# An Apropos Signal Report and Adaptive Period (ASAP) Scheme for Fast Handover in the Fourth-Generation Wireless Networks

Jenhui Chen[a,1], Zhuxiu Yuan[b], Lei Wang[b,*]

[a] *Computer Science and Information Engineering, School of Electrical and Computer Engineering, College of Engineering, Chang Gung University, Kweishan, Taoyuan, Taiwan 33302, R.O.C.*
[b] *Software School, Dalian University of Technology, Dalian, Liaoning, China*

**Abstract**

In this paper, a medium access control (MAC) layer handover scheme is proposed to reduce the handover processing time. Two critical points of investigation are examined to improve the performance of the handover process. The first point is the exact point of time when the handover process should be triggered. This factor is highly related to the signal-to-noise ratio (SNR) the mobile station (MS) receives at that moment. The second point is the moving direction and current moving velocity of the MS. Based on these two factors, an *apropos signal report and adaptive period* (ASAP) handover scheme is proposed to reduce the handover processing time. ASAP is based on the concept that the received SNR is highly related to the point of time the MS reports and the circumstance the MS stays. If the variation between two reported SNR values is large, the time interval of receiving the next SNR report must be reduced, or increased when the variation is small, to accurately reflect the SNR change. ASAP adjusts the periodic ranging period dynamically according to the MS moving velocity and no additional control messages are introduced into ASAP. ASAP is a base station (BS) initiated handover scheme. Detailed comparisons reveal that ASAP can achieve a handover delay reaching 15 ms low at most, which is substantially smaller than the fastest legacy network-assisted handover mechanism (50 ms).

*Keywords:* Delay, handover, mobility, signal-to-noise ratio (SNR), velocity, wireless

## 1. Introduction

Handover schemes can be classified into two types: the hard handover and the soft handover [1, 2]. Determining an accurate trigger time for handovers is difficult when using the hard handover. The transmission should be continuous without being affected by the delay caused by the long term process of ranging, reassociation, and reauthorization. Soft handover, however, is investigated in this study to remedy this drawback of hard handover. Soft handover connects the radio link to the target base station (TBS) before breaking with the serving BS (SBS). Nevertheless, soft handover also leads to other drawbacks such as the radio resource being occupied by one mobile station (MS) on both sides during the handover time.

In previous studies, various feasible solutions were proposed to enhance handover by diminishing the number of turn-around times during the handover procedure [3], or by enabling the ability of BS to monitor the position of MSs [4, 5]. Using the global positioning system (GPS) is one common solution to measure the locations of devices, but the GPS may suffer from serious interfering problems caused by extreme weather conditions [6, 7]. Recently, numerous mechanisms without using the GPS to determine the location of devices were studied [8, 9, 10, 11]. Numerous research studies regarding the handover among femtocells were investigated [12, 13, 14]. These studies [15, 16, 17] considered the signal power into their proposed methods, but they neglected the effect of the MS's moving velocity. Some studies considered the mobility management but neglected using the signal power to assist handover decision making [18, 19].

In soft handover, because the MS holds the radio resources of cell $i$ and cell $i+1$ simultaneously during the handing over period, the handover processing time $E[\tau_i]$ (i.e., the handover preparation time plus the handover execution time) must be reduced. Two problems will cause handover failure:

1. The handover processing time is longer than the duration of the point of time, say $t_2$, that the received signal-to-noise ratio (SNR) value is below the SNR threshold minus the point of time, say $t_1$, that the MS starts the handover process, i.e., $E[\tau_i] > t_2 - t_1$.
2. The starting time of proceeding handover is too late.

*Corresponding author: lei.wang@dlut.edu.cn; Tel: +86-411-87571585.

*Email addresses:* jhchen@mail.cgu.edu.tw (Jenhui Chen), zhuxiu.yuan@gmail.com (Zhuxiu Yuan), lei.wang@dlut.edu.cn (Lei Wang)

[1] This work was supported in part by the National Science Council, Taiwan, R.O.C., under Contract NSC102-2221-E-182-032, and the High Speed Intelligent Communication (HSIC) research center, Chang Gung University, Taoyuan, Taiwan, R.O.C.

Thus, a handover mechanism that reflects the received SNR value (both in the MS side or the BS side), the MS moving velocity, and the network load of the TBS has been invented to improve the handover efficiency. This study reveals that the SNR value does not drastically change if the MS is stationary or moves slowly. Therefore, the received SNR is closely related to the location environment the MS stays. According to the observations that have been made in this study, an *apropos signal report and adaptive period* (ASAP) handover scheme has been invented to improve handover efficiency.

ASAP uses a periodic ranging function, which is a mandatory signaling function for maintaining connections, to measure signals reported by the MS. The reported signals can be used to determine

1. the time point when the handover is triggered by checking if the SNR value of the signal is under the threshold;
2. the location of the MS estimated by the SNR value and the position estimation function [9];
3. the moving velocity of the MS by calculating the difference between two consecutive MS positions over the time variance between the corresponding reported time points.

Based on the reported signal measurement, the reported time period between the two consecutive positions, must be maintained appropriately and dynamically. Thus, when the MS moving velocity is high, the period between the two time points of periodic ranging must be shortened to reflect the high velocity for making a more accurate handover time point. Conversely, when the MS velocity is low, the period has to be lengthened to reduce the monitoring overheads (i.e., reduce the cost of periodic ranging). By obtaining this information, the SBS can perform ranging, reassociation, and reauthorization with the TBS in advance according to the signals, velocity, and measured probability of handover to reduce the handover time. The contributions of this paper are highlighted as follows:

1. ASAP can dynamically change the periodic ranging period $T_r$ according to the MS velocity, which significantly increases the success probability of handover as well as reduces the cost of signal reporting;
2. ASAP can pre-coordinate the MS with the TBS for fast handover based on the information of the MS appearing in the overlap area and MS moving vector;
3. the velocity-aware ASAP scheme achieves an accurate handover without using GPS information.

The rest of this paper is organized as follows: Section 2 describes two standards of the fourth-generation mobile communication protocol in which ASAP can be applied; Section 3 presents the system model; the ASAP scheme is introduced in Section 4; the performance analysis of ASAP is given in Section 5; a series of numerical experiments on various performance metrics is reported in Section 6; and several conclusions are presented in Section 7.

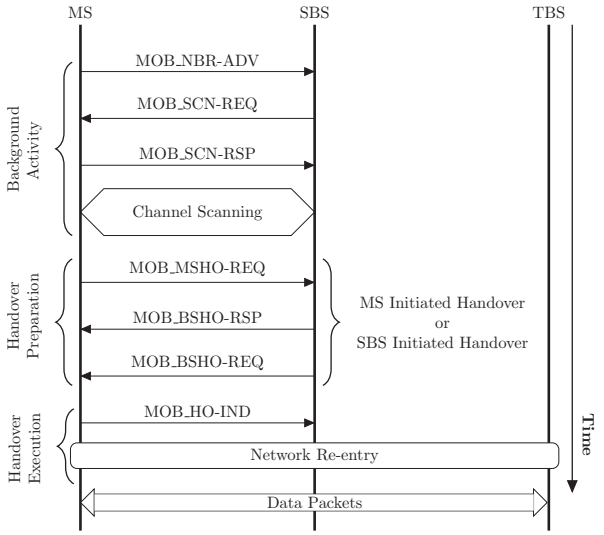



Figure 1: Handover procedure of the IEEE 802.16m protocol.

## 2. Background

In the fourth-generation long term evolution advanced (LTE-A) networks, the user equipment (UE) assists in the handover decision making by measuring neighboring cells and reporting the measurements to the serving eNodeB (SeNB), which then decides on handover timing and the target eNodeB (TeNB). The X2 handover procedure is used to handover the UE from a SeNB to a TeNB by using the X2 interface when the mobility management entity (MME) and serving gateway (SGW) are unchanged. This is only possible if direct connectivity with the X2 interface exists between the SeNB and TeNB. In this paper, only intra LTE-A handovers using X2 signaling (i.e., handovers in the intra-MME/SGW) are considered.

The X2 handover procedure is performed without the evolved packet core (EPC) involvement, i.e., preparation messages are directly exchanged between the SeNB and TeNB. The release of the resources at the serving side during the handover completion phase is triggered by the TeNB. As described in the evolved universal terrestrial radio access (E-UTRA) [21], the SeNB sends the measurement control message to request the UE to report the received SNR (the downlink SNR) in time. ASAP uses this compulsory action of the MS downlink SNR report to accomplish the velocity estimation if it is applied in the LTE-A networks.

Fig. 1 depicts the handover procedure contained in the IEEE 802.16m standard. The concept of the IEEE 802.16m handover procedure is similar to that of the LTE-A handover procedure. The only dissimilarity between the two is that IEEE 802.16m allows MS-initiated handover by using mobile MS handover request (MOB_MSHO-REQ) and mobile BS handover response (MOB_BSHO-RSP), or BS-initiated handover by using mobile BS handover request (MOB_BSHO-REQ). To perform the proactive handover, ASAP adopts the BS-initiated handover method.

In addition, the SNR measurement of the MS is obtained by using a periodic ranging action provided by the IEEE 802.16m standard. The two types of standards possess the same mechanism for reporting messages periodically to allow the SBS to receive the SNR value from the MS.

