Hindawi Publishing Corporation
EURASIP Journal on Wireless Communications and Networking
Volume 2009, Article ID 372185, 11 pages
doi:10.1155/2009/372185

Research Article

Cost-Based Vertical Handover Decision Algorithm for WWAN/WLAN Integrated Networks

KunHo Hong,[1] SuKyoung Lee,[1] LaeYoung Kim,[2] and PyungJung Song[3]

[1] *Department of Computer Science, Yonsei University, Seoul 120-749, South Korea*
[2] *LG Electronics, South Korea*
[3] *Electronics and Telecommunications Research Institute (ETRI), Daejon 305-700, South Korea*

Correspondence should be addressed to SuKyoung Lee, sklee@cs.yonsei.ac.kr

Received 31 October 2008; Revised 5 March 2009; Accepted 28 May 2009

Recommended by Mohamed Hossam Ahmed

Next generation wireless communications are expected to rely on integrated networks consisting of multiple wireless technologies. Heterogeneous networks based on Wireless Local Area Networks (WLANs) and Wireless Wide Area Networks (WWANs) can combine their respective advantages on coverage and data rates, offering a high Quality of Service (QoS) to mobile users. In such environment, multi-interface terminals should seamlessly switch from one network to another in order to obtain improved performance or at least to maintain a continuous wireless connection. Therefore, network selection algorithm is important in providing better performance to the multi-interface terminals in the integrated networks. In this paper, we propose a cost-based vertical handover decision algorithm that triggers the Vertical Handover (VHO) based on a cost function for WWAN/WLAN integrated networks. For the cost function, we focus on developing an analytical model of the expected cost of WLAN for the mobile users that enter the double-coverage area while having a connection in the WWAN. Our simulation results show that the proposed scheme achieves better performance in terms of power consumption and throughput than typical approach where WLANs are always preferred whenever the WLAN access is available.

1. Introduction

Beyond 3G mobile networks are being defined as the integration of diverse access technologies [1, 2]. One promising way to construct integrated wireless networks is to integrate IEEE 802.11 Wireless Local Area Networks (WLANs) [3] and Wireless Wide Area Networks (WWANs), such as Universal Mobile Telecommunications System (UMTS) by 3rd Generation Partnership Project (3GPP) [4] and Code Division Multiple Access 2000 (CDMA2000) by 3GPP2 [5]. These networks have characteristics that perfectly complement each other in terms of both their data rates and coverage. Additionally, they are both widely deployed already. As a result, WWAN/WLAN-integrated networks can provide their users with a wider range of services at a higher quality. In the integrated wireless networks; a Mobile Node (MN) can roam dynamically between different wireless networks. Thus, one of the most important issues for integrated wireless networks is the Vertical Handover (VHO), which refers to a handover between systems with different air interfaces.

VHOs are classified into two categories: a downward VHO (DVHO) is a handover to a wireless network with a smaller cell size and (usually) higher bandwidth per unit area; an upward VHO (UVHO) is a handover to a wireless network with a larger cell size and (usually) lower bandwidth per unit area [6, 7].

The natural trend for VHOs has been to use small-coverage, high-bandwidth networks, such as WLAN, whenever they are available, and to remain connected to the network for as long as possible. Hence, the switch to low-bandwidth networks, such as UMTS and CDMA2000, occurs only when a high-bandwidth network is not available [6]. This type of network selection scheme for WWAN/WLAN-integrated networks is called the WLAN-first scheme, where WLANs are always preferred by all services whenever the WLAN access is available to possibly increase bandwidth and reduce costs compared to WLANs. The studies in [8, 9] propose the WLAN-first scheme, in which the VHO decision is based mainly on Received Signal Strength (RSS).

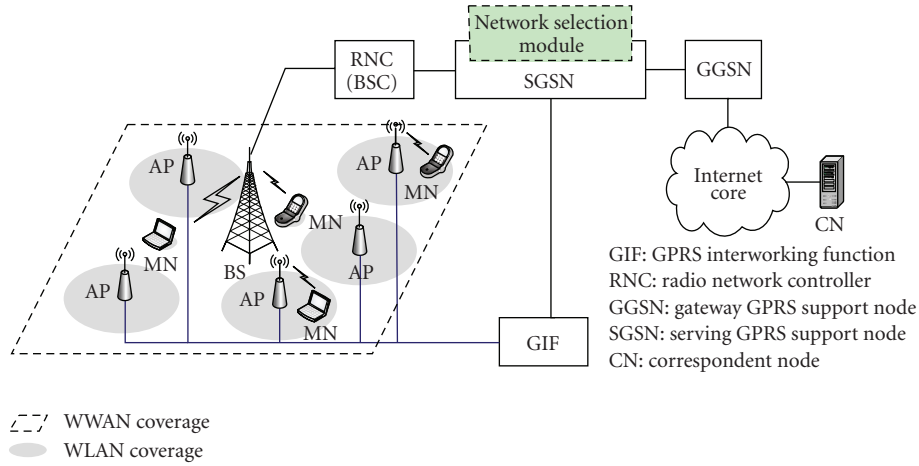

Figure 1: An overall heterogeneous networking system architecture for the proposed vertical handover decision algorithm.

The studies in [10–16] propose cost-based VHO schemes that consider several factors, such as user preferences and network conditions, in order to perform flexible VHOs. In the cost-based VHO schemes, MN selectively changes its connection from WWAN to WLAN based on the cost function, even though the WLAN access is available. The authors of [10] consider higher level parameters which fall in the transport and application layers in addition to the available access networks' characteristics. In [11], the handoff decision is made by the system, aiming to achieve a better performance in the blocking probability. The authors of [12] address the main issues that arise while implementing the interoperability mechanisms between two different radio access networks. Although these VHO decision algorithms have certain advantages, they do not consider power efficiency, which is one of the most important aspects of wireless networks. This is especially true of integrated wireless networks since MNs with more than one wireless interface consume more power than those with one wireless interface. The studies in [13–16] consider the power consumption as one of the metrics for deciding VHO. However, they do not provide detailed analytical models to estimate network conditions such as available bandwidth and average user throughput in addition to power consumption.

In this paper we propose a new vertical handover decision algorithm based on a cost function for WWAN/WLAN-integrated networks. In the proposed VHO decision algorithm, when an MN moves from the WWAN-only area to an area that is covered by both WLAN and WWAN, DVHO is performed only if the cost of WLAN is less than that of WWAN. Our cost function considers data throughput, power consumption, and monetary cost. In addition, user preferences for these factors are considered to select an optimal network for each MN. To estimate how much throughput is obtained and how much power is drained to send and receive a bit in the target WLAN, we develop an analytical model based on Automatic Rate Fallback (ARF) [17], which is generally used for adaptive modulation schemes in IEEE 802.11 WLAN. Adaptive modulation schemes have been studied extensively and advocated at the physical layer in order to enhance throughput by adaptively matching transmission parameters to time-varying wireless channel conditions [18, 19]. Therefore, in order to benefit from the use of adaptive modulation, we assume that our VHO decision algorithm operates on IEEE 802.11 WLAN, which employs adaptive modulation.

The rest of this paper has the following structure. Section 2 explains the details of our proposed scheme, including an analysis of the expected available bandwidth and the expected power consumption for WLAN with the adaptive modulation technique. Section 3 shows the effectiveness of the proposed scheme using the results from thorough simulations of various scenarios. Section 4 presents our conclusions.

