FAST INTER-AP HANDOFF USING PREDICTIVE AUTHENTICATION SCHEME IN A PUBLIC WIRELESS LAN

SANGHEON PACK AND YANGHEE CHOI

School of Computer Science and Engineering, Seoul National University, Seoul, Korea
E-mail: shpack@mmlab.snu.ac.kr and yhchoi@snu.ac.kr

Recently, wireless LAN systems have been widely deployed for public mobile Internet services. Public wireless LAN system can provide high speed Internet connectivity using portable devices such as laptop computers, Personal Digital Assistants (PDA), etc. In the public wireless LAN systems, reliable user authentication and mobility support are essential issues. However, re-authentication during handoff procedures causes long handoff latency and this affects the quality of service in real-time multimedia applications. In this paper, we proposed a fast Inter-AP handoff scheme based on a predictive authentication method. In our scheme, a mobile host entering the area covered by an AP, performs authentication procedures for multiple APs, rather than just the current AP. These multiple APs are selected using a Frequent Handoff Region (FHR) selection algorithm, which takes into account users' mobility patterns, service classes, etc. Since a mobile host is registered and authenticated for an FHR in advance, handoff latency resulting from re-authentication can be minimized. Simulation results show that the proposed scheme is more efficient than other schemes in terms of handoff delay and buffer requirements.

1 Introduction

Public wireless LAN systems based on IEEE 802.11 are becoming popular in hot spot areas such as convention centers, airports, shopping malls, and so on. Unlike existing wireless Internet services, public wireless LAN system can provide high-speed Internet connectivity of up to 11Mbps using portable devices such as laptop computers and Personal Digital Assistances (PDA). Originally, wireless LAN was only designed for indoor network solutions, so that deploying wireless LAN in public areas requires a system of access control for unauthorized users. In addition, public wireless LAN systems should support different user mobility patterns. Therefore, user authentication and handoff support between Access Points (AP) are two of the most important issues to consider for public wireless LAN systems. Generally, since user authentication should be performed at each AP, when a mobile host (MH) moves into the area covered by a new AP, it should perform a new user authentication procedure and receive a new Wired Equivalent Privacy (WEP) key, which encrypts the transmitted data in the wireless link.

These authentication mechanisms impact network and device performance, quality of service, etc. Because mobile hosts need to be authenticated during and after handoff, the authentication mechanisms need to be responsive to the handoff time-scale required in micro-mobility environments [3]. Furthermore, since Authentication, Authorizing, and Accounting (AAA) servers are located at locations

far away from the AP, the handoff scheme based on the current AAA protocols cannot meet certain requirements for real-time multimedia application.

Recently, several Technical Groups (TG) within IEEE, specifically TGi and TGf, are focusing on the standardization of a secure authentication scheme and an inter-AP handoff mechanism. In terms of security, it is suggested that the authentication scheme based on IEEE 802.1x standard be used [1]. IEEE 802.1x is a network-to-client authentication mechanism utilizing EAP (Extensible Authentication Protocol) [2] as the encapsulation protocol for upper-layer authentication information. In IEEE 802.1x, a separate AAA server performs user authentications. In addition, this protocol supports secure and dynamic key management. For the inter-AP handoff, Inter-Access Point Protocol (IAPP) is proposed. Since IAPP specifies the information to be exchanged between APs located within the same subnet, it can be used for performing context transfer during handoff. However, in IEEE 802.1x and IAPP, handoff latency caused by message delivery and server processing for re-authentication is unavoidable.

In this paper, we propose a fast handoff scheme that minimizes the re-authentication latency in public wireless LANs. When a mobile host sends an authentication request, the AAA server authenticates not only the currently used AP, but also multiple other APs, and sends multiple WEP keys to the mobile host. These multiple APs are selected by the Frequent Handoff Region (FHR) selection algorithm. This centralized algorithm is based on traffic pattern and user mobility characteristics within the public wireless LAN. The rest of this article is organized as follows. Section II describes related fast handoff schemes. In Section III, we propose a fast handoff scheme using FHR selection. Section IV shows the simulation results. Section V concludes this paper.

