A Survey on the Location Management Problem in Mobile Networks

Ting-Yu Lin and Sheng-Po Kuo

***Abstract*—The question "Where is X?" is one of the commonly asked questions in our daily life. In mobile networks, this is connected with a fundamental location management problem, whose goal is to maintain a centralized/distributed database which maps X to its current location. X could be a person, a mobile host, an IP address, or a telephone number. X's location is a conceptual term which may be reflected by a physical or relative location, a logical area, a subnet, or a cell ID. In this article, we review this essential issue in three major types of network architectures: cellular networks, IP networks, and mobile ad-hoc networks. Although each particular network has its own concerns, the problem can be analyzed from two aspects: update and page. Finally, We summarize the article by introducing recent location management standardization reports and suggesting a cross-domain framework for different networks to share location information.**

***Index Terms*—cellular networks, location management, mobile ad hoc networks, mobile computing, mobile IP, peer-to-peer networks, wireless networks, location management.**

I. Introduction

MOBILE computing and wireless communications are perhaps the fastest growing areas in recent years. Not only have we seen a variety of emerging wireless networking technologies (such as GSM, GPRS, WCDMA, cdma2000, IEEE 802.11 WLAN, and Bluetooth), but also are there numerous portable computing devices widely available (such as laptops, tablet PCs, PDAs, and handsets). The marriage of these two fields has made ubiquitous computing and communications possible.

The commonly asked question, "Where is X?", in our daily life has a strong connection to mobile networks. In human life, X can be a movable person or object, and the answer to the question can be a physical, a logical, or a relative location (e.g., "45.7th mile of highway I-95", "at the north of Miami", or "under the table"). In mobile networks, "Where is X?" is connected to a fundamental *location management* problem, whose goal is to maintain a centralized/distributed database which maps X to its current location. In different networks, X could be a person, a mobile host, an IP address, or a telephone number. X's location is a conceptual term to reflect X's residency in the network, such as a physical location, a relative location, a logical area, a subnet, or a cell ID. Table I shows the location management issue in different wireless and mobile networks. In GSM networks, the problem is to map a telephone number (Mobile Station ISDN) to a *location area (LA)*, which is a set of base stations. Also being a cellular network, the third generation wireless telecommunication system cdma2000 incorporates other positioning strategies to further enhance location tracking precision beyond LAs. Those positioning techniques include A-GPS (Assisted-GPS), TDOA (Time Difference of Arrival), and E-OTD (Enhanced Observed Time Difference of Arrival), etc. Many of recently released cdma2000 1X mobile handset models are claimed to support A-GPS features. In the Internet society, the Mobile IP is accepted as a standard to support IP mobility. The "home" of a mobile host must maintain the current subnet that the host is currently visiting so as to correctly deliver packets to it. Many services in the emerging ad hoc/sensor networks also count on locations of hosts or objects. In such networks, positions are typically reflected by 2D/3D physical coordinates. A number of indoor positioning systems will be introduced in Section III-A.

In this article, we review important works of the location management problem in three types of mobile networks: cellular networks, IP networks, and mobile ad hoc networks. Although each particular network has its own concerns, the problem can be analyzed from two aspects: *update* and *page/query*. The updating process notifies the location servers of the current locations of mobile stations. In search of a mobile station, the paging process queries the servers to identify the exact/possible locations of a mobile station before the actual search. This avoids the potentially high costs of doing a global search. Updating and paging costs are tradeoffs. More frequent updates can preserve the freshness of the information in location servers, thus reducing the paging costs. On the contrary, less frequent updates can save updating costs, but may incur higher paging costs, especially for highly mobile stations.

This article has two purposes: (a) to serve as a review and (b) to act as a bridge among different mobile networks for the location management problem. Section II reviews related literature articles in infrastructure-based networks, including cellular and IP networks. Section III reviews the location management protocols in mobile ad hoc/sensor networks. In Section IV, we introduce the OSA (Open Service Access) concept promoted by 3GPP (3rd Generation Partnership Project) [4]. The OMA (Open Mobile Alliance) [6] and former LIF (Location Inter-operability Forum) will be introduced as well. Then, a unified framework for reusing location resources across different network domains is suggested. Section V draws our conclusions.

Manuscript received March 6, 2011

T.-Y. Lin is with the Department of Communication Engineering, National Chiao Tung University, Hsin-Chu 30010, Taiwan (e-mail: tylin@csie.nctu.edu.tw).

S.-P. Kuo is with Telcordia Applied Research Center, Suite E-453, 4F, No. 19-13, SanChung Rd., Taipei 115, Taiwan (e-mail: kuopo@research.telcordia.com).

TABLE I
LOCATION MANAGEMENT ISSUES IN DIFFERENT MOBILE NETWORKS.

	identity	location	tracking strategies
GSM/ cdma2000	MSISDN	location area (LA)/ enhanced Cell ID/ <latitude, longitude>	time/movement/distance-based, BS-initiated, centralized servers (HLR/VLR), A-GPS/TDOA/E-OTD centralized servers
IP networks	IP address	subnet	time-and-movement-based, terminal-initiated, centralized servers (home agent)
ad hoc/sensor networks	IP/MAC address	2D/3D coordinates	time/distance-based, host-initiated, distributed servers (virtual home zone, grid, etc.)

II. LOCATION MANAGEMENT IN INFRASTRUCTURE-BASED NETWORKS

A. Location Management in Cellular Networks

For cellular networks, we raise the instance GSM telecommunication system [23]. GSM stands for *Global System for Mobile Communications*, whose architecture is shown in Fig. 1(a). Defined by the radio coverage areas of *Base Transceiver Stations* (BTSs), the operating region is partitioned into *cells*. Multiple BTSs can be controlled by a *Base Station Controller* (BSC). The location management unit in GSM is called a *Location Area* (LA). Each LA consists of at least one BSC, but cells of a BSC may belong to different LAs. A *Mobile Switching Center* (MSC) is responsible for directing calls to subscribers in one or more than one LA. Calls originating from or terminating in the fixed network (PSTN/ISDN) are handled by the *Gateway Mobile Switching Center* (GMSC). In GSM, the configuration of LAs, BSCs, and MSCs are left open to system providers. Hence, this can be directed to an optimization problem.

