. In this scheme, service areas are created with each area considered a LA. Only when a mobile host moves from one LA to another that an update to its location in the *Location Database* is taken place.
2. **Reporting Cells:** Also referred to as *Reporting Centers*. In this scheme, updates take place at specific centers (cells) in the network. Only when a mobile host gets re-located to one of these centers that an update takes place.

The main drawback to *Static Update Strategies* is that they don't accurately account for user mobility and frequency of incoming calls.

3.2 Dynamic Update Strategies

In this strategy, a mobile host determines when an update should take place based on its movement, frequency of incoming messages, signal strength … and other factors. A natural approach to dynamic strategies is to extend the *Static Update Strategies* to integrate call and mobility patterns. Several proposed *Dynamic Update Strategies* include:

1. Depending on the incoming call arrival rate and mobility, the size of a mobile host's LA is determined. Analytical results for this approach have shown that this strategy is an improvement over *Static Update Strategies* when call arrival rates are user-dependent or time-dependent.
2. An asymmetric distance-based cell boundary system with cell search order optimization that uses a combination of information of the most recent update that took place along with the direction of motion.
3. Time-based location updates that take place every T seconds.
4. Movement-based location updates that take place after every M cell crossings.
5. Distance-based location updates that take place whenever the distance covered exceeds D. This kind of strategy is the toughest to implement since it requires information about the topology of the wireless network. However, it has been shown that for memory-less movement patterns on a ring topology, this strategy outperforms both Time-based and Movement-based schemes.

Having said enough about location updating, the next step is to get to know more about location finding.

4. Location Finding (Paging)

As mentioned earlier, *Paging* is the process of which the network initiates a query for an end-point's location. This process is implemented by the system sending beacons to all cells so that one of the cells could locate the user. This might also result in an update to the location register. Various paging schemes have been devised and implemented in real life, and some of the new schemes integrate mobile IP into them.

4.1 Paging Schemes

Several paging schemes have been developed, some of these include:

1. A method for balancing both *Location Updating* and *Paging*. Probability distribution techniques along with the lower bounds on the average cost of paging and Poisson incoming-call arrival model are used to formulate the paging and registration optimization problem in terms of timeout parameters.
2. Minimizing the amount of bandwidth expanded in locating a mobile host by using probability distribution on user location.
3. A mobile user mechanism that incorporates a distance-based location updates scheme and a paging mechanism that satisfies predefined delay requirements.
4. A mobility tracking mechanism that combines a movement-based location update policy with a *Selective Paging* scheme. The *Selective Paging* scheme decreases the location tracking cost under a small increase in the allowable paging delay.
5. The application of multiple steps paging strategy.
6. Locating mobile users by using paging costs and delay bounds. Paging costs are associated with bandwidth utilization while delay bounds influences call setup time. Reverse, semi-reverse, and uniform paging schemes provide a simple way for portioning the service area and decrease the paging costs based on each mobile user's location probability distribution.
7. *Reverse Paging* is a scheme designed for a situation where the called mobile host is most likely a few PAs away.
8. *Semi-reverse Paging* is a scheme where a set of PAs is created in a non-increasing order of location probabilities.
9. *Uniform Paging* is a scheme in which a LA is partitioned into a series of PAs with each PA consisting of approximately equal number of mobile cells.
10. *Intelligent Paging* will be described in section 4.2.

A few of the paging schemes mentioned above are applied in industry. Despite their widespread use, some disadvantages have been found in these schemes, and a few new intelligent paging schemes have been devised to overcome these disadvantages.

4.2 Intelligent Paging Schemes

It is mathematically proven that the movement of mobile hosts is modeled according to a probability process called the *Ergodic Stochastic Process*. When an incoming call comes to a mobile host roaming in a certain LA, paging is initially performed within that portion of the LA (called PA as mentioned in section 1). Intelligent paging is a multi-step strategy that aims at determining the proper PA within which the mobile host currently roams. Its strategy maps the PAs inside the LA into a probability line at the time of the arrival of the incoming call. This mapping depends on factors such as the mobility of the mobile device, its speed profile, the incoming calls statistics, and the state of the mobile host at that instant. Three scenarios occur here regarding the paging request:

1. If it is detached, the request is cancelled.
2. If it is busy, a relation between the mobile user and the network already exists, and therefore paging isn't required.
3. If it is free, the network proceeds for paging upon receipt of a paging request.

The application of intelligent paging includes the event of paging failures due to wrong predictions of the locations of the called mobile host. In such cases, the called mobile host will be paged in other PAs. Continuous unsuccessful paging attempts may lead to unacceptable paging delays and thus increase the paging cost.

4.3 Intersystem Paging

In a multi-tier wireless service area consisting of dissimilar systems, it is desirable for a mobile host to be able to communicate with the various systems and be able to roam between them efficiently and with no problems. Roaming between different systems can be one of two types:

1. **Intrasystem Roaming:** Refers to a mobile host that moves between different LAs within the same system.
2. **Intersystem Roaming:** Refers to a mobile host that moves between different systems.

