This paper was presented as part of the Mobility Management in the Networks of the Future World (MobiWorld) Workshop at

A Timer-based Session Setup Procedure in Cellular-WLAN Integrated Systems

Gwangwoo Park, Younghyun Kim, and Sangheon Pack
School of Electrical Engineering
Korea University, Korea
Email: {kwoo17, m.s.yhkim, shpack}@korea.ac.kr

***Abstract*—The multiple registration scheme can reduce signaling overhead at the expense of increased session setup latency in cellular-WLAN integrated systems. In this paper, we propose a timer-based session setup procedure, which can reduce the session setup latency by looking for a user based on the estimation of the WLAN residence time. We investigate the optimization of the timer under realistic WLAN residence time model. Numerical results demonstrate that the timer-based session setup procedure can reduce the session setup latency by means of adaptive timer setting.**

***Index Terms*—Timer-based session setup procedure, adaptive timer, multiple registration, dual home agent, performance analysis.**

I. INTRODUCTION

Wireless local area networks (WLANs) are very popular in public areas because they can provide high data rates up to 11Mbps (IEEE 802.11b) or 54Mbps (IEEE 802.11a/g). However, the service coverage of WLAN is limited to a few tens of meters and thus the WLAN cannot be used for ubiquitous wireless access. To complement this shortage, the integration of WLANs and cellular networks (e.g., 2G/3G and High Speed Downlink Packet Access (HSDPA)) is actively discussed, and extensive works have been reported in different areas such as interworking architecture, mobility/resource management, and quality of service (QoS) support [1]. In particular, the emergence of smartphones requires frequent vertical handoffs between WLAN and cellular systems, and thus efficient location management, for keeping track of the current location of a mobile terminal (MT) and establishing a call with the MT, becomes a critical issue in cellular-WLAN integrated systems.

For efficient location management in cellular/WLAN integrated systems, we have proposed a multiple registration (MR) scheme [2] with a concept of dual home agent (DHA), which is analogy to the multiple home location register (MHLR) [3]. In the MR scheme, the DHA can maintain multiple location information: one is for WLAN and the other is for cellular systems. By means of MR, no registration procedure is performed for a movement from a WLAN area to a cellular area, and therefore signaling traffic can be reduced. However, this reduced signaling traffic can be obtained at the expense of the increased session setup latency. In other words, since the DHA maintains multiple registration information, ambiguity may occur before establishing a session with a called MT. Reducing the session setup latency is critical especially for delay-sensitive interactive multimedia applications (e.g., voice over IP (VoIP)) in wireless networks [4].

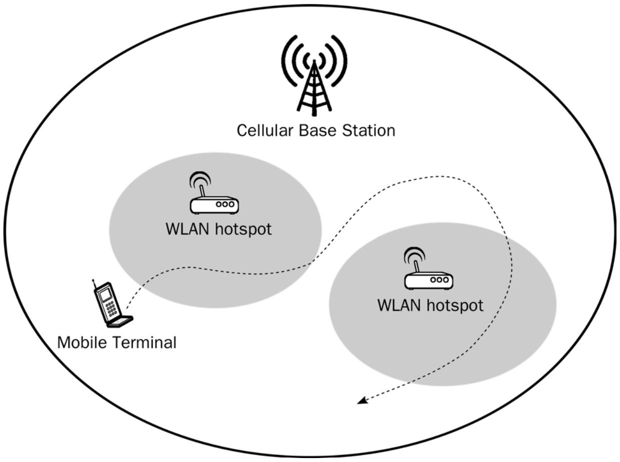


Fig. 1. Cellular-WLAN integrated networks.

In this paper, we propose a timer-based session setup procedure to reduce the session setup latency incurred in the MR scheme. In the proposed procedure, the residence time at the WLAN area is estimated and a timer is set to the estimated value. For an arriving session setup request, the WLAN area is first searched if the timer does not expire. On the contrary, the cellular area is first checked for a session setup request arriving after the timer expiration. By developing the analytical model, we investigate the optimization of the timer and present a practical binary search algorithm for the optimal timer under realistic WLAN residence time. The impacts of the timer and session arrival rate are investigated, which demonstrates that the proposed timer-based session setup procedure outperforms conventional procedures (i.e., WLAN-first and Cellular-first) in diverse environments by setting an adaptive timer.