2 Related Work

Currently, handoff between two APs belonging to the same subnet is partially supported. However, existing schemes are not suitable for meeting requirements of real-time multimedia applications because of its long delay in handling handoff.

To overcome these drawbacks, Koodli and Perkins proposed a fast handover scheme based on context transfer [4]. They proposed this fast handover to reduce connectivity latency and reception latency. The key idea behind minimizing the connectivity latency is to allow a mobile host to configure a new IP address even before it connects to its new access router (AR). The process of improving connectivity latency includes sending messages to indicate handover to the mobile host, allowing it to form a new IP address, and negotiating between the ARs to support this new IP address. In the process, the access routers also set up a suitable forwarding path for the packets destined for the MN's previous IP address. In this scheme, the forwarding path from the previous AR to the new AR must not be enabled until the mobile host explicitly authorizes the previous AR to do so.

Therefore, the mobile host sends this indication using a “Fast” Mobile IPv6 Binding Update message to the previous AR only after receiving this information. And it starts forwarding the packets on the tunnel established earlier using the HI and Hack message sequence. Although the fast handover scheme described in [4] can reduce connectivity and reception latencies, it requires link layer triggers to support the functions. Namely, it is dependent on link layer technologies.

In additions, Choyi et al proposed a fast handoff scheme in wireless LANs for real-time applications [5]. The scheme uses the Explicit Multicast (XCAST) needing a layer-2 re-association prior trigger. XCAST is a new packet transmission technique. It has no unique multicast group address and just uses the unicast address of the group members to route packets. Namely, the XCAST source inserts unicast addresses into the XCAST header and then the packets are delivered to the specified hosts. In XCAST, XCAST capable routers process XCAST headers, and duplicate or just relay packets, depending upon the routing table entries. In this scheme, the layer 2 trigger takes place before re-association. At first, either the old AP or the mobile host initiates an XCAST join message with the new AP's IP address. Once re-association is performed, the new AP sends a re-association complete message to the XCAST capable access router and the XCAST router removes the old AP's IP address from the MN's XCAST entry. During handoff, the XCAST capable access router duplicates packets and sends them over the distributed system to both of the APs listed the XCAST table. Unlike regular multicast, XCAST does not need to keep state information for each mobile host at each XCAST router. Also, duplication of packets takes place only during handoff, so once handoff is complete it functions in the same way as regular unicast. This scheme based on XCAST can provide fast and smooth handoffs, but it does not consider any user authentication and security issues. Therefore, it is not a suitable solution for public wireless LAN requiring secure user authentication mechanisms. Furthermore, this solution requires XCAST capable access routers.

3 Fast Inter-AP Handoff

In this paper, we proposed the fast inter-AP handoff scheme. The objective of this scheme is to minimize the handoff latency caused by authentication procedures at the new AP. In the proposed scheme, a mobile host performs authentication procedures not only for the current AP but also for neighboring APs, when initial registration is required. The key issue in this scheme is how to select the neighboring APs to be authenticated in advance. The simplest method is to select all of the APs adjacent to the current AP. However, this is inefficient because it doesn't take into account any movement patterns or the APs' geographical location. Since wireless LAN is a solution usually employed within an administrative network domain, such as for a campus network, centralized AP configuration and management are possible, whereas they are not possible in the case of cellular networks. Therefore,

we present a centralized neighbor selection algorithm, called the Frequent Handoff Region (FHR) selection algorithm.

3.1 Frequent Handoff Region (FHR) Selection

The Frequent Handoff Region (FHR) is a set of adjacent APs. It is determined by factors such as the APs' location in a wireless LAN service area and users' movement pattern. Namely, the FHR is comprised of APs which mobile hosts are likely to move to in the near future. Although there are a lot of APs in a public wireless LAN, the movement ratio between different APs are not the same. For example, if two APs are installed within the same large conference room, users may move from one AP to another AP quite frequently. However, if some obstacles (e.g. thick walls or lakes) are situated between APs, users will seldom move along that path. Since wireless LAN is a solution for an administrative network domain, this information can be collected and processed by network administrators in a centralized fashion.

