QoS Guaranteed Handover Scheme for Global Roaming in Heterogeneous Proxy Mobile IPv6 Networks

Kwangsub Go¹, Misun Kim², Kyujin Lee³ and Youngsong Mun³

¹KREONET Center, KISTI, Seoul, KOREA
²Administrative Management Division, Anti-Corruption and Civil right Commission, Seoul, KOREA
³School of Computing, Soongsil University, Seoul, KOREA

1. Introduction

Traditional Internet service does not consider issues with host mobility and QoS such as transmission delay, packet loss ratio, and bandwidth. However, providing secure and seamless mobility and guaranteed QoS become more critical issues in wireless mobile services. Furthermore, the services that are now being used on mobile networks, including Internet broadcasting, teleconferencing, interactive role playing, and telemedicine, tend to require a predefined QoS.

In the near future, wireless networks are going to fully integrate different wireless access technologies to enable their users to exploit advantages of these various technologies and satisfy the QoS requirements of new applications. Various wireless communication protocols developed for different purposes are integrating and converging to ubiquitous communications, which request QoS-guaranteed, fast, and seamless roaming services. The increasing demands for ubiquitous and mobile computing will require the integration of various wireless access technologies such as WLAN (wireless LAN), 3GPP (3rd generation partnership project), 3GPP2 (3rd generation partnership project 2), and IEEE 802.16. To visualize these access technologies, wireless networks have been converging to the Internet Protocol (IP). For instance, the mobile IPv4 (MIPv4) [1] and mobile IPv6 (MIPv6) [2] protocols were already standardized by IETF (Internet Engineering Task Force). In addition, the IEEE 802.16e amendment [3] enhances IEEE 802.16 with mobility support for users moving at vehicular speeds, and the WiMAX ( Worldwide Interoperability for Microwave Access) Forum has adopted IP mobility [4]. An IP level mobility protocol is needed to enable the users of these future wireless networks to roam freely between various access networks. Global roaming, guaranteed QoS, and vertical handover will be prevalent in the near future.

The Internet protocol IPv4 has been slowly progressing toward IPv6, and during this period both protocols have become a part of the Internet service infrastructure. The existing wired core Internet architecture that is linked to wireless access networks is going to evolve into wireless Internet environments. A large number of different wireless access networks are linked using a multi-hop infrastructure, including notebook PCs, PDAs, and small sensors, and are being served actively. The MIP (mobile Internet protocol) has been developed to provide mobile Internet services based on host mobility. Although it is a stable and mature technology, there are some obstacles to overcome to be widely deployed. It is too heavy to be implemented on small mobile devices and complicated message exchange procedures must be handled by the mobile host itself. As a network-based localized mobility protocol, PMIPv6 has been proposed to overcome the problems with MIP’s host-based mobility and long handover latency [5]. The goal of PMIPv6 is specifying a simple extension of MIPv6 that would support network-based mobility for IPv6 hosts, while reusing the signaling and features of MIPv6. As an alternative to MIP, PMIPv6 reuses the MIPv6 entities and concepts as much as possible, but the mobility management procedures are carried by the network devices. Moreover, the mobility infrastructure in the PMIPv6 domain can provide mobility to an MN operating in IPv4, IPv6, or dual mode, even if the transport network is not an IPv4 or IPv6 network. As a result, a new standard for supporting an IPv4 host and IPv4 network in PMIPv6 was released [6] and a new protocol for supporting MIPv6 hosts attached in PMIPv6 is actively being discussed[7]. PMIPv6 is
handover delay in PMIPv6, but the performance was analyzed separately for each mode [11-12]. To enhance the handover performance, K. Lee [13] proposed a cross-layering mechanism combined with IEEE 802.16e networks. For a more accurate performance evaluation in FMIPv6, S. Ryu et al. proposed combining the two modes and considered the probability of predictive mode failure (PPMF), which is affected by the radius of a cell, velocity of a mobile host, and L2 triggering time [14]. To enhance the handover performance in PMIPv6, they proposed optimizing the authentication delay and predicting the optimized route in [15, 16], respectively. Although PMIPv6 performs better than the typical host-based MIPv6 and its extensions in terms of handover performance [11, 12], it has a long handover delay when an MN moves away from its local area.