2. Vertical Handover Decision Algorithm

2.1. Overall Architecture for the Vertical Handover Decision Algorithm. Figure 1 illustrates an overall heterogeneous networking system architecture for the proposed VHO decision algorithm. As shown in Figure 1, our system model is based on the generic architecture for integration of WWAN and WLAN defined in [20]. In this architecture, the WLAN is connected to the Serving General Packet Radio Service (GPRS) Support Node (SGSN) via the GPRS Interworking Function (GIF) which provides a standardized interface to the GPRS core network and virtually hides the WLAN peculiarities. The primary function of the GIF is to make the SGSN consider the WLAN as a typical GPRS access system. That is, GIF is in charge of interfacing the WLANs with the WWAN by connecting to all Access Points (APs) in the WLANs as well as to SGSN in the WWAN. In the system, MN utilizes Packet Data Protocol (PDP) to manage its ongoing sessions. When the MN enters the heterogeneous wireless networking system, a PDP address (usually an IP address) is allocated to the MN by a Dynamic Host Configuration Protocol (DHCP) server for IP connection. The PDP context

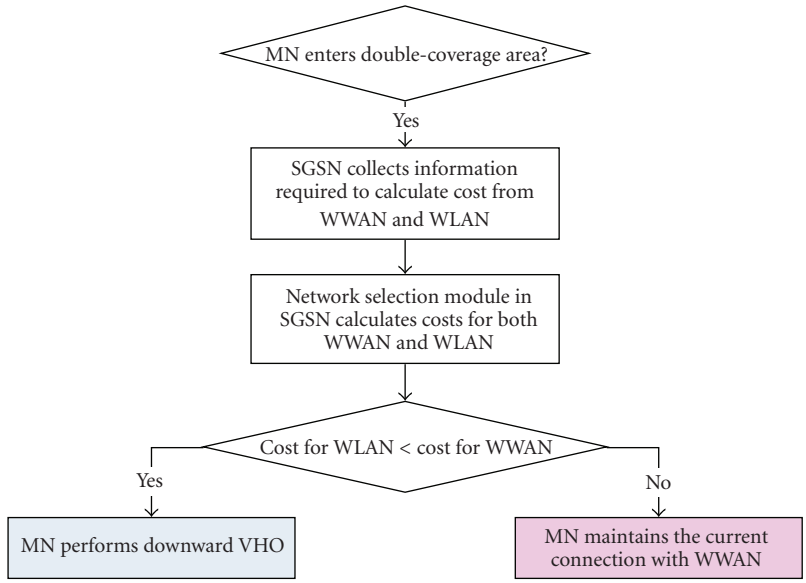

Figure 2: The flow chart of the proposed vertical handover decision algorithm.

can be maintained in the system shown in Figure 1 when the MN changes an access technology. Thus, when a VHO occurs, the packets destined to the MN can be rerouted at the SGSN by using the intra/inter-SGSN Routing Area Update (RAU) procedure defined in [20] without going through the reauthentication process.

We add the *network selection module* in SGSN to implement the proposed VHO decision algorithm. In our integrating architecture, the *network selection module* in the SGSN not only collects all the information related to the vertical handover but also makes the decision about what network to which an MN connect-based on the results of a cost function. Therefore, GIF must provide the SGSN with information about its WLAN, which is required to calculate the cost for WLAN.

Figure 2 shows the flow chart for the proposed vertical handover decision algorithm. As shown in Figure 2, when an MN enters the double-coverage area that is covered by both WLAN and WWAN, SGSN collects all the required information for calculating cost from both WWAN and WLAN. The collected information includes the current power consumption, data rate, and monetary cost of each bit for WWAN as well as the available modulation levels, Packet Error Rate (PER) at each modulation level, expected power consumption, and monetary cost of each bit for the target WLAN. Based on this information, SGSN calculates the cost for both networks using the proposed cost function and then decides whether the MN will perform DVHO or not. That is, if the cost for WLAN is less than that for WWAN, the MN will switch from WWAN to WLAN. The result of this decision is delivered to the MN, allowing the MN to connect to the best available network. On the other hand, UVHO is performed when the MN moves from the double-coverage area to the WWAN-only area based on RSS.

We now explain the construction of the cost function. The factors that are taken into account for our cost function are data throughput, power consumption, and monetary cost. The cost function for network I is defined as

$$C(I) = \mu_t f_t(I) + \mu_p f_p(I) + \mu_m f_m(I) \quad \text{for } I \in \{\text{wlan}, \text{wwan}\}, \tag{1}$$

where $f_t(I)$, $f_p(I)$, and $f_m(I)$ denote the normalized variables for data rate, power consumption, and monetary cost in network I, respectively. In (1), μ_t, μ_p, and μ_m denote the weights of each factor, which are set according to the user preference. The constraint pertaining to μ_t, μ_p, and μ_m is given by $\mu_t + \mu_p + \mu_m = 1$. Note that network I improves with lower values of $C(I)$.

$f_t(I)$ is obtained as follows:

$$f_t(I) = \frac{1}{e^{\alpha_t D(I) + \beta_t}}, \tag{2}$$

where α_t and β_t are variables to adjust between $f_t(I)$, $f_p(I)$, and $f_m(I)$ of (1). Because the three factors have different ranges of values, each factor has adjusting variables. $D(I)$ is the throughput when using network I. Because the MN uses WWAN as its current network, it knows the value of $D(\text{wwan})$ by measurement. However, the MN does not measure the value of $D(\text{wlan})$ directly, and so the value is obtained as follows:

$$D(\text{wlan}) = \begin{cases} \hat{B}(\text{wlan}), & \hat{B}(\text{wlan}) < D(\text{wwan}), \\ \rho \cdot \hat{B}(\text{wlan}) + (1-\rho) \cdot D(\text{wwan}), & \hat{B}(\text{wlan}) \geq D(\text{wwan}), \end{cases} \tag{3}$$

where $\hat{B}(\text{wlan})$ is the expected available bandwidth of network wlan. If $\hat{B}(\text{wlan})$ is smaller than the current throughput, $D(\text{wwan})$, then the expected throughput, D(wlan), should

be $\widehat{B}(\text{wlan})$ because the network cannot send or receive more than $\widehat{B}(\text{wlan})$. If the available bandwidth is larger than current throughput, however, then the MN can send or receive more than $D(\text{wwan})$. This means that the value of $D(\text{wlan})$ is between $\widehat{B}(\text{wlan})$ and $D(\text{wwan})$. To apply this feature of the heterogeneous network to applications with different bandwidth requests, we introduce ρ, the value that represents dependency between the throughput requested by application and the network bandwidth. ρ is given between 0 and 1. It is dependent on current application that is using the network. If the MN uses real-time (rt) UDP applications, then the value is close to 0. If it uses greedy Best-Effort (BE) TCP applications, then the value is close to 1. For example, rt UDP applications such as VoIP or video call do not need to increase network throughput, although they are connected to a high-bandwidth network. That is, such applications occupy certain bandwidth as addressed in [21]. Thus, ρ is close to 0. If the MN has been using an rt UDP application in the WWAN-only area, that means that the WWAN admitted the MN's rt traffic by considering the WWAN's load status and checking whether the WWAN can allocate a certain bandwidth requested by the application (i.e., $D(\text{wwan})$). If BE TCP applications like FTP are used, however, then the throughput, that can be achieved by the applications, is dependent on the available bandwidth [21], and so ρ is close to 1. On the other hand, if the MN' BE TCP traffic was admitted in the WWAN with high load, $D(\text{wwan})$ must be small for the MN, and hence, the throughput for the MN's BE application may be increased by handing over to the WLAN. The value of $\widehat{B}(wlan)$ is calculated in Section 2.2.

