Wireless Personal Communications (2006) 36: 195–212
DOI: 10.1007/s11277-006-0378-y

A Fault-Tolerant Scheme for Mobility Management in PCS Networks

MING-JENG YANG[1] and YAO-MING YEH[2]

[1]*Department of Information Technology, Takming College, Taipei, 114 Taiwan*
E-mail: mjyang@ieee.org
[2]*Department of Information & Computer Education, National Taiwan Normal University, Taipei, 106 Taiwan*
E-mail: ymyeh@ice.ntnu.edu.tw

Abstract. One of the most important and challenging issues in the design of personal communication service (PCS) systems is the management of location information. In this paper, we propose a new fault-tolerant location management scheme, which is based on the cellular quorum system. Due to quorum's salient set property, our 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, our scheme stores/retrieves the MH location information in the location servers of a quorum set of the local region as much as possible to avoid long delays caused by the possible long-distance of VLR and HLR. Thus, it yields better connection establishment and update delay.

Keywords: PCS networks, fault tolerant, location tracking, quorum system, *Legion*.

1. Introduction

The third generation networks, called the personal communication service (PCS) networks, can provide wireless communication services to users on the move. Typically, PCS networks have a cellular architecture. An important issue in PCS networks is 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. 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.

Two standards currently exist for PCS location management: the IS-41 [1] and the GSM MAP [2]. Both schemes use a two-tier system of home location register (HLR) and visitor location register (VLR) databases.

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 the new 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 called MH, and then the HLR will query the VLR of the called MH. Upon receiving the caller's location, the HLR will forward the location information to the calling VLR.

Many strategies have been proposed for location management of PCS networks [3–5], but they all assume the databases of the systems to be fault free. We assume that both the database (VLR or HLR) and the network link may fail. Due to the frequent changes of location information for a mobile in the VLR or HLR, the location information may be corrupted. Another problem of database fault may be caused by power down or hardware failure. When the VLR or HLR fails, calls are dropped as the called MH can not be queried, thus reducing the quality of service (QoS) provided by the network. A link failure in the network could partition the network resulting in a loss of location updates and query. Only a few papers pay attention to fault-tolerance issues on location databases. Fault-tolerant issues are addressed in [6–8]. In the IS-41 scheme, the success of a call connection requires the HLR and the called MH's current VLR to be failure-free. In Xiao's paper [6], seven backoff strategies for demand re-registration were proposed. Liu et al. in [7] proposed schemes to tolerate the failure of the VLRs.

In our previous work [9], the *Legion* structure was defined and a *LegRing* scheme was developed. In this paper, based on the *Legion* structure, we design a new cellular quorum scheme and develop a new region-based quorum approach for location update and query management. In two-tier networks, if the VLR or HLR fails, the subscribers' services will be seriously degraded due to the loss or corruption of location information. Our proposed scheme tolerates the failures of VLRs and the HLR. In addition, our scheme has fast query and update response in comparison with the traditional two-tier scheme.

In the next section, a description of the system is provided. In Section 3, we propose a cellular quorum construction, which is used to design a fault-tolerant mobility management system. In Section 4, a design approach for the system is described. In Section 5, the latency of query and update is analyzed. Finally, in Section 6, we draw conclusions based on our research.

2. System Description

The framework of the system network is shown in Figure 1. The system is modeled as a geographical area that consists of many cells. The cells are aggregated into contiguous geographical areas called location areas (LAs), which are symbolized by the big hexagons shown in Figure 1. The mobile node, referred to as the mobile host (MH), is a part of only one cell at a time. A fixed base station, called the mobile support station (MSS), supports each cell. The communication between MH and MSS is through radio waves or infrared waves which are wireless. The mobile support station is static and connected through a dedicated wire-line link to a mobile switch center (MSC). An MSC, which typically provides switching and can be viewed as a bridge for connecting the wireless network and the wired network, serves one LA. A visitor location register (VLR) is maintained at each LA, which stores the temporary records of the current location information of the MH while it is not at "home". Several LAs are grouped into a fault-tolerant region (FTR). All the VLRs in one FTR are formed as a local fault-tolerant system, which can tolerate the failures of one or more VLRs. The mechanism of the fault-tolerance will be described in Section 4. All MSCs are finally connected to a Public Switched Telephone Network (PSTN) that is a wired backbone network. The system is a two-tier database architecture consisting of a home location register (HLR) and visitor location registers (VLRs). The HLR cooperates with VLRs to track and find the locations of MHs. The information, recorded in the HLR database, can assist the system in finding the FTR

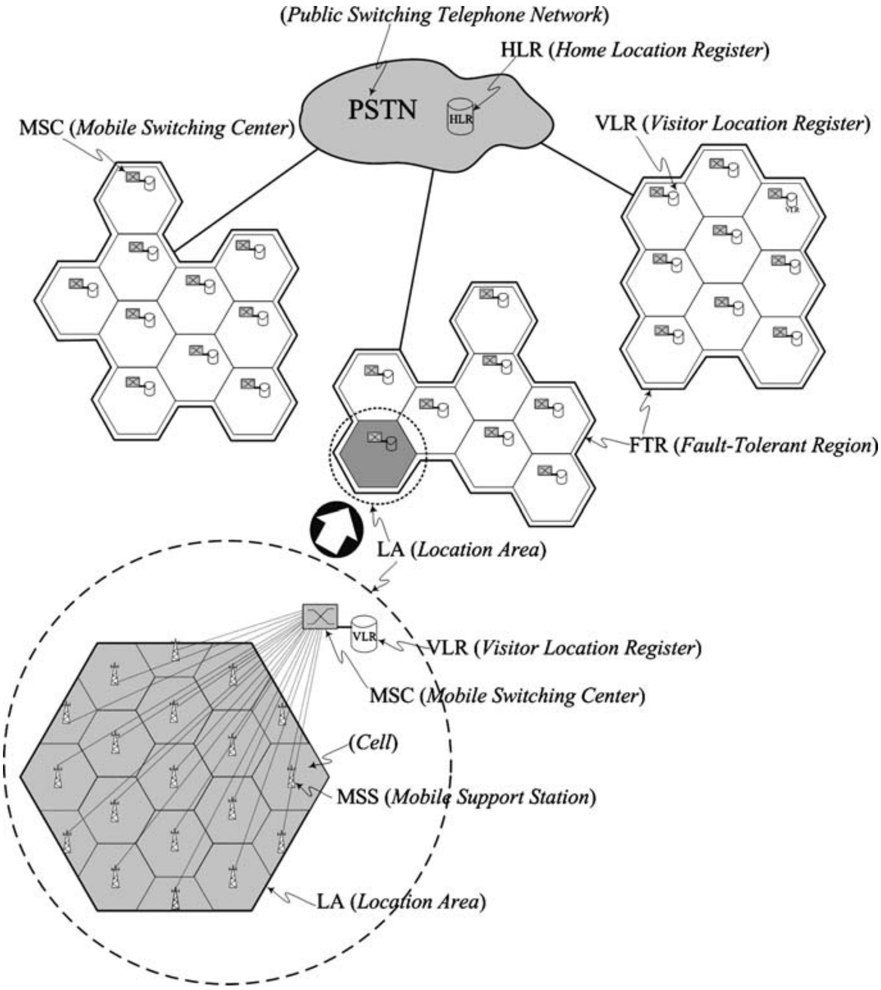


Figure 1. System architecture.

where the MH is currently located. Then, the LA where the MH resides is tracked. In the next section, we will describe the cellular quorum approach that will be used in our system design.