In GSM, two essential databases are exercised to achieve location management and call control: *Home Location Register* (HLR) and *Visitor Location Register* (VLR). The HLR stores service profiles of all registered subscribers, with MSISDN as the searching key. Each MSC is accompanied by a VLR database, which keeps record of the LA where each mobile subscriber is currently located within its service area. In search of a mobile subscriber, say MS1 in Fig. 1(a), the HLR is first interrogated, which will indicate that MSC1 is being visited. Then MSC1 checks VLR1 and finds that MS1 is now within the service area of LA1. Finally, mobile stations in LA1 are all paged and MS1 is reached.

There are two basic operations to determine a user's exact location: *location update* and *paging*. To keep the database in HLR up-to-date, an important aspect of location management in GSM is the location update strategies. Frequent update activities ensure the freshness of location information in databases, so as to reduce the paging misses and thus the paging cost. Paging and updating are tradeoffs, thus leading to an optimization problem. In the literature, solutions to the problem can be classified into *time-*, *location-*, *movement-*, and *distance-based*.

For the third generation wireless telecommunication systems, cdma2000 standard is expected to deliver more location-based services by incorporating several positioning technologies (A-GPS/TDOA/E-OTD for instance) to increase location tracking precision. In the US, the Federal Communications Commission (FCC) has started the E911 (Enhanced 911) effort since 1996. The E911 services mandate telecommunication operators to provide caller location information in response to emergency requests. Many of the major cdma2000 operators chose A-GPS solution to meet the E911 requirements. A-GPS can provide more precise location information and detect weaker signals than those that conventional GPS receivers require. However, in some urban settings (such as dense multi-story large buildings, underground subway stations, and steel/concrete indoors, etc.), A-GPS dose not work well. According to previous reports, A-GPS is superior to pure GPS for outdoor environments and inferior to GPS when used indoors. Other alternatives beyond A-GPS include TDOA and E-OTD, which had been adopted by some operators. These techniques may provide location fixes in some harsh indoor environments, though they lack 3D positioning. With similar purposes, in July 2000, the European Commission (EC) initiated the LOCUS (Location of Cellular Users for Emergency Services) project to assist the European Union (EU) on implementing E112 (Enhanced 112) emergency call services. According to the LOCUS assessment report, there is no single positioning technology sufficient in meeting all service requirements. A combination of various technologies, for example A-GPS+E-OTD, may yield better performance at the cost of higher implementation complexities.

B. Location Management in IP Networks

Originally designed for stationary hosts, the Internet assumes that hosts always have fixed points of attachments. Hosts' identities, i.e., IP addresses, are in fact "location-dependent," in the sense that an IP address cannot move from subnet to subnet. The Mobile IP is designed to support host mobility without changing hosts' IP addresses. With Mobile IP, a mobile host can be associated with a permanent IP, despite its visiting areas, and thus IP addresses become "location-independent."

The design philosophy behind Mobile IP is similar to that of GSM networks. There is a *Home Agent* (HA), which maintains a *Location Directory* (LD) for all mobile hosts within its subnet. When a mobile host, say MH in Fig. 1(b), travels to the area of a *Foreign Agent* (FA), it should first obtain a *care-of address* (CoA). A CoA could be the FA's address or one freshly obtained from a local DHCP server. Then it will register with the HA its current CoA. In the future, packets from any Corresponding Host (CH) for MH will be routed to HA first via standard IP routing. Then HA will interrogate its

