An MDP-based Vertical Handoff Decision Algorithm for Heterogeneous Wireless Networks

Enrique Stevens-Navarro, *Student Member, IEEE*, Yuxia Lin, *Student Member, IEEE*, and
Vincent W.S. Wong, *Senior Member, IEEE*

Abstract— **The architecture for the Beyond 3rd Generation (B3G) or 4th Generation (4G) wireless networks aims to integrate various heterogeneous wireless access networks. One of the major design issues is the support of vertical handoff. Vertical handoff occurs when a mobile terminal switches from one network to another (e.g., from WLAN to CDMA 1xRTT). The objective of this paper is to determine the conditions under which vertical handoff should be performed. The problem is formulated as a Markov decision process with the objective of maximizing the total expected reward per connection. The network resources utilized by the connection are captured by a link reward function. A signaling cost is used to model the signaling and processing load incurred on the network when vertical handoff is performed. The value iteration algorithm is used to compute a stationary deterministic policy. For the performance evaluation, voice and data applications are considered. The numerical results show that our proposed scheme performs better than other vertical handoff decision algorithms, namely: Simple Additive Weighting (SAW), Technique for Order Preference by Similarity to Ideal Solution (TOPSIS), and Grey Relational Analysis (GRA).**

Index Terms— **Vertical handoff, handoff decision, Markov decision processes, heterogeneous wireless networks, network selection.**

I. INTRODUCTION

The architecture for the Beyond 3rd Generation (B3G) or 4th Generation (4G) wireless networks aims to integrate various heterogeneous wireless access networks over an IP (Internet Protocol) backbone. Currently, there are various standardization bodies working towards this vision. Examples include the 3rd Generation Partnership Project (3GPP) [1], 3GPP2 [2], and the IEEE 802.21 Media Independent Handover (MIH) working group [3]. 3GPP and 3GPP2 have standardized the interconnection requirements between 3G wireless cellular systems and wireless local area networks (WLANs) to provide mobility support to users roaming between both systems. Several levels of integration have been proposed ranging from simple common billing and customer care to seamless mobility and session continuity [4], [5].

In order to provide seamless mobility, one of the design issues is the *vertical handoff* support. Vertical handoff occurs when connections switch from one network to another (e.g., from WLAN to Code Division Multiple Access 1x Radio Transmission Technology (CDMA 1xRTT)). It is different from conventional horizontal handoff in which the mobile terminals just move from one base station to another within the same access network. In the literature, vertical handoff is also referred to as *intersystem* handoff, while horizontal handoff is referred to as *intrasystem* handoff [6]. In B3G/4G wireless networks, mobile terminals are envisioned to be equipped with multiple interfaces to establish connections with different types of wireless access networks. Thus, seamless vertical handoff is an important network operation.

The vertical handoff process involves three main phases [7], [8], namely *system discovery*, *vertical handoff decision*, and *vertical handoff execution*. Different access networks can be *collocated* within the same coverage area. During the system discovery phase, the mobile terminal determines which networks can be used and what services are available in each network. These networks may also advertise the supported data rates and Quality of Service (QoS) parameters for different services. Since the users are mobile, the available collocated networks depend on the location of the user. The traffic load in each network may also change with time. Thus, this phase may be invoked periodically.

In the vertical handoff decision phase, the mobile terminal determines whether the connections should continue using the existing selected network or be switched to another network. The decision may depend on various parameters including the type of the application (e.g., conversational, streaming, interactive, background), minimum bandwidth and delay required by the application, access cost, transmit power, current battery status of the mobile terminal, and the user's preferences.

During the vertical handoff execution phase, the connections are re-routed from the existing network to another one in a seamless manner. This phase also includes the authentication, authorization, and the transfer of context information. Since the mobile terminal may still be communicating via the existing network while handoff execution takes place, this provides enough time for the network to perform the necessary functions while minimizing any service disruptions.

