

Signal Threshold Adaptation for Vertical Handoff in Heterogeneous Wireless Networks

Ahmed H. Zahran

Department of Electrical and Computer Engineering
University of Toronto
10 King's College Road
Toronto, Ontario, M5S 3G4, Canada
Email: zahran@comm.utoronto.ca
Tel: 1-416-978-4829, Fax: 1-416-978-4425

Ben Liang

Department of Electrical and Computer Engineering
University of Toronto
10 King's College Road
Toronto, Ontario, M5S 3G4, Canada
Email: liang@comm.utoronto.ca
Tel: 1-416-946-8614, Fax: 1-416-978-4425

Aladin Saleh

Bell Canada/ Wireless Technology
Floor 5N - 5099 Creekbank Road
Mississauga, Ontario, L4W 5N2, Canada
Email: aladdin.saleh@bell.ca
Tel: 1-905-282-3264, Fax: 1-905-282-3337

*1

¹To appear in *ACM/Springer Mobile Networks and Applications (MONET)* journal, this article is the extended version of a paper presented in IFIP Networking 2005.

Abstract

The convergence of heterogeneous wireless access technologies has been envisioned to characterize the next generation wireless networks. In such converged systems, the seamless and efficient handoff between different access technologies (vertical handoff) is essential and remains a challenging problem. The heterogeneous co-existence of access technologies with largely different characteristics results in handoff asymmetry that differs from the traditional intra-network handoff (horizontal handoff) problem. In the case where one network is preferred, the vertical handoff decision should be carefully executed, based on the wireless channel state, network layer characteristics, as well as application requirements. In this paper, we study the performance of vertical handoff using the integration of 3G cellular and wireless local area networks as an example. In particular, we investigate the effect of an application-based signal strength threshold on an adaptive preferred-network lifetime-based handoff strategy, in terms of the signalling load, available bandwidth, and packet delay for an inter-network roaming mobile. We present an analytical framework to evaluate the converged system performance, which is validated by computer simulation. We show how the proposed analytical model can be used to provide design guidelines for the optimization of vertical handoff in the next generation integrated wireless networks.

Keywords: heterogeneous wireless networks, seamless integration, vertical handoff, application signal strength threshold, 3G cellular, wireless LAN

1 Introduction

Wireless technologies are evolving toward broadband information access across multiple networking platforms, in order to provide ubiquitous availability of multimedia applications. Recent trends indicate that wide-area cellular networks based on the 3G standards and wireless Local Area Networks (WLANs) will co-exist to offer multimedia services to end users. These two wireless access technologies have characteristics that perfectly complement each other. By strategically combining these technologies, a converged system can provide both universal coverage and broadband access. Therefore, the integration of heterogeneous networks is expected to become a main focus in the development toward the next generation wireless networks [1–3].

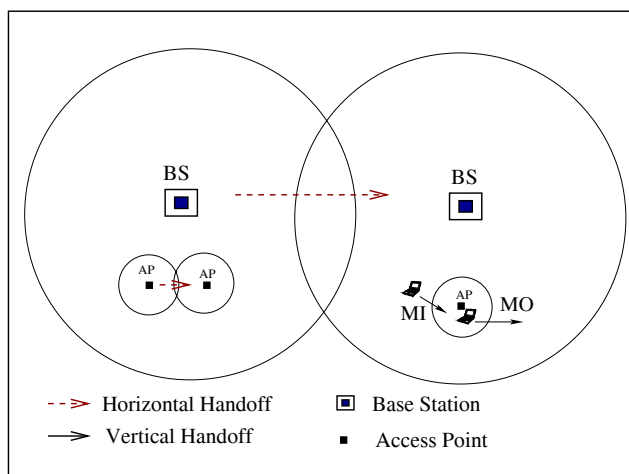


Figure 1: Mobile handoff in heterogeneous wireless system

Mobility management is a main challenge in the converged network. It addresses two main problems: location management and handoff management [4,5]. Location management tracks the Mobile Terminals (MT) for successful information delivery. For this purpose, Mobile IP (MIP) enables seamless roaming and is expected to be the main engine for location management in the next generation networks. Handoff management maintains the active connections for roaming mobile terminals as they change their point of attachment to the network. Handoff management is the main concern of this paper.

In the converged network, both intra-technology handoff and inter-technology handoff take place as illustrated in Figure 1. Intra-technology handoff is the traditional Horizontal Handoff (HHO) process in which the mobile terminal hands-off between two Access Points (AP) or Base Stations (BS) using the same access technology. On the other hand, inter-technology handoff, or Vertical Handoff (VHO), occurs when the MT roams between different access technologies. The main distinction between VHO and HHO is symmetry. While HHO is a symmetric process, VHO is an asymmetric process in which the MT moves between two different networks with different characteristics. This introduces the concept of a preferred network, which is usually the underlay WLAN that provides better throughput performance at lower cost, even if both networks are available and in good condition for the user.

There are two main scenarios in VHO: moving out of the preferred network (MO) and moving into the preferred network (MI) [6]. In the converged model,

it is highly desirable to associate the MT with the preferred network, as long as the preferred network satisfies the user application. This can improve the resource utilization of both access networks, as well as improving the user perceived QoS. Furthermore, this handoff should be seamless with minimum user intervention, while dynamically adapting to the wireless channel state, network layer characteristics, and application requirements.

In this work, we present an adaptive lifetime-based VHO (ALIVE-HO) algorithm which takes into consideration the wireless signal strength, handoff latency, and application QoS and delay tolerance. It can satisfy the system handoff signalling load, as well as different application requirements by the tuning of an application-based signal strength threshold (ASST). We further propose an analytical model to evaluate the performance of adaptive VHO. This analytical framework is then applied to show how the VHO decision and the ASST choice can be optimized based on multiple conflicting criteria including vertical handoff signaling, user available bandwidth, and encountered packet delay. Hence, the optimal ASST value is determined for different QoS requirements.

The rest of this paper is organized as follows. Section 2 provides an overview for the handoff algorithms in wireless heterogeneous networks and the related literature work. In Section 3, we present a vertical handoff algorithm that incorporates cross-layer adaptation to terminal mobility, channel state, and application demand. In Section 4, we propose an analytical framework to study the effect of cross-layer adaptation. Numerical and simulation results are provided in Section 5, where we show how the ASST can be tuned to optimize VHO decision. Concluding remarks are presented in Section 6.

2 Related Work

The traditional HHO problem has been studied extensively in the past. Several approaches have been considered in cellular networks using the Received Signal Strength (RSS) as an indicator for service availability from a certain point of attachment. Additionally, several handoff initiation strategies have been defined based on the comparison between the current attachment point RSS and that of the candidate attachment points as shown in [7]:

- RSS: handoff takes place if the candidate attachment point RSS is higher than the current attachment point RSS ($RSS_{new} > RSS_{cur}$).

- RSS plus threshold: handoff takes place if the candidate attachment point RSS is higher than the current attachment point RSS and the current attachment point RSS is less than a pre-defined threshold T ($RSS_{new} > RSS_{cur}$ and $RSS_{cur} < T$).
- RSS plus hysteresis: handoff takes place if the candidate attachment point RSS is higher than the current attachment point RSS with a pre-defined hysteresis margin H . ($RSS_{new} > RSS_{cur} + H$).
- A *dwel timer* can be added to any of the above algorithms. In this case, the timer is started when one of the above conditions is satisfied, and the MT performs a handoff if the condition is satisfied for the entire dwell timer interval.

In VHO, the RSSs are incomparable due to VHO's asymmetrical nature. However, they can be used to determine the availability as well as the condition of different networks. If the MI decision is based only on the preferred network availability, the MT should start the MI process as it discovers the WLAN. In addition, if more than one WLAN APs are available, the MT should associate itself with the one having the strongest RSS as it does in HHO². When the MT is associated with the preferred network, it enjoys all the preferred network advantages before moving out. Therefore, in the ideal MO scenario, the MT performs no more than one handoff at the WLAN edge when the network is expected to be unavailable. This ideal MO decision usually cannot be achieved. Thus, the main design requirements of a VHO algorithm are

- minimizing the number of unnecessary handoffs to avoid overloading the network with signaling traffic,
- maximizing the underlay network utilization,
- providing active application with the required degree of QoS,
- prioritizing handoff to the underlay network over MO to the overlay network,
- avoiding MI to a congested network, and

²If other criteria such as available bandwidth are considered, the MT may not move instantaneously to a WLAN, but may consider other factors such as QoS, user preference, cost, and power consumption.

