A Key Caching Mechanism for Reducing WiMAX Authentication Cost in Handoff

Shih-Feng Hsu and Yi-Bing Lin, *Fellow, IEEE*

***Abstract*—The IEEE 802.1X is utilized in mobile Worldwide Interoperability for Microwave Access (WiMAX) authentication. This procedure incurs a long delay in WiMAX handoff. To resolve this issue, this paper proposes a key caching mechanism to eliminate the nonnecessary IEEE 802.1X authentication cost in WiMAX handoff. This mechanism is investigated through analytic and simulation modeling. Our study indicates that the key caching scheme can effectively speed up the handoff process.**

***Index Terms*—Authentication, authorization, and accounting (AAA), handoff, mobile Worldwide Interoperability for Microwave Access (WiMAX).**

I. Introduction

THE IEEE 802.16e mobile *Worldwide Interoperability for Microwave Access* (WiMAX) provides broadband wireless services with wide service coverage, high data throughput, and high mobility. To support security network access, the *authentication, authorization, and accounting* (AAA) mechanism is exercised in WiMAX [3]. Fig. 1 shows the AAA architecture and protocol stack for WiMAX. In this architecture, the *access service network* [see ASN; Fig. 1 (2)] consists of *base stations* [see BSs; Fig. 1 (4)] and *ASN gateways* [see ASN-GWs; Fig. 1 (5)]. An ASN-GW controls several BSs. A BS provides WiMAX radio access for *mobile stations* [see MSs; Fig. 1 (1)] after the MSs are authenticated by the AAA server [see Fig. 1 (6)] in the *connectivity service network* [see CSN; Fig. 1 (3)]. In the WiMAX AAA architecture, the ASN-GW serves as the authenticator for the MS. The authenticator is responsible for forwarding authentication messages between the MS and the AAA server and for maintaining the MS-related information (e.g., encryption keys) after authentication. We assume that the *subscriber identity module* (SIM)-based *extensible authentication protocol* (EAP) is utilized for AAA [1]. Note that this approach reuses the authentication mechanism in mobile telecommunications [6]. In the authentication procedure, an EAP-SIM message [see Fig. 1(a)] is encapsulated in an EAP message [see Fig. 1(b)]. The MS then encapsulates the EAP message in *privacy key management protocol version 2* [see PKMv2; Fig. 1(c)] before it is transmitted to the BS. The BS exercises the *authentication relay protocol* [see AuthRelay; Fig. 1(d)] to forward the received EAP message to the authenticator (i.e., the ASN-GW). Upon receipt of an EAP message, the authenticator translates it into a *remote authentication dial-in user service* [see RADIUS; Fig. 1(e)] message. Then, the RADIUS message is sent to the AAA server. Upon receipt of the RADIUS message, the AAA server utilizes the *mobile application part* [see MAP; Fig. 1(f)] of the *signaling system number 7* [see SS7; Fig. 1(g)] protocol to communicate with the *home location register* (HLR)/*authentication center* (AuC) [see Fig. 1 (7)]. The HLR is the mobility database of the GSM/UMTS mobile telecommunication networks [2], [6]. The AuC maintains the secret keys of the MSs and provides the authentication information to the AAA server.

II. WiMAX Initial Network Entry Process

By using the protocols described in Fig. 1, the WiMAX authentication works as follows. Supposing that an MS first connects to the WiMAX network, the following steps are executed for the initial network entry process (see Fig. 2).

Step 1) The MS, the BS, and the ASN-GW (authenticator) negotiate the security policy (i.e., to select the encryption and decryption algorithms) and the authorization policy, specifically to select the *message-authentication-code* (MAC) type.

Step 2) The authenticator sends an EAP request message to the MS. This message initiates the IEEE 802.1X authentication procedure by requesting the user identity.

Steps 3) and 4) The MS replies with an EAP response message with the user identity to the authenticator. The user identity consists of two elements: the AAA server address AAA-addr and the user account User-acct. In the SIM-based EAP authentication, the user account is set to the *international mobile subscriber identity* (IMSI) of the MS [2]. According to AAA-addr, the authenticator forwards the EAP response message to the AAA server.

Step 5) Upon receipt of the user identity, the AAA server performs the SIM-based EAP authentication with the MS as follows.

Step 5.1) The AAA server issues an EAP request message with type “Start” to the MS.

Step 5.2) The MS replies with the EAP response message containing a random number, i.e., MS-RAND. This random number is used to derive the encryption keys in Steps 5.3 and 5.5 below.

Manuscript received September 27, 2008; revised January 23, 2009 and March 14, 2009. First published May 2, 2009; current version published October 2, 2009. The work of Y.-B Lin was supported in part by the National Science Council, Taiwan, under Grant NSC 97-2221-E-009-143-MY3 and Grant NSC 97-2219-E-009-016, by Intel, by Chunghwa Telecom, by the Industrial Technology Research Institute/National Chiao Tung University Joint Research Center, and by the Ministry of Education under the Aim for Top University project. The review of this paper was coordinated by Prof. J. Li.

S.-F. Hsu is with the Department of Computer Science, National Chiao Tung University, Hsinchu 300, Taiwan (e-mail: sfhsu@pcs.csie.nctu.edu.tw).

Y.-B. Lin is with the Department of Computer Science, National Chiao Tung University, Hsinchu 300, Taiwan, and also with the Institute of Information Science, Academia Sinica, Taipei 115, Taiwan (e-mail: liny@csie.nctu.edu.tw).