In this paper, the focus of our work is in the vertical handoff decision phase. To this end, we propose a vertical handoff decision algorithm for heterogeneous wireless networks. The problem is formulated as a Markov decision process (MDP). There is a link reward function associated with the QoS received by the mobile connection. There is also a signaling cost function associated with the signaling overhead and processing

Manuscript received August 28, 2006; revised May 16, 2007. This work was supported by Bell Canada, the Natural Sciences and Engineering Research Council (NSERC) of Canada, and the Programa de Mejoramiento del Profesorado (PROMEP) from Mexico. Part of this paper was presented at the *IEEE Wireless Communications and Networking Conference (WCNC'07)*, Hong Kong, China, March 2007. The review of this paper was coordinated by Prof. C.-J. Chang.

The authors are with the Department of Electrical and Computer Engineering, University of British Columbia, Vancouver, BC V6T 1Z4, Canada, e-mail: {enriques, yuxial, vincentw}@ece.ubc.ca.

load incurred when vertical handoff execution is performed. The objective is to determine the policy which maximizes the expected total reward per connection. A stationary policy is obtained when the connection termination time is geometrically distributed. The contributions of this paper are as follows:

1) The proposed model is adaptive and applicable to a wide range of conditions. Different link reward functions can be assigned to various applications and networks with different QoS requirements. Different signaling cost functions can be used based on the complexity of the re-routing operation and the signaling load incurred on the network.
2) We provide guidelines for the implementation of our proposed vertical handoff algorithm. All the parameters in our proposed algorithm can be obtained via the services provided by the handoff-enabling functions in the IEEE 802.21 standard.
3) Performance of our proposed algorithm is evaluated under two types of traffic: voice and data. Numerical results show good performance improvement of our proposed scheme over several vertical handoff decision algorithms, including SAW (Simple Additive Weighting) [9], TOPSIS (Technique for Order Preference by Similarity to Ideal Solution) [9], GRA (Grey Relational Analysis) [10], and two other heuristic policies.

The rest of the paper is organized as follows. The related work is summarized in Section II. The MDP model formulation is presented in Section III. The vertical handoff decision algorithm, the optimality equations, and the value iteration algorithm are described in Section IV. Implementation issues are discussed in Section V. Extensions of the vertical handoff algorithm to include other QoS parameters, user's preferences, and horizontal handoffs are described in Section VI. Numerical results, sensitivity analysis, and the structure of the policy are presented in Section VII. Conclusions and future work are given in Section VIII.

II. RELATED WORK

In this section, we provide an overview of recent work on vertical handoff decision in heterogeneous wireless networks.

A policy-enabled vertical handoff model is proposed in [11]. The model considers the preference of the user and the tradeoff between different characteristics of the networks (e.g., bandwidth, access cost, and power consumption). In [9], the vertical handoff decision is formulated as a fuzzy multiple attribute decision making (MADM) problem. Fuzzy logic is used to represent the imprecise information of some attributes of the networks and the preferences of the user. The fuzzy MADM method consists of two steps. The first step converts the fuzzy data into a real number. The second step uses classical MADM methods [12] to determine the ranking of the candidate networks. Two MADM ranking methods are proposed in [9]: Simple Additive Weighting (SAW) and Technique for Order Preference by Similarity to Ideal Solution (TOPSIS).

In [10], the network selection for vertical handoff is modeled by the Analytic Hierarchy Process (AHP) and the Grey Relational Analysis (GRA). AHP decomposes the network selection problem into several sub-problems and assigns a weight value to each sub-problem. Then, GRA is used to rank the candidate networks and selects the one with the highest ranking. AHP and GRA are also used for network selection in [13], where a mobile-controlled three-step vertical handoff prediction algorithm is proposed.

In [14], we investigate the performance among SAW, TOPSIS and GRA regarding the vertical handoff decision. Another MADM ranking algorithm called the Multiplicative Weighting Exponent (MEW) is also studied. The performance comparison considers different types of traffic and network parameters such as bandwidth, packet delay, jitter and bit error rate. In [15], the handoff decision mechanism is formulated as an optimization problem. Each candidate network is associated with a cost function, which depends on the bandwidth, delay, and power consumption. An appropriate weight factor is assigned to each parameter to account for its importance on the vertical handoff decision. An application oriented vertical handoff decision mechanism is proposed in [16]. Each candidate network is associated with a utility function. The selected network is the one which provides the highest utility value calculated from a weighted sum of the QoS parameters. Such parameters are provided by a location service server.