- keeping fast users connected to the overlay network.

As far as we are aware, there exist very few works dealing with VHO beyond simple extensions to the common techniques for HHO. Three main directions for VHO algorithms are recorded in the literature.

The first approach is based on the traditional strategies of using the RSS that may be combined with other parameters such as network loading. In [8], Hatami et. al use the dwelling timer as a handoff initiation criterion to increase the WLAN utilization. They combine simulation and analysis to show that associating the MT with the WLAN for the longest possible duration improves user throughput even during the transition period in which the RSS oscillates around the receiver sensitivity level. However, they did not define a clear mechanism for choosing the dwelling timer value. In [9], Ylianttila et al. present an algorithm to compute an optimization policy for the dwelling timer according to the available data rates in both networks. The main result is that the optimal value for the dwelling timer is highly dependent on the difference between the available rates in both networks. In [10], Ylianttila et al. extend the same analytical framework of [8] to include multiple radio network environments. Their main results show that the handoff delay effect seems to be dominant even with the dwelling timer optimal choice as in [9]. In [11], Park et al. propose using a similar dwelling timer-typed approach for both MI and MO by performing the VHO if a specific number of the received beacons exceed or go below a predefined MI or MO threshold respectively. Additionally, the authors propose adapting the performance of their algorithm based on the application requirements by using two different numbers of beacons for real-time and non-real-time services. Although the dwelling timer approach seems to be an attractive approach for the VHO in order to maximize the underlay network usage, the proper dwelling timer choice is a critical decision because a large dwelling timer may result in undesirable service interruption periods for real-time applications. In our approach, interruptions are avoided by the proper choice of an application specific signal threshold to satisfy the requirements of the applications.

The second approach uses artificial intelligence techniques combining several parameters such as network conditions and MT mobility in the handoff decision. In [12], Ylianttila et al. present a general framework for the vertical handoff process based on fuzzy logic and neural networks. In [13], Pahlavan et al. present a neural network-based approach to detect signal decay and making handoff decision. In [14], Majlesi and Khalaj present a fuzzy logic based adaptive algorithm that varies the hysteresis margin and averaging window size based on MT velocity

and WLAN traffic. It is worth mentioning that some of these artificial intelligence based algorithms are complex and may be difficult to implement in practical systems. It is possible to extend our work to include improvement using similar artificial intelligence approaches. However, this is outside the scope of this paper and will be left for future work.

The third direction combines several metrics such as access cost, power consumption, and bandwidth in a cost function estimated for the available access networks, which is then used in the MT handoff decision. Wang et al. introduce the policy enabled handoff in [15], which was followed by several papers on similar approaches. In [15], the authors propose policies considering different parameters such as monetary cost, power consumption, network available bandwidth, and other parameters that differ among different heterogeneous networks. For each policy, a cost function is defined as a weighted sum of normalized policy parameters. These weights vary according to user preferences and the MT status (e.g., power reserve). In this scheme, the MT periodically compares the cost of different networks and then is handed off to the one with the minimum cost. Additionally, the authors introduce the programming model and software architecture of their solution. In [16], Zhu and McNair present cost functions that account for the dynamic values that are inherent to vertical handoff and incorporate a network elimination factor to potentially reduce delay and processing power in the handoff calculation. They introduce two cost-based policies for VHO decision considering the available bandwidth and RSS of the available networks. The collective handoff policy estimates one cost for all the services, while the prioritized multi-network handoff policy estimates the cost for each service independently. Also, Chen et al. [17] introduce a smart decision model using a handoff control center module in the MT. This module monitors the available interfaces and the system resources to collect information required for the handoff decision. This decision is based on a score function that considers the usage expenses, link capacity, and power consumption for the available access technologies. The MT uses the network that achieves the largest score. One main difficulty of the cost approach is its dependence on some parameters that are difficult to estimate, especially in large cellular networks, such as the available bandwidth, the channel condition, and the network user density, all of which change dynamically.

In ALIVE-HO, we adopt the first approach of using the RSS as a unique input for the algorithm, to estimate the duration through which the WLAN usage will be beneficial for the active applications.

3 Application Life-Time Adaptation

3.1 System Model

We study the overlapping of 3G cellular and WLAN networks. The cellular network is assumed to provide universal coverage, while WLAN availability is indicated by the presence of the WLAN beacons [13] that are periodically transmitted by the WLAN APs. Mobile-IP is assumed for mobility management.

The European Telecommunication Standards Institute (ETSI) proposed two mobility management architectures for the next generation wireless networks: tightly coupled and loosely coupled [18]. In tight coupling, a WLAN gateway emulates the functions of a cellular Radio Network (RN), while in loose coupling the WLAN gateway helps authenticate the users, obtains their service profile at the session beginning only, and then uses its own resources to route the subscriber data. The latter approach is preferred for several reasons, including flexibility that enables the integration of the third party wireless Internet service providers and independent implementation of WLAN and 3G networks.

We assume that WLAN hotspots implement loosely coupled connection with the 3G network using WLAN gateways. These gateways perform several tasks including serving as Mobile-IP agents and possibly providing QoS in the form of multiple service classes defined within the WLAN. However, it is worth mentioning that end-to-end QoS support requires other mechanisms such as differentiated services to be implemented over the entire network path. The details of such implementation is unimportant to the proposed VHO algorithm and mathematical analysis. We are mainly concerned about the resultant VHO delay values.

The MT is equipped with dual interfaces that allow it to communicate with both networks. However, since Mobile-IP provides only one IP tunnel, the MT can connect to only one network at a time. In addition, multi-interface mobility client software is installed on the MT. This software performs Mobile-IP signaling with the foreign and home agents. It periodically scans the available interfaces and measures the observed RSS. Then it intelligently selects the best access network according to the predefined VHO algorithm.

Within the WLAN, a log-linear path loss channel propagation model with shadow fading is used [19]. The RSS is expressed in dBm as

$$RSS = P_T - L - 10n \log(d) + f(\mu, \sigma), \quad (1)$$

where P_T is the transmitted power, L is a constant power loss, n is the path loss exponent and usually has values between 2 - 4, d represents the distance between

the MT and the WLAN AP, and $f(\mu, \sigma)$ represents shadow fading which is modelled as Gaussian with mean $\mu = 0$ and standard deviation σ with values between 6-12 dB depending on the environment. We assume that when the RSS is below a certain interface sensitivity level, α , the MT is unable to communicate with the AP.

3.2 Adaptive Preferred-Network Life-Time Vertical Handoff (ALIVE-HO)

For the MO scenario when the MT is within a WLAN, we use the RSS to estimate the expected duration after which the MT is unable to maintain its connection with the WLAN. We take into consideration the handoff delay due to MIP tunnelling, authentication, and service initiation. We further consider an Application Signal Strength Threshold (ASST), which is the required level of RSS for the active application to perform satisfactorily.

The ASST is an application dependent parameter which represents a composite of the channel bit error rate, application error resilience, and application QoS requirements. We present here how the ASST can be incorporated into the VHO decision. We further discuss in the next section how the ASST can be adjusted to optimize the overall system performance.