We now calculate the values of $f_p(I)$ and $f_m(I)$. Let $pw(I)$ and $m(I)$ be the power consumption and monetary cost for a user in network I, respectively. The expected amount of power consumption for the target WLAN, $pw(\text{wlan})$, is derived analytically in a later section, and $m(I)$ can be obtained by $\text{Price}(I)/D(I)$, where $\text{Price}(I)$ indicates the price to transceive data in network I per unit time. We can then determine the normalized amount of power consumption as follows:

$$f_p(I) = \frac{1}{e^{\alpha_p/pw(I)+\beta_p}}. \tag{4}$$

The normalized monetary cost is defined as

$$f_m(I) = \frac{1}{e^{\alpha_m/m(I)+\beta_m}}. \tag{5}$$

α_p, β_p, α_m, and β_m in (4) and (5) represent the adaptation variables. Note that we take the inverse number for $m(I)$ and $pw(I)$ in (4) and (5), respectively, since lower values are better.

Whenever an MN enters the double-coverage area of WWAN and WLAN in the proposed vertical handover decision algorithm, SGSN analyzes the user preference so that μ_t, μ_p, and μ_m in (1) are set. Based on these weights as well as the information collected from WLAN and WWAN, then SGSN calculates the costs for both WLAN and WWAN, $C(\text{wlan})$ and $C(\text{wwan})$, respectively. The MN is allowed to perform DVHO if the estimated cost for WLAN is less than that for WWAN. Otherwise, the MN maintains a connection with the WWAN.

2.2. Expectation for Available Bandwidth and Power Consumption. In this section, we develop an analytical model to estimate the amount of power that is drained to send or receive a bit in the target WLAN. We start by establishing an analytical model of the available bandwidth for WLAN assuming that n-level modulation is allowed. We then estimate the amount of power consumption using this analytical model.

2.2.1. Available Bandwidth Expectation for WLAN. In order to estimate the available bandwidth for WLAN, we determine the probability that the modulation is escalated after the transmission of k packets. This probability can be expressed as a function of the probability that a packet transmission succeeds, q, and the number of transmitted packets, k. Let $f_u(k,q)$ be this probability. Similarly, let $f_d(k,p)$ denote the probability that the modulation is lowered after the transmission of packet k, where p is the probability that a packet transmission fails, that is, packet error rate (PER). For analysis, we assume that the wireless channel is a Gaussian channel, in which each bit has the same bit error probability, and bit errors are identically and independently distributed over the whole frame as in [22].

Let N be the number of successfully transmitted packets that is required to elevate the modulation level, and let K be the number of unsuccessfully transmitted packets that is required to degrade the modulation level. That is, the next higher modulation level is selected after transmitting consecutive N frames successfully while the next lower modulation level is selected after failing to transmit consecutive K frames. Depending on the value of k, $f_u(k,q)$ can be decomposed into four cases as follows.

(1) When $k < N$. The modulation remains unchanged since consecutive N packets are not successfully transmitted, resulting in $f_u(k,q) = 0$.

(2) When $k = N$. All N packets in the transmission should succeed so that $f_u(k,q)$ is computed as q^N.

(3) When $k = N+1$. In this case, the first packet transmission should fail, followed by N consecutive successful packets, resulting in pq^N.

(4) When $k > N+1$. As shown in Figure 3, we can divide this case into two components. In the first component, the modulation should remain unchanged during the transmission of the first $k-N-1$ packets. Let $A_u(k-N-1,q)$ be this probability. In the second component, the $(k-N)$th packet transmission should fail, followed by the successful transmission of N consecutive packets, resulting in pq^N. Thus, $f_u(k,q) = A_u(k-N-1,q)pq^N$ is the joint probability. Note that the $(k-N)$th packet failure is necessary. Otherwise, the modulation can be elevated after the $(k-N)$th packet:

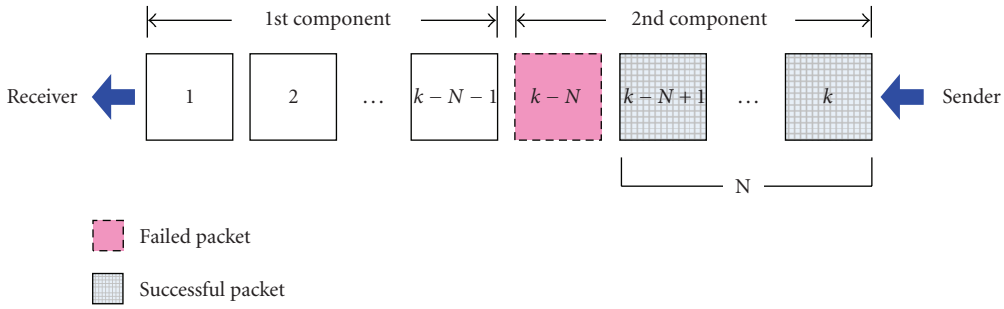

Figure 3: The case that the modulation is elevated after the transmission of k packets when $k > N + 1$.

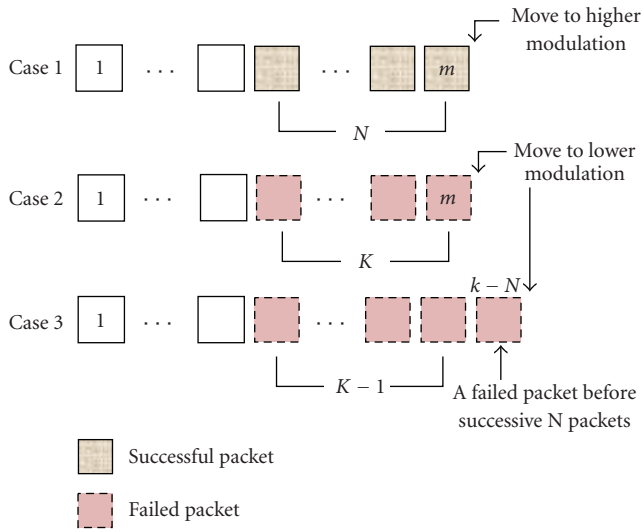

Figure 4: Three complement cases for $A_u(k - N - 1, q)$.

$$f_u(k,q) = \begin{cases} 0, & k < N, \\ q^N, & k = N, \\ pq^N, & k = N+1, \\ A_u(k-N-1,q)\,pq^N, & k > N+1. \end{cases} \tag{6}$$

In order to obtain $A_u(k - N - 1, q)$ in (6), we subtract the complement probabilities from 1 instead of calculating them directly. The complement events, which mean that the modulation is changed during the transmission of the first $k-N-1$ packets, consist of the three cases shown in Figure 4.