The remainder of this paper is organized as follows. In Sections II, the timer-based session setup procedure is described. Section III presents an analytical model and a binary search algorithm for finding the optimal timer. Section IV illustrates numerical results in different environments. Finally, Section V concludes this paper.

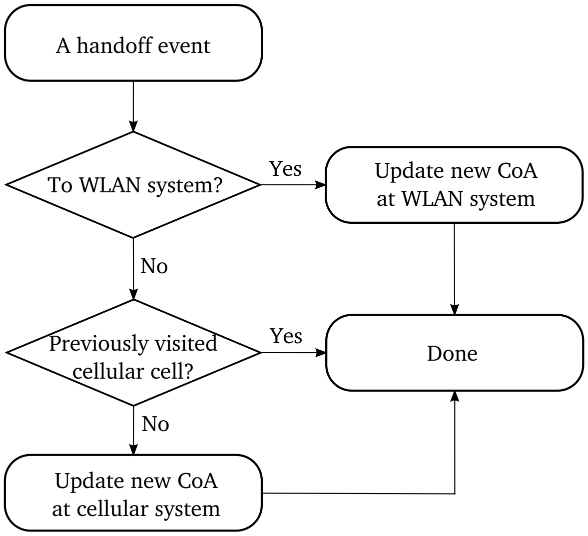


Fig. 2. Location update procedure in the MR scheme.

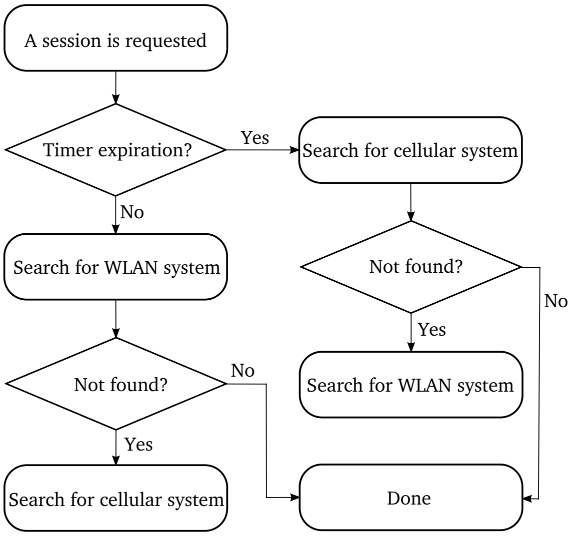


Fig. 3. Timer-based session setup procedure.

II. TIMER-BASED SESSION SETUP PROCEDURE

As shown in Figure 1, we consider a WLAN/cellular integrated network where WLAN hotspots are sparsely deployed within a cellular area. Since WLANs provide much higher data rates than cellular systems, an MT will perform vertical handoff from a cell to a WLAN hotspot whenever it enters the WLAN hotspot [5]. In addition, the MT conducts location update procedure with the DHA.

Figure 2 shows the flow chart for location update procedure in the MR scheme [2]. If the occurred handoff is destined to the WLAN area, the MT should configure a new care of address (CoA) and update the CoA with the DHA. On the other hand, for a handoff toward the cellular system, it should be checked whether the target cellular cell is the same as the previously registered one. If the target cell has been previously visited, no CoA configuration and update procedure are needed and therefore the location update cost can be saved. Otherwise, a new CoA is configured and it is notified to the DHA.

Due to the location update mechanism with the DHA, the exact location of an MT is not known after entering a WLAN area if another handoff occurs within the same cellular cell. This is because the MT does not inform the DHA of its location until it moves into a new WLAN area, and the DHA maintains two location information. That is, the DHA has two CoAs for an MT; one is for cellular systems and another is for WLAN systems. Although this location update procedure is of benefit to reduce the location update cost, it incurs another challenge in determining the current location (or CoA currently used by the MT) during the session setup procedure.