To utilize the collected information in the neighbor selection algorithm, there are some considerations. Firstly, as mentioned above, the geographical location of the APs is very important. Although two APs may be adjacent in a distributed system (e.g. Ethernet), if they are installed on different floors and the user cannot easily move from one floor to the other, then they cannot be considered to be neighboring APs. Secondly, the users' service class and moving pattern are also important factors. Some users may be satisfied with their service in spite of session disconnection occurring during handoff. They are willing to re-connect to the wireless LAN system by means of new initiation procedures. However, other users may want seamless connectivity without data loss during handoff. Moreover, some users may move from one AP to another more frequently than others. Furthermore some users move around faster than others. To support those users, whose movement frequency and moving velocity are high, a higher number of neighboring APs should be pre-authenticated. In order to take these factors into account, we propose the following FHR selection algorithm.

The proposed algorithm consists of two parts. The first step is to draw a weighted bi-directional graph of AP placement. In the graph, each edge has a weight denoted by $w(i, j)$. Eq. (1) shows the weight value between AP(i) and AP(j).

$$w(i,j) = \begin{cases} 0 & (i = j) \\ A & (i \neq j,\ AP(i)\ and\ AP(j)\ are\ adjacent) \\ \infty & (AP(i)\ and\ AP(j)\ are\ \text{not}\ adjacent) \end{cases} \quad \textbf{(1)}$$

Since traffic patterns are asymmetric in the real world, $w(i, j)$ and $w(j, i)$ are generally not equal. The weight value, A, is determined by traffic information and

mobility pattern and is assigned in each link separately based on (*Theorem 1),* by the centralized system. Generally, A is reversely proportional to the handoff ratio in a path from AP(i) to AP(j) in the observation period [9] .

(Theorem 1) If the weight in a path from AP(i) to AP(j) is higher than that of the path from AP(m) and AP(n), then $w(i, j)$ *should be less than* $w(m, n)$.

Using Eq. 1, we obtain an N by N weight matrix, *W.* (N denotes the number of APs) Then, the next step is to select frequent handoff region (FHR) using a weight bound, *D* and a maximum hop count, *H*. The value, *D* represents the corresponding user's service class level. In other words, if the value of *D* is high, the user will authenticate more neighboring APs for a better seamless service. The value, H determines the handoff scope of a mobile host. In most cases, handoffs occur between two adjacent APs, but, if a mobile host moves with very high speed, we should consider handoff possibilities between two non-adjacent APs. Namely, H provides for the possibility of multi-hop handoff.

Below we show the *SelectFHR* procedure in the case where H is 2. *M* denotes a column vector for results of *SelectFHR*. Each element, $m(j)$ of *M*, is initialized as *False*. If the sum of the weights of AP(j) is less than or equal to the weight bound (D), $m(j)$ is set as True and AP(j) is selected as FHR. This nested loop is repeated according to the hop count, H.

```
SelectFHR(W : Weight Matrix, N : Number of APs,
i : Index of current AP, D : Weight bound value, M : FHR Matrix)
1    for j = 1 to N
2        m(j) = False;
3    end
4    for j = 1 to N
5      if ( w(i, j) ≤ D )
6          then m(j) = True;
7          for k = 1 to N
8            if ( w(j, k) ≤ D - w(i, j) ) then m(k) = True;
9            end if
10       end
11     end if
12   end
```

The following figures show an example of the FHR selection procedure. The marked point in the Fig. 1 means the current AP where a mobile host is located.