Case 1. The modulation can be elevated during the transmission of the first $k - N - 1$ packets. This probability can be recursively computed by $\sum_{m=1}^{k-N-1} f_u(m, q)$.

Case 2. In contrast with Case 1, the modulation can be lowered before the $(k - N)$th packet, though it is still computed by $\sum_{m=1}^{k-N-1} f_d(m, p)$, as in Case 1.

Case 3. It is also possible that the modulation remains unchanged after transmitting $k - N - 1$ packets, but after sending the $K - N$th packet, the modulation degrades. In this case, the failure of K sequential packets occurs after the first phase, as shown in Figure 4. It is not included in Cases 1 and 2. $f_d(k - N, p)$ is the probability that modulation falls at $k-N$th packet. However, the last packet, or $k-N$th packet, is not included in the first phase, and so the probability of this case is $(1/p) f_d(k - N, p)$.

By summing up all of the above three cases, we have $A_u(k - N - 1, q) = 1 - \sum_{m=1}^{k-N-1} \{f_u(m, q) + f_d(m, p)\} - (1/p) f_d(k - N, p)$.

Using the opposite derivation process, we can also obtain $f_d(k, p)$ as follows:

$$f_d(k,p) = \begin{cases} 0, & k < K, \\ p^K, & k = K, \\ qp^K, & k = K+1, \\ A_d(k-K-1,p)\,qp^K, & k > K+1, \end{cases} \tag{7}$$

where $A_d(k-K-1, p) = 1 - \sum_{m=1}^{k-K-1} \{f_u(m, q) + f_d(m, p)\} - (1/p) f_u(k - N, p)$.

Using (6) and (7), we can calculate the average number of packets until the modulation is altered when the current modulation is given. Let $g_u(q)$ and $g_d(p)$ denote the average numbers of packets until the modulation is escalated and lowered, respectively. Along with $f_u(k, q)$ and $f_d(k, p)$, we have

$$\begin{aligned} g_u(q) &= \sum_{k=1}^{\infty} k f_u(k,q), \\ g_d(p) &= \sum_{k=1}^{\infty} k f_d(k,p). \end{aligned} \tag{8}$$

A key observation found in the ARF algorithm is that the transmission rate is always switched to adjacent one, so that the rate adaptation procedure of ARF could be expressed via a birth-death Markov chain as in [23]. Figure 5 shows the Markov chain where the state i represents the modulation level i of the single target station. Then, we can model the n-level modulation technique as a simple queueing system with the state of the modulation level as shown in Figure 5, where λ_i and μ_i denote the rates that the modulation is elevated from M_i to M_{i+1} and lowered from M_{i+1} to M_i, respectively.

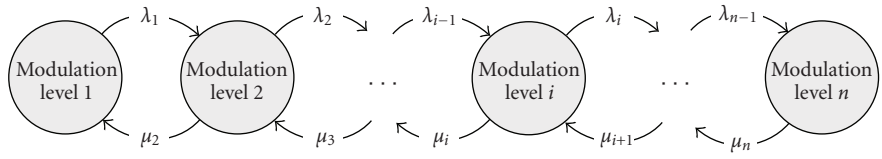

Figure 5: The queueing system with the state of modulation level.

λ_i and μ_i can then be expressed as functions of $g_u(q_i)$ and $g_d(p_i)$, respectively:

$$\lambda_i = \begin{cases} \dfrac{1}{g_u(q_i)}, & 1 \le i < n, \\ 0, & i = n, \end{cases} \tag{9}$$

$$\mu_i = \begin{cases} 0, & i = 1, \\ \dfrac{1}{g_d(p_i)}, & 1 < i \le n, \end{cases} \tag{10}$$

where q_i and p_i represent the success and failure probabilities, respectively, when the modulation M_i is used.

From (9) and (10), the steady-state probability that the modulation level i is used, $f(i)$, can be derived as

$$f(i) = \frac{\prod_{j=1}^{i-1}\lambda_j}{\prod_{j=2}^{i}\mu_j} f(1), \quad \text{for } 1 < i \le n, \tag{11}$$

where $\sum_{i=1}^{n} f(i) = 1$.

We assume that B_i and p_i are the maximum bandwidths provided by the target WLAN and the PER at the modulation level i, respectively. Finally, we obtain B_{wlan}, the expected bandwidth of the WLAN, as follows:

$$B_{\text{wlan}} = \sum_{i=1}^{n} (1 - p_i) f(i) B_i. \tag{12}$$

Let θ_{wlan} be the current usage of the WLAN. θ_{wlan} is defined as the fraction of time during which the AP finds that the channel is busy for unit time: it can be obtained by the AP of the WLAN as in [21]. Then, we obtain the expected available bandwidth of the WLAN, $\hat{B}(\text{wlan})$, as follows:

$$\hat{B}_{\text{wlan}} = (1 - \theta_{\text{wlan}}) \cdot B_{\text{wlan}}. \tag{13}$$

2.2.2. Estimation of Power Consumption. In this section we estimate the amount of power consumption for both WLAN and WWAN. Let D_{in} and D_{out} denote the receiving and sending bit rates in WWAN, respectively. We then obtain the average amount of power consumption required to send or receive a single bit in WWAN, $pw(\text{wwan})$, by measuring the average amount of power consumption per unit time, δ_{wwan}, as follows:

$$pw(\text{wwan}) = \frac{\delta_{\text{wwan}}}{(D_{\text{in}} + D_{\text{out}})}. \tag{14}$$

As done in WWAN, we denote the receiving and sending bit rates in WLAN by D'_{in} and D'_{out}. Let δ^r_{wlan}, δ^t_{wlan}, and δ^i_{wlan} denote the amount of power consumed per unit time to receive, send, and stay-in-idle mode in WLAN, respectively. We can then obtain the average amount of power consumption in WLAN as follows:

$$pw(\text{wlan}) = \frac{\delta^r_{\text{wlan}} D'_{\text{in}} + \delta^t_{\text{wlan}} D'_{\text{out}} + \delta^i_{\text{wlan}} (B_{\text{wlan}} - (D'_{\text{in}} + D'_{\text{out}}))}{(D'_{\text{in}} + D'_{\text{out}})}, \tag{15}$$

where B_{wlan} is the available effective bandwidth for WLAN and can be obtained as derived in Section 2.2.1.

Here we assume that the bit rate in WLAN is affected by the difference in overall network bandwidth between WLAN and WWAN as well as the characteristic of ongoing applications. The amount of data throughput in WLAN, D', is then computed as

$$D' = D'_{\text{in}} + D'_{\text{out}} = (D_{\text{in}} + D_{\text{out}}) \left\{ \rho \left(\frac{B_{wlan}}{B_{wwan}} - 1 \right) + 1 \right\}, \tag{16}$$

where ρ represents the implementation variable based on the characteristic of application, as mentioned earlier ($0 \le \rho \le 1$). Assuming that D'_{in} and D'_{out} are adapted to the increased/decreased bandwidth with the same ratio, our proposed scheme estimates a reduction in power consumption by using (14), as well as (15), if the MN performs DVHO.

