Diminishing Signaling Traffic for Authentication in Mobile Communication System

Chi-Chun Lo and Kuen-Liang Sue
Institute of Information Management
National Chiao Tung University
Hsinchu, Taiwan
cclo@cc.nctu.edu.tw, klsue@csie.nctu.edu.tw

Abstract

To identify a mobile station (MS) and validate legal service requests, authentication functions are utilized in the location registration, call origination and call termination procedures. In GSM, the VLR requests the HLR for assistance in authenticating the visiting user. The authentication center (AuC) in the home network generates 3-tuples and sends them back to the VLR for subsequent authentications during the user's residence. If these 3-tuples are used up before the MS's leaving, another request is issued by the VLR. The request is expensive, because it needs to access the HLR/AuC. Traditionally, a fixed-K strategy is used. That is, K 3-tuples are sent to the VLR for each request. Lager K is preferred to reduce the number of the expensive requests for 3-tuples. However, much waste of 3-tuples is observed, especially when an inactive user is considered. Hence, K value should be determined based on the usage pattern of the user. We propose a dynamic-K (DK) strategy to reduce the waste and diminish the signaling traffic for authentication. Simulation results show that the DK strategy can effectively determine the appropriate K value. Not only the waste but also the number of requests are diminished efficiently.

1. Introduction

In Global System for Mobile (GSM) communications, the Mobile Station (MS) communicates with the Base Station (BS) by radio. BS are connected to the Mobile Switching Center (MSC) which communicates with the Public Switched Telephone Network (PSTN). There are two kinds of mobility databases for handling service provision. One is the Home Location Register (HLR), and the other is the Visitor Location Register (VLR). The HLR contains the information of all subscribers. The VLR stores the roaming users' records that are used to handle the location update, call origination and call termination [1]. The Authentication Center (AuC) generates security parameters upon the request from the HLR. The AuC may be co-located with the HLR. A simplified system architecture is shown as Fig. 1.

One of the core security issues in personal communication system is user authentication for providing the contracted services. Another one is session key agreement for securing subsequent sessions. Both of the security issues are accomplished by the 3-tuples (RAND, SRES, K_c) generated from AuC. When a roaming user arrives in a location area (LA), the VLR of the LA will need the assistance of the HLR to authenticate the roamer. In general, the HLR asks AuC to generate K 3-tuples and forwards them to the VLR for the subsequent utilization. Each of the location update, call origination and call termination needs one 3-tuple for authentication operation. If the 3-tuples are used up, the VLR will request again and the HLR offers K new 3-tuples. Such arrangement is called fixed-K strategy. The 3-tuples request is expensive, because it needs to access the HLR/AuC. Hence, an larger K is preferred to reduced the number of requests. A very similar mechanism is adopted by Universal Mobile Telecommunication (UMT) which is the third generation mobile service technology evolved from General Packet Radio Service (GPRS). In 3GPP TS 29.002, K=5 is recommended [2].

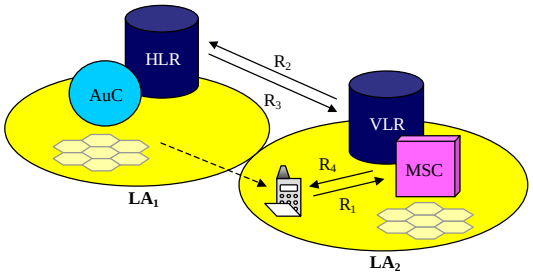

Figure 1: A simplified GSM architecture

The 3-tuples transmissions from AuC to VLR may occupy much network bandwidth if larger K is used. Moreover, many 3-tuples are not used before he/she leaves the LA or becomes the victim user. These unused 3-tuples are discarded and becomes a waste. The victim user is the one whose VLR record is replaced due to VLR overflow [3]. The 3-tuples in the record will be lost under such condition. Especially, the inactive users are the more possible targets to be selected as a victim [4]. Because they seldom make/receive calls, the waste of 3-tuples is more serious.