In the timer-based session setup procedure, the DHA initializes a timer T after the location update at a WLAN area. The timer is set to an estimated value of the WLAN area residence time. To estimate the residence time, an exponentially weighted moving average (EWMA) scheme is assumed. Let $T_W[i]$ and T_M be the ith estimated WLAN area residence time and the measured WLAN area residence time, respectively. Then, $T_W[i+1]$ is updated as

$$T_W[i+1] = \alpha \cdot T_W[i] + (1-\alpha) \cdot T_M, \tag{1}$$

where α is a weighting parameter and $0 \leq \alpha \leq 1$.

When a session setup request (e.g., INVITE message in session initiation protocol (SIP) or the first packet in mobile IP (MIP)) arrives at the DHA, the DHA evaluate the timer T. Before the timer expiration, it is likely that the MT still remains in a WLAN area, and therefore the WLAN area is first looked for. On the contrary, the cellular area is first checked for a session setup request arriving after the timer expiration. This procedure is illustrated in Figure 3.

The performance of the timer-based session setup procedure is highly dependent on the timer, mobility model at WLAN areas, and session arrival patterns. Therefore, we will investigate the effect of these parameters in the next section.

III. PERFORMANCE ANALYSIS

To analyze the performance of the timer-based session setup procedure, we consider the following six cases (see Figure 4).

- Case A: A session setup request arrives when an MT resides in the WLAN area before the timer expiration. The timer is smaller than the WLAN residence time.
- Case B: A session setup request arrives when an MT resides in the WLAN area after the timer expiration. The timer is smaller than the WLAN residence time.
- Case C: A session setup request arrives when an MT resides in the cellular area after the timer expiration.. The timer is smaller than the WLAN residence time.
- Case D: A session setup request arrives when an MT resides in the WLAN area before the timer expiration. The timer is equal to or larger than the WLAN residence time.
- Case E: A session setup request arrives when an MT resides in the cellular area before the timer expiration.

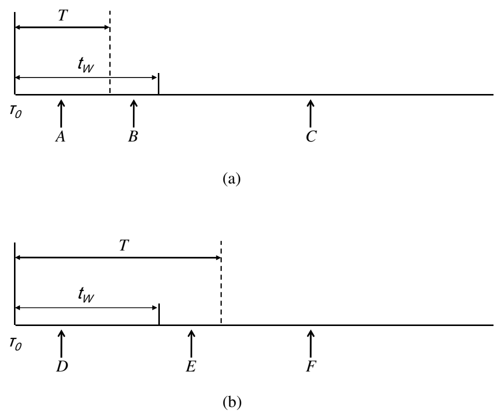


Fig. 4. Timing diagram (t_W: WLAN residence time).

The timer is equal to or larger than the WLAN residence time.

- Case F: A session setup request arrives when an MT resides in the cellular area after the timer expiration. The timer is equal to or larger than the WLAN residence time.

To compute the probability for case X, P_X, we assume that the WLAN subnet residence time t_W follows a two-stage hyper-exponential distribution with mean $1/\mu_W$, which is simple but well captures the high variability in WLAN hotspots [6]. The probability density function (PDF) of t_W is then given by

$$f_W(t) = \frac{a}{a+1}\frac{1}{\frac{1}{a\mu_W}}e^{-a\mu_W t} + \frac{1}{a+1}\frac{1}{a\frac{1}{\mu_W}}e^{-\frac{1}{a}\mu_W t}. \quad (2)$$

Increasing the parameter a results in t_W with a higher variability. The cumulative distribution function (CDF) of t_W is

$$\begin{aligned} F_W(t) &= \frac{a}{a+1}\left(1 - e^{-a\mu_W t}\right) + \frac{1}{a+1}\left(1 - e^{-\frac{1}{a}\mu_W t}\right) \\ &= 1 - \frac{a}{a+1}e^{-a\mu_W t} - \frac{1}{a+1}e^{-\frac{1}{a}\mu_W t}. \quad (3) \end{aligned}$$

On the other hand, the inter-session arrival time t_A follow an exponential distribution with rate λ_A. Let τ_0 be the time instant when the MT enters a WLAN area and completes its location update to the DHA. Then, the time period from τ_0 to the session arrival epoch, t_A, follows the same exponential distribution with rate λ_A by the random observer property [8].