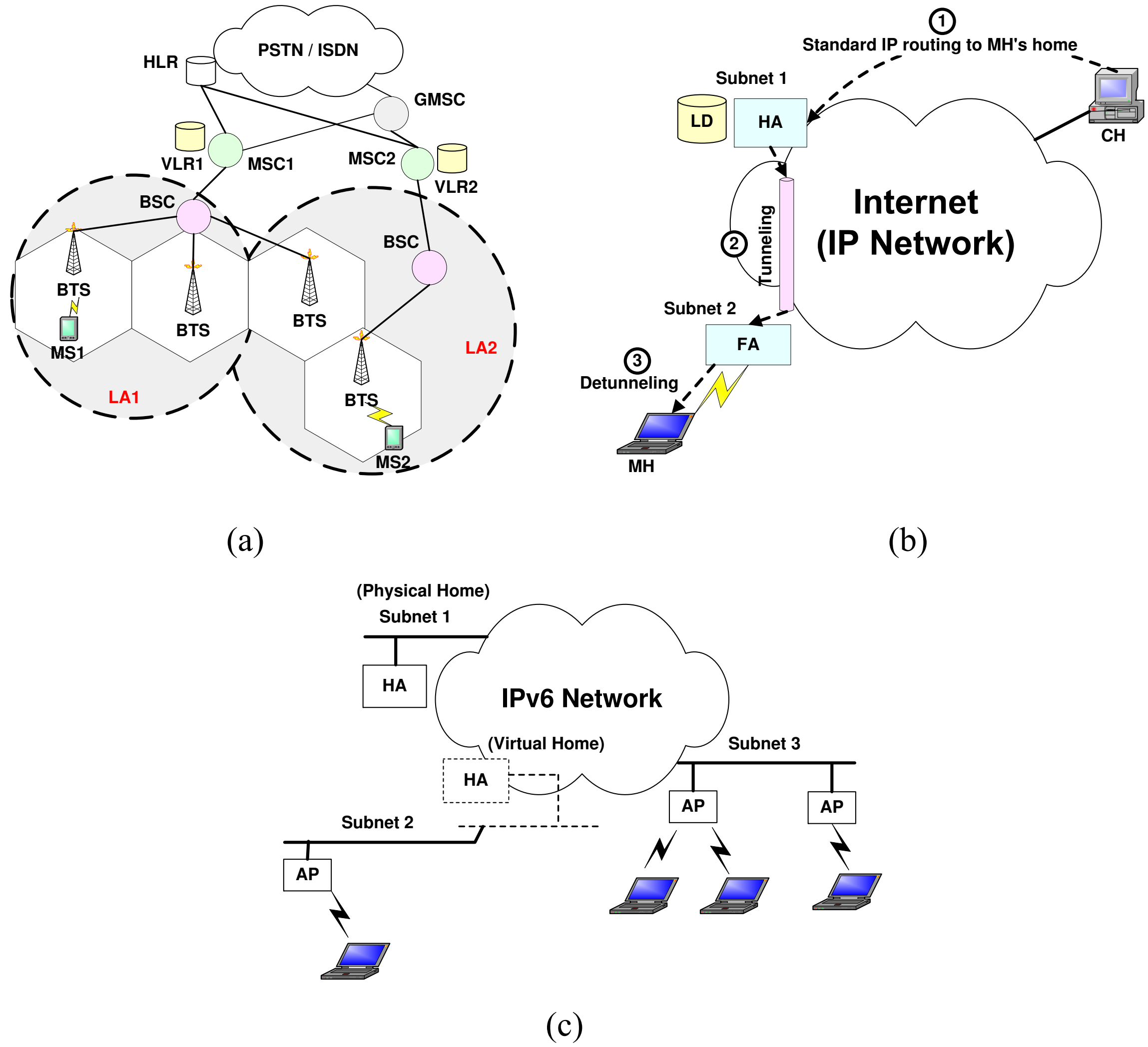


Fig. 1. Location management in (a) GSM system, (b) Mobile IP architecture, and (c) Mobile IPv6 architecture.

LD to obtain MH's CoA. To provide protocol transparency, Mobile IP adopts a *tunneling* mechanism by encapsulating each IP packet with another IP header carrying MH's CoA. The packet can be decapsulated either at FA or at MH. Such an IP-in-IP technique requires no modification to the original IP layer.

The mobility management overhead in Mobile IP consists of several costs. First, the mobility agents (HA and FA) need to periodically advertise their existence, which can be considered as the paging cost. A host may miss packets after it enters a new subnet but before it successfully registers with its HA. This depends on the frequency of advertisements and registration delays, and the packet loss cost can also be added to the paging cost. The update strategy of mobile hosts follows a mixture of time-based and movement-based approaches. On hearing an advertisement different from previous ones (which happens as entering a new subnet or returning home), a mobile host should register. When a mobile host does not change its point of attachment but its previous registration is close to expiration, it should refresh its registration too. Several variants of mobility management strategies for Mobile IP also exist.

1) Lookup Services in Peer-to-Peer Networks: While Mobile IP requires an explicit destination identification to set up a data flow, content-based location lookup (without dictated identification) is implemented in peer-to-peer networks, in search of interested data/information. Peer-to-peer (P2P) networks enable two or more peers to collaborate spontaneously in a network of equals (peers) by using appropriate information and communication systems without the necessity for central coordination. The P2P network is dynamic where peers come and go (i.e., leave and join the group). Peer-to-peer network models such as Gnutella [2], Freenet [1], and Napster [3] have become popular for sharing information and data through direct exchange.

The location management in P2P networks is also called the lookup problem. Specifically, how can we find any given data item in a large P2P network in a scalable manner, without any centralized servers or hierarchy? In serverless approaches, flooding-based search mechanisms are used such as DFS with depth limit D (in Freenet) or BFS with depth limit D (in Gnutella), where D is the system-wide maximum TTL of a message in hops. In limited server-based approaches, location information is kept to a limited region of nodes such as within a predefined hops as in local indices [37]. In NEVRLATE [11], nodes are organized in a logical 2-D grid with a set of servers enabling registration (publish) in one "horizontal" dimension and lookup in the other "vertical" dimension.

In the server-based approaches, each node acts as a server for a subset of data items. The operation lookup (key) is supported, which returns the identity (e.g., the IP address) of the node storing the data item with that key. The values of the node could be actual data items, or could be pointers to where the date items are currently stored. Each data item is associated with a key through a hashing function. Nodes have identifiers, taken from the same space as the keys. Each node maintains a routing table consisting of a small subset of nodes in the system. In this way, a overlay network is constructed that captures logical connections between nodes. Usually, the logical network is regular such as a ring, tree, or mesh. When a node receives a query for a key for which it is not responsible, the node routes the query to the neighbor that makes the most "progress" (normally defined in terms of "distance" between source and destination identifies) towards resolving the query. The above approach supports a distributed hash table (DHT) functionality to provide a general-purpose interface for location-independent naming services. Several representative projects on scalable P2P system through DHT include CAN [27], Chord [31], Pastry [29], and Tapestry [38].

III. Location Services in Infrastructure-less Ad Hoc/Sensor Networks

The emerging *Mobile Ad Hoc Network* (MANET) architecture has attracted much attention. Sensor networks also adopt the similar architecture. In a MANET, hosts act equally as routers and cooperate to relay packets. When used outdoors, MANET is particularly attractive in providing location-based services. It is commonly assumed that mobile hosts are attached to GPS receivers to determine their own positions. Due to the absence of fixed infrastructure in a MANET, it is difficult to apply a centralized mechanism that always tracks the current whereabouts of its member mobile hosts. The infrastructure-less nature of MANETs poses a unique technical challenge for location management compared to other networks. Furthermore, for indoor environments where satellite signal cannot reach the GPS receiver, or due to GPS cost consideration, mobile hosts may need certain non-GPS positioning techniques to identify its own location. Below we briefly review representative indoor positioning systems in Section III-A. Once the location information is readily available at each mobile host, Section III-B provides state-of-the-art location management protocols in MANETs (particularly distributed algorithms will be reviewed).

A. Indoor Positioning Techniques

Positioning in an indoor environment is a challenge research issue. According to the positioning processes with or without distance measurement, we can briefly categorize them into two classes. The one which needs distance measurement is called *range-based* positioning systems; the other one is *range-free*. In the following, we will discuss two famous positioning systems for each category.

