2009 IEEE 34th Conference on Local Computer Networks (LCN 2009)
Zürich, Switzerland; 20-23 October 2009

Evaluation of Authentication Signaling Loads in 3GPP LTE/SAE Networks

Chan-Kyu Han, Hyoung-Kee Choi*
School of Information and Communication Engineering
Sungkyunkwan University
Suwon, South Korea
{hedwig,hkchoi}@ece.skku.ac.kr

Jung Woo Baek, Ho Woo Lee
Department of Systems Management Engineering
Sungkyunkwan University
Suwon, South Korea
{rainbeak, hwlee}@skku.edu

Abstract—**The integrated core network architecture and various mobile subscriber behaviors can result in a significant increase of signaling loads inside the evolved packet core network proposed by 3GPP in Release 8. Consequently, an authentication signaling analysis can provide insights into reducing the authentication signaling loads and latency, satisfying the quality-of-experience. In this paper, we evaluate the signaling loads in the EPS architecture via analytical modeling based on the renewal process theory. The renewal process theory works well, irrespective of a specific random process (i.e. Poisson). This paper considers various subscribers patterns in terms of call arrival rate, mobility, subscribers' preference and operational policy. Numerical results are illustrated to show the interactions between the parameters and the performance metrics. The sensitivity of vertical handover performance and the effects of heavy-tail process are also discussed.**

Keywords- 3GPP authentication and key agreement (AKA), long term evolution (LTE), evolved packet system (EPS), renewal process theory, mobile network security, performance evaluation

I. INTRODUCTION

Scarce radio resources and the ambition to increase the number of mobile customers are providing an impetus for mobile communication systems to evolve from 3G wireless systems into 4G networks with heterogeneously integrated networks and convergent service. The 3GPP long term evolution/system architecture evolution (LTE/SAE) has proposed evolved packet system (EPS) in Release 8 in order to support higher bandwidth, wider coverage, full interworking between heterogeneous systems on the All-IP architecture. The EPS architecture is exposed to a risk of unlawful access, malicious modification, willful tampering, and denial-of-service. Hence, 3GPP has announced a new security standard [12] supporting a more elaborate key hierarchy and security context switching that has been improved from [1]. In the meantime, Release 5 introduced the Internet protocol multimedia subsystem (IMS) to define an IP-based core network. However, the introduction of IMS can result in higher complexity on the signaling plane [10]. This is due to the following facts: (1) Release 5 signaling processing is managed by more network components than in Release 99; (2) Signaling in IMS is based on the session initiation protocol (SIP), which is significantly heavier in terms of message size and computational complexity than SS7 in Release 9. Consequently, an accurate analysis in the authentication signaling can provide insights into reducing loads and latency in the authentication signaling.

The three characteristics of beyond Release 5 (also including EPS) are as follows: (1) Various packet data services (*i.e.* instant messaging, multimedia streaming, mobile games, mobile internet, etc) are engaged, while only circuit-based voice service (*i.e.* call) was available in the past; (2) The multiple-access network can be integrated within the EPS architecture, thus handover across networks should be frequent; and (3) The authentication and key agreement system should be reconsidered due to changes of the key hierarchy and manipulation. To the best of our knowledge, there are no previous studies evaluating authentication signaling loads in the EPS architecture in 3GPP Release 8. We present an analytical framework for evaluating authentication signaling traffic with respect to numerous authentication arrival processes, subscriber's pattern and operational strategy. We adapt the renewal reward theorem to analyze authentication signaling loads normalized according to the residence time in the steady state form. In the aforementioned studies, for the sake of simplicity, the Poisson process is assumed for authentication requests. However, due to the busy-line effect, the Poisson call arrival process may not be proper [6].

The main contribution of this paper is fourfold: (1) Introducing the EPS and the revised security architecture in Release 8; (2) Relaxing the Poisson assumption in authentication request arrival process; (3) Exploring new authentication triggers (e.g. horizontal and vertical handover, multimedia requests) other than call arrival and location updates; and (4) Ramifying authentication signaling costs according to their origination and its elucidation.

Our analytical results show that the total signaling cost function is convex as either the number of authentication vectors or the lifetime of cryptographic key is increased. The increasing arrival rate holding the largest signaling cost composes distinct growing signaling loads while the effect of increasing the communication delay for querying to authentication server is insignificant. We examine a variety of authentication events which have different random processes, and explore the impact of exponential-based process such as Erlang distribution, hyper-exponential distribution, as well as Pareto distribution and log2-normal distribution which is very