Digital Object Identifier 10.1109/TVT.2009.2022079

0018-9545/$26.00 © 2009 IEEE

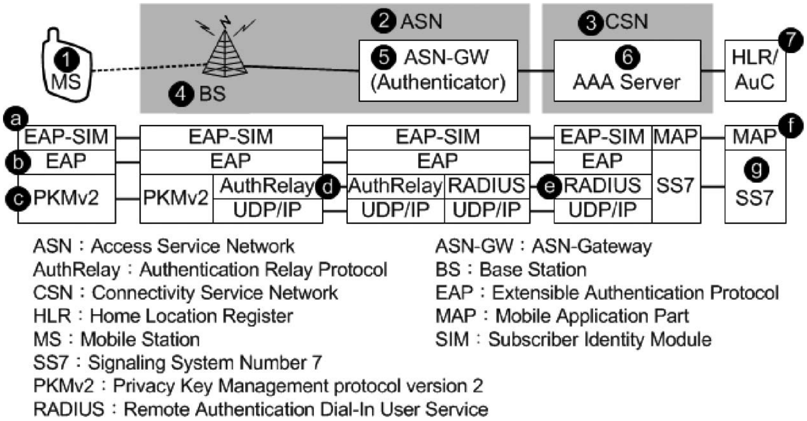

1 MS
2 ASN
4 BS
5 ASN-GW (Authenticator)
3 CSN
6 AAA Server
7 HLR/AuC
a EAP-SIM
b EAP
c PKMv2
EAP-SIM
EAP
PKMv2
AuthRelay
UDP/IP
d
EAP-SIM
EAP
AuthRelay
UDP/IP
RADIUS
UDP/IP
e
EAP-SIM
EAP
RADIUS
UDP/IP
MAP
SS7
f MAP
g SS7
ASN : Access Service Network
AuthRelay : Authentication Relay Protocol
CSN : Connectivity Service Network
HLR : Home Location Register
MS : Mobile Station
SS7 : Signaling System Number 7
PKMv2 : Privacy Key Management protocol version 2
RADIUS : Remote Authentication Dial-In User Service
ASN-GW : ASN-Gateway
BS : Base Station
EAP : Extensible Authentication Protocol
MAP : Mobile Application Part
SIM : Subscriber Identity Module

Fig. 1. WiMAX AAA architecture and protocol stack.

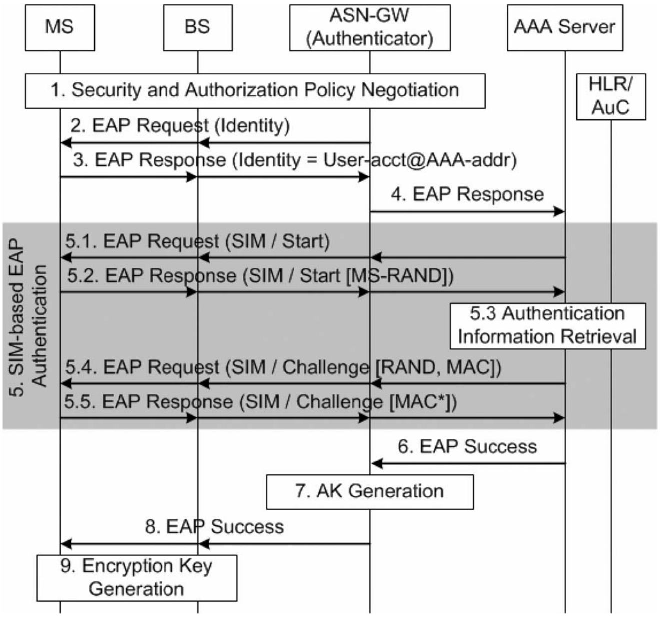

MS
BS
ASN-GW (Authenticator)
AAA Server
HLR/ AuC
1. Security and Authorization Policy Negotiation
2. EAP Request (Identity)
3. EAP Response (Identity = User-acct@AAA-addr)
4. EAP Response
5. SIM-based EAP Authentication
5.1. EAP Request (SIM / Start)
5.2. EAP Response (SIM / Start [MS-RAND])
5.3 Authentication Information Retrieval
5.4. EAP Request (SIM / Challenge [RAND, MAC])
5.5. EAP Response (SIM / Challenge [MAC*])
6. EAP Success
7. AK Generation
8. EAP Success
9. Encryption Key Generation

Fig. 2. WiMAX initial network entry process.

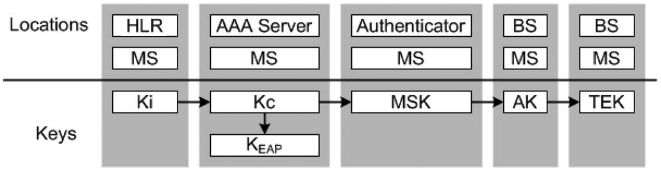


Fig. 3. WiMAX key derivation tree.

Step 5.3) Based on the IMSI received in Step 4), the AAA server communicates with the HLR/AuC to obtain the authentication information, including a random number RAND, a signed result SRES, and a cipher key Kc. Both the MS and the HLR/AuC utilize the RAND and the secret key Ki (stored in the SIM card and the HLR/AuC) to execute the A3 and A8 algorithms to derive the signed result SRES and the cipher key Kc [9]. Then, the AAA server utilizes Kc and MS-RAND (received in Step 5.2) to derive the *master session key* (MSK) and the EAP integrity key $\mathrm{K}_{\mathrm{EAP}}$.

Step 5.4) The AAA server sends a challenge EAP request message with the RAND and the MAC. This MAC is derived from $\mathrm{K}_{\mathrm{EAP}}$ and is used to ensure the integrity of this message.