3. Cellular Quorum Constructions

In the FTR, all the VLRs play the role of a local fault-tolerant system. Based on a quorum-based scheme, the fault-tolerant system can tolerate the failures of VLRs. In this section, in order to prove the correctness of cellular quorum system, we introduce some definitions and theories, such as quorum, set system [10], and *Legion* structure [9].

A set system [10] is composed of a number of quorums.

Definition 1. A set system [10] $C = \{Q_1, Q_2, \ldots, Q_n\}$, $1 \leq n$, is a collection of nonempty subsets $Q_i \subseteq U$ of a finite universe U.

Each element Q_i of C in Definition 1 is called a **quorum.**

Take the set $C = \{\{a, b, c\}, \{d, e\}, \{f, k, h\}\}$ as an example. C is a set system, and any element in C, for example, $\{f, k, h\}$ is a quorum.

In the following, we introduce a definition of *Legion* structure, which is our previous work [9]. A *Legion* is constructed from set systems.

Definition 2. A Legion structure $\{C_i, C_j\}$ is a collection of two set systems that has the following properties:

I $C_i = \{Q_1, Q_2, \ldots, Q_n\}$ and $C_j = \{Q_1, Q_2, \ldots, Q_m\}$ are set systems. $(1 \leq n, m)$
II For any pair of quorums $Q_s \in C_i$ and $Q_t \in C_j$, there is $Q_s \cap Q_t \neq \emptyset$. That is, Q_s and Q_t have at least one common element. $(1 \leq s \leq n, 1 \leq t \leq m)$

Theorem 1. *The Legion structure $\{C_i, C_j\}$ defined in Definition 2 can be used as a mathematic model for quorum-based location management in PCS networks.*

Proof: According to the Definition 2, any two quorums of a pair (Q_s, Q_t) have at least one common element, where Q_s is a quorum in C_i and Q_t is a quorum in C_j. This structure can be applied to develop a location management scheme for PCS systems. In PCS systems, if one of the servers requires information from the other, it suffices to query one server from an appropriate quorum. While using this quorum-based location scheme, we can assign quorums in C_i as update-quorums, and quorums in C_j as query-quorums. According to the definition of *Legion*, the set of queried servers is bound to contain at least one server that belonged to the quorum that received the latest update (i.e. the update-quorum and query-quorum have at least one common server). Hence, the newest location information is extracted from the common server(s). □

In the following, we introduce a cellular quorum approach which is based on the hexagonal location areas (LAs). Consider a fault-tolerant region (FTR) with continuous location areas (LAs) shown in Figure 2. Each LA can be identified with position coordinates (x, y). We denote the LA(x, y) as the position of LA that is located at the position coordinates (x, y), where x denotes the column number of the FTR and y denotes the sequential order where it appears in its column. Some sequences of patterns in the FTR are employed as Update-quorum (U-quorum) and Query-quorum (Q-quorum).

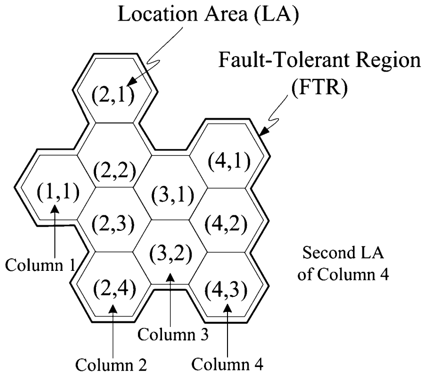


Figure 2. A fault-tolerant region (FTR) and its position coordinate.

Definition 3. Let N be the total number of columns in a FTR region and M_i be the total number of rows in the column i of the FTR region. In the cellular quorum construction of one FTR region, the Update-set (U-set) and Query-set (Q-set) are defined as follow:

$$U\text{-set} = \{U_i \mid 1 \leq i \leq N, \quad U_i = \{(i, j) \mid 1 \leq j \leq M_i\}\};$$
$$Q\text{-set} = \{Q_s \mid 1 \leq s, \quad Q_s = \{(i, r_i^s) \mid 1 \leq ig \leq N\}\},$$

where $r_i^s(1 \leq r_i^s \leq$ M $_i)$ is the row number randomly selected with variable s in column i; i, j, s, N, M $_i$ and r_i^s are all natural numbers. Each element of U-set and Q-set is called an U-quorum and a Q-quorum, respectively.

For example, consider a fault-tolerant region (FTR) with continuous location areas (LAs), which are identified with position coordinates (x, y) shown in Figure 3. Then, we have U-set $= \{U_1, U_2, U_3, U_4\} = \{\{(1, 1)\}, \{(2, 1), (2, 2), (2, 3), (2, 4)\}, \{(3, 1), (3, 2)\}, \{(4, 1), (4, 2), (4, 3)\}\}$, where U_i, for example, $\{(2, 1), (2, 2), (2, 3), (2, 4)\}$, is an update quorum ($U$-quorum) shown in Figure 3(a); moreover, according to the random selection (i.e., randomly pick one LA from each column), we may have the Q -set $= \{Q_1, Q_2, \ldots, Q_n, \ldots\} = \{\{(1, 1), (2, 4), (3, 1), (4, 2)\}, \{(1, 1), (2, 1), (3, 2), (4, 2)\}, \ldots, \{(1, 1), (2, 4), (3, 3), (4, 1)\}, \ldots\}$, where Q_i, for example, $\{(1, 1), (2, 4), (3, 1), (4, 2)\}$, is an query quorum ($Q$-quorum) shown in Figure 3(b).