In discrete time, the RSS is expressed as

$$RSS[k] = \mu_{RSS}[k] + N[k], \quad (2)$$

where k is the time index, $\mu_{RSS}[k] = P_T - L - 10n \log(d[k])$, and $N[k] = f(\mu, \sigma)$. The averaged RSS, $\overline{RSS}[k]$, can be estimated using a moving average

$$\overline{RSS}[k] = \frac{1}{W_{av}} \sum_{i=0}^{W_{av}-1} RSS[k-i]. \quad (3)$$

The RSS rate of change, $S[k]$, can be obtained by

$$S[k] = \frac{M_1[k] - M_2[k]}{W_S T_S}, \quad (4)$$

where

$$M_1[k] = \frac{2}{W_S} \sum_{i=0}^{\frac{W_S}{2}-1} \overline{RSS}[k - W_S + 1 + i], \quad (5)$$

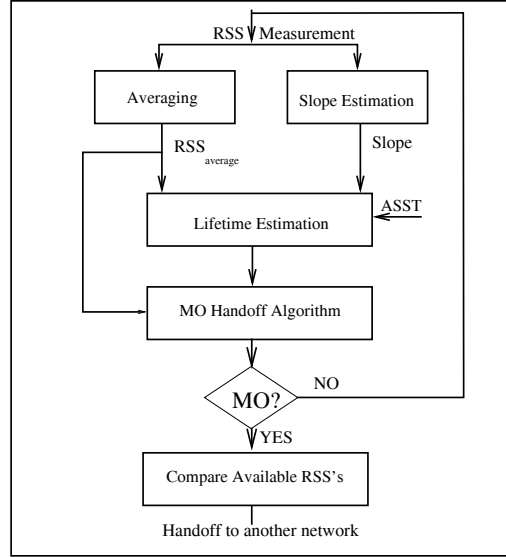


Figure 2: MO handoff algorithm

$$M_2[k] = \frac{2}{W_S} \sum_{i=\frac{W_S}{2}}^{W_S-1} \overline{RSS}[k - W_S + 1 + i], \quad (6)$$

and W_S and T_S denote the slope estimator window size and the RSS sampling interval respectively.

Then, we estimate the MT lifetime within the WLAN, $EL[k]$, as follows.

$$EL[k] = \frac{\overline{RSS}[k] - \gamma}{S[k]}, \quad (7)$$

where γ denotes the ASST. Thus, $EL[k]$ represents the application specific time period in which the WLAN is likely to remain usable to the MT. Figure 2 depicts the MO scenario block diagram.

Once the VHO decision is taken, the available cellular RSS from different base stations are compared to determine the base station with which the MT will associate itself.

Based on the measured and estimated parameters, the MT will initiate the MO handoff at time k if the averaged received signal strength is less or equal to a predefined MO threshold, MOT_{WLAN} , and the estimated lifetime is less than or equal to the handoff delay threshold, T_{HO} . The first condition prevents unnecessary handoffs near the access point resulting from short lifetime estimate due

to fast signal decay; additionally, the lifetime part tunes the handoff instant according to users mobility to benefit from WLAN resources. The MOT_{WLAN} is usually chosen to be a few dB above the wireless interface sensitivity. T_{HO} can be set to the expected handoff delay between the two access technologies. This delay includes several signaling delay components such as discovery delay, authentication delay, and registration delay. These delays vary depending on the adopted approach for location management, whether it is Mobile-IP [2] or an end-to-end approach [20, 21].

Clearly, the window sizes have significant effect on the lifetime-based algorithm performance. In general, a larger window size results in better estimation but also larger delay in handoff performance [7]. Using variable window sizes that adapt to the MT mobility can improve handoff performance. For example, W_{av} and W_S can be determined as follows.

$$W_{av} = \max \left(10, \left\lfloor \frac{D_{av}}{VT_S} \right\rfloor \right), \quad (8)$$

$$W_S = 2 * \max \left(50, \left\lfloor \frac{D_s}{VT_S} \right\rfloor \right), \quad (9)$$

where D_{av} and D_s represent the averaging and slope distance windows respectively, $\lfloor \cdot \rfloor$ represents the greatest lower integer function, and V is the MT velocity away from the AP, which can be obtained by many velocity estimators proposed in the literature, for example [22]. Hence, better estimates due to larger windows are obtained for slower users; which safely improves the handoff performance and enables maximizing the benefits of WLANs.

In the MI scenario, several factors need to be considered. The main one is the WLAN availability, which can be determined by the WLAN RSS. In addition, the QoS, specified in terms of the available bandwidth, is a key factor in the handoff decision. Other factors such as security, user preference can be considered. In this work, we consider a simplified model where the MT performs MI to the WLAN if $\overline{RSS}[k] > MIT_{WLAN}$ and the available bandwidth is greater than the required bandwidth. The available bandwidth can be estimated based on observing the Network allocation vector [23] or can be incorporated within the AP beacon or MIP foreign agent advertisement to decrease the delay between the WLAN discovery and the MI initiation. In our simulation and analysis, we assume that the WLAN is always in good condition, so that the MT always perform an MI after an unnecessary MO.

4 Performance Analysis

In this section, we provide an analytical framework for evaluating the performance of the cross-layer ALIVE-HO algorithm.

4.1 Transition probabilities

The calculation of the transition probabilities is based on recursive computation of the handoff probabilities similar to [24]. In the integrated heterogeneous networking model, the following probabilities are required for handoff algorithm analysis.

- $P_W[k]$: Pr{MT is associated with the WLAN at instant k}
- $P_C[k]$: Pr{MT is associated with the 3G network at instant k}
- $P_{W|C}[k]$: Pr{MT associates itself with the WLAN at instant k given that it is associated with the cellular network at instant k-1}
- $P_{C|W}[k]$: Pr{MT associates itself with the 3G network at instant k given that it is associated with the WLAN at instant k-1}

In our model, the MT is assumed to be attached to the WLAN at the beginning; hence $P_W[0] = 1$ and $P_C[0] = 0$. $P_W[k]$ and $P_C[k]$ can be calculated recursively as follows.

$$P_W[k + 1] = P_{W|C}[k + 1]P_C[k] + (1 - P_{C|W}[k + 1])P_W[k], \quad (10)$$

$$P_C[k + 1] = P_{C|W}[k + 1]P_W[k] + (1 - P_{W|C}[k + 1])P_C[k]. \quad (11)$$

The Conditional probabilities $P_{C|W}[k + 1]$ and $P_{W|C}[k + 1]$ depend on the handoff algorithm initiation strategy. For the proposed cross-layer ALIVE-HO algorithm, $P_{C|W}[k + 1]$ is determined by

$$P_{C|W}[k + 1] = Pr \left\{ \overline{RSS}[k + 1] < MOT_{WLAN}, EL[k + 1] < T_{HO|W}[k] \right\} \quad (12)$$

where $W[k]$ represents the event that the MT is associated with the WLAN at time k . In practice, WLANs are designed for low mobility users. The lifetime part of the MO condition becomes more significant for low mobility users. Hence the

MO condition can be reduced to $EL[k] < T_{HO}$. Consequently, one can determine $P_{C|W}[k+1]$ as follows:

$$\begin{aligned} P_{C|W}[k+1] &= Pr\{EL[k+1] < T_{HO} | EL[k] > T_{HO}\} \\ &= Pr\{\overline{RSS}[k+1] - T_{HO}S[k+1] < \gamma | \overline{RSS}[k] - T_{HO}S[k] > \gamma\} \end{aligned} \quad (13)$$

Let $Z[k] = \overline{RSS}[k] - T_{HO}S[k]$. Then we have

$$\begin{aligned} P_{C|W}[k+1] &= Pr\{Z[k+1] < \gamma | Z[k] > \gamma\} \\ &= \frac{Pr\{Z[k+1] < \gamma, Z[k] > \gamma\}}{Pr\{Z[k] > \gamma\}}. \end{aligned}$$

Clearly, since $RSS[k]$ is a Gaussian process, the processes $\overline{RSS}[k]$ and $S[k]$ are Gaussian, and hence $Z[k]$ is Gaussian too. Let its mean be $\mu_Z[k]$ and standard deviation be $\sigma_Z[k]$. It can be shown that

$$\mu_Z[k] = \mu_{\overline{RSS}}[k] - T_{HO} \mu_S[k] \quad (15)$$

where

$$\mu_{\overline{RSS}}[k] = \mu_{RSS}[k] + \frac{1}{W_{av}} \sum_{i=0}^{W_{av}-1} 10n \log\left(1 - \frac{iVT_S}{d[k]}\right),$$

and

$$\mu_S[k] = \frac{E\{M_1[k]\} - E\{M_2[k]\}}{W_S T_S},$$

and furthermore

$$\sigma_z^2[k] = \sigma_{\overline{RSS}}^2[k] + T_{HO}^2 \sigma_S^2[k] + \frac{4T_{HO} \sigma_{RSS}^2 \sum_{h=0}^{W_{av}-1} (W_{av} - |h|)}{W_S^2 T_S W_{av}^2},$$

where

$$\sigma_{\overline{RSS}}^2[k] = \frac{\sigma^2}{W_{av}},$$

and

$$\sigma_S^2[k] = \frac{4\sigma^2}{(T_S W_S^2 W_{av})^2} \left[W_{av} W_S + \sum_{h=1}^{W_{av}-1} (W_{av} - |h|)(2W_S - 6|h|) \right].$$

Additionally, $Z[k]$ and $Z[k-1]$ are jointly Gaussian with correlation coefficient $\rho_{Z[k], Z[k-1]}$ as derived in the Appendix, which defines their joint PDF $f_{Z[k], Z[k-1]}(z_1, z_2)$ [25].