3. Performance Evaluation via Simulation

We have developed the simulator using C++ to investigate the effectiveness of our proposed scheme in comparison with that of the existing WLAN-First Scheme (WFS) and a cost-based scheme [14] where a cost function of power consumption, available bandwidth, and monetary cost, is used for network selection. The proposed scheme takes account of the MN's RSS and ARF when computing the cost for each network, whereas the cost-based scheme does not. Our simulation model accounts for the throughput and power consumption as performance metrics. The number of VHOs is also evaluated.

3.1. Simulation Environment. In our simulation, we use the network topology illustrated in Figure 1 in which five APs are deployed in a single WWAN coverage of $500 \times 500\,\text{m}^2$. As shown in Figure 1, each AP in the network topology covers a circular area with a radius of 100 m and employs IEEE 802.11a standard protocol [24], including an adaptive modulation technique.

Table 1: The characteristics of modulation for 802.11A.

Modulation level (i)	1	2	3	4	5
Bandwidth (B_i)	6 Mbps	12 Mbps	18 Mbps	36 Mbps	54 Mbps
Modulation	BPSK	QPSK	QPSK	16-QAM	64-QAM
Coding Rate	1/2	1/2	3/4	3/4	3/4
a_i	274.7229	90.2514	67.6181	53.3987	35.3508
g_i	7.9932	3.4998	1.6883	0.3956	0.0900
$\gamma_{pi}(dB)$	− 1.5331	1.0942	3.9722	10.2488	15.9784

In our simulation, the bandwidth for WLAN is set to a maximum of 54 Mbps, while the WWAN uses the bandwidth of High-Speed Downlink Packet Access (HSDPA), 14 Mbps. The data rate for WLAN changes adaptively based on an adaptive modulation technique through which five modulation levels are available. Table 1 shows the characteristics of the five modulation levels for 802.11a adopted in our simulation [19, 24]. Both K and N for adaptive modulation are set to 10.

In our simulation, the PER at each modulation level i for WLAN is modeled in [19, Equation (5)] as follows:

$$p_i(\gamma) \approx \begin{cases} 1, & \text{if } 0 < \gamma < \gamma_{pi} \\ a_i e^{-g_i \gamma}, & \text{if } \gamma \geq \gamma_{pi}, \end{cases} \tag{17}$$

where γ represents the received signal-to-noise ratio (SNR). The level-dependent parameters a_i, g_i, and γ_{pi} are obtained by fitting (17) to the exact PER. Supposing that a packet length is 1080 bits, then Table 1 provides the fitting parameters for transmission modes. Assuming that the path-loss exponent is 3, the received SNR, γ, is computed as follows [18]:

$$\gamma = P_t - PL(d0) - 30 \log_{10} d + G(t), \tag{18}$$

where P_t is the transmit power, $PL(d0) = 46.7$ dB is the mean path loss at a close reference distance of $d0 = 1$ m, d is the distance in meters, and $G(t)$ is a time-variant Rayleigh-distributed fading gain, for which the well-known Jakes' Doppler spectrum is assumed.

The Correspondent Node (CN) generates two different classes of traffic: BE TCP traffic ($\rho = 1$) and UDP traffic ($\rho = 0$). The data session arrival rate follows a Poisson process with mean 1/300. The packet size is 1000 bytes. In the simulation, TCP Reno has been implemented, where the initial and maximum window sizes are set to 1 and 1024 packets, respectively. Each session generates a file with the size varying from 1 to 11 Mbytes. To simulate the UDP traffic, On/Off traffic is generated, where the On and Off periods are exponentially distributed with means of 100 seconds and 300 seconds, respectively. During On period, traffic is generated with a rate of 20 to 200 Kbytes/s. We obtain the values of (α_t, β_t), (α_p, β_p), and (α_m, β_m) in (2), (4), and (5) from the simulation such that they normalize the value of each factor, that is, $f_t(I)$, $f_p(I)$, and $f_m(I)$. Table 2 gives these values. In our simulation, the amounts of power consumed to transmit, receive, and stay in an idle state are set to 1.828 W, 1.0494 W, and 0.6699 W, respectively, for the WLAN interface. These amounts are set to 2.805 W, 0.495 W, and 0.066 W, respectively, for the WWAN interface. The ratio of Price(wwan) and Price(wlan) is set to 5 to 1. The simulation time in the simulator is 100 hours.

Table 2: Adaptation variables for each factor.

α_t	0.003	β_t	0
α_p	0.6	β_p	0
α_m	2	β_m	0

Table 3: Parameter values for mobility model.

Parameter	Input 1	Input 2
$v_{\max}$	0.66 m/s	1.54 m/s
$a_{\min}$	-0.18 m/s^2	-0.42 m/s^2
$a_{\max}$	0.12 m/s^2	0.28 m/s^2
T_v	125 s	
$T_{\varphi_{\text{new}}}$	600 s	

For the mobility of MNs, we use the Smooth Random Mobility model, in which there is a maximum speed, $v_{\max}$ m/s, and a set of two preferred speeds, $v_{\text{pre f1}} = (3/5)v_{\max}$ and $v_{\text{pre f2}} = v_{\max}$ [25]. In our simulation, a total of six MNs move at pedestrian speeds. The initial velocity of each MN is chosen from the two preferred speeds, $v_{\text{pre f1}}$, $v_{\text{pre f2}}$, and the range $[0, v_{\max}]$ with probabilities of 0.2, 0.5, and 0.3, respectively. We assume a uniform distribution on the range of $[0, v_{\max}]$. All the MNs move with the selected speed until a new target speed is set again. The target speed is updated at every time interval which follows an exponential distribution with a mean of T_v s. Whenever a new target speed is decided, the MN decelerates or accelerates until the target speed is reached or a new target speed is decided again. $a_{\min}$ and $a_{\max}$ denote the maximum possible deceleration and acceleration in m/s^2, respectively. The MN selects a random acceleration/deceleration value from $[a_{\min}, a_{\max}]$. It also decides whether to change its movement direction or to maintain it with probabilities of 1/6 and 5/6, respectively. This is done at every time interval with an exponential distribution, the mean of which is $T_{\varphi_{\text{new}}}$ s.

All the parameter values for the mobility model are obtained from [25, 26] and are given in Table 3. In the simulation, MNs move in accordance with two input sets that represent two different levels of mobility, as shown in Table 3. That is, half of the MNs move according to Input 1, and the other half of the MNs move according to Input 2.

3.2. Simulation Results. We first run simulation tests by setting the weights (μ_m, μ_t, μ_p) in (1), which represent user preference, all to 1/3 for the proposed scheme. To determine the impact of the weights in (1) on the performance of the proposed scheme in terms of throughput, power consumption, and monetary cost, we repeat the simulation with

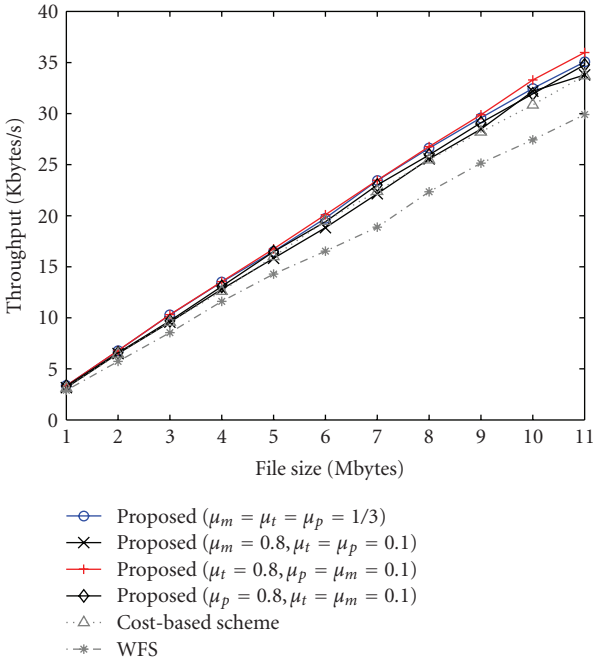

FIGURE 6: File size per session versus throughput when the TCP traffic ($\rho = 1$) is generated.