Theorem 2. *The U-set and Q-set defined in Definition 3 satisfy the properties of Legion structure $\{C_i, C_j\}$.*

Proof: According to Definition 2, the properties of *Legion* structure $\{U$-set, Q-set$\}$ are: (I) U-set and Q-set are set systems. (II) Any pair of U-quorum in U-set and Q-quorum in Q-set have joint elements. We need to prove these properties. First, according to Definition 3 and Definition 1, it is easy to see that U-set and Q-set are set systems. Second, according to Definition 3, the U-quorum's elements are picked from the same column's LAs of FTR and one LA from each column of FTR is randomly picked as the Q-quorum's element. Hence, the U-quorum's elements are all in the same column and one of them must be chosen as one member in the Q-quorum. Thus, there is one intersectional element of each pair of U-quorum and Q-quorum. This satisfies property (II). Hence, the U-set and Q-set of cellular quorum system satisfy the properties of *Legion* structure. □

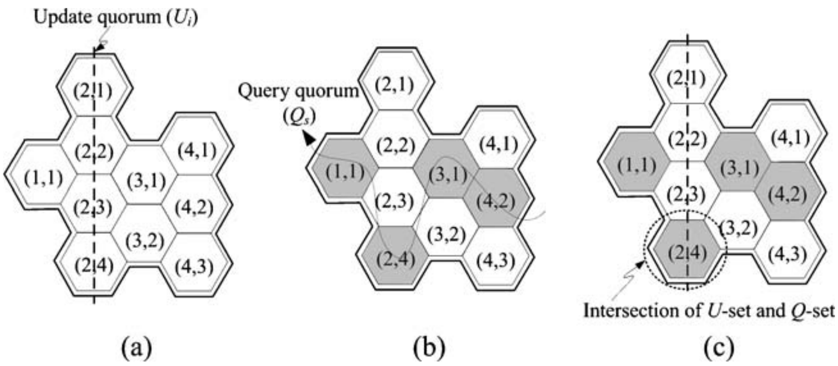


Figure 3. The construction and intersectional property of U-quorum and Q-quorum.

Consider the previously described example. The U-quorum, for example, $U_2 = \{(2, 1), (2, 2), (2, 3), (2, 4)\}$, is composed from all the LAs of column 2 of FTR (Figure 3(a)). The Q-quorum, for example, $Q_1 = \{(1, 1), (2, 4), (3, 1), (4, 2)\}$, is composed from the LAs that are picked from each column (Figure 3(b)). Thus, there is one intersectional element (2, 4) of the U-quorum and Q-quorum (Figure 3(c)).

4. The System Design

In this section, we use the cellular quorum approach described in Section 3 to devise a fault-tolerant location tracking system. Based on the intersectional property of the U-quorum and Q-quorum, the location information is disseminated to VLRs of the U-quorum and can be extracted from one of them by using the Q-quorum even though one or more location servers fail. In cellular PCS systems, location management is achieved by querying and updating. A query occurs when a host needs to communicate with another mobile host whose location is unexpected, and an update occurs when a mobile host changes its location. A detailed description of update and query procedures will be discussed in the Sections 4.1 and 4.2.

In the cellular quorum approach, several LAs are grouped into a fault-tolerant region (FTR) and identified with a region number (FTR_no). All the VLRs in one FTR (with same FTR_no) are formed as a local fault-tolerant system, which can tolerate the failures of one or more VLRs. In the next section, the location update procedure will be described.

4.1. Location Update

We divide the location update procedure into two types: region update (Figure 4) and home update (Figure 5). When a mobile host moves from one LA to others in the same FTR (i.e., with the same FTR_no), the region update is invoked and its location information has to be updated locally in the same FTR region. When an MH roams to another FTR, a home update procedure will be triggered (i.e., involves the update of the HLR). In a traditional location management scheme, an update procedure that would need to coordinate between HLR and VLRs will occur whenever an MH moves to a new LA. In our region update approach, when an MH moves to a new LA in the same FTR, instead of updating information to the HLR, it merely reports its location to the VLRs of U-quorum in the local FTR. By the user mobility

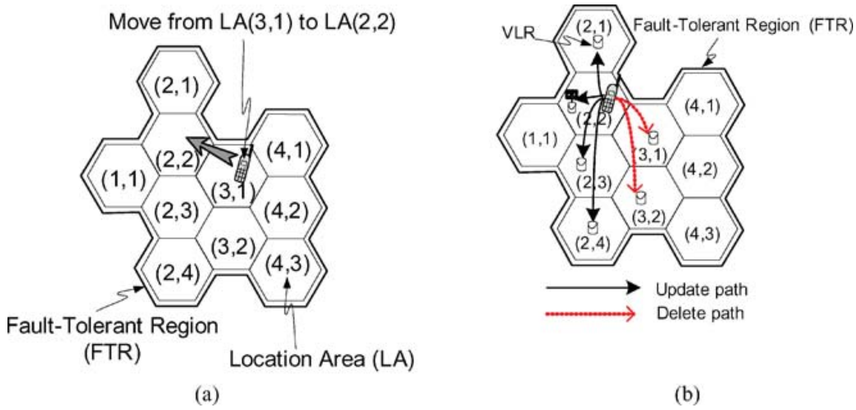


Figure 4. Region Update. (When a MH moves from one LA to another LA in the same FTR).

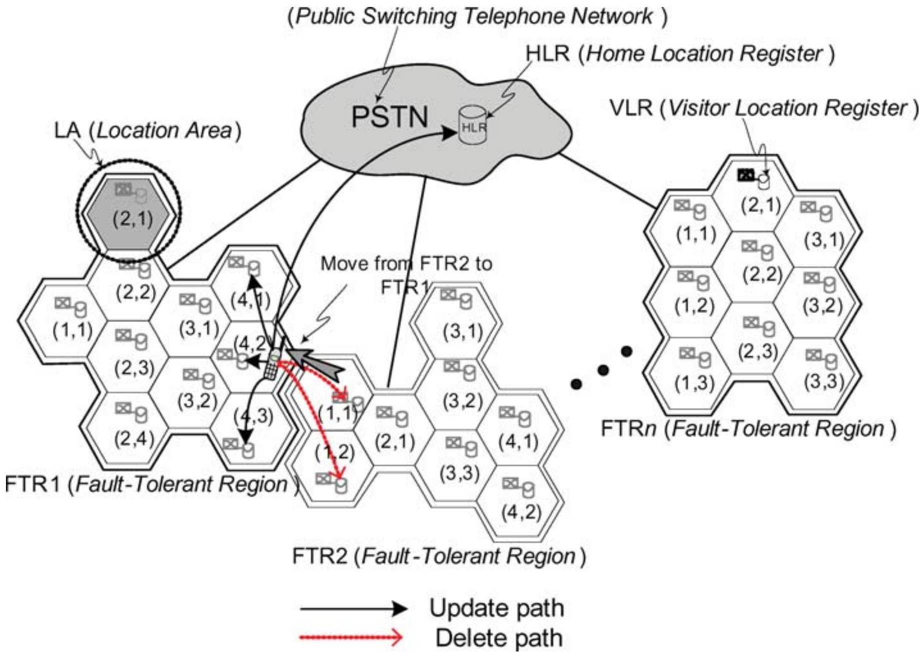


Figure 5. Home update. (When a MH moves from one LA to another LA in different FTR).

behavior model, we could suitably choose the coverage areas of one FTR. Then, with high probability, the users would move around in the same FTR. Hence, in most conditions, the update procedures will be handled locally.

