Secured Fast Handoff in 802.11-Based Wireless Mesh Networks for Pervasive Internet Access

Gang YAO^{†a)}, Jiannong CAO[†], Ye YAN^{††}, Nonmembers, and Yusheng JI^{†††}, Member

SUMMARY Authentication issue has been mostly ignored to ensure fast handoff in 802.11 Wireless Mesh Network (WMN). With the proliferation of WMNs in recent years for practical deployment, secured fast handoff has drawn much attention to enforce authenticated access while reduce the extra delay caused by enabling authentication operations. In this paper, we present an overview on the state-of-the-art advance in this field and tackle the problem from a practical perspective based on experiments and analysis on our real-world testbed HAWK. We propose a novel fast handoff scheme Network-assisted Radio Signature to eliminate probing delay by taking advantage of the characteristic of the actual dynamic topology about mesh routers in WMN. Moreover, we apply an optimistic authentication mechanism Dual Re-authentication to counteract the authentication delay while providing the secured wireless access. In this manner, we have reduced the end-to-end handoff delay of WMN back again to a level below 50 ms to achieve secured handoff and support time-sensitive applications. We describe detailed mechanisms, simulation, implementation and experimental results. To our best knowledge, we are the first to achieve such an optimal performance of secured fast handoff.

key words: handoff management, fast handoff, secured fast handoff

1. Introduction

Provision of pervasive wireless Internet access to citizens anywhere and anytime has been a hectic drive for many governments and operators across the globe. To serve this purpose, IEEE 802.11-based Wireless Mesh Network (WMN) has just been matured to a point of wide acceptance as a cost-efficient technology. Hence, authentication issue naturally becomes a significant concern and ongoing effort when it comes to the stage of commercial deployment, though it has been mostly ignored in early studies on Layer 2 handoff. Extra delay is introduced into handoff process by enabling authentication operations like challenge-response, multiple way handshakes and key distribution etc. so that it severely undermines prior efforts towards fast handoff in 802.11 networks.

We have been carrying out empirical studies on mobility management regarding multiple radio technologies under HAWK (wmn.comp.polyu.edu.hk) project. In this paper, we address the challenges of existing IEEE 802.11 handoff mechanisms. By identifying the bottleneck and the direction

Fig. 1 Origins of handoff delay on multiple layers.

for improvement, we propose specific and applicable solutions for fast and authenticated handoff to support Internet access.

IEEE 802.11-based WMN consists of a self-organized and self-configurable mesh backbone and conventional IEEE 802.11 mobile clients (MCs). The mesh routers of the backbone communicate with each other via multi-hop wireless links while serving as Access Point (AP) for MC. Although a WMN shares similarities as ad hoc network, it distinguishes itself with a static topology or trivial mobility of mesh routers. In early deployments of WMNs, open access was a common practice which incurs public concerns on security later. So subsequently authentication measures are enforced which include basic WEP, WPA and WPA2. However, although the security objective is fulfilled, the performance of WMNs is influenced, particularly the handoff.

Basically, the fundamental performance bottleneck of handoff in IEEE 802.11 WMN is caused by interruptions of communication occur in the event of MC changing its point of attachment, e.g. AP. Under the circumstance, the time-sensitive applications such as streaming media demand a delay of less than 300 ms while VoIP demands one less than 50 ms, specified by the long accredited International Telecommunication Union (ITU) standard [1]. Such requirements impose difficulties upon the primitive IEEE 802.11 network mechanisms.

As illustrated in Fig. 1, the handoff involves multiple dimensions of operations from the bottom physical layer to the upmost application layer. Conventionally, research focuses on Layer 2 and Layer 3 due to that they contribute the most of the delay. Concerning Layer 3 handoff, it covers IP acquisition, route update, etc. and has been long investigated. Mobile IP (MIP) and its derivatives [9], [10] have been standardized as major solutions. The situation about Layer 2 handoff is somewhat different and far from

Manuscript received July 19, 2009.

Manuscript revised October 9, 2009.

[†]The authors are with the Department of Computing, Hong Kong Polytechnic University, Hong Kong.

^{††}The author is with the Department of Computer and Electronics Engineering, University of Nebraska-Lincoln, U.S.A.

^{†††}The author is with the National Institute of Informatics, Tokyo, 101-8430 Japan.

a) E-mail: csgyao@comp.polyu.edu.hk

DOI: 10.1587/transinf.E93.D.411

being mature. It involves operations including probing for candidate APs, authentication of MC and AP, and association with AP. The delay caused by association is determined by hardware. It is relatively fixed and simple similar to channel switch on Layer 1. However, on the contrary, there are diversified probing and authentication mechanisms existed in vendors' implementation. Early research targets probing delay with assumption of open authentication which means no authentication at all; currently secured fast handoff tackles with probing delay with authentication enabled. It is obvious that varied authentication methods, e.g. WEP/802.11X, WPA/802.11i/WPA2, involve various ways of handshaking and exchanges of authentication packets to interrupt communication. The delay accumulates. The longer authentication procedure is, the performance of secured handoff deteriorates more. In order to achieve an optimal overall performance in real world, we have proposed an integrated solution, namely a novel Layer 2 fast handoff scheme Network-assisted Radio Signature (NRS) and an optimistic authentication mechanism Dual Re-authentication. The NRS scheme eliminates the probing delay by taking advantage of the characteristic of the actual topology of a WMN, that is the actually topology of mesh routers is much less dynamic. We assume that probing for candidate APs can be conducted proactively and the outcome of probing can be valid as long as the topology remains unchanged. In this sense, the delay caused by reactive probing is totally eliminated. Moreover, the Dual Re-authentication mechanism counteracts the authentication delay when providing secured wireless Internet access by utilizing optimistic Access-before-Authenticate approach. Based on the total solution, we have eventually achieved an end-to-end secured handoff delay which is around 50 ms.