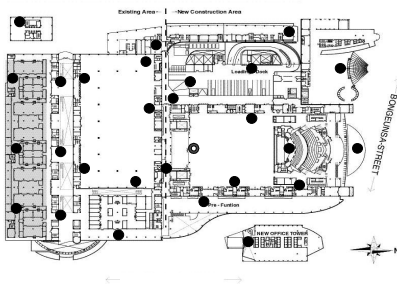

Figure 1. Initial AP Placement

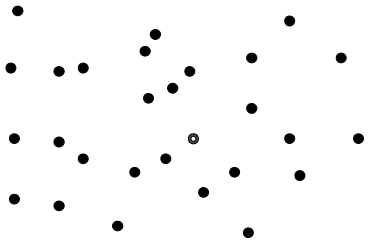

Figure 2. AP Placement

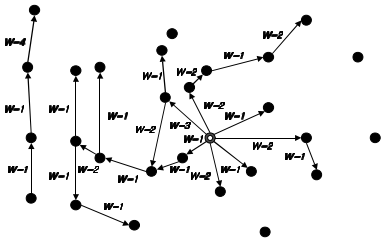

Figure 3. Weighted Directed Graph

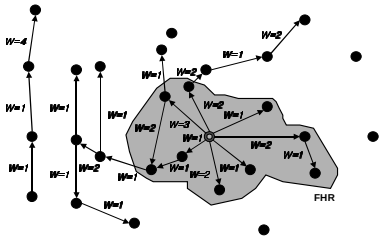

Figure 4. FHR Selection

3.2 Key Management

After selecting the FHR, a mobile host requests the authentication procedures. Firstly, the mobile host sends an "Access Request" message to the AAA server. Specific message formats are dependent on the AAA protocols being used, such as RADIUS, Diameter, etc. The AAA server receiving an "Access Request" message performs an authentication procedure on the user and sends an "Access Reply" message to the corresponding host. The "Access Reply" message contains several pieces of information: i.e. the session key, WEP key, etc.

In our scheme, the key information should be distributed to all corresponding APs located in the FHR of the mobile host being authenticated. Since the current standard supports only one-to-one key distribution, a modified key management system is required. Fig. 5 shows the initial authentication procedure in the proposed scheme. Although a mobile host sends only one access request through the current AP, the AAA server sends multiple authentication responses to all APs belonging to the FHR. Although multicast can be used for multiple key distributions, since the overhead resulting from the dynamic selection of the FHR is high, unicast is more

suitable. After receiving a response, the APs, with the exception of the current AP, maintain this authentication information in the *soft state*. Namely, if there are no handoff events within a specific time period, the key information is deleted. Unlike other authentication responses, the response relayed to the mobile host through to the current AP should contain a variety of information such as the session key, multiple WEP keys, etc.

Fig. 6 shows the case of re-authentication after handoff. When the mobile host hands off to any other AP, since the new AP receives session information in advance, further message exchanges are not needed. The relocated mobile host can obtain all information from the new AP and it is not necessary to send an "Access Request" request to the AAA server. Generally, since the AAA server is often located in a remote domain for more scalable service, the delay in the path from the AP to the AAA server is a critical factor in the total handoff latency. All of these functions can be implemented by using various attributes available in the current AAA protocol.

To support the proposed scheme, each AP should maintain session information in *soft state*. To this end, IEEE 802.1x [1] can be utilized. IEEE 802.1x enables authenticated access to IEEE 802 media, including Ethernet, Token Ring, and 802.11 wireless LANs. IEEE 802.1x provides a network port access control scheme. In IEEE 802.1x, there are two types of network ports: the controlled port and the uncontrolled port. The uncontrolled port is used for transmission of the "Access-Request" and "Access-Reply" messages. On the other hand, the controlled port is used for date transmission. A mobile host can obtain access to the controlled port only after performing the user authentication and receiving session and WEP keys. In our scheme, since an authenticated mobile host possesses multiple session and WEP keys, corresponding to the different APs included in the FHR, and each AP has already received the result of authentication from the AAA server, the mobile host can obtain access to the controlled port for data transmission without further re-authentication.

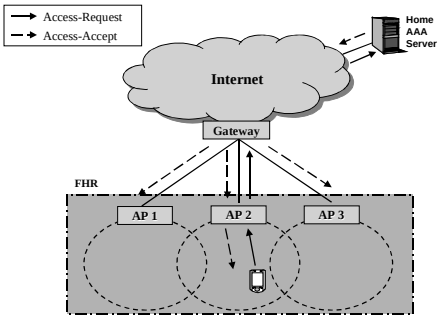

Figure 5. Initial Authentication

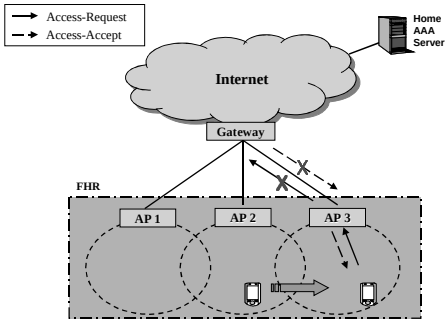

Figure 6. Re-Authentication after Handoff

4 Performance Evaluation

4.1 Simulation Environment

To evaluate the performance of the proposed fast handoff scheme, we used the simulation environment shown in Fig. 7. This figure shows the weight values for each link. In this environment, AP(4) is the current AP receiving an authentication request from a mobile host. Fig. 8 represents the weighted graph in the form of a matrix used in a *SelectionFHR* procedure. The matrices shown below present the results of the *SelectionFHR* procedure in the cases where D=1, D=2, and, D=3 respectively. (1: True, 0: False)