Furthermore, to reduce the handover time, also known as the *service disruption time*, the IEEE 802.16 working group proposed three types of soft handover to achieve this goal [22]. The WiMAX Forum [23] classifies the three types of handover mechanisms as the MS-initiated, contention-based handover without coordination, called uncoordinated association (Level 0), denoted as $L_0$; the BS-assisted MS association handover, called coordinated association (Level 1), denoted as $L_1$; and the network assisted association reporting (Level 2), denoted as $L_2$. The achieved minimum handover time of the three soft handover mechanisms are 280 ms ($L_0$), 230 ms ($L_1$), and 60 ms ($L_2$) [24]. However, the minimum handover time of $L_2$ is still too lengthy to satisfy the maximum tolerable delay time[2] [24, 25]. This is because when $L_2$ sends required handover messages to the TBS for the MS handover coordination, it cannot avoid the backhaul network transmission time, $T_d$. Although the delay time is slight, it causes $L_2$ to obtain a longer handover time, which does not satisfy the real-time services requirements.

Table 1: Summary of Important Notations Used

| Notation | Definition |
|---|---|
| $A_{i,i+1}$ | The overlap area between cell $i$ and cell $i+1$ |
| $D_s$ | The distance for crossing the overlay area |
| $d_{\max}$ | The radius of the cell |
| $x_i$, $y_i$ | The dwelling time and association time with the SBS in cell $i$ |
| $\lambda_n$, $\lambda_h$ | The arrival rate of the new call and the handover call |
| $t_i^s$, $t_i^e$ | The start time and the end time of receiving signals from cell $i$ |
| $\omega_i$ | The overlap time in overlap $A_{i,i+1}$ |
| $z_i$ | The time an MS stays in the non-overlap area of cell $i$ |
| $\tau$ | The network response time for a handover call |
| $T_h$ | The call holding time |
| $\eta$, $\zeta$, $\beta$, $\mu$ | The rates of exponential distributions $\omega_i$, $z_i$, $\tau$ and $T_h$ |
| $p_b$ | The new call blocking probability |
| $p_a$ | The probability that a handover call is blocked because no radio resource is available |
| $p_t$ | The probability that a handover call is blocked because the network response time is too long |
| $p_f$ | The forced termination probability that a handover call is blocked because no radio channel is available or the network response time is too long |
| $p_x$ | The handover call incompletion probability |
| $\phi$, $\delta$ | The handover preparation and execution time |
| $p_n$ | The steady-state probability of state $s(n)$ in an M/M/c birth-death process |
| $E[L_q]$, $E[W_q]$ | The expected queue length and the expected waiting time of the handover calls |
| $d(t_i)$ | The distance between an MS and the SBS |
| $D(t_i)$ | The Euclidean distance from time $t_{i-1}$ to $t_i$ |
| $v$ | The average moving velocity of an MS |
| $p_x[\cdot]$ | The handover failure probability |

## 3. System Model

### *3.1. The Handover Model*

The system is modeled as *cellular networks*, where MSs travel among *cells* over wireless fading channels. Each cell $i$ is dominated by one SBS residing inside the center of the cell. The MS processes handover procedures when it moves across the boundary of the serving cell to another cell (one of neighboring cells). Let $A_{i,i+1}$ denote the overlap area between cell $i$ and the neighboring cell $i+1$. The call arrivals to and from the MSs are assumed to be a Poisson process. The new call arrival rate to a cell is $\lambda_n$ and the handover call arrival rate to a cell is $\lambda_h$. Table 1 lists the notations used in this paper.

The timing diagram of the handover model is depicted in Fig. 2, where $x_i$ represents the dwelling time that an MS can receive signals from cell $i$, i.e., the period of time the MS resides in cell $i$. Let $t_i^s$ and $t_i^e$ be the starting and ending time of receiving signals from cell $i$. Assume an MS must enter an overlap area $A_{i,i+1}$ first before it enters cell $i+1$ from cell $i$. Let $\omega_i = t_i^e - t_{i+1}^s$ denote the overlap time that an MS can receive both signals from cell $i$ and cell $i+1$, i.e., the MS stays in $A_{i,i+1}$. And let $z_i$ be the time an MS stays in the non-overlap area of cell $i$, i.e., the MS can only receive signals from cell $i$. Let $\tau$ be the network response time for a handover call. Suppose $\omega_i$, $z_i$, and $\tau$ are exponentially distributed with rates $\eta$, $\zeta$, and $\beta$. Then let the expected values be $E[\omega_i] = 1/\eta$ and $E[z_i] = 1/\zeta$. The expected association time that an MS is served by the SBS in cell $i$ can be expressed as

$$E[y_i] = E[z_i] + E[\omega_i] = \frac{1}{v} = \frac{1}{\zeta} + \frac{1}{\eta} = \frac{\eta + \zeta}{\eta\zeta},$$

where $v$ is the average moving velocity of the MS. In a previous study, Lin and Pang [26] demonstrated that the MS's association time $y_i$ with the SBS in cell $i$ reveals an exponential distribution with the density function

$$a(y_i) = ve^{-vy_i}.$$

Assume the call holding time $T_h$ is exponentially distributed with the mean $1/\mu$. Let $p_a$ denote the handover call blocking probability and $p_b$ denote the new call blocking probability because no radio resources are available, respectively. According to the specifications of the IEEE 802.16m standard [22], a handover call receives a higher priority than a new call, i.e., $p_b > p_a$. Let $p_f$ be the handover failure probability that a handover call is blocked because no radio resources are available or the network response time is too long. Therefore, $p_f$ can be obtained by

$$p_f = 1 - (1 - p_t)(1 - p_a),$$

where $p_t$ is the probability that a handover call is blocked because the network response time $\tau$ is too long. Lin and

[2] Banerjee and his co-authors [25] concluded that a handover time of 50 ms is sufficient for media streams, and an interruption of 200 ms is generally acceptable. However, a handover time of 500 ms causes a perceptible interruption, which is unacceptable.

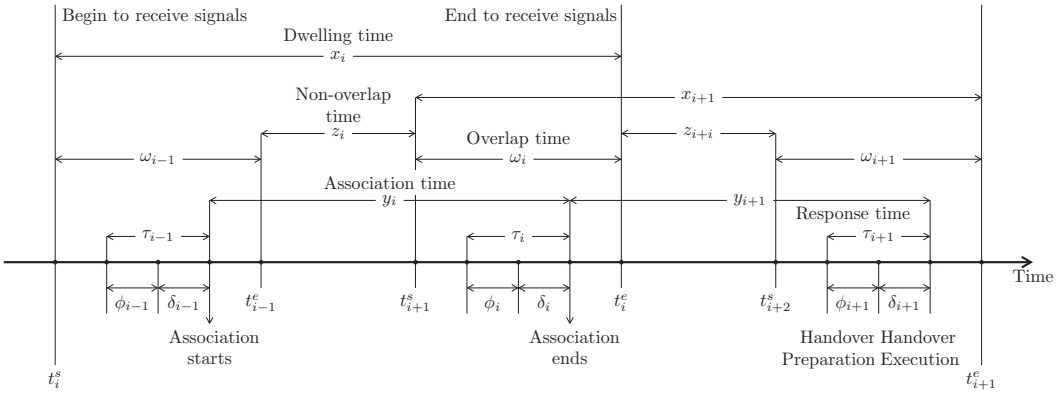



Figure 2: Time diagram for the IEEE 802.16 handover model.

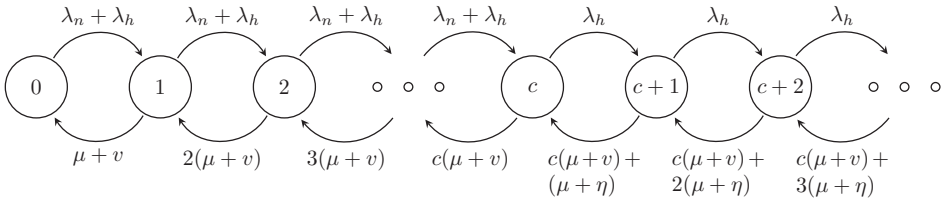



Figure 3: Rate transition diagram for the $M/M/c$ queue.

Pang [26] derived the probability $p_t$, which is

$$p_t = \eta/(\eta+\beta). \tag{1}$$

In another previous study [27], Lin derived the relationship between the handover call incompletion probability and $p_a$, $p_b$, and $p_t$ as

$$p_x = p_b + \frac{v(1-p_b)[1-(1-p_t)(1-p_a)]}{\mu+v[1-(1-p_t)(1-p_a)]}. \tag{2}$$

A network response time $\tau$ comprises a handover preparation time $\phi$ and a handover execution time $\delta$. To compute the expected handover execution time $E[\delta_i]$, the handover queue is modeled as an $M/M/c$ birth-death process [28, 29] as shown in Fig. 3. When the state is in $s(k)$ $(0 \le k \le c)$, there are $k$ busy channels, and the arrival rate is a constant of $\lambda_n+\lambda_h$. The state transition rate from $s(k)$ to $s(k-1)$ is $k(\mu+v)$, $0 < k \le c$, because the service-completion rate is $\mu+v$. When the state is in $s(c+j)$, $j \ge 0$, the arrival rate is a constant of $\lambda_h$ and any new call is dropped immediately. The service-completion rate (channel release rate from the $c$ connected calls) is $c(\mu+v)$ because all channels are busy. For those $j$ handover calls that wait for the second connections, before the second connections are available, the calls may leave the system if one of two cases occurs: 1) an MS leaves the overlap area with rate $\eta$, or 2) a call is completed with rate $\mu$. Thus, the service rate from $s(c+j)$ to $s(c+j-1)$ is $c(\mu+v)+j(\mu+\eta)$ for $j > 0$.