We claim our contributions as follows:

- We propose and implement a total solution of secured fast handoff which includes NRS and Dual Re-authentication mechanisms in an 802.11-based WMN. To our best knowledge, we are the first to achieve such a performance of secured fast handoff.
- NRS is the first Layer 2 fast handoff scheme to realize proactive probing by making the most of the static topology of a WMN in real world.
- The Dual Re-authentication mechanism improves the performance of secured handoff in a manner of optimistic access, meanwhile, it does not comprise the security by providing its unique design of one time ticket (OTT).

The remainder of this paper is organized as the below. We give an overview to state-of-the-art of the related work in Sect. 2. The Sect. 3 presents our novel Layer 2 fast handoff scheme Network-assisted Radio Signature (NRS). The featured Dual Re-authentication mechanism is introduced in Sect. 4. We also provide the experimental and evaluation results respectively. In the end, we conclude the current progress and future work in Sect. 5.

2. Related Work

The secured fast handoff of 802.11 WMN involves two main aspects of operations: namely probing and authentication. In this section, we present an overview on the existing probing schemes and fast authentication mechanisms for Layer 2 fast handoff in Table 1. A simple analysis on pros and cons of the related works is offered subsequently. Note that the probing and scanning are interchangeable in the following section. Since the probing delay is one of the two main contributors to the overall Layer 2 handoff delay, most fast handoff schemes aim to reduce this lengthy process. Some schemes rely on reducing the number of scanning channels, the time taken on each channel, scanning-related timers [2]. The tuning technique targets to find an optimal value for the time taken on each channel, i.e. MinChannelTime and MaxChannelTime to reduce active scanning delays [3]. Intelligent channel scanning aims to minimize the probing and waiting time on each channel [4]. SyncScan synchronizes MCs and APs and instructs MCs to scan the channel by switching channels at the exact moment when a beacon is about to arrive with periodic beacon broadcasting from APs [5]. Instead of probing all available channels individually, selective scanning reduces the number of channels required to discover APs. A number of selective scanning approaches have been proposed, such as selective scanning plus AP caching methods [6] designed to reduce L2 hand-

Table 1 Comparison of fast probing and authentication mechanisms.

	Features	Pros	Cons
Timer tuning	Optimize values of scanning timers	Simple and effective	Limited improvement
Intelligent channel scanning	Minimize the probing and waiting time on channels	Simple and effective	Limited improvement
SyncScan	Periodically probe channels in case of communication	Reduce the delay significantly	Large overhead
MultiScan	Use one radio to probe channels while another radio for communication	Minimize the delay to close to 0 ms	Extra hardware is required
Neighbor graph	Reduce the probed channels and probe-waiting time by using NG	Systematic Enhancement of timer tuning and channel selection	Large overhead, limited improvement
Procedure optimization	Merge overlapping procedures in different authentication mechanisms	Simple and applicable	Limited Improvement on current mechanisms

off delay to a level where VoIP communication becomes seamless [7]. Another typical scheme is Neighbor Graphs approaches, which aim to reduce the total number of probed channels and the probe-waiting time on each channel via selective scanning as well as using caching techniques to solve the problem of packet loss [8]. MultiScan uses multiple radio interfaces equipped on MC to search proactively for alternate APs while being associated with an AP and interleaving data communication [17].

Any authentication mechanism will introduce additional delay and network overhead into the network. Three major approaches to optimize and realize fast authentication in secure fast handoff have been proposed. The first category of approaches is to integrate and merge the overlapping procedures in authentication to reduce the total time delay. In [11], Optimized Integrated Registration Procedure for MIP and SIP with AAA operations is proposed and attempts to reduce the delay by combining the round trips between the MC and the home AAA server when MIP and SIP are both deployed. Another similar case is MPA [12]. Both of them employ cross-layer approach to reduce redundant changes of messages in authentication procedures as one of basic technique to optimize authentication delay, nevertheless they cannot eliminate the delay. Proactive authentication approaches constitute another direction for fast authentication, e.g. Shadow Registration [13], which is proposed to optimize secure handoffs by proactively conducting the authentication before the handoff really occurs. A security association is established between the MC and every neighboring APs before the former associates with one of them. Similar mechanisms are employed in AAA context transfer [14] and P2P context transfer [15]. Such proactive authentication mechanisms can further reduce the authentication delay but may cause the problem of heavy traffic and waste of resources induced by redundant information transfer. The third approach is called optimistic access which postpones the real authentication to a later moment as addressed in [16]. During the handoff process, the network allows an optimistic access by the MC to the new mesh router using a light-weight authentication instead of executing a normal authentication. The full authentication is executed in a later stage. If the MC cannot pass the full authentication, the current access will be terminated. This technique makes a trade-off between security and performance. The possibility of suffering from vulnerability is reduced by the small time window before the full authentication completes.