4.4 IP Micro-mobility and Paging

Recent research in mobile IP has proposed that IP should take support from the underlying wireless network architecture to achieve good performance for handover and paging protocols. Recent IETF work defines requirements for layer 2 (the *Data Link Layer* of the *OSI Model*) to support optimized layer 3 (the *Network Layer* of the *OSI Model*) handover and paging protocols. The *Data Link Layer* can send notifications to the *Network Layer* that a certain event has happened or is about to happen.

Paging triggers aid a mobile host in entering the dormant mode in a graceful manner and make the best use of paging provided by the underlying wireless architecture.

Having spoken enough about paging and its various flavors, the next section tends to focus on general issues in LM such as security, group-based LM, and Distributed LM. These issues apply to all kinds of Wireless Data Networks and Cellular Networks as well. For interested readers, please refer to [Mukherjee03] for more information on location updating and paging.

5. General Issues in LM

Thus far, we've spoken about the two-stage process of LM: Registration and Paging. Current issues for LM involve database architecture design, transmission of signaling between various components of a signaling network, security, dynamic database updates, querying delays, terminal paging methods, and paging delays. The following are some LM schemes:

Without Location Management

1. Is referred to as the *Level 0 method*.
2. The system doesn't track any mobile devices.
3. Searching for a user is done over the complete radio coverage area and within a specific time limit.
4. Is also referred to as the *Flooding Algorithm*.
5. It is used in:
   a. Paging systems because of the lack of uplink channel allowing mobile hosts to inform their whereabouts to the network.
   b. Small private mobile wireless networks because of their small coverage areas and user populations.
6. Its main advantage is that it is simple to implement because of the absence of a special database.
7. Its main disadvantage is that it doesn't fit into large networks dealing with high number of users and high incoming data exchange rates.

Manual Registration in Location Management

1. Is referred to as the *Level 1 method*.
2. The system is relatively simple to manage because it only requires the management of an indicator that stores the current location of the user.

3. The mobile is relatively simple since its task is limited to scanning the channels to detect paging messages. An example of such a system is telepoint cordless systems.
4. The main disadvantage of this method is that users have to re-register each time they move.

Automatic Location Management using LAs

1. Is referred to as the *Level 2 method*.
2. Widely used and deployed in 1G and 2G cellular systems.
3. Since this method is a LA based method, a home database and several visitor databases are included in the network architecture.

Memoryless-Based Location Management Methods

These methods depend mainly on the processing capabilities of the system. They are based on algorithms and the network architecture.

Memory-Based Location Management Methods

The design of memory-based location management methods has been motivated by the fact that systems perform many repetitive actions that can be avoided if predicted.

Location Management in Next Generation Systems

The next generation in mobility management will enable different mobile networks to interoperate with each other to ensure terminal and personal mobility and global portability of network services. However, in order to ensure global mobility, the deployment and integration of both wired and wireless components is necessary. These future systems will all depend on the usages of Mobile IP. For example, the aim of 4G cellular networks is to deploy Mobile IP in its infrastructure so that users can switch between different access technologies.

5.1 Security

As LM may be viewed as a specific service discovery mechanism, any attack on it is consequently an attack on service discovery. [Sethom05-1] presents S-PALMA, a security mechanism that protects the LM messages. It provides security services such as access control, authentication, message integrity and authentication, confidentiality, and anti-replay. It is intended to cope with identity usurpation, man-in-the-middle, and replay attacks at the edge of the overlay network. It relies on the P2P distributed architecture to provide additional availability.

The S-PALMA architecture consists of a set of Distributed Lookup Servers (DLS), interconnected via heterogeneous networks. These DLSs are organized into an overlay network to publish location information to each other for storage, and to collaboratively resolve queries from mobile hosts.

Applicative interconnections are maintained using Tapestry node insertion and neighbor notification algorithms. In this new virtual space, each DLS is assigned a new identifier generated from its IP address using a hash function (SHA-1). Mobile nodes communicate with DLSs to advertise their presence or submit queries. When a node issues a query to a DLS, it receives the response from that particular DLS.

The secure messages are designed in such a way to cope with mobility constraints (particularly delay) by avoiding heavy public-key operations.

5.2 Group-based Location Management

[Lam04] proposes a *group-based location-updating* scheme (GBL). It is based on the assumption that the number of high cost location update messages from mobile hosts to the location server can be reduced by clustering mobile hosts with similar mobility into a set of groups. A single location report for the whole group is sent to the location server. A leader will be selected to perform location updating on behalf of the whole group to the moving object database. A positive consequence is that mobile hosts no longer need to possess the long-range communication capability with the remote server; location information can be reported via the group leader.

Owing to the dynamic movement of mobile hosts, some of them will occasionally leave their current group and join other groups. These changes in group membership dictate a group management mechanism for a group leader or cluster-head to handle the leave and join events for a mobile host. The more interesting step lies in the selection of an appropriate group and its leader for a wandering mobile host before triggering a join event. This is referred to as the *group finding process*, and the mobile host looking for a group to join as the *group seeker*. Based on this observation, [Lam04] proposes two new *join procedures* for a mobile host to locate for an appropriate joining group.