Mobile Internet traffic has recently increased at an exponential rate, and the new class of applications increases the need for QoS guarantees, which pose big challenges for the current wireless communication environments. The gap between QoS provisioning and demand has been significantly enlarged. Integrated services (Intserv) [17] and Diffserv [18] were proposed to solve the QoS problem in IP networks [19, 20]. Most QoS-providing models in wired and wireless networks focused on the buffer scheduling mechanisms based on traffic types. Various QoS provisioning schemes have been proposed for mobile IP networks [21-24]. However, these studies did not consider the MN’s movement. In this paper, we propose a new scheme that allows an MN to move wherever it wants while receiving a guaranteed QoS. A priority queue model is used to provide differentiated QoS in our model. Because the BE traffic has the lowest priority, the highest and medium priorities are allocated to the EF class and AF class, respectively. To enable a performance analysis of the priority queue, we use the M/G/1 [25] queuing model.

2. Related Works

In MIPv4, an MN is identified by its home address, regardless of its current point of attachment in the network, and the MN is associated with a care-of-address (CoA) when it is away from home [1]. This triangular routing causes significant delay that degrades the handover performance. It is important for the mobility management to support fast and seamless handover with negligible delay, which enables active services without disruption. Thus, improvements were made and incorporated in a newer version of MIP called mobile IPv6 (MIPv6) to overcome some of the drawbacks [2]. It has been discovered that mobility can be more efficiently handled if the mobility management is divided into global mobility management and localized mobility management. Extensions of MIPv6 such as hierarchical mobile IPv6 (HMIPv6) [8] and fast handover for mobile IPv6 (FMIPv6) [9] have been proposed by IETF for efficient localized mobility management. The main goal of these localized mobility management protocols is to reduce the handover delay by localizing the registration of an MN [8] so that seamless service continuity can be achieved during roaming across wireless networks. Handover latency is mainly the result of delays caused by the discovery, configuration, authentication, and binding update procedures associated with a mobility event. Most of the recently proposed mobility management schemes have been host based, that is, the MN is directly involved in mobility-related signaling. Studies on handover cost evaluation to reduce these delays have been performed in various MIPv6 networks [10-16]. Handover anticipation based on layer 2 (L2) trigger information is used to reduce the registration delay [10-14]. M. Lopez et al. proposed a proactive handover scheme [10] that only considers the predictive mode. Both the predictive and reactive modes are considered to optimize the handover delay in PMIPv6, but the performance was analyzed...
sends a proxy binding update message to register the MN's location to the LMA. After receiving the proxy binding update message, the LMA sends a proxy binding acknowledge message, including the MN's home prefix information, and creates a bi-directional tunnel to the MAG. It also manages a routing table for transferring data to the MN and maintains the MN's reachability. Then, the MN sets up its IP address through a router advertisement procedure to get information about its home network prefix and address configuration method. When an LMA receives external packets from the PMIPv6 domain for the MN, it forwards them through the tunnel to the MAG, which will eventually forward them to the MN.

3.2. QoS Guaranteed Handover Scheme for Global Roaming

PMIPv6 provides mobility without the participation of the host, but its mobility management scope is restricted within an LMA domain. However, the host needs inter-LMA and inter-Internet service provider (ISP) movement, as well as global roaming. In this section, we propose a roaming scheme which includes global roaming. Based on the MN's movement scope, which might be intra-LMA, inter-LMAs, or inter-ISP domains, we analyze the QoS guaranteed handover costs.