Thus, it is desirable to select an appropriate K value to reduce the waste and minimize the signaling traffic for authentication. In this paper, a dynamic-K strategy is inves-

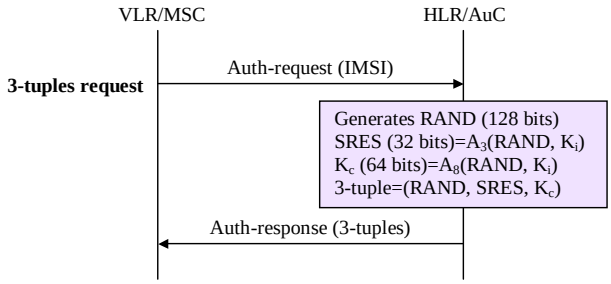

Figure 2: The 3-tuples request protocol between VLR and HLR/AuC

tigated. The authentication protocol in GSM is described in section II. Section II also explains the relationship between the waste of 3-tuples and the VLR overflow. The dynamic-K strategy is introduced in section III. According to various system parameters, the numerical analysis is investigated in section IV.

2. User Authentication and VLR Overflow

When a roaming user visits an LA, the service is provided by the system in the visited area. The visited system should verify the identity of the user to provide the contracted services. The verification is implemented with an authentication protocol between the MS and the VLR/MSC. For simplicity, we assume VLR is co-located with the MSC. For the security of the user, the authentication must be executed under the assistant of the user's HLR. The VLR cannot authenticates the roaming user itself. The details of the authentication is introduced in this section. Moreover, the VLR overflow control scheme may affect the authentication information. The affection is also described briefly in the subsection.

2.1 The Authentication Protocol in GSM

In GSM, each SIM card has a unique International Mobile Subscriber Identity (IMSI) and a secret key (K_i). One-way functions A_3 and A_8 are used during the authentication. Fig. 2 diagrams the 3-tuples request protocol which is executed as the roaming user arrives in the LA or the 3-tuples are used up. The aim of the request is to gain K 3-tuples computed by the HLR/AuC of the roaming user. The first 3-tuple is for the location registration that confirms the arriving of the user. Whenever he/she wants to take advantage of system services (call origination or termination), the remaining 3-tuples will be used for the authentication of the roaming user. One authentication for services needs one tuple.

Fig. 3 shows the authentication protocol between MS and VLR/MSC. To verify the identity of the roaming user, the VLR/MSC sends RAND to the user. RAND is a 128-bits random number. The roamer uses its secret key K_i, RAND, and A_3 algorithm to compute SRES* which is sent back to the VLR. If the stored SRES matches the received SRES*, the authentication is successful and services will

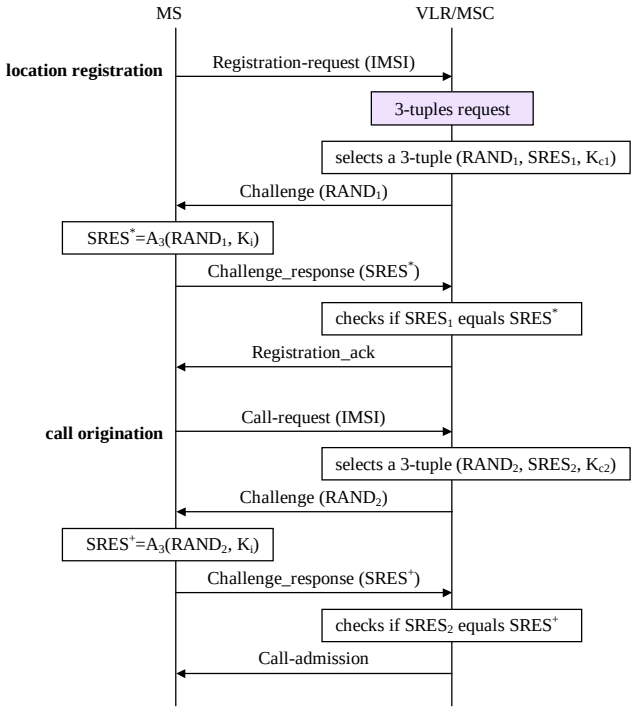

Figure 3: The authentication protocol between MS and VLR/MSC

be provided to the roaming user. Besides, the MS uses K_i, RAND, and A_8 algorithm to compute the session key K_c. The K_c is utilized for encrypting/decrypting the subsequent session between MS and MSC.