In the following, we use the cellular quorum approach, defined in Definition 3, to implement the location update algorithms. Therefore, the following steps are performed as the location update procedure:

1. The procedure sends the UPDATE message to the serving MSC/VLR which forwards this message to other VLRs in the U_i quorum of the FTR, where i is the column in which the MH resides currently. Meanwhile, the FTR_no is checked. If new and old VLR are in different FTR, then the home update (Figure 5) is invoked and the procedure concurrently sends the REDIRECT pointer (i.e., a forwarding route information including the FTR_no) to the HLR.
2. Upon receiving the UPDATE message or REDIRECT pointer, the VLRs or HLR add the new information received to their databases and send back the ACK message.
3. If the procedure does not receive all the ACK messages from all VLRs in the quorum during a given period of time, then it randomly selects other quorum U_j, sends the UPDATE message to all VLRs in the new quorum again, and goes to step 2; otherwise, it goes to the next step.
4. The DELETE message is sent to the current MSC which forwards this message to all the VLRs in the U_k quorum of the FTR, where k is the column in which the MH resided previously.
5. Upon receiving the DELETE message, the VLRs delete the old location information in their databases and send back the ACK message.
6. If the procedure does not receive all the ACK messages from all VLRs in the quorum during a given period of time, then it resends the DELETE message to VLRs that did not send back ACK messages, and goes to step 5; otherwise, it stops.

Take a region update as an example. Consider a fault-tolerant region (FTR) with continuous location areas (LAs), which can be identified with position coordinates (x, y) shown in Figure 4. According to the Definition 3, the U-set is $\{U_1, U_2, U_3, U_4\} = \{\{(1, 1)\}, \{(2, 1), (2, 2), (2, 3), (2, 4)\}, \{(3, 1), (3, 2)\}, \{4, 1\}, (4, 2), (4, 3)\}\}$. If a mobile host h moves from LA(3, 1) to LA(2, 2), new location information is sent to all the VLRs of LA(2, 1), LA(2, 2), LA(2, 3), and LA(2, 4) in the U_2 quorum of the FTR and the ACK messages are sent back. Then, the DELETE message is sent to all the VLRs of LA(3, 1) and LA(3, 2) in the U_3 quorum of the FTR, where the column three is the column in which the MH resided previously, and the ACK messages are sent back.

In the IS-41 location management scheme, if the distance between the visited area and the HLR is large, the signaling delay for the location update is long. Our approach, a region update, is a way to reduce the signaling delay for performing updates. This is especially useful when the MH always moves within a particular FTR.

Take a home update as an example. Consider n fault-tolerant regions (FTRs) each with continuous location areas (LAs), which can be identified with position coordinates (x, y) shown in Figure 5. According to the Definition 3, the U-set of FTR1 is $\{U_1^1, U_2^1, U_3^1, U_4^1\} = \{\{(1, 1)\}, \{(2, 1), (2, 2), (2, 3), (2, 4)\}, \{(3, 1), (3, 2)\}, \{4, 1\}, (4, 2), (4, 3)\}\}$ and the U-set of FTR2 is $\{U_1^2, U_2^2, U_3^2, U_4^2\} = \{\{(1, 1), (1, 2)\}, \{(2, 1)\}, \{(3, 1), (3, 2), (3, 3)\}, \{4, 1\}, (4, 2)\}\}$. If a mobile host h moves from LA(1, 1) of FTR2 to LA(4, 2) of FTR1, new location information is sent to the all the VLRs of LA(4, 1), LA(4, 2), and LA(4, 3) in the U_4^1quorum of the FTR1 and the ACK messages are sent back by VLRs. Meanwhile, the REDIRECT pointer (i.e., a forwarding route information including the FTR identifier that indicates the FTR in which the MH currently resides) is sent to the HLR and the ACK message is sent back by the HLR. Then, the DELETE message is sent to all the VLRs of LA(1, 1) and LA(1, 2) in the U_1^2 quorum of the FTR2 and the ACK messages are sent back by VLRs.

As described in the above procedure, the VLRs in the selected quorum of the current FTR maintain the *id* of the current MSC serving the MH. For the HLR, it maintains the *id* of the FTR where the MH currently resides.

When a mobile host wishes to communicate with another host whose location is unknown, the query procedure is invoked. Again, the fault-tolerant cellular quorum approach is used in the location query process. A detailed description of location query will be discussed in the next section.

4.2. Location Query

We have designed the query procedure as two types: region query (Figure 6) and home query (Figure 7). A region query occurs when the calling and called MH are in the same FTR. The query is handled locally in the same FTR. Figure 6 shows the region query processes. If the calling and called host are in different FTRs, the communication invokes the home query procedure. Figure 7 shows the home query processes. The location query procedure is described as follows:

1. The calling MH queries the location information of the called MH from its local VLR (Figure 6(a)). If the information is found in the local VLR, then the local VLR forwards the routing request (ROUTREQ) message to the called MH's current MSC/VLR (Note that the called MH and the calling MH may be located in the same LA) and sends the ACK message back to the procedure. Finally, the current MSC/VLR launches a paging process

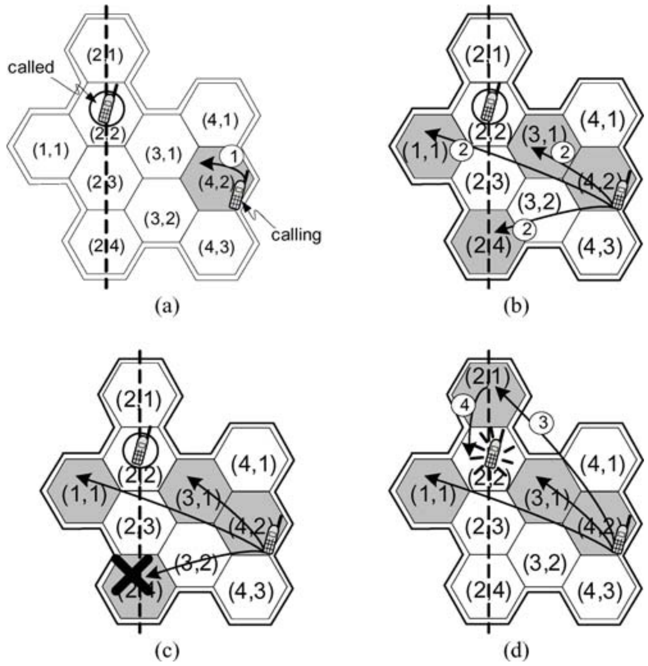


Figure 6. Region query.

