Mobility Management Incorporating Fuzzy Logic for a Heterogeneous IP Environment

P. M. L. Chan, R. E. Sheriff, and Y. F. Hu, University of Bradford, United Kingdom
P. Conforto and C. Tocci, Alenia Spazio, Italy

ABSTRACT

The next generation in mobility management will enable different mobile networks to interoperate with each other to ensure terminal and personal mobility and global portability of network services. However, in order to ensure global mobility, the deployment and integration of both satellite and terrestrial components are necessary. This article is focused on issues related to mobility management in a future mobile communications system, in a scenario where a multisegment access network is integrated into an IP core network by exploiting the principles of Mobile IP. In particular, attention is given to the requirements for location, address, and handover management. In a heterogeneous environment, the need to perform handover between access networks imposes particular constraints on the type of information available to the terminal and network. In this case, consideration will need to be given to parameters other than radio characteristics, such as achievable quality of service and user preference. This article proposes a new approach to handover management by applying the fuzzy logic concept to a heterogeneous environment. The article concludes with the presentation of mobility management signaling protocols.

INTRODUCTION

Mobility management essentially consists of two main components: location management, including address management, and handover management. This article presents an approach for managing intersegment mobility, concluding with the design of signaling protocols using Mobile IP. Three different wireless access segments are considered, namely General Packet Radio Service (GPRS), Universal Mobile Telecommunications System (UMTS) and a mobile satellite system based on the EuroSkyWay network (M-ESW) [1]. By utilizing Mobile IP for integration, any modification to the existing segment procedures can be minimized, thus making the possibility of providing a truly global access network commercially attractive and technically feasible.

The architecture of the system under consideration is shown in Fig. 1. GPRS involves overlaying a packet-based air interface onto the existing circuit-switched GSM network [2]. GPRS introduces two new network nodes: the serving GPRS support node (SGSN) and the gateway GPRS support node (GGSN). As shown in Fig. 1, the GGSN is physically linked to an edge router (ER) of an Internet subnetwork. A common packet domain core network, consisting of both SGSN and GGSN, is used for both GPRS and UMTS.

Although the concepts described in this article are applicable to any mobile satellite system capable of supporting packet-oriented services, in the following, the emphasis is placed on a geostationary satellite system. The satellite system has multiple points of attachment to the Internet network through fixed earth stations (FESs) connected to the ERs. A network interface unit (NIU) is placed between the FES and the ER, which allows interworking between IP and satellite-specific protocols. The network operation center (NOC) is responsible for performing synchronization, registration, authentication, and connection admission control.

A multimode terminal (MMT) capable of operating in all access segments is also required. The MMT considered in this article comprises terminal equipment (TE), such as a laptop or any other handheld terminal, a terminal interworking unit (T-IWU), and the three segment-specific mobile terminals (SS-MTs): the Sat MT, GPRS mobile station (MS), and UMTS MS.

The T-IWU, as depicted in Fig. 2, is involved in all intersegment mobility procedures and cooper-
ates with the access segment specific mechanisms for the provision of quality of service (QoS) over the multisegment access network. Each functional block groups similar functionality of the T-IWU.

The correct execution of the procedures in the T-IWU is achieved by using suitably devised commands to trigger internal routines and pass necessary parameters. In Fig. 2, these commands are schematically indicated by arrows linking the different modules. Bold arrows are used to represent the flow of IP packets.

The interaction between the T-IWU and the SS-MT is performed through two service access points (SAPs):

- SAP for data transfer: IP packets go through this access point.
- SAP for commands: Specific commands are exchanged through this access point in order to trigger segment-specific procedures.

In addition to the T-IWU, a service node is also envisaged in the system. Both are designed to bridge the gap between the requirements of Mobile IP and the services provided by segment-specific protocols. The service node assists the T-IWU in executing several mobility procedures such as user authentication and registration. In this respect, it acts as an authentication server at the network side. The service node is also responsible for performing any resource management functionalities during handover.