Then we can compute $P_{C|W}[k+1]$ by

$$P_{C|W}[k+1] = \frac{\int_{-\infty}^{\gamma} \int_{\gamma}^{\infty} f_{Z[k+1]|Z[k]}(z_1, z_2) dz_1 dz_2}{Q\left(\frac{\gamma - \mu_{Z[k]}[k]}{\sigma_{Z[k]}[k]}\right)},$$

where $Q(x)$ is the complementary error function. Similarly, $P_{W|C}[k+1]$ can be determined by

$$P_{W|C}[k+1] = Pr\{\overline{RSS}[k+1] > MIT | \overline{RSS}[k] < MIT\} \quad (16)$$

$$= \frac{Pr\{\overline{RSS}[k+1] > MIT, \overline{RSS}[k] < MIT\}}{Pr\{\overline{RSS}[k] < MIT\}}. \quad (17)$$

where, similar to the $(Z[k+1], Z[k])$ tuple, the $(\overline{RSS}[k+1], \overline{RSS}[k])$ tuple is jointly Gaussian. These transition probabilities are used to calculate the performance metrics in the next subsections.

4.2 Handoff Probabilities and the Number of Handoffs

The number of handoffs has major impact on the signaling traffic, which may overload the network resulting in degradation in the overall performance. The number of handoffs, denoted N_{HO} , is defined as the sum of MOs and MIs between WLAN and 3G network as the MT roams across the network boundary. Hence, it is a random variable that depends on the instantaneous move out/in probabilities, which can be calculated by

$$P_{MO}[k+1] = P_{C|W}[k+1]P_W[k], \quad (18)$$

$$P_{MI}[k+1] = P_{W|C}[k+1]P_C[k]. \quad (19)$$

The MT movement between the two networks can be modeled by a two-state non-homogeneous Markov chain. Each state represents the network with which the MT is associated. The transition probabilities are $P_{MO}[k]$ and $P_{MI}[k]$ as shown in Figure 3. Hence, by using binary impulse rewards for the handoff transition as shown in [26], we calculate the average accumulated rewards for MO and MI transitions, which are equivalent to the expected number of MOs, N_{MO} , and the expected number MIs, N_{MI} , respectively. Hence, the expected number of handoffs is

$$E\{N_{HO}\} = E\{N_{MO}\} + E\{N_{MI}\} \quad (20)$$

$$= \sum_{k=1}^{k_{max}} (P_{MO}[k] + P_{MI}[k]). \quad (21)$$

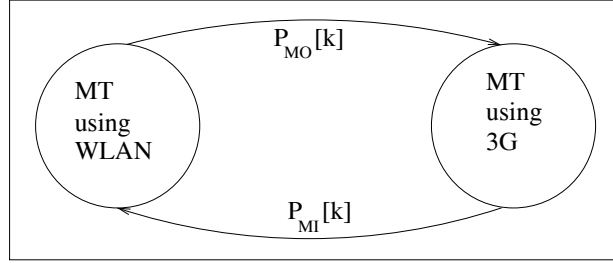


Figure 3: VHO Markov chain model

4.3 Available Bandwidth

The available bandwidth to the MT depends on the proportion of time that the MT stays in the WLAN and the 3G network, as well as the WLAN state when the MT is connected to the WLAN. To the MT, the WLAN is in one of two states: WLAN Up and WLAN Down. The WLAN Up state represents the event that the WLAN signal received at the MT is above the sensitivity level α . WLAN Down is the reverse case. Let $p[k]$ be the probability that the WLAN is in the Up state at time k . Clearly

$$\begin{aligned} p[k] &= Pr\{RSS[k] > \alpha\} \\ &= Q\left(\frac{\alpha - \mu[k]}{\sigma}\right). \end{aligned}$$

In the adopted handoff algorithm, the MO distance varies; consequently the captured WLAN Up durations does too. For the rest of the analysis, we are interested in evaluating the system performance during the *transition region*, which is defined as the range of distance between the point when the RSS starts to oscillate around the interface sensitivity and the WLAN edge. The transition region determination is equivalent to a long-standing complex level crossing problem that is analytically tractable only for a few simple cases and is usually solved numerically for complex cases. Here, we obtained the transition region starting point, denoted k_{start} , from rough estimates based on simulation results.

Then, the WLAN efficiency, ζ_{LT} , defined as the percentage of the WLAN up duration over the MT lifetime in the WLAN, can be estimated as

$$\zeta_{LT} = \sum_{k=k_{start}}^{k_{max}} \overline{P_{MO}}[k] \frac{\sum_{h=1}^k p[h]}{k},$$

where $\overline{P_{MO}}$ is a scaled version of P_{MO} to represent a valid PDF within interval $[1, k_{max}]$, and k_{max} represents the time index at which the MT reaches the WLAN edge and is determined by the planned coverage area.

Hence, the MT available bandwidth, BW_{Av} , assuming R_W and R_C as the effective data rates in WLAN and cellular networks respectively, can be computed as

$$BW_{Av} = \frac{\zeta_{LT} R_W (\overline{k_{MO}} - k_{start}) + R_C (k_{max} - \overline{k_{MO}})}{(k_{max} - k_{start})}, \quad (22)$$

where $\overline{k_{MO}}$ denotes the average time to MO.

4.4 Packet Delay

In addition to the MT available bandwidth, RSS degradation in the transition region impacts on the head of line (HoL) packet delay probability. To study this, we assume a threshold, θ_D for packet delay in the current hop as a part of the end-to-end delay budget for the real-time application packet from the source to the destination. A packet is considered *excessively delayed*³ if its HoL delay exceeds θ_D . Consequently, the average packet delay probability, D , can be estimated as

$$D = \frac{\sum_{k=k_{start}}^{\overline{k_{MO}}} P_D[k]}{(\overline{k_{MO}} - k_{start} + 1)},$$

where $P_D[k]$ represents the probability that a packet will be excessively delayed, which is equal to the probability of WLAN Down runs whose duration is equal to the delay threshold. Here we have performed an approximation by using $\overline{k_{MO}}$, instead of using k_{MO} and then applying conditional expectation. As shown in the next section, this approximation produces accurate results over a wide range of system parameters.

5 Numerical Results and the Optimization of ASST

5.1 Simulation Model

In addition to the above analysis, we have simulated the VHO algorithms using MATLAB. Table 1 shows the simulation parameter values. The WLAN parameters are used as in [14], which are suitable to model outdoor suburban (e.g. with

³Note that this does not necessarily mean that the packet is lost.

Table 1: *Simulation parameter values*

Parameter	Value	Parameter	Value
P_T	100 mWatt	T_S	0.01 sec
n	3.3	MOT_{WLAN}	-85 dBm
σ	7 dB	MIT_{WLAN}	-80 dBm
S	28.7 dB	$T_{handoff}$	1 sec
D_{av}	0.5 m	R_W	6 Mbps
D_s	5 m	R_C	0.6 Mbps
α	-90 dBm		

tree and low buildings along the road side which is similar to the characteristics of the commercial WLAN services) and indoor locations with wide areas (such as hotel lobbies and campuses). These parameters result in a WLAN coverage of 100 meters approximately.

The data rates shown in the table are used for performance evaluation only and have no effect on the handoff decision. Currently, IEEE802.11b WLANs are widely deployed and support rates vary from 11 Mbps to 1 Mbps depending on the distance between the MT and the WLAN AP. On the other hand, cellular service providers are still deploying their first phase of the 3G network that supports rates up to 144 Kbps for CDMA1x⁴. In the future, IEEE802.11a [27] and CDMA20001x-EV [28] are expected to be widely deployed. The former support rates that vary from 54 Mbps to 6 Mbps, while the latter supports a peak rate of 2.4 Mbps on the forward link with an average throughput of 600 kbps. Hence, these values show that the service rate of WLANs is generally approximately one order of magnitude larger than that of the cellular network.