Based on the survey on the related work, we observe that no significant outcome has been announced in literature to further achieve optimal performance of secured fast handoff in term of end-to-end delay due to the complicity of system implementation. Since HAWK project values practical and overall system performance via an empirical approach, so we attempt to provide a solution which combines mechanisms of fast handoff with authentication for secured Internet access, namely Network-assisted Radio Signature and Dual Re-authentication.

3. Secured Fast Handoff Solution

3.1 Network-assisted Radio Signature (NRS) Scheme for Layer 2 Fast Handoff

3.1.1 Motivation

If we look into the handoff schemes which are summarized in Sect. 2, it is unobvious that there are certain assumptions behind their common practice: first the design philosophy beneath all probing schemes is that the network topology keeps changing, so that the topology information needs to be collected in a on-the-fly manner and keeps updating; secondly, the schemes are mainly client-initiated, i.e. the handoff intelligence is employed at MC side instead of network side, which means the network intelligence has not been utilized. We notice that such assumptions do not always hold when it comes to a WMN in the real world.

Hence, we argue that if the network topology is relatively static, the handoff decision can be made not only in a predetermined manner but also stand long. The delay caused by probing can be totally avoided and the collected topology information can be reused as long as the mesh routers remain unchanged in the life cycle of the deployment. In order to proactively collect the predetermined network topology information, i.e. candidate APs for handoff, we decouple a training phase from the operational phase in the life cycle of the WMN. In order to comply with the fact of single radio for most MCs (otherwise the schemes which utilize extra radios like MultiScan [17] could be applied), multiple APs have to be coordinated to monitor and manage the handoff. Based on the above arguments and analysis, we present the details of NRS scheme in next section.

3.1.2 Overview

The fundamental design philosophy of the proposed NRS scheme is that since the effort of continuous and proactive probing deals with highly dynamic network scenario which is not necessarily right in all occasions, we tradeoff such an advantage to gain the improvement of handoff performance.

As illustrated in Fig. 2, when a WMN equipped with NRS is put into operation, an MC moves to a certain location in the network. The most preferred AP to associate for the client is predefined in the training phase. The radio characteristics of the client are monitored by multiple APs: e.g. AP1, AP2, AP3, in the network. For an instance, the on-the-fly measurement of the Received Signal Strength Indicator (RSSI) is transmitted by the MC. It constitutes the radio signature of the MC on the scene. The mapping relationship between the preferred AP and the corresponding radio signature of an MC is stored in a knowledge base in advance. Thus the knowledge base is looked up to find the most matched radio signature stored against the radio signature monitored by the APs. According to the match, a

Fig. 2 The illustration of NRS scheme.

corresponding AP, which is also the preferred AP to associate with, is determined. After the lookup, the handoff is initiated and managed by the network side. In this manner, handoff delay is totally eliminated since the MC does not probe for candidate APs anymore when handoff occurs.

How to build such a radio signature knowledge base? In line of the operational phase, the MC needs to traverse the coverage area of the network to collect topology information in another training phase. The coverage area is separated into subarea with certain granularity. The MC is used for measurement in each subarea one by one. In a subarea, the radio signal transmitted by the MC is recorded, processed and stored in the knowledge base by AP1, AP2, AP3. The most preferred AP is arbitrary and determined by the designer of NRS system according to specific application requirements, for an instance, the AP which provides highest data rate by the actual measurement. This implies that the most preferred AP can be either of AP1, AP2, or AP3, but not necessarily be one of them. After the preferred AP is determined, the mapping between this AP and the radio signature is recorded into the knowledge base.

3.1.3 NRS Operations and Algorithms

A. Training Phase

The objective of the training phase is to build the radio signature knowledge base. In practice, we use a notebook PC which is equipped with an 802.11 STA to traverse the subarea of the WMN. The radio characteristics of the STA, which is the so-called radio signature, is measured, processed and stored by the functional entities resided on APs. As in Fig. 3, the process is illustrated step by step as follows:

- **Step 1:** The most preferred AP is determined by a pre-defined metric, e.g. data rate, RSSI, or Signal Interference Ratio (SIR). To simplify the implementation, we adopt RSSI as the indicator. It is noticeable that the implication of this RSSI is different from the RSSI used as radio signature. We aim to establish the relationship: $\max\{RSSI_1, RSSI_2, \dots, RSSI_i\} \rightarrow AP_{preferred}$
- **Step 2:** The area of 802.11 WMN is divided into $m \times n$

```

{ //MC traverses the network area
  for each subarea  $A_i \in \text{Area}_{\text{network}} A$  do
  {
    for each  $AP_i \in \text{AP}_{\text{NRS}}$  do
    { //Monitor and calculate the real time radio signature.
      //Process the radio signature with application-specific fuzzy functions.
       $\mu_{AP_i\_RSSI\_high} = f_{\text{fuzzy}}(AP_i\_RSSI)$ ;
      RS = ( $AP_1\_RSSI\_high, \mu_{AP_1\_RSSI\_high}$ ),
      ( $AP_2\_RSSI\_high, \mu_{AP_2\_RSSI\_high}$ ), ..., ( $AP_i\_RSSI\_high, \mu_{AP_i\_RSSI\_high}$ )
      // Associate RS with the most preferred AP according to data rate or RSSI etc., write in database KBD.
    }
  }
}

```

Fig. 3 The algorithm of NRS for operational stage.

subareas: $A_1, A_2, \dots, A_i$.