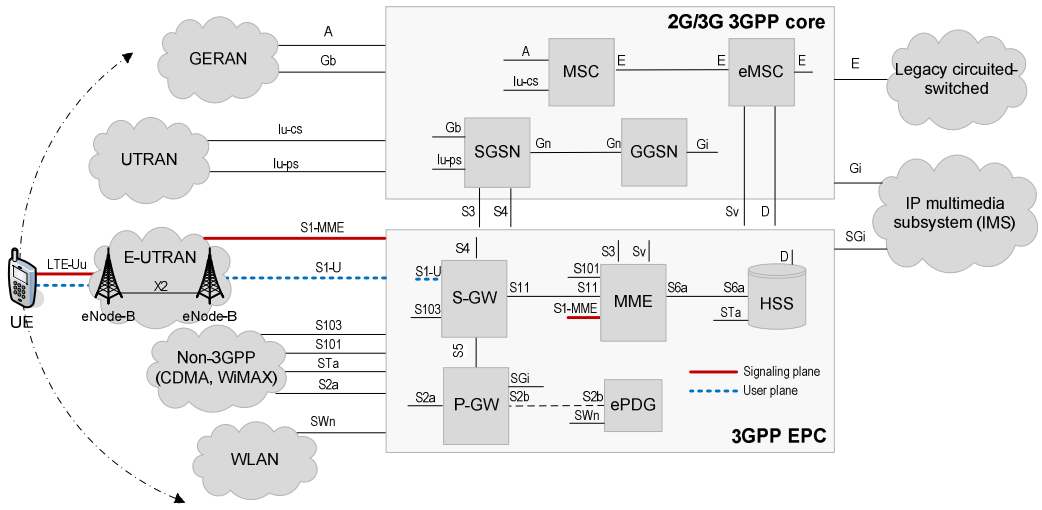

Figure 1. SAE/LTE System Architecture and EPC Network Elements

suitable for multimedia process due to its heavy-tail characteristic. The result shows that the baseline distribution without Pareto suffers from underestimation. Finally, we discuss the consequence of newly proposed mechanism by 3GPP Release 8, key re-use scheme.

The remainder of the paper is organized as follows. We review related works about authentication signaling traffic loads in Section 2. Section 3 specifies the EPS architecture and authentication procedures, and various handover scenarios. In Section 4, we present the proposed analytical framework used to study the performance of EPS authentication via the renewal process theorem. We provide various numerical analyses of the security and performance of EPS authentication and key agreement in Section 5. Finally, we present our conclusions along with discussion on future work in Section 6.

II. Related Works

In authentication and key agreement of 3GPP mobile networks, the authentication server generates multiple authentication vectors and sends them to the visited network. This mechanism reduces the amount of signaling traffic between the visited network and the authentication server. On the other hand, it results in extra storage overhead at the visited network [2][3].

Numerous analytical models have been proposed to investigate the impact of the number of authentication vectors [4][8], the waiting probability of a new authentication vector [5], and the time interval needed to maintain an unused authentication vector [7]. Yi-Bing Lin *et al*. have done pioneering work analyzing authentication signaling traffic in terms of the number of authentication vectors [4]. A quite naive algorithm has also been suggested to determine the value of the number of authentication vectors in assisting the authentication server. Ja'afer Al-Saraireh *et al*. improved Lin's scheme by pointing out that the algorithm for reducing signaling traffic should be optimized [8]. The analytical modeling methodologies of Lin and Al-Saraireh are quite similar in the sense that they are based on the Poisson arrival process and an exponential cell residence time. Yan Zhang pointed out that a subsequent authentication event after all previous authentication vectors have been used must wait until the authentication vector arrays have been fetched from the authentication server [5]. The authors showed that their proposed pre-authentication scheme decreases the authentication delay with minor increased signaling overhead. An appropriate reservation timeout period for maintaining remained authentication vector was investigated in [7]. It was concluded that a long reservation timeout period results in fewer accesses to the authentication server at the cost of extra storage in the visited network.

In the aforementioned studies, for the sake of simplicity the Poisson process is assumed for authentication requests, except in [6]. Recently, Yan Zhang in [6] has considered the impact of different mobility models and various random arrival processes other than Poisson arrival. But an aggregated arrival process was assumed for simplicity. This assumption has a limitation not to reflect various forms of authentication request arrival process. 3GPP released EPS authentication and key agreement (EPS-AKA) in February 2008 [12]. Soon after, Sankaran C. B. introduced the EPS architecture and a tutorial on network access security in the EPS architecture [11].

III. EPS and EPS-AKA

In this Section, we describe the EPS of 3GPP LTE/SAE in Release 8, including its authentication and key agreement procedures.

A. 3GPP LTE/SAE Architecture

3GPP LTE and SAE aim to offer a high data rate with OFDM technologies and the flat IP-based architecture. The EPS architecture is proposed to attain to the goals of 3GPP LTE/SAE. Fig. 1 shows the system architecture of EPS along with important interfaces. EPS consists of the access network and the core network. The access network is called the evolved universal terrestrial radio access network (E-UTRAN), and the core network is called the evolved packet core (EPC). The EPS