Additionally, a simple mobility model is assumed in which a MT moving away from the WLAN access point in a straight line at a constant speed V . As shown in [29], this model is suitable for evaluating the performance of signal strength based algorithms with log-normally distributed shadow fading environments as in our case. Additionally, the proposed algorithm will function with any mobility pattern since the algorithm dynamically adapts to the MT velocity, and the algorithm time resolution is sufficient to track mobility pattern variation, especially for low speed MTs.

⁴Effective data rates are much lower than this value.

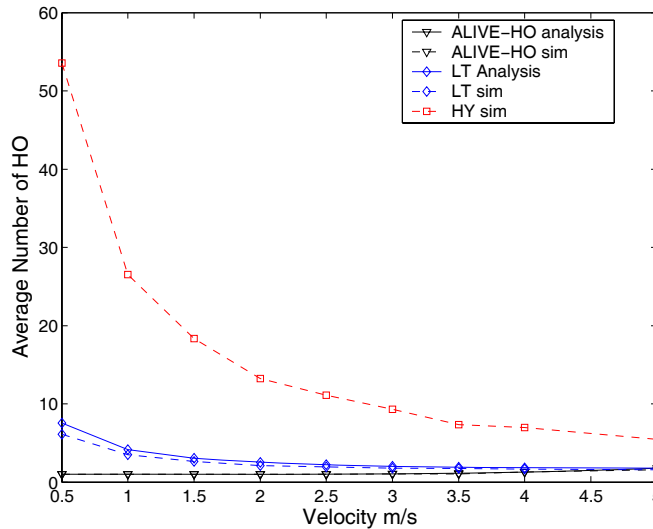


Figure 4: Number of handoffs ($\gamma = -90$ dBm).

5.2 Performance Comparison

We compare the performance of ALIVE-HO with traditional hysteresis VHO, which is used in [2]. In hysteresis based algorithms, there are two different thresholds MIT_{WLAN} and MOT_{WLAN} for the MI and MO respectively. The MT performs a MI if the $\overline{RSS}[k]$ is larger than MIT_{WLAN} and performs a MO if $\overline{RSS}[k]$ is smaller than the predefined MOT_{WLAN} . Usually, MIT_{WLAN} is chosen larger than MOT_{WLAN} to decrease the number of unnecessary handoffs known as ping-pong effect. We also consider a non-adaptive WLAN lifetime based VHO algorithm, where the lifetime estimation does not adapt to MT mobility or application demand, and hence a fixed RSS averaging window of ten samples is used.

Figures 4, 5, and 6 illustrate the number of handoffs, available bandwidth, and the packet delay probability for the VHO handoff algorithms. In all figures, HY denotes hysteresis VHO, LT denotes non-adaptive lifetime VHO, and ALIVE-HO denotes the adaptive lifetime VHO. All figures show good match between analysis and simulation.

Figure 4 shows that the introduction of the adaptive lifetime approach to the traditional HY VHO algorithm results in significant decrease of the number of unnecessary handoffs. Figure 5 demonstrates the improvement on the available bandwidth by using adaptive lifetime estimation. Clearly, from a pure bandwidth point of view, it is preferable for the MT to perform MO handoff only once at the

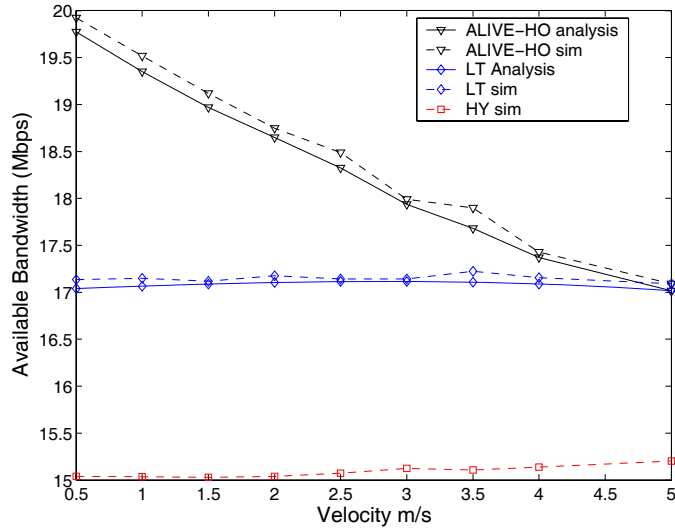


Figure 5: Available bandwidth ($\gamma = -90$ dBm).

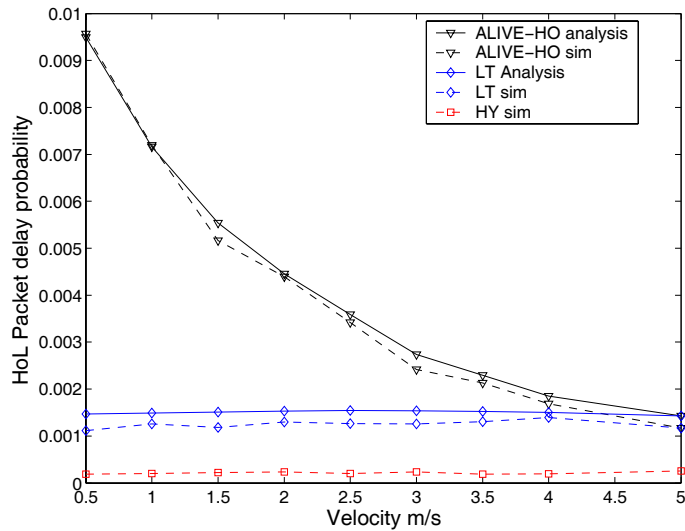


Figure 6: HoL packet delay rate ($\gamma = -90$ dBm, $\theta_D = 30$ ms).

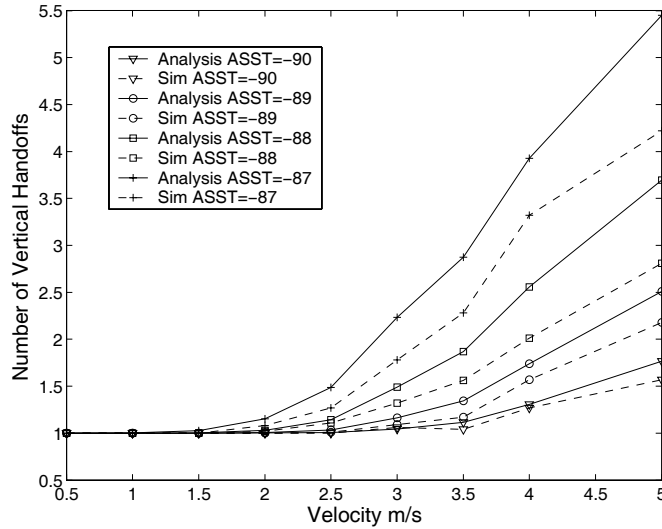


Figure 7: Number of handoffs vs. ASST.

WLAN edge, even though the RSS can temporarily go below the MT sensitivity level in the transition region. However, a drawback of increasing the lifetime of the MT within the WLAN is increasing the packet delay resulting from channel condition degradation. As shown in Figure 6, the packet delay probability using the adaptive approach can be much higher than that when the traditional hysteresis algorithm is used. This may be critical if the MT is running real-time application. However, by properly tuning the ASST as shown in the next subsection, ALIVE-HO can adapt to the active real-time application requirements in the MT.

5.3 Application Signal Strength Threshold Adaptation

Figures 7, 8, and 9 illustrate the effect of the ASST on the number of handoffs, available bandwidth, and packet delay probability. They show that the number of handoffs decreases when the ASST is reduced, since reducing the ASST allows the MT to remain in the WLAN for a longer duration. For the same reason, the available bandwidth to the MT increases when the ASST is reduced. However, at the same time, the packet delay probability is increased, since signal outage is more severe near the WLAN edge. Hence, there is a clear trade off among the handoff signaling load, available bandwidth, and packet delay.