The *ad hoc positioning system (APS)* designs a distributed algorithm, called *DV-Hop* [24], based on the concept of distance vector used in many routing algorithms. A small fraction of nodes which are assumed to be aware of their locations are called landmarks. They will periodically flood out their coordinates. Similar to distance vector routing scheme, each node maintains a table to collect updates from neighbors and forwards these updates by the manner of hop-by-hop dissemination. The update packets includes the coordinates and the corresponding hop counts to the landmarks. When a landmark receives another landmark's updates, it can compute a correction which denotes the average distance for one hop. The correction packets are also flooded out. Then each node can estimate its own location when it receives any correction and more than 3 updates from different landmarks by the trilateration algorithm. Another range-free positioning system is *Active Badge*. It is a cell-based positioning system [35]. Several infrared receivers are equipped at some specific locations to receive signals. Users wear the badges which can periodically emit an unique identification. The system can identify each user's current location according to the infrared cell where the user currently stay.

The *Active Bat* system locates users by a trilateration algorithm after distance measurement [7]. Similar to the Active Badge, a number of receivers are mounted on the ceiling. A central coordinator will control the receivers to periodically emit RF signals first. Then, users carrying wireless transmitters (Bats) send ultrasonic signals back and nearby receivers measure the distances by signal traveling time. According to these distance measurements, the trilateration algorithm is performed to compute the exact users' locations. In experiments, the Active Bat positioning system can provide accuracy within $3cm$ in a three dimensional space. However, its deployment cost is higher than other positioning systems.

The RADAR is another range-based positioning system [26] but it does not measure distances directly. Alternatively, it collects the signal patterns at a set of training locations. These signal patterns will be stored at a central database. To locate users' current location, it compares their received signal patterns with the ones in the database. The final location estimation is the most similar one according to a pattern matching algorithm. This system is sensitive to environmental noise; thus, its accuracy is around 3 meters. However, this system only relies on signal strengths to locate users so it can be integrated with an existed communication infrastructure without extra hardware cost. Some enhanced algorithms can further improve the accuracy via tracking techniques [18] and performance via optimization skills [19]. Furthermore, although RADAR is originally implemented based on WiFi signals and infrastructures, some extended researches reveal that similar pattern-matching techniques can be applied in GSM/LTE networks [25].

Table II summarizes these positioning systems according to their categories, used techniques, positioning accuracy, and extra hardware cost.

TABLE II
THE SUMMARY OF FOUR INDOOR POSITIONING SYSTEMS.

	DV-Hop	Active Badge	Active Bat	RADAR
Category	Range-free	Range-free	Range-based	Range-based
Technique	Trilateration	Cell ID	Trilateration	Pattern Matching
Accuracy	Low (Depend on Topology	Medium	High ($< 3cm$)	Medium ($\approx 3m$)
Extra Hardware Cost	No	Medium	High	Low

B. Location Management

Location management in MANETs can be classified as *serverless* or *server-based*. In serverless approaches, each mobile host tries to maintain location information of others either proactively or reactively; whereas in server-based approaches, the locations of hosts are managed by a subset of hosts. In the literature, several distributed strategies have been devised [9], [12]–[14], [20]. Below we review those proposals in more detail.

As a routing protocol, the DREAM scheme [9] tries to maintain a simple location table in each host to track the approximated locations of other hosts. Each location entry contains a host's ID, moving direction, and distance to that host, and a timestamp indicating the freshness of the information. Location information is broadcasted periodically by each host. To reduce traffic overheads, the *distance effect* is considered: the farther two hosts are separated, the less often they need to exchange with each other their location tables. To realize the distance effect, each broadcast packet is assigned a *lifetime*, which reflects the geographic distance the packet can travel. A majority of packets are short-lived, and will "die" after traveling a short distance. Long-lived packets can travel longer distances, and are sent less frequently. The frequency with which a host broadcasts is a function of its mobility rate. Since information in location tables is not precise, a host, when intending to page (and thus send packets to) another host, needs to flood the packet in a "cone" area that is likely to reach the host. However, this approach is not scalable to large networks.

A rather interesting concept called *virtual home region (VHR)* is proposed in [12]. Each host x periodically updates its current position to all hosts currently residing in its own VHR, which is a geographical region. Every host in the VHR should store x's current location. The VHR can be determined by a globally known hash function with x's identity (such as IP or MAC address) as the input. A host who would like to locate x can page any host in x's VHR. So hosts only need to keep locations of those hosts whose VHRs cover itself. A simple way to define VHR is a center point with a fixed radius. It is also suggested in [12] that the radius can be dynamically adjusted based on the host density near the VHR. In general, VHR can be of any shape. The idea is illustrated in Fig. 2. More information can be found in [32].

A similar work *geographical hash table (GHT)* is addressed in [28]. Its original purpose is provided for the applications of data-centric storage. All nodes share an unique hash function $h(\cdot)$ to map an identity to a location and all data of the same data type d will be stored at the *home node* which is closest to $h(d)$. If we regard the identity of each node as an unique

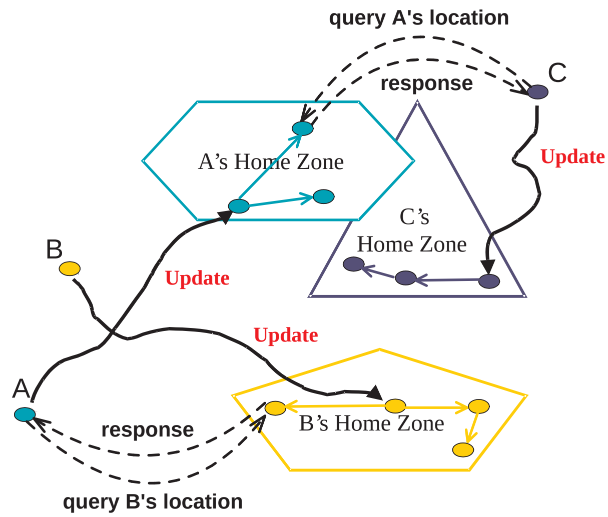
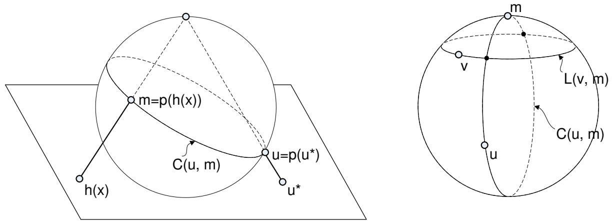