Fig. 1 illustrates the proposed network architecture for QoS-guaranteed, fast, global roaming in heterogeneous PMIPv6. In this model, we adopt new entities: a QoS agent (QA) to support differentiated service and global MAG (G-MAG) for fast handover when an MN moves between ISP domains. This scheme could provide QoS guaranteed services depending on a subscriber's service level agreement (SLA) in PMIPv6 networks. More details of the procedures for an MN’s movement and SLA management are provided in [26, 27]. We assume that an ISP network is composed of various access infrastructures according to its QoS level and locality. An ISP has a global QoS agent (GQA or QA) that acts as a global QoS manager, like a bandwidth broker in a DiffServ network. Neighboring GQAs can communicate with each other to establish an inter-domain QoS association such as SLAs. Each subnet of an ISP may have its own local QoS agent (LQA or QA) and HA for QoS and mobility management within its local area. Each LQA can manage the resources within its subnet, and serves the MN with a service profile.

Three MN movement scope cases are shown in Fig. 1. In the first two cases, an MN moves around within the same ISP domain, while it moves into a neighboring ISP domain in the other case. In this figure, (1) indicates intra-LMA movement, which means that an MN attached to MAG1 moves into the new access point (AP) linked to MAG2, where both MAGs are managed by the same LMA1. Movement (2) represents inter-LMA movement. This means that the MN attached to MAG2 moves into MAG3, where the MAGs are managed by different LMAs, but both of them belong to the same ISP domain. Movement (3) shows inter-ISP movement, which means that an MN moves into another ISP domain's access point. This movement is called global roaming, which causes complicated procedures for handover. It causes long delays and QoS degradation. A global MAG (G-MAG) may solve this problem. It is geographically located on the border between ISP domains and used to connect the domains with a security association (SA). Thus, the G-MAG can manage the LMAs, AAAs, and QAs of both domains. In addition, it is able to transfer its profile and authentication information from one LMA to the other LMA during the pre-inter-domain handover. The G-MAG is able to estimate the MN’s location and detect its movement by tracing the access point where the MN is attached. Thus, it can properly predict the point of the inter-domain handover execution. If the MN’s inter-domain handover is imminent, the G-MAG performs a pre-inter-domain handover in advance of the inter-domain handover between the previous MAG and new MAG while the MN is still connected to the G-MAG.

3.3. Performance Analysis

This section describes a QoS-guaranteed handover cost analysis model that depends on an MN’s movement scope. The handover cost of PMIPv6 $C_{HO}$ can be expressed as the sum of movement detection latency $T_{MD}$, proxy binding update latency $T_{PBND}$, and router advertisement latency $T_{RA}$:

$$C_{HO} = T_{MD} + T_{PBND} + T_{RA},$$  

where $T_{PBND}$ can be expressed as the sum of authentication delay $T_{AUTH}$, QoS management delay $T_{QoS}$, and address configuration delay $T_{CONF}$.

$$T_{PBND} = T_{AUTH} + T_{QoS} + T_{CONF}. $$

Fig. 2 illustrates the QoS-guaranteed fast handover cost analysis model based on Fig. 1, which shows all three MN movement scenarios. The notations (1), (2), and (3) represent intra-LMA,
inter-LMA, and inter-ISP movements, respectively.

As shown in (1) of Fig. 1, when an MN moves to MAG2 from MAG1, MAG2 should register the MN to its LMA. Since the handover procedure is performed by the MN’s movement detection, when an MN attaches to a new AP that is covered by a new MAG (MAG2), the AP sends L2 handover messages to the MAG2. This causes MN movement detection delay \( T_{\text{MD}} \). When MAG2 detects the MN’s movement, it sends authentication and QoS profile request messages for the MN to the AAA and QA servers, respectively. After receiving the authentication and QoS acknowledgement messages for the above queries, MAG2 sends a proxy binding update (PBU) message to its LMA. If the LMA can allocate the home network prefix to the MN, it sends a binding acknowledgement (PBA) message to MAG2. MAG2 sends a router advertisement (RA) message to the MN when it receives a PBA message from the LMA. The intra-LMA handover in PMIPv6 is now completed. In this case, both MAGs (MAG1 and MAG2) are in the same region of LMA1, and the MN’s registration procedures follow the conventional PMIPv6 protocol [5]. The QoS-guaranteed handover cost of intra-LMA movement in PMIPv6 network \( T_{\text{P\_HO}} \) can be described in expression (2):

\[
T_{\text{P\_HO}} = T_{\text{P\_MD}} + T_{\text{P\_AUTH}} + T_{\text{P\_CONF}} = 2(t_r + t_a + t_q) + t_{am} + t_{sr},
\]

(2)

An inter-LMA handover cost analysis is shown in (2) of Fig. 2. Since PMIPv6 is designed to manage local area movement, if an MN is away from its current LMA domain, registration procedures for the MN should be carried out through its home network. This causes a long delay and QoS degradation. To reduce the handover delay through the MN’s home network, we propose the use of a pre-handover mechanism before detaching the current MAG (MAG2). When an MN receives a new router advertisement message with a network prefix that is different from its current one, it sends a request to its current MAG (MAG2) to send its profile to the new MAG (MAG3) before detaching from MAG2.

In this message, the MN’s profile is included for movement management without link disruption. Since MAG3 receives the MN’s profile from MAG2, it does not need to request authentication and QoS properties from its home network entities. Thus, we can significantly reduce the handover cost. When an MN moves to a new MAG that belongs to a new LMA domain, the proposed fast handover cost is represented by equation (3):

\[
T_{\text{P\_HO}} = T_{\text{P\_MD}} + T_{\text{P\_PRE\_REG}} = 2(t_r + t_a + t_q + t_{am}) + t_{sr} + 6t_{am},
\]

(3)

where \( t_{am} \) is the delay for the MN’s profile transfer between the two MAGs and pre-handover delay \( T_{\text{P\_PRE\_REG}} \), which occurs before the MN’s movement to support for the MN’s movement between LMAs, can be neglected. An inter-ISP handover cost analysis is shown in (3) of Fig. 2. If it has to be executed in the conventional PMIPv6 schemes, the MN suffers from a long service disruption because of the home registration of the MN. When the MN moves from MAG3 to MAG4, the inter-ISP handover latency becomes too long to support seamless service continuity. To reduce this long inter-ISP handover latency, we propose a new scheme that adopts a network entity called G-MAG, as shown in Fig. 1. The G-MAG is located between ISP domains and is connected to both domains with a security association (SA). Upon predicting an imminent inter-domain handover for the MN, it performs the inter-domain handover while the MN is connected to the previous MAG (MAG3) in the previous domain (ISP A). Since the G-MAG has dual connections between ISP domains, it can gather and manage the LMA information of both domains. It can also transfer the MN’s profile and authentication information from one LMA to the other LMA during the pre-inter-domain handover. The G-MAG is able to estimate the MN’s location and movement direction by tracing the AP where the MN attaches. Consequently, the G-MAG can properly predict the inter-domain handover execution point. Thus, because the new MAG (MAG4) receives the MN’s profile from the G-MAG, it does not need to request authentication and setup QoS properties from its home network.

Therefore, the proposed scheme can reduce the inter-domain handover latency by avoiding the signaling to the MN’s home network. When an MN moves to a new MAG that belongs to a new ISP domain, the handover cost of the proposed scheme is represented by equation (4):

\[
T_{\text{P\_HO}} = T_{\text{P\_MD}} + T_{\text{P\_PRE\_REG}} = 2(t_r + t_a + t_q + t_{am}) + t_{sr} + 6t_{am},
\]

(4)

where \( t_{am} \) is the delay between MAGs that belong to
different ISP domains.

To show our model’s handover cost effectiveness, we present a handover cost analysis model for the conventional PMIPv6 in Fig. 3. It shows a cost evaluation for inter-LMA and inter-ISP handovers in the conventional PMIPv6. As can be seen in this figure, the intra-LMA handover cost is identical to (1) of Fig. 2. Thus, the intra-LMA handover cost in the conventional PMIPv6 is the same as expression (1). However, when an MN moves to LMA2 from its previous LMA (LMA1), the new MAG (MAG3) and LMA2 have no information about the MN if these events occur in conventional PMIPv6 networks. The delay associated with acquiring the MN's profile, authentication, and QoS information through the MN's home network should be taken into account, in addition to the handover costs for inter-LMA and inter-ISP movements in the conventional PMIPv6.

![Fig. 3 Handover cost analysis in conventional PMIPv6.](image)

In Fig. 3, (2) and (3) represent the inter-LMA and inter-ISP handover cost evaluations. When an MN moves into a new LMA or new ISP domain that does not belong to its previous network, there are no benefits to PMIPv6 because it was developed as a local movement management scheme. Thus, all of the mobility procedures have to be carried out like in MIPv6 networks. When a new MAG such as MAG2 or MAG3 detects an MN’s movement, it has to acquire the MN’s profile, authentication, and QoS information through the MN’s home network. After receiving the MN’s profile from its home network, the new MAG sends a PBU message to its LMA (LMA2 or LMA3). After receiving the PBA message from the new LMA or new ISP domain that does not belong to its home network, the new MAG sends an RA message to the MN, and the inter-LMA or inter-ISP handover is completed. The handover costs for these two cases are the same because the handover procedures for inter-LMA and inter-ISP movements are identical. The handover cost is expressed by (5):

\[
\begin{align*}
C_{\text{INTER-LMA,ISP}} &= \tau_{P_{\text{MD}}} + \tau_{P_{\text{MD}}} + \tau_{P_{\text{AUTH}}} + \tau_{P_{\text{CONF}}} + \tau_{P_{\text{RA}}} \\
&= 2(t_a + t_q + t_a + t_q + t_m) + 6t_{ah} + t_{af}.
\end{align*}
\]

\[ (5) \]

where \( t_{ah} \), \( t_a \); and \( t_q \) are the delays for the MN’s registration, authentication, and QoS management through its home network, respectively.

### 4. Cost Evaluation Results

In this section, we compare the handover costs of the conventional scheme and our proposed models in heterogeneous PMIPv6 networks by service types and the MN’s movement scope. The parameters used in this handover cost analysis are defined as follows. Let’s assume that a CN generates data packets destined for an MN at mean rate \( \alpha \) and an MN moves from one subnet to the other at a mean rate \( \beta \). The packet to Mobility Ratio (PMR), \( \rho = \alpha/\beta \), is defined as the mean number of packets received by an MN from a CN per movement. The parameters \( l_1 \) and \( l_2 \) are defined as the average length of a control packet and data packet, respectively. Then, their ratio, \( l \), can be defined as \( l_1/l_2 \). The cost of transmitting a control packet is given by the distance between a sender and a receiver. The cost of transmitting a data packet is 10 times greater than the average cost of processing a control packet at any host (and forwarding data packets at an HA). The average delay time and packet loss probability are very important factors for QoS. In DiffServ networks, ingress edges classify the traffic by SLA. We assume three service levels: EF, AF, and BE. We also assume that an access router accommodates different types of data packets from numerous connections. To serve these three types of traffic, we use 3 buffers in the output buffer module of each node. In this paper, each node has an M/G/1 queuing model for evaluating the performance of prioritized packets. Packets are summed to arrive in the queue according to a Poisson process with mean rates \( \lambda_1, \lambda_2, \) and \( \lambda_3 \) for EF, AF, and BE packets, respectively. The service times for packets from each traffic class follow exponential distributions, with mean rates of \( 1/\mu_1, 1/\mu_2, \) and \( 1/\mu_3 \) for EF, AF, and BE packets, respectively. The mean offered load of the EF, AF, and BE packets in the buffer are \( \rho_1 = \lambda_1/\mu_1, \rho_2 = \lambda_2/\mu_2, \) and \( \rho_3 = \lambda_3/\mu_3 \), respectively. The packet scheduling at the buffer module is as follows. First, the server visits an EF buffer. If packets exist in the EF buffer, it serves them until the buffer is empty. Otherwise, the server visits an AF buffer, and serves the packets in that buffer. After the service is finished for the AF buffer, BE packets are served.

To compare our QoS-guaranteed model with the existing extensions of MIPv6, we use the wired and wireless experimental results in [11, 12]. The values are \( t_{ccw} = t_{um} = t_{ic} = 10 \text{ms}, t_a = t_q = t_{ah} = t_{af} = 20 \text{ms}, t_{ra} = 2 \text{ms}, t_{ang} = 6 \text{ms}, \) and \( t_{aml} = 4 \text{ms}\). The control packet size and data packet size are assumed to be 100 bytes and 1,024 bytes, respectively, and the buffer size, \( K \), is assumed to be 100. The traffic arrival rates are \( \lambda_1 = 0.5, \lambda_2 = 0.3, \) and \( \lambda_3 = 0.2 \) under a work-load of 0.5.

Figs. 4, 5, and 6 show the handover costs according to service type and MN movement scope. Fig. 4 shows the intra-
LMA handover costs of MIPv6 and PMIPv6. In this case, we represent the cost of handover from the conventional PMIPv6 (cPMIPv6) and proposed fast PMIPv6 (fPMIPv6) as PMIPv6 because when an MN moves around in an LMA domain, the handover procedures are identical, and thus the handover costs are the same. As can be seen, the difference between the handover costs of MIPv6 and PMIPv6 is significant because of the MN’s CoA registration in MIPv6 networks. It also shows the service types are a significant element influencing the handover costs. As shown in Fig. 4, the main factor for the handover cost is the service policy. PMIPv6 shows very good performance compared to MIPv6. In the PMIPv6 network, the handover costs for the EF, AF, and BE services are 9.64 times smaller than MIPv6’s for all of the traffic.

The inter-LMA and inter-ISP handover costs in PMIPv6 are shown in Fig. 5 and Fig. 6, respectively. In these figures, we omit the MIPv6 costs because the results are the same as those of Fig. 4. Thus, we compare the results of the conventional PMIPv6 and proposed fast PMIPv6 schemes. When an MN moves between LMAs, the handover costs of fPMIPv6 are significantly lower than those of cPMIPv6. This is because when an MN moves to a different LMA in a cPMIPv6 network, additional handover procedures to acquire the MN’s profile are needed, as in MIPv6. All of the fPMIPv6 service types show better performance than the AF service of cPMIPv6. The handover costs of fMIPv6 are 2.15 times smaller than those of cPMIPv6 for all of the traffic. This shows that our scheme is effective for inter-LMA movement.

Fig. 6 shows the inter-ISP handover costs. In this case, we adopt a G-MAG to eliminate the MN authentication process and QoS setup delay involving its home network. The handover costs are shown for BE, AF, and EF traffic in the conventional and fast PMIPv6. Our scheme reduces the handover cost by 1.65 times compared to cPMIPv6. The results show that fPMIPv6 is effective under global roaming circumstances and could be adopted in future global networks that require an MN that is fast and QoS-guaranteed roaming.
5. Conclusions

Various applications and portable communication devices demand various QoS and mobility levels under different circumstances. To provide a guaranteed QoS and fast mobility, we propose procedures for mobility and QoS management in heterogeneous PMIPv6 networks. When Diffserv is deployed in a mobile IP network, various problems are encountered. The most significant problem is related to the acquisition of the service profile for the MN when it moves into a different ISP domain. Since the first-hop router does not have any information about the newly attached MN, a guaranteed QoS could not be provided. To solve this problem, we present procedures for acquiring the MN’s service profile and additional information based on the MN’s movement scope. Through a guaranteed-QoS handover cost analysis and evaluation, we reduce the handover cost by 9.64 times compare to MIPv6 for intra-LMA movement. Moreover, our scheme shows handover cost efficiencies for inter-LMA and inter-ISP movements that are 2.15 and 1.64 times better than the conventional PMIPv6, respectively. Consequently, our scheme is effective under global roaming circumstances and could be adopted in future global networks that require a fast speed, guaranteed QoS, and roaming. In future work, we will refine the message types and registration procedures to manage the MN’s mobility more effectively and reduce the total communication cost. An authentication mechanism for seamless handover will also be considered to construct a service that is more efficient and secure, with a guaranteed QoS in heterogeneous PMIPv6 networks.

6. References