In [17], a framework is proposed to compare different vertical handoff algorithms. The framework includes a path loss channel model between the mobile terminal and the access point, and a Markov chain that models the user's movement between different access networks. A multi-layer framework for vertical handoff is proposed in [18]. A rules engine combined with several threshold parameters is used to monitor the decision parameters while the handoff policies are stored in a database. The framework allows the trigger of the vertical handoff by either changes in applications, variations of network's conditions, or preferences of the users.

In [19], a proactive end-to-end mobility management system is proposed for IEEE 802.11 WLANs and wireless wide area networks. The system relies on various components to support transparent mobility management and continuity of the connection among access networks. In [20], a utility-based strategy for network selection is proposed. Several utility functions are evaluated based on the economic concepts of consumer surplus and risk. In [21], the vertical handoff decision is evaluated via a handoff cost function and a handoff threshold function which can be adapted to changes in the network environment dynamically. In [22], a vertical handoff decision algorithm based on dynamic programming is introduced. It considers the movement and location information, which is provided by the location service server.

Although there have been various vertical handoff algorithms proposed in the literature, our work is motivated by two particular aspects. First, the connection duration needs to be taken into account during the vertical handoff decision. Second, the processing and signaling load during the vertical handoff execution also needs to be taken into consideration. Our work aims to incorporate these two aspects in the model formulation of vertical handoff decision.

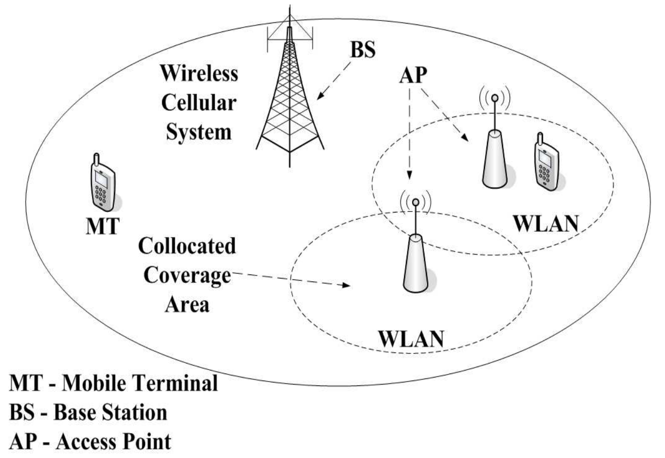


Fig. 1. Collocated heterogeneous wireless networks.

III. MODEL FORMULATION

Each mobile connection may experience a number of vertical handoffs during its connection lifetime. The envisioned heterogeneous wireless environment considered is shown in Fig. 1, where WLANs are collocated inside the coverage of a wireless cellular system. These specific areas are referred to as the *collocated coverage areas*. The mobile terminal is assumed to receive information from the collocated networks within its receiving range periodically. The advertised information from each network may include, among other parameters, the available bandwidth and the average delay. Details about how the information is obtained, and the system discovery phase will be discussed in Section V. In each time period, the mobile terminal *decides* whether the connection should use the current selected network or be re-routed to another network, which can provide better performance (e.g., lower cost, higher throughput). The re-routing of the connection during the vertical handoff execution phase is a complex process. It increases the processing and signaling load of the network. Thus, there is an important tradeoff between the QoS of the mobile connection, and the processing and signaling load incurred on the network.

We now describe how to formulate the above *vertical handoff decision problem* as a Markov decision process (MDP). The notations that we use follow those described in [23]. An MDP model consists of five elements: *decision epochs*, *states*, *actions*, *transition probabilities*, and *rewards*. The mobile terminal has to make a decision whenever a certain time period has elapsed. Referring to Fig. 2, the sequence $T = \{1, 2, \dots, N\}$ represents the times of successive decision epochs. The random variable N denotes the time that the connection terminates.