First, P_A can be obtained as

$$\begin{aligned} P_A &= \Pr(t_A \le T \le t_W) \quad (4) \\ &= F_A(T)(1 - F_W(T)) \\ &= \left(1 - e^{-\lambda_A T}\right)\left(\frac{a}{a+1}e^{-a\mu_W T} + \frac{1}{a+1}e^{-\frac{1}{a}\mu_W T}\right). \end{aligned}$$

On the other hand, P_B can be obtained from

$$\begin{aligned} P_B &= \Pr(T < t_A \le t_W) \quad (5) \\ &= e^{-\lambda_A T} - \lambda_A \int_T^\infty e^{-\lambda_A \tau} F_W(\tau)\, d\tau \\ &= \frac{a}{a+1}\frac{\lambda_A}{\lambda_A + a\mu_W}e^{-(\lambda_A + a\mu_W)T} \\ &+ \frac{1}{a+1}\frac{\lambda_A}{\lambda_A + \frac{1}{a}\mu_W}e^{-\left(\lambda_A + \frac{1}{a}\mu_W\right)T}. \end{aligned}$$

Similar to P_B, P_C can be computed as

$$\begin{aligned} P_C &= \Pr(T \le t_W < t_A) \quad (6) \\ &= \int_T^\infty f_W(\tau)\, e^{-\lambda_A \tau} d\tau \\ &= \frac{a}{a+1}\frac{a\mu_W}{\lambda_A + a\mu_W}e^{-(\lambda_A + a\mu_W)T} \\ &+ \frac{1}{a+1}\frac{\frac{1}{a}\mu_W}{\lambda_A + \frac{1}{a}\mu_W}e^{-\left(\lambda_A + \frac{1}{a}\mu_W\right)T}. \end{aligned}$$

By its definition, P_D is given by

$$\begin{aligned} P_D &= \Pr(t_A \le t_W < T) \quad (7) \\ &= F_W(T) - \int_0^T f_W(\tau)\, e^{-\lambda_A \tau} d\tau \\ &= 1 - \frac{a}{a+1}\left[e^{-a\mu_W T} + \frac{a\mu_W}{\lambda_A + a\mu_W}\left(1 - e^{-(\lambda_A + a\mu_W)T}\right)\right] \\ &- \frac{1}{a+1}\left[e^{-\frac{1}{a}\mu_W T} + \frac{\frac{1}{a}\mu_W}{\lambda_A + \frac{1}{a}\mu_W}\left(1 - e^{-\left(\lambda_A + \frac{1}{a}\mu_W\right)T}\right)\right]. \end{aligned}$$

Similarly, P_E can be obtained as

$$\begin{aligned} P_E &= \Pr(t_W < t_A < T) \quad (8) \\ &= \lambda_A \int_0^T e^{-\lambda_A \tau} F_W(\tau)\, d\tau \\ &= 1 - e^{-\lambda_A T} - \frac{a}{a+1}\frac{\lambda_A}{\lambda_A + a\mu_W}\left(1 - e^{-(\lambda_A + a\mu_W)T}\right) \\ &- \frac{1}{a+1}\frac{\lambda_A}{\lambda_A + \frac{1}{a}\mu_W}\left(1 - e^{-\left(\lambda_A + \frac{1}{a}\mu_W\right)T}\right). \end{aligned}$$