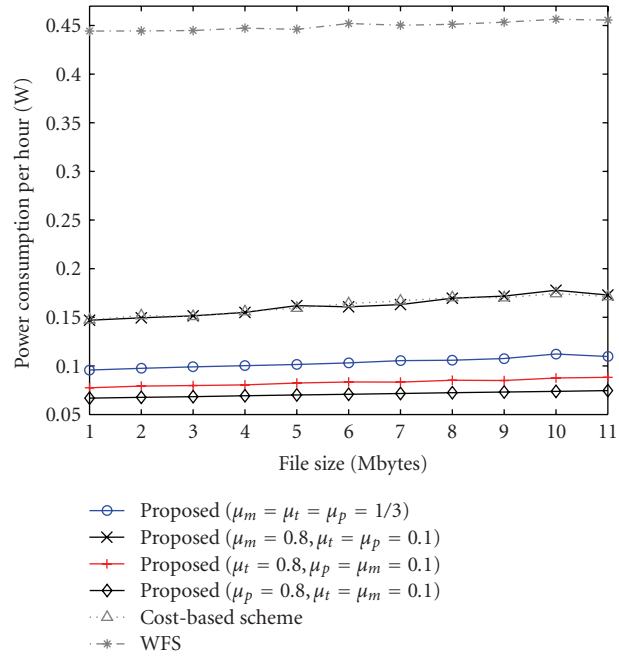

FIGURE 7: File size per session rate versus power consumption per minute when the TCP traffic ($\rho = 1$) is generated.

different combinations of μ_m, μ_t, and μ_p. We run simulation tests for the proposed scheme with another three weight sets that make the data throughput, the power consumption, and the monetary cost most important, respectively, among the three factors in our cost function. The three weight sets are $(\mu_m, \mu_t, \mu_p) = (0.8, 0.1, 0.1)$, $(0.1, 0.8, 0.1)$, and $(0.1, 0.1, 0.8)$. The values for performance metrics in simulation results are those averaged over all the MNs.

Figure 6 shows the throughput performance for different file sizes when the TCP traffic ($\rho = 1$) is generated. We observe from Figure 6 that both the proposed and the cost-based schemes achieve much better throughput performance than WFS over the whole range of the file size. Specifically, the throughput improvement by our scheme over WFS is 17.3% in average. Table 4(a) shows the relative throughput performance gains of our scheme with the four weight sets and the cost-based scheme over WFS when ρ is 1. We see that when a higher weight is put on the throughput (i.e., $\mu_t = 0.8$), our scheme shows the highest throughput improvement. It can be also seen from Table 4(a) that the throughput is also improved by the proposed scheme over the cost-based scheme.

Figure 7 shows the power consumption per minute in Watts versus the different file sizes when the TCP traffic ($\rho = 1$) is generated. We see from Figure 7 that both the proposed and the cost-based schemes show less power consumption than WFS for all the values of the file size. It is also observed from Figure 7 that our scheme consumes less power than the cost-based scheme except for the case of $\mu_m = 0.8$, in which the power consumption for the proposed scheme is almost same as that for the cost-based scheme since 80% of the weight is put on the monetary cost in the proposed scheme. Specifically, Table 4(b) shows the relative performance gains of our scheme with the four weight sets and the cost based scheme over WFS in terms of the power consumption when ρ is 1. As expected, the proposed scheme with $\mu_p = 0.8$ shows the most performance gain. Our scheme decreases the power consumption per minute from those of WFS and the cost-based scheme by 76.71% and 35.35%, respectively, in average.

Figure 8 shows the number of VHOs per hour for the different file sizes when the TCP traffic ($\rho = 1$) is generated. We see that the WFS yields a larger number of VHOs, compared to both the proposed and the cost-based schemes. In the proposed scheme, the average number of VHOs performed by each MN is 85.51% less than that of WFS in average. This is because WFS lets each MN perform DVHO whenever the MN enters the double-coverage area. The cost-based scheme shows almost the same number of VHOs as the proposed scheme with $\mu_m = 0.8$, whereas our scheme performs 73.7% less VHOs than the cost-based scheme, in average.

From Figures 6, 7, and 8, we can see that when the TCP traffic is generated, our scheme outperforms WFS in terms of the throughput and the power consumption for the different file sizes, while yielding a smaller number of VHOs than WFS. It can be also seen that the proposed scheme achieves better performance than the cost-based scheme in terms of the throughput, the power consumption, and the number of VHOs, except for the case in which 80% of weight is put on the monetary cost.

We now evaluate the performance of our VHO decision algorithm when the UDP traffic ($\rho = 0$) is generated. Table 5 shows the throughput performance for different

Table 4: Relative performance gain of the proposed and the cost-based schemes over WFS when the TCP traffic ($\rho = 1$) is generated.

(a) Relative throughput gain

File size (Mbyte)	Proposed scheme with $\{\mu_m, \mu_t, \mu_p\}$ = $\{1/3, 1/3, 1/3\}$	$\{0.8, 0.1, 0.1\}$	$\{0.1, 0.8, 0.1\}$	$\{0.1, 0.1, 0.8\}$	Cost-based scheme
1	15.71%	8.92%	15.78%	14.43%	7.73%
2	18.48%	13.54%	18.56%	14.76%	12.75%
3	20.30%	11.98%	20.57%	13.50%	12.19%
4	16.36%	10.30%	16.65%	12.64%	8.58%
5	15.50%	10.83%	17.29%	15.32%	12.04%
6	19.54%	13.86%	21.54%	17.38%	17.72%
7	24.18%	17.30%	24.34%	21.84%	18.61%
8	19.10%	14.41%	19.80%	16.10%	14.05%
9	17.65%	13.25%	18.84%	15.74%	12.22%
10	18.36%	17.37%	21.39%	16.44%	12.49%
11	17.23%	12.97%	20.30%	16.27%	12.78%

(b) Relative power consumption gain

File size (Mbyte)	Proposed scheme with $\{\mu_m, \mu_t, \mu_p\}$ = $\{1/3, 1/3, 1/3\}$	$\{0.8, 0.1, 0.1\}$	$\{0.1, 0.8, 0.1\}$	$\{0.1, 0.1, 0.8\}$	Cost-based scheme
1	78.45%	66.94%	82.57%	84.96%	66.82%
2	78.06%	66.38%	82.18%	84.78%	65.86%
3	77.73%	65.95%	82.05%	84.63%	66.08%
4	77.59%	65.36%	82.00%	84.52%	65.18%
5	77.27%	63.68%	81.53%	84.30%	64.28%
6	77.20%	64.45%	81.54%	84.35%	63.69%
7	76.61%	63.81%	81.51%	84.09%	62.96%
8	76.56%	62.43%	81.11%	83.96%	62.28%
9	76.28%	62.12%	81.25%	83.87%	62.47%
10	75.40%	61.07%	80.82%	83.82%	61.82%
11	75.93%	62.06%	80.63%	83.64%	62.40%

Table 5: Data transmission rate (Kbyte/s) versus throughput when the UDP traffic ($\rho = 0$) is generated.

