Chapter 1

Introduction

1.1 The Mobile Networks

Advances in network technology have provided extensive use of portable computers and enabled the on-line services through wireless communication channels. This kind of computing paradigm is called mobile computing. Mobile computing and communication devices will have an enormous impact on our lifestyle over the next several decades. In D.P. Agrawal’s descriptions [1], the mobile wireless systems have evolved over time, and the chronological development of first-generation (1G) and second-generation (2G) system is developed from 1970 to 1984 and 1982 to 1999, respectively. The 2-G mobile wireless systems have been designed for both indoor and vehicular environments with an emphasis on voice communication. An increased acceptance of mobile communication networks for conventional services has led demands for high bandwidth wireless multimedia service. These ever-growing demands require a new generation, third-generation (3G), of high-speed mobile infrastructure network that can provide the needed for high traffic volumes as well as flexibility in communication bandwidth or services.

The 3G systems (International Mobile Telecommunications 2000) need to support real-time data communication while maintaining compatibility with 2G systems. As indicated by Agrawal [2], mobile technology has improved substantially, making mobile devices remarkably convenient and affordable. R. Malladi and D.P. Agrawal [3] drew a table to provide an overview of different mobile wireless technologies (Table 1-1).

Technology	**Services / Features**	**Coverage Area**
Cellular	Voice and data through hand held phones	Continuous coverage
Wireless LAN (WLAN)	Traditional LAN with wireless interface	Only in local environment
GPS	Determines three dimensional position, and velocity	Any place on Earth
Satellite-based PCS	Mainly for paging	Almost any place on Earth
Ad hoc networks	Group of people come together for short time to share data	Similar to local area networks
Sensor networks	Tiny sensors with wireless capabilities	Small terrain
Bluetooth	All digital devices can be connected without any cable	Private ad-hoc groupings away from fixed network infrastructures

Table 1-1. Wireless technologies and associated characteristics.

The cellular mobile system is modeled as a geographical area, which consists of many hexagonal cells. A fixed base station, called the mobile support station (MSS), supports each cell. The mobile support station is static and connected through a dedicated wire-line link to an existing wired backbone. The mobile node, referred to as the mobile host (MH), is a part of one and only one cell at a time. Each mobile host can only communicate through the MSS of a cell with any particular node, whose position has been located. Communication between MH and MSS occurs through radio waves or infrared waves, which are wireless. The illustration of a cell with a MSS and MHs is shown in Fig. 1-1. In this dissertation, our fault-tolerant schemes are designed based on the cellular mobile networks described above.

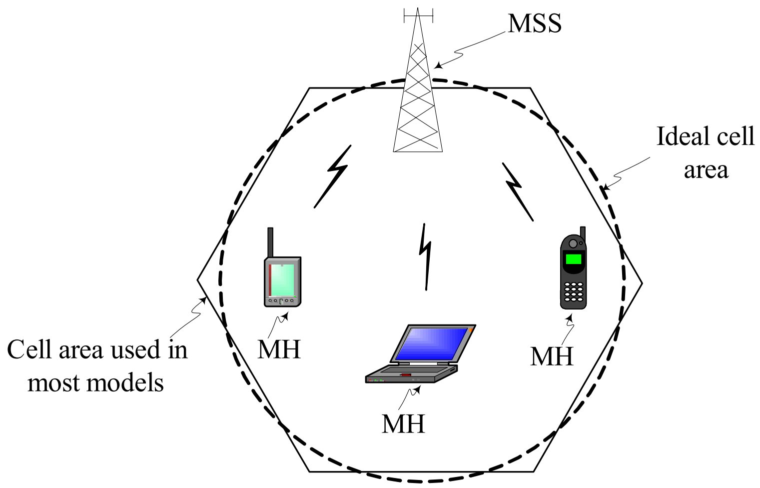

Fig. 1-1. Illustration of a cell with a MSS and mobile hosts (MHs).

1.2 Location Tracking in Mobile Networks

The mobile networks can provide wireless communication services to users on the move. Typically, mobile networks have a cellular architecture. An important issue in mobile networks is the location tracking problem (also called the location management problem). The movement of the mobile host (MH) can cause changes in the physical topology of the network over time. The location of a mobile host must be identified before a call to the mobile host can be connected. Generally location management involves two kinds of activities, called location updates and location queries. Many location management schemes have been proposed for PCS cellular networks with mobility databases [4-16]. When a mobile host changes its location, it should inform one or more location registers of its position. On the other hand, when a mobile host wishes to communicate with another host whose location is unknown, a query sequence is invoked.