Clearly, the ASST should not depend on the application QoS alone. Rather, it

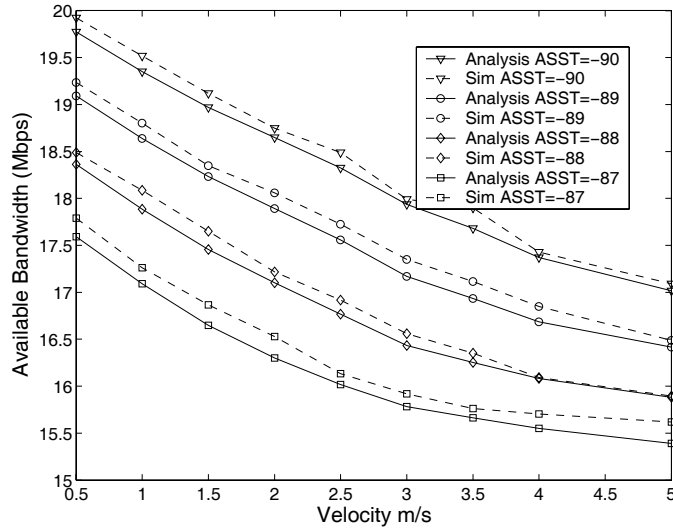


Figure 8: Available bandwidth vs. ASST.

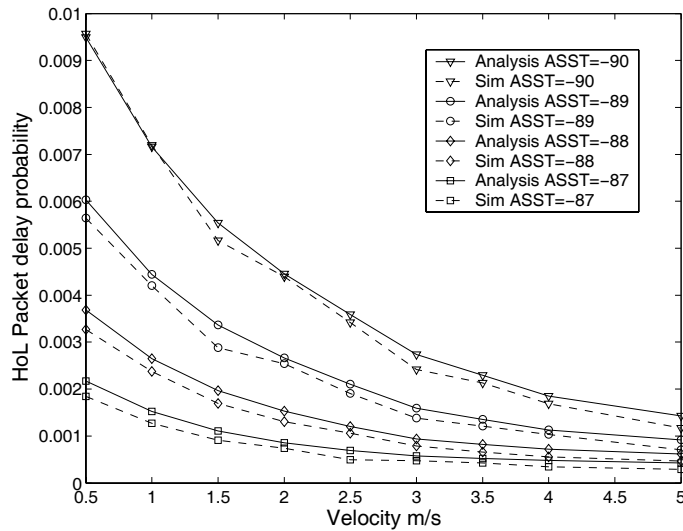


Figure 9: HoL packet delay rate vs. ASST ($\theta_D = 30\text{ms}$).

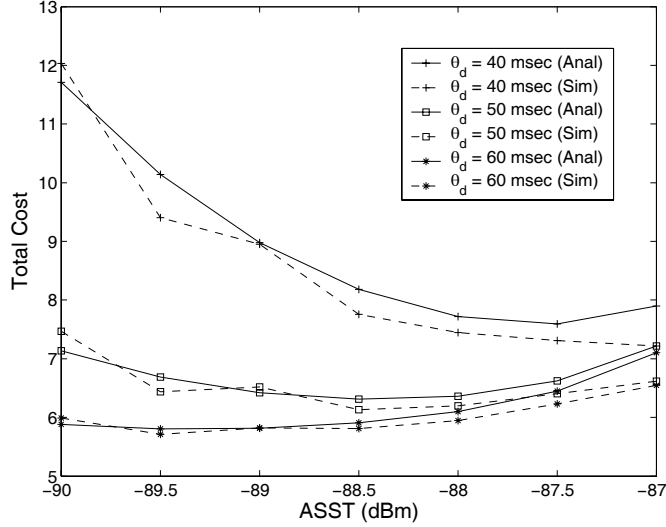


Figure 10: Total cost vs. ASST.

can be optimally tuned based on the various conflicting criteria of VHO. Likewise, the optimal VHO decision can be made adaptive to the RSS variation, network delay characteristics, and application QoS demands, through a properly chosen ASST value. The proposed analytical framework provides a means to carry out this optimization.

As an example, a possible cost function to aggregate the multiple VHO criteria may be

$$C_{total} = \frac{c_H E\{N_{HO}\} + c_D D}{BW_{av}},$$

where c_H represents the signaling cost per handoff, c_D represents the penalty factor for packet delay, and C_{total} is normalized to cost per Mbps of data bandwidth⁵.

Figure 10 plots C_{total} over different ASST values, for $V = 2$, $c_H = 100$, and $c_D = 10000$, where each curve represents a delay threshold value of 40 ms, 50 ms, and 60 ms, respectively. Clearly, the optimal ASST increases as the delay threshold decreases. In particular, when $\theta_D = 40ms$, an ASST of -87.5 dBm strikes the optimal balance to minimize the total cost, but when $\theta_D = 60ms$, the optimal ASST is -89.5 dBm.

⁵We emphasize here that this is only one of many possible cost functions, whose suitability depends on practical application goals and system constraints.

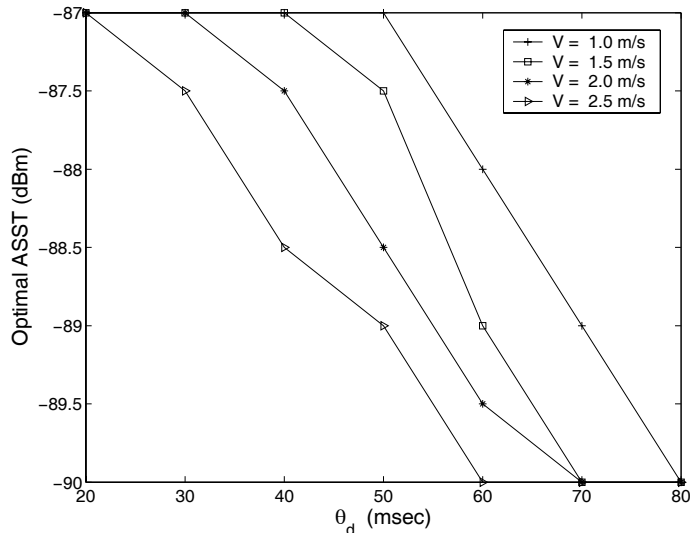


Figure 11: Optimal ASST vs. delay budget threshold and velocity ($c_H = 100$, $c_D = 100000$).

5.4 Optimal Application Signal Strength Threshold Values

To further study how the optimal ASST is affected by the system parameters, in Figures 11 and 12, we present numerical analysis results obtained for the optimal ASST values given various system parameters.

Figure 11 shows the optimal ASST over different delay budgets and MT velocities, where $c_H = 100$ and $c_D = 100000$. It is clear that as the delay constraint is relaxed, the optimal ASST value decreases approximately linearly in dB (exponentially in linear scale), and consequently, the MT WLAN lifetime increases. Additionally, as the MT velocity increases, the optimal ASST decreases. For example, if two MTs, MTa and MTb moving at 1.5 m/s and 2 m/s respectively, are running a real-time application with a 40 ms delay budget in the WLAN, MTa should set its ASST to -87 dBm while MTb should set it to -87.5 dBm. Clearly, as the MT velocity increases, the signal decay rate will increase. Hence, the decrease in the optimal ASST for the faster MT compensates this to make both MTs handoff at a similar distance, in order to satisfy the required delay constraint.

Figure 12 plots the optimal ASST over different handoff signalling costs and packet delay penalties, where $V = 2$ and $\theta_D = 50$. Clearly, as the signalling cost increases, the optimal ASST decreases sub-linearly in dB. With high handoff cost, the MT is pushed to perform handoff nearer the WLAN edge, and hence reducing

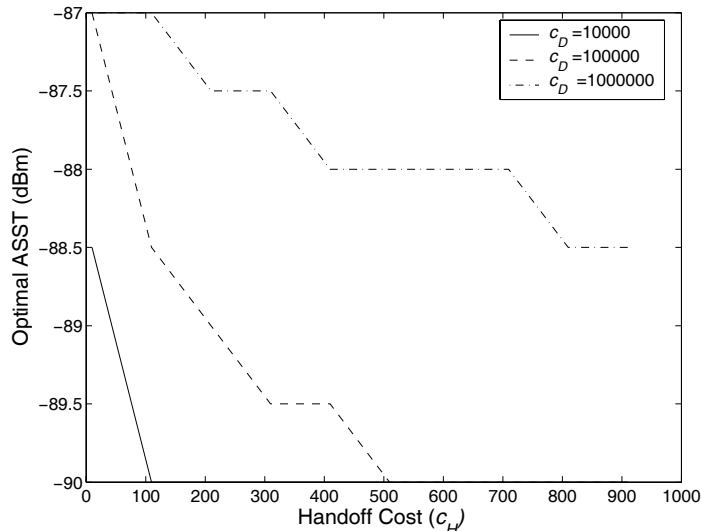


Figure 12: Optimal ASST vs. handoff cost and delay penalty ($V = 2, \theta_D = 50$).

the number of unnecessary handoffs. The same figure also shows that as the packet delay penalty increases, the optimal ASST increases. Hence, the MT is allowed to handoff earlier to avoid the deteriorating channel condition as it approaches the WLAN edge.