Step 5.5) Upon receipt of the EAP request message, the MS utilizes RAND, MS-RAND (generated in Step 5.2), and Ki (stored in the SIM card) to generate SRES*, Kc, MSK, and $\mathrm{K}_{\mathrm{EAP}}$. With $\mathrm{K}_{\mathrm{EAP}}$ and the received RAND, the MS verifies the received MAC. If the MAC is correct, the AAA server is successfully authenticated by the MS. Then, the MS replies with a challenge EAP response message with a code MAC* derived from $\mathrm{K}_{\mathrm{EAP}}$ and SRES*.

Step 6) The AAA server verifies MAC* by using $\mathrm{K}_{\mathrm{EAP}}$ (generated in Step 5.3) and SRES (received in Step 5.3). If MAC* is correct, the MS is successfully authenticated by the AAA server. The AAA server sends the EAP success message to the authenticator containing MSK (generated in Step 5.3), the MSK lifetime, and the MS authorization profile (e.g., service restrictions and supplementary services). The MSK lifetime is the period that the MS is authorized to access the ASN-GW. When the MSK lifetime has expired, the MS should execute the IEEE 802.1X authentication with the AAA server again.

Step 7) The ASN-GW stores MSK, the MSK lifetime, and the authorization profile. Then, it derives the *authentication key* (AK) by using the MSK and the BS address. This AK is shared between the MS and the BS.

Step 8) The ASN-GW forwards the EAP success message to the BS with AK. The BS passes the EAP success message to inform the MS that the authentication is successful. Upon receipt of this message, the MS generates its version of AK.

Step 9) The BS generates the final encryption key, i.e., the *traffic encryption key* (TEK). This encryption key is used to provide data integrity and confidentiality for a communication session between the MS and the BS. The BS passes the generated TEK (encrypted by AK) to the MS.

The relationship of WiMAX encryption keys and the locations maintaining these keys are shown in Fig. 3.

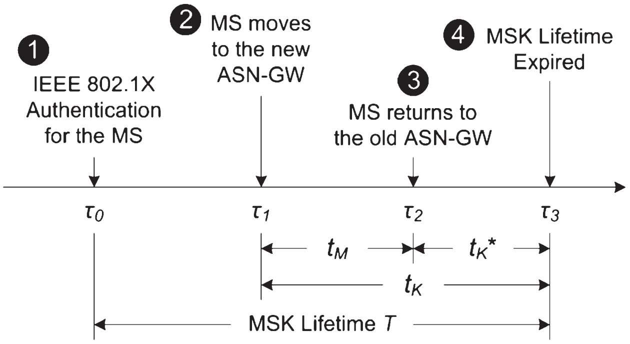

1 IEEE 802.1X Authentication for the MS
2 MS moves to the new ASN-GW
3 MS returns to the old ASN-GW
4 MSK Lifetime Expired
τ_0
τ_1
τ_2
τ_3
t_M
t_K^*
t_K
MSK Lifetime T

Fig. 4. Relationship of the MSK lifetime and the MS movement.

If the MS moves from the old BS to the new BS connecting to a different authenticator (ASN-GW), a new MSK must be generated in this inter-ASN-GW handoff process, which is the same as the initial network entry process described in Fig. 2. In this case, the authenticator (ASN-GW) of the old BS will remove the MS key record (i.e., MSK, the MSK lifetime, and the MS authorization profile). When the MS moves back to the old ASN-GW again, another inter-ASN-GW handoff process should be performed, which may incur a long delay.

III. Key Caching Mechanism

To speed up the inter-ASN-GW handoff process, we propose a *key caching* mechanism. The idea is simple: When the MS moves from the old ASN-GW to the new ASN-GW, the old ASN-GW still keeps the MS key record. If the MS returns to the old ASN-GW before the MSK lifetime expires, it can reuse the MSK without executing the IEEE 802.1X authentication. That is, only Steps 1) and 9) in Fig. 2 are executed to speed up the inter-ASN-GW handoff process. In Fig. 2, Step 1) contains two message exchanges, and Step 9) contains five message exchanges [3]. Therefore, the caching mechanism speeds up the process by saving 50% $(= 7/14)$ of the message exchanges between the MS and the BS.

Although the key caching mechanism may effectively avoid the execution of IEEE 802.1X authentication, it consumes extra storage to keep the MS key records at the old ASN-GW, where a stored key record includes 512 or 1024 bits for the MSK, 32 bits for the MSK lifetime, and 512 or 1024 bits for the MS authorization profiles. Therefore, it is desirable to select an appropriate MSK lifetime to eliminate the IEEE 802.1X authentication without consuming too much extra storage in the ASN-GW. We investigate the effect of the MSK lifetime on the caching performance by an analytic model described below.

Fig. 4 illustrates the relationship between the movement of an MS and its MSK lifetime. In this figure, the IEEE 802.1X authentication is executed at time τ_0 [see Fig. 4 (1)], and the MSK lifetime expires at time τ_3 [see Fig. 4 (4)]. At time τ_1 [see Fig. 4 (2)], the MS moves from the old ASN-GW to the new ASN-GW. The residual MSK lifetime is $t_K = \tau_3 - \tau_1$. If the MS does not return to the old ASN-GW before the MSK lifetime expires, we call this t_K period the unused key period. At time τ_2 [see Fig. 4 (3)], the MS returns to the old ASN-GW. Let $t_M = \tau_2 - \tau_1$ be the period between when the MS leaves the old ASN-GW and when it returns. If the MS returns before the MSK lifetime expires, the MS can reuse the MSK for period $t_K^* = t_K - t_M$ without executing the IEEE 802.1X authentication. Period t_K^* is referred to as the reused key period. We make the following assumptions.