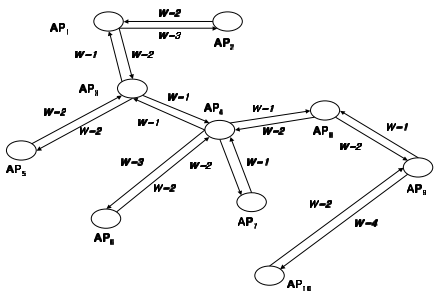

Figure 7. Simulation Environment

$$W=\begin{pmatrix} 0 & 3 & 2 & \infty & \infty & \infty & \infty & \infty & \infty & \infty \\ 2 & 0 & \infty & \infty & \infty & \infty & \infty & \infty & \infty & \infty \\ 1 & \infty & 0 & 1 & 2 & \infty & \infty & \infty & \infty & \infty \\ \infty & \infty & 1 & 0 & \infty & 3 & 2 & 1 & \infty & \infty \\ \infty & \infty & 2 & \infty & 0 & \infty & \infty & \infty & \infty & \infty \\ \infty & \infty & \infty & 2 & \infty & 0 & \infty & \infty & \infty & \infty \\ \infty & \infty & \infty & 1 & \infty & \infty & 0 & \infty & \infty & \infty \\ \infty & \infty & \infty & 2 & \infty & \infty & \infty & 0 & 2 & \infty \\ \infty & \infty & \infty & \infty & \infty & \infty & \infty & 2 & 0 & 4 \\ \infty & \infty & \infty & \infty & \infty & \infty & \infty & \infty & 2 & 0 \end{pmatrix}$$

Figure 8. Weighted Graph

M(1) = [0011000100] *M(2)* = [1011001100] *M(3)* = [1011111110]

4.2 Mobility Model

Generally, the random walk model is widely used as the micro-mobility model [8]. In the random walk model, a user moves in one direction in a random manner. However, since we assumed that the mobility weights for each path could be determined by a centralized method, we used the independent and identically distributed (*i.i.d.*) mobility model. In this model, time is divided into fixed intervals and a mobile user can make at most one move during one interval. If a user is in cell i at the beginning of a time slot, then during this slot he moves to cell i +1 with probability p, moves to cell i −1 with probability q, or remains in cell i with probability 1-p-q, independent of his movements in other time slots. Each transition probability, p and q, can be found based on Eq. (1). However, we did not assign the weight value in the case that a user stays at the same AP in the next time slot, so that

we used the stability factor, α. If $\alpha = 0$, the mobile host hands off to another AP with probability 1. On the other hand, if $\alpha = \infty$, the mobile host stays at the current AP with probability 1. $P(i, j)$ is the transition probability from AP(i) to AP(j) and G is the normalization constant.

$$P(i, j) = \begin{cases} \frac{1}{G} \cdot \frac{1}{w(i, j)} & (i \neq j) \\ \frac{1}{G} \cdot \alpha & (i = j) \end{cases} \qquad \textbf{(2)}$$

$$G = \sum_{i \neq j} \frac{1}{w(i, j)} + \alpha$$

4.3 Simulation Result

In this section, we compare the handoff latency in both the proposed fast handoff scheme and the general handoff scheme. The total latency during handoff can be measured by the summation of the latencies in both the wireless link and the wired link (Eq. (3)). Each latency in the wireless and wired links is proportional to the number of exchanged authentication messages. Each latency can be measured using general network tool such as "ping" and "traceroute".