At each decision epoch, the mobile terminal has to decide whether the connection should use the current chosen network, or be re-routed to another network (i.e., execute a vertical handoff). Let M denote the total number of collocated networks in the coverage area of interest. The action set $A = \{1, 2, \dots, M\}$, and the random variable Y_t denotes the action chosen at decision epoch t.

The mobile terminal chooses an action based on its current *state* information. The state space is denoted by S. For each state $\mathbf{s} \in S$, the state information includes the address or identification number of the network that the mobile terminal is currently connected to, the available bandwidth and the average delay provided by all the available collocated networks in the area. The random variable X_t denotes the state at decision epoch t. Given that the current state is $\mathbf{s}$ and the chosen action is a, the *state transition probability function* for the next state $\mathbf{s}'$ is denoted by $P[\mathbf{s}' \mid \mathbf{s}, a]$. This function is Markovian since the state transition depends on the current state and action but not the previous states.

The *link reward function* $f(X_t, Y_t)$ reflects the QoS provided by the chosen network to the mobile connection within the time interval $(t, t+1)$. The *signaling cost function* $g(X_t, Y_t)$ captures the processing and signaling load incurred when the connection switches from one network to another. If the connection remains using the same network during the interval $(t, t+1)$, then $g(X_t, Y_t)$ is equal to zero. For convenience, we define the *reward function* as $r(X_t, Y_t) = f(X_t, Y_t) - g(X_t, Y_t)$.

A *decision rule* prescribes a procedure for action selection in each state at a specified decision epoch. Deterministic Markovian decision rules are functions $\delta_t : S \to A$, which specify the action choice when the system occupies state $\mathbf{s}$ at decision epoch t. A *policy* $\pi = (\delta_1, \delta_2, \dots, \delta_N)$ is a sequence of decision rules to be used at all decision epochs.

Let $v^\pi(\mathbf{s})$ denote the *expected total reward* between the first decision epoch till the connection termination, given that the policy π is used with initial state $\mathbf{s}$. We have,

$$v^\pi(\mathbf{s}) = E_{\mathbf{s}}^{\pi} \left[E_N \left\{ \sum_{t=1}^{N} r(X_t, Y_t) \right\} \right], \tag{1}$$

where $E_{\mathbf{s}}^{\pi}$ denotes the expectation with respect to policy π and initial state $\mathbf{s}$, and E_N denotes the expectation with respect to random variable N. Note that different policy π and initial state $\mathbf{s}$ will change the chosen action a. This will also cause a different state transition probability function $P[\mathbf{s}' \mid \mathbf{s}, a]$ to be used in the expectation $E_{\mathbf{s}}^{\pi}$. The random variable N, which denotes the *connection termination time*, is assumed to be geometrically distributed with mean $1/(1-\lambda)$. As shown in the Appendix, (1) can be written as:

$$v^\pi(\mathbf{s}) = E_{\mathbf{s}}^{\pi} \left\{ \sum_{t=1}^{\infty} \lambda^{t-1}\, r(X_t, Y_t) \right\}, \tag{2}$$

where λ can also be interpreted as the discount factor of the model, and $0 \le \lambda < 1$.

Since our optimization problem is to maximize the expected total discounted reward, we define a policy π^* to be *optimal* in Π if $v^{\pi^*}(\mathbf{s}) \ge v^\pi(\mathbf{s})$ for all $\pi \in \Pi$. A policy is said to be *stationary* if $\delta_t = \delta$ for all t. A stationary policy has the form $\pi = (\delta, \delta, \cdots)$; for convenience we denote π simply by δ. Our objective is to determine an *optimal stationary deterministic policy* δ^*, which maximizes the expected total discounted reward given by (2). For the rest of the paper, we refer to (2) as the *expected total reward*. We refer to δ^* as the MDP optimal policy. Note that δ^* is optimal under the expected total discounted reward optimality criterion [23].