2.2 The Waste of 3-tuples in VLR Overflow

When a mobile user roams from his/her home system to another system, a temporary record for the user is created in the VLR of the visited system. The record includes information for handling calls to or from the user. In order to generate the important VLR record, the roamer must inform the VLR about his coming by registration. When the user leaves the visited system, deregistration is necessary to cancel the corresponding record in the VLR [5]. The number of VLR records changes dynamically when the visitor moves in or out. Due to the limited capacity, the VLR may become full if too many mobile users enter a VLR during a short period. As a VLR is full, the arriving mobile users will fail to register in the VLR. In such condition, these users cannot receive any service. The phenomenon is called *VLR overflow*. These users are called *overflow users*.

To resolve the VLR overflow problem, an overflow control scheme selects a record from the overflow VLR as a victim. The victim will be replaced with the record of the overflow user. The approach enables overflow users to receive services even when a VLR overflows. The inactive users are the mobile subscribers with low probability to issue/receive calls. Hence, the inactive users are the better targets to be selected as victims. [3]

The 3-tuples are stored in the VLR record. Once the

record is replaced, the remaining 3-tuples will be lost. When the victim user would like to make/receive calls, his/her record must be reconstructed again. Another 3-tuples request is issued from VLR to HLR for K new 3-tuples. Even if the record of an inactive user is never replaced, the remaining 3-tuples are still discarded when he/she leaves the LA. The waste of the 3-tuples is considerable for the inactive users. In next section, we propose a dynamic-K strategy to diminish the waste of 3-tuples and the authentication traffic load in signaling network.

3. The Dynamic-K Strategy

The main idea of the dynamic-K (DK) strategy is that the number of 3-tuples should be determined according to the call frequency of the roaming user. A user with high call frequency should be given more 3-tuples. On the other hand, less 3-tuples are enough for the inactive user. In this method, a simple call counter is maintained in the VLR record of the user. When a call origination/termination occurs, the call counter *cn* increases by one. If the user has no call for a period of time, the call counter *cn* decreases by one. As a roaming user arrives an LA, the registration procedure enables the VLR to request 3-tuples from the HLR/AuC of the user. Because no historic analysis of the user's call activity, three 3-tuples are generated and transferred to the VLR. If the user is an active user, the three 3-tuples will be used up soon. Then, K is computed with $K = cn + m$, where m is the system parameter preset by the system manager. The next request for 3-tuples will be responded with K 3-tuples. To compare the DK strategy and fixed-K strategy, we designed simulation models for both of them. For approximating true condition, the overflow control scheme with second chance replacement policy is also added in our simulation models [3]. We made the following assumptions in our experiments [6, 7].

- There are two classes of mobile users. Class 1 users have a low call connection rate $\lambda_{c,1}$, and class 2 users have a high call connection rate $\lambda_{c,2}$.
- The user arrival rates of the class 1 and class 2 users are $\lambda_{u,1}$ and $\lambda_{u,2}$, respectively.
- The call connections are Poisson processes [8].
- The user arrivals are assumed to be Poisson processes, too.
- The residence time distributions of all users have the Gamma distribution with mean $1/\lambda_m$ and variance V_m [9].

The second and third assumptions can be easily relaxed to fit general distributions in our simulation experiments. For the stable simulation result, 1,000,000 roaming users are simulated in each simulation run [10]. The output of the simulations are the number of waste 3-tuples N_w and the number of the request for 3-tuples N_r. The analysis of the DK strategy and the fixed-K strategy is investigated according the simulation results.