1) We consider two distributions for the MSK lifetime T. That is, T is either an exponential period with rate μ or a fixed period.
2) The MS residence time t_M in new ASN-GWs has the density function $f(t_M)$ with mean $1/\lambda$ and variance V_M.

Three output measures are evaluated in our study:

1) α: the probability that the MS returns to the old ASN-GW before the MSK lifetime expires;
2) $E[t_K|t_M \geq t_K]$: the expected unused key period under the condition that the MS does not return to the old ASN-GW before the MSK lifetime expires (therefore, the cached MSK will not be reused);
3) $E[t_K^*|t_M \leq t_K]$: the expected reused key period under the condition that the MS returns to the old ASN-GW before the MSK lifetime expires (the cached MSK is reused).

We derive the above output measures for exponentially distributed t_M with fixed T and then generalize the derivation for generally distributed t_M with exponentially distributed T.

A. Derivation for Exponentially Distributed t_M and Fixed T

Suppose that the departure of the MS from the old ASN-GW is a random observer to the MSK lifetime. For the fixed MSK lifetime T, from the residual life theorem [4], t_K has a uniform distribution over $0 \leq t_K \leq T$. Then, α is derived as

$$\begin{aligned}\alpha &= Pr[t_M \leq t_K] \\ &= \int_{t_K=0}^{T} \left(\frac{1}{T}\right) \times \left(\int_{t_M=0}^{t_K} \lambda e^{-\lambda t_M} dt_M\right) dt_K \\ &= \frac{e^{-\lambda T} + \lambda T - 1}{\lambda T}. \end{aligned} \tag{1}$$

$E[t_K|t_M \geq t_K]$ is expressed as

$$E[t_K|t_M \geq t_K] = \frac{E[t_K \text{ and } t_M \geq t_K]}{Pr[t_M \geq t_K]} \tag{2}$$

where

$$\begin{aligned} &E[t_K \text{ and } t_M \geq t_K] \\ &\quad= \int_{t_M=0}^{T} \lambda e^{-\lambda t_M} \times \left[\left(\int_{t_K=0}^{t_M} t_K \times \left(\frac{1}{T}\right) dt_K\right)\right] dt_M \\ &\qquad+ \int_{t_M=T}^{\infty} \lambda e^{-\lambda t_M} \times \left[\left(\int_{t_K=0}^{T} t_K \times \left(\frac{1}{T}\right) dt_K\right)\right] dt_M \\ &\quad= \frac{1-e^{-\lambda T}}{\lambda^2 T} - \frac{e^{-\lambda T}}{\lambda}. \end{aligned} \tag{3}$$

From (1)–(3), we have

$$\begin{aligned} E[t_K|t_M \geq t_K] &= \frac{E[t_K \text{ and } t_M \geq t_K]}{Pr[t_M \geq t_K]} \\ &= \left(\frac{1-e^{-\lambda T}}{\lambda^2 T} - \frac{e^{-\lambda T}}{\lambda}\right) \times \left(\frac{1}{1-\alpha}\right) \\ &= \frac{1}{\lambda} - \frac{Te^{-\lambda T}}{1-e^{-\lambda T}}. \end{aligned} \tag{4}$$

Similarly, $E[t_K^*|t_M \leq t_K]$ is expressed as

$$E\left[t_K^*|t_M \leq t_K\right] = \frac{E\left[t_K^* \text{ and } t_M \leq t_K\right]}{Pr[t_M \leq t_K]} \tag{5}$$

where

$$\begin{aligned} &E\left[t_K^* \text{ and } t_M \leq t_K\right] \\ &\quad= \int_{t_K=0}^{T} \left(\frac{1}{T}\right) \times \left[\int_{t_M=0}^{t_K} (t_k - t_M)\lambda e^{-\lambda t_M} dt_M\right] dt_K \\ &\quad= \frac{T}{2} - \frac{1}{\lambda} + \frac{1-e^{-\lambda T}}{\lambda^2 T}. \end{aligned} \tag{6}$$

From (1), (5), and (6), we have

$$\begin{aligned} E\left[t_k^*|t_M \leq t_K\right] &= \left(\frac{T}{2} - \frac{1}{\lambda} + \frac{1-e^{\lambda T}}{\lambda^2 T}\right) \times \left(\frac{1}{\alpha}\right) \\ &= \frac{\lambda T^2}{2(\lambda T + e^{-\lambda T} - 1)} - \frac{1}{\lambda}. \end{aligned} \tag{7}$$

B. Derivation for Generally Distributed t_M and Exponential T

Since the departure of the MS from the old ASN-GW is a random observer to the MSK lifetime, from the residual life theorem, t_K is exponentially distributed with mean $E[T] = 1/\mu$. Let t_M have an arbitrary distribution with density function $f(t_M)$ and Laplace transform $f^*(s)$. Then, α is derived as

$$\alpha = \int_{t_K=0}^{\infty} \mu e^{-\mu t_K} \times \left[\int_{t_M=0}^{t_K} f(t_M) dt_M\right] dt_K = f^*(\mu). \tag{8}$$

$E[t_K \text{ and } t_M \geq t_K]$ is expressed as

$$\begin{aligned} &E[t_K \text{ and } t_M \geq t_K] \\ &\quad= \int_{t_M=0}^{\infty} f(t_M) \times \left(\int_{t_K=0}^{t_M} t_K \mu e^{-\mu t_K} dt_K\right) dt_M \\ &\quad= \frac{1}{\mu} + \left[\frac{df^*(s)}{ds}\right]\Bigg|_{s=\mu} - \frac{f^*(\mu)}{\mu}. \end{aligned} \tag{9}$$