- **Step 3:** In the subarea A_i , manually associate MC with $AP_1, \dots, AP_i$, measure the RSSI, to determine the most preferred $AP_{preferred}$. e.g. when $\max\{RSSI_1, RSSI_2, \dots, RSSI_i\}$ is fulfilled, MC is associated with AP_i , then we determine AP_i is the most preferred $AP_{preferred}$.
- **Step 4:** In the meantime, measure the radio signature of the MC, e.g. RSSI of MC at $AP_1, \dots, AP_i$, e.g. $AP_1_RSSI = -30 \text{ dbm}$, $AP_2_RSSI = -50 \text{ dbm}$, $\dots$, a multiple tuple $(AP_1_RSSI, AP_2_RSSI, \dots, AP_i_RSSI)$ is constructed to represent the raw numeric value of a radio signature.
- **Step 5:** We further process the radio signatures by employing functions according to specific application requirements. We apply fuzzy functions in our implementation since the fuzzification deals with the vagueness of measurement introduced by granularity of subareas, e.g. Radio Signature Process Function1:

$$F_{\text{fuzzy}}(AP_i_RSSI) = \begin{cases} 0 & AP_i_RSSI < 10 \text{ dB} \\ 1 + \log \frac{AP_i_RSSI}{10} & AP_i_RSSI \geq 10 \text{ dB} \end{cases} \quad (1)$$

Without loss of generosity, we use $\mu_{AP_i_RSSI_high}$ to denote the numeric result of the function.

$$\begin{aligned}
F_{\text{fuzzy}}(AP_1_RSSI) &= \mu_{AP_1_RSSI_high} \\
F_{\text{fuzzy}}(AP_2_RSSI) &= \mu_{AP_2_RSSI_high} \\
&\vdots \\
F_{\text{fuzzy}}(AP_i_RSSI) &= \mu_{AP_i_RSSI_high}
\end{aligned}$$

- **Step 6:** We combine the outcomes and the original radio signatures to build a knowledge base.

$(AP_1_RSSI, high, \mu_{AP_1_RSSI \text{ high}})$
 $(AP_2_RSSI, high, \mu_{AP_2_RSSI \text{ high}})$
 $\vdots$
 $(AP_i_RSSI, high, \mu_{AP_i_RSSI \text{ high}})$

e.g. $(AP_1_RSSI, high, \mu_{AP_1_RSSI \text{ high}})$ means the signature of MC heard by AP_1 is $\mu_{AP_1_RSSI} = 0.8$.

- **Step 7:** We have an entry in the knowledge base for the subarea A_i , which is $f KB(A_i) \rightarrow AP_{\text{preferred}}$, e.g. $\{(AP_1_RSSI, high, 0.8), (AP_2_RSSI, high, 0.7), (AP_2_RSSI, high, 0.5)\} \rightarrow AP_{\text{preferred}}$.
- **Step 8:** We move MC to the next position A_2 , repeat the above procedures.
- **Step 9:** After MC traverses the whole area, the knowledge base is built. e.g. $KBD = \{KB(A_1), KB(A_2), \dots, KB(A_i)\}, (1 < i < n)$.

B. Operational Phase

Basically, the operational phase is the reversal process of the training phase. In the operational phase, the real time radio characteristics of the STA is monitored and processed by APs, which is considered as the radio signature on the “crime scene”, the captured radio signature will be compared with the radio signature in the knowledge base. The similarity between the captured one and the records will be measured and the matched radio signature will be identified. In case of no identical match, the most similar one of all records will be identified. Corresponding to the identified radio signature, the most preferred AP to be associated is determined. By this means, the handoff is achieved proactively and without any on-the-fly operation of STA. As illustrated in Fig. 4, the algorithm of the operational stage is listed below:

- **Step 1:** When the 802.11 WMN is put into operation, MT moves in the network. $AP_1, \dots, AP_i$, monitor the movement of MC.
- **Step 2:** At given time i , $AP_1, \dots, AP_i$ measure the radio signature of MC, e.g. when MC is in subarea AP_1 , the measurement is $(AP_1_RSSI, AP_2_RSSI, \dots, AP_i_RSSI)$.
- **Step 3:** We apply the same fuzzy functions as those employed in the training stage for further processing, e.g. $RS(\text{current}) = \{(AP_1_RSSI, high, 0.5), (AP_2_RSSI, high, 0.5), (AP_i_RSSI, high, 0.7)\}$.
- **Step 4:** The radio signature is looked up in the knowledge base, calculate the match degree between the radio signature on the scene and the entries in the knowledge base. If we discover: e.g. $RS(\text{current})(A1) = KB(A1)$, which is, $\{(AP_1_RSSI, high, 0.8), (AP_2_RSSI, high, 0.7), (AP_2_RSSI, high, 0.5)\} \rightarrow AP_{\text{preferred}}$.
- **Step 5:** We conclude that, AP_1 is $AP_{\text{preferred}}$, then instruct the MC to handoff to the AP_1 .