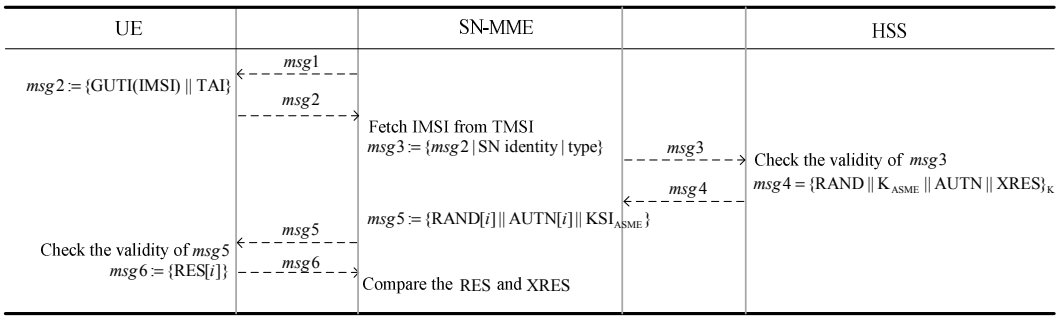

Figure 2. EPS authentication and key agreement in NAS and AS plane

is defined to work across multiple access networks (both 3GPP access networks include E-UTRAN, UTRAN, and GERAN; non-3GPP access networks include CDMA-2000, WiMAX and WiFi) [13]. All radio access protocols must be managed in one node, namely eNodeB. The signaling control plane and the user data plane are established on the interfaces of S1-MME and S1-U, respectively, between E-UTRAN and EPC based on different interfaces. All 3GPP-specific access technologies are connected through the serving gateway (S-GW), and all non-3GPP specific access technologies are typically connected through either the packet data network gateway (P-GW) or the evolved packet data gateway (ePDG). The mobility management entity (MME) is equivalent to the GERAN/UTRAN serving GPRS support node (SGSN), and handles control functions including authentication, signaling security and mobility. Further detailed information will be available elsewhere [11][13][14].

B. Extended Authentication and Key Agreement (EPS-AKA)

To eliminate the vulnerability of the All-IP EPC network and lack of interoperability with various access technologies, a new authentication and key agreement (AKA), referred to as EPS-AKA, was proposed by amending the UMTS-AKA. The description of EPS-AKA is shown in Fig. 2.

The user equipment (UE) sends its permanent identity, namely the international mobile subscriber identity (IMSI), in clear text as a response regarding *msg*1. Or the UE may send its globally unique temporary identity (GUTI) with tracking area identity (TAI) if the UE already has the GUTI that the previous MME transmitted after the EPS-AKA procedure. The MME passes the second message (*msg*2) to the HSS with the serving network identity including the MME's mobile network code and network type (*e.g.* E-UTRAN). If the IMSI is valid, the HSS generates and sends an array of *K* authentication vectors (EPS-AVs) to the MME in *msg*4. An authentication vector consists of a random number chosen by the HSS (RAND), an expected response from the UE in the sixth message (XRES), an authentication token including an message authentication code (AUTN) and a derived key (K_{ASME}). K_{ASME} is derived from hash function of CK, IK and serving network identity. The MME selects one authentication vector from among *K* vectors and sends RAND[i] and AUTN[i] to the UE in *msg*5. Further, the MME can include KSI_{ASME}, which is an index of K_{ASME}. The UE authenticates the MME by checking the message authentication code contained in AUTN, and then it generates CK, IK and K_{ASME}. The UE also stores K_{ASME} along with the received KSI_{ASME} and the GUTI. Upon receipt of *msg*6, the MME compares the received RES[i] with XRES[i]. If they match, the MME considers the UE as valid, and authentication is successful. After the EPS-AKA, the ciphering and integrity algorithms and security capabilities are exchanged to negotiate the security mode command (SMC). In EPS, two SMC procedures are defined: one on the non-access stratum (NAS) level between UE and MME and the other on the access stratum (AS) level between UE and eNodeB.

C. Comparison of EPS-AKA and UMTS-AKA

The differences between UMTS-AKA and EPS-AKA are summarized in Table I. False base station (BS) attacks can be eluded by strengthening the mutual authentication between the UE and the MME. Note that the UE could identify the MME using serving network identity. It was recommended that the MME fetches only one authentication vector at a time, as the need to perform the EPS-AKA was reduced because of a more elaborate key hierarchy. In particular, service requests can be authenticated using a cached K_{ASME} without the need to perform a full-blown EPS-AKA. Furthermore, the re-synchronization problems caused by interleaving of batches of authentication vectors are eliminated. Henceforth, we call the prior authentication methodology as described in Section 3.B *authentication vector-based* EPS-AKA, while other methodologies are referred to as cached K_{ASMI} *key reuse-based* EPS-AKA. Moreover, NAS signaling is ciphered, and its integrity is checked at the two points, one in the NAS and the other within the AS plane.

TABLE I. SECURITY FUNCTIONALITY COMPARISON BETWEEN UMTS-AKA AND EPS-AKA

Criteria	UMTS-AKA	EPS-AKA
Vulnerability	Vulnerable to redirection, false BS attack	Prevented by including network identity
Authentication	Only vector-based AKA	(1) Either Vector-, or (2) key reuse- based AKA
ID protection	TMSI, but limited	GUTI, but limited
Key material	VLR recognize CK, IK	MME hold only K_{ASME}, not CK and IK
Confidentiality	Ciphering only at the AS level	Ciphering at the AS (both signaling and data) and NAS level (signaling only)
Integrity	Integrity protection only at the AS level	Integrity protection at the AS and NAS level
Handover support	(1) SGSN handover, (2) roaming to GSM	(1) X1, (2) S1 handover, (3) Inter RAT and (4) interworking non-3GPP

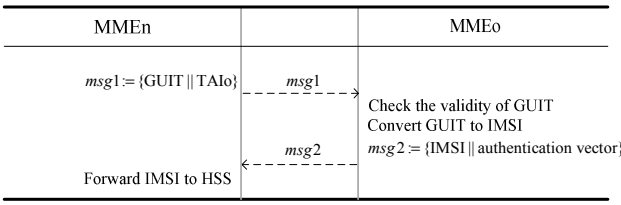

Figure 3. Distribution of IMSI and security context within one serving network domain

Handovers occurs more frequently than ever because eNodeB's coverage gets smaller and more access network technologies are introduced. There are four types of handover supported by EPS-AKA.