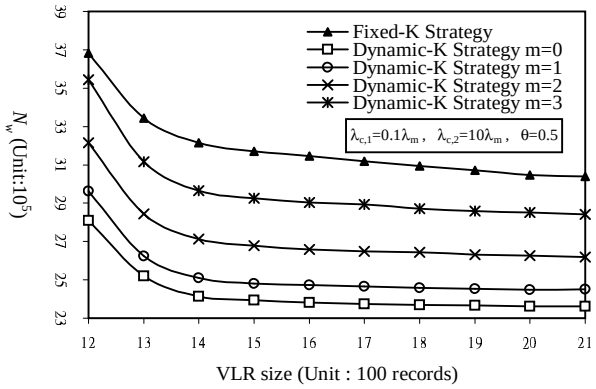

Figure 4: Effects of V on N_w

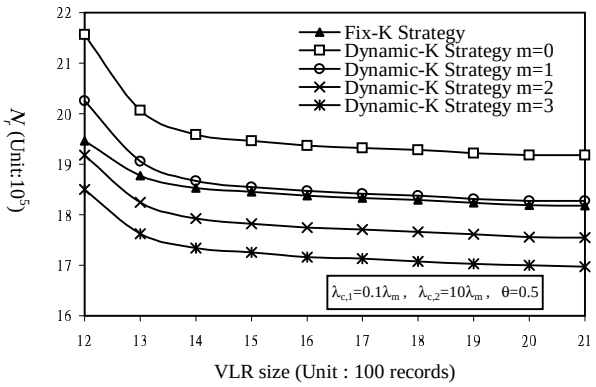

Figure 5: Effects of V on N_r

4. Numerical Analysis

As mentioned above, the VLR residence time distributions for both classes of users are assumed to have the Gamma distribution with mean $1/\lambda_m$ and variance V_m. In our simulation, V_m=$1/\lambda_m^2$, $\lambda_{c,1} = 0.1\lambda_m$ and $\lambda_{c,2} = 10\lambda_m$ are considered. We also assume that the total user arrival rate λ_u to a VLR area is a fixed value $2000\lambda_m$, where $\lambda_u = \lambda_{u,1} + \lambda_{u,2}$. V represents the VLR size.

Fig. 4 plots N_w as functions of VLR size V. In this experiment, we assume that $\theta = 0.5$, where $\theta = \frac{\lambda_{u,1}}{\lambda_{u,1}+\lambda_{u,2}}$. That is, half of the roaming users are class 1 and the other half are class 2. Compared with the fixed-K strategy, this figure demonstrates that the DK strategy can always reduce the waste of the 3-tuples. As VLR size increases, the waste decreases for both DK strategy and fixed-K strategy. Due to the larger VLR size, the frequency of discarding the 3-tuples drops off with the probability of record replacement. Another interesting phenomenon is that greater system parameter m seems to reduce the performance of the DK strategy. A straightforward judgment is that $m = 0$ is the best choice. However, the answer will be quite different when the number of 3-tuple requests N_r is considered below.

The relationship between N_r and VLR size is illustrated in Fig. 5. For the same reason in previous discussion, N_r decreases as VLR size increases. Moreover, another im-

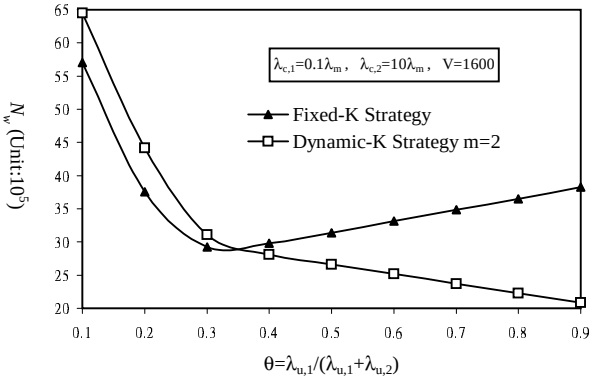

Figure 6: Effects of θ on N_w

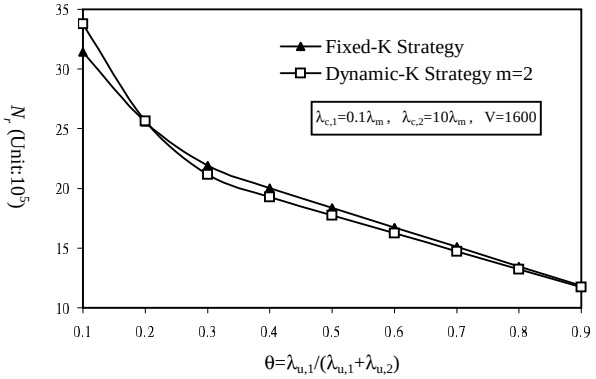

Figure 7: Effects of θ on N_r

portant phenomenon is noticed. When m is greater than 1, the DK strategy needs less N_r than the fixed-K strategy. The request for 3-tuples is expensive, because it needs to access the HLR/AuC. Therefore, any strategy to reduce the waste of the 3-tuples should not increase N_r. From Fig. 4 and Fig. 5, we know that n=2 is a good choice for DK strategy which can diminish both N_w and N_r.