The probing delay is eliminated until this phase. However, after the preferred AP is determined, as explained in Sect. 1, the authentication still takes time to grant the Internet access for the MC. We further present the solution in the next

```

{ //MC moves to a sub-area of the network coverage
  for each sub-area subarea  $A_i \in \text{Area}$  do
  {
    for each  $AP_i$  do
    { //Calculate the real time radio signature.
      RS(current) =  $(AP_1\_RSSI, high, \mu_{AP_1\_RSSI \text{ high}}),$ 
 $(AP_1\_RSSI, high, \mu_{AP_2\_RSSI \text{ high}}), \dots, (AP_i\_RSSI,$ 
 $high, \mu_{AP_i\_RSSI \text{ high}})$ 
    }
  }
  //Look up and compare with each entries in the
  knowledge base and select the one match.
  for each  $KB_i \in KBD$  do
  { //Deal with user preference and compare the real
    time measurement with the stored radio signature.
    if (exist new preferred  $AP_j$ )
    then select new  $AP_j$ .
    else
    if RS (current) ( $A_i$ ) =  $KB(A_k)$ 
    {
       $A_k$  is the  $AP_{\text{preferred}}$ 
    }
  }
}

```

Fig. 4 The algorithm for operational stage.

section.

3.2 Dual Re-authentication Mechanism for Secured Fast Handoff

3.2.1 Overview

As analyzed in the related work, most recent efforts of fast authentication focus on proactive approaches by transferring security context to potential target mesh routers before handoff occurs. However, the efficiency and effectiveness of such approaches are significantly affected by the density of potential targets and the accuracy of handoff prediction which is in fact a very difficult task. So we propose a reactive scheme to reduce re-authentication delay which consists of two steps to exploit the tradeoff between system performance and security.

The fundamental design philosophy of our proposed mechanism is Access-before-Authenticate, which means to grant the user an “immediate” access based on a lightweight authentication while a strong authentication is executed afterwards.

The original authentication process is separated to two phases, namely Immediate Authentication (IA) and Full Authentication (FA). With IA an MC has a particular evidence which can prove it have associated with a trusted AP within a granted period. After that, the new AP allows the MC to access the network temporarily for a certain time of duration. The design of evidence makes it to be validated quickly without any assistance from remote Authentication Server

(AS). Full Authentication (FA) is executed after that. If the strong authentication is failed, the access will be terminated.

3.2.2 IA and FA Procedures

To simplify the implementation, our scheme is based on IEEE 802.11i standard which adopts IEEE 802.1X framework. The FA phase adopts standard IEEE 802.1X authentication. The IA phase uses One-Time Ticket (OTT) as the authentication evidence.

A. One-Time Ticket

OTT provides a method to validate whether an MC has associated with another AP dominated by the same AS before.

In the IEEE 802.11i, when an MC enters the network initially, the MC should conduct IEEE 802.1X authentication to decide whether data access is allowed or not. In our scheme, an OTT is delivered from AS to the MC accompanying the initial IEEE 802.1X authentication. When the MC leaves the current AP's domain, the MC sends its OTT to new AP through OTT-request message. The new AP can compare the OTT with the entries of OTT cache table which it maintains. If the OTT is valid, the new AP trusts the MC for a certain while and initiates a four-way handshake to establish the session keys used for traffic encryption. Otherwise, AP notifies the network access request is denied through OTT-response. The MC has to conduct IEEE 802.1X authentication and another four-way handshake within a given duration contained in the received OTT-response. The OTT is established by hash chain [10].

Let $H_{s,k}^n$ denote a set of hashed keys established by a seed, s . k is a secret key, and n is the number of hashed keys. $H_{s,k}^n$ is formulated by

$$H_{s,k}^n = \{h_i | h_i = \text{HMAC}_k(h_{i-1}), h_0 = s, 0 < i \leq n\} \quad (2)$$

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

B. Immediate Authentication Procedure

A seed s , randomly chosen by AS, has to be shared between authorized 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 authorized APs regularly through *Seed-advertisement* message (P1 in Fig. 5). k is used for securing hash function. Particularly, t_c is the current time stamp used for time synchronization between AS and the authorized AP. In the active mode, when an AP boots up, the AP requests s from the AS through *Seed-solicitation* message.

The AS replies the AP through *Seed-advertisement*. IEEE 802.11i assumes that AP and AS can established a

Fig. 5 Proposed dual re-authentications scheme.

secure channel such as IEEE802.1X, we can assume s is delivered securely. The *Seed-advertisement* contains seed number s , the number of keys n , unit time l , usage starting time of the seed number t_s , and OTT valid duration time t_v .

Since s and k are shared at both APs and AS sides, they can generate common outcome of $H_{s,k}^n$ individually. When an MC attempts to access the network initially, IEEE802.1X authentication and four-way handshake are conducted consequently. While the EAP procedure at the AAA server such as RADIUS server, AS sends the MC the current time t_r and an appropriate OTT (P2 in Fig. 5). It is the same as IEEE 802.1X authentication and four-way handshake procedure except an OTT attached inside.

The MC should keep the received OTT and t_r . When the MC leaves from the current serving AP, the MC starts to probe available wireless channels and associates with a new AP, AP_n . In order to reduce the re-authentication delay, the MC sends its OTT to AP_n through *OTT-request* (P3 in Fig. 5). Since the wireless channel between MC and AP is insecure, $\{f_h(t_r || ID || OTT), t_r, ID\}$ are transmitted from MC to AP_n as an *OTT-request*, where f_h is a one-way hash function such as MD5 and SHA-1. ID , and the MC's identity such as MAC address, are used for preventing attacks such as replay and man-in-the-middle. When receiving the *OTT-request*, AP_n is able to check immediately whether the MC has ever associated with other APs also dominated by the same AS through comparing the validation of received OTT to the OTT cache table maintained itself (P4 in Fig. 5). If the OTT from the MC is valid, AP_n notifies MC's admission through *OTT-response* (P5 in Fig. 5), and initiates a four-way handshake to establish session key used for data traffic encryption (P6 in Fig. 5).

