Dual Authenticaions for Fast Handoff in IEEE 802.11 WLANs: A Reactive Approach

SeongWoo Kim and SeungWoo Seo
Department of Electrical Engineering and Computer Sciences
Seoul National University
Email: swkim@cnslab.snu.ac.kr, sseo@snu.ac.kr

Abstract—Although the mobility between APs (Access Point) was initially not a major concern of IEEE 802.11, the inter-AP mobility becomes an essential issue in WLAN toward the paradigm of ubiquitous computing. Supporting inter-AP mobility, however, incurs handoff latency including discovery and reauthentication delay. In particular, most efforts to reduce the reauthentication delay have focused on proactive approaches, which transfer security contexts to candidate network entities via an inter-AP protocol before handoff occurs. These proactive approaches have a number of restrictions such as target prediction and inter-AP communication. The selection of a candidate network inherently has a probabilistic in-deterministicity. Implementation and deployment of inter-AP communication have not been successful so far, and even been withdrawn from IEEE 802.11 standardization. In this paper, we propose a novel deterministic reactive authentication scheme to achieve fast handoff in IEEE 802.11 which does not require inter-AP communication. The proposed protocol is divided into two steps: Immediate Authentication (IA) and Full Authentication (FA). IA enables the AP receiving an authentication request to allow a mobile node (MN) to temporally access the network, if the MN has trustworthy evidence which the AP can validate promptly. In the FA step, the AP fully authenticates the MN for reducing the optimisticy of IA’s temporal authentication. The performance evaluation and security analysis show the proposed scheme can reduce reauthentication delay enough to support seamless inter-AP mobility without a significant sacrifice of secrecy in practical and realistic scenarios.

I. INTRODUCTION

Since IEEE 802.11 has originally been developed to supplement or replace the wired LAN with wireless, inter-AP mobility was initially not a major concern of IEEE 802.11. The inter-AP mobility, however, becomes an essential issue in WLAN toward the paradigm of ubiquitous computing. Supporting mobility necessitates a handoff procedure which may take excessive latency affecting delay-sensitive applications such as streaming audio/video and IP telephony. For example, the overall latency should not exceed 50 ms to avoid excessive jitter in VoIP applications [1]. However, as shown in some experiments, traditional IEEE 802.11 may suffer from more than 1000 ms handoff latency [4]. Thus, to overcome this problem, many proposals have been made for seamless session mobility in WLAN.

The handoff of IEEE 802.11 consists of discovery and reauthentication procedures. In the discovery step, a mobile node (MN) needs to find a potential AP with which to associate. This discovery step gives the most significant contribution to the total handoff latency [11]. Several schemes such as Neighbor-Graph pruning scheme [14], SyncScan [13], and Multiscan [5] have been proposed and effectively reduce the channel probing latency to less than 20-30 ms. In the reauthentication step, an MN tries to associate with new AP, and the AP authenticates the MN with the help of Authentication Server (AS). Most efforts to reduce the reauthentication delay have focused on proactive approaches such as pre-authentication or proactive key distribution [10], [12], which transfer security contexts to candidate network entities, e.g., neighboring APs, via inter-AP communications before handoff occurs. In contrast to proactive approaches, reactive handoff operations are initiated in response to the completion of a discovery step.

The proactive approaches obviously accommodate prompt and secure handoff because there is no need for reauthentication after handoff and security contexts are transferred through the backbone network. In particular, the proactive authentication schemes are suitable for centralized network systems. In centralized systems, e.g., cellular networks, the central systems keep track of all information related to handoff, and even make a handoff decision including target indication. Therefore, the central system can transfer MN’s information to the target network deterministically before handoff.

Although the proactive manners achieve two goals at the same time, they have a number of weaknesses for IEEE 802.11 WLAN such as target prediction and inter-AP communication. In distributed systems, e.g., IEEE 802.11, a handoff is initiated when the connection with the current AP is closed. Nobody knows the exact target network before handoff. In order to transfer security contexts before handoff, the selection of potential target networks is indispensable. When the signal strength from current AP is lower than a certain threshold, either the MN or the current AP transfers the MN’s security contexts to the selected potential APs. Since transferring security contexts to other APs has a problem of scaling up, the proactive approaches adopt the prediction of candidate APs towards which the MN will move. These predictive approaches, however, have inherently probabilistic in-deterministicity, which makes it hard to be adopted in commercial markets. Another significant weakness of proactive manners in IEEE 802.11 is that they assume inter-AP communications such as the usage of Inter Access Point Protocol (IAPP) [15]. However, implementation and deployment of inter-AP communications
have not been successful so far. Moreover, it was withdrawn during standardization and is not currently included in current IEEE 802.11 standard.