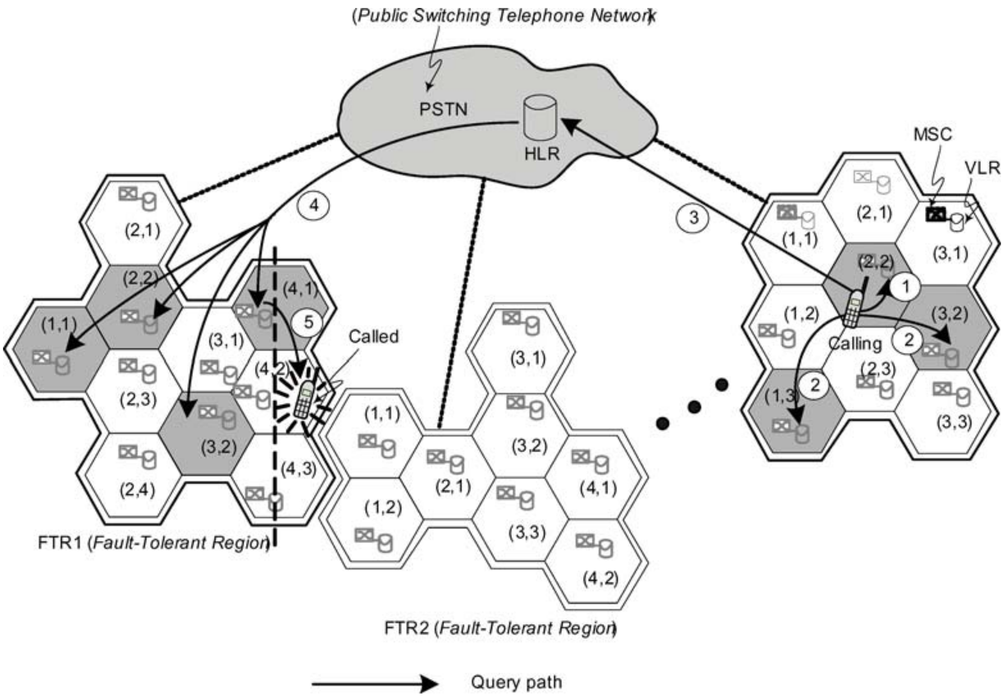


Figure 7. Home query.

to find the called MH's current MSS. At this point the call is established and then the query procedure is stopped. Otherwise, it goes to next step.
2. The procedure randomly selects a quorum Q_s and sends a QUERY message to all the VLRs (except the VLR which locates at the same column with the calling MH's local VLR) in the quorum (Figure 6(b)).
3. Upon receiving the QUERY message, the VLR, which does not have a copy of queried information, sends the null ACK message back to the procedure. On the other hand, if the information is found in the VLR, then it forwards the routing request (ROUTREQ) message to the called MH's current MSC/VLR and sends the ACK message back to the procedure. Then, the current MSC/VLR launches a paging process to find the called MH's current MSS. At this point the call is established and then the query procedure is stopped. Otherwise, it goes to next step.
4. If the call is not established after a given period of time, then the procedure resends the QUERY message to the randomly selected VLRs, which located at the same column with the VLRs that did not send back the ACK messages and goes to step 3 (Figure 6(d)). Otherwise, it stops.
5. If the calling and called MH are in different FTRs (with different FTR_no), then the null ACK messages, which issued by all VLRs in the quorum of the local FTR, would be received by the procedure.
6. Then, the HLR is inquired for the location information of the called MH.
7. According to the REDIRECT pointer in the database, the HLR forwards the QUERY message to the VLRs in the randomly selected quorum, Q_j, of the remote FTR in which the called MH resides.
8. Upon receiving the QUERY message, the VLR, which has not a copy of queried information, sends the null ACK message back to the procedure. On the other hand, if the information is found in the VLR, then it forwards the routing request (ROUTREQ) message to the called MH's current MSC/VLR and sends the ACK message back to the procedure. Then, the current MSC/VLR launches a paging process to find the called MH's current MSS. At this point the call is established and then the query procedure is stopped. Otherwise, it goes to next step.
9. If the call is not established after a given period of time, then the procedure resends the QUERY message to the VLRs, which are located at the same column with the VLRs that did not send back the ACK messages and goes to step 8. Otherwise, it stops.

Take a region query as an example. Consider a fault-tolerant region (FTR) with position coordinates (x, y) shown in Figure 6. According to the Definition 3, we may have the Q-set $= \{Q_1, Q_2, \ldots, Q_n, \ldots\} = \{\{(1, 1), (2, 4), (3, 1), (4, 2)\}, \{(1, 1), (2, 1), (3, 2), (4, 2)\}, \ldots, \{(1, 1), (2, 4), (3, 3), (4, 1)\}, \ldots\}$. Assume a called host h stored its information at the VLRs of LA(2, 1), LA(2, 2), LA(2, 3), and LA(2, 4) in the U_2 quorum of the FTR. When a calling host h' at LA(4, 2) of the same FTR wants to communicate with host h, it first acquires the information from the local VLR of LA(4, 2). Since the information is not found in the local VLR, then the procedure randomly select a quorum, for example, Q_1, and sends a QUERY message to all the VLRs of LA(1, 1), LA(2, 4), and LA(3, 1) in the quorumQ_1 (except the local VLR of LA(4, 2)). But, the only VLR that keeps the information of h in quorum Q_1 has failed. So, the procedure resends the QUERY message to the randomly selected VLR of LA(2, 1), which located at the same column with the VLR of LA(2, 4). The VLR of LA(2, 1), which has the location information of called MH, forwards the routing request

(ROUTREQ) message to the called MH's current MSC/VLR. Then the query procedure is completed.

The completion of the region FTR query procedure is quick, since it is done locally. Compared to the query in the IS-41 scheme, which accesses the HLR that may be far away from the calling MH, our region FTR query is more time effective since the access to HLR could involve two long-distance legs.

Take a home query as an example. Consider n fault-tolerant regions (FTRs) system with position coordinates (x, y) shown in Figure 7. According to the Definition 3, the Q-set of FTR1 may be $\{Q_1^1, Q_2^1, Q_3^1, \ldots\} = \{\{(1, 1), (2, 4), (3, 1), (4, 2)\}, \{(1, 1), (2, 1), (3, 2), (4, 2)\}, \{(1, 1), (2, 2), (3, 2), (4, 1)\}, \ldots\}$ and the Q-set of FTRn may be $\{Q_1^n, Q_2^n, Q_3^n, \ldots\} = \{\{(1, 3), (2, 2)\}, \{(3, 2)\}, \{(1, 1), (2, 1), (3, 2)\}, \{(1, 3), (2, 3), (3, 3)\}, \ldots\}$. Assume a called host h stored its information at the VLRs of LA(4, 1), LA(4, 2), and LA(4, 3) in the quorum U_4^1 of the FTR1. When a calling host h' at LA(2, 2) of FTRn wants to communicate with host h, it first acquires the information from the local VLR of LA(2, 2). Since the information is not found in the local VLR, then the procedure randomly select a quorum, for example, Q_1^n, and sends a QUERY message to all VLRs of LA(1, 3), and LA(3, 2) in the quorum Q_1^n (except the local VLR of LA(2, 2)). Since the called MH is not located at FTRn, all VLRs in Q_1^n have not the location information of called MH and send null ACKs back. Then, the HLR is inquired for the location information of the called MH. According to the REDIRECT pointer in the database, the HLR forwards the QUERY message to the VLRs of LA(1, 1), LA(2, 2), LA(3, 2), and LA(4, 1) in the randomly selected query quorum Q_3^1 of the FTR1 in which the called MH resides. Since (4, 1) is the intersectional element of U_4^1 and Q_3^1, the VLR of LA(4, 1) has the location information of called MH and forwards the routing request (ROUTREQ) message to the called MH's current MSC/VLR. Then the home query procedure is completed.