Fig. 2. Location management by the Virtual Home Region (VHR) scheme.

data type, we can apply GHT to the location management in MANETs. There are two major differences between VHR and GHT. First, GHT is built on the top of a geographic routing protocol GPSR [15] such that the update and query mechanisms are fully integrated with the underlying routing protocol. Using the perimeter-mode of GPSR, the update packets for host x can be easily routed to the home node after traversing the perimeter nodes enclosing $h(x)$. Second, considering the instability issue in MANETs, GHT proposes a perimeter refresh protocol to provide more reliable location management by periodically replicating location information at the perimeter nodes enclosing $h(x)$ besides the home node.

To extend GHT, a double ruling scheme proposed in [30] can further provide query locality. This feature makes the routing distance of query packets depend on the distance between the node sending query packets and the corresponding location server. Hence, the node can obtain the location information quickly if it is near to the location server. Instead of the geographical routing algorithm GPSR used in GHT, the routes of packets are delivered by the concept of the projection of circles as follows. Given a field with boundaries, we place a sphere on this field and map each location on the sphere to the field. The projective mapping can be illustrated by Fig. 3(a). If the south pole is tangent to the field, we place a luminary at the north pole of the sphere. For each point u on the sphere, we can project an unique point u^* on the field. In the inverse direction, we can also map a point u^* on the field to the point u on the sphere through the same projective path. Thus, there is a one-to-one projective mapping function $p(\cdot)$ such



(a) (b)

Fig. 3. (a) The projective mapping mechanism in the double rulings scheme and (b) an illustration of a replica curve $C(u, m)$ and a retrieval curve $L(v, m)$ in the distance-sensitive retrieval scheme.

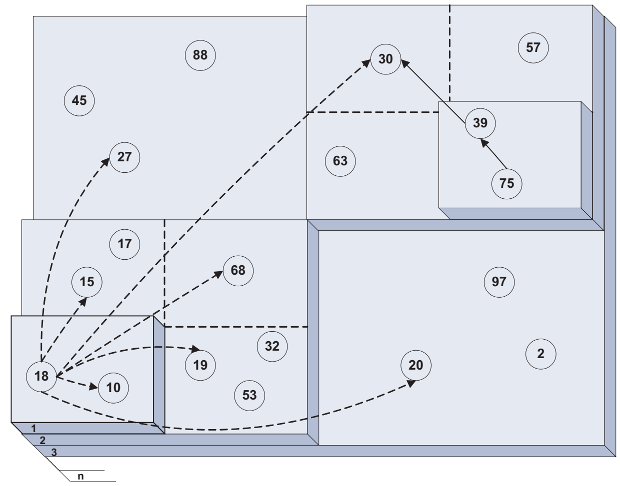


Fig. 4. Location management in the GLS (Grid's Location Service) scheme.

that given a point u^* on the field, we can have an unique point $u = p(u^*)$ on the sphere. This projective mapping preserves two features. First, any circle on the sphere is mapped to a circle on the field. Seconde, for any two points u^* and v^* on the field, their distance is smaller than the distance of their mapping points $p(u^*)$ and $p(v^*)$ with a constant coefficient.

Based on this projective mapping model, each node x updates its current location u^* to the nodes whose mapping points on the sphere are closed to a *replica curve*. This curve is a great circle $C(u, m)$ passing two mapping points $u = p(u^*)$ and $m = p(h(x))$, where the hash function h is similar to the one in GHT. On the other hand, another node y which wants to query x's current location can follow the distance-sensitive retrieval scheme proposed in [30]. First, the sphere is rotated such that m is at the north pole. Because replica curves are great circles, $C(u, m)$ is one of longitude curves. Then, we generate another *retrieval curve* $L(v, m)$ which is a latitude curve passing the mapping point $v = p(v^*)$ of y's location v^* on the sphere with m as its north pole. An example is shown in Fig. 3(b). Follow this curve $L(v, m)$, it is guaranteed that the distance traversed by the query packets is bounded by $O(d)$ where d denotes the distance between u and v (detailed proof can be found in [30]).

A hierarchical, grid-based strategy called *Grid's Location Service (GLS)*, originally motivated by [8], is proposed in [20]. A host only needs to maintain location information of some nodes. GLS partitions the network region into a hierarchy of squares called *grids*, as shown in Fig. 4. The smallest grids are *order-1* squares. An *order-n* square consists of four neighboring *order-$(n-1)$* squares. For example, in Fig. 4, an *order-2* grid covers four *order-1* squares, while an *order-3* grid consists of four *order-2* squares. This leads to a property that for any $i \geq 1$, each host is resident in exactly one order-i square, which has three *sibling* squares of the same order. Each host x is assumed to have a unique ID and will distributedly and dynamically select a set of hosts as its location servers. Within an order-1 square, x maintains the location information of all other hosts in the same square. For all other order-i squares, $i \geq 1$, host x recruits one host within each of the three sibling order-i squares of the order-i square where x is currently resident as its location servers. The recruiting rule is such that the server's ID must be "closest" to x's ID within the considered order-i square, where "closest" is defined to be the least ID greater than x's by regarding numbers in a circular manner. For example, in Fig. 4, host 18 serves as its own location server within its own order-1 grid, and selects hosts 15, 68, and 19 in the three sibling order-1 squares as its location servers. Hosts 27, 30, and 20 are selected as 18's location servers in the three sibling order-2 squares of 18's own order-2 square. Note that by "circular", host 15 is considered "closer to" 18 than 17 in the square where 15 serves as the server. So hosts will have equal opportunity to serve as servers.