Thus, the propose numerical analysis can provide general guidelines for the optimal operation of lifetime-based VHO, adapting to various system conditions through the ASST value. To implement this in practice, a lookup table for the optimal ASST can be built based on the above analysis results.

6 Conclusions

In converged wireless systems, efficient vertical handoff management between heterogeneous networks is critical to the overall system performance. We have presented an application-specific signal strength tuning mechanism to a cross-layer adaptive VHO approach, which takes into account the wireless channel variation, network layer latency, and application QoS demands. We have proposed an analytical framework to evaluate the performance of VHO based on multiple criteria. The adaptive VHO approach has been shown to improve the system resource utilization by increasing the reliance of the MT on the WLAN, as well as conserving the resources of the 3G network for users located outside the WLAN.

More importantly, the proposed application signal threshold adaptation provides a means for flexible system design. Given a predefined priority policy, it can be used to optimize the tradeoff between handoff signalling, available bandwidth, and packet delay. Since the ASST can be optimally tuned for any access network based on practical system characteristics and requirements, it may have a significant role in future generation wireless networks where access technologies with vastly differing characteristics are expected to seamlessly co-exist and efficiently inter-operate.

Appendix: $Z[k]$ Statistics

Variance

$$\begin{aligned}\sigma_z^2[k] &= E\{Z^2[k]\} - E^2\{Z[k]\} \\ &= E\{RSS^2[k]\} + T_{HO}^2 E\{S^2[k]\} - 2T_{HO} E\{\overline{RSS}[k]S[k]\} - E^2\{Z[k]\}\end{aligned}\quad (23)$$

Since

$$\begin{aligned}E\{\overline{RSS}[k]S[k]\} &= E\{\overline{RSS}[k] \frac{2 \sum_{i=0}^{\frac{W_S}{2}-1} \overline{RSS}[k - W_S + 1 + i]}{W_S^2 T_s}\} - \\ &E\{\overline{RSS}[k] \frac{2 \sum_{i=\frac{W_S}{2}}^{W_S-1} \overline{RSS}[k - W_S + 1 + i]}{W_S^2 T_s}\} \quad (25)\end{aligned}$$

$$\begin{aligned}&= \frac{2 \sum_{i=0}^{\frac{W_S}{2}-1} \mu_{\overline{RSS}}[k] \mu_{\overline{RSS}}[k - W_S + 1 + i]}{W_S^2 T_s} - \\ &\frac{2 \sum_{i=\frac{W_S}{2}}^{W_S-1} \mu_{\overline{RSS}}[k] \mu_{\overline{RSS}}[k - W_S + 1 + i]}{W_S^2 T_s} - \\ &\frac{\sigma^2 \sum_{i=0}^{W_{av}-1} (W_{av} - |h|)}{\frac{W_S}{2} W_S T_s W_{av}^2} \quad (26)\end{aligned}$$

$$= \mu_{\overline{RSS}}[k] \mu_S[k] - \frac{2\sigma^2 \sum_{h=0}^{W_{av}-1} (W_{av} - |h|)}{W_S^2 T_s W_{av}^2}, \quad (27)$$

We have,

$$\sigma_z^2[k] = \mu_{\overline{RSS}}^2[k] + \sigma_{\overline{RSS}}^2[k] + T_{HO}^2 \mu_S^2[k] + T_{HO}^2 \sigma_S^2[k] - 2T_{HO} \mu_{\overline{RSS}}[k] \mu_S[k] +$$

$$\frac{4T_{HO}\sigma^2 \sum_{h=0}^{W_{av}-1} (W_{av} - |h|)}{W_S^2 T_S W_{av}^2} - E^2\{Z[k]\} \quad (28)$$

$$= \sigma_{RSS}^2[k] + T_{HO}^2 \sigma_S^2[k] + \frac{4T_{HO}\sigma^2 \sum_{i=0}^{W_{av}-1} (W_{av} - |h|)}{W_S^2 T_S W_{av}^2}. \quad (29)$$

One-Step Autocorrelation Coefficient

By definition,

$$\rho_{Z[k]Z[k-1]} = \frac{Cov(Z[k], Z[k-1])}{\sigma_{Z[k]}\sigma_{Z[k-1]}}. \quad (30)$$

Since $Z[k] = \overline{RSS}[k] - T_{HO} * S[k]$, we have

$$\begin{aligned} Cov(Z[k+1], Z[k]) &= E\{(Z[k+1] - \mu_Z[k+1])(Z[k] - \mu_Z[k])\} \\ &= R_{\overline{RSS}}[k+1, k] - T_{HO}E\{\overline{RSS}[k]S[k+1]\} - T_{HO}E\{S[k]\overline{RSS}[k+1]\} + \\ &\quad T_{HO}^2 R_S[k+1, k] - \mu_Z[k+1]\mu_{\overline{RSS}}[k] + T_{HO}\mu_Z[k+1]\mu_S[k] + \quad (31) \\ &\quad \mu_Z[k+1]\mu_Z[k] - \mu_Z[k]\mu_{\overline{RSS}}[k+1] + T_{HO}\mu_Z[k]\mu_S[k+1]. \quad (32) \end{aligned}$$

It can be shown that

$$R_{\overline{RSS}}[k+1, k] = \mu_{\overline{RSS}}[k+1]\mu_{\overline{RSS}}[k] + \frac{\sigma^2(W_{av} - 1)}{W_{av}^2}, \quad (33)$$

and from (4)

$$E\{\overline{RSS}[k]S[k+1]\} = \mu_{\overline{RSS}}[k+1]\mu_S[k] - \frac{2\sigma^2(2W_{av} - 1 + \sum_{h=1}^{W_{av}-1} (W_{av} - |h|))}{W_{av}^2 W_S^2 T_S} \quad (34)$$

$$E\{S[k]\overline{RSS}[k+1]\} = \mu_{\overline{RSS}}[k+1]\mu_S[k] - \frac{2\sigma^2 \sum_{h=1}^{W_{av}-1} (W_{av} - |h|)}{W_{av}^2 W_S^2 T_S} \quad (35)$$

and

$$\begin{aligned} R_S[k+1, k] &= \mu_S[k+1]\mu_S[k] + \frac{4\sigma^2}{W_{av}^2 W_S^4 T_S^2} * \\ &\quad \left(W_{av}(W_S - 2) + \sum_{h=1}^{W_{av}-1} (W_{av} - |h|)(2W_S - 4|h|) - (W_{av} + \right. \\ &\quad \left. \sum_{h=1}^{W_{av}-1} (W_{av} - |h|)(2|h|)) \right) \quad (36) \end{aligned}$$

$$\begin{aligned}
&= \mu_S[k+1]\mu_S[k] + \frac{4\sigma^2}{W_{av}^2 W_S^4 T_S^2} * \\
&\quad (W_{av}(W_S - 3) + \sum_{h=1}^{W_{av}-1} (W_{av} - |h|)(2W_S - 6|h|)). \tag{37}
\end{aligned}$$

By direct substitution from (15), (34), and (37) in (30), we get

$$\begin{aligned}
Cov(Z[k], Z[k-1]) &= \frac{\sigma^2}{W_{av}^2} [(W_{av} - 1) + \frac{4T_{HO} \sum_{h=1}^{W_{av}-1} (W_{av} - |h|)}{W_S^2 * T_S} + \\
&\quad \frac{4 * T_{HO}^2}{W_S^4 T_S^2} (W_{av}(W_S - 3) + \\
&\quad \sum_{h=1}^{W_{av}-1} (W_{av} - |h|)(2W_S - 6 * |h|))], \tag{38}
\end{aligned}$$

and consequently the $\rho_{Z[k]Z[k-1]}$ can be obtained by direct substitution from (29) and (38) in (30).