- *X2 handover*: Handovers are possible directly between eNodeBs. eNodeBs establish a handover chain via an X2 interface (See Fig. 1). The previous eNodeB transfers key materials including K_{eNB} to the current eNodeB. To guarantee backward secrecy, K_{eNB} is updated to the newly created K_{eNB}^* via a one-way hash function. In addition, the MME can assist in guaranteeing forward secrecy.
- *S1 handover*: The distribution of IMSI and the security context within one serving network domain is defined in EPS-AKA [1][12]. The procedure is shown in Fig. 3. The S1 handover procedure is invoked by the new MMEn, wherein the UE is identified via GUITo and TAIo under the jurisdiction of a previously visited MMEo. The MMEo finds the user data, and sends the IMSI back to MMEn, which can optionally include a number of unused EPS-authentication vectors or current security context.
- *Inter RAT*: The UE may be registered in both SGSN in UMTS and MME in EPS simultaneously. MME converts K_{ASME} into the new CK' and IK', and sends them to SGSN within the context response message. The target SGSN replaces the stored parameters CK and IK with CK' and IK' received from the MME.
- *Interworking with non-3GPP network*: There are three options for mobility between 3GPP EPS and non-3GPP networks; Proxy mobile IP, mobile IPv4 and dual stack mobile IPv6. IPsec tunnel (with ePDG) is established in case the non-3GPP network is an untrusted network such as a wireless LAN.

IV. ANALYTICAL MODELING

In this section, we introduce the renewal process and the renewal reward theorem as a means of analyzing, modeling and evaluating signaling loads caused by EPS-AKA.

A. System model and Analysis

A Poisson process is known as a counting process for which the inter-arrival times between events are independent and identically distributed (*i.i.d*) exponential random variables. One possible generalization is to consider a counting process for which the times between successive events are *i.i.d* with an arbitrary distribution. Such a counting process is called a renewal process [9]. If the sequence of nonnegative random variables, $\{X_1, X_2, \ldots\}$, where X_n denote the time between the $(n\text{-}1)$th and the nth event of a counting process, the counting process $\{N(t), t\geq 0\}$ is said to be a *renewal process*. At every renewal epoch $\{S_n, n\geq 0\}$, the renewal process is completely recursive. For a renewal process with renewal interval times $\{X_n, n\geq 0\}$, equation (1) holds:

$$S_n = \sum_{i=1}^{n} X_i = S_{n-1} + X_n, \text{ where } S_0 = 0 \tag{1}$$

Assume that a reward is earned at the time of renewal epoch or renewal interval. The reward could be the cost or profit attached to the renewal. Let R_n denote the reward earned at the time of the n-th renewal X_n, with a common mean $E[R]$. R_n are *i.i.d* with allowing dependability on X_n. Then, the total reward earned by t, $R(t)$, is given by equation (2). The new process $\{R(t), t\geq 0\}$ is called a *renewal reward process* [9].

$$R(t) = \sum_{n=1}^{N(t)} R_n, \quad \text{where } N(t) = \max\{n : S_n \leq t\} \tag{2}$$

Suppose that $\{R(t), t\geq 0\}$ is a renewal *reward* process with $E[R]<\infty$ and $E[X] <\infty$. Then, the long-run reward rate is given by equation (3). The equation (3) holds by proofing of strong law of large number.

$$\lim_{t\to\infty} \frac{R(t)}{t} = \lim_{t\to\infty} \frac{\sum_{n=1}^{N(t)} R_n}{N(t)} \cdot \frac{N(t)}{t} = \frac{E[R]}{E[X]} \tag{3}$$

B. Signaling Cost in vector-based scheme

We deduce that the EPS-AKA procedure can be interpreted as a renewal process for the following reasons. Note that the renewal epoch indicates the spending all authentication vectors fetched from HSS (vector-based EPS-AKA) or expiry of the key-reuse timer (key reuse-based EPS-AKA). Furthermore, the renewal inter-arrival time may not follow an exponential distribution because of mash-up of multiple authentication triggers (*i.e.* call arrival/termination, X2 handover, S1 handover, multimedia request, etc.). Renewal process has a tolerance to arbitrary random process not limited to Poisson arrival process. Besides, renewal reward theorem can calculate

TABLE II. NOTATIONS AND DESCRIPTIONS

Notation	Description
K	Number of fetched EPS authentication vector
T	Reuse expiration period of K_{ASME}
M	Number of authentication trigger (M=5; call or multimedia request, X2 handover, S1 handover, and Inter RAT)
p_i	Probability of occurrence for i-th authentication trigger ($1\leq i\leq M$)
X_i	Inter-arrival time of i-th authentication trigger
$F(x_i)$	$\Pr[X_i\leq x_i]$, $f(x_i)= (d/dx)F(x_i)$
$T(x_i)$	$\Pr[T\leq x_i]$, $t(x_i)= (d/dx)T(x_i)$
Y	Permutation of multiple of X_i (any one of X_i)
C_i	Total signaling cost for completion of i-th authentication trigger
D	SIP signaling load and processing time for generating authentication vector