**Mobility Management Procedures**

**Location Management**

Due to the different types of architecture envisaged in the multi-access system under consideration, three levels of location management procedures can be envisaged:

- *Internet network location management*: Identifies the point of access to the Internet network.
- *Intrasegment location management*: Executed by segment-specific procedures when the terminal moves within the same access network.
- *Intersegment location management*: Executed by system-specific entities when the terminal moves from one access network to another.

In Mobile IP [3], each mobile node is assigned a pair of addresses. The first address is used for identification, known as the home IP address, which is defined in the address space of the home subnetwork. The second address is used to determine the current position of the node and is known as the care-of-address (CoA), which is defined in the address space of the visited/foreign subnetwork.

The continuous tracking of the subscriber’s CoA allows the Internet to provide subscribers with roaming services. The location of the subscriber is stored in a database, known as a binding table, in the home agent (HA) and in the corresponding node (CN). By using the binding table, it is possible to route the IP packets toward the Internet point of access to which the subscriber is connected.

The MMT can be seen from the Internet perspective as a mobile node. Once the MMT selects an access segment (either GPRS/UMTS/ M-ESW), the access point to the Internet network is automatically defined. Such an access point is represented by the ER to which the segment is connected (in the case of the satellite segment, the point is represented by the ER connected to the FES). The MMT is therefore identified by a home address of the home subnetwork and by a CoA of the access segment.

In the target system, location management in the Internet network is based on the main fea-
Using certain parameters, including the condition of the radio coverage, and QoS perceived by the user, the MMT continuously executes procedures with the objective of selecting the most suitable access segment. Any modifications to these parameters could therefore lead to a change of access segment.

Figure 2. A T-IWU functional block diagram.

...
ADDRESS MANAGEMENT

Two classes of address can be identified in the system as a result of the network architecture and the mobility management scheme devised for the system. These addresses are defined as follows:

- **Underlying network addresses**: These addresses are valid and recognized by the components of the access segments. Therefore, GPRS, UMTS, and M-ESW specific addresses belong to this class.
- **IP addresses**: IPv6 addresses with global validity.

The design of address management in the system takes into account the presence of both of these addresses and their mutual relationships.

In the satellite segment, a Sat MT is identified by the:

- **Sat_Terminal_Id**: which identifies the satellite terminal within the satellite system.

A Sat MT is capable of serving several different overlaying users, which exploit the satellite connectivity to access the corresponding overlaying network (OLN). In this article, the OLN is represented by the Internet network, whereas the OLN user is analogous to the TCP/IP-based terminal equipment present in the MMT. The OLN user is assigned the following identifier:

- **OLN_User_Id**: Identifies, within the OLN, the OLN user requesting satellite services.

It is used for connection setup and addressing purposes. This corresponds to the IP address of TCP/IP-based terminal equipment in the MMT.

The correspondence between the OLN_User_Id of a given OLN user and the Terminal_Id is stored in a localization table in the NOC, together with the identifier of the spotbeam serving the Sat MT. When a “calling” OLN (i.e., IP) user intends to communicate with a “called” OLN (IP) user in the MMT through the satellite segment, a connection has to be setup between the FES and Sat MT in the terminal. The setup message contains the OLN_User_Id of the called user. Using the localization table, the NOC is able to determine the Terminal_Id of the Sat MT serving the called OLN user and the spotbeam where it is roaming. Consequently, the call can be routed. OLN packet datagram units (PDUs) (IP packets) can therefore be exchanged between the called and calling terminals defined above, and are transmitted over the satellite connection.

The rationale followed in devising a possible mechanism for the IP address assignment is that the satellite network can be configured as an organization able to manage its own address space like each Internet organization.

Handover can normally be separated into three phases: initiation, decision, and execution. This article, the fuzzy logic concept is applied to the initiation and decision phases. During handover initiation, the T-IWU collects information on the user profile, QoS perceived by the user, and radio link availability. Based on this information, together with information on segment availability, a handover can be initiated. In terms of QoS, handover can be initiated when the QoS perceived by the user for a particular application degrades and falls below a specified threshold.