According to the $M/M/c$ process, the steady-state probabilities $p_n$ for $s(n)$ can be

$$p_n = \begin{cases} \dfrac{(\lambda_n+\lambda_h)^n}{n!(\mu+v)^n}p_0, & \text{for } n < c \\ \dfrac{(\lambda_n+\lambda_h)^c\lambda_h^{n-c}}{(c-1)!(\mu+v)^{c-1}\displaystyle\prod_{0\le j\le n-c}[c(\mu+v)+j(\mu+\eta)]}p_0, & \\ \text{for } n \ge c. & \end{cases}$$

Because $p_0+p_1+\cdots+p_n+\cdots=1$, the following equation is obtained

$$p_0 = \left(\sum_{n=0}^{c-1}\frac{(\lambda_n+\lambda_h)^n}{n!(\mu+v)^n} + \sum_{n=c}^{\infty}\frac{(\lambda_n+\lambda_h)^c\lambda_h^{n-c}}{(c-1)!(\mu+v)^{c-1}F(n)}\right)^{-1}, \tag{3}$$

where

$$F(n) = \prod_{0\le j\le n-c}[c(\mu+v)+j(\mu+\eta)].$$

The derivation of the handover call rate, as Lin indicated in his study [27], is

$$\lambda_h = \frac{v(1-p_b)\lambda_n}{\mu+v[1-(1-p_t)(1-p_a)]}. \tag{4}$$

### *3.2. Derivation for $E[W_q]$*

Let $E[L_q]$ be the expected queue length when the handover call must wait for the channel release. When $n \ge c$

and $m = n - c$, then

$$E[L_q] = \sum_{n=c+1}^{\infty} (n-c)p_n = \frac{(\lambda_n+\lambda_h)^c}{c!(\mu+v)^c} p_0 \sum_{m=1}^{\infty} \frac{m\lambda_h^m}{\prod_{1\le j\le m}[c(\mu+v)+j(\mu+\eta)]}. \quad (5)$$

Because the arrival rate to a state is $\lambda_h$ for $n \ge c$, the Little's formula is used to obtain

$$E[W_q] = \frac{E[L_q]}{\lambda_h}. \quad (6)$$

Substituting (4) and (5) into (6) results in

$$E[W_q] = \frac{p_0[\mu + v(p_t + p_a + p_t p_a)]}{\lambda_n v c!(1-p_b)} \left(\frac{\lambda_n+\lambda_h}{\mu+v}\right)^c \cdot \sum_{m=1}^{\infty} \frac{m\lambda_h^m}{\prod_{1\le j\le m}[c(\mu+v)+j(\mu+\eta)]}, \quad (7)$$

where

$$p_b = \sum_{n\ge c}^{\infty} p_n, \quad (8)$$

and, according to Lin and Pang [26],

$$p_a = \sum_{0\le j<\infty} \frac{p_{c+j}\eta(j+1)}{c(\mu+v)+(j+1)(\mu+\eta)}. \quad (9)$$

## 4. ASAP Scheme

In this section, ASAP is presented in detail. The flows of handover messages and the procedures of ASAP in time sequence are illustrated in Fig. 4.

**Step 1: Distance Estimation.** The SBS obtains the distance from the SBS to MS via the periodic ranging messages reported by MSs. Let $d_i$ denote the distance between an SBS and an MS at a point of time $t_i$, $i \in \mathbb{N}$, in which $t_{i-1} \prec t_i \prec t_{i+1}$. The SBS uses the distance estimation function based on the received SNR value [9] to obtain the distance $d_i$ at time $t_i$ as

$$d_i = \ell 10 \exp\left\{\left[P_t + G_t + G_r - 20\log\left(\frac{4\pi d_0 f}{c_L}\right) - X_\sigma - C_f - C_h - P_0 - \mathrm{SNR}(t_i) - N_0\right] \Big/ (10\rho_0)\right\}, \quad (10)$$

where $\ell$ is the free-space path loss (PL) with a known selection in the reference distance; $P_t$ is the transmitted power; $G_t$ and $G_r$ are the transmitter and receiver antenna gains; $f$ is the operating frequency; $X_\sigma$ is a zero-mean Gaussian distributed random variable; $c_L = 3 \times 10^8$ meters/sec; $C_f$ and $C_h$ are the frequency correction factor and the receiver antenna height correction factor; $P_0$ and $\mathrm{SNR}(t_i)$ are the system loss factor and the received SNR value at time $t_i$; $N_0 = -174(\mathrm{dBm}) + 10\log B + F(\mathrm{dB})$ is the noise power in dBm, where $B$ is the carrier bandwidth and $F$ is the receiver noise figure; and $\rho_0$ is the path loss exponent, where $\rho_0 = 2$ for free space, and is generally higher for wireless channels.

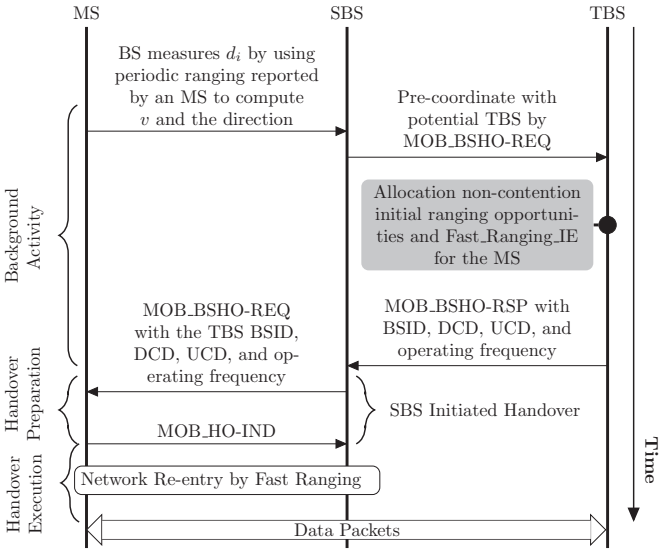
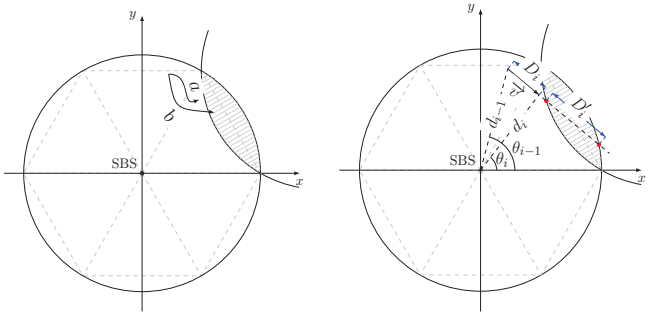



Figure 4: The sequence diagram of the ASAP scheme.




(a) Two traces of MSs. (b) Estimation of trace $a$.

Figure 5: MS movement estimation.

When $d_i > d_{i-1}$, the monitored MS is moving farther from the SBS. When $d_i < d_{i-1}$, the monitored MS is moving closer to the SBS. According to this comparison, the SBS can use the distance variation information to assist the handover preparation works.

**Step 2: Movement Estimation.** The SBS estimates the movement of the MS to improve the handover efficiency because the MS may move across the overlap area quickly or move around the overlap area without requiring a handover. Fig. 5(a) displays the original mobility tracking of an MS. Let the coordinate of SBS be $(0, 0)$. To estimate the moving distance from time $t_{i-1}$ to $t_i$, the Euclidean distance between the two locations, denoted as

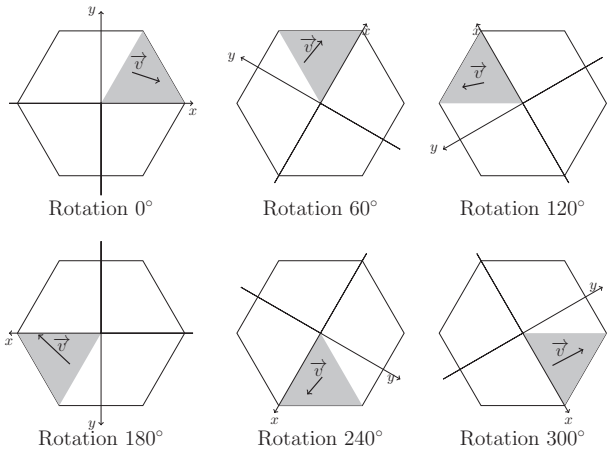



Figure 6: Six sectors of a cell and their corresponding coordinates.

$D_i$ in Fig. 5(b), is obtained by

$$D_i = \sqrt{d_{i-1}^2 + d_i^2 - 2d_{i-1}d_i\cos(\theta_i - \theta_{i-1})}. \quad (11)$$

Then the current MS moving velocity $v_i = D_i/(t_i - t_{i-1})$. To accurately determine the moving direction of MS, the MS motion vector is calculated by

$$\begin{aligned}\overrightarrow{v} &= (d_i\cos\theta_i - d_{i-1}\cos\theta_{i-1}, d_i\sin\theta_i - d_{i-1}\sin\theta_{i-1}) \\ &= (|\overrightarrow{v}|\cos\gamma, |\overrightarrow{v}|\sin\gamma), \end{aligned} \quad (12)$$

where $\theta_i$, $\theta_{i-1}$ can be estimated by using the smart antenna system [30] and $\gamma$ is the included angle of $\overrightarrow{v}$ with axis $x$, then

$$\gamma = \arctan\left(\frac{d_i\sin\theta_i - d_{i-1}\sin\theta_{i-1}}{d_i\cos\theta_i - d_{i-1}\cos\theta_{i-1}}\right). \quad (13)$$

Let $E[d]$ be the mean distance for crossing the overlap area $A_i$ and $R$ be the radius of a cell, then the mean distance across $A_i$ can be obtained by

$$E[d] = \int_{\theta=0}^{\pi/3}\frac{3}{\pi}\left(2R\sin\frac{\theta}{2}\right)d\theta. \quad (14)$$

Therefore, the next point of time that the MS should report a periodic ranging message is set as

$$t_{i+1} = t_i + \frac{E[d]}{v_i}. \quad (15)$$

Based on (15), the next ranging period $T_r = t_{i+1} - t_i$.