In the GBL system model proposed in [Lam04], each mobile host *m* is assumed to possess a unique ID and a GPS sensor for keeping track of its existing location and its movement information. The current location of *m* is denoted by $\{x_m, y_m\}$, while the movement information is maintained and represented as a vector $\vec{v}_m = \{v_{xm}, v_{ym}\}$, being resolved into the x and y components. Two mobile hosts are considered as *neighbors* if the Euclidean distance between the two hosts is smaller than the transmission range of these two hosts (i.e., they can communicate in an Ad-hoc mode).

In the GBL scheme proposed in [Lam04], there are two levels of location update occurring: local update and group update. The first level, termed as *local location update*, is about the strategy for reporting location and movement information to the leader of the group by its members. The second level, termed as *group location update*, is about the strategy for reporting the group location information to the stationary location server via the uplink channel.

In the basic join procedure, group seeker *m* must first request for the group information from its neighbors, by broadcasting a "FIND GROUP" messages. A neighbor receiving the message will reply with a "GROUP INFO" message containing group information pertaining to its current group. Group information received from the neighbors includes leader's host ID, predicted group location and velocity according to the latest updated time of this group information, and the group range **r**. The hosts that are the neighbors of the group seeker will report their group information back after they receive the request message. As the mobile host population increases, the number of neighbors surrounding the group seeker increases. Thus, many reply messages containing information about the relevant groups are produced for a single request message originated from a single group seeker. Among the reply messages, a significant portion of them, i.e., those originated from the neighbors belonging to the same group, are basically identical. As a result, unnecessary reply messages are generated from the neighbors. This situation should be rectified.

The first improved join procedure restricts the set of mobile hosts that are eligible to replying the group information back to the *group seeker*. In particular, only a group leader or its delegate is eligible to replying the "FIND GROUP" request message. This is called the *leader-only join procedure*, since only the leader will reply.

To further reduce the number of reply messages, a leader that receives the request message can make a judicial decision to reply or not, by evaluating the degree of affinity, s_m, G, between the potential group G and the group seeker *m*. If the degree of affinity is larger than a predefined *filtering threshold*, θ, the leader will reply the group information back to the group seeker. Otherwise, the request will be ignored. As a result, groups with low degree of affinity are further filtered out, since they would unlikely be selected in the end. This constitutes the second proposed join procedure by [Lam04], termed *leader-filter join procedure.*

5.3 Distributed Location Management

The design of location directories whose content changes dynamically raises important issues. Some of them are as follows: when should the location directory get updated? Should the location directory be maintained at a centralized site, or should it be distributed? And in the latest case how should the information be distributed among the location servers? Should the information be replicated across multiple location servers?

Numerous location strategies have been proposed in the recent years. One simple approach is to maintain a central database (home location server HLS) mapping host identification to its current location. The centralized approach does not scale well with highly distributed and heterogeneous systems. On one hand, to lookup an object, the possibly distant HLS must be contacted first. On the other hand, even a move to a nearby location must be registered at a potentially distant home location. Thus, local movement and communications are not well considered. Another drawback of this approach is that it presents reliability problems that make it vulnerable to attacks on the centralized database.

Another approach is for the current node to broadcast a message to all its neighbors with a request for the mobile node's IP address (IP paging). When a node receives such a request, it checks its local database. If it has the information, it responds. Otherwise, it forwards the request to its neighbors that execute the same protocol. However, this "broadcast" approach doesn't scale because of the bandwidth consumed by broadcast messages and necessary computation cycles.

To reduce the cost of broadcast messages, nodes in the network can be organized into a hierarchy, like the Internet Domain Name System (DNS) does. Searches start at the top of the hierarchy and by following forwarding references from node to node, traverse a single path down to the node that administrates the desired data.

DNS is an excellent system for identifying static nodes in the Internet and at the time the web was first used. However, as the Internet evolves towards a more mobile and pervasive environment, a system that statically maps well-known names to locations may not be a good choice. Moreover, DNS requires significant expertise to administer since name servers are difficult and time-consuming to configure.

Recently, new functionalities have been added to the DNS in order to deal with these challenges and handle host mobility such as Dynamic update mechanism. The major reason for discarding this approach is that it has demonstrated relatively poor performance in particular due to the inefficiency of caching.

Rethinking the LM architecture allows us to address some of the current shortcomings in current architectures. An ideal LM scheme should adopt a ZEROCONF-like approach i.e. it would allow dynamic networking in the absence of configuration and administration. The network will be able to insert or remove nodes without any centralized administration. Peer to peer distributed lookup algorithms seem to be a good alternative that meet these goals and avoid the drawbacks of

the previous approaches. Unlike the hierarchical scheme, no node plays a special role - a search can start at any node, and each node is involved in only a small fraction of the search path in the system. As a result, no node consumes an excessive amount of resources while supporting searches. These new algorithms are designed to scale well - they require each node to only maintain information about a small number of other nodes in the system, and they allow the network to self-organize into an efficient overlay structure with little effort.

[Sethom05-2] presents a new LM architecture based on a de-centralized overlay network. It is designed to achieve fast and scalable update/lookup operations in an environment where all nodes equally share the same responsibility. To achieve this goal, the architecture relies on an efficient peer to peer algorithm.