In the handover decision phase, the T-IWU is responsible for selecting the most suitable target segment. The selection of the segment depends mostly on the user profile. A standard user profile contains information such as the minimum and maximum cost and the list of segments with the highest and lowest priority. If the user does not specify any preference, the default segment is assumed to be the terrestrial segment, provided this is available. In this phase, the T-IWU checks the availability of resources in the other segments in order to negotiate the degradation or improvements in the QoS offered to the user.

During handover execution, two main procedures are performed. The first procedure is used to establish IP connectivity to the core network of the access segment. In other words, the terminal is required to obtain an IP address from the selected segment. This is known as

The rationale followed in devising a possible mechanism for the IP address assignment is that the satellite network can be configured as an organization able to manage its own address space like each Internet organization.
PDP context activation in terrestrial mobile systems, such as GPRS. Upon completion of this procedure, standard Mobile IPv6 (MIPv6) procedures are executed to register the new CoA to the HA. At this stage, the terminal will have two active IP addresses, since a mobile node may use more than one CoA at a time [4]. In this case, the terminal will have a primary CoA and a secondary CoA. Once the new CoA has been established, the old CoA should still be used as a secondary CoA for a specified time. This is to reduce packet loss in the system; ensuring handover is as smooth as possible. Consequently, any packets routed to the secondary address before the completion of the handover procedure can be rerouted to the primary CoA once the registration procedure for the primary CoA is established. The old resources and IP address will only be released when the registration is complete.

Application of the Fuzzy-Logic-Based Handover Concept — The proposed solution to initiate handover utilizes the fuzzy logic concept, where a robust mathematical framework for dealing with imprecision and nonstatistical uncertainty is introduced [5]. This is advantageous in the target system because a fuzzy logic system is flexible and capable of operating with imprecise data, and can therefore be used to model nonlinear functions with arbitrary complexity. Furthermore, since most algorithms related to fuzzy logic use a relativity approach, it reduces some problems faced by designers when comparing dissimilar systems such as the different propagation delays and link quality measurements used by the systems. Users are also generally given the option to influence the result of handover by specifying the QoS required, the charging model, and the application used. The block diagram describing the methodology and criteria used in the handover is shown in Fig. 3.

For handover initiation, four different criteria are used: bit error rate; network coverage; perceived QoS; and signal strength (SS). This could be changed depending on the requirements of the system. For handover initiation, the algorithm is separated into three different stages. In the first stage, the parameters of the system are fed into a fuzzifier, which will transform the real-time measurements into fuzzy sets. Fuzzy sets contain elements that have a varying degree of membership in a set. Therefore, it is different from an ordinary or crisp set, where elements will only be considered members of a class if they have full membership in the class. For example, if signal strength is considered in a crisp set, the signal can only be considered to be either strong or weak, not both simultaneously, whereas in a fuzzy set the signal can be classed as quite weak, not so
strong, or medium. This indicates that an element in a fuzzy set can have membership in more than one set. The membership values are obtained by mapping the values obtained for a particular parameter onto a membership function. This function is a curve or line that defines how each data or value is mapped onto a membership value. It is represented graphically in Fig. 4a, where the three lines represent the range available for weak, medium, and strong. SS1, SS2, and SS3 represent the threshold for weak, medium, and strong, respectively. Then, by mapping the position of the received signal strength onto the graph of the membership function, \( \mu \), the signal will be allocated with a membership value in each set ranging from 0 to 1. Therefore, if a signal falls between SS1 and SS2, it could be assigned a value of 0 in the strong set, 0.6 in the medium set, and 0.4 in the weak set. This is shown in Fig. 4a for the membership function related to signal strength. Figures 4b, c, and d refer to the perceived QoS, bit error rate, and network coverage, respectively. These membership values will later be used to obtain the handover initiation factor, which determines whether a handover is initiated.