From (2), (8), and (9), we have

$$\begin{aligned} &E[t_K|t_M \geq t_K] \\ &\quad= \frac{E[t_K \text{ and } t_M \geq t_K]}{Pr[t_M \geq t_K]} \\ &\quad= \left\{\frac{1}{\mu} + \left[\frac{df^*(s)}{ds}\right]\Bigg|_{s=u} - \frac{f^*(\mu)}{\mu}\right\} \times \left[\frac{1}{1-f^*(\mu)}\right]. \end{aligned} \tag{10}$$

$E[t_K^* \text{ and } t_M \leq t_K]$ is derived as

$$\begin{aligned} &E\left[t_K^* \text{ and } t_M \leq t_K\right] \\ &\quad= \int_{t_K=0}^{\infty} \mu e^{-\mu t_K} \times \left[\int_{t_M=0}^{t_K} (t_k - t_M) f(t_M) dt_M\right] dt_K \\ &\quad= \frac{f^*(\mu)}{\mu}. \end{aligned} \tag{11}$$

Therefore, from (5), (8), and (11), we have

$$\begin{aligned} E\left[t_K^*|t_M \leq t_K\right] &= \frac{E\left[t_K^* \text{ and } t_M \leq t_K\right]}{Pr[t_M \leq t_K]} \\ &= \left[\frac{f^*(\mu)}{\mu}\right] \times \left[\frac{1}{f^*(\mu)}\right] = \frac{1}{\mu}. \end{aligned} \tag{12}$$

Equation (12) says that $E[t_K^*|t_M \leq t_K]$ is not affected by the t_M distribution.

To further investigate (8) and (10), we assume that t_M has a Gamma distribution, which has been used in telecommunication modeling [5], [7]. The Gamma-distributed t_M has mean $1/\lambda$, variance V_M, and Laplace transform

$$f^*(s) = \left(\frac{1}{\lambda V_M s + 1}\right)^{\frac{1}{\lambda^2 V_M}}.$$

Then, from (8)

$$\alpha = f^*(\mu) = \left(\frac{1}{\lambda \mu V_M + 1}\right)^{\frac{1}{\lambda^2 V_M}} \tag{13}$$

and $E[t_K|t_M \geq t_K]$ is expressed as

$$\begin{aligned} &E[t_K|t_M \geq t_K] \\ &\quad= \left\{\frac{1}{\mu} + \left[\frac{df^*(s)}{ds}\right]\Bigg|_{s=\mu} - \frac{f^*(\mu)}{\mu}\right\} \times \left[\frac{1}{1-f^*(\mu)}\right] \\ &\quad= \left[\frac{1}{\mu} - \frac{1}{\lambda} \times \left(\frac{1}{\lambda \mu V_M + 1}\right)^{\frac{1}{\lambda^2 V_M}+1}\right. \\ &\qquad\quad \left. - \frac{1}{\mu} \times \left(\frac{1}{\lambda \mu V_M + 1}\right)^{\frac{1}{\lambda^2 V_M}}\right] \\ &\qquad \times \left[\frac{1}{1 - \left(\frac{1}{\lambda \mu V_M + 1}\right)^{\frac{1}{\lambda^2 V_M}}}\right]. \end{aligned} \tag{14}$$

TABLE I
COMPARISON OF ANALYTIC AND SIMULATION RESULTS. (a) α (EXPONENTIAL t_M). (b) $E[t_k|t_M \geq t_k]$ (GAMMA t_M AND EXPONENTIAL T). (c) $E[t_k^*|t_M \leq t_k]$ (GAMMA t_M AND EXPONENTIAL T)

$E[T]$	Fixed T			Exponential T		
(1/λ)	Analytic	Simulation	Errors (%)	Analytic	Simulation	Errors (%)
10^{-2}	0.005	0.005	0.2187	0.0099	0.0098	0.7171
10^{-1}	0.0484	0.0484	0.0068	0.0909	0.0909	0.0444
10^{0}	0.3679	0.3676	0.0645	0.5	0.4998	0.0492
10^{1}	0.9	0.9	0.005	0.9091	0.9092	0.0122
10^{2}	0.99	0.99	0.0008	0.9901	0.9901	0.0007

(a)

V_M	$E[T]=0.1/\lambda$			$E[T]=10/\lambda$		
$(1/\lambda^2)$	Analytic	Simulation	Errors (%)	Analytic	Simulation	Errors (%)
10^{-1}	0.0995	0.0996	0.0575	0.5364	0.5363	0.0261
10^{0}	0.0909	0.0909	0.0372	0.9091	0.91	0.0947
10^{1}	0.0831	0.0831	0.0257	3.0336	3.0425	0.2914
10^{2}	0.086	0.0861	0.0505	6.2541	6.229	0.4019
10^{3}	0.0892	0.0893	0.1398	7.8596	7.8467	0.1651

(b)

V_M	$E[T]=0.1/\lambda$			$E[T]=10/\lambda$		
$(1/\lambda^2)$	Analytic	Simulation	Errors (%)	Analytic	Simulation	Errors (%)
10^{-1}	0.1	0.1006	0.5986	10	10.005	0.0501
10^{0}	0.1	0.1001	0.1492	10	10.001	0.0137
10^{1}	0.1	0.1001	0.1074	10	10.012	0.1206
10^{2}	0.1	0.1	0.0222	10	10.003	0.032
10^{3}	0.1	0.1	0.0104	10	9.9999	0.001

(c)