[Lee04] proposes a novel LM architecture that relies on a hierarchy of location database agents to help provide a flexible and distributed system wherein agents rapidly replicate to other base-stations in response to change in user patterns. [Lee04] also proposes an analytical cost model to formulate the LM optimization problem, shows it is NP-complete, and suggests an approximation algorithm (service ability) to solve the optimization.

In Figure 5.1, agent A_0 can store locations of many other agents, resulting in a tree-like structure between different agents: A_5 is a child of A_1, and A_1 is the parent of A_5. The number of children b an agent can have is fixed. This gives a tree structure with branching factor b. The leaf of the tree stores the locations of the mobiles. The upper level nodes store the locations of the location databases directly below them. Each agent in the tree has an individual unique net-mask. Mobile hosts that match its net-mask will be in the sub-tree of this particular agent. An agent can move freely within the network, by binding to the base-station it is migrating to. It is possible that more than one agent is bound to a single base-station.
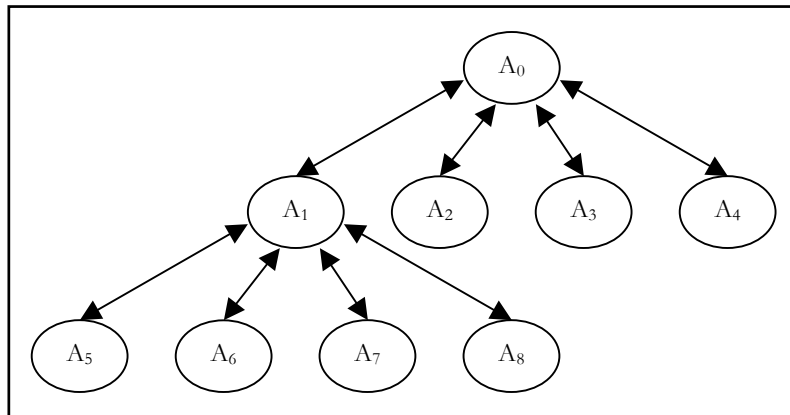

Figure 5.1 Relationship Between Various Agents

Now that we have created the basis necessary for LM, the next sections will explore the advances on LM in Wireless Networks, Mobile Ad-hoc Networks, 802.11 WiFi, 802.15 Bluetooth, and LM with IP.

6. LM in Wireless Networks

Most of the initial research on LM was based on Wireless Networks in general. Despite the fact that the wireless world has become huge and diverse, there is still lot of research done in this field. In this section, some of the advancements in this particular field are presented.

6.1 Heterogeneous Wireless Networks

In heterogeneous wireless networks, mobile users are able to move from one subsystem to another while maintaining access capability to their subscribed services, which refers to global mobility or global roaming. One of the most challenging problems in global roaming management is LM that consists of keeping track of mobile users who leave their home network and roam into foreign networks that use different protocols. In this context, locating a user requires interoperability between several fixed and mobile subsystems that do not necessarily implement the same technology, which may increase the signaling traffic and decrease the network performance. [Assouma05] proposes a model that improves the efficiency of LM in heterogeneous wireless networks in terms of signaling traffic generated during global roaming. Such a model essentially consists of adding at the boundary location area between two different subsystems a specialized equipment called LR-ING (*Location Register and Internetworking Gateway*) that is connected to the *Home Location Register* (HLR) of both subsystems. Numerical results reveal that the proposed scheme enables to reduce the signaling cost generated by the databases by about 45% when compared with other proposed architectures such as the *Boundary Location Register* (BLR) protocol.

6.2 Replication in Wireless Networks

[Hwang05] focuses on the hierarchical scheme with user profile replication. The two-phase algorithm proposed in previous work, though simple, does not provide insights on whether or why it works well. It also discusses the nature of the replica assignment problem in the context and proposes an optimal solution to it. As the optimal solution takes a long time to compute, further assumptions are made to simplify the problem and then are solved via dynamic programming. Finally, rather than determining the replica assignment on a per-user basis, [Hwang05] proposes to first cluster mobile users based on their calling and moving patterns and then perform the replica assignment for each group. This will further improve the efficiency of replica assignment, in addition to reducing the storage requirements. A preliminary experimental result shows that the dynamic programming approach returns better replica assignment in most cases. To further reduce the overhead of storage requirements and execution complexity, clustering techniques that group mobile users with similar mobility behavior are incorporated.

In this section, the focus was on wireless networks. Heterogeneous wireless networks and replication in wireless networks were discussed. The next section focuses on LM in Mobile Ad-hoc networks.

7. LM in Wireless Mobile Ad-hoc Networks

A major challenge faced in Mobile Ad-hoc Networks (MANET) is locating devices for communication, especially with high node mobility and sparse node density. Present solutions provided by Ad-hoc routing protocols range from flooding the entire network with route requests, to deploying a separate LM scheme to maintain a device location database.

LM in wireless Ad-hoc networks can be classified mainly as *Soft LM* and *Hard LM*. Soft LM schemes do not require strict LM strategies and is thus computationally less expensive than standard *Hard LM* schemes. The subsequent subsections explain what *Soft LM* and *Hard LM* schemes are as well as a few examples of what is implemented in today's Ad-hoc networks.