Similar to (4), P_F is given by

$$\begin{aligned} P_F &= \Pr(t_W < T < t_A) \quad (9) \\ &= F_W(T)\, e^{-\lambda_A T} \\ &= e^{-\lambda_A T} - \frac{a}{a+1}e^{-(\lambda_A + a\mu_W)T} \\ &- \frac{1}{a+1}e^{-\left(\lambda_A + \frac{1}{a}\mu_W\right)T}. \end{aligned}$$

Then, the average session setup latency can be written as

$$\begin{aligned} L &= P_A D_W + P_B(D_C + D_W) + P_C D_C \\ &+ P_D D_W + P_E(D_W + D_C) + P_F D_C. \quad (10) \end{aligned}$$

where D_W and D_C are the session setup latencies in a WLAN and cellular areas, respectively. It is assumed that D_W and D_C are constant values.

To derive the optimal T minimizing the average session setup latency L, the first-order derivative of (10) with respect to T is obtained as

$$\frac{dL}{dT} = D_W\left[f_A(T)\, F_W(T)\right] + D_C\left[f_A(T)\left(F_W(T) - 1\right)\right]. \quad (11)$$

When $\frac{dL}{dT}$ equals 0, the average session setup latency can be minimized. Therefore, we have $D_W\left[f_A(T)\, F_W(T)\right] + D_C\left[f_A(T)\left(F_W(T) - 1\right)\right] = 0$, which can be reduced to $F_W(T) = \frac{D_C}{D_W + D_C}$. When the WLAN residence time is drawn from a hyper-exponential distribution, we have

$$1 - \frac{a}{a+1}e^{-a\mu_W T} - \frac{1}{a+1}e^{-\frac{1}{a}\mu_W T} = \frac{D_C}{D_W + D_C}. \quad (12)$$

However, it is not easy to find out T satisfying (12). Therefore, we introduce an indirect approach for optimal T. We first derive the optimal timer T_{EXP} when the WLAN residence time is exponentially distributed, i.e., $a = 1$. When a is 1, (12) becomes $e^{-\mu_W T} = \frac{D_W}{D_W + D_C}$. Consequently, the optimal timer T_{EXP} is given by

$$T_{EXP} = \frac{1}{\mu_W} \ln \frac{D_W + D_C}{D_W}.$$

By plotting (10), the trend of optimal T depending on a can be observed as shown in Figure 5. Apparently, the optimal T decreases as a increases. Consequently, the optimal T for arbitrary $a \geq 1$ can be expressed as αT_{EXP} where $0 < \alpha \leq 1$. Then, the problem for finding the optimal T can be solved by seeking to α satisfying (12). Since (12) can be rewritten as

$$ae^{-a\mu_W T} = (a+1)\frac{D_W}{D_W + D_C} - e^{-\mu_W T/a}, \qquad (13)$$

we need to balance $ae^{-a\mu_W T}$ and $(a+1)\frac{D_W}{D_W+D_C} - e^{-\mu_W T/a}$ for the optimal timer. Moreover, $ae^{-a\mu_W T}$ decreases whereas $(a+1)\frac{D_W}{D_W+D_C} - e^{-\mu_W T/a}$ increases with the increase of T. Therefore, we devise a binary search algorithm for the optimal T as shown in Algorithm 1.