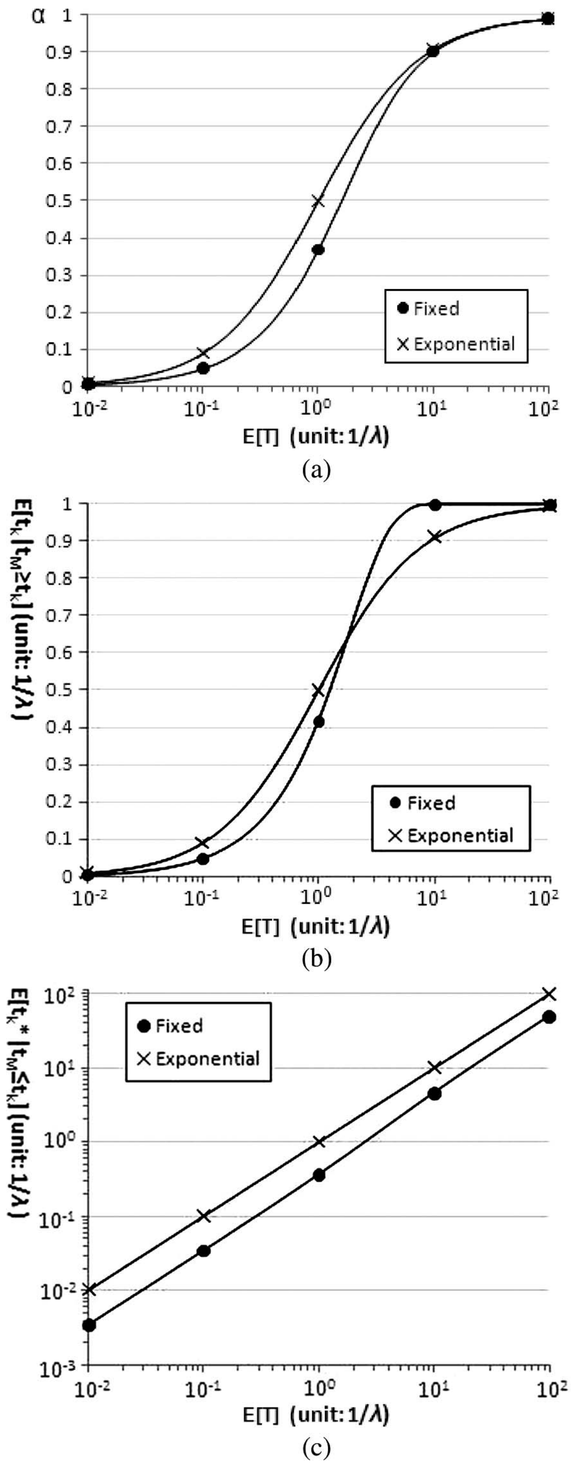
When t_M is exponentially distributed (i.e., $V_M = 1/\lambda^2$), (13) is rewritten as

$$\alpha = \left(\frac{1}{\lambda\mu V_M + 1}\right)^{\frac{1}{\lambda^2 V_M}} = \frac{\lambda}{\lambda+\mu} \tag{15}$$

and (14) is expressed as

$$\begin{aligned} E[t_K|t_M \geq t_K] = &\left[\frac{1}{\mu} - \frac{1}{\lambda} \times \left(\frac{1}{\lambda\mu V_M + 1}\right)^{\frac{1}{\lambda^2 V_M}+1} \right. \\ &\left. - \frac{1}{\mu} \times \left(\frac{1}{\lambda\mu V_M + 1}\right)^{\frac{1}{\lambda^2 V_M}}\right] \\ &\times \left[\frac{1}{1-\left(\frac{1}{\lambda\mu V_M+1}\right)^{\frac{1}{\lambda^2 V_M}}}\right] \\ = &\frac{1}{\lambda+\mu}. \end{aligned} \tag{16}$$

Equations (1), (4), (7), (12), (15), and (16) provide the mean value analysis to show the "trends" of the output measures. These equations are also used to validate the simulation experiments. Table I shows that the simulation is consistent with the analytic analysis, and all errors are within 1%.



(a)

(b)

(c)

Fig. 5. Effect of μ. (a) Effect of E[T] on α. (b) Effect of $E[T]$ on $E[t_K|t_M \geq t_K]$. (c) Effect of $E[T]$ on $E[t_K^*|t_M \leq t_K]$.

IV. NUMERICAL EXAMPLES

According to the analytic and the simulation models, we use numerical examples to investigate how the MSK lifetime T affects the performance of the key caching mechanism. Fig. 5 plots the results for exponential t_M. Fig. 5(a) plots α against $E[T]$. The figure indicates that α is an increasing function of $E[T]$. It is intuitive that if $E[T]$ is large, then it is more likely that the MS will return before the MSK lifetime expires. From (1)

$$\lim_{E[T]\to\infty} \alpha = \lim_{T\to\infty}\left(\frac{e^{-\lambda T}+\lambda T-1}{\lambda T}\right) = 1.$$

Since $E[T] = 1/\mu$, from (15), we have

$$\lim_{E[T]\to\infty} \alpha = \lim_{E[T]\to\infty}\left[\frac{\lambda}{\lambda+(1/E[T])}\right] = 1.$$

This figure also shows that the exponential T outperforms the fixed T in terms of α.