The session keys are derived from the pairwise mas-

ter key (PMK), Authenticator address (AA), Supplicant address (SPA), Authenticator nonce (ANonce), and Supplicant Nonce (SNonce). IEEE 802.11i also specifies a 256 bit Pre-Shared Key (PSK) shared beforehand between AP and MC, which derives PMK through Pseudo Random Function (PRF) without AS. IA uses the OTT as a PSK. Consequently, IA requires no modifications to the four-way handshake procedures in IEEE 802.11i. After P6 is completed, the MC is able to access the network on secure channel (D3) established by the procedure P6. by that way, the re-authentication delay of our proposal IA is only from P3 to P6. Within these procedures, there are no information exchanging required therefore packets exchanged with AP_o or AS so far. Therefore, no packets transmission are needed. Conventional IEEE 802.1X authentication and four-way handshake will be accomplished within a given expired time for restraining IA's optimism (P7 in Fig. 5). After that, OTT of MC is updated (P8 in Fig. 5). AS should broadcast *Seed-advertisement* within the Ticket Grant Lifetime (P9 in Fig. 5).

C. OTT Validation

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

$$\begin{aligned} \widehat{H_{t_r, ID}} &= \{\xi_i | \xi_i = f_h(t_r || ID || h_i), \\ h_i &\in H_{s,k}^n, \zeta_l \leq n - i < \zeta_h \} \end{aligned} \quad (3)$$

where $\zeta_l = \max(0, \lfloor (\tilde{t}_c - t_v - t_s) \times l^{-1} \rfloor)$, $\zeta_h = \lfloor (\tilde{t}_c - t_s) \times l^{-1} \rfloor$ and $\tilde{t}_c$ is the current time. Let h_o be $f_h(t_r || ID || OTT)$ contained in *OTT-request* message. If $h_o \in \widehat{H_{t_r, ID}}$, this *OTT-request* is valid, thus AP notifies MC is allowed to access the network through *OTT-response* (P5 in Fig. 5). *OTT-response* contains the time (allowed FA guaranteed duration) within which indicates the MC must accomplish conventional IEEE 802.1X authentication and four-way handshake. Otherwise, it is probably that the MN is not a normal user such as an abused reusing cached old OTTs (replay or man-in-the-middle attacks). The AP can adjust an allowed FA time bound to correspondent OTT in order to mitigate the affect of OTT abuse. Even if OTT is revealed to attackers, the access of the network is only within the allowed FA guaranteed duration.

4. Experiments and Evaluation

In order to investigate the feasibility and effectiveness of the proposed mechanisms, we implement them in our HAWK integrated experiment environment which is consisted with prototype testbed and simulation tools to test and evaluate the system performance:

4.1 Layer 2 Fast Handoff Delay

We compare the end-to-end delay of a VOIP application considering two cases without and with our proposed NRS scheme in the same simulation scenario. As illustrated in

Fig. 6 The latency due to handoff delay without vs. with employing NRS scheme.

Fig. 1, an MC roaming through three neighboring APs connecting through wireless mesh backbone. Three APs collaboratively monitor the radio signature of MC, the signature is compared with the entries recorded in the knowledge base which has been prepared. The preferred AP is calculated and stored in the list for MC to associate in case handoff will be initiated.

We run a VOIP application on the local MC and the remote correspondent. We arbitrarily make MC to hand-off by setting handoff threshold and observe the time delay. We keep measuring the end-to-end delay of packets between MC and the remote correspondent. When we manipulate Layer 2 handoff to occur by moving MC, the end-to-end delay will vary and indicate the start and the end of the handoff. Referring to Fig. 6, we can see that our proposed NRS scheme can reduce the time-consuming scanning delay during the handoff procedure greatly. Under the normal circumstance without NRS scheme, the delay is around 150 ms. By enabling NRS scheme, the total handoff delay is shortened to 50 ms, which are accumulated by open authentication and association excluding probing. It is obvious that the proposed mechanism works. However, the accuracy of the handoff becomes our concern. The accuracy issue of NRS schemes majorly comes from the dividing the area into subareas. The granularity of the subarea determines whether a handoff occurs according to the training phase. If a handoff occurs in the operational phase while it is not expected in the training phase, we consider it inaccurate. In general, smaller the granularity is, more unnecessary hand-off occurs. Nevertheless, if the granularity is too large, it loses the meaning of handoff in the first place and only an ideal AP with the unlimited coverage can support that. We use false negative (FN) and the false positive (FP) rate to represent the accuracy of the proposed NRS scheme. We apply the two-ray ground reflection model as the wireless channel propagation representation in our experiments. We set the size of the whole area is 100 m × 100 m. As the size of the subarea which is a square as the basic union of territory increases, the FN in Fig. 7 and FP in Fig. 8 go to their theoretical values fluctuating around 21% and 36% respectively. We can notice that when the size of subarea goes over 10 m, the rate of FN and FP are relatively stable which can be set as a recommended system parameter of our proposed NRS

Fig. 7 The false positive rate of handoff decision.

Fig. 8 The false negative rate of handoff decision.

scheme in implementation.

4.2 Dual Re-Authentication

This section provides performance analysis of the proposed scheme in terms of authentication delay. We also make a quantification model to measure the degree of security for our proposal.