In this paper, we propose a deterministic reactive authentication scheme to achieve fast handoff compatible to IEEE 802.11i [16] without inter-AP communications. The IEEE 802.11i specifies the use of the IEEE 802.1X1 [17] and four-way handshake, which offers a general framework to build authentication services and key distributions. In order to reduce the post-handoff IEEE 802.11i authentication latency, we introduce Immediate Authentication (IA). If an MN has particular evidence which proves that the MN is associated with an old AP within an acceptable time, new AP allows the MN to access the network temporarily. It should be possible for the evidence to be checked quickly without any help from AS or other APs. Since the access to the network is temporarily allowed, the MN has to be fully authenticated by performing the original IEEE 802.1X and four-way handshake within a constrained time. This step is called Full Authentication (FA). As a result, the proposed protocol is divided into two steps: optimistic but prompt IA, and FA which is equal to the reauthentication procedure of the conventional IEEE 802.11i. Performance evaluations show that our scheme can reduce reauthentication delay less than 41% without a significant sacrifice of security in realistic environments compared to the minimum delay of original IEEE 802.11i reauthentication.

The remainder of this paper discusses our work in more detail. We present a concrete algorithm and protocol for dual authentications scheme in Section II. Section III evaluates our approach using analysis and simulation. Section IV discusses security, performance, and deployment issues of our proposal. Section V summarizes our work.

II. DUAL AUTHENTICA TIONS

In this section, we describe our proposal in detail. To reauthenticate an MN quickly through IA, we propose to use One-Time Ticket (OTT). FA is equal to IEEE 802.11i authentication except delivering OTT. The following sections will give mechanisms to achieve the objective of our paper by using OTT, IA, and FA.

A. One-Time Ticket

In the IEEE 802.11i, when an MN enters a network initially, the MN has to perform the IEEE 802.1X authentication for AS

1The standard includes three entities: Supplicant, Authenticator, and Authentication Server. The Supplicant and the Authenticator in the standard are corresponding to MN and AP, respectively. Although they can be used interchangeably, we prefer to using MN and AP in this paper.

to decide whether the MS’s access is allowed or not. IEEE 802.1X authentication contributes the most significant latency to authentication delay. For example, in the case of EAP-MD5 without IPSec, the initial association with an AP takes 400 ms through IEEE 802.1X. However, when an MN performs a handoff, IEEE 802.1X authentication takes 1700 ms [2]. To decrease the latency for IEEE 802.1X authentication, we propose One-Time Ticket (OTT), providing a method to promptly check whether an MN had ever associated with an AP dominated by the same AS. OTT is delivered from AS to the MN, piggybacked on the initial IEEE 802.1X message. The detailed procedure of OTT delivery is explained in Section II.D.

The OTT is established by hash chain [9]. Let $H^n_{s,k}$ denote a set of hashed keys established by a seed $s$. Here, $k$ is a secret key, and $n$ is the number of hashed keys. $H^n_{s,k}$ can be formulated by

$$H^n_{s,k} = \{ h_i | h_i = HMAC_k(h_{i-1}), h_0 = s, 0 < i \leq n \}$$

where $HMAC_k$ is a keyed-Hash Message Authentication Code [8]. $b_i$ is an OTT value and is bound to the time when the MN is guaranteed. For example, if one OTT is guaranteed for $l$ sec, $H^n_{s,k}$ is valid for $n \times l$ sec. If an MN tries to associate with an AP at time $t_o$, the corresponding OTT is $h_j$, where $j = n - \left\lfloor (t_o - t_o) \times l^{-1} \right\rfloor$, and $t_o$ denotes the time when the seed was distributed. Since $H^n_{s,k}$ is bound to time, APs and AS should be time synchronized.

Fig.1 shows the timing diagram of OTTs. For instance, if an MN tries to associate with new AP at $t_o$, the MN receives $h_{n-2}$ as its OTT. $h_{n-2}$ are valid until $t_1$. Even if the given OTT is expired, it can access the network through the current AP. However, if the MN does handoff, it has to start from the original IEEE 802.1X reauthentication.