**Step 3: Pre-coordinate with TBS.** The SBS checks whether the MS reaches the overlap area at $t_i$. Since each cell is divided into 6 sectors, ASAP uses 6 corresponding coordinates, as shown in Fig. 6, to monitor the MS moving direction when the MS appears in different sectors. Let the distance from the SBS to the edge of overlap area along $d_i$ be $d_e$ as shown in Fig. 7. Using the law of consines, we have

$$\begin{aligned}R^2 = d_e^2 + (2R\cos(\pi/6))^2 \\ - 4d_eR\cos(\pi/6)\cos(\theta_i - \pi/6), \end{aligned} \quad (16)$$

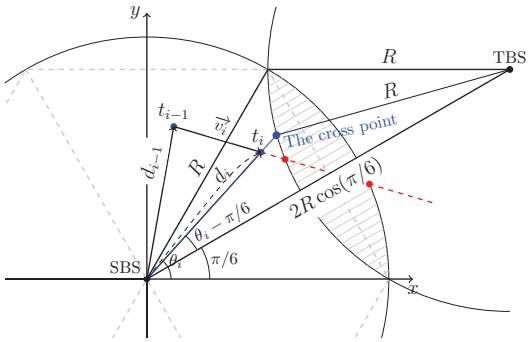



Figure 7: Diagram of checking the MS in overlap area or not.

where $R$ is the radius of a cell. If the MS is not in the overlap area, it must satisfy the condition of $d_i < d_e$; based on (10) and solving (16) for $d_e$, we have

$$\begin{aligned}d_i < 2R\cos(\pi/6)\cos(\theta_i - \pi/6) \\ + R\sqrt{1 + 4\cos^2(\pi/6)[1 + \cos^2(\theta_i - \pi/6)]}. \end{aligned} \quad (17)$$

If the monitored MS reaches the overlap area (i.e., $d_i \geq d_e$) and $\cos(\gamma + \pi/6) > 0$, the SBS sends MOB_BSHO-REQ to the TBS for the upcoming handover usage.

**Step 4: TBS Reply.** If the TBS resources are available, the TBS agrees to the handover request by sending a handover respond (MOB_BSHO-RSP) message, which includes the TBS BSID, operating channel, DCD, UCD, and CQICH_ID information [22], to the SBS to inform the MS. In addition, the TBS sends CQICH_Alloc_IE and a Fast_Ranging_IE information message by using an UL-MAP for a time period $w_i$. This time period is set to avoid system resources waste if the MS cannot finish the handover procedure.

**Step 5: SBS-initiated Handover.** The SBS sends MOB_BSHO-REQ with the TBS BSID, DCD, UCD, operating channel, and CQICH_ID to the MS for fast handover usage.

**Step 6: Fast Handover.** After receiving MOB_BSHO-REQ from the SBS, the MS immediately sends a handover indication (MOB_HO-IND) message to the SBS and then performs the fast handover with the TBS, as indicated in MOB_BSHO-REQ.

**Termination.** ASAP terminates here.

In the following, an example of MS monitoring and tracing scenario is presented to demonstrate the ASAP process in a cell. Fig. 8 illustrates the tracks of a moving MS traced by ASAP with corresponding points of time. Suppose that the diameter of the cell is 1000 meters long. The monitored information obtained by using provided estimation functions is presented in Table 2. When the SBS detects that the MS enters the overlap area, it checks whether $\sin\gamma > 0$. According to the direction of the MS measured from $t_4$ to $t_5$, the example demonstrates that $\sin\gamma > 0$, and then the SBS sends MOB_BSHO-REQ to the possible TBS for the upcoming handover usage. The SBS also computes an estimated time period $T_r = t_6 - t_5 = E[d]/v$ by

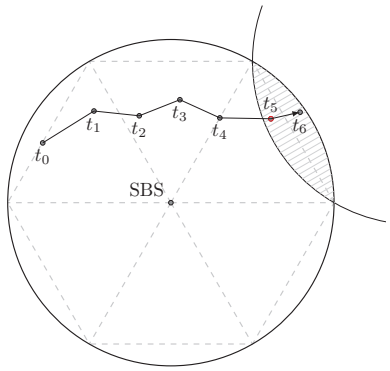



Figure 8: An MS moving and tracing scenario by using ASAP.

Table 2: Periodic Monitoring Parameters

| | $d_i$ (m) | $d_{i-1}$ (m) | $\theta_i$ (rad) | $\theta_{i-1}$ (rad) | $T_r$ (sec) |
|---|---|---|---|---|---|
| $t_0$ | 433.3 | 0 | 2.705 | 0 | 10 |
| $t_1$ | 366.7 | 433.3 | 2.269 | 2.705 | 5.4 |
| $t_2$ | 283.3 | 366.7 | 1.92 | 2.269 | 0.116 |
| $t_3$ | 316.7 | 283.3 | 1.484 | 1.92 | 0.30 |
| $t_4$ | 300 | 316.7 | 1.047 | 1.484 | 0.38 |
| $t_5$ | 400 | 300 | 0.698 | 1.047 | 0.301 |
| $t_6$ | 483.3 | 400 | 0.611 | 0.698 | 0.059 |

(15) for fast handover usage. If the MS does not send any handover indication (MOB_HO-IND) message to the SBS before the end of $w$, the TBS cancels the resource which is prepared for the handover of the MS to avoid unwanted resource occupation (wasting).

## 5. Theoretical Analysis

Fig. 9 illustrates the required time for handling the handover call process by adopting various schemes $L_0$, $L_1$, $L_2$, and ASAP. In this system, the network response time includes the handover preparation time and handover execution time.

### *5.1. Handover Incompletion Probability of $L_0$*

$L_0$ is a non-association ranging handover mechanism. The ranging code transmission is based on the contention-based resolution scheme with a random backoff contention window $W_i = 2^{i+5}$, $i = 0, 1, \ldots, 5$, where $W_{\min} = W_0$ and $W_{\max} = W_5$. The ranging process (including the new call and the handover call) is a Poisson process with a mean arrival rate $\lambda_n + \lambda_h$, therefore, the probability that there are no arrivals in time period $t$ is

$$p_0(t) = e^{-(\lambda_n+\lambda_h)t}. \tag{18}$$

Therefore, the ranging call arrival probability in time period $t$ is

$$p_r(t) = 1 - p_0(t) = 1 - e^{-(\lambda_n+\lambda_h)t}. \tag{19}$$

According to the memoryless property of the exponential distribution, the arrival probability of the ranging calls is equal to $1-e^{-(\lambda_n+\lambda_h)t}$, which is independent of the random backoff mechanism.

Let $N$ be the mean number of MSs in a cell. The mean number of MSs that attempt to perform the ranging process can be obtained by

$$N_r = Np_r(t) = N(1 - e^{-(\lambda_n+\lambda_h)t}). \tag{20}$$

A simple modification is made to the findings of a previous study [31] to calculate the expected contention delay of the $L_0$ handover scheme as

$$E[T_C] = \frac{1}{2}\left[\frac{1}{1-p_c} - \frac{(M+1)p_c^{M+1}}{1-p_c^{M+1}} + \frac{W_0}{1-r}\left(1 - \frac{r(1-p_c)}{1-p_c^{M+1}}\sum_{n=0}^{M}(rp_c)^n\right)\right] - 1, \tag{21}$$

where $M$ is the maximum retry times, $r$ is the binary exponential backoff, and the collision probability at each attempt is

$$p_c = 1 - (1 - p_r(t))^{N_r-1}. \tag{22}$$

There is no handover preparation in $L_0$, therefore, the excepted network response time $E[\tau_{L_0}]$ is the excepted handover execution time

$$E[\tau_{L_0}] = \frac{1}{\beta_0} = E[\delta_{L_0}] = E[T_C] + E[W_q] + T_{\text{RRQ}} + T_{\text{RA}} + T_{\text{RR}} + T_{\text{SYN}}, \tag{23}$$

where $T_{\text{RRQ}}$ is the mean time for RNG-REQ, $T_{\text{RA}}$ is the mean time for reauthorization during the handover, $T_{\text{RR}}$ is the mean time for re-registration during the handover, and $T_{\text{SYN}}$ is the mean time for the frame synchronization of a new downlink. By applying (1), (2), (8), (9), and (23), the handover incompletion probability of $L_0$ can be obtained by

$$p_x[L_0] = p_b + \frac{v(1-p_b)[1-(1-\eta/(\eta+\beta_0))(1-p_a)]}{\mu + v[1-(1-\eta/(\eta+\beta_0))(1-p_a)]}. \tag{24}$$

### *5.2. Handover Incompletion Probability of $L_1$*

Similarly, let $E[\tau_{L_1}]$ be the expected network response time of $L_1$, then

$$E[\tau_{L_1}] = \frac{1}{\beta_1} = E[\phi_{L_1}] + E[\delta_{L_1}] = T_{\text{MS}\leftrightarrow\text{SBS}} + T_{\text{SBS}\leftrightarrow\text{TBS}} + T_R + T_d + T_{\text{RA}} + T_{\text{RR}} + T_{\text{SYN}}, \tag{25}$$

where $T_{\text{MS}\leftrightarrow\text{SBS}}$ is the mean transmission time from the MS to the SBS, and $T_{\text{SBS}\leftrightarrow\text{TBS}}$ is the mean transmission time from the SBS to the TBS. $T_d$ is the mean backhaul network transmission delay required to send handover messages to the TBS. The first frame immediately followed by the rendezvous time is denoted as $T_R$, which includes the