To query the location of another host y, a host simply searches its own location database for the entry that is closest to y, say y'. If $y' = y$, this is done. Otherwise, the host sends a query to y' (in which case the host knows the location of y'). Again, y' searches its own location database for the entry that is closest to y, say y''. Either $y'' = y$ or y'' is closer to y than y'. In the latter case, a query is sent to y'' recursively. By repeating this process, it is guaranteed that the location of y can be found. For example, Fig. 4 shows the path for host 75 to look for 18's location, which will end up at host 30, which is 18's location server. The correctness of this protocol relies on the fact that in any order-i sibling square, there must exist a server of y with the least ID. Since the server also needs to select its own location servers, the above searched hosts $y', y'', ...$ must have IDs larger than the server's and there must exist a chain connecting to the server. Location update can be done similarly to querying. A host x who intends to update its location in a square can query any host in that square. The queried host then searches for x following the same query procedure as above. This will end up with a query leaving this square. The last queried host is in fact x's server in this square.

An extension of GLS is discussed in [36]. Instead of selecting the host in a partition with the closest ID to host x as the location server of x, a hash function is used to map x directly to the location(s) of the server(s). In addition, different logical network partitions are used so that locations of hosts can be represented at different accuracy levels. Only a small set of location servers needs to be updated when a host moves.

Using the same grid configuration, reference [14] proposes a location dissemination scheme called *Geographical Region Summary Service (GRSS)*. Instead of keeping exact information, GRSS only tries to maintain rough locations of mobile hosts. From a summary, one can easily tell whether a host is within a region with high level of confidence. Within each order-1 square, a host needs to know complete information of all other hosts, produce an order-1 summary, and forward the summary to all its three sibling order-1 squares. On receiving order-1 summaries from all sibling squares, an order-1 square needs to produce an order-2 summary and sends the summary to all its three sibling order-2 squares. This is repeated recursively for higher-order squares. Summaries are distributed in a square by flooding.

To save space and reduce communication overhead, a summary is computed as a fixed-length bit vector. To look up a host in a summary, we first apply n hash functions on its ID and get n bit positions. If any of the n bit positions in the vector is not 1, the host is not in the summary. Otherwise, the host is in the summary with high probability. It is called a "false positive" if the host is not in the corresponding region when all n bits are 1's. As a result, each host keeps location information of *all* other hosts, but only knows roughly in which regions they currently reside. This is different from GLS, in which each host keeps exact location information, but only for part of the network. Routing is also discussed in [14] (which is beyond the scope of this paper and is omitted here).

Many applications have been proposed for MANETs based on location services. The LAR (location-aided routing) protocol is proposed in [17] to facilitate route discovery in a MANET by restricting the flooding area of route requests. Packet delivery by greedy forwarding, without going through the route discovery procedure, is proposed in [15], which was originally motivated by [10]. In [15], methods to route around local dead ends (i.e., local minimum) are addressed. A survey of position-based routing is in [22]. The concept of geocasting which delivers packets to a group of nodes in a specific area is discussed in [16], [21], [33]. Location awareness and energy concerns for MANETs are addressed together under an integrated architecture in [34].

IV. Framework for Cross-Domain Location Resources Access

Location-based services allow mobile users to receive services based on their geographical positions. Unfortunately, as reviewed, different mobile networks usually adopt different positioning models and techniques. In order to realize cross-domain mobile services, several leading groups, such as 3GPP, 3GPP2, OMA, GSMA, and CDG, have dedicated efforts to interoperable services. Below, we discuss the OSA (Open Service Access) framework promoted by 3GPP and suggest possible integration with IP-based networks. The emerging OMA (Open Mobile Alliance) group, which has consolidated over 300 companies, will also be introduced. Our discussion will focus on the related location data sharing mechanisms.

First released in year 2000, OSA [4] was defined by 3GPP to enable fast third-party application developments. To shorten the time-to-market of 3G applications, such as multi-media messaging, m-commerce, and location-oriented services, network operators typically invite third-party vendors to join the development. To speed up the process, providers need to develop programming interfaces for vendors to access core networks. As a result, 3GPP has initiated the Open Service Access (OSA) standard, which consists of a collection of open network APIs by which third-party programmers can make use of underlying network functionalities, including call control, terminal status, user location, and messaging delivery.

The OSA architecture is illustrated in Fig. 5. The OSA gateway is provided by the operator. Inside the gateway, there are several Service Capability Features (SCFs), each is implemented in a Service Capability Server (SCS). For application developers, the core network capabilities can be accessed transparently through the OSA interfaces. In this way, third-party developers can focus on applications themselves, and end-users can access various services by connecting to operators' portals as usual. Such cooperation between operators and software vendors is advantageous for both parties.

Among the various services defined in OSA, the Mobility SCF provides an interface for querying a user's status and location. The OSA User Location Service specifies the interface for tracking a mobile subscriber's location at specific/periodical time, or requesting location-triggered reports. The location responses delivered by the network can be geographical positions (with an indication of accuracy) or radio cell IDs.

With similar goals to OSA, the Open Mobile Alliance (OMA) [6] was initiated in June 2002 as an industry organization targeted at realizing user-centric mobile services across countries, operators, and mobile terminals. By integrating the whole mobile services value chain, all parties, including mobile operators, device and network vendors, content providers, and application developers, can work together to foster a single common framework with open standard interfaces. OMA aims to provide end-users with seamless mobile services. By April 2003, OMA has consolidated the Location Inter-operability Forum (LIF), SyncML, MMS Interoperability Group (MMS-IOP), and Wireless Village. Among the many working groups, the Location Working Group (LOC) aims at developing interoperable standards for Mobile Location Services.