The second stage involves feeding the fuzzy sets into an inference engine, where a set of fuzzy rules is applied. Fuzzy rules can be defined as a set of possible scenarios utilizing a series of IF-THEN rules, which decides whether handover is necessary. Following this, a set of different handover decisions can be obtained. The decision set could be classed into four different sets: yes (Y), probably yes (PY), probably no (PN), and no (N). An example of the IF-THEN rule is as follows:

IF signal strength is weak, and QoS is undesirable, and bit error rate is high, and network coverage is bad, THEN handover=YES

However, since a measurement usually falls into more than one category, as shown in Fig. 4, more than one decision set can be obtained following the implementation of the IF-THEN rule. For example, there might be a few rules that result in handover \( r = Y \), and a few with PY, N, and PN. At this stage, it is useful to convert the resultant fuzzy decision sets into a precise quantity. This is implemented in the third stage, which is also known as defuzzification. Here, the membership values and decision sets are used to obtain the handover factor using a method known as the centroid method [6]. For further information on other methods of defuzzification, refer to [7, 8].

When considering the handover decision algorithm, inputs from both the system and the user are required. The main purpose of the handover decision algorithm is to select a segment for a particular service that can satisfy the following objectives: low cost, good signal strength, optimum bandwidth, low network latency, high reliability, and long battery life, while taking into account the preferred segment of the user. There are two different stages in the handover decision algorithm. The fuzzy ordinal ranking procedures and the weighting of the criteria is the first stage and could be performed before handover is initiated. The algorithm used is a combination of the model introduced by Yager [9] and the model described in [6]. In addition to the fuzzy sets, there also exists a classical/crisp set, which is used to influence the weighting of the criteria. The crisp set consists of information on segment

---

IEEE Communications Magazine • December 2001

47
to each criterion by performing a pairwise comparison between the designated criterion and the other criteria. For example, if \( C_1 \) is the designated criterion, the set of values of relative importance for \( C_1 \) in comparison with the other criteria can be written as \( (a_{12}, a_{13}, \ldots, a_{1p}) \) where \( a_{ij} \) is determined by comparing the importance of criteria \( C_i \) and \( C_j \). The values of \( a_{ij} \) is determined by the system designer based on the network capability and service requirements, and the user’s preferences. Then a matrix, \( B \), of dimension \( p \times p \) can be created where

1. \( b_{ii} = 1 \)
2. \( b_{ij} = a_{ij} \)
3. \( b_{ij} = 1/b_{ji} \).

An example of the comparison is shown in the following matrix. Here, criterion \( C_1 \) is first compared to \( C_2, C_3 \), and so on. If the value given to a criterion is higher than 1, this indicates a higher importance rating. Using this matrix, the unit eigenvector of the matrix, \( W \), coresponding to the maximum eigenvalue of \( B \) is then determined to obtain the cardinal ratio scale of the criteria. To obtain the final weighting matrix, \( \alpha \), used in the decision process, \( W \) is multiplied with the number of criteria, \( n \):

\[
\begin{bmatrix}
C_1 & C_2 & C_3 & C_4 & C_5 & C_6 & C_7 \\
C_1 & 1 & 3 & 3 & 1/3 & 1/3 & 1/3 & 1 \\
C_2 & 1/3 & 1 & 1/3 & 1/3 & 1/3 & 1/3 & 1 \\
C_3 & 1/3 & 3 & 1 & 1/3 & 1/3 & 1/3 & 1 \\
C_4 & 1/3 & 3 & 3 & 1 & 1/3 & 1/3 & 1 \\
C_5 & 3 & 3 & 3 & 3 & 1 & 1/3 & 1 \\
C_6 & 3 & 3 & 3 & 1 & 1/3 & 1/3 & 1 \\
C_7 & 1 & 1 & 1/3 & 1/3 & 1/3 & 1/3 & 1
\end{bmatrix}

\Rightarrow

\begin{bmatrix}
C_1 & C_2 & C_3 & C_4 & C_5 & C_6 & C_7 \\
0.3284 & 0.1276 & 0.1959 & 0.1959 & 0.7083 & 0.5108 & 0.1911 \\
2.2988 & 0.8932 & 1.3711 & 1.3711 & 49582 & 3.5756 & 1.3375
\end{bmatrix}

The weighting is then applied to the criteria, which have earlier been passed through the ordinal ranking procedures. The chosen segment is the best segment that could satisfy the objectives outlined by the user and the service provider.