Data transmission Rate (Kbytes/s)	Proposed scheme with $\{\mu_t, \mu_p, \mu_m\}$ = $\{1/3, 1/3, 1/3\}$	$\{0.8, 0.1, 0.1\}$	$\{0.1, 0.8, 0.1\}$	$\{0.1, 0.1, 0.8\}$	WFS	Cost-based scheme
20	4.67	4.67	4.96	4.57	4.64	4.50
40	9.39	9.19	10.01	9.28	9.14	9.16
60	14.37	13.97	14.84	13.82	13.80	13.73
80	18.46	18.64	19.56	18.53	18.35	18.52
100	23.45	23.53	25.04	22.96	22.48	22.89
120	28.18	27.80	30.39	27.70	27.17	27.66
140	32.95	32.92	35.18	32.12	32.06	32.36
160	37.07	37.91	39.00	36.52	37.36	37.83
180	42.52	42.12	45.11	40.65	41.47	41.55
200	44.84	45.97	49.89	45.57	45.63	45.12
220	51.98	51.54	55.86	50.73	50.43	51.16

values of the data transmission rate when $\rho = 0$. We run simulation tests for the proposed scheme with the four weight sets, as in the simulation for the TCP. We observe from Table 5 that the proposed and the cost-based schemes and WFS have similar throughput performances, because when ρ is 0, the amount of data throughput achieved by WWAN is same as that achieved by WLAN, as known from (16).

Figure 9 depicts the power consumption per minute for the different values of the data transmission rate when the

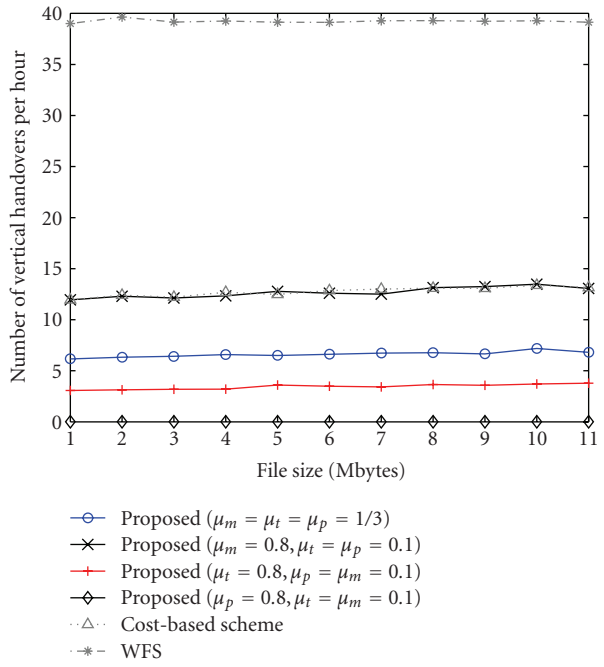

Figure 8: File size per session versus number of VHOs per hour when the TCP traffic ($\rho = 1$) is generated.

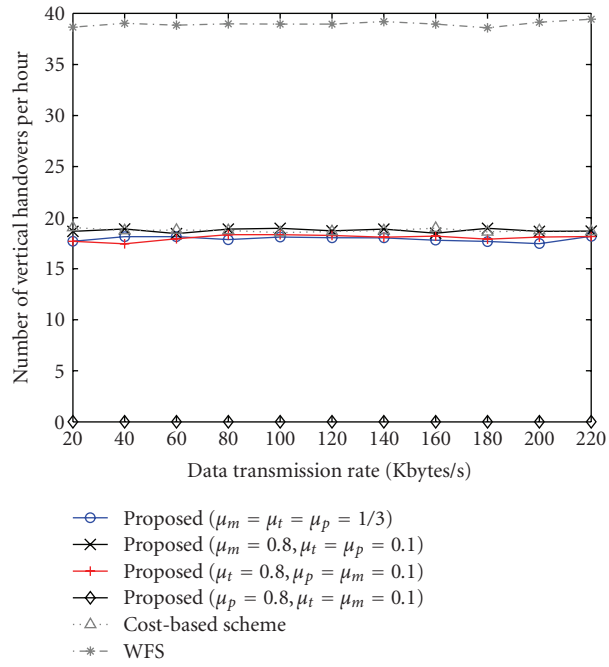

Figure 10: Data transmission rate versus number of VHOs per hour when the UDP traffic ($\rho = 0$) is generated.

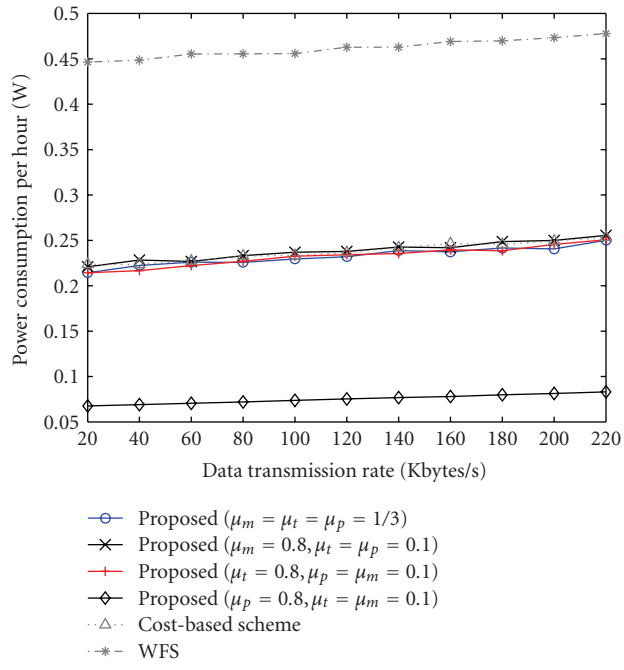

Figure 9: Data transmission rate versus power consumption per minute when the UDP traffic ($\rho = 0$) is generated.

UDP traffic ($\rho = 0$) is generated. We see from Figure 9 that our scheme with $\mu_p = 0.8$ achieves the greatest improvement over the WFS and the cost-based scheme in terms of the power consumption. Specifically, our scheme with $\mu_p = 0.8$ decreases the power consumption per minute 83.71% and 68.35% from those of WFS and the cost-based scheme, respectively, in average. For the other three weigh sets (i.e., $\mu_t = 0.8$, $\mu_m = 0.8$, and all the weights are 1/3), our scheme consumes 1.33% less power than the cost-based scheme in average, because less weight is put on the power cost, compared to the case of $\mu_p = 0.8$.

Figure 10 shows the number of VHOs per hour for the different values of the data transmission rate when the UDP traffic ($\rho = 0$) is generated. We see that WFS, in which the MN performs DVHO whenever it enters the double-coverage area, shows the greatest number of VHOs. The number of VHOs for the cost-based scheme is 2.8% more than our scheme with the three weight sets (excluding $\mu_p = 0.8$), in average. However, for our scheme with $\mu_p = 0.8$, no VHO is performed over the whole range of the data transmission rate in order to comply with the user's request and to save power as much as possible.