Fig. 5(b) plots the unused key period $E[t_K|t_M \geq t_K]$ as a function of $E[T]$. The figure shows that the unused key period increases as $E[T]$ increases. From (4), we have

$$\lim_{E[T]\to\infty} E[t_K|t_M \geq t_K] = \lim_{T\to\infty}\left(\frac{1}{\lambda} - \frac{Te^{-\lambda T}}{1-e^{-\lambda T}}\right) = \frac{1}{\lambda}$$

and from (16)

$$\lim_{E[T]\to\infty} E[t_K|t_M \geq t_K] = \lim_{E[T]\to\infty}\left[\frac{1}{\lambda+(1/E[T])}\right] = \frac{1}{\lambda}.$$

Therefore, the maximum unused key period is $E[t_M] = 1/\lambda$. When $E[T]$ is small (e.g., less than $1/\lambda$), the fixed T outperforms the exponential T. When $E[T]$ is large, the exponential T yields better performance in terms of the unused key period.

Fig. 5(c) plots the reused key period $E[t_K^*|t_M \leq t_K]$ as a function of $E[T]$. The figure indicates that the key reused period increases as $E[T]$ increases. From (7), we have

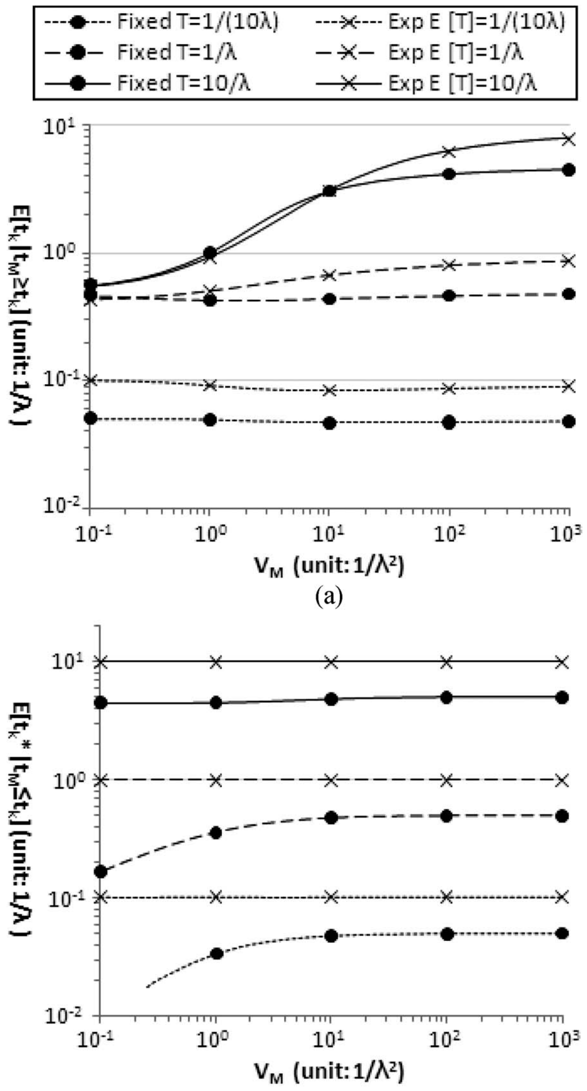
$$\lim_{E[T]\to\infty} E\left[t_K^*|t_M \leq t_K\right] = \lim_{T\to\infty}\left[\frac{\lambda T^2}{2(\lambda T + e^{-\lambda T}-1)} - \frac{1}{\lambda}\right] = \infty$$

and from (12)

$$\lim_{E[T]\to\infty} E\left[t_K^*|t_M \leq t_K\right] = \lim_{E[T]\to\infty}\left(\frac{1}{1/E[T]}\right) = \infty.$$

The figure also indicates that the exponential T outperforms the fixed T in terms of the reused key period.

Fig. 6(a) plots the unused key period $E[t_K|t_M \geq t_K]$ against $E[T]$ and V_M. When $E[T] \geq 1/\lambda$, the unused key period increases as V_M increases. This phenomenon is explained as follows. As V_M increases, more long and short t_M are observed. Since a random observer (an MS movement) tends to observe long t_M, short t_M will not contribute to $E[t_K|t_M \geq t_K]$. Therefore, more long t_K are observed as V_M increases. From (14), we have the equation shown at the bottom of the page. When $E[t_K] < 1/\lambda$, $E[t_K|t_M \geq t_K] \fallingdotseq E[t_K]$, which is not sensitive to V_M.



(a)

(b)

Fig. 6. Effect of V_M. (a) Effect of V_M on $E[t_K|t_M \geq t_K]$. (b) Effect of V_M on $E[t_K^*|t_M \leq t_K]$.

Fig. 6(b) plots the reused key period $E[t_K^*|t_M \leq t_K]$ against $E[T]$ and V_M. For the exponential T, according to (12), $E[t_K^*|t_M \leq t_K] = E[T]$. This phenomenon is explained as follows. Since the residual MSK lifetime t_K is exponentially distributed, the arrival of the MS to the old ASN-GW is a random observer to t_K. Thus, from the residual life theorem, t_K^* is also exponentially distributed with the mean $E[T]$. For the fixed T, $E[t_K^*|t_M \leq t_K]$ increases as V_M increases. Since we only consider the case when $t_M \leq t_K$, as V_M increases,

$$\lim_{V_M\to\infty} E[t_K|t_M \geq t_K] = \lim_{V_M\to\infty}\left\{\frac{\left[\frac{1}{\mu} - \frac{1}{\lambda}\times\left(\frac{1}{\lambda\mu V_M+1}\right)^{\frac{1}{\lambda^2 V_M}+1} - \frac{1}{\mu}\times\left(\frac{1}{\lambda\mu V_M+1}\right)^{\frac{1}{\lambda^2 V_M}}\right]}{1-\left(\frac{1}{\lambda\mu V_M+1}\right)^{\frac{1}{\lambda^2 V_M}}}\right\} = \frac{1}{\mu}$$

short t_K periods are observed, and long t_K will not contribute to $E[t_K^*|t_M \leq t_K]$. Thus, for the fixed T, the reused key period increases as V_M increases and eventually approaches $E[t_K] = T/2$.