**Signalling Protocols**

**System Procedures**

Three classes of procedures are defined for the system: one concerning generic procedures, such as registration; another for functionalities related to location management; and the third for the functionalities related to handover management. It is important to note that the procedures defined here are not segment-specific and are introduced mainly to collect information from all the differ-
ent segments. This information is subsequently used to manage the mobility of the system.

Registration procedures such as obtaining access to the network, service provider selection, and authentication are executed when the terminal is switched on, whereas deregistration is executed when the terminal is switched off and when each segment has to be deregistered.

Five procedures are envisaged for location management: system location update, segment selection, segment reselection, session establishment, and session release. The main purpose of the system location update procedure is to observe the availability of segments as the user roams from one segment to another. Based on this information, the T-IWU can decide whether to select a better segment or initiate a handover.

Segment selection is performed immediately after segment-specific registration and system location update procedures. The criteria and algorithm for segment selection are similar to the algorithm described for the handover decision.

Segment reselection occurs when the MMT is on STANDBY or READY mode and a new segment that was previously not available becomes available, or when a new segment has to be chosen due to changes in the system parameters or user preferences. The segment reselection procedure is also invoked when a segment becomes unavailable and when the registration of a CoA expires.

Session establishment refers to procedures that establish IP connectivity from terminal equipment (TE) to the IP backbone using the selected segment (e.g., PDP context activation in

---

3 The term IP connection establishment is strictly related to the need to set up a wireless link supporting the exchange of IP packets.
When the MMT is in the ACTIVE mode, an inter-segment handover is performed rather the segment reselection procedure when a change of segment is required. This is one of the most important mobility management procedures as it is executed during an active IP session.

LOCATION REGISTRATION

Figure 5 shows the signaling procedures for registration, segment selection, and session establishment. When the MMT is first switched on, segment-specific registration is performed by the SS-MT. At the same time, the T-IWU sends a segment availability request to the SS-MT so that the status of the segments can be probed. If a segment is available, the SS-MT will then send a message back to the T-IWU, confirming the identity and status of the segment. This information is then updated in the T-IWU's database.

Next, the T-IWU has to check the user's preferences and the quality of the radio link in all the available segments before deciding on the most suitable segments. Upon completion of segment selection, the T-IWU will request IP connectivity up to the ERs in order to obtain a CoA. This is then followed by the Mobile IP registration, where the T-IWU sends binding updates to the HA and its CN. This information is then passed to the TE to be stored in a local database.

HANDOVER

Figure 6 shows a proposed approach for the information exchanged among entities for handover. Before handover can be initiated, the T-IWU needs to retrieve certain information on the ACTIVE segment, such as the updated radio link measurements, user profile, and QoS information. Based on this information, together with the information on segment availability, an intersegment handover can be initiated if deemed necessary.

In the handover decision phase, the user profile is checked again with the information obtained from the target segments. At this stage, the T-IWU will try to negotiate the degradation or improvements in the QoS offered to the user. Once the target segment is chosen, a handover execution request is issued.

The handover execution process is similar to the session establishment procedure described for location registration. First, the T-IWU will request IP connectivity in order to obtain a CoA before registering the CoA with the HA and CN. The difference between this procedure and location registration is the fact that for handover, there are two active session establishment procedures during handover execution to ensure that the handover is as smooth as possible. The T-IWU will only release the resources of the old segment after the completion of the Mobile-IP registration in the new segment.