4. Conclusion

The coexistence of different wireless access technologies including WLAN and WWAN (e.g, UMTS and CDMA2000) has enabled for multi-interface MNs to roam freely from one network to another, being able to maintain their network connection and the QoS required by higher layer applications. In the integrated WWAN/WLAN network, the MNs should hand over from one network to another in order to obtain improved performance or at least to maintain a continuous wireless connection. Therefore, in this paper we proposed a cost-based vertical handover decision algorithm that operates based on a cost function for WWAN/WLAN integrated networks. In the proposed

scheme, when an MN moves to the double coverage area, DVHO is performed only if the cost expected in the WLAN is less than that in the WWAN. The cost function is based on data throughput, power consumption, and monetary cost. In addition, user preferences for these cost factors are considered to select an optimal network for each MN. Our simulation results showed that the proposed scheme outperforms the typical WFS scheme in terms of throughput and power consumption by selecting a lower cost network over all user preferences.

Acknowledgments

This work was supported in part by the IT R&D program of MKE/IITA (2009-F-043-01, Development of user-centric terminal-controlled seamless mobility technology) and in part by the Information Technology Research Center (ITRC) support program supervised by the MKE/IITA, South Korea (IITA-2009-C1090-0902-0005).

References

[1] N. Nasser, A. Hasswa, and H. Hassanein, "Handoffs in fourth generation heterogeneous networks," *IEEE Communications Magazine*, vol. 44, no. 10, pp. 96–103, 2006.

[2] Y. Kim, B. J. Jeong, J. Chung, et al., "Beyond 3G: vision, requirements, and enabling technologies," *IEEE Communications Magazine*, vol. 41, no. 3, pp. 120–124, 2003.

[3] IEEE 802.11, "Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications," IEEE Standard, 1999.

[4] 3GPP TR 23.234 v7.7.0, "3GPP System to WLAN Interworking; System Description (Release 7)," March 2008, http://www.3gpp.org/Specifications.

[5] 3GPP2, "CDMA2000 WLAN Interworking," S.R0087-A, v1.0, March 2006, http://www.3gpp2.org.

[6] K. Pahlavan, P. Krishnamurthy, A. Hatami, et al., "Handoff in hybrid mobile data networks," *IEEE Personal Communications*, vol. 7, no. 2, pp. 34–47, 2000.

[7] M. Lott, M. Siebert, S. Bonjour, D. von Hugo, and M. Weckerle, "Interworking of WLAN and 3G systems," *IEE Proceedings: Communications*, vol. 151, no. 5, pp. 507–513, 2004.

[8] M. Ylianttila, R. Pichna, J. Vallstrom, et al., "Handoff procedure for heterogeneous wireless networks," in *Proceedings of the IEEE Global Telecommunications Conference (GLOBECOM '99)*, vol. 5, pp. 2783–2787, December 1999.

[9] Q. Zhang, C. Guo, Z. Guo, and W. Zhu, "Efficient mobility management for vertical handoff between WWAN and WLAN," *IEEE Communications Magazine*, vol. 41, no. 11, pp. 102–108, 2003.

[10] A. Calvagna and G. Modica, "A cost-based approach to vertical handover policies between WiFi and GPRS," *Wireless Communications and Mobile Computing*, vol. 5, no. 6, pp. 603–617, 2005.

[11] W. Shen and Q.-A. Zeng, "A novel decision strategy of vertical handoff in overlay wireless networks," in *Proceedings of the 5th IEEE International Symposium on Network Computing and Applications (NCA '06)*, pp. 227–230, 2006.

[12] T. Al-Gizawi, K. Peppas, D. I. Axiotis, E. N. Protonotarios, and F. Lazarakis, "Interoperability criteria, mechanisms, and evaluation of system performance for transparently interoperating WLAN and UMTS-HSDPA networks," *IEEE Network*, vol. 19, no. 4, pp. 66–72, 2005.

[13] J. McNair and F. Zhu, "Vertical handoffs in fourth-generation multinetwork environments," *IEEE Wireless Communications*, vol. 11, no. 3, pp. 8–15, 2004.

[14] L.-J. Chen, T. Sun, B. Chen, V. Rajendran, and M. Gerla, "A smart decision model for vertical handoff," in *Proceedings of the ANWIRE International Workshop on Wireless Internet and Reconfigurability*, 2004.

[15] Q. Song and A. Jamalipour, "A time-adaptive vertical handoff decision scheme in wireless overlay networks," in *Proceedings of the IEEE International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC '06)*, 2006.

[16] R. Tawil, G. Pujolle, and O. Salazar, "A vertical handoff decision scheme in heterogeneous wireless systems," in *Proceedings of the IEEE Vehicular Technology Conference*, pp. 2626–2630, 2008.

[17] A. Kamerman and L. Monteban, "WaveLAN®-II: a high-performance wireless LAN for the unlicensed band," *Bell Labs Technical Journal*, vol. 2, no. 3, pp. 118–133, 1997.

[18] P. Chevillat, J. Jelitto, A. N. Barreto, and H. L. Truong, "A dynamic link adaptation algorithm for IEEE 802.11a Wireless LANs," in *IEEE International Conference on Communications (ICC '03)*, vol. 2, pp. 1141–1145, 2003.

[19] Q. Liu, S. Zhou, and G. B. Giannakis, "Cross-layer combining of adaptive modulation and coding with truncated ARQ over wireless links," *IEEE Transactions on Wireless Communications*, vol. 3, no. 5, pp. 1746–1755, 2004.

[20] 3GPP, "General Packet Radio Service(GPRS); Service Description Stage 2 (Release 7)," Technical Specifications 3GPP TS 23.060 v7.4.0, March 2007.

[21] H. Zhai, J. Wang, and Y. Fang, "Providing statistical QoS guarantee for voice over IP in the IEEE 802.11 wireless LANs," *IEEE Wireless Communications*, vol. 13, no. 1, pp. 36–43, 2006.

[22] Q. Ni, T. Li, T. Turletti, and Y. Xiao, "Saturation throughput analysis of error-prone 802.11 wireless networks," *Wireless Communications and Mobile Computing*, vol. 5, no. 8, pp. 945–956, 2005.

[23] Q. Pang, V. C. M. Leung, and S. C. Liew, "A rate adaptation algorithm for IEEE 802.11 WLANs based on MAC-layer loss differentiation," in *Proceedings of the 2nd International Conference on Broadband Networks (BroadNets '05)*, pp. 709–717, October 2005.

[24] IEEE 802.11a, "Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications: High-speed Physical Layer in the 5 GHz Band," IEEE Standard, 1999.

[25] C. Bettstetter, "Smooth is better than sharp: a random mobility model for simulation of wireless networks," in *Proceedings of the 4th ACM International Workshop on Modeling, Analysis and Simulation of Wireless and Mobile Systems (MSWiM '01)*, pp. 19–27, July 2001.

[26] W. Chung and S. Lee, "Cost-effective MAP selection in HMIPv6 networks," in *Procweedings of the IEEE International Conference on Communications (ICC '07)*, pp. 6026–6031, June 2007.