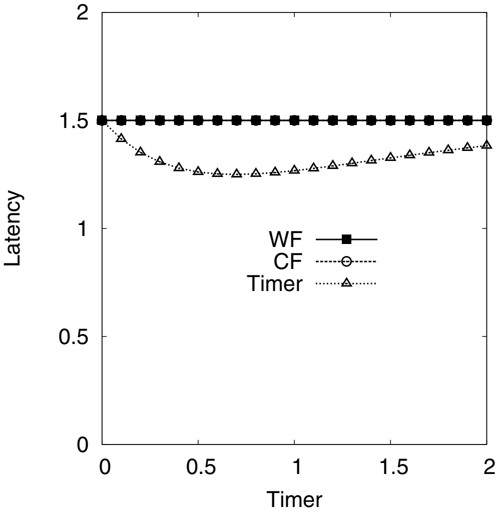
First, left and right bounds (i.e., $Left$ and $Right$) for binary search are set to 0 and T_{EXP} in lines 1 and 2, respectively. At line 4, T_{opt} is set to the middle value of $Left$ and $Right$. After that, two terms, denoted by $Term_A$ (i.e., the left-side term of (13)) and $Term_B$ (i.e., the right-side term of (13)), are evaluated at T_{opt} in line 7. If $Term_A$ is larger than $Term_B$, $Left$ is reset as T_{opt} to check the right half range of $(0, 1]$. Otherwise, $Right$ is reset as T_{opt} to evaluate the left half range of $(0, 1]$. These steps (from line 3 to line 12) are iterated until the optimal timer balancing $Left$ and $Right$ with a sufficiently small error δ (e.g., 10^{-3}) [7] is found.

Algorithm 1 Binary search algorithm for optimal T

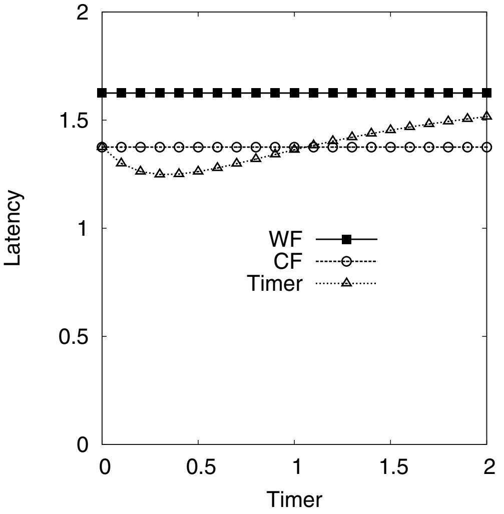
```
1:  $Left \leftarrow 0$ ;
2:  $Right \leftarrow \ln(1 + D_C/D_W) \ / \ \mu_W$ ;
3:  while $|Left - Right| > \delta$ do
4:      $T_{opt} \leftarrow (Left + Right) \ / \ 2$ ;
5:      $Term_A \leftarrow ae^{-a\mu_W T_{opt}}$ ;
6:      $Term_B \leftarrow (a+1)\frac{D_W}{D_W+D_C} - e^{-\mu_W T_{opt}/a}$;
7:      if $Term_A > Term_B$ then
8:          $Left \leftarrow T_{opt}$ ;
9:      else
10:         $Right \leftarrow T_{opt}$ ;
11:     end if
12: end while
```

IV. NUMERICAL RESULTS

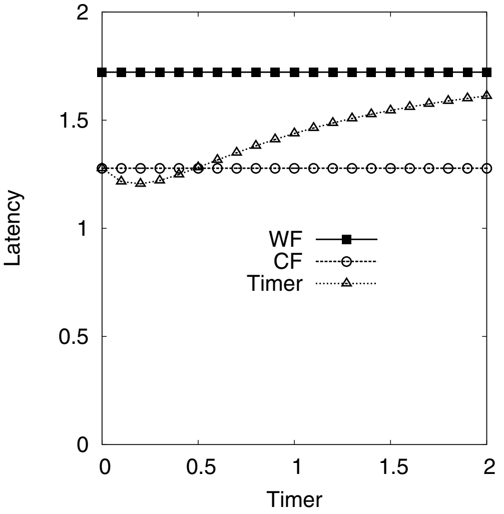
In this section, we compare the timer-based procedure against cellular-first (CF) and WLAN-first (WF) procedures. CF and WF procedures first check a cellular area and a WLAN area for an incoming session setup request, respectively. D_W and D_C are dependent on communication systems and both



(a) a=1



(b) a=3



(c) a=5

Fig. 5. Effect of timer.