Since LOC continues most of the work originated by the former LIF, it is worth elaborating its services/protocols. LIF sets its goal on offering global location-based services on different networks and terminals. The Mobile Location Protocol (MLP) [5] is proposed to provide a simple yet secure API to access the location server. All exchanging messages are defined by XML Document Type Definitions (DTD). These messages can be transferred through a variety of transport protocols, including HTTP, SOAP, and SSL/TLS. Five types of location services defined in MLP provide simple query and reporting protocols. Relied on the highly extensible property of XML, MLP is suitable to support cross-domain location-based services.

Motivated by the powerful architecture of OSA and extensible message format of MLP, we propose that IP-based networks follow similar steps to realize resource sharing as well. The idea, as illustrated in Fig. 5, lies in a standardized interface for outside interactions. For Mobile IP networks,

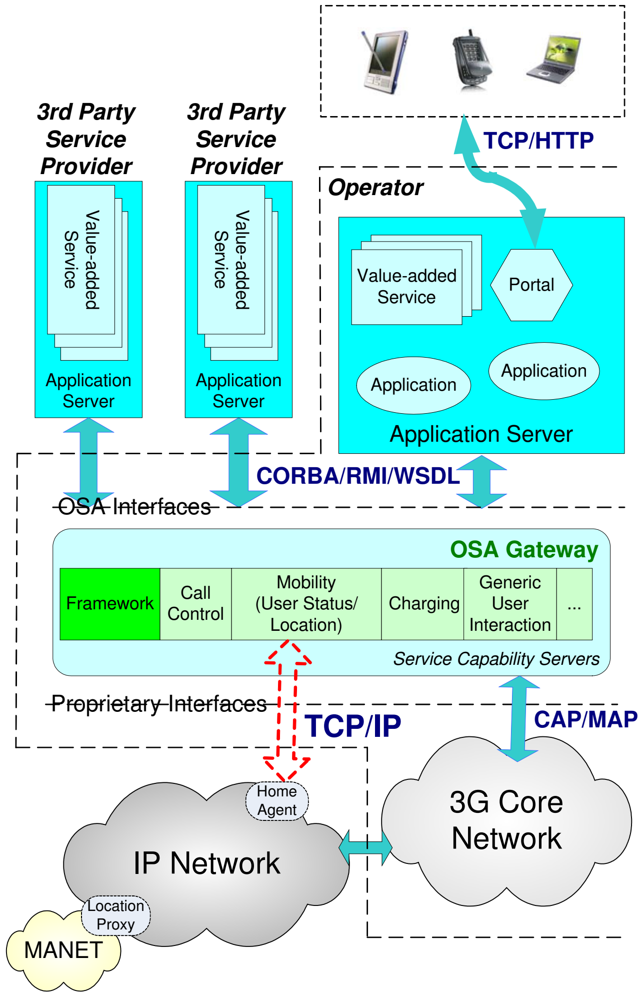


Fig. 5. The suggested cross-domain location service framework based on Open Service Access (OSA) architecture.

the HA (Home Agent) should act as the access gateway by providing a query interface for client applications. The location responses would be the registered care-of addresses. For ad hoc networks, the geographical positions (coordinates) may also be useful for outside applications. One possible way is to designate a representative ad hoc node as the interface gateway (location proxy). Interested applications may request for relevant location information if the ad hoc network is recognized by the Internet world. Through unified OSA interfaces and mobile user profiles, location-oriented applications can always retrieve the location information whatever network the mobile subscriber is attached to. The location information exchanging for different networks can be handled by MLP. We believe that such frameworks integrating the aids of OSA and OMA's MLP are attractive trends for efficient location-based application deployments across various network domains.

V. Conclusion

We have reviewed the design issues of location management in several different mobile networks. How to integrate location databases of different networks to provide more attractive location-based services poses a new challenge. We have introduced 3GPP's OSA and OMA's MLP technologies and suggested a unified framework to enable the cross-domain location-based services. We note that the emerging peer-to-peer (P2P) and overlay networks also have the location management problem (called the *lookup problem*), whose goal is to search for an object/file in a logical network with thousands or millions of active users. All these interesting issues are closely related to a simple question: "Where is X?".

REFERENCES

[1] Freenet website, http://freenet.sourceforge.net.

[2] Gnutella website, http://www.gnutella.com.

[3] Napster website, http://www.napster.com.

[4] Open Service Access, Application Programming Interface (Release 4). *3G TS 29.198, V4.2.0, 2001-09*, 2001.

[5] Mobile Location Protocol (MLP), LIF TS 101 Specification v3.0.0. *Location Inter-operability Forum (LIF)*, Jun. 2002.

[6] OMA Release Program and Specifications, http://www.openmobilealliance.org. *Open Mobile Alliance (OMA)*, 2002.

[7] M. Addlesee, R. Curwen, S. Hodges, J. Newman, P. Steggles, A. Ward, and A. Hopper. Implementing a sentient computing system. *Computer*, 34(8):50–56, 2001.

[8] K. Amouris, S. Papavassiliou, and M. Li. A Position-Based Multi-Zone Routing Protocol for Wide Area Mobile Ad-Hoc Networks. In *IEEE Int'l Conf. on Vehicular Technology Conference (VTC)*, pages 1365–1369, 1999.

[9] S. Basagni, I. Chlamtac, and V. R. Syrotiuk. A Distance Routing Effect Algorithm for Mobility (DREAM). In *ACM Int'l Conf. on Mobile Computing and Networking (MobiCom)*, pages 76–84, 1998.

[10] P. Bose, P. Morin, I. Stojmenovic, and J. Urrutia. Routing with Guaranteed Delivery in Ad Hoc Wireless Networks. In *Int'l Conf. on Discrete Algorithms and Methods for Mobile Computing and Communications*, pages 48–55, Aug. 1999.

[11] A. Chander, S. Dawson, P. Lincoln, and D. Stringer-Calvert. Nevrlate: Scalable Resource Discovery. *Proceedings of the 2nd IEEE/ACM Int'l Symp. on Cluster Computing and the Grid (CC-GRID)*, pages 352–358, 2002.

[12] S. Giordano, M. Hamdi, J.-P. Hubaux, J.-Y. Le Boudec, and L. Blazevic. Issues on Mobile Ad-Hoc WANS. In *IEEE Int'l Conf. on Multimedia and Expo*, pages 1261–64, 2000.