The effects of $\lambda_{u,1}$ and $\lambda_{u,2}$ are shown in Fig. 6. It plots N_w as functions of the ratio θ. As θ increases, the class 1 users increases. It means that the average number of the calls per user decreases. A significant phenomenon is that the waste caused by fixed-K strategy increases for $\theta > 0.3$. Obviously, it is inefficient to set K=5 fixedly. A contrasting result gained by the DK strategy. The waste always decreases as θ increases. For $\theta > 3.5$, the DK strategy outperforms the fixed-K strategy. If $\theta < 3.5$, most of the user are active and their K value will be increased by DK strategy. Once they leaves, the average of waste is K/2 which has high probability to be larger than 5/2. Under such condition, the system manager can adjust m=1 or 0 to get better performance than fixed-K strategy.

The relationship between N_r and θ is illustrated in Fig. 7. $\theta < 0.1$ is the extreme case. Under such condition, almost all users have high probability to make/receive calls. The N_r becomes very high. In other words, most users need to make more than one request for 3-tuples. Because three 3-tuples are given for the first request in the DK strategy, it has higher probability to require the second request than the fixed-K strategy which gives five 3-tuples. However, this is an abnormal condition and seldom happens. For $\theta > 0.2$, the DK strategy always needs fewer 3-tuples requests than the fixed-K strategy.

5. Conclusion

The 3-tuples (RAND, SRES, K_c) are generated from the HLR/AuC, which enable the VLR to authenticate the roaming user. A similar protocol is adopted by the third generation mobile service technology. The fixed-K strategy is used in GSM and 3G communication systems. Since the cost for accessing HLR/AuC is expensive, K=5 is recommended for the fixed-K strategy to prevent too many requests for the 3-tuples. It means that five 3-tuples are sent back to the VLR from HLR/AuC for each request. However, this strategy is inflexible. Much waste of the 3-tuples is observed. In this paper, we propose a dynamic-K (DK) strategy to diminish the waste. Most importantly, the DK strategy can reduce the number of the expensive requests for 3-tuples at the same time.

The main idea of the DK strategy is that K should be adjusted dynamically based on the call pattern of each user. The active user makes/receives calls often, so higher K is appropriate for them. On the other hand, the inactive user needs only smaller K value to prevent the waste of 3-tuples. The DK strategy can provide appropriate K value for each user. Simulation results show that the DK strategy gains better performance than the fixed-K strategy in most scenarios investigated in our study. The DK strategy reduces not only the waste of 3-tuples but also the number of the expensive requests for 3-tuples. Furthermore, a parameter m is provided to the system manager for adjusting the tradeoff between the waste of 3-tuples and the number of requests according to various population of resident users. The DK strategy is efficient and flexible to diminish the signaling traffic for authentication in GSM and 3G mobile communication systems.

References

[1] I. Chlamtac, T. Liu, and J. Carruthers, "Location Management for Efficient Bandwith Allocation and Call Admission Control," *Proc IEEE Wireless Commu. and Networking Conf.*, Sept. 1999..

[2] 3GPP, "3rd Generation Partnership Project; Technical Specification Core Network; Mobile Application Part specicfication for Release 1999," Technical Spec. 3G TS 29.002 version 3.7.0 (2000-12), 3GPP, 2000.

[3] C.-C. Lo and K.-L. Sue, "Second Chance Replacement Policy for Mobile Database Overflow," To be appeared in *Proc. IEEE GLOBECOM'02*.

[4] Y.-B. Lin, "Overflow control for cellular mobility database," *IEEE Trans. Veh. Technol.*, vol. 49, no. 2, pp. 520-530, Mar. 2000.

[5] I. F. Akyildiz, J. Mcnair, J. S. M. Ho, H. Uzunalioglu, and W. Wang, "Mobility management in Next-Generation

Wireless Systems," *Proc. IEEE*, vol. 87, pp. 1347-1384, Aug. 1999.

[6] Y. Fang and I. Chlamtac, "Teletraffic Analysis and Mobility Modeling for PCS Networks," *IEEE Trans. on Commu.*, vol. 47, no. 7, pp. 1062-1072, Jul. 1999.

[7] Y.-B. Lin, "Modeling techniques for large-scale PCS networks," *IEEE Trans. on Comput.*, vol. 50, no. 4, pp. 356-370, Apr. 2001.

[8] S.M. Ross, "Stochastic Processes Volume I-Theory," John Wiley & Sons, Inc., 1983.

[9] S.M. Ross, "Introduction to Probability Models," Academic Press, Inc., 1993.

[10] J. Banks, J. Carson II, and B. Nelson, "Discrete-Event System Simulation," Prentice Hall, 1996.