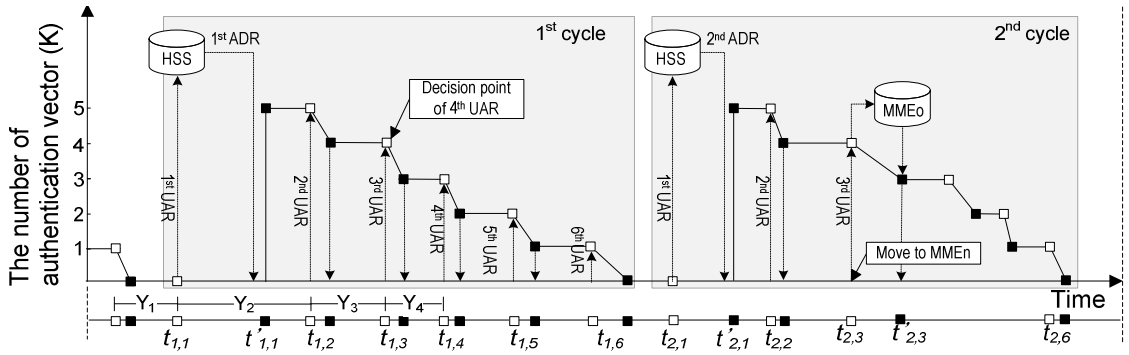

Figure 4. Authentication event timing diagram as a form of renewal process

the overall cost (reward) without observation during long period, rather than infer it by considering a cycle of renewal process. Consequently, we analyze the authentication signaling loads of EPS-AKA by adapting the renewal reward theorem. The notation used throughout this paper is listed in Table II.

Fig. 4 depicts the timing diagram of vector-based EPS AKA with the interpretation as a form of renewal process. Assume that the number of authentication vectors (K) would be six. Also note that $t_{x,y}$ and Y_i represent the time when y-th user authentication request (UAR) occurs within x-th ADR and duration of each authentication event, respectively. At $t_{1,1}$, the UE firstly invokes the UAR message. Then it fetches the authentication vectors from HSS, and terminates at $t'_{1,1}$ through authentication data request/response (ADR). The filled and blank rectangle represents the starting and terminating point of each UAR event, respectively. At time $t'_{1,1}$, the five authentication vectors are valid. Assume that an S1 handover occurs at time $t_{2,3}$, and then a context transfer between MMEs (MMEo and MMEn) as described in Section 3 is invoked. Assume that dice with pentagonal faces (M=5) are cast every time the n-th UAR starts to decide the kind of next (n+1)-th UAR. For example, the fourth UAR has been decided at the starting point of the third UAR, $t_{1,3}$. The expected value of the renewal interval, $[t_{n,K}, t_{n+1,K}]$, can be decided by a combination of Y_i and D, where D is the computation delay of HSS and the communication delay of the SIP signaling cost up to HSS. Y_i has a probability density function $f(Y)$ as shown in equation (4), because it follows the Bernoulli distribution with p_i for each authentication trigger X_i.

$$f(Y) = \sum_{i=1}^{M} p_i \cdot f(X_i), \text{ where } \sum_{i=1}^{M} p_i = 1 \qquad (4)$$

According to the renewal reward theorem, the cost per unit time is calculated as the expected value of the reward, $E[R]$, divided by expected value of the renewal interval, $E[N]$. First, renewal interval is determined by $E[D]$ and the (K-1) accumulation of $E(Y)$ as shown in equation (5).

$$E[N] = (K-1) \cdot E[Y] + E[D] \qquad (5)$$

Second, the expected renewal (cost) would be calculated as shown in equation (6), where C_i is the total signaling cost of completion of i-th authentication trigger X_i.

$$\begin{aligned} E[R] &= C_0 + (K-1) \cdot E[Y] \cdot C_{i(1 \le i \le M)} \\ &= C_0 + \sum_{j=1}^{K-1} j \cdot (\sum_{i=1}^{M} p_i E[X_i] C_i) \end{aligned} \qquad (6)$$

Thus, the average cost function, $C(K)$, is given by,

$$\begin{aligned} C(K) &= \frac{E[R]}{E[N]} \\ &= \frac{C_0 + \sum_{j=1}^{K-1} j \cdot (\sum_{i=1}^{M} p_i E[X_i] C_i)}{K \cdot E[Y] + E[D]} \end{aligned} \qquad (7)$$

However, we must consider the exceptional event C_ε and the probability p_ε. The corresponding exceptional authentication trigger occurs when UE departs the renewal process because of power-off, S1 handover to a different serving domain, and roaming to Non-3GPP access technologies. The cost is bounded until the exceptional case occurs, thus it is affected by a geometric random variable where $p_{\varepsilon,k} = p_\varepsilon \cdot (1 - p_\varepsilon)^k$. The total authentication signaling cost for an arbitrary inter-arrival time is expressed as shown in equation (8).

$$C(K) = \frac{C_0 + \sum_{k=1}^{K-1} p_{\varepsilon,k} \cdot \left(\sum_{i=1}^{M} (K-1) p_i E[X_i] C_i \right) \cdot C_\varepsilon}{K \cdot E[Y] + E[D]} \qquad (8)$$

C. Signaling Cost in key reuse-based scheme

Because CK and IK are never sent to MME, re-authentication with current K_{ASME} would be possible. In case of re-authentication, only SMC is processed via the same K_{ASME}, without invoking a new EPS-AKA procedure. The purpose of the KSI_{ASME} is to be used to allow re-use of the K_{ASME} during subsequent connection set-ups. Assume that UE has a policy that it invokes a new EPS-AKA procedure as soon as K_{ASME} either is compromised or reaches the time period of T. No matter whether K_{ASME} is compromised or expired, the renewal process should be established by the strong law of large number. The renewal epoch is the moment when UE replace previous K_{ASME} with new one by re-starting the key reuse-based EPS-AKA procedure.