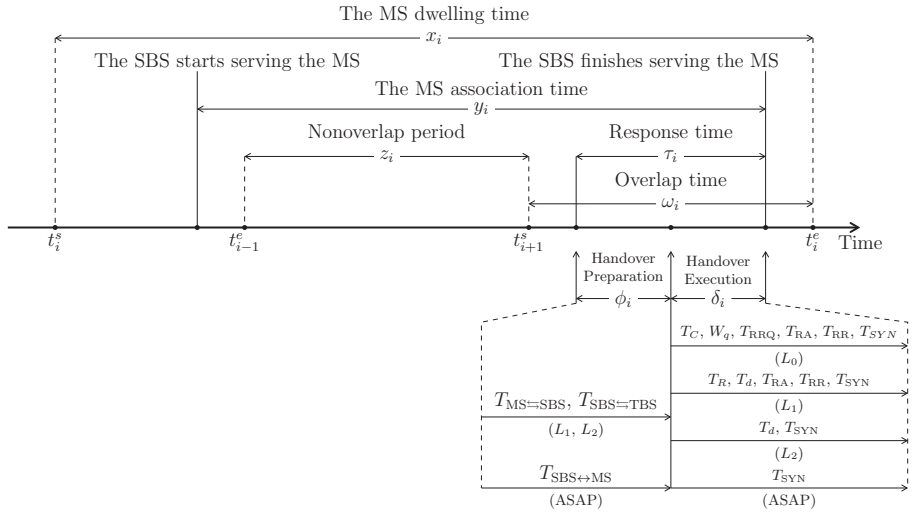



Figure 9: Timing diagram for analyzing the performance of handover mechanisms.

uplink map (UL-MAP) transmitting time. The typical rendezvous time is around 0–500 ms [32]. $T_{\mathrm{RA}}$, $T_{\mathrm{RR}}$, and $T_{\mathrm{SYN}}$ are the same as those in $L_0$.

In accordance with (24), the handover incompletion probability of $L_1$ is

$$p_x[L_1] = p_b + \frac{v(1-p_b)[1-(1-\eta/(\eta+\beta_1))(1-p_a)]}{\mu + v[1-(1-\eta/(\eta+\beta_1))(1-p_a)]}. \tag{26}$$

### 5.3. Handover Incompletion Probability of $L_2$

Similarly, let $E[\tau_{L_2}]$ be the expected network response time of $L_2$, then

$$\begin{aligned} E[\tau_{L_2}] &= \frac{1}{\beta_2} \\ &= E[\phi_{L_2}] + E[\delta_{L_2}] \\ &= T_{\mathrm{MS}\leftrightarrow\mathrm{SBS}} + T_{\mathrm{SBS}\leftrightarrow\mathrm{TBS}} + T_d + T_{\mathrm{SYN}}. \end{aligned} \tag{27}$$

By applying (24), the handover incompletion probability of $L_2$ is

$$p_x[L_2] = p_b + \frac{v(1-p_b)[1-(1-\eta/(\eta+\beta_2))(1-p_a)]}{\mu + v[1-(1-\eta/(\eta+\beta_2))(1-p_a)]}. \tag{28}$$

### 5.4. Handover Incompletion Probability of ASAP

Let $E[\tau_A]$ be the expected network response time of ASAP, then

$$\begin{aligned} E[\tau_A] &= \frac{1}{\beta_A} \\ &= E[\phi_A] + E[\delta_A] \\ &= T_{\mathrm{SBS}\leftrightarrow\mathrm{MS}} + T_{\mathrm{SYN}}. \end{aligned} \tag{29}$$

The corresponding handover incompletion probability of ASAP is

$$p_x[A] = p_b + \frac{v(1-p_b)[1-(1-\eta/(\eta+\beta_A))(1-p_a)]}{\mu + v[1-(1-\eta/(\eta+\beta_A))(1-p_a)]}. \tag{30}$$

### 5.5. Delay Estimation

According to (7), (23), (25), (27), and (29), the expected handover delays of $L_0$, $L_1$, $L_2$, and $L_A$, denoted as $E[T_0]$, $E[T_1]$, $E[T_2]$, and $E[T_A]$, are as follows:

$$E[T_0] = E[\tau_{L_0}]; \tag{31}$$

$$E[T_1] = E[\tau_{L_1}] + E[W_q]; \tag{32}$$

$$E[T_2] = E[\tau_{L_2}] + E[W_q]; \tag{33}$$

$$E[T_A] = E[\tau_A] + E[W_q]. \tag{34}$$

## 6. Numerical Result

In this section, various numerical results are presented to demonstrate how the mechanism achieves a fast handover. The simulation model-specific parameters follow the WiMAX Forum [23] and are listed in Table 3. The orthogonal frequency-division multiple access (OFDMA) frame length is 5 ms long. Any message transmission must follow frame by frame, and the time of a one-way transmission cannot be less than 5 ms.

Table 3: Parameters

| Parameter | Value |
|---|---|
| Avg. Number of MSs | 100 |
| No. of subchannels ($c$) | 32 |
| Max. retry times ($M$) | 6 |
| Binary exponential backoff ($r$) | 2 |
| Avg. time of contention ranging ($T_{\mathrm{CR}}$) (ms) | 120 |
| Avg. time of re-auth. ($T_{\mathrm{RA}}$) (ms) | 175 |
| Rendezvous time ($T_{\mathrm{R}}$) (ms) | 50 |
| Avg. time of re-reg. ($T_{\mathrm{RR}}$) (ms) | 35 |
| Avg. time of RNG-REQ ($T_{\mathrm{RNG}}$) (ms) | 25 |
| Avg. time of frame synchronize ($T_{\mathrm{SYN}}$) (ms) | 5 |
| Avg. time of MS $\leftrightarrow$ SBS ($T_{\mathrm{MS}\leftrightarrow\mathrm{SBS}}$) (ms) | 10 |
| Avg. time of SBS $\leftrightarrow$ TBS ($T_{\mathrm{SBS}\leftrightarrow\mathrm{TBS}}$) (ms) | 10 |
| Avg. backhaul network transmission time ($T_d$) (ms) | 50 |

### 6.1. Handover Delay

The effect on the handover delay of four handover schemes, $L_0$, $L_1$, $L_2$, and ASAP, by the network load, the

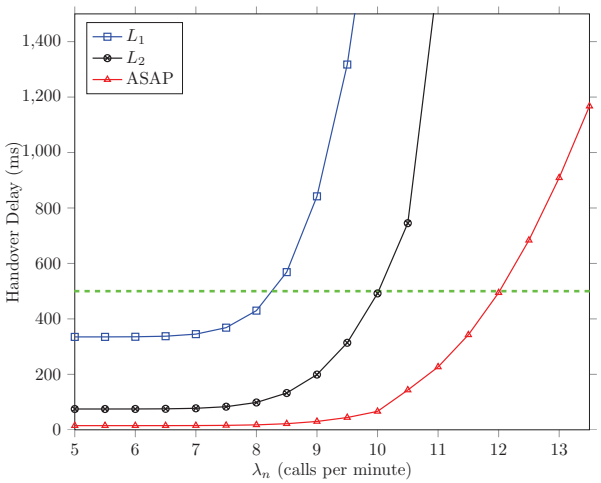



Figure 10: The effect of the network load on the handover delay with $\eta = 10$ seconds.

Table 4: Minimum Network Response Time

| Handover level | Notation | Response time (ms) |
|---|---|---|
| $L_0$ | $\beta_0$ | 530 |
| $L_1$ | $\beta_1$ | 335 |
| $L_2$ | $\beta_2$ | 75 |
| ASAP | $\beta_A$ | 15 |

overlap time ($\omega_i$), and the non-overlap time ($z_i$) is demonstrated. The minimum network response time of the four handover levels is also obtained, as shown in Table 4. The minimum network response time of ASAP is only 15 ms, which is substantially lower than that of the other three schemes and the typical rendezvous time of 500 ms. The effect of $L_0$ does not require further discussion because the minimum network response time of $L_0$ is greater than 500 ms.

### *6.1.1. Effects of Network Load*

The increment of network load, i.e., the number of new calls and the number of handover calls to the cell, can affect the handover delay, which leads to a longer waiting time. Therefore, the network load is a critical factor in handover call delay. To investigate the effects of the network load on handover delay, the handover delay in various network loads is discussed, and other factors that simultaneously affect the handover delay are controlled, as shown in Table 5, where $E[T_h]$ represents the mean call holding time. In this study, the network load is measured by using the new call arrival rate ($\lambda_n$), and increases when $\lambda_n$ increases. Fig. 10 illustrates the handover delay as a function of $\lambda_n$. Apparently, ASAP demonstrates a substantially lower handover delay than $L_1$ and $L_2$ under any network load, and slowly increases when $\lambda_n$ increases. Specifically, when $\lambda_n$ is less than 12 calls/minute, ASAP can provide an acceptable handover delay and $L_1$ and $L_2$ demonstrate considerably higher handover delays. ASAP can provide a faster handover compared with that of $L_1$ and $L_2$ because ASAP is an SBS-initial handover mechanism that exchanges necessary information between the SBS and the TBS in the handover preparation process and obtains SNR value in time, which expedites the handover process and reduces the waiting time.