B. Immediate Authentication Procedure

A seed $s$, randomly chosen by AS, has to be shared between APs and AS in advance. $s$ can be distributed through two modes: passive and active. In the passive mode, AS broadcasts the seed to all APs regularly through Send-advertisement message (P1) in Fig.2. Table I shows the elements of Send-advertisement. $k$ is used for securing hash function. Particularly, $t_o$ is a current datetime used for time synchronization between APs and AS. In the active mode, when an AP boots up, the AP requests $s$ and $k$ from the AS through Seed-
solicitation message. The AS replies the AP through Seed-advertisement.

Since $s$ and $k$ are shared, APs and AS can generate common $H_{s,k}^{\text{IA}}$. When an MN attempts to access the network initially, IEEE 802.1X authentication and four-way handshake are performed. As a result of IEEE 802.1X authentication, AS sends the MN an OTT correspondent to the current time $t_v$, and its expire time $t_e$ (P2). $t_e$ is set to $t_v + t_c$, where $t_v$ is the valid duration of an OTT in Table I. This procedure is equal to IEEE 802.1X authentication and four-way handshake procedure except an OTT and $t_c$ attached. When the MN leaves from the current AP at time $t_v$, the MN starts to probe available channels and tries to be reauthenticated by new AP, $AP_n$. In order to reduce reauthentication delay, the MN sends its OTT to $AP_n$ through OTT-request (P3). Since the channel between MN and AP is insecure, $f_h(MA||OTT)$, $MA$ are transmitted from MN to $AP_n$ as an OTT-request. $f_h$ is an one-way hash function such as MD5 and SHA-1. $MA$ is MN’s address, which is not explicitly transmitted as a data field of OTT-request but implicitly obtained from the message header of the received OTT-request. This implicit address transmission is used for the generation of a temporal encryption key of IEEE 802.11i [16]. If $t_v > t_e$, the MN performs not IA but FA.

When receiving the OTT-request, $AP_n$ is able to check immediately whether the MN has ever associated with other AP dominated by the same AS through comparing with the OTT cache table maintained in $AP_n$ (P4). If OTT from the MN is valid, $AP_n$ notifies MN’s admission through OTT-response (P5), and initiates a four-way handshake to establish session keys used for traffic encryption (P6). The session keys are derived from the Pairwise Master Key (PMK), Authenticator Address (AA), Supplicant Address (SPA), Authenticator Nonce (ANonce), and Supplicant Nonce (SNonce). MN and AP already share PMK as a result of IEEE 802.1X authentication. MN knows its own address (SPA) and generates SNonce. AP is vice versa. Thus, the both share the five elements if the AP sends ANonce to the MN, and the MN sends SNonce to the AP. AA and SPA are extracted from the message header of Nonce transmissions. The procedure of sharing these five elements consists of four messages. Since IA does not generates PMK, IA uses OTT as a PMK. IEEE 802.11i also specifies 256 bit Pre Shared Key (PSK) shared beforehand between AP and MN, which derives PMK through Pseudo Random Function (PRF) without AS. Consequently, since OTT can be used for PSK, IA requires no modification of the four-way handshake in IEEE 802.11i. The degree of security of encryption key generation with OTT will be analyzed in detail in Section III.B.

After P6 is completed, the MN is able to access the network on secure channel (D3) established by the procedure P6. The reauthentication delay of our proposal is from P3 to P6. No packet is exchanged with $AP_n$ or AS so far. The original IEEE 802.1X authentication and four-way handshake have to be performed within a given constrained time for restraining IA’s optimism (P7). At this moment, MN’s OTT is updated (P8). AS should broadcast Seed-advertisement within the Ticket Grant Lifetime (P9).

C. OTT validation

AP should maintain OTT-cache table which contains valid OTTs. OTT entries of the cache table is formulated by

$$H_{MA} = \{\xi_l | \xi_l = f_h(MA||h_i), h_i \in H_{s,k}^{\text{IA}} : \xi_l < \beta - i < \xi_h\}$$

(2)

where $\xi_l = \max(0, [(\tilde{t}_c - t_v - t_s) \times l^{-1}])$, $\xi_h = [(\tilde{t}_c - t_s) \times l^{-1}]$ and $\tilde{t}_c$ is the current time. Let $h$ be $f_h(MA||OTT)$ contained in OTT-request message. If $h \in H_{MA}$, this OTT-request is valid, thus AP notifies MN is allowed to access the network through OTT-response (P5). OTT-response contains the time (constrained IA time) within which indicates the MN must perform the original IEEE 802.1X authentication and four-way handshake. Otherwise, there is a possibility that the MN is not a normal user such as an abuser reusing cached old OTTs. AP can decrement a constrained IA time bound to correspondent OTT in order to mitigate the affect of OTT abuse.