7.1 Soft LM Schemes

[Ghosh04] proposes a novel scheme called *Acquaintance Based Soft Location Management* (ABSLM) in MANET. In ABSLM, nodes make use of the real life concept of making acquaintances

and keeping in touch with them regarding each other's current locations. ABSLM has a twofold aim: to avoid the overhead of flooding and to use a 'soft' location management setup that does not require strict location management strategies and is thus computationally less expensive than standard 'hard' location management schemes. A main difference between this scheme from the LM schemes mentioned in [Philip04-1] and [Philip04-2] is that no assumptions are made regarding either the shape or size of the network terrain or the density of the nodes. Simulation results show that ABSLM not only outperforms existing flooding schemes in terms of throughput, overhead and location discovery latency, but also achieves performance comparable to 'hard' grid based LM schemes with a much lower control overhead.

7.2 Hard LM Schemes

Most of the *Hard LM* schemes are grid based. Grid based means that given a rectangular region of area A, the topography of this area is divided into G logical unit regions (also known as *Order-1* regions) where each node is aware of the size of the topography as well as the size of a unit region. [Philip04-1] and [Philip04-2] propose three *Soft LM* schemes that are grid based: *Scalable Update Based Routing Protocol* (SLURP), *Scalable Location Management* (SLALoM), and *Hierarchical Grid Location Management* (HGRID). However, they all differ from each other in how the logical division is used in LM. Figure 7.1 shows a 3-level hierarchy grid.
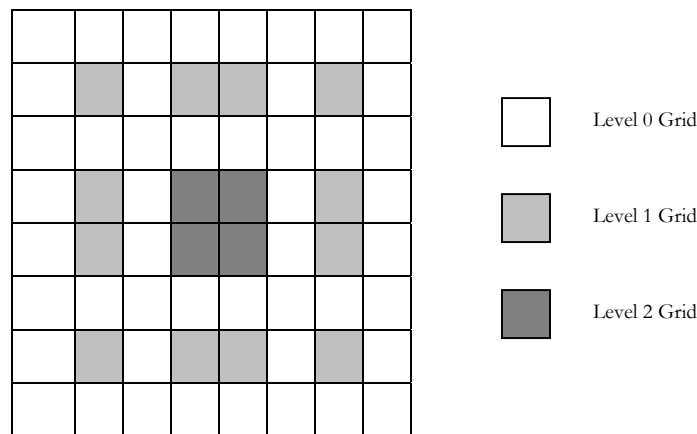

Figure 7.1 A 3-level Hierarchy Grid

In SLURP, each mobile node selects exactly one unit region as its home region by using a mapping function *f*, which uniquely (and randomly) maps its address to the selected home region. The mapping function allows any node to discover another node's home region simply by knowing its address [Philip04-1].

SLALoM combines K^2 Order-1 regions to form Order-2 regions. Each node selects a home region in each Order-2 region via *f* that maps roughly the same number of nodes to each Order-1 region in an Order-2 region. Hence, every node has $O\left(\dfrac{A}{K^2}\right)$ home regions in *A* (note that since the original square cannot be perfectly tiled with Order-2 regions, it is possible that some nodes may not have home regions in the Order-2 regions adjacent to the boundary of A). Also, if a node *u* is present in an Order-1 region R_i, which lies in an Order-2 region Q_i, then all home regions of *u* that lie in or adjacent to Q_i are considered near home regions, while the rest are considered far home regions [Philip04-1].

HGRID defines a hierarchy of K levels (L_1... L_k) on the unit grid regions. Each L_i+1 quadrant is composed of four L_i quadrants. At each level, the leader selection is as follows: for level i ($1 \leq i \leq k - 1$), the top rightmost L_i-1 leader is the ith hierarchical leader of the bottom left L_i grid, top leftmost L_i-1 leader is the hierarchical leader of the bottom right L_i grid, bottom rightmost L_i-1 leader is the hierarchical leader of the top left L_i grid, and the bottom leftmost L_i-1 leader is the hierarchical leader of the top right L_i grid. The top of the hierarchy, (L_k), is defined by the four L_k-1 grids. A node that is physically located in an ith hierarchical grid is responsible for the duty of an ith hierarchical location server [Philip04-1] [Philip04-2].

The results shown in [Philip04-1] indicate that all protocols perform well, and only affect the performance of geographic routing minimally. In particular, (HGRID) performs the best for all practical purposes. While the network size has to be asymptotically large for SLALoM to perform better than SLURP, this may not be realizable in practice for most applications envisaged for MANETs. The results show that key considerations for designing an efficient LM protocol are low control overhead, close servers in proximity, and quick location discovery.

7.3 The Use Of TCP In MANETs

The main LM proposals in MANETs have as a common characteristic two distinct phases: the location query of the position of a destination node and the transmission of a flow toward the destination node. [Ziviani05] proposes to send the initial packet of a flow to learn the position of its destination instead of adopting a dedicated query packet. Such an approach specially benefits TCP flows.

Based on the transport protocol of the flows to be transmitted, a source selects to apply either the TCP tailored approach (for the majority of flows composed by the TCP flows) or the conventional approach (for the remaining flows). As each cooperative node in the network seeks to adopt the method having the lowest cost at each transmission, the adoption of the TCP-tailored approach provides a performance improvement of the LM.

