A Fault-Tolerant and Region-Based Scheme for Mobility Management

Ming-Jeng Yang, Yao-Ming Yeh
Department of Information & Computer Education
National Taiwan Normal University, Taipei, 106 Taiwan
E-mail: {mjyang, 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 and region-based 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). 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 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.

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. Only a few papers pay attention to fault-tolerance issues on location databases. Fault-tolerant issues are addressed in [6-8]. 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.

Based on the Legion structure that we developed in previous work [9], we propose a fault-tolerant scheme for mobility 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 one or more VLRs. In addition, our scheme has fast query response and can improve the performance of queries 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 connection establishment delay 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 MHs. The information, recorded in the HLR database, can assist the system in finding the MHs. The information, recorded in the HLR database, can assist the system in finding the MHs.

In the following, we introduce a cellular quorum approach that will be used in our system design. 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 Legion is constructed from set systems.

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.

In the following, we introduce a definition of a 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. \( 1 \leq s \leq n, 1 \leq t \leq m \).

Claim 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.

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.

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).

Definition 3. Let \( N \) be the total number of columns in a FTR region and \( M \) be the total number of rows in the column \( i \) of the FTR region. In the cellular

Figure 1. System architecture.

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].

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.

In the following, we introduce a definition of
quorum construction of one FTR region, the Update-set (U-set) and Query-set (Q-set) are defined as follow:

\[
\text{U-set } = \{U_i | 1 \leq i \leq N, U_i = \{(i, j) | 1 \leq j \leq M_i \}\};
\]

\[
\text{Q-set } = \{Q_i | 1 \leq s, Q_i = \{(i, r_i^s) | 1 \leq i \leq N \}\},
\]

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.

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, \(Q_2 = \{(1,1),(2,4),(3,1),(4,2)\}\), is composed from the LAs that are picked from each column (Figure 3(b)). Thus, there is one intersectional element of \((2,4)\) 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 to 4.4.

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 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 VLR of U-quorum in the local FTR. We divide the update procedure into two types: region update and home update. When an MH roams to another FTR, a home update procedure will be triggered. By the user mobility 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.

4.1. Region Update

In the following, we use the cellular quorum approach, defined in Definition 3, to implement the region update algorithms. When a mobile host moves from one LA to others in the same FTR, its location information has to be updated locally (Figure 4). Therefore, the following steps are performed as the region update procedure:

1. The UPDATE message is sent to the current MSC, which forwards this message to all the VLRs in the \(U_i\) quorum of the FTR, where \(i\) is the column that the MH resides currently.
2. Upon receiving the UPDATE message, the VLRs add the new location 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 another $U_j$, where $1 \leq j \leq N$, sends the UPDATE message to all VLRs in the new quorum again, and goes to step 2; otherwise, it goes to next step.

4. The DELETE message is sent to the current MSC which forwards this message to all the VLRs in the $U_i$ 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.

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

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 $\{U_1, U_2, U_3, U_4\} = \{(1,1), (2,1), (2,2), (2,3), (2,4), (3,1), (3,2), (4,1), (4,2), (4,3), (4,4)\}$. If a mobile host $h$ moves from LA(3,1) to LA(2,2), new location information is sent to the 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.

### 4.2. Home Update

When roaming around the whole network, an MH may go through different FTRs. If an MH enters another FTR, a home update procedure will be triggered (Figure 5). The procedure is described as follows:

1. The procedure sends the UPDATE message to the current MSC which forwards this message to all the VLRs in the $U_i$ quorum of the FTR, where $i$ is the column in which the MH resides currently, and concurrently sends the REDIRECT pointer (i.e., a forwarding route information including the FTR identifier) 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 or the ACK message form the HLR during a given period of time, then it randomly selects other quorum $U_j$, where $1 \leq j \leq N$, sends the UPDATE message to all VLRs in the new quorum again or resends the REDIRECT pointer to the HLR, 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.

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 $\{U_1, U_2, U_3, U_4\} = \{(1,1), (2,1), (2,2), (2,3), (2,4), (3,1), (3,2), (4,1), (4,2), (4,3)\}$ and the $U$-set of FTR2 is $\{U_1, U_2, U_3, U_4\} = \{(1,1), (1,2), (2,1), (3,1), (3,2), (3,3), (4,1), (4,2)\}$. 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 $U^1_4$ quorum 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 $U^1_1$ 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. We have designed the query procedure as two types: region query and home query, and describe them in Section 4.3 and 4.4, respectively.

### 4.3. Region Query

A region query occurs when the calling and called MH are in the same FTR. Figure 6 shows the region query processes. The procedure is described in the following steps:

1. The calling MH queries the location information of the called MH form 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 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 $Q_i$ 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.

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 quorum $Q_1$ (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.

### 4.4. Home Query
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 procedure is described as follows:

1. First, the region query procedure (as described in the Section 4.3) is processed in the local FTR in which the calling MH resides. Since the calling and called MH are in different FTRs, the null ACK messages, which issued by all VLRs in the quorum, would be received by the procedure.

2. Then, the HLR is inquired for the location information of the called MH.

3. According to the REDIRECT pointer in the database, the HLR forwards the QUERY message to the VLRs of the randomly selected quorum, $Q^n$, of the FTR in which the called MH resides.

4. 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.

5. 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 4. Otherwise, it stops.

Consider $n$ fault-tolerant regions (FTRs) system with position coordinates $(x_i,y_i)$ shown in Figure 7. According to the Definition 3, the $Q^n$-set of FTR1 may be $\{Q^n_1, Q^n_2, Q^n_3, \ldots\} = \{(1,1),(2,4),(3,1),(4,2)\}, \{(1,1),(2,1),(3,2),(4,2)\}, \{(1,1),(2,2),(3,2),(4,1)\}, \ldots \}$.}

and the $Q^n$-set of FTR$n$ may be $\{Q^n_1, Q^n_2, Q^n_3, \ldots\} = \{(1,3),(2,2)\}, \{(3,2)\}, \{(1,1),(2,1),(3,2)\}, \{(1,3),(2,3)\}, \{(3,3)\}, \ldots \}$. 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 $U^n_1$ of the FTR1. When a calling host $h'$ at LA(2,2) of FTR$n$ 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, $Q^n_1$, and sends a QUERY message to all VLRs of LA(1,1), and LA(3,2) in the quorum $Q^n_1$ (except the local VLR of LA(2,2)). Since the called MH is not located at FTR$n$, all VLRs in $Q^n_1$ 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 $Q^n_1$ of the FTR1 in which the called MH resides. Since $(4,1)$ is the intersectional element of $U^n_1$ and $Q^n_1$, 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.

### 5. Connection Establishment Delay

The evaluations and comparisons of the query 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 $L_l : L_{paging} = 1:5$ and $L_l : L_{paging} = 1:0.2$ are used. Further, we assume the latency ratio of $L_{VLR} : L_{HLR} = 1:5$.

Figure 8 shows the relation between call to mobility ratio (CMR) and the query latency for proposed and 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 more CMR ratio we have, the less 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 9 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.

6. Conclusions

This paper proposed a new scheme with cellular quorum construction to tolerate the failures of the 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