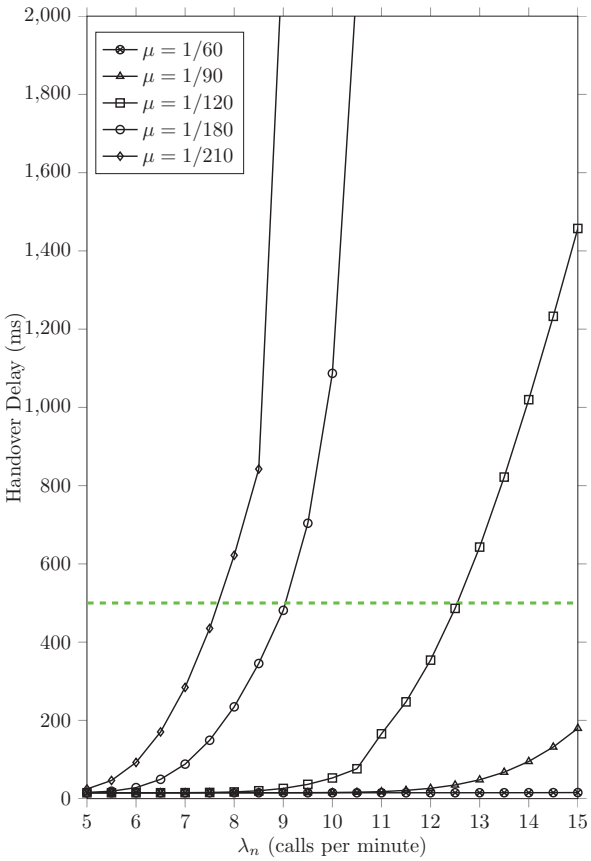



Figure 11: The handover delay of ASAP against $\lambda_n$.

### *6.1.2. Effects of Call Holding Time*

The available number of channels for new arriving calls, regardless of whether handover calls or new calls, greatly affects the handover delay and the handover success probability because the network resources, i.e., radio channels, are limited. The handover delay of ASAP $E[T_A]$ is plotted against $\lambda_n$ under the expected call holding time $E[T_h] = 1/\mu$ in Fig. 11. In this experiment, the handover delay increases as $\lambda_n$ and $E[T_h]$ increase. Two facts are derived from the obtained results. First, the impact of $\lambda_n$ on $E[T_A]$ is small because the handover calls receive higher priority than new calls. This explains why each $E[T_A]$ is almost same when $E[T_h] = 60$ seconds. Second, the handover delay increases sharply when $E[T_h]$ increases. This is because the longer the mean call holding time $E[T_h]$ is, the less channel release occurs. Therefore, the handover delay is greatly affected by the available number of channels. Furthermore, when $\lambda_n = 7.5$ calls per minute, the handover delay of ASAP occurs within the typical rendezvous time (i.e., 500 ms) when the call holding time is 210 seconds ($\mu = 1/210$). This result indicates that ASAP

Table 5: Scenario Settings.

| Performance | Scenarios | Parameters | | | |
|---|---|---|---|---|---|
| | | $E[T_h] = 1/\mu$ (sec) | $E[w_i] = 1/\eta$ (sec) | $E[z_i] = 1/\zeta$ (sec) | $\lambda_n$ |
| Handover Delay | Network Load | $\mu = 1/120$ | $\eta = 20\mu$ | $\zeta = 0.5\mu$ | variable |
| | | $\mu = 1/180$ | $\eta = 20\mu$ | $\zeta = 0.5\mu$ | variable |
| | Overlap Time | $\mu = 1/120$ | variable | $\zeta = 0.5\mu$ | $\lambda_n = 30\mu$ |
| | | $\mu = 1/180$ | variable | $\zeta = 0.5\mu$ | $\lambda_n = 30\mu$ |
| | Non-Overlap Time | $\mu = 1/120$ | $\eta = 20\mu$ | variable | $\lambda_n = 20\mu$ |
| | | $\mu = 1/180$ | $\eta = 20\mu$ | variable | $\lambda_n = 20\mu$ |
| Failure Probability | Network Load | $\mu = 1/120$ | $\eta = 20\mu$ | $\zeta = 0.5\mu$ | variable |
| | Overlap Time | $\mu = 1/120$ | variable | $\zeta = 0.5\mu$ | $\lambda_n = 30\mu$ |
| | Non-overlap Time | $\mu = 1/120$ | $\eta = 20\mu$ | variable | $\lambda_n = 20\mu$ |

can obtain low handover delay when $E[T_h]$ is extended.

#### *6.1.3. Effects of Overlay Time*

The handover delay is affected by the overlay time, which is related to the MS moving velocity in the overlay area, because ASAP is only triggered if the MSs enter the overlap area. In Fig. 12, the handover call rate and the handover delay are plotted against the overlap time $E[\omega_i]$. Fig. 12(a) indicates that the overlap time substantially affects the handover rate of ASAP. In other words, ASAP is particularly sensitive to MS mobility in the overlap area. Fig. 12(a) demonstrates that ASAP dramatically increases handover rate as $E[\omega_i]$ decreases, and decreases the handover rate as $E[\omega_i]$ increases. This result reveals that ASAP performs the handover according to the velocity change. In addition, the effect caused by ASAP prevents the ping-pong effect when the BS manages the handover requests. By comparing Fig. 12(a) and (b), ASAP is revealed to achieve a fast handover during a small overlap time, even if the network load increases.

#### *6.1.4. Effects of Non-overlap Time*

Fig. 13 illustrates the impact of $E[z_i]$ on handover rate and handover delay. The values of various parameters used in this experiments are listed in Table 5. Fig. 13(a) demonstrates that the handover rate increases when $E[z_i]$ decreases. The handover rate must increase because the $E[z_i]$ value is small. However, the effect of $E[z_i]$ on the handover delay of ASAP is not noticeable, as shown in Fig. 13(b). This result reveals that ASAP demonstrates optimal handover delay performance compared with that of the $L_1$ and $L_2$ handover schemes. The handover delay achieved by ASAP is only 86 ms, which is considerably lower than the 618 ms achieved by $L_2$ and lower than the threshold of 500 ms. This result shows that the handover delay achieved by ASAP is not affected by the $E[z_i]$ because the SBS monitors the MS and estimates the MS velocity periodically.

### *6.2. Handover Failure Probability*

In this subsection, the effect of $\lambda_n$ and $E[\omega_i]$ on the handover failure probability (incompletion probability $p_x[\cdot]$) and the new call blocking probability $p_b[\cdot]$ is discussed.

#### *6.2.1. Effects of Network Load*

Because the handover delay increases as the network load increases, the effect of the network load on $p_x[\cdot]$ is first investigated. As shown in Fig. 14(a), $p_x[\cdot]$ increases slightly as $\lambda_n$ increases. This result indicates that the new call arrival rate affects the success probability of handover, despite the priority of new calls being lower than that of handover calls. Fig. 14(b) demonstrates that the $p_b[\cdot]$ of ASAP is higher than that of $L_1$ and $L_2$. However, the new calls blocking probability caused by ASAP does not substantially increase. The result reveal that ASAP triggers the handover procedure for the MS in time according to the variation of the SNR value and the estimated MS velocity. This proactive monitoring scheme allows the SBS to make the correct handover decision at an accurate point of time. This design greatly reduces the possibility of unnecessary handover events, which cause numerous handover overheads because of the ping-pong effect. Fig. 14(a) demonstrates that $p_x[A]$ is lower than $p_x[L_2]$ and $p_x[L_1]$. Thus, the handover failure probability can be decreased without increasing the new call blocking probability if one handover mechanism can accurately trigger the handover at the correct time point and prevent the ping-pong effect. This result indicates that ASAP can allow MS handover calls to receive the radio resource in time, which takes advantage of obtaining the latest SNR of the monitored MS to estimate an accurate time point to trigger the MS handover.

#### *6.2.2. Effects of Overlap Time*

A longer overlap time indicates that the MS moving velocity is low when the MS is moving across the overlap area. In this experiment, the effect of overlap time (the MS velocity across the overlap area) on $p_x[\cdot]$ and $p_b[\cdot]$ is investigated. Fig. 15(a) demonstrates that $p_x[A]$ remains constant when $E[\omega_i]$ increases. This result reveals that ASAP can reflect the MS velocity to trigger the handover procedures and, thus, keep $p_x[A]$ low. Although ASAP achieves a relatively lower $p_x[\cdot]$ compared with that of $L_1$ and $L_2$, the gap between the $p_b[\cdot]$ values of ASAP and $L_1$ is about 3/11 of the gap between the $p_x[\cdot]$ values of ASAP and $L_1$. This result indicates that ASAP is more effectively able to fulfill the demand for fast handover requests.

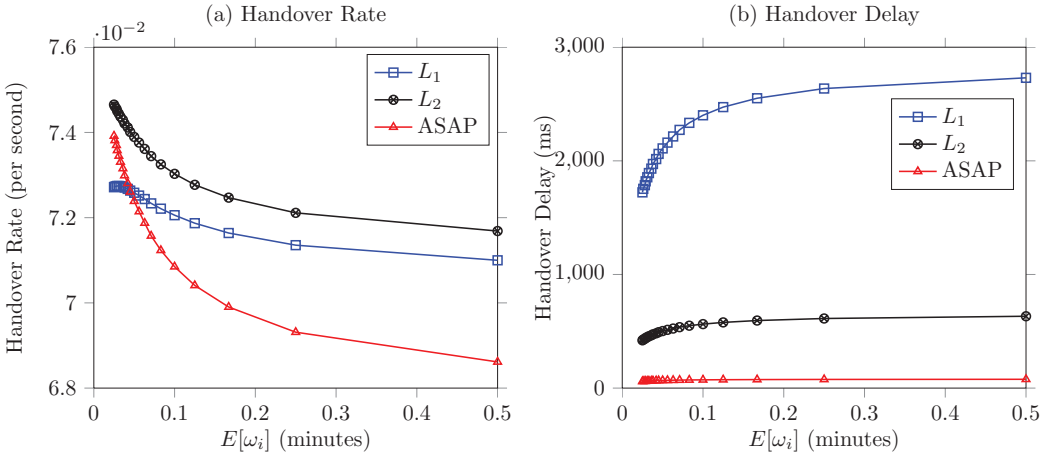



Figure 12: The effect of the overlap time $E[\omega_i]$.