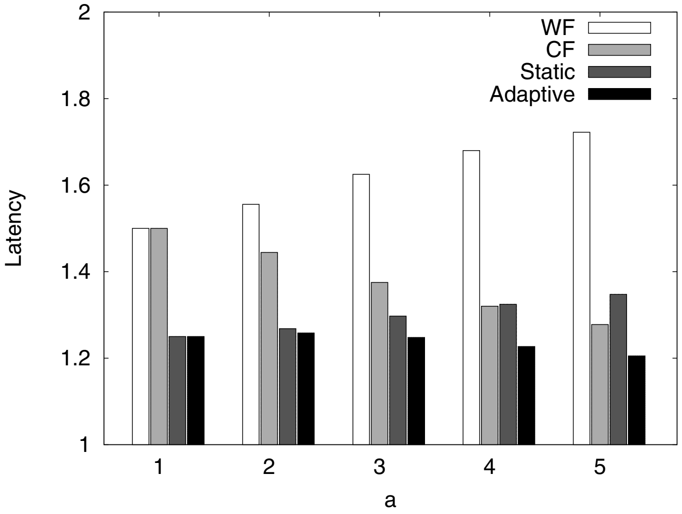


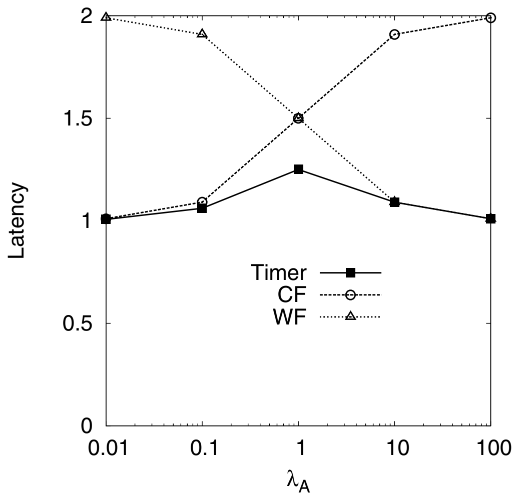
Fig. 6. Comparison of session setup latency.

values are set to 1.0 for simplicity. μ_W is normalized as 1.0 whereas different values of λ_A are evaluated.

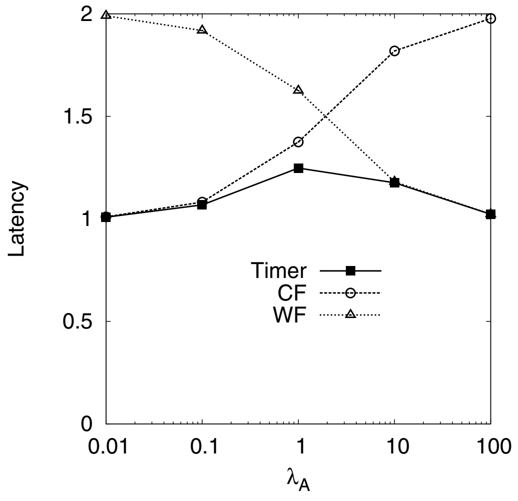
Figure 5 shows the effects of timer under different a. Since the CF and WF procedures are independent of the timer, they have constant session setup latencies regardless of the timer. Also, as a increases, CF outperforms WF because it is more likely that MTs stay in a cellular system for a longer time. When $a = 1$, the timer-based procedure is better than CF and WF even though non-optimal timers are set. However, the timer should be determined more carefully as a increases; otherwise, the timer-based procedure has a longer session setup latency than CF.

Figure 6 compares the average session setup latencies in CF, WF, and two timer-based procedures with static and adaptive timers. In the static timer, the optimal timer derived from the exponential WLAN residence time is used. On the other hand, the timer is adaptively set depending on a in the adaptive timer. It can be seen that the static timer works well when a is small. However, the static timer shows poor performance when a is large due to inappropriate setting of the timer. Unlike the static timer, the adaptive timer seeks to the most appropriate timer by Algorithm 1 and thus the session setup latency can be minimized for all values of a.