4.3. Dealing with the Failure of the HLR

In some accident situations, the HLR may be down for a period of time. In this section, we will discuss the approach to deal with the situation when the HLR fails.

If the HLR is down, then the home query can not be performed. When a mobile host wishes to communicate with another host whose location is not in the same FTR, the connection can not be achieved. In this situation, the procedure is switched to multicast the QUERY message to all the VLRs in the randomly selected Q-quorum of each FTR (Figure 8). Upon receiving the QUERY message, the VLR, which has not a copy of queried information, sends the null ACK message back to the procedure. On the other hand, if the information is found in the VLR, then it forwards the routing request (ROUTREQ) message to the called MH's current MSC/VLR and sends the ACK message back to the procedure. Then, the current MSC/VLR launches a paging process to find the called MH's current MSS. At this point the call is established and then the query procedure is stopped.

5. Latency Analysis

Since our scheme retrieves the MH location information in the location servers of a quorum set of the local region as much as possible to avoid long delay caused by the possible long-distance of VLR and HLR, we will analyze and compare the query and update latency between proposed scheme and the IS-41 scheme. We introduce parameters for analyses in the following:

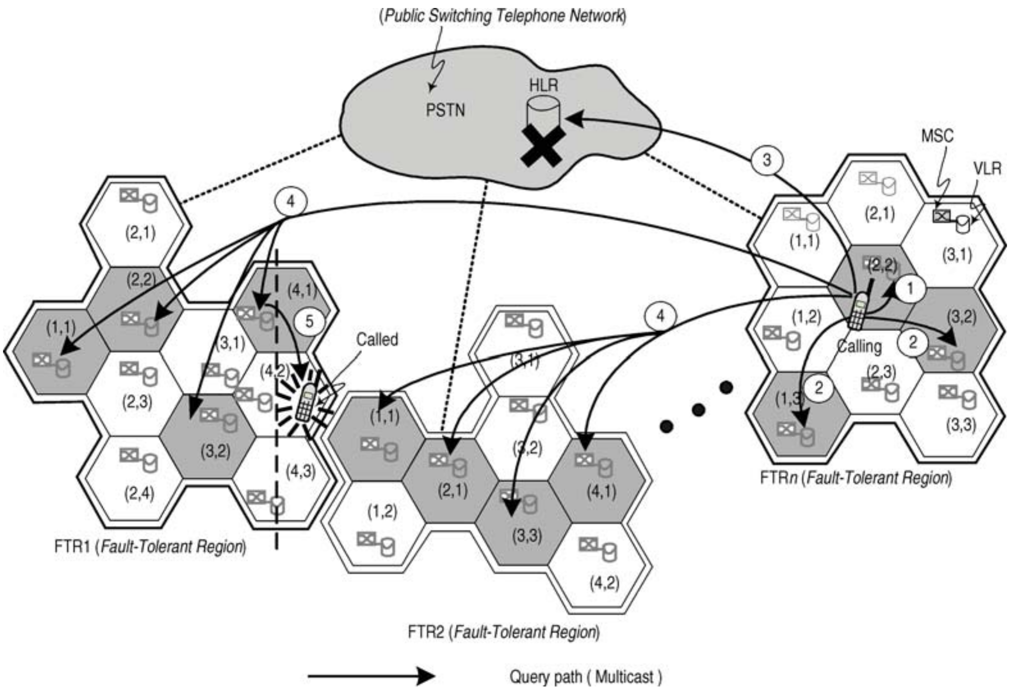


Figure 8. Dealing with the failure of the HLR.

- L_l: latency for transmitting a message through the local link (within the FTR).
- L_r: latency for transmitting a message through the remote link (out of the FTR).
- L_{VLR}: latency for a query to the VLR.
- L_{HLR}: latency for a query to the HLR.
- L_{paging}: the latency to page a user in a LA.
- h: region hit ratio (the ratio of hit query that the location information of called MH is found in the FTR where the calling MH currently resides).
- p: boarder-crossing probability out of an FTR (the probability that the MH moving to another LA out of FTR).

5.1. Query Latency

We first calculate the query latency for our proposed scheme, which is associated with the hit ratio, h. The h is assumed to be a region hit ratio. The query latency is divided into two parts: region query latency, L_{Region}, and home query latency, L_{Home}. That is, L_{Region} and L_{Home} are the query latencies of the protocols when there is a hit query (which means the location information is found in the FTR where the calling MH currently resides) and a miss query, respectively.

In addition to latencies for sending messages through the local link, there are database access latencies of VLRs. In hit case, we have a query latency equation

$$L_{\text{Region}} = 6L_l + 2L_{\text{VLR}} + L_{\text{paging}}. \tag{1}$$

In the case of a miss, the HLR and LAs of the remote FTR are inquired. Therefore, additional latencies for sending messages through the remote links, the latencies for querying the redirect pointer of the HLR, and latencies for access to the VLRs are needed. Thus, we have a query latency equation of a miss case

$$L_{\text{Home}} = 6L_l + 4L_r + 3L_{\text{VLR}} + L_{\text{HLR}} + L_{\text{paging}}. \tag{2}$$

The average query latency L_{Query} is then given by

$$L_{\text{Query}} = hL_{\text{Region}} + (1 - h)L_{\text{Home}}. \tag{3}$$

In the IS-41 approach, a query message is transmitted to the HLR to acquire the called MH's location information; and the routing-request message is issued from HLR to the called VLR. Then, the location information is sent to the calling MSC. Therefore, the average query latency for IS-41 scheme is

$$L_{\text{Query}}^{IS-41} = 2L_l + 4L_r + 2L_{\text{VLR}} + L_{\text{HLR}} + L_{\text{paging}}. \tag{4}$$

In [11], the relation between the hit ratio and call to mobility ratio (*CMR*) was calculated as

$$CMR = h/(1 - h), \tag{5}$$

where the *CMR* was expressed as the ratio of the call arrival rate to the mobility rate.

5.2. Update Latency

In this section, we will analyze the average update latency both for our and IS-41 schemes. We first calculate the update latency for our proposed scheme, which is associated with the probability, p. The p is assumed to be a boarder-crossing probability out of an FTR region (i.e., the probability that the MH moving to another LA out of FTR). The update latency is divided into two parts: region update latency, L'_{Region}, and home update latency, L'_{Home} (i.e., involves the latency of the update of HLR).