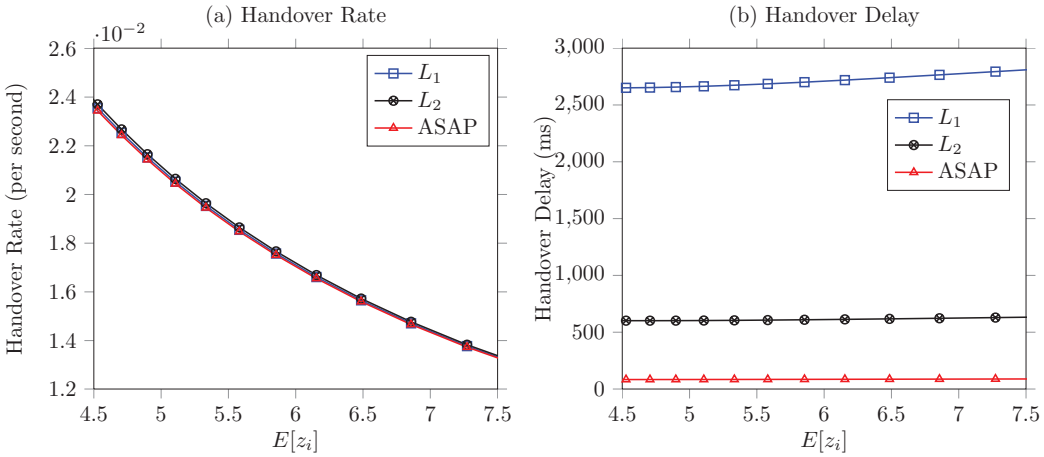



Figure 13: The effect of the non-overlap time $E[z_i]$ on handover rate and handover delay.

*6.2.3. Effects of Handover Call Rate*

The handover call rate $\lambda_h$ affects the handover success probability directly because the larger the handover call rate is, the higher the handover failure probability is. In this study, the handover success probability is plotted against the handover call rate in Fig. 16. The handover success probability falls when the handover call rate increases. The handover success probability falls more sharply in a light new call arrival rate (e.g., $\lambda_n = 5$ calls per minute) than a heavy one. When the handover call rate is small (i.e., $\lambda_h < 1$ calls per second or 60 calls per minute), the new call arrival rate ($\lambda_n$) affects the success probability because a heavy network load ($\lambda = \lambda_n + \lambda_h$) means a higher handover call rate. When the handover call rate is large, the handover processing time becomes the most important factor and contributes directly the success probability.

*6.3. Cost of Ranging Overhead*

To evaluate the handover cost caused by ASAP, a simple simulation is added to demonstrate the ranging overhead of ASAP, $L_1$, and $L_2$. The radius of a cell is assumed as $R = 1000$ meters and the distance between two neighboring BSs is 1800 meters. The mean length for crossing an overlap area is $E[d] = 511.74$ meters. The default period of periodic ranging is set as 10 seconds for necessary connection maintenance. Since $L_1$ and $L_2$ do not change the ranging period adaptively, they have a fixed ranging overhead, 6 times per minutes as shown in Fig. 17, no matter how the MS moving velocity is. In ASAP, since the ranging period is changed adaptively according to the velocity, the ranging overhead is proportional to the MS moving velocity. This result reveals a fact that ASAP enhances the handover efficiency as well as reduces the cost of handover.

## 7. Conclusion

In this paper, the ASAP handover scheme for macro BS, which periodically monitors moving MSs and prepares CDMA ranging codes for handover in advance is investigated. Numerical results indicate that the handover failure probability and the handover delay can be substan-

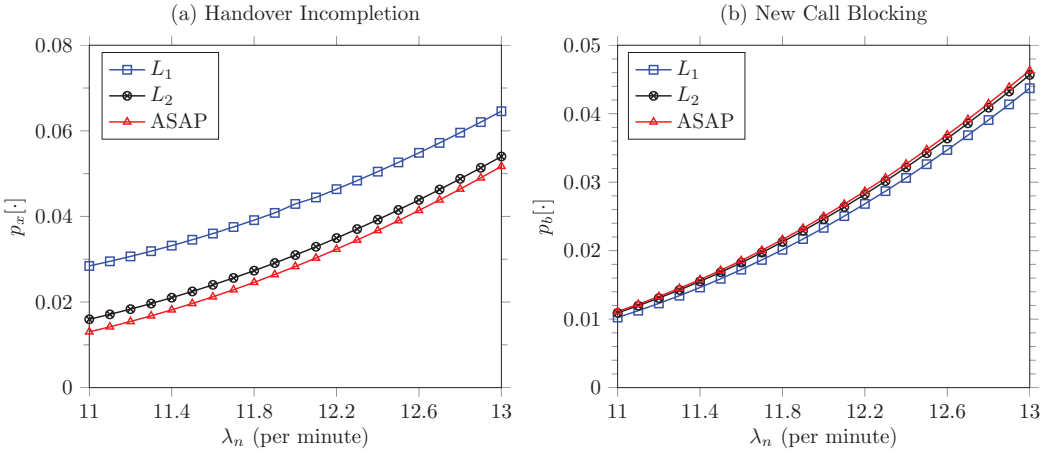



Figure 14: The effect of $\lambda_n$ on $p_x[\cdot]$ and $p_b[\cdot]$.

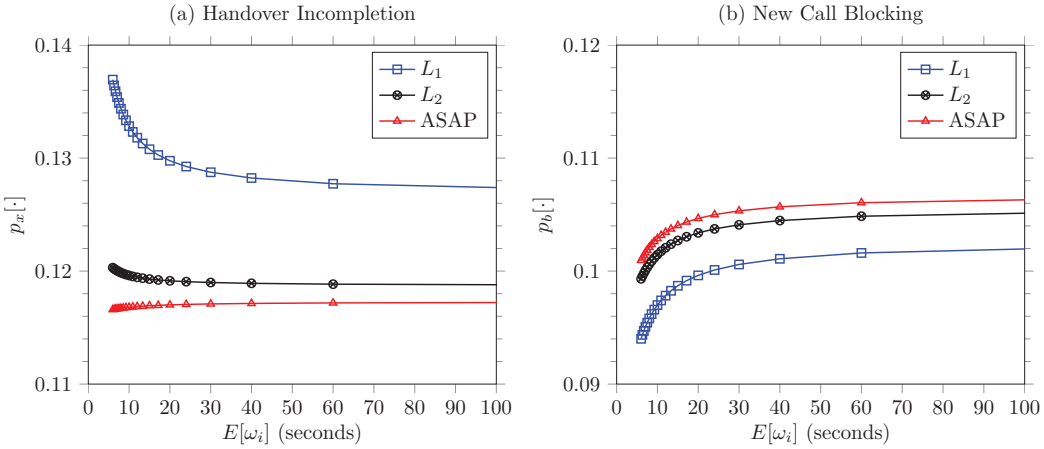



Figure 15: The effect of $E[\omega_i]$ on $p_x[\cdot]$ and $p_b[\cdot]$.

tially decreased if the handover mechanism can reflect the status of the MS SNR and MS velocity to trigger the handover procedures when MSs cross the overlap area. The performance of ASAP in terms of handover failure probability, handover delay, and handover rate under factors $E[\omega_i]$, $\lambda_n$, and $\lambda_h$ is demonstrated. By monitoring the MS SNR variation periodically and estimating the MS velocity to dynamically change the time period for ranging transmission, the SBS can determine an accurate handover time to avoid performing a handover at the incorrect point of time. Simulation results revealed that ASAP decreases the average handover time and efficiently decreases the handover failure probability of MSs in highly competitive circumstances. Based on the derived performance analysis, ASAP achieves a low handover time as low as 15 ms.

To achieve ASAP, an algorithm based on applying the IEEE 802.16m protocol [22] as an example for implementation is proposed. We emphasize that ASAP is a handover procedure that incorporates parts of MS SNR monitoring function, the MS moving velocity estimation function, and the trigger time of handover estimation function. ASAP can be applied to any type of protocol, such as the IEEE 802.16m or the E-UTRA [21] protocols. In other words, ASAP does not need to be modified and is ready for disposal over the existing network infrastructure. ASAP measures device locations without the GPS application which substantially reduce system cost. Finally, considering the mobility of MS, the ability of ASAP to support real-time QoS among femtocells or macrocells can be investigated further.

## References


[1] B. V. Quang, R. V. Prasad, and I. Niemegeers, "A Survey on Handoffs Lessons for 60 GHz Based Wireless Systems," *IEEE Commun. Surveys & Tutorials*, vol. 14, no. 1, pp. 64–86, 1st Quarter 2012.

[2] I. Al-Surmi, M. Othman, and B. M. Ali, "Mobility Management for IP-based Nextgeneration Mobile Networks: Review, Challenge, and Perspective," *J. Network Computer Applications*, vol. 35, no. 1, pp. 295–315, January 2012.

[3] V. Erceg, *et al.*, "A Model for the Multipath Delay Profile of Fixed Wireless Channels," *IEEE J. Select. Areas Commun.*, vol. 17, no. 3, pp. 399–410, March 1999.

[4] S. Tekinay and B. Jabbari, "A Measurement-Based Prioritization Scheme for Handovers in Mobile Cellular Networks," *IEEE J. Sel. Area Comm.*, vol. 10, pp. 1343–1350, October 1992.