References

- [1] R. Berezdivin, R. Breinig, and R. Topp, "Next-generation wireless communications concepts and technologies," *IEEE Commun. Mag.*, vol. 40, no. 3, pp. 108 – 116, March 2002.
- [2] M. M. Buddhikot, G. Chandranmenon, S. Han, Y. W. Lee, S. Miller, and L. Salgarelli, "Integration of 802.11 and third generation wireless data networks," in *Proc. of IEEE INFOCOM*, San Francisco, US, Apr. 2003, pp. 503 – 512.
- [3] A. K. Salkintzis, C. Fors, and R. Pazhyannur, "WLAN-GPRS integration for next-generation mobile data networks," *IEEE Wireless Commun.*, vol. 9, no. 5, pp. 112 – 124, Oct. 2002.
- [4] I. F. Akyildiz, J. McNair, J. Ho, H. Uzunalioglu, and W. Wang, "Mobility management in current and future communications networks," *IEEE Network*, vol. 12, no. 4, pp. 39 – 49, Jul/Aug 1998.

- [5] B. Liang and Z. J. Haas, "Predictive distance-based mobility management for multi-dimensional pcs networks," *IEEE/ACM Transactions on Networking*, vol. 11, no. 5, pp. 718–732, October 2003.
- [6] J. Makela, M. Ylianttila, and K. Pahlavan, "Handoff decision in multi-service networks," in *Proc. of 11th IEEE International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC'00)*, vol. 1, London, UK, Sep. 2000, pp. 655 – 659.
- [7] G. P. Pollini, "Trends in handover design," *IEEE Commun. Mag.*, vol. 34, no. 3, pp. 82 – 90, March 1996.
- [8] A. Hatami, P. Krishnamurthy, K. Pahlavan, M. Ylianttila, J. Makela, and R. Pichna, "Analytical framework for handoff in non-homogeneous mobile data networks," in *Proc. of (PIMRC'99)*, Osaka, Japan., Sep. 1999, pp. 760 – 764.
- [9] M. Ylianttila, M. Pande, J. Makela, and P. Mahonen, "Optimization scheme for mobile users performing vertical hand-offs between IEEE 802.11 and GPRS/EDGE networks," in *Proc. of IEEE Global Telecommunications Conference GLOBECOM'01*, vol. 6, San Antonio, Texas, USA, Nov. 2001, pp. 3439 – 3443.
- [10] M. Ylianttila, J. Makela, and P. Mahonen, "Supporting resource allocation with vertical handoffs in multiple radio network environment," in *Proc. of IEEE International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC'02)*, Lisbon, Portugal, Sep. 2002, pp. 64 – 68.
- [11] H. Park, S. Yoon, T. Kim, J. Park, M. Do, , and J. Lee, "Vertical handoff procedure and algorithm between IEEE802.11 WLAN and CDMA cellular network." *Proc. CDMA Int'l Conf.*, 2002, pp. 103–112.
- [12] M. Ylianttila, R. Pichna, J. Vallstram, J. Makela, A. Zahedi, P. Krishnamurthy, and K. Pahlavan, "Handoff procedure for heterogeneous wireless networks," in *Proc. of IEEE Global Telecommunications Conference (GLOBECOM'99)*, vol. 5, Dec. 1999, pp. 2783 – 2787.
- [13] K. Pahlavan, P. Krishnamurthy, A. Hatami, M. Ylianttila, J. P. Makela, R. Pichna, and J. Vallstron, "Handoff in hybrid mobile data networks," *IEEE Commun. Mag.*, vol. 7, no. 4, pp. 34 – 47, Apr. 2000.

- [14] A. Majlesi and B. H. Khalaj, "An adaptive fuzzy logic based handoff algorithm for hybrid networks," in *Proc. of 6th International Conference on Signal Processing*, vol. 2, Aug. 2002, pp. 1223 – 1228.
- [15] H. Wang, R. H. Katz, and J. Giese, "Policy-enabled handoffs across heterogeneous wireless networks," in *Proc. of the Second IEEE Workshop on Mobile Computer Systems and Applications*, New Orleans, Louisiana, Feb. 1999, p. 51.
- [16] F. Zhu and J. McNair, "Optimizations for vertical handoff decision algorithms," in *Proc. of IEEE Wireless Communications and Networking Conference (WCNC)*, vol. 2, March 2004, pp. 867 – 872.
- [17] L. Chen, T. Sun, B. Chen, V. Rajendran, and M. Gerla, "A smart decision model for vertical handoff." The 4th Int'l Workshop on Wireless Internet and Reconfigurability (ANWIRE'04), May 2004.
- [18] M. M. Buddhikot, G. Chandranmenon, S. Han, Y. Lee, S. Miller, and L. Salgarelli, "Design and implementation of a WLAN/CDMA2000 interworking architecture," *IEEE Commun. Mag.*, vol. 41, no. 11, pp. 90 – 100, Nov. 2003.
- [19] T. S. Rappaport, *Wireless Communications: Principles and Practice*. Prentice Hall, July 1999.
- [20] K. Murakami, O. Haase, J. Shin, and T. LaPorta, "Mobility management alternatives for migration to mobile internet session-based services," *IEEE Journal on Selected Areas in Communications*, vol. 22, no. 5, pp. 834 – 848, June 2004.
- [21] C. Guo, Z. Guo, Q. Zhang, and W. Zhu, "A seamless and proactive end-to-end mobility solution for roaming across heterogeneous wireless networks," *IEEE Journal on Selected Areas in Communications*, vol. 22, no. 5, pp. 834 – 848, Jun. 2004.
- [22] C. Tepedelenlioglu and G. Giannakis, "On velocity estimation and correlation properties of narrow-band mobile communication channels," *IEEE Trans. on Vehicular Technology*, vol. 50, no. 4, pp. 1039 – 1052, Jul. 2001.
- [23] Q. Zhang, C. Guo, Z. Guo, and W. Zhu, "Efficient mobility management for vertical handoff between WWAN and WLAN," *IEEE Commun. Mag.*, vol. 41, no. 11, pp. 102 – 108, Nov. 2003.

- [24] N. Zhang and J. M. Holtzman, “Analysis of handoff algorithms using both absolute and relative measurements,” *IEEE Transactions on Vehicular Technology*, vol. 45, no. 1, pp. 174 – 179, Feb. 1996.
- [25] A. Papoulis and S. Pillai, *Probability, Random Variables and Stochastic Processes*, 4th ed. McGraw-Hill, 2002.
- [26] G. Bolch, S. Greiner, H. de Meer, and K. S. Trivedi, *Queuing networks and Markov Chains: Modeling and Performance Evaluation with Computer Science Applications*, 2nd ed. Wiley, August 1998.
- [27] IEEE Standard, “Part 11: Wireless LAN medium access control (MAC) and physical layer (PHY) specifications,” IEEE, 1999.
- [28] 3GPP2 Standard, “CDMA2000 wireless ip network standard: Packet data mobility and resource management,” 3GPP2 X.S0011-003-C v1.0, Sep. 2003.
- [29] N. Tripathi, J. Reed, and H. VanLandingham, “Handoff in cellular systems,” *IEEE Personal Commun.*, vol. 5, no. 6, pp. 26– 37, Dec. 1998.

List of Figures

1	Mobile handoff in heterogeneous wireless system	3
2	MO handoff algorithm	10
3	VHO Markov chain model	15
4	Number of handoffs ($\gamma = -90$ dBm).	18
5	Available bandwidth ($\gamma = -90$ dBm).	19
6	HoL packet delay rate ($\gamma = -90$ dBm, $\theta_D = 30$ ms).	19
7	Number of handoffs vs. ASST.	20
8	Available bandwidth vs. ASST.	21
9	HoL packet delay rate vs. ASST ($\theta_D = 30$ ms).	21
10	Total cost vs. ASST.	22
11	Optimal ASST vs. delay budget threshold and velocity ($c_H = 100$, $c_D = 100000$).	23
12	Optimal ASST vs. handoff cost and delay penalty ($V = 2$, $\theta_D = 50$).	24

List of Tables

1	<i>Simulation parameter values</i>	17
---	--	----