In addition to latencies for sending messages through the local link, there are database update latencies of VLRs. In region update case, we have an update latency equation

$$L'_{\text{Region}} = 4L_l + 2L_{\text{VLR}}. \tag{6}$$

In contrast to the region update procedure, the home update procedure must transmit update messages through the local and remote links to the local VLRs and HLR concurrently since the MH roams to another FTR. Therefore, additional latencies for sending messages through the remote links, the latencies for updating the redirect pointer of the HLR, and latencies for access to the VLRs are needed. Thus, we have a update latency equation of a home update case

$$L'_{\text{Home}} = 2L_l + 2L_r + L_{\text{VLR}} + L_{\text{HLR}}. \tag{7}$$

The average update latency L_{Update} is then given by

$$L_{\text{Update}} = (1 - p)L'_{\text{Region}} + pL'_{\text{Home}}. \tag{8}$$

In the IS-41 scheme, the update message is forwarded to the MH's HLR from the MH's MSC/VLR to update its current location and the ACK message is then sent back to the MH's MSC/VLR. Thus the average update latency for IS-41 scheme is

$$L^{IS-41}_{\text{Update}} = 2L_l + 2L_r + L_{\text{VLR}} + L_{\text{HLR}}. \tag{9}$$

5.3. Numerical Results

The evaluations and comparisons of the query and update delay for the proposed scheme and the conventional IS-41 scheme are presented in this section. We normalize the value of L_l to one for evaluating the latency. The values of the latency ratios of L_l:L_r = 1:5 and L_l:L_{paging} = 1:0.2 are used. Further, we assume the latency ratio of L_{VLR}: L_{HLR} = 1:5.

Figure 9 shows the relation between call to mobility ratio (*CMR*) and the ratio (Q_p/Q_{IS-41}) of query latency for proposed (Q_p) and IS-41 (Q_{IS-41}) schemes. Three data sets are considered when L_l:L_{VLR} = 1:5, 1:10, and 1:15. This figure indicates that by using our proposed scheme the query latency is reduced, if the *CMR* ratio is more than 0.125. Since the region (hit case) queries are quickly processed locally without inquiring the database of the HLR, the average query latency can be decreased. This figure illustrates the higher *CMR* ratio we have, the smaller query latency we get. For example, when *CMR* = 1, the query latency of the proposed scheme is about 30% less than that of the IS-41 scheme.

Figure 10 represents the query latency for the varying hit ratio (h) between the proposed and the IS-41 scheme. The figure shows the linear relation between the query latency ratio and the hit ratio. If the hit ratio is more than 0.2, then the proposed scheme outperforms the IS-41 scheme in query latency.

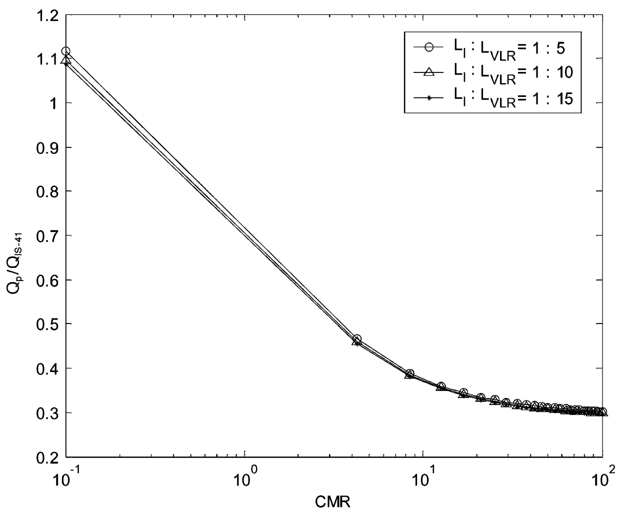


Figure 9. Comparisons of query latency under various *CMR*.

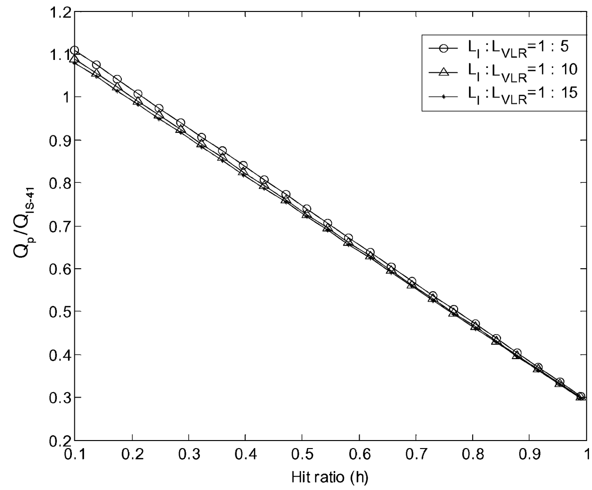


Figure 10. Comparisons of query latency under various hit ratio.

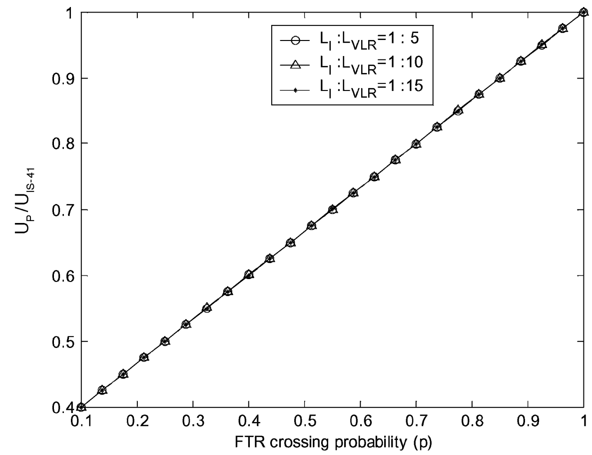


Figure 11. Comparisons of update latency under various FTR crossing probability.

In Figure 11, we show the update latency between our proposed scheme (U_p) and the IS-41 (U_{IS-41}) scheme under the FTR crossing probability (p), where p is assumed to be the probability that the MH moving to another LA out of FTR. If the MH always moves within a particular FTR, then it has a small FTR crossing probability and the remote updates are fewer. In the IS-41 scheme, if the distance between the visited area and the HLR is large, the signaling delay for the location update is long. Since some updates are done locally (i.e., within the FTR) in our scheme, the average update latency of our proposed scheme is smaller than that of IS-41.

Figure 12 shows the comparisons of total latency (i.e., the query latency plus update latency) between our proposed scheme (T_p) and the IS-41 (T_{IS} -41) under various hit ratio (h) and FTR crossing probability (p). It illustrates that the higher hit ratio and the smaller FTR crossing probability we have, the smaller query latency we get. In our scheme, the higher hit ratio and smaller FTR crossing probability mean that more query and update are handled locally within the FTR. Hence, we get the smaller total delay when compare to the IS-41 scheme. Only when the hit ratio is smaller than 0.2, we get the higher total latency than IS-41 scheme.