D. Full Authentication

FA is equal to standard IEEE 802.11i authentication except delivering OTT. IEEE 802.11i defines two procedures for enhancing the security of IEEE 802.11: IEEE 802.1X for authentication, and four-way handshake for temporal encryption key generation. IEEE 802.1X provides a port-based network access control. AP blocks all traffics except 802.1X messages before MN is authenticated. In order to control the network access of MN, IEEE 802.1X uses Extensible Authentication Protocol (EAP) [3]. IEEE 802.1X consists of two EAPs: EAP between MN and AP, and EAP between MN and AS. In P7 of Fig.2, the EAP over WLAN and the EAP over RADIUS
III. PERFORMANCE EVALUATIONS

A. Handoff Delay Evaluation

Let $T_r$ denote the response time from the point when reassociation is initiated to the point when four-way handshake is completed. $T_r$ is formulated by

$$T_r = T_{ras} + T_{ath} + T_{fwh}$$

(3)

where $T_{ras}$ is a delay for reassociation, and $T_{fwh}$ is a delay for four-way handshake. In IEEE 802.11i, $T_{ath}$ is a delay for IEEE 802.1X authentication, which is represented by

$$T_{ath} = T_{1X} = T_{col} + T_{cor}$$

(4)

where $T_{col}$ is the time for EAP over LAN and consists of three delay components as follows:

$$T_{col} = n_1T_{ma} + T_{ac1}$$

(5)

where $n_1$ is the number of transmissions between MN and AP during EAP over LAN, $T_{ma}$ is the time taken for an IEEE 802.11 packet transmission, and $T_{ac1}$ is the computational time for processing EAP messages on AP. In (4), $T_{cor}$ is the time taken for UDP-based EAP over RADIUS, and consists of six delay components as follows:

$$T_{cor} = n_2T_{ma} + n_3T_{aa} + T_{ac2} + T_{sc}$$

(6)

where $n_2$ is the number of transmissions between MN and AP during EAP over LAN, $n_3$ is the number of transmissions between AP and AS, and $T_{aa}$ is the time taken for a UDP packet transmission, $T_{ac2}$ is the computational time for processing EAP messages on AP, and $T_{sc}$ is the computational time for processing EAP messages such as database searching on AS.

In the proposed scheme, the delay can be formulated by

$$T_{ath} = T_{ott} = 2T_{ma} + T_{ac3}$$

(7)

where $T_{ac3}$ is the computational time for processing OTT-request message on AP. The factor 2 of $T_{ma}$ means two OTT messages. Now, the condition of $T_{1X} > T_{ott}$ can be derived from (4), (5), (6), and (7).

$$\alpha T_{ma} + n_3T_{aa} + T_{ac1} + T_{ac2} + T_{sc} > T_{ac3}$$

(8)

where $\alpha = n_1 + n_2 - 2$ and its value is between 4 and 5. $T_{ma}$ and $T_{aa}$ are time-varying functions with several parameters such as the number of users, traffic patterns, interferences and noises. Since $T_{1X} \gg 2T_{ma}$, the left term of (8) is approximately equal to $T_{1X}$. Many measurement data of $T_{1X}$ have been published and the majority of results is around 1 s. One of the fastest measurement of $T_{1X}$ is 250 ms measured in fast resume mode by the author of [4].

The dominant operation $T_{ac3}$ is OTT validation introduced in Section II.C. $T_{ac3}$ is determined by $t_v$ and $l$, which are delivered to AP through Seed-Advertisement and described in Fig.1 and Table I. The total possible number of keys assigned to each MN per AP is formulated by $\eta = \frac{T_{1X}}{l}$. $\eta$ also represents the maximum number of iterations for OTT validation. Since $T_{ath}$ is the time to perform a linear searching based on hashing and comparing, $T_{ath}$ is proportional to $\eta$. The OTT validation takes $O(\eta)$ times. The expectation of $T_{ath}$ is defined as

$$E[T_{ath}] = (\eta \times T_h)/2$$

(9)

where $T_h$ is $f_h$ computation process time for OTT validation. Our analysis adopts SHA-1 as $f_h$. Table II shows these parameters in detail. $T_h$ is measured by Benchmarks of Crypto++ 5.5 Library [6]. Conclusively, reauthentication delay in our architecture, i.e., $T_r$, can be written as a function of $\eta$:

$$T_r = f(\eta)$$

(10)

where $f(\eta)$ is a reauthentication delay function.