Results show that the proposed approach reduces the transmission costs in LM for the TCP flows, specially favoring short-lived TCP flows. This improvement depends on the existing dispersion between the position of the source and destination nodes with respect to the position of the location server of the destination node.

This section focused on LM in MANETs. Soft LM and Hard LM schemes were presented and explained. The next section will talk about LM in Wireless LANs.

8. LM in WiFi (802.11)

The widespread deployment of wireless LANs and the increasing popularity of light-weight mobile computing devices have lead to an increased interest in location-aware applications and services. The goal of these applications and services is to enable the user to interact more effectively with his environment. Examples of such value added service include services like displaying the map of the immediate surroundings and guiding a user inside a building. Several LM schemes for 802.11 systems have been proposed.

[Agiwal04] introduces LOCATOR, a radio frequency based system for location estimation of users in indoor wireless networks. The system works by building a radio map of the network site, which involves taking signal strength samples at various points in the wireless network, and then using this radio map, the system estimates the user's current location from the value of his current

observed signal strength. Although the use of an RF-based technique for location estimation is not new, LOCATOR is unique in the way it builds and manages the radio map to process location queries. Moreover, in addition to using ideas of probability to model the problem, [Agiwal04] has devised a multi-level clustering based algorithm that work in conjunction with the Lagrange interpolation scheme for a more efficient and accurate location estimation. The system was tested on 802.11b wireless network testbeds and has been able to achieve an accuracy of location estimation to within 4 feet of the actual user location with 90% probability. The core idea of LOCATER could be used in other wireless technologies as well.

[Da Silva04] presents a user position management system for WiFi networks. The location system is based on access point coverage range and beacon devices. The implementation was made over an authentication system extended to manage SNMP (Simple Network Management Protocol) in order to handle network events, and SLP (Service Location Protocol) to provide and collect service usage information per user. The work described was initially developed for restricted environments, such as a museum or a smart room.

The ability to determine a user's location through an existing 802.11 wireless network has vast implications in the area of context-aware and pervasive computing. Such abilities have been developed mainly in the Linux environment to date. To maximize its usefulness, a location determination system was developed in [Calvert05] for the more dominant Windows OS. While being able to operate outdoors as well as indoors, this system succeeds where traditional GPS fail, namely indoor environments. The system could benefit the large number of existing wireless networks and requires no additional hardware; only a few simple software downloads. Results show that just over half (52%) of all trials returned the exact location.

[Li] investigates the problem of user location over an 802.11 Ad-hoc network. The contribution analyzes the various factors that could affect signal strengths in an 802.11 network in Ad-hoc mode. The investigation was done by building a radio propagation model that could express the variance of obstruction/orientation. A new location algorithm was designed that incorporates the guess of the degree of obstruction/orientation. The approach is only applied to an outdoor environment without buildings, forests… etc among the peers. This means the effects of reflection and diffraction could be ignored and the main obstructions are human bodies. Another assumption was that the peers work collaboratively and uniformly. For example, all the subjects need to hold their PDAs in front of their chests (not in their bags) and there are no issues of security or privacy taken into account. Results indicate that even though providing location information in Ad-hoc wireless network is feasible, it is very hard to obtain location information of higher accuracy for reasons mentioned in [Li].

Having said all that is there to be said on LM in WiFi, the next section points at the latest advancements on LM in Bluetooth.

9. LM in Bluetooth

LM in 802.15 devices such as Bluetooth, Zigbee, and Infrared hasn't been much of a research issue because these devices operate in areas of less than 1 m radius. In addition, these technologies were initially designed without any location sensing in mind. Another issue to mention here is that there is a signaling interference between 802.11 devices and 802.15 devices since both operate in the 2.4 GHz unlicensed frequency band.

[Patel05] proposes some enhancements that Bluetooth would have to undergo so that it is best suited for indoor location sensing. Some of these enhancements include:

1. Reducing the time taken by a Bluetooth device to detect other Bluetooth devices in its vicinity.
2. The Bluetooth core specifications do not require that signal strength values be available to higher-level software. If this information is made available, it can greatly aid in accurate location sensing. With signal strength information, some applications such as the *Nearest Neighbor Concept* mentioned in [Patel05] can be applied.
3. The Bluetooth spec specifies 3 power levels at which Bluetooth devices can operate. The minimum power level (1 mW in class 1) gives a coverage range of roughly 10 m. If future implementations can give a shorter range, then location sensing can be made more accurate. Since Bluetooth devices are expected to be cheaper in the future, we can use more sensors covering very short area thus improving accuracy.

[Gonzalez-Castano03] analyzes survivability issues of auxiliary Bluetooth Location Networks (BLN) for location-aware or context-driven mobile networks. Among assumptions made is that there exist *service servers* that need to know user location in real-time to send context-oriented information to user handhelds when necessary. The BLN transmits position information to service servers, without user participation. It is not subject to line-of-sight constraints and is supported by existing commercial handhelds. BLN users carry either a Bluetooth-enabled handheld or any mobile data terminal and a Bluetooth *badge*. The BLN is composed by wireless Bluetooth nodes that establish a spontaneous network topology at system initialization. The BLN can coexist with other Bluetooth systems that aren't part of the location system such as printers and headphones.