4.2.1 Handoff Delay

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

$$T_r = T_{reassoc} + T_{auth} + T_{fwh} \quad (4)$$

where $T_{reassoc}$ is a delay for re-association, and T_{fwh} is a delay for four-way handshake. In IEEE 802.11i, T_{auth} is a delay for IEEE 802.1X authentication. In the proposed scheme, T_{auth} denotes operation time for OTT validation in Sect. 2. C. T_{auth} is determined by t_v and l . The total possible number of keys assigned to each MC per AP is formulated by $\eta = \frac{l}{t_v}$. Since T_{auth} is the time to conduct a linear

searching based on hashing and comparing, T_{auth} is stochastically proportional to η . The OTT-Validation takes $O(\eta)$ times. The expectation of T_{auth} is defined as

$$E[T_{auth}] = (\eta \times T_h)/2 \quad (5)$$

where T_h is the time for f_h computation process time. Our analysis adopts SHA-1 as f_h . T_h is measured by Benchmarks of Crypto++ 5.5 Library [19]. Conclusively, T_r , re-authentication delay in our scheme can be represented as a function of η :

$$f(\eta) = T_r \quad (6)$$

where $f(\eta)$ is the re-authentication delay function. Given that delays for re-association, IEEE 802.1X, and four-way handshake average 2, 250, and 60 ms, respectively, re-authentication delay in the conventional IEEE 802.11i is totally 312 ms. 250 ms for IEEE 802.1X authentication is measured in fast resume mode [18]. Although full IEEE 802.1X authentication takes 1000 ms, we compares with the fixed version for unbiased comparison.

4.2.2 Quantification for the Degree of Security

As l is increased, the probability that the same OTT is assigned to a number of MCs are increased. Let λ_a denote the total arrival rate, i.e., the total association rate. λ_a is defined as $\lambda_1 + \lambda_2$. Let μ_d denote the total departure rate. μ_d is defined as $\mu_1 + \mu_2$. Let $P_l(N = n)$ denote the probability that the number of associated MNs are n within a period l . The probability that the same OTT is assigned to a number of MCs before handoff is equal to the probability that at least two MCs arrive within a period l , i.e., $P_l(N \geq 2)$.

$$P_l(N = 0) = 1 - \lambda_a e^{-\lambda_a l} \quad (7)$$

$$- \int_0^l \lambda_a e^{-\lambda_a t} \times \mu_d e^{-\mu_d(l-t)} dt \quad (8)$$

where the integral part indicates the case that an MC leaves within a period $l - t$, even though arrives at $t(\leq l)$.

$$P_l(N = 1) = \lambda_a e^{-\lambda_a l} \quad (9)$$

$$P_l(N \geq 2) = 1 - P_l(N = 0) - P_l(N = 1) \\ = \int_0^l \lambda_a e^{-\lambda_a t} \times \mu_d e^{-\mu_d(l-t)} dt. \quad (10)$$

The duplicated OTT problem during initial association does not threaten the security of network system critically, because OTT is used after handoff. The case of after-handoff duplication should be considered.

If MN_0 moves into AP_n , the probability that MCs assigned the same OTT to MN_0 exist is

$$p_d = 1 - \left(1 - \frac{1}{\eta}\right)^{m_n} \quad (11)$$

This implies a tradeoff between OTT acquisition speed and security. The reduced re-authentication delay decreases

Fig. 9 Probability of existence of duplicated OTTs after handoff.

the degree of security measured by p_d , comparing with original IEEE 802.11i. Additionally, even though OTT is revealed to a later-coming MC, since the hash function for OTT chain generation varies depending on a secret key k , the hash chain is difficult to be derived by an adversary. Figure 9 shows the relationship between p_d and m_n in terms of different η . The black dot indicates that η is 180. Likewise, the white dot indicates η is 360. In Fig. 9, we find out that the probability of OTT duplication is decreased faster, as η increasing. p_d is also defined as a function as follows:

$$g(m_n, \eta) = p_d \quad (12)$$

where $g(m_n, \eta)$ is a security measure function.

5. Conclusion

In this paper, we present the detailed mechanisms, simulation, implementation and experimental results of a total solution for secured fast handoff for Internet Access. Network-assisted Radio Signature is a novel Layer 2 fast handoff scheme to realize proactive probing by making the most of the static topology of an actual WMN and the Dual Re-authentication mechanism improves the security of optimistic access by providing one time ticket while effectively reduces the authentication delay. To our best knowledge, we are the first to achieve such an optimal performance of secured fast handoff at no cost of extra hardware or overhead, comparing with other pioneering projects.

The detailed information about our testbed system is also available at the project website: wmn.comp.polyu.edu.hk.

Acknowledgements

This work is supported by Innovation and Technology Funding of Hong Kong (Grant No: GHP/049/06).

References

- [1] CCITT, 1988, International Telecommunication Union, General Characteristics of International Telephone Connections and International Telephone Circuits.