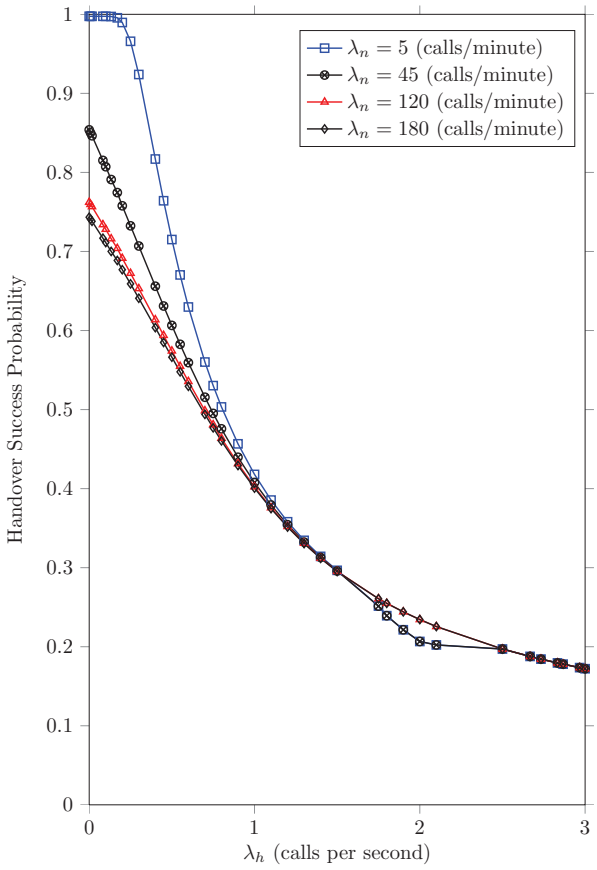



Figure 16: The effect of $\lambda_h$ on handover success probability.

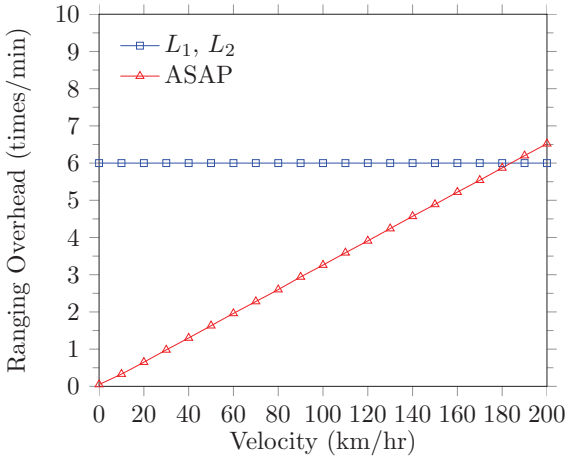



Figure 17: Ranging overhead versus MS moving velocity when $R = 1000$ and $E[d] = 511.74$ meters.


[5] A. E. Xhafa and O. K. Tonguz, "Dynamic Priority Queueing of Handover Calls in Wireless Networks: An Analytical Framework," *IEEE J. Sel. Area Comm.*, vol. 22, no. 45, pp. 904–916, June 2004.

[6] S Hong, M.H. Lee, H.-H. Chu, S.-H. Kwon, and J.L. Speyer, "Experimental Study on the Estimation of Lever Arm in GPS/INS," *IEEE Trans. Veh. Technol.*, vol. 55, no. 2, pp. 431–448, March 2006.

[7] S.-S. Hwang and J. J. Shynk, "Blind GPS Receiver with A Modified Despreader for Interference Suppression," *IEEE Trans. Aero. Elec. Sys.*, vol. 42, no. 2, pp. 503–513, April 2006.

[8] I. F. Akyildiz and W. Wang, "The Predictive User Mobility Profile Framework for Wireless Multimedia Networks," *IEEE/ACM Trans. Network.*, vol. 12, no. 6, pp. 1021–1035, December 2004.

[9] J. Chen and W.-K. Tan, "Predictive Dynamic Channel Allocation Scheme for Improving Power Saving and Mobility in BWA Networks," *Mobile Netw. Appl.*, vol. 12, no. 1, pp. 15-30, February 2007.

[10] P. N. Pathirana, A. V. Savkin, and S. Jha, "Location Estimation and Trajectory Prediction for Cellular Networks with Mobile Base Stations," *IEEE Trans. Veh. Technol.*, vol. 53, no. 6, pp. 1903–1913, November 2004.

[11] Y. L. Chen and S. L. Tsao, "A Low-Latency Scanning with Association Mechanism for Real-Time Communication in Mobile WiMAX," *IEEE Trans. Wirel. Commun.*, vol. 11, no. 10, pp. 3550–3560, October 2012.

[12] H. Tang, P. Hong, and K. Xue, "HeNB-Aided Virtual-Handover for Range Expansion in LTE Femtocell Networks," *IEEE J. Commun. Netw.*, vol. 15, no. 3, pp. 312–320, June 2013.

[13] N. W. Sung, N. T. Pham, T. Huynh, and W. J. Hwang, "Predictive Association Control for Frequent Handover Avoidance in Femtocell Networks," *IEEE Commun. Lett.*, vol. 17, no. 5, pp. 924–927, May 2013.

[14] T. Guo, A. Quddus, and R. Tafazolli, "Seamless Handover for LTE Macro-Femto Networks Based on Reactive Data Bicasting," *IEEE Commun. Lett.*, vol. 16, no. 11, pp. 1788–1791, November 2012.

[15] A. Rath and S. Panwar, "Fast Handover in Cellular Networks with Femtocells," in *Proc. IEEE ICC 2012*, pp. 2752–2757, Ottawa, Canada, June 2012.

[16] A. Roy, J. Shin, and N. Saxena, "Multi-Objective Handover in LTE Macro/Femto-Cell Networks," *J. Commun. Netw.*, vol. 14, no. 5, pp. 578–587, October 2012.

[17] Y. Yu and D. Gu, "The Cost Efficient Location Management in the LTE Picocell/Macrocell Network," *IEEE Commun. Lett.*, vol. 17, no. 5, pp. 904–907, May 2013.

[18] T. Guo, A. Quddus, N. Wang, and R. Tafazolli, "Local Mobility Management for Networked Femtocells Based on X2 Traffic Forwarding," *IEEE Trans. Veh. Technol.*, vol. 62, no. 1, pp. 326–340, January 2013.

[19] J. Astorga, M. Aguado, N. Toledo, and M. Higuero, "A High Performance Link Layer Mobility Management Strategy for Professional Private Broadband Networks," *J. Network Computer Applications*, vol. 36, no. 4, pp. 1152–1163, July 2013.

[20] Y. Shen, T. Luo, and M. Z. Win, "Neighboring Cell Search for LTE Systems," *IEEE Trans. Wirel. Commun.*, vol. 11, no. 3, pp. 908–919, March 2012.

[21] 3GPP, "Evolved Universal Terrestrial Radio Access (E-UTRA) and Evolved Universal Terrestrial Radio Access Network (E-UTRAN): Overall Description, Stage 2," ETSI TS 36.300 V11.4.0, Release 11, January 2013.

[22] IEEE 802.16 Working Group, "IEEE Standard for Local and Metropolitan Area Networks–Part 16: Air Interface for Broadband Wireless Access Systems, Amendment 3," IEEE Std. 802.16m–2011, May 2011.

[23] WiMAX Forum Working Group, "WiMAX Forum Mobile System Profile Specifications," Release 1–IMT-2000 Edt., WMF-T23-007-R010v02, Aug. 2009.

[24] S. Choi, G.-H. Hwang, T. Kwon, A.-R. Lim, and D.-H. Cho, "Fast Handover Scheme for Real-Time Downlink Services in IEEE 802.16e BWA System," in *Proc. IEEE VTC 2005-Spring*, vol. 3, pp. 2028–2032, Stockholm, Sweden, June 2005.

[25] N. Banerjee, K. Basu, and S. Das, "Handoff Delay Analysis and Measurement in SIP-Based Mobility Management in Wireless Networks," in *Proc. Int'l. Parallel and Distrib. Processing Symp. 2003*, pp. 224–231, Nice, France, April 2003.

[26] Y.-B. Lin and A.-C. Pang, "Comparing Soft and Hard Handoffs," *IEEE Trans. Veh. Technol.*, vol. 49, no. 3, pp. 792–798, May 2000.

[27] Y.-B. Lin, "Impact of PCS Handoff Response Time," *IEEE*

*Commun. Lett.*, vol. 1, no. 6, pp. 160–162, Novemeber 1997.

[28] D. Gross, J. F. Shortle, J. M. Thompson, and C. M. Harris, *Fundamentals of Queueing Theory*, 4th ed., Wiley, Hoboken, NJ, 2008.

[29] J. Chen and C.-C. Wang, "Cross-Layer Cut-Through Switching Mechanism for IEEE 802.16d/e Wireless Networks," *IEEE Commun. Lett.*, vol. 13, no. 10, pp. 779–781, October 2009.

[30] A. Kavak, M. Torlak, W. J. Vogel, and G. Xu, "Vector Channels for Smart Antennas-Measurements, Statistical Modeling, and Directional Properties in Outdoor Environments," *IEEE Trans. Microw. Theory*, vol. 48, no. 6, pp. 930–937, June 2000.

[31] B.-J. Kwak, N.-O. Song, and L. E. Miller, "Performance Analysis of Exponential Backoff," *IEEE/ACM Trans. Networking*, vol. 13, no. 2, pp. 343–355, April 2005.

[32] W. K. Lai and J. C. Chiu, "Improving Handoff Performance in Wireless Overlay Networks by Switching between Two-Layer IPv6 and One-Layer IPv6 Addressing," *IEEE J. Sel. Areas Comm.*, vol. 23, no. 11, pp. 2129–2137, November 2005.