Fig. 6 shows that the exponential T outperforms the fixed T in terms of the reused key period. On the other hand, for the unused key period, the fixed T outperforms the exponential T in most cases. Another advantage of the exponential T over the fixed T is that the reused key period $E[t_K^*|t_M \leq t_K]$ performance is not affected by the variance V_M. This stability property is important for a telecom-grade system.

V. Conclusion

This paper has proposed a key caching mechanism to speed up the inter-ASN-GW handoff for mobile WiMAX. With this mechanism, when an MS leaves the old ASN-GW, the MS key record (e.g., the MSK) is cached in the old ASN-GW. If the MS returns to the old ASN-GW before the MSK lifetime expires, it can reuse the MSK without executing the IEEE 802.1X authentication. On the other hand, the old ASN-GW consumes extra storage to maintain the MS key records when the MS leaves the old ASN-GW. This paper has investigated how the period T of the MSK lifetime affects the key caching performance by an analytic model and simulation experiments. Three output measures are evaluated: the key reuse probability, the unused key period, and the reused key period. We have shown that the caching mechanism can effectively speed up the inter-ASN-GW handoff. We also observed that the exponential T outperforms the fixed T in most cases. Moreover, for the reused key period, the exponential T is not affected by the variance of the MS residence period in new ASN-GWs and is more suitable for telecommunication systems. As a final remark, the operator uses our study and the number of serving MSs to calculate the storage budget at an ASN-GW. Our study indicates that $E[T] > 10/\lambda$ will not improve performance. Therefore, if $E[T] < 10/\lambda$ is selected, the extra storage can be computed from Little's law, i.e., $N = xE[T] < 10x/\lambda$, where N is the extra storage (the number of MS key records), and x is the rate of the MSs leaving the ASN-GW [10].

References

[1] H. Haverinen and J. Salowey, *Extensible authentication protocol method for Global System for Mobile Communications (GSM) Subscriber Identity Modules (EAP-SIM)*, Jan. 2006. RFC 4186.

[2] Y.-B. Lin and A.-C. Pang, *Wireless and Mobile All-IP Networks*. Hoboken, NJ: Wiley, 2005.

[3] WiMAX Forum, *WiMAX Forum Network Architecture, Stage 3: Detailed Protocols and Procedures*, Mar. 2007. Rel. 1.1.0.

[4] S. M. Ross, *Stochastic Processes*. Hoboken, NJ: Wiley, 1996.

[5] S.-R. Yang, "Dynamic power saving mechanism for 3G UMTS system," *Mobile Netw. Appl. (MONET)*, vol. 12, no. 1, pp. 5–14, Jan. 2007.

[6] M.-F. Chang, L.-Y. Wu, and Y.-B. Lin, "Performance evaluation of a push mechanism for WLAN and mobile network integration," *IEEE Trans. Veh. Technol.*, vol. 55, no. 1, pp. 380–383, Jan. 2006.

[7] W. Ma, Y. Fang, and P. Lin, "Mobility management strategy based on user mobility patterns in wireless networks," *IEEE Trans. Veh. Technol.*, vol. 56, no. 1, pp. 322–330, Jan. 2007.

[8] B. Rong, Y. Qian, K. Lu, H.-H. Chen, and M. Guizani, "Call admission control optimization in WiMAX networks," *IEEE Trans. Veh. Technol.*, vol. 57, no. 4, pp. 2509–2522, Jul. 2008.

[9] Y.-B. Lin and I. Chlamtac, *Wireless and Mobile Network Architectures*. Hoboken, NJ: Wiley, 2001.

[10] L. Kleinrock, *Queueing System*, vol. 1. Hoboken, NJ: Wiley, 1975.

Shih-Feng Hsu received the B.S. degree in computer science from the National Tsing Hua University, Hsinchu, Taiwan, in 2001. He is currently working toward the Ph.D. degree in computer science and information engineering with the Department of Computer Science, National Chiao Tung University, Hsinchu.

His current research interests include personal communication services, mobile computing, and wireless security in wireless local area network, Worldwide Interoperability for Microwave Access, and Universal Mobile Telecommunication Systems.

Yi-Bing Lin (M'95–SM'95–F'03) received the B.S. degree in electrical engineering from the National Cheng Kung University, Tainan, Taiwan, in 1983 and the Ph.D. degree in computer science from the University of Washington, Seattle, in 1990.

He is the Dean and the Chair Professor of the Department of Computer Science, National Chiao Tung University, Hsinchu, Taiwan. He is also with the Institute of Information Science, Academia Sinica, Taipei, Taiwan. He is the author of the books *Wireless and Mobile Network Architecture* (Wiley, 2001), *Wireless and Mobile All-IP Networks* (Wiley, 2005), and *Charging for Mobile All-IP Telecommunications* (Wiley, 2008).

Prof. Lin is a Fellow of the Association for Computing Machinery, the American Association for the Advancement of Science, and the Institution of Engineering and Technology. He is the recipient of numerous research awards, including the 2005 NSC Distinguished Researcher award and the 2006 Academic Award of the Ministry of Education. He is a Senior Technical Editor of IEEE Network. He serves on the editorial boards of the IEEE Transactions on Wireless Communications and the IEEE Transactions on Vehicular Technology. He has served as the General or Program Chair for many prestigious conferences, including the 2002 ACM MobiCom. He has been a Guest Editor for several first-class journals, including the IEEE Transactions on Computers.