Fig.3 shows the average $T_{ath}$ by 70 measurements according as $\eta$ increases. The first bar, 250 ms, indicates the minimum delay of $T_{1X}$ in the conventional IEEE 802.1X [4]. $T_{ath}$ is proportional to $\eta$. The variation of the histogram in Fig.3 comes from linear searching characteristics of OTT validation. Fig.3 shows that the proposed scheme can reduce reauthentication delay averagely less than 41% (100 ms) compared to the minimum delay of the conventional IEEE 802.11i reauthentication, if $\eta \leq 400$. Here, since this reauthentication delay reduction is achieved at the sacrifice of secrecy, the sacrifice should be considered for practical and deployable applications. Security considerations of our proposal, especially with the case of $\eta \leq 400$ for $T_{ath} < 100ms$, will be given in the following section.

B. Quantification for the Degree of Security

A malicious MN is not able to know other MNs’ OTT because OTT is securely delivered through EMSK of EAP. However, it is possible that there are MNs that have the same OTT of a compromised MN. If an MN moves into new AP, the probability that MNs assigned the same OTT to the MN exist

$$p_d = 1 - \left(1 - \frac{1}{\eta} \right)^{m_n}$$

(11)
where \( m_n \) denotes the total number of MNs associated with the visited AP. We can find out that the probability of post-handoff OTT duplication decreases, as \( \eta \) increases. Conversely, (9) shows that the reauthentication delay of IA is proportional to \( \eta \). In order to reduce reauthentication delay, \( \eta \) should be smaller. This implies a tradeoff between speed and secrecy. The reduced handoff delay decreases the degree of security by as much as \( p_d \) compared with IEEE 802.11i. For example, as shown in Section III.A, the expected reauthentication delay can be less than 100 ms, if \( \eta \leq 400 \). In this case, if less than 7 users are connecting the visited AP, \( p_d \) is less than 0.017369. If the number of users is less than 4, \( p_d \) is less than 0.009962. That is a trivial sacrifice considering the amount of the reduced reauthentication. Furthermore, since FA is performed immediately within a constrained IA time after MN attaching new AP, the effect of \( p_d \) is significantly mitigated.

IV. DISCUSSIONS

In this section, we discuss security, performance, and deployment issues of our proposal. We can consider the possibility that an authenticated MN can regenerate a partial hash chain with its own OTT, even though the MN does not know seed \( s \). The partial hash chain is defined as \( H_{h, k}^n \in H_{s, k}^n \) where \( h \) is an OTT of a malicious or compromised MN. If an MN knows \( k \), it can generate a partial hash chain. However, since IEEE 802.11i assumes that AP and AS have established a secure channel [16], we can assume \( s \) and \( k \) are delivered securely. Eventually, no MN can regenerate a partial hash chain.

The basic idea behind IA is to support seamless handoff through reducing time for reauthentication immediately after handoff. However, as much as IA reduces the delay, the system has to perform an additional operation, i.e., FA after IA, to pay for the benefit. In the long term perspective, the total reauthentication delay of our proposal is longer than conventional approach by as much as \( T_{ott} \). These characteristics inherently come from the purpose of fast reactive reauthentication. However, network operations after attaching new AP can be performed in parallel rather than ones during handoff, because a connection is already established. Therefore, the delay can be avoidable with two possible approaches: FA under background, and FA during idle time. In addition, although the amount of secrecy decrease from the IA's delay benefit, \( p_d \), is trivial in realistic scenarios as we consider in section III.B, FA mitigates \( p_d \) as much as negligible.

Seed-Advertisement can also incur overhead which does not exist in the conventional approaches. In practice, we expect that Seed-Advertisement may be triggered at most once in a day. This may be an acceptable overhead. Finally, AP which MN visits can be the AP which is not dominated by the same AS. Our approach can not be applied in this case. Several research groups including HOKEY [7] investigate such issues.

V. CONCLUSIONS

We propose a deterministic reactive reauthentication mechanism for fast handoff in IEEE 802.11 wireless networks to support seamless services, which consists of optimistic but prompt IA, and FA which is equal to the reauthentication procedure of the conventional IEEE 802.11i. Performance evaluation and security analysis show that the proposed scheme can reduce reauthentication delay enough to support seamless inter-AP mobility without a significant sacrifice of the degree of security in practical and realistic scenarios.

VI. ACKNOWLEDGEMENT

This work was supported by Seoul R&BD program. The authors would like to thank Yen Ye and JuHyung Son for helpful suggestions and discussions.

REFERENCES