Expected length of the renewal cycle, $E[N]$, is as shown in equation (9), where Y_i is i-th replacement moment of re-starting the EPS-AKA. Note that $T(x_i)$ is the cumulative density function of lifetime of K_{ASME} in days.

$$\begin{aligned} E[N] &= \int_0^\infty E(Y \mid X_i = x_i) dT(x_i) \\ &= \int_0^T x_i \cdot t(x_i) dx_i + \int_T^\infty T \cdot t(x_i) dx_i \\ &= \int_0^T [1 - T(x_i)] dx_i \end{aligned} \quad (9)$$

Suppose that C_{-1} is additional cost needed when K_{ASME} is compromised thus all EPS network entities should be rekeyed.

$$\begin{aligned} E[R] &= (C_{-1} + C_i) \cdot \int_0^T t(x_i) dx + C_i \cdot \int_T^\infty t(x_i) dx \\ &= (C_{-1} + C_i) \cdot T(T) + C_i \cdot [1 - T(T)] \end{aligned} \quad (10)$$

Based upon equation (9) and equation (10), the total authentication signaling cost for an arbitrary inter-arrival time is expressed as shown in equation (11). For the simplicity, we exclude the exceptional event with probability $p\varepsilon$ in this scheme, while it is self-explanatory as in the vector-based scheme.

$$C(T) = \sum_{i=1}^{M} p_i \cdot \frac{(C_{-1} + C_i) \cdot T(T) + C_i \cdot [1 - T(T)]}{\int_0^T [1 - T(x_i)] dx_i} \quad (11)$$

D. Cost function for each authentication trigger

We need to clarify and quantify the cost function of each authentication trigger, C_i ($0 \le i \le M$). For each authentication event, the message exchanges between EPS entities and the service procedure is fully specified in [13]. Table III shows the authentication trigger and the corresponding message exchange expressed as $C_{A\text{-}B}$ (the EPS entity A communicated with the EPS entity B). We exclude uplink and downlink data, communication between gateways and communication with policy servers. Generally in the LTE standard, the RRC signaling latency and the user control plane latency are delayed by up to 10ms and 100ms, respectively. We normalized three cases of message exchanges between UE and eNodeB: within the core network, and across the network, as α, β, and γ, respectively. The RTT in the core network tends to be slower than the RTT in wireless access networks as the number of hops increases. Thus $\gamma \ge \beta \gg \alpha$ expresses the relationship between the normalized asymptotic signaling costs.

C_1 depends on whether it originated from a network (call termination) or from a UE (call origination). In an outgoing call, additional radio bearer establishment and bearer update are expensive, while only paging-related signaling is exchanged in an ingoing call. Multimedia requests are associated with third parties. We presume that multimedia broadcast multicast service (MBMS) is used for this authentication trigger. Along with the handover (C_3, C_4 and C_5), the tracking area update procedure (C_{TAU}) must be performed without the S-GW relocation procedure.

V. Numerical Results

We relax the Poisson assumption by adapting the renewal process in order to analyze the authentication signaling loads. Besides, we divide the authentication event (*trigger*) into five different service requests (*i.e.* call, multimedia, X2, S1 and UMTS handover). In this section, we will present the representative examples to show the authentication traffic load via evaluation based on the proposed analytic model.

A. Analysis of baseline cost function with exponential distribution

We analyze the baseline cost function, $C(K)$, with an exponential distribution for all authentication triggers. We assume that i-th authentication trigger follows an exponential distribution for its inter-arrival process with rate λ_i. The probability of occurrence for i-th event is uniformly distributed. Based on these assumptions, the expectation of the renewal interval is given by equation (8), where the authentication fetch latency follows an exponential distribution with rate γ.

Fig. 5 shows the total signaling cost function $C(K)$ with an exponential distribution for all authentication triggers. The total signaling cost function is convex with increasing K. The optimized K' value minimizes $C(K)$. If the arrival rate is increased from λ_i=2 (Case 1) to λ_i =5 (Case 2), the total signaling cost is minimal at the K' values between 10.21 and 16.71. The effect of increasing the setup distribution rate (γ) is insignificant as shown in Case 2 and Case 3. The increasing arrival rate with the largest signaling cost (4th authentication trigger) composes a distinct growing signaling cost.

Fig. 6 examines the sensitivity of the distribution pattern of pi with same mean rate. We compare three groups with various variances (*std*). When there is lower variation, the optimal K value is inclined to decrease, while the minimized cost function $C(K)$ tends to increase. The rationale behind this phenomenon is as follows: if a subscriber has propensity to be incline to a certain event (*e.g.* popularity of multimedia service among teen-ager), surprisingly the subscriber less agitates the whole EPS system than a polyphagous subscriber.