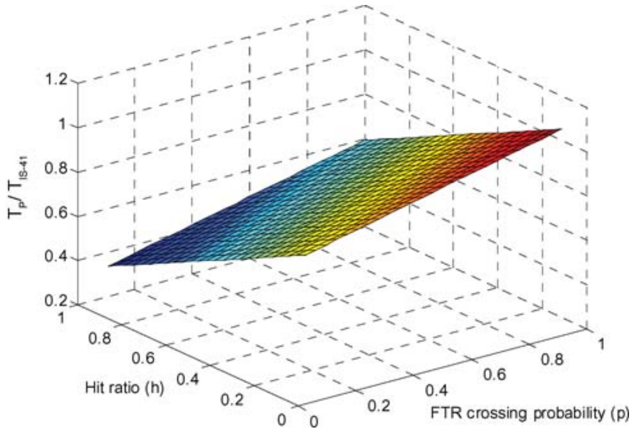


Figure 12. Comparisons of total latency under various hit ratio and FTR crossing probability.

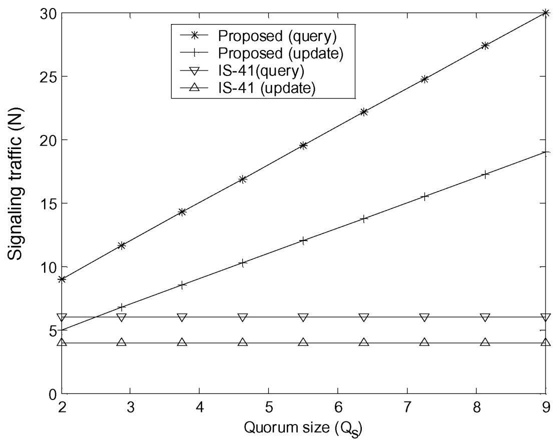


Figure 13. Number of signaling traffic under various average quorum size.

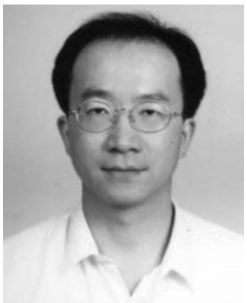
In Figure 13, we show the number of signaling traffic (N) for each query or update between our proposed scheme and the IS-41 under various average quorum size (Q_S). The hit ratio $h = 0.5$ and FTR crossing probability $p = 0.5$ are assumed. Since the IS-41 scheme is not based on the quorum scheme (i.e., without relation to quorum size), it has the constant number of signaling traffic for each query or update. In our proposed scheme, the numbers of signaling traffic of query and update are 3 and 2 times the quorum size, respectively, due to the quorum access. Although the signaling traffic is increased in our quorum-based scheme, the reliability is improved compared to the IS-41 scheme.

6. Conclusions

This paper proposed a new scheme with cellular quorum construction to tolerate the failures of the HLR and VLRs in two-tier networks. Based on the intersectional property of the U-quorum and Q-quorum, the location information is disseminated to VLRs of the U-quorum's set and can be extracted from one of them by using the Q-quorum even though one or more location servers fail. Thus, without adding or changing the hardware of the systems in the two-tier networks, our scheme provides fault tolerance for the system. Meanwhile, with region-based approach, our scheme stores/retrieves the MH location information in the location servers of a quorum set of the local region as much as possible to avoid long delays caused by the possible long-distance of VLR and HLR. Hence, our scheme is not only fault-tolerant but also connection establishment effective.

References

1. EIA/TIA, "Cellular Radio Telecommunications Intersystem Operation", PN-2991, Nov. 1995.
2. M. Mouly and M.B. Pautet, *The GSM System for Mobile Communications*, Palaiseau: France, 1992.
3. G.H. Li, K.Y. Lam, T.W. Kuo, and S.W. Lo, "Location Management in Cellular Mobile Computing Systems with Dynamic Hierarchical Location Databases", *Journal of Systems and Software*, Vol. 69, pp. 159–171, 2004.
4. I.R. Chen and B. Gu, "Quantitative Analysis of a Hybrid Replication with Forwarding Strategy for Efficient and Uniform Location management in Mobile Wireless Networks", *IEEE Transactions on Mobile Computing*, Vol. 2, pp. 3–15, 2003.
5. Z. Mao and C. Douligeris, "Location Tracking in Mobile Networks: A Scheme and an Analysis framework", *Electronics Letters*, Vol. 39, pp. 1406–1408, 2003.
6. Y. Xiao, "Backoff Strategies for Demand Re-registration in PCS Database Failure Recovery", *Computer Communications*, Vol. 27, pp. 400–411, 2004.
7. C. Liu, D. Yu, H. Qiu, and J. Wu, "Research of VLR Mobility Database Failure Recovery and Performance Analysis for CDMA2000 System", in *IEEE Proceedings on Personal, Indoor and Mobile Radio Communications*, pp. 1501–1505, 2003.
8. K. Ratnam, S. Rangarajan, and A.T. Dahbura, "An Efficient Fault-Tolerant Location Management Protocol for Personal Communication Networks", *IEEE Transactions on Vehicular Technology*, Vol. 49, pp. 2359–2369, 2000.
9. M.J. Yang, Y.M. Yeh, and Y.M. Chang, "Legion Structure for Quorum-Based Location Management in Mobile Computing", *Journal of Information Science and Engineering*, Vol. 20, pp. 191–202, 2004.
10. M. Naor and A. Wool, "Access Control and Signatures via Quorum Secret Sharing", *IEEE Transactions on Parallel and Distributed Systems*, Vol. 9, pp. 909–922, 1998.
11. K. Wang and J. Huey, "A Cost Effective Distributed Location Management Strategy for Wireless Networks", *Wireless Networks*, Vol. 5, pp. 287–297, 1999.



Ming-Jeng Yang received the M.S. degree in computer science from the Syracuse University, New York, in 1991, and the Ph.D. degree in computer science from National Taiwan Normal University, Taiwan, in 2004. He is an associate professor in the Department of Information Technology, Takming College, Taiwan. His research interests include wireless networks, mobile computing, fault-tolerant computing, and distributed computing. He is a member of the IEEE Computer Society and the ACM.

Yao-Ming Yeh received the B.S. degree in computer engineering from National Chiao-Tung University, Taiwan, in 1981, and the M.S. degree in computer science and information engineering from National Taiwan University, Taiwan, in 1983. In August 1991, he received the Ph.D. degree in the Department of Electrical and Computer Engineering, The Pennsylvania State University, Pa., U.S.A. He is a professor in the Department of Information and Computer Education, National Taiwan Normal University, Taiwan. His research interests include fault-tolerant computing, web and XML computing, and distributed computing.