$$L_{Total} = L_{Wireless} + L_{Wired} \qquad \textbf{(3)}$$

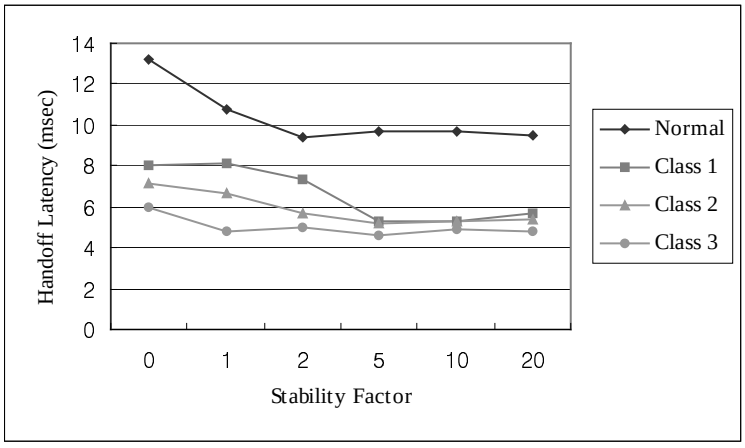

Figure 9. Handoff Latency (Local AAA Server)

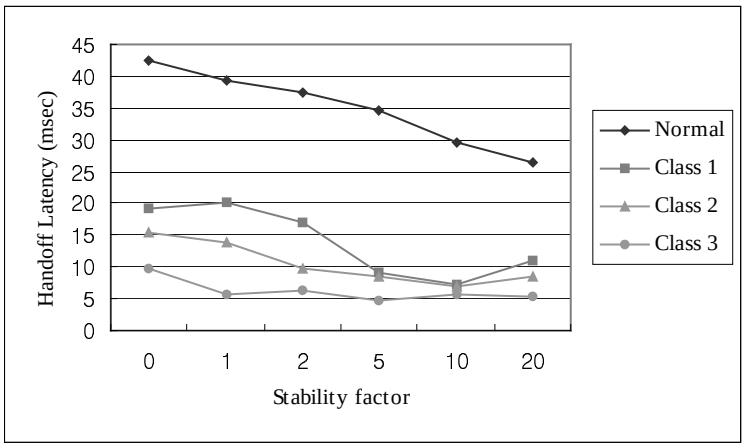

Figure 10. Handoff Latency (Remote AAA Server)

Fig. 9 shows the average handoff latency when the AAA server is located in the local domain. Generally, since the latency of local AAA server is not high, the average latency is also not high. However, we find that the average latency is lower when the proposed scheme is used. Also, since user 3, belonging to class 3, has a larger weight bound value, his latency is lower than that of a user belonging to another class, especially when user mobility is high. Fig. 10 shows the simulation result in the case of a remote AAA server. The overall pattern is very similar to that of Fig. 9. However, the average handoff latency is more than 40msec when the fast handoff scheme is not used. Since the end-to-end latency bound in multimedia applications is generally about 150msec, this value is not appropriate for multimedia applications. However, the latency obtained in the case of the fast handoff scheme is less than 20msec, more appropriate for multimedia applications and other situations requiring reduced latency times.

In mobility management, buffering at the previous AP is necessary for smooth handoff. The buffer size is affected by handoff latency. To determine the required buffer size at the previous AP, we used four traffic sources. Each traffic source has a constant bit rate and packet size. Table 1 shows the parameters for each source. Fig. 11 shows the necessary buffer size of each source for the cases where fast handoff is used or not. According to Fig. 11, the buffer size is substantially reduced when the proposed fast handoff scheme is used. Of course, in the case where fast handoff is not used, the buffer size here is less than 25Kbytes, however, this size is only for the case of one mobile host where each traffic source has bit rate less than 2Mbps. Furthermore, in this simulation, we didn't include any server processing time consumed during user authentication. Therefore, we can conclude that an increased buffer size is required for smooth handoff in real AP systems.

Table 1. Four Traffic Sources.