As can be seen, not much research has been carried out on Bluetooth and 802.15 technology in general. The next section will talk about the relationship between LM and IP. This topic is considered a hot research topic since most of the research is focusing on the deployment of IPv6 in next generation systems.

10. LM and IP

The integration of IP, Mobile IP, IPv6, and Mobile IPv6 into next generation LM systems has been an active area of research for the past 2 decades. The idea of locating users using Mobile IP has been more of a dream a few years ago. However, researchers are getting there, and the dream is becoming more of a reality.

Mobile IP is a common standard to support global mobility of mobile hosts. One of the major problems for the Mobile IP is frequent location update and high signaling overhead. To solve this problem, a regional registration scheme was proposed to employ the hierarchy of the foreign agents (FAs) and the gateway foreign agents (GFAs) to localize registration operation. The system performance is largely dominated by the ability of a GFA and its reliability. [Kim05] proposes a new LM scheme for Mobile IP that reduces signaling burdens configuring regional network dynamically according to the characteristics of users and network. [Kim05] also proposes a cost model to calculate the total signaling cost for cost comparison. Simulation demonstrates that the proposed scheme performs better than other schemes when compared the total signaling cost and the regional network size. Also, the scheme enhances the system robustness and the load distribution. In addition, the proposed scheme has advantages in a dynamic network environment.

[Sharma04] describes an IPv6 based protocol for localized mobility management in next generation wireless networks. The protocol exploits the hierarchical IPv6 addressing for LM and uses specialized mechanisms for handoff management. It eliminates packet redirection along correspondent node (CN) - mobile node (MN) path, rendering greater scalability and simplifying routing. Unlike other micro-mobility proposals, this scheme takes an end-to-end approach to micro-

mobility by ensuring that MN packets arrive at topologically correct address from communicating nodes. [Sharma04] discuss the low-latency handoff mechanism that is critical to support real-time applications in wide area networks as well as WLAN networks. The results demonstrate that auto-update performs better than base Mobile IPv6 with lesser packet loss, shorter handoff duration, and improved throughput for TCP connections.

[Kumagai04] proposes a *Hierarchical Mobile IPv6* (HMIPv6) scheme to improve the performance capability of Mobile IPv6 at handover. In HMIPv6, local entities named *Mobility Anchor Points* (MAPs) are distributed throughout a network to localize the management of intra-domain mobility. In particular, multilayered MAP has been proposed to improve performance. MAPs reduce the number of *Binding Updates* (BA) to the *Home Agent* (HA) and improve the communication quality at handover. However, these conventional methods that manage a multi-layered MAP cannot select an appropriate MAP because they use the virtual mobility speed. As a result, they increase the signaling traffic in a multi-layered MAP. Moreover, they may cause the load to concentrate at a specific MAP. [Kumagai04] also propose a LM method for HMIPv6 using the mobile node's mobile history. In this method, when a mobile node performs a handover, the *Access Router* calculates the area-covered rate of each upper MAP from the mobile node's mobile history and selects the MAP that best manages the mobile node in accordance with its rate. Thus, the proposed method reduces both the number of Binding Updates (BU) to the Home Agent and the signaling traffic.

Summary

LM is a key factor for wireless mobile networks. Without a good strategy for LM, mobile communication and computing cannot exist.

Location Based Services are services that integrate a mobile device's location or position with other information so as to provide added value to a user. LBS are classified as Reactive and Proactive. Applications of LBS are classified as military and government industries, emergency services, and the commercial sector.

LM functions such as location updating and paging have to fulfill services and the requirements of users and operators. One of these requirements is cost efficiency, which could be reached by minimizing the signaling traffic both on radio links and on fixed network links. What we aim for is a LM scheme that will provide efficient searches and updates transparent to the user.

Current research in LM for wireless data networks includes Wireless Networks, Ad-hoc networks, and WiFi. Not much research is done for LM in Bluetooth. Integration of IPv6 and Mobile IP into LM is another interesting research topic that has a promising future.

References

[Mukherjee03] Amitava Mukherjee, Somprakash Bandyopadhyay, Debashis Saha, "**Location Management and Routing in Mobile Wireless Networks**", Artech House, 2003, pp. 69-113

[Kupper05] Axel Kupper, "**Location-based Services: Fundamentals and Operation**", John Wiley and Sons Ltd, 2005

[Schiller04] Jochen Schiller, Agnes Voisard, "**Location-Based Services**", Morgan Kaufmann, 2004

[Luo05] Xinwei Luo, Tracy Camp, William Navidi, "**Predictive Methods for Location Services in Mobile Ad-hoc Networks**", IEEE, 2005, 6 pages

[Karimi03] Hassan A. Karimi, Xiong Liu, "**A Predictive Location Model for Location-Based Services**", ACM, 2003, pp. 126-133

[Sethom05-1] Kaouthar Sethom, Khaled Masmoudi, Hossam Afifi, "**A Secure P2P Architecture for Location Management**", ACM, May 2005, pp. 22-26

[Lam04] Gary Hoi Kit Lam, Hong Va Leong, Stephen Chi Fai Chan, "**Reducing Group Management in Group-based Location Management**", Proceedings of the 15th International Workshop on Database and Expert Systems Applications (DEXA '04), 2004, 5 pages