TABLE III. Signaling cost for each authentication trigger

Cost	Authentication trigger	Asymptotic signaling cost
C_0	Authentication fetch to HSS	$2\alpha+2\beta+4\gamma+ K \cdot C_{SHA-1}$
C_1	Call origination	$4\alpha+9\beta$
	Call termination	$2\alpha+5\beta$
C_2	Multimedia request	$4\alpha+4\beta+4\gamma$
C_3	X2 handover	$4\alpha+10\beta+3\gamma$
C_4	S1 handover	$6\alpha+14\beta+21\gamma$
C_5	Inter RAT preparation	$\alpha+5\beta+6\gamma$
	Inter RAT execution	$\alpha+2\beta+7\gamma$
C_{TAU}	Tracking area update	$4\alpha+6\beta+3\gamma$
C_{-1}	Revocation / Rekeying	$C_0+2\alpha+2\beta+2\gamma$

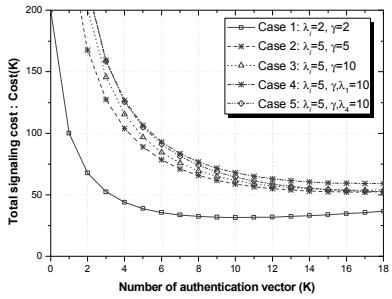

Figure 5. Total signaling cost with respect to exponential inter-arrival rate

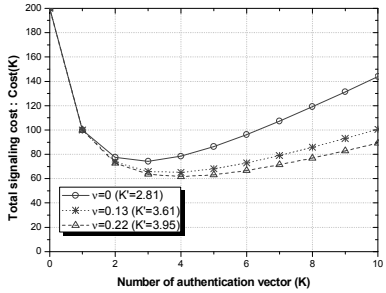

Figure 6. Effect of the cvariance v for probability of occurrence (λ_i=1)

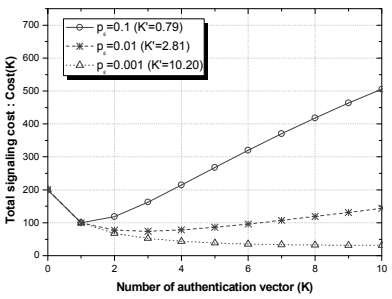

Figure 7. The sensitivity of exceptional conditions (λ_i=1)

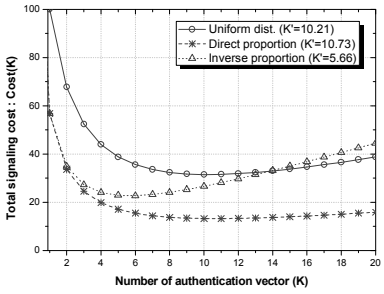

Figure 8. The effect of cost function ratio with respect to the arrival rate (p_i=0.2, p_ε=0.001)

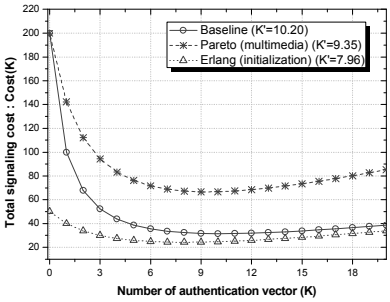

Figure 9. The effect of the Pareto and Erlang distribution

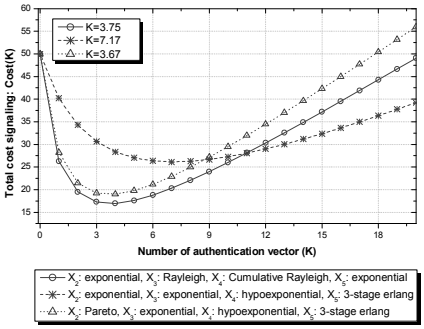

Figure 10. The various combinations of continuous random processes for handover

p_ε is the probability of satisfying the escape condition from the renewal process. Fig. 7 shows the effect of exceptional conditions. When there is a higher p_ε, the total signaling cost is obviously increased and the K value should be minimized to alleviate a waste of system resource. When a UE habitually undergoes power-off or roams into other access technologies with a probability of 10%, the vector-based EPS-AKA is meaningless (K'<1).

Fig. 8 represents the three different arrays of the arrival rate, λ_i, according to the cost function ratio. The cost function introduced obeys the relation, $C1<C_2<C_3<C_5<C_4$. We assume the arrival rate of an exponential distribution follows the direct proportion of the original sequence, and the inverse proportion of the original sequence with preserving same mean.

B. The effect of various random processes

We consider the following typical different inter-arrival time distributions having the fixed mean: exponential, hyper-exponential, Erlang (gamma), Rayleigh distribution. The traditional exponential distribution for call inter-arrival may not be valid. Even if the Poisson process would not be implicated, the comparison indicates that there is insignificant difference in employing exponential, hyper-exponential, or Erlang distributed call inter-arrival distributions [6]. Instead, the author in [15] has presented a number of analytic random variable (*e.g.* Pareto) models for describing the characteristics of various Internet-based applications such as telnet, smtp, and ftp connections. Accordingly, the Pareto distribution is suitable for simulations of multimedia requests. Also log2-normal and log2-extreme distribution has heavy-tail characteristics with self-similarity, thus they are suitable for call arrival request.