CONCLUSIONS

The future mobile environment will incorporate cellular and satellite networks, operating seamlessly together in order to provide the optimum service to the mobile user. This article has presented a proposal for mobility management in a packet-oriented multisegment system using Mobile IP. The signaling protocols for location, address, and handover management have been presented, demonstrating how the use of Mobile IP concepts can be used to facilitate integration between disparate networks. The introduction of a heterogeneous network architecture with the need to perform handover between networks requires careful consideration of the initiation and decision criteria. It has been demonstrated how the use of fuzzy logic concepts can be used to combine the parameters associated with each network so that a rational approach to handover management can be adopted based on network, radio, and user-defined parameters. Similar concepts can also be applied to initial segment selection upon activating a mobile device.

ACKNOWLEDGMENT

This work has been supported by the European Union in the framework of the IST program (project SUITED).

REFERENCES


BIOGRAPHIES

P. M. L. Chan (p.m.l.chan@bradford.ac.uk) received an electronic systems and information engineering degree from Sheffield Hallam University, United Kingdom, in 1999 following the award of a JPA Scholarship from the Malaysian Government. She is currently in the final year of her Ph.D. at the University of Bradford, United Kingdom, where she is researching the design of satellite network procedures for heterogeneous packet-oriented environments. She participates in the EC's IST SUITED project and was previously involved in the COST 253 Action.

R. E. Shurief (r.e.shurief@bradford.ac.uk) graduated from the University of Leeds, United Kingdom, in 1986, and spent the next four and a half years in the satellite communications industry. In 1991 he took up a lecturership position at the University of Bradford.
with the University of Bradford, where he was awarded a Ph.D. in 1995 and appointed reader in mobile communications in 2000. During his time at Bradford, he has participated in a number of influential satellite communications projects, principally funded under the EC’s R & D programs. He is co-author of the book Mobile Satellite Communications Networks (Wiley). He is an Executive Committee member of the IEE’s Satellite Systems & Applications Professional Network and a U.K. Delegate on the Management Committee of the COST 272 Action.

Y. F. Hu (y.f.hu@bradford.ac.uk) received a 1st Class B.Sc. degree in mathematical sciences and a Ph.D. degree in information systems engineering, both from the University of Bradford. In 1990 she joined the University of Leeds as a research fellow after working two years in the satellite communications industry. In 1994, she was appointed as a lecturer at the University of Bradford. She is now a senior lecturer in the same university. She has been involved in several defense, European Space Agency, and European Union funded projects, including SAINT, SINUS, SUMO, VANTAGE, SECOMS, and SUITED. She is co-author of the book Mobile Satellite Communications Networks (Wiley). She was national delegate to the EU COST 253 and COST 256 Actions. She is now a national delegate to the EU COST 272 Action and a member of the Executive Committee of the IEE Satellite Systems & Applications Professional Network.

P. CONFORTO (p.confarto@roma.alespazio.it) received an M.S. degree in electronic engineering from the University of Rome “La Sapienza” in 1998. That year he received a post-degree grant from the University of Rome “La Sapienza” Dipartimento di Informatica e Sistemistica and worked in the scope of several ACTS projects funded by the European Commission. Since March 1999 he has worked in Alenia Spazio where he is currently involved in the IST SUITED project sponsored by the European Commission. Since March 2000 he is also a Ph.D. candidate at the University of Rome “La Sapienza” Dipartimento di Informatica e Sistemistica. His research interests mainly focus on the integration of satellite systems with wireless and fixed IP-based networks.

C. Tocci (c.tocci@md250b.roma.alespazio.it) received a summa cum laude M.Sc. degree in electronic engineering from the University of Rome “La Sapienza” in 1999, with a final thesis on the integration of ATM-based multimedia access networks, developed in the framework of the ACTS project ACCORD. In 1999 she received a post-degree grant issued by the Dipartimento di Informatica e Sistemistica, University of Rome “La Sapienza,” and continued working on the ACTS project ACCORD in synergy with Alenia Spazio. Since October 1999 she has been with Alenia Spazio and is currently involved in the system design of a global mobile broadband system for the provision of IP-based mobile and QoS-sensitive services developed in the framework of IST project SUITED.

The introduction of a heterogeneous network architecture with the need to perform handover between networks requires careful consideration of the initiation and decision criteria.