[Sethom05-2] Kaouthar Sethom, Hossam Afifi, Guy Pujolle, "**A Distributed Architecture for Location Management in Next Generation Networks**", IEEE, 2005, pp. 1181-1186

[Lee04] Kevin Lee, Hong Wing Lee, Sanjay Jha, Nirupama Bulusu, "**Adaptive, Distributed Location Management in Mobile, Wireless Networks**", IEEE Communications Society, 2004, pp. 4077-4081

[Assouma05] Abdoul D. Assouma, Ronald Beaubrun, Samuel Pierre, "**A Location Management Scheme for Heterogeneous Wireless Networks**", IEEE, 2005, pp. 51-56

[Hwang05] San-Yih Hwang, Jeng-Kuen Chiu, "**Toward Optimal Replication for Hierarchical Location Management in Wireless Systems**", 2005, 12 pages

[Ghosh04] Joy Ghosh, Sumesh Philip, Chunming Qiao, "**Acquaintance Based Soft Location Management (ABSLM) in MANET**", IEEE Communications Society, 2004, pp. 166-171

[Philip04-1] Sumesh Philip, Joy Ghosh, Swapnil Khedekar, Chunming Qiao, "**Scalability Analysis of Location Management Protocols for Mobile Ad-hoc Networks**", IEEE Communications Society, 2004, pp. 183-188

[Philip04-2] Sumesh Philip, Chunming Qiao, "**MobiHoc Poster: Hierarchical Grid Location Management for Large Wireless Ad hoc Networks**", IEEE Communications Society, 2004, pp. 33-34

[Ziviani05] Artur Ziviani, Serge Fdida, Jose F. de Rezende, Otto Carlos Duarte, "**A TCP-Tailored Approach to Location Management in Mobile Ad-hoc Networks**", IEEE, 2005, pp. 1043-1045

[Agiwal04] Ankur Agiwal, Parakram Khandpur, Huzur Saran, "**LOCATER – Location Estimation for Wireless LANs**", ACM, 2004, pp. 102-109

[Da Silva04] Andrei Oliveira da Silva, Paulo Henrique de Souza Schneider, Fabricio D'Avila Cabral, Ana Cristina da Silva, Joao Batista de Oliveira, Eduardo Augusto Bezerra, "**Towards Service and User Discovery on Wireless Networks**", ACM, 2004, pp. 79-82

[Calvert05] Travis Calvert, Steven Case, "**Wireless Location Determination: Using Existing 802.11 Wireless networks to Determine a User's Location**", 2005, 8 pages

[Li] Song Li, Gang Zhao, Lin Liao, "**User Location Service over an 802.11 Ad-hoc Network**", 13 pages

[Patel05] Abhishek Patel, Dan Kim, Lionel Ni, "**A Study of Frequency Interference and Indoor Location Sensing with 802.11b and Bluetooth Technologies**", IEEE, 2005, 10 pages

[Gonzalez-Castano03] F. J. Gonzalez-Castano, J. Garcia-Reinoso, "**Survivable Bluetooth Location Networks**", IEEE, 2003, pp. 1014-1018

[Kim05] Ha Won Kim, Sung Je Hong, Jong Kim, "**Adaptive Location Management Scheme for Mobile IP**", IEEE, 2005, 5 pages

[Sharma04] Aurbin Sharma, A. L. Ananda, "**A Protocol for Micromobility Management in Next Generation IPv6 Networks**", ACM, October 2004, pp. 72-78

[Kumagai04] Takashi Kumagai, Takuya Asaka, Tatsuro Takahashi, "**Location Management Using Mobile History for Hierarchical Mobile IPv6 Networks**", IEEE, 2004, pp. 1585-1589

List of Acronyms

ABSLM: Acquaintance Based Soft Location Management
BLN: Bluetooth Location Networks
BLR: Boundary Location Register
CN: Correspondent Node
DLS: Distributed Lookup Servers
DNS: Domain Name Server
FA: Foreign Agent
FCC: Federal Communications Commission
GBL: Group-Based Location Updating
GFA: Gateway Foreign Agent
GPS: Global Positioning System
GLS: Grid Location Service
HGRID: Hierarchical Grid Location Management
HLR: Home Location Register
HLS: Home Location Server
HMIPv6: Hierarchical Mobile IPv6
LA: Location Area
LBS: Location Based Services
LEAP: Legend Exchange and Augmentation Protocol
LM: Location Management
LR-ING: Location Register and Internetworking Gateway
LU: Location Update
IP: Internet Protocol
IPv6: Internet Protocol version 6
MAP: Mobility Anchor Points
MANET: Mobile Ad-hoc Networks
MH: Mobile Host
MT: Mobile Terminal
PA: Paging Area
PLM: Predictive Location Model
PLS: Prediction Location Service
RC: Reporting Cell
PR: Paging Request
OSI: Open System Interconnection
SLALoM: Scalable Location Management
SLP: Service Location Protocol
SLS: Simple Location Service
SLURP: Scalable Update Based Routing Protocol
SNMP: Simple Network Management Protocol
TCP: Transmission Control Protocol
WiFi: Wireless Fidelity (802.11)
WLAN: Wireless Local Area Network