The Pareto distribution is widely used to accurately approximate Web traffic [6]. The parameters α and β used in the Pareto distribution were set to 1.2 and $(\alpha\text{-}1)/\alpha\cdot\lambda_i$, respectively, where i=2 (in case of multimedia request). Further, we examined the Erlang distribution for the initial authentication vector fetch phase for D. The results in Fig. 9 show that the baseline distribution without Pareto suffers from underestimation. A Pareto distribution in multimedia requests increases overall authentication signaling, on the other hand the number of authentications required is decreased.

Fig. 10 examines the various combinations of continuous random variables that are only applicable for the multimedia request (X_2), and handover process (X_3, X_4 and X_5). Rayleigh random variables can arise when a two-dimensional vector has elements that are normally distributed, are uncorrelated and have equal variance. These characteristics of Rayleigh process could explain the handover, which consists of the subscriber's velocity and cell coverage by eNodeB. The fitting and modeling of various user patterns can be found via Fig. 10. We do not argue that one specific combination is better than another. Given to various combination of authentication event distribution, we could find the optimal value of the number of authentication vectors.

VI. CONCLUSION

Integrated convergent 4G services are required to support various handover and increase authentication signaling costs. This paper has given an introduction to the authentication procedures and types of support for handover that have been used in the EPS network. We have proposed a sophisticated mathematical framework for analyzing and evaluating the

authentication signaling load. The effects of various performance parameters were investigated. The result indicates that the optimal K value is inclined to decrease, while the minimized cost function tends to increase with lower variation. The Pareto distribution in multimedia requests increases overall authentication signaling, on the other hand, the number of authentications required is decreased. We examined the Rayleigh distribution required for adequate eNodeB handover. Our result establishes the necessity of studying mobility management, security policy, and various random arrival processes. It is important for finding the means of determining system parameters based on the proposed model and results. Finally, there are several aspects that still need be investigated in order to improve the feasibility of our system for real-world scenarios (*e.g*. random walk mobility model, correlations with each authentication trigger, non-homogeneous poisson process).

REFERENCES

[1] Third Generation Partnership Project; Technical Specification Group Services and System Aspects; 3G Security, "Security Architecture (Release 7)," 3GPP TS 33.102 version 7.1.0, December 2006.

[2] Chung-Ming Huang and Jian-Wei Li, "Authentication and key agreement protocol for UMTS with low bandwidth consumption," Proceedings of the *19th International Conference on Advanced Information Networking and Applications* (AINA), March 2005.

[3] Ja'afer Al-Saraireh and Sufian Yousef, "A new authentication protocol for UMTS mobile networks," *EURASIP Journal on Wireless Communications and Networking*, Vol.2006, Issue 2, April 2006.

[4] Yi-Bing Lin and Yuan-Kai Chen, "Reducing authentication signaling traffic in third-Generation mobile network," *IEEE Transactions on Wireless Communications*, Vol.2, No.3, May 2003.

[5] Yan Zhang and Masayuki Fujise, "An improvement for authentication protocol in third-generation wireless networks," *IEEE Transactions on Wireless Communications*, Vol.5, No.9, September 2006.

[6] Yan Zhang, Shaoqiu Xizo, Mingtuo Zhou and Masayuki Fujise, "Authentication traffics modeling and analysis in next generation wireless networks," *Wiley Wireless Communications and Mobile Computing*, Vol. 8, No. 5, January 2007

[7] Lin-Yi Wu and Yi-Bing Lin, "Authentication vector management for UMTS," *IEEE Transactions on Wireless Communications*, Vol.6, No.11, November 2007.

[8] Ja'afer Al-Saraireh and Sufian Yousef, "Analytical model for authentication transmission overhead between entities in mobile networks," *Elsevier Computer Communications*, Vol.30, Issue 8, June 2007.

[9] Sheldon M. Ross, "Introduction to probability models," Academic Press, 8th ed., December 2002.

[10] Dario S. Tonesi, Luca Salgarelli, Yan Sun and Thomas F. La Porta, "Evaluation of signaling loads in 3GPP networks," *IEEE Wireless Communications*, February 2008.

[11] Sankaran C. B., "Network access security in next-generation 3GPP systems: A Tutorial," *IEEE Communications Magazine*, February 2009.

[12] Third Generation Partnership Project; Technical Specification Group Services and System Aspects; 3GPP System Architecture Evolution (SAE), "Security Architecture Security (Release 8)," 3GPP TS 33.401 version 8.2.1, December 2008.

[13] Third Generation Partnership Project; Technical Specification Group Services and System Aspects. "General Packet Radio Service (GPRS) Enhancements for Evolved Universal Terrestrial Radio Access Network (E-UTRAN) Access," 3GPP TS 23.401 version 8.4.1, December 2008.

[14] Third Generation Partnership Project; Technical Specification Group Radio Access Network. "Evolved Universal Terrestrial Radio Access (E-UTRA) and Evolved Universal Terrestrial Radio Access Network (E-UTRAN); Overall description; Stage 2," 3GPP TS 36.300 version 8.7.0, December 2008.

[15] Vern Paxson, "Empirically derived analytic models of wide area TCP connections," *IEEE/ACM Transactions on Networking*, Volume: 2 Issue: 4, August 1994.