In IS-41 and GSM schemes [17,18], they utilize home location registers (HLR) and visitor location registers (VLR) to keep track of the location information of the MHs. The latest location information of the MHs is recorded in the HLR/VLR databases through the location update procedure. When a call to a mobile host is made, the information is retrieved from the HLR/VLR databases through the query procedure. The basic HLR/VLR location tracking scheme is shown in Fig. 1-2. In a cellular structure, an MH needs to communicate with the MSS of the cell where the MH is currently located, and the MSS acts as a gateway to the rest of the world. Therefore, to provide a link, the MH needs to be in the area of one of the cells so that mobility of the MH can be supported. Several MSSs are connected through hardwires and are controlled by a mobile switch center (MSC). Several MSCs are interconnected to a public switched telephone network (PSTN).

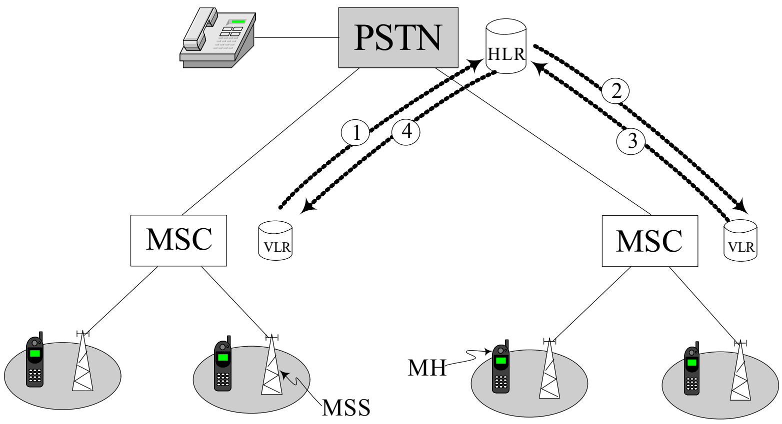

Fig. 1-2. The basic cellular system infrastructure and location tracking scheme.

1.3 Centralized and Distributed Location Tracking

Much research has been done on the issue of location management. These include both distributed [19-23] and centralized schemes [17,18]. Each scheme has its advantages and disadvantages. A centralized scheme is simpler to implement and manage, but it is neither robust, nor scalable. The distributed scheme may provide fault tolerance, load balance, scalability, and modularity at the expense of increased control traffic and connection delay.

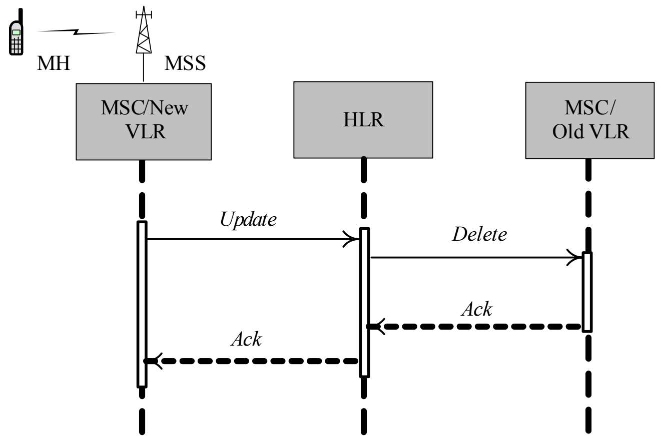

Fig. 1-3. Signal flow of location update for IS-41scheme.

Two standards currently exist for cellular mobile location management: the IS-41[17] and the GSM MAP [18]. Under the basic IS-41 centralized scheme, an MH is permanently registered with a home location register (HLR). When an MH moves into a new location area (LA), it reports to the new visitor location register (VLR) of this area. This VLR forwards the message to the HLR, which updates the location information of the MH. Then, the HLR issues a location deletion message to the old serving VLR. In the query sequence, the VLR queries the HLR for the callee, and then the HLR will query the callee's VLR. Upon receiving the callee's location, the HLR will forward the location information to the calling VLR. The signal flow of the location update and query are shown in Fig. 1-3 and Fig. 1-4, respectively.

Based on distributed schemes, Lin et al. [24,25] employed hierarchical databases to store location information. This tree-like architecture can reduce signaling message traveling distance, but it increases multiple query operations and causes unbalanced load for some servers. Also, most hierarchical mobility schemes suffer from reliability issues because of single points of failure at each level of the network hierarchy.