Sources	Bit rate (kbps)	Packet Size (bytes)
1	64	64
2	384	128
3	1024	512
4	2048	512

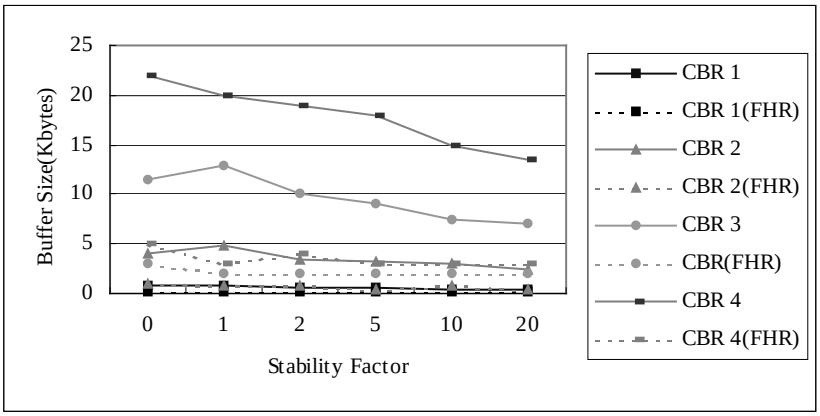

Figure 11. Buffer requirement at the previous AP

4.4 Overhead Analysis

In our scheme, since an "Access Reply" message that a mobile host receives through the authentication procedures contains multiple WEP keys, the length of the message is longer than that of the message used in the general scheme. The length of the message is proportional to the number of pre-authenticated APs calculated according to the *FHR selection* algorithm. However, no other information besides the WEP key needs to be duplicated, so the overhead induced by the increased message length is not a critical factor. Furthermore, the number of selected APs can be adjusted according to the users' service class, and the number of selected APs is usually quite small due to the users' locality in a local network environment.

Generally, the resources available in a wired link are much richer than those of the wireless link. Therefore, it is important to minimize the amount of resources being wasted in the wireless link. Although multiple APs are pre-authenticated in our scheme, there is only one message delivery in the wireless link between the mobile host and the current AP. Therefore, no additional resources are used in the wireless link.

5 Conclusion

In this paper, we proposed a fast handoff scheme for public wireless LAN. Since inter-AP handoff and authentication procedures are essential in public wireless LAN, we focused our attention on minimizing authentication latency during handoff. In our scheme, multiple APs, selected by the *FHRSelect* algorithm, are authenticated in advance. The *FHRSelect* algorithm utilizes traffic patterns and user characteristics, which are collected and managed by the centralized system. Simulation results show that the total handoff latency of the proposed handoff scheme is much less than that of the general handoff scheme. Furthermore, as regards smooth handoff, when the fast handoff scheme is used, less buffering is required at the previous AP. Therefore, the proposed fast handoff scheme should enable public wireless LAN systems to support a variety of multimedia applications.

6 Acknowledgements

This work was supported in part by the Brain Korea 21 project of the Ministry of Education, and in part by the National Research Laboratory project of Ministry of Science and Technology, 2002, Korea.

References

1. IEEE Standards for Local and Metropolitan Area Networks: Port based Network Access Control, IEEE Std 802.1x-2001, June 2001.
2. Blunk, L., and J. Vollbrecht, "PPP Extensible Authentication Protocol (EAP)", RFC 2284, March 1998.
3. A. T. Campbell and J. Gomez, "IP Micro-Mobility Protocols," ACM Mobile Computing and Communication Review, October 2000.
4. R. Koodli and Charles E. Perkins, “Fast Handovers and Context Transfers in Mobile Networks,” ACM Computer Communication Review, September 2001.
5. V. Kumar Choyi et al., “Fast Handoff Scheme in Wireless LANs For Real-Time Systems,” The Third IEEE WLANs Workshop, September 2001.
6. Pat R. Calhoun et al., ”Diameter Base Protocol,” Internet draft, draft-ietf-aaa-diameter-10.txt, April 2002.
7. D. Mitton et al., "Authentication, Authorization, and Accounting: Protocol Evaluation," Internet RFC 3127. June 2001.
8. Christian Bettstetter, “Smooth is Better than Sharp: A Random Mobility Model for Simulation of Wireless Networks,” ACM MSWiM 2001, July 2001.
9. S. K. Sen et al.,” A Selective Location Update Strategy for PCS Users,” ACM/Baltzer J. Wireless Networks, September 1999.