[13] Z. J. Haas and B. Liang. Ad Hoc Mobility Management with Uniform Quorum Systems. *IEEE/ACM Transactions on Networking*, 7(2):228–240, 1999.

[14] P.-H. Hsiao. Geographical Region Summary Service for Geographical Routing. *ACM Mobile Computing and Communications Review (MC2R)*, 5(4):25–39, 2001.

[15] B. Karp and H. T. Kung. GPSR: greedy perimeter stateless routing for wireless networks. In *ACM Int'l Conf. on Mobile Computing and Networking (MobiCom)*, pages 243–254, New York, NY, USA, 2000. ACM Press.

[16] Y.-B. Ko and N. H. Vaidya. Geocasting in mobile ad hoc networks: Location-based multicast algorithms. In *IEEE Workshop on Mobile Computing Systems and Applications (WMCSA)*, pages 101–110, 1999.

[17] Y.-B. Ko and N. H. Vaidya. Location-Aided Routing (LAR) in Mobile Ad Hoc Networks. *Wireless Networks*, 6(4):307–321, 2000.

[18] S.-P. Kuo and Y.-C. Tseng. A Scrambling Method for Fingerprint Positioning Based on Temporal Diversity and Spatial Dependency. *IEEE Trans. on Knowledge and Data Engineering*, 20(5):678–684, 2008.

[19] S.-P. Kuo and Y.-C. Tseng. Discriminant Minimization Search for Large-scale RF-based Localization Systems. *IEEE Trans. on Mobile Computing*, 10:291–304, 2011.

[20] J. Li, J. Jannotti, D. S. J. D. Couto, D. R. Karger, and R. Morris. A Scalable Location Service for Geographic Ad Hoc Routing. In *ACM Int'l Conf. on Mobile Computing and Networking (MobiCom)*, pages 120–130, 2000.

[21] W.-H. Liao, Y.-C. Tseng, K.-L. Lo, and J.-P. Sheu. GeoGRID: A Geocasting Protocol for Mobile Ad Hoc Networks Based on GRID. *Journal of Internet Technology*, 1(2):23–32, 2000.

[22] M. Mauve, J. Widmer, and H. Hartenstein. A Survey on Position-Based Routing in Mobile Ad-Hoc Networks. *IEEE Network*, 15(6):30–39, 2001.

[23] M. Mouly and M.-B. Pautet. *The GSM System for Mobile Communications*. Telecom Publishing, 1992.

[24] D. Niculescu and B. Nath. Ad hoc positioning system (APS). In *IEEE Global Telecommunications Conf. (GLOBECOM)*, volume 5, pages 2926–2931, 2001.

[25] V. Otsason, A. Varshavsky, A. LaMarca, and E. de Lara. Accurate GSM Indoor Localization.

[26] Paramvir Bahl and Venkata N. Padmanabhan. Radar: an in-building rf-based user location and tracking system. In *INFOCOM*, volume 2, pages 775–784, 2000.

[27] S. Ratnasamy, P. Francis, M. Handley, R. Karp, and S. Shenker. A Scalable Content-addressable Network. *Proceedings of ACM SIGCOMM*, pages 161–172, 2001.

[28] S. Ratnasamy, B. Karp, L. Yin, F. Yu, D. Estrin, R. Govindan, and S. Shenker. GHT: a geographic hash table for data-centric storage. In *ACM Int'l workshop on Wireless Sensor Networks and Applications (WSNA)*, pages 78–87, 2002.

[29] A. Rowstron and R. Druschel. Pastry: Scalable, Distributed Object Location and Routing for Large-scale Peer-to-peer Systems. *Proceedings of the 18th IFIP/ACM Int'l Conf. on Distributed Systems Platform (Middleware 2001)*, 11:329–350, 2001.

[30] R. Sarkar, X. Zhu, and J. Gao. Double rulings for information brokerage in sensor networks. In *ACM Int'l Conf. on Mobile Computing and Networking (MobiCom)*, pages 286–297, 2006.

[31] I. Stoia, R. Morris, D. Karger, M. F. Kaashoek, and H. Balakrishnan. Chord: A Scalable Peer-to-peer Lookup Service for Internet Application. *Proceedings of ACM SIGCOMM*, 31(4):149–160, 2001.

[32] I. Stojmenovic. Home Agent Based Location Update and Destination Search Schemes in Ad Hoc Wireless Networks (a chapter in Advances in Information Science and Soft Computing edited by A. Zemliak and N.E. Mastorakis). *WSEAS Press*, pages 6–11, 2002.

[33] I. Stojmenovic, A. P. Ruhil, and D. K. Lobiyal. Voronoi diagram and convex hull based geocasting and routing in wireless networks: Research articles. *Wireless Communication and Mobile Computing*, 6(2):247–258, 2006.

[34] Y.-C. Tseng and T.-Y. Hsieh. Fully power-aware and location-aware protocols for wireless multi-hop ad hoc networks. In *Int'l Conf. on Computer Communications and Networks (ICCCN)*, pages 608–613, 2002.

[35] R. Want, A. Hopper, V. Falcão, and J. Gibbons. The active badge location system. *ACM Transactions on Information Systems*, 10(1):91–102, 1992.

[36] Y. Xue, B. Li, and K. Nahrstedt. A scalable location management scheme in mobile ad-hoc networks. In *IEEE Annual Conf. on Local Computer Networks (LCN)*, pages 102–111, 2001.

[37] B. Yang and H. Garcia-Molina. Improving Search in Peer-to-peer Networks. *Proceedings of the 22nd Int'l Conf. on Distributed Computing Systems (ICDCS)*, pages 5–14, 2002.

[38] B. Zhao, J. Kubiatowicz, and A. Joseph. Tapestry: An Infrastructure for Fault-tolerant Wide-area Location and Routing. *Tech. Rep. UCB/CSD-01-1141, University of California at Berkeley, Computer Science Department*, 2001.