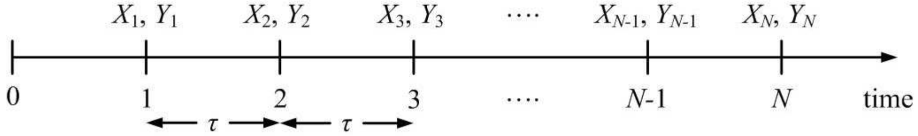


Fig. 2. Timing diagram of a Markov Decision Process (MDP).

IV. VERTICAL HANDOFF DECISION ALGORITHM

In this section, we begin by describing how the MDP model formulated in Section III can be used to analyze the vertical handoff decision algorithm. Then, the optimality equations and the value iteration algorithm are introduced.

A. State, Reward Function, and Transition Probability

In our proposed vertical handoff decision algorithm, the state space S is defined as:

$$S = \{1, 2, \cdots, M\} \times B^1 \times D^1 \times B^2 \times D^2 \cdots \times B^M \times D^M,$$

where M denotes the number of available collocated networks, B^m and D^m denote the set of the available bandwidth and delay from network m (with $m = 1, 2, \ldots, M$), respectively.

To reduce the number elements in the state space, we assume that the bandwidth information is provided in a *multiple* of units of bandwidth. Specifically, we define:

$$B^m = \{1, 2, 3, \ldots, b^m_{max}\}, \qquad m = 1, 2, \ldots, M,$$

where b^m_{max} denotes the maximum available bandwidth provided to a connection by network m. As an example, each unit of bandwidth in a WLAN and in a wireless cellular system can be 500 $kbps$ and 16 $kbps$, respectively.

Similarly, the delay information is also provided in a *multiple* of units of delay. That is,

$$D^m = \{1, 2, 3, \ldots, d^m_{max}\}, \qquad m = 1, 2, \ldots, M,$$

where d^m_{max} denotes the maximum delay provided to a connection by network m. As an example, each unit of delay in a WLAN and in a wireless cellular system can be 50 ms and 20 ms, respectively.

Given the current state s and the chosen action a, the link reward function $f(\mathbf{s}, a)$ is defined as:

$$f(\mathbf{s}, a) = \omega \; f_b(\mathbf{s}, a) + (1 - \omega) \; f_d(\mathbf{s}, a), \tag{3}$$

where $f_b(\mathbf{s}, a)$ denotes the bandwidth reward function, $f_d(\mathbf{s}, a)$ denotes the delay reward function, and ω is the *weight* (or *importance factor*) given to the available bandwidth with $0 \leq \omega \leq 1$.

Let $\mathbf{s} = [i, b_1, d_1, \ldots, b_M, d_M]$ denote the current state vector where i denotes the current network used by the connection. The bandwidth and delay reward functions are defined as follows:

$$f_b(\mathbf{s}, a) = \begin{cases} 1, & b_a \geq U_B, \\ (b_a - L_B)/(U_B - L_B), & L_B < b_a < U_B, \\ 0, & b_a \leq L_B, \end{cases} \tag{4}$$

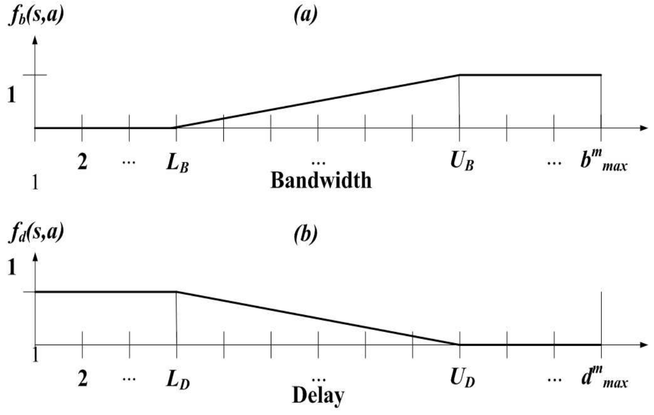


Fig. 3. Functions $f_b(\mathbf{s}, a)$ and $f_d(\mathbf{s}, a)$ from the link reward function.