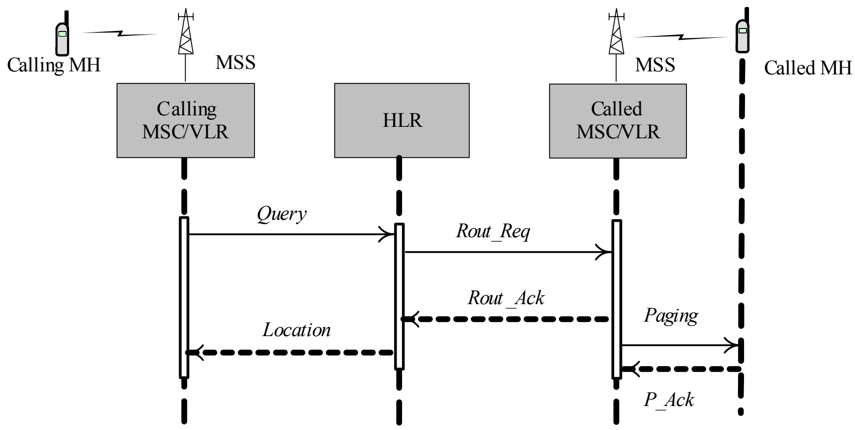

Fig. 1-4. Signal flow of location query for IS-41scheme.

Some other distributed location management schemes [20,22] fully replicate location information in all the location registers (LRs) of a local area. In Ratnam's scheme [22], the LRs replace the centralized VLRs and HLRs. When a mobile host updates with an LR, the new location information is disseminated to all other LRs in the network. In the CDLM scheme of Hassan [20], the originating basestation broadcasts a location query message to all basestations in the network. Both schemes introduce heavy cost with information fully replicated or message broadcasted in the network.

1.4 Quorum-Based and Fault-Tolerant Scheme

Some distributed location management schemes [23,26,27] employ quorum-based servers to store location information. In Prakash et al. [23], a dynamic hashed location management scheme with grid-based, reduced overhead grid-based, and CWlog quorum construction was proposed. With these three quorum constructions, the quorum size is $2\sqrt{N}-1$, $\sqrt{N}$, and d /2+log d (for a crumbing wall of d rows), respectively, but they are not constructed to be a symmetric quorum with any arbitrary numbers of nodes.

In Prakash's paper [26], a dynamic load-balanced location management scheme with an iterative and grid-based quorum construction was proposed. In Prakash's scheme, N location servers are divided into quorums of cardinality of $0.97N^{0.63}$ and $2\sqrt{N}-1$, respectively. Any two quorums of servers have at least one common server. Upon receiving a location update request, the update procedure uses a hashing function, which takes the mobile's ID and its current location into account to select the quorum for updating. When it receives a call request, the search procedure uses the hashing function along with the caller's location and the called mobile's ID to select the quorum for querying. From the common server, finding location information is guaranteed.

Ihn-Han Bae in [27] proposed a distributed location management scheme that uses the quorum, which is based on a triangular configuration of location servers. Without using virtual identity, Bae's algorithm is not only simpler than Prakash's [26] algorithm, but also less expensive. Since the quorum structure of the triangular configuration is not symmetric, Bae's scheme is neither load-balanced, nor fully tolerant.

Many strategies have been proposed for location managements in PCS networks, but they all assume the databases of the systems to be fault free. Only a few papers pay attention to fault-tolerance issues on location databases. Fault-tolerant issues are addressed in [28-30]. In IS-41 scheme, the success of a call connection requires the HLR and the callee's current VLR to be failure-free. In Xiao's paper [28], seven backoff strategies for demand re-registration were proposed. In Liu et al. [29], there were proposed schemes to tolerate the failure of the VLRs.

1.5 Thesis Outline

In this dissertation, we propose a new structure called *Legion* and develop a simple distributed location tracking scheme based on this new structure. This simple scheme, called *LegRing*, can achieve fault-tolerant and efficient location updating and querying. We also propose another fault-tolerant location management scheme, which is based on the cellular quorum system. This scheme can tolerate the failures of one or more location server(s) without adding or changing the hardware of the systems in the two-tier networks. Meanwhile, with a region-based approach, it yields better connection establishment delay.

The rest of this dissertation is organized as follows: Chapter 2 introduces our new *Legion* structure and describes a new location tracking scheme. In Chapter 3, we present a quorum-based scheme for distributed location tracking. Chapter 4 describes our approach to providing a fully distributed fault-tolerant location management for PCS networks. Chapter 5 proposes a new scheme with quorum system to tolerate the failures of the VLRs and HLR in centralized two-tier networks, without adding or changing the hardware of the systems. In chapter 6, another new scheme with cellular quorum construction and region-based approach is proposed to tolerate the failures of the HLR and VLRs. Chapter 7 concludes this dissertation based on our research.