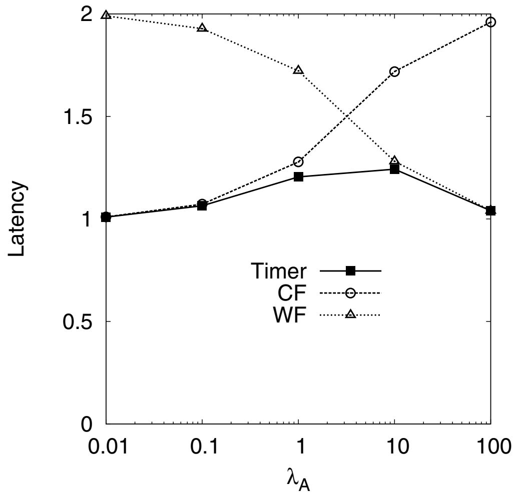
Figure 7 illustrates the effect of session arrival rate λ_A. As λ_A increases, the probability that an MT resides at the WLAN area for an incoming session setup request increases. Therefore, the performance of WF becomes better with the increase of λ_A. On the contrary, the session setup latency of CF increases drastically with λ_A. Also, it can be shown that the timer-based procedure outperforms CF and WF regardless of λ_A. This is because an adaptive timer is assumed in the timer-based procedure. In short, even when the WLAN residence time or the traffic pattern varies drastically, the proposed timer-based procedure can outperform WF and CF by adaptively setting the timer.



(a) a=1



(b) a=3



(c) a=5

Fig. 7. Effect of λ_A.

V. CONCLUSIONS

In this paper, we have proposed a timer-based session setup procedure in cellular-WLAN integrated systems for reducing the session setup latency. An analytical model on the session setup latency has been developed and the optimal timer for realistic WLAN residence time has been derived. Numerical results demonstrate that the proposed procedure with the adaptive timer can effectively reduce the session setup latency, and it can be extended to other types of heterogeneous wireless networks.

ACKNOWLEDGEMENT

This work was supported in part by the KRF grant (KRF-2008-314-D00354) and in part by MEST, Korea under WCU (World Class University) Project (R33-2008-000-10044-0).

REFERENCES

[1] Y. Xiao, K. Leung, Y. Pan, and X. Du, "Architecture, mobility management, and quality of service for integrated 3G and WLAN networks," *Wiley Wireless Communications and Mobile Computing*, vol. 5, no. 7, pp. 805-823, October 2005.

[2] S. Pack and W. Lee, "Dual Home Agent (DHA)-based Location Management Scheme in Cellular-WLAN Integrated Networks," *Elsevier Computer Networks*, vol. 57, no. 12, pp. 3273-3282, December 2008.

[3] Y. Lin and I. Chlamtac, "Heterogenesou Personal Communications Services: Integration of PCS Systems," *IEEE Communications Magazine*, vol. 3, no. 9, pp. 106-113, September 1996.

[4] S. Pack and H. Lee, "Call Setup Latency Analysis in SIP-based Voice over WLANs," *IEEE Communication Letters*, vol. 12, no. 2, pp. 103-105, February 2008.

[5] W. Song, H. Jiang, and W. Zhuang, "Performance analysis of WLAN-first scheme in cellular/WLAN interworking," *IEEE Transactions on Wireless Communications*, vol. 6, no. 5, pp. 1932-1952, May 2007.

[6] W. Song, Y. Cheng, and W. Zhuang, "Improving Voice and Data Services in Cellular/WLAN Integrated Network by Admission Control," *IEEE Transactions on Wireless Communications*, vol. 6, no. 11, pp. 4025-4037, November 2007.

[7] S. Pack, H. Jung, T. Kwon, and Y. Choi, "SNC: A Selective Neighbor Caching Scheme for Fast Handoff in IEEE 802.11 Wireless Networks," *ACM Mobile Computing and Communications Review*, vol. 9, no. 4, pp. 39-49, October 2005.

[8] S. Ross, *Stochastic Processes*. Wiley, 1995.