- [2] A. Mishra, M. Shin, and W. Arbaugh, "An empirical analysis of the IEEE 802.11 MAC layer handoff process," *ACM SIGCOMM Computer Communications Review*, vol.33, no.2, pp.93–102, April 2003.
- [3] H. Velayos and G. Karlsson, "Techniques to reduce the IEEE 802.11b handoff time," *Proc. IEEE ICC 2004*, pp.3844–3848, Paris, France, June 2004.
- [4] K. Kwon and C. Lee, "A fast handoff algorithm using intelligent channel scan for IEEE 802.11 WLANs," *Proc. ICACT 2004*, pp.46–50, Phoenix Park, Republic of Korea, Feb. 2004.
- [5] I. Ramani and S. Savage, "SyncScan: Practical fast handoff for 802.11 infrastructure networks," *Proc. IEEE INFOCOM 2005*, pp.675–684, Miami, FL, USA, March 2005.
- [6] S. Pack, J. Choi, T. Kwon, and Y. Choi, "Fast-handoff support in IEEE 802.11 wireless networks," *IEEE Communications Surveys & Tutorials*, vol.9, no.1, pp.2–12, Jan. 2007.
- [7] S. Shin, A.G. Forte, A.S. Rawat, and H. Schelzrinne, "Reducing MAC layer handoff latency in IEEE 802.11 wireless LANs," *Proc. ACM MobiWac'04*, pp.19–26, Philadelphia, PA, USA, Oct. 2004.
- [8] M. Shin, A. Mishra, and W.A. Arbaugh, "Improving the latency of 802.11 hand-offs using neighbor graphs," *Proc. ACM MobiSys'04*, pp.70–83, Boston, MA, USA, June 2004.
- [9] D. Johnson, C. Perkins, and J. Arkko, "Mobility support in IPv6," *IETF RFC3775*, June 2004.
- [10] R. Ramjee, K. Varadhan, L. Salgarelli, S.R. Thuel, S.-Y. Wang, and T. La Porta, "HAWAII: A domain-based approach for supporting mobility in wide-area wireless networks," *IEEE/ACM Trans. Netw.*, vol.10, no.3, pp.396–410, June 2002.
- [11] P. Xu, J. Liao, X. Wen, and X. Zhu, "Optimized integrated registration procedure of mobile IP and SIP with AAA operations," *20th Int'l. Conf. Advanced Info. Networking and Applications*, vol.1 (AINA'06), pp.926–931, 2006.
- [12] A. Dutta, T. Zhang, Y. Ohba, K. Taniuchi, and H. Schulzrinne, "MPA assisted optimized proactive handoff scheme," *Proc. 2nd Annual Int'l. Conf. Mobile and Ubiquitous Systems: Networking and Services 2005 (MobiQuitous 2005)*, pp.155–165, July 2005.
- [13] T. Kwon, M. Gerla, and S. Das, "Mobility management for VoIP: Mobile IP vs. SIP," *IEEE Wireless Commun. Mag.*, vol.9, no.5, pp.66–75, Oct. 2002.
- [14] M. Georgiades, N. Akhtar, C. Politis, and R. Tafazolli, "AAA context transfer for seamless and secure multimedia services over All-IP infrastructures," *5th European Wireless Conf.*, Spain, Feb. 2004.
- [15] T. Braun and K. Hahnsang, "Efficient authentication and authorization of mobile users based on peer-to-peer network mechanisms," *Proc. 38th Annual Hawaii Int'l. Conf. System Sciences (HICSS '05)*, p.3066, Jan. 2005.
- [16] T. Aura and M. Roe, "Reducing reauthentication delay in wireless networks," *1st Int'l. Conf. Security and Privacy for Emerging Areas in Commun. Networks (SecureComm 2005)*, pp.139–148, Sept. 2005.
- [17] Brik, V. Mishra, and S. Banerjee, "Eliminating handoff latencies in 802.11 WLANs using multiple radios: Applications, experience, and evaluation," *Proc. ACM Internet Measurement Conf. 2005*, pp.27–37, Oct. 2005.
- [18] Y. Yan, J. Cao, C. Liu, S.W. Kim, and W. Wu, "A dual re-authentication scheme for fast handoff in IEEE 802.11 wireless mesh networks," *Wireless Communications and Networking Conference, 2009. WCNC 2009*, pp.1–5, April 2009.
- [19] Crypto++ 5.5 Benchmarks. [Online]. Available: <http://www.cryptopp.com/benchmarks.html>

Gang Yao is a Ph.D. student in the Department of Computing at the Hong Kong Polytechnic University. His research interests include seamless communication and mobility management, mobile computing, wireless networks and mobile positioning. Gang Yao received his BSc in computer science from Wuhan University.

Jiannong Cao is a professor in the Department of Computing at Hong Kong Polytechnic University. His research interests include parallel and distributed computing, networking, mobile computing, fault tolerance, and distributed software architecture and tools. Cao received his BSc from Nanjing University, China, and his MSc and Ph.D. from Washington State University.

Ye Yan is a Ph.D. student in Department of Computer & Electronics Engineering at the University of Nebraska-Lincoln. His research interests include network architecture, protocols, traffic security and performance analysis for quality of service provisioning in wireless networks. Yan received his B.Eng. and M.Eng. in Jimei University in China and Chonbuk National University in South Korea, respectively. He is a student member of IEEE.

Yusheng Ji received the B.Eng., M.Eng. and D.Eng. in Electrical Engineering from the University of Tokyo in 1984, 1986, and 1989, respectively. She joined the National Center for Science Information Systems in 1990. She is currently an Associate Professor at the National Institute of Informatics and the Graduate University. Her research interests include network architecture, protocols, traffic controls, and performance analysis for quality of service provisioning in wired and wireless networks. She is a

member of IPSJ and IEEE.