and

$$f_d(\mathbf{s}, a) = \begin{cases} 1, & 0 < d_a \leq L_D, \\ (U_D - d_a)/(U_D - L_D), & L_D < d_a < U_D, \\ 0, & d_a \geq U_D, \end{cases} \tag{5}$$

where the constants L_B and U_B in (4) denote the minimum and maximum bandwidth required by the connection, respectively. On the other hand, the constants L_D and U_D in (5) denote the minimum and maximum delay required by the connection, respectively. The bandwidth reward function $f_b(\mathbf{s}, a)$ is shown in Fig. 3(a), and the delay reward function $f_d(\mathbf{s}, a)$ is shown in Fig. 3(b). Note that the available bandwidth is a utility parameter, while the delay is a cost parameter.

The signaling cost function $g(\mathbf{s}, a)$ is defined as:

$$g(\mathbf{s}, a) = \begin{cases} K_{i,a}, & i \neq a, \\ 0, & i = a, \end{cases} \tag{6}$$

where $K_{i,a}$ is the switching cost from the current network i to the new network a.

Between two successive vertical handoff decision epochs, the reward function $r(\mathbf{s}, a)$ can be defined as:

$$r(\mathbf{s}, a) = f(\mathbf{s}, a) - g(\mathbf{s}, a). \tag{7}$$

Finally, given that the current state $\mathbf{s} = [i, b_1, d_1, \ldots, b_M, d_M]$ and the chosen action is a, the probability function that the next state $\mathbf{s}' = [j, b_1', d_1', \ldots, b_M', d_M']$ is given by:

$$P[\mathbf{s}' \mid \mathbf{s}, a] = \begin{cases} \prod_{m=1}^{M} P[b_m', d_m' \mid b_m, d_m], & j = a, \\ 0, & j \neq a. \end{cases} \tag{8}$$

In (8), we assume that the joint bandwidth and delay probability function of each network is independent. This is due to the fact that although the networks are collocated in the same service area, they are managed by different network operators and use different wireless access technologies.

B. Optimality Equations and the Value Iteration Algorithm

Let $v(\mathbf{s})$ denote the maximum expected total reward given the initial state $\mathbf{s}$. That is,

$$v(\mathbf{s}) = \max_{\pi \in \Pi} \; v^{\pi}(\mathbf{s}). \tag{9}$$

From [23], the *optimality equations* are given by:

$$v(\mathbf{s}) = \max_{a \in A} \left\{ r(\mathbf{s}, a) + \sum_{\mathbf{s}' \in S} \lambda \; P[\mathbf{s}' \mid \mathbf{s}, a] \; v(\mathbf{s}') \right\}. \tag{10}$$

The solutions of the optimality equations correspond to the maximum expected total reward $v(\mathbf{s})$ and the MDP optimal policy $\delta^*(\mathbf{s})$. Note that the MDP optimal policy $\delta^*(\mathbf{s})$ indicates the decision as to which network to choose from given that the current state is $\mathbf{s}$.

There are various algorithms available to solve the optimization problem given by (10). Examples include the value iteration, policy iteration, and action elimination algorithms [23]. The following value iteration algorithm (VIA) determines a stationary deterministic optimal policy and the corresponding expected total reward.

Value Iteration Algorithm (VIA):

1) Set $v^0(\mathbf{s}) = 0$ for each state $\mathbf{s}$. Specify $\varepsilon > 0$ and set $k = 0$.
2) For each state $\mathbf{s}$, compute $v^{k+1}(\mathbf{s})$ by

$$v^{k+1}(\mathbf{s}) = \max_{a \in A} \left\{ r(\mathbf{s}, a) + \sum_{\mathbf{s}' \in S} \lambda \; P[\mathbf{s}' \mid \mathbf{s}, a] \; v^{k}(\mathbf{s}') \right\}.$$

3) If $\| v^{k+1} - v^{k} \| < \varepsilon(1-\lambda)/(2\lambda)$, go to step 4. Otherwise, increase k by 1 and return to step 2.
4) For each $\mathbf{s} \in S$, compute the stationary optimal policy

$$\delta(\mathbf{s}) = \arg \; \max_{a \in A} \left\{ r(\mathbf{s}, a) + \sum_{\mathbf{s}' \in S} \lambda \; P[\mathbf{s}' \mid \mathbf{s}, a] \; v^{k+1}(\mathbf{s}') \right\},$$

and stop.

There are a number of definitions for the norm function $\|.\|$. In this paper, the norm function is defined as $\|v\| = \max |v(\mathbf{s})|$ for $\mathbf{s} \in S$. Convergence of the VIA is ensured since the operation in step 2 corresponds to a contraction mapping. Thus, the function $v^k(\mathbf{s})$ converges in norm to $v(\mathbf{s})$. Note that the convergence rate of the VIA is linear. Additional details about the VIA are discussed in the following sections.

V. VERTICAL HANDOFF IMPLEMENTATION ISSUES

In this section, we first summarize the system discovery phase standardized within the IEEE 802.21 Media Independent Handover (MIH) working group. We then describe how the information provided by the 802.21 standard can be used to implement our proposed vertical handoff decision algorithm.

A. System Discovery Phase

In order to estimate the network conditions in heterogeneous wireless networks, the IEEE 802.21 MIH working group [3] is currently developing standards to enable efficient handoff operations, and to provide inter-operability between heterogeneous access networks. This includes both 802 and non-802 networks (e.g., 3GPP and 3GPP2). The proposed draft for the IEEE 802.21 standard [24] relies on a set of handoff-enabling functions within the mobility management protocol, and on a new network entity called the MIH Function (MIHF). The MIHF provides three services: Media Independent Event Service (MIES), Media Independent Command Service (MICS), and Media Independent Information Service (MIIS).

The MIES provides event classification, event filtering and event reporting corresponding to dynamic changes in link characteristics, status and quality. One event may indicate changes in state and transmission behavior of the physical, data link and logical link layers. The MICS provides a set of commands that enables MIH users to issue commands for handoff control and mobility. Finally, the MIIS provides the capability for obtaining the necessary information to make *effective handoff decisions*. Such information includes details on the characteristics and services provided by the serving and collocated networks in the area. The mobile terminal can access the relevant information via its current active network interface. Since the other interfaces do not need to be turned on simultaneously, the battery lifetime can be preserved.

The MIIS defines a set of information elements (IEs) to provide access to static/dynamic information and higher layer services supported by the networks. The IEs are classified into three specific groups. The first group gives an overview of the collocated networks within a coverage area. Such information can include the list of available networks and operators, roaming agreements, access costs, and security capabilities. The second group provides information about base stations and/or access points for each collocated network. The IE includes addressing information, location of the base stations and/or access points, supported data rates, and an extended set of link and QoS parameters (e.g., data rate, delay, jitter). It also includes higher layer services offered by the networks (e.g., IP multimedia such as Voice over IP). Finally, the last group includes other relevant information which is vendor specific.

B. Implementation Details

Based on the IEEE 802.21 standard [24], a mobile terminal which implements our proposed vertical handoff decision algorithm is able to periodically obtain information about the collocated networks in its receiving range by using its current network interface. The information provided by the MIIS of the MIH function is used to estimate the parameters of the link reward function in (3) and the signaling cost function in (6). The information about the available bandwidth and average delay in the networks can be estimated by following the standardized procedures for performance metrics for Internet services defined by the IETF IP Performance Metrics (IPPM) working group [25]. These procedures are designed in such a way that they can be implemented by network operators to provide accurate and unbiased quantitative measures of such metrics. Examples of such standardized metrics are: connectivity, packet delay and loss, packet delay variation, and link bandwidth capacity.

For the link reward function in (3), the mobile terminal can map the values of bandwidth and delay offered by network m into the bandwidth and delay units defined by B^m and D^m, respectively. Additionally, the maximum and minimum values required by (4) and (5