Multi-level Threshold Handover for Homogeneous Wireless Mobile Networks

Thanachai Thumthawatworn and Anjum Pervez
London South Bank University
Faculty of Engineering, Science and The Built Environment
103 Borough Rd. London SE1 0AA
thumthat@lsbu.ac.uk and perveza@lsbu.ac.uk

Abstract—Handover latency is one of the major problems in wireless mobile networks since it prevents users from seamless mobility. When a mobile node moves from one wireless access point to another in a homogeneous network environment, it has to perform horizontal handover. During the handover procedure the mobile node can neither send nor receive any data packets. This results in packet delay and in some cases packet loss. These two factors severely affect the Quality of Service (QoS), particularly for real-time applications. To address this problem, we have proposed a multi-level threshold handover concept. An algorithm has been developed to verify the proposed concept. The results show that the proposed approach has an improvement (reduction in handover latency) of 67 – 98% (based on different scenarios) when compared with the standard FMIPv6.

I. INTRODUCTION

In homogeneous wireless mobile networks, the handover process occurs when a mobile node moves from one wireless access point (AP) to another wireless AP. During the handover process, the mobile node is unable to send or receive data packets since it is disconnected from the former AP and is trying to be connected to the new AP. The length of period that the mobile node cannot send or receive data packets is called handover latency. The handover process introduces a significant amount of latency, which prevents seamless services to the mobile user, especially for real-time applications such as video conferencing and mobile video streaming.

To deal with the handover latency problem, we have introduced a multi-level threshold handover concept, which enhances the performance of horizontal handover. An algorithm has been developed to evaluate the performance of the proposed concept. The results show that it helps reduce the handover latency up to 75% when compared to Fast Mobile IPv6 (FMIPv6) standard. Additionally, when the existing approach to link layer handover reduction is incorporated into our work, the handover latency can be reduced up to 98%.

The rest of this paper is organised as follows. Section 2 discusses the handover latency problem and the existing handover approaches. In section 3, we describe the concept of multi-level threshold handover and its algorithm. Evaluation and results are shown in section 4. Section 5 concludes this paper.

II. BACKGROUND AND RELATED WORK

In wireless mobile networks, when the mobile node moves from one AP to the other AP, there is a period of time when it cannot send and receive data packets. This is referred to as the handover latency, $T_{HL}$. It can be classified into link layer latency ($T_{L2}$) and network layer latency ($T_{L3}$). So $T_{HL}$ is given by:

$$T_{HL} = T_{L2} + T_{L3}$$

where $T_{L2}$ is the time interval from the point when the mobile node is disconnected from the current AP to the point when the mobile node is connected to the new AP. It consists of three important phases: 1) scanning, 2) authentication, and 3) association. It has found to be as high as 550 ms [1]. $T_{L3}$ is caused by the operations in the network layer, namely, router discovery, address configuration, duplicate address detection (DAD), security session-authentication, authorization and accounting (AAA), QoS provisioning and binding update [2]. $T_{L3}$ depends mainly on the time required by the above operations. Figure 1 illustrates the mobile node’s handover procedure in standard Mobile IPv6 (MIPv6) specification.
The resultant $T_{HL}$ is unacceptably large for real-time applications [3]. To improve this situation, FMIPv6 standard has been specified [4]. In this standard, a number of operations such as a router discovery (an exchange of control messages-RtSolPr and PrRtAdv as shown in figure 2 and described in table 1) and DAD are executed before the mobile node is disconnected from the current AP. The use of L2 trigger allows the mobile node to start the above activities prior to the handover execution. As a result, the handover latency introduced by the network layer operations is reduced. Therefore, the overall handover latency is decreased.

In FMIPv6, $T_{L2}$ does not differ from that of MIPv6. $T_{pre}$ (in figure 2) is the time interval from the point when the mobile node starts to acquire a new IPv6 address to the point when it receives the acknowledgement which contains necessary address information. The control messages (FBU, HI, HAck and FBAck as shown in figure 2) are exchanged prior to L2 link termination. The handover procedure in FMIPv6 could be either in reactive or predictive mode as shown in figure 2 and 3, respectively.

### TABLE 1

<table>
<thead>
<tr>
<th>CONTROL MESSAGES DESCRIPTION</th>
</tr>
</thead>
<tbody>
<tr>
<td>FBU: Fast Binding Update</td>
</tr>
<tr>
<td>HI: Handover Initiation</td>
</tr>
<tr>
<td>RA: Router Advertisement</td>
</tr>
<tr>
<td>BU: Binding Update</td>
</tr>
<tr>
<td>BU_Ack: Binding Update Acknowledgement</td>
</tr>
<tr>
<td>HAck: Handover Acknowledgement</td>
</tr>
<tr>
<td>FBAck: Fast Binding Acknowledgement</td>
</tr>
<tr>
<td>PrRtAdv: Proxy Router Advertisement</td>
</tr>
<tr>
<td>RtSolPr: Router Solicitation for Proxy</td>
</tr>
</tbody>
</table>

A number of ideas concerned with the horizontal handover have been proposed in the literature [5, 6]. These are based on the comparison between the received signal strength (RSS) of the current AP and that of the candidate APs.

In general, the horizontal handover is initiated when the RSS of the candidate AP is higher than that of the current AP. In [7], dwell timer has been incorporated into the RSS-based handover decision strategy. The dwell timer allows the mobile node to stay with the current AP as long as possible resulting in efficient utilisation of WLAN bandwidth. In [8], the dwell timer can be adapted based on the available data rate of the current AP and candidate APs. There are further variants of RSS as described in [9].

### III. MULTI-LEVEL THRESHOLD CONCEPT

We have developed a new approach to enhance the performance of horizontal handover, which is based on a multi-level threshold concept.

The setting of the threshold level is a function of RSS. This implies that the handover procedure may be completed in $n$ stages, each stage associated with one of the $n$ threshold levels, as shown in figure 4. In this situation the handover process may be terminated at any stage if the RSS value remains constant or begins to increase.

A three-level threshold handover approach has been developed to deal with the handover latency problem. The idea is to allow certain network layer operations to be executed at each threshold level.
The control messages of FMIPv6 specification are divided into two message groups as follows:

- Group 1 contains RtSolPr and PrRtAdv messages
- Group 2 contains FBU, HI, Hack, 2 of FBAck messages

These two message groups will be transmitted separately when the mobile node moves across different threshold levels as explained below:

The mobile node maintains its RSS measurement (every 100 ms) whilst on the move. If the RSS remains higher than the first threshold level ($Th_{L1}$), the mobile node does nothing, but keeps measuring the RSS. If the RSS is found to be lower than $Th_{L1}$, the mobile node has to determine availability of candidate APs by sending probe requests (scanning phase). If there is no response from any APs, it means that a horizontal handover cannot be realised and the mobile node will be out of service as it moves beyond its current AP’s range.

If there is a response from any one of the APs, it sends group 1 messages and performs DAD. If the mobile node later on crosses $Th_{L2}$, it performs network layer (L3) security session. When it crosses $Th_{L3}$, it sends group 2 messages and carries out QoS provisioning. The mobile node is then disconnected from the current AP. At this stage, the mobile node cannot send and/or receive data packets until it is connected to the candidate (new) AP. Thus, by activating the network layer operations in advance, the handover latency is minimised. Three-level threshold handover procedure is shown in figure 5. An algorithm has been developed to evaluate the performance of the proposed handover procedure.

IV. EVALUATION

Three-level threshold handover is evaluated using the algorithm in figure 6 and assuming the test network scenario as shown in figure 7. There are two network domains, namely, domain A and domain B. Each domain has its own AAA server (for security sessions between the two domains).
In this scenario, the mobile node is moving out of the range of AP1 and moving into the service area of AP2. At some point, the handover will be initiated.

Based on the FMIPv6 standard, the handover latency ($T_{HL_{FMIPv6}}$) is given by:

$$T_{HL_{FMIPv6}} = T_{link-layer} + T_{network-layer}$$

where $T_{link-layer}$ is the link layer latency and is a constant. $T_{network-layer}$ is the network layer latency which includes processing time of DAD ($T_{DAD}$), security session ($T_{security}$) and QoS provisioning ($T_{QoS}$). The $T_{network-layer}$ is given by:

$$T_{network-layer} = T_{DAD} + T_{security} + T_{QoS}$$

where $T_{DAD}$ is a constant, and $T_{security}$ and $T_{QoS}$ are given by the following expressions:

$$T_{security} = RTT_{MN-AR} + 2RTT_{server-AR} + \sum_{i=1}^{3} (RTT_{AR}) + \sum_{i=1}^{3} (P_{server}) + \sum_{i=1}^{3} (P_{router})$$

and

$$T_{QoS} = RTT_{MN-AR} + RTT_{CN-AR} + \sum_{i=1}^{3} (RTT_{AR}) + \sum_{i=1}^{3} (P_{router})$$

where

- $RTT_{MN-AR}$ is an average round-trip time between MN and AR.
- $RTT_{server-AR}$ is an average round-trip time between AAA server and AR.
- $RTT_{CN-AR}$ is an average round-trip time between CN and AR.
- $RTT_{AR}$ is an average round-trip time between ARs.
- $P_{server}$ and $P_{router}$ are the processing delay of the AAA server and AR, respectively.

We have evaluated four cases based on the three-level threshold handover approach as follows.

Case 1: If DAD, security session and QoS provisioning have been completely executed prior to the handover is initiated, the total handover latency of the three-level threshold handover ($T_{HL_{SLTH}}$) is

$$T_{HL_{SLTH}} = T_{link-layer}$$

Case 2: If the mobile node cannot complete QoS provisioning before the handover is initiated (the velocity of the mobile node is too high), QoS provisioning has to be executed again after the mobile node is connected to the new AP. Then, $T_{HL_{SLTH}}$ is

$$T_{HL_{SLTH}} = T_{link-layer} + T_{QoS}$$

Measurements in [2] revealed that the mobile node needs as high as 90% of the link layer latency for scanning channels. A number of ideas have been proposed in the literature [10, 11, 12] to further reduce the link layer latency. They try to reduce the time spent in the scanning phase. Here, we have incorporated the above idea into our procedure.

Case 3: The approach in [12] is incorporated in the procedure of case 1. In this model, $T_{HL_{SLTH}}$ is

$$T_{HL_{SLTH}} = T_{link-layer-reduced}$$

Case 4: The approach in [12] is incorporated in the procedure of case 2. In this model, $T_{HL_{SLTH}}$ is

$$T_{HL_{SLTH}} = T_{link-layer-reduced} + T_{QoS}$$

where $T_{link-layer-reduced}$ is a link layer latency when the approach in [12] is incorporated into three-level threshold handover procedure.

$T_{HL_{FMIPv6}}$ and $T_{HL_{SLTH}}$ (all cases) are evaluated based on the following parameter values:

- $T_{link-layer}$ is 550 ms [2].
- $T_{link-layer-reduced}$ is 50 ms [12]
- $RTT_{MN-AR}$, $RTT_{server-AR}$, $RTT_{CN-AR}$ and $RTT_{AR}$ are 30 ms, 30 ms, 100 ms and 100 ms, respectively [13].
- $T_{DAD}$ is 1000 ms [14].
- $P_{server}$ and $P_{router}$ are negligible [15].

A comparison between the FMIPv6 standard and the three-level threshold handover (the four cases) is shown in figure 8.

![Performance comparison between FMIPv6 and three-level threshold handover](image-url)
From the results it is clear that the proposed three-level threshold handover procedure has a significant advantage over the standard FMIPv6 when considering the overall handover latency.

The FMIPv6 requires 2,250 ms to complete the handover process. In comparison, the three-level threshold handover requires 550 ms and 240 ms, in case 1 and case 2 respectively, giving a reduction in handover latency of 75.5% and 67% corresponding to the two cases.

In the above two cases, the link layer latency is the main contributor to the overall handover latency. The results in case 3 and 4 suggest that further improvements are possible if the procedure proposed in [12] is also included in the three-level threshold handover. The resultant latency values are 50 ms and 240 ms, for case 3 and 4 respectively, giving a reduction in handover latency of 98% and 89% corresponding to these two cases.

It is well understood that the handover latency causes packet delay, which in turn gives rise to packet loss. During the handover process, the mobile node cannot receive data packets transmitted by the correspondent node. Those packets are dropped at the router, to which the mobile node is communicating with, since they cannot be forwarded to the mobile node (if the router’s buffer is full). The lost packets will have to be re-transmitted after the mobile node is connected to the new AP. This produces an unacceptable level of packet delay, particularly for real time applications, which results in degradation of QoS.

Clearly, our approach has the potential to reduce the overall handover latency compared with the FMIPv6 standard, and hence can provide an improved QoS.

V. CONCLUSION

In this paper, a new approach to enhance the performance of horizontal handover has been developed. The approach is based on a multi-level threshold handover decision concept. A three-level threshold decision model has been presented and an algorithm has been developed to evaluate the performance of the proposed concept. It allows network layer operations such as DAD, security session and QoS provisioning to be executed prior to the termination of the link layer. The approach has been evaluated and results have been compared with the three-level threshold handover show that it has an advantage of 67% - 75.55% in terms of handover latency reduction compared with the standard FMIPv6. If the existing approach to reduce link layer latency is implemented, the total handover latency can be reduced up to 98% when compared to the standard FMIPv6. A reduced latency greatly minimises packet loss and hence improves QoS, especially for real-time applications.

An extension of this work would be to introduce a dynamically varying multi-level threshold policy. In such a policy the number of threshold may depend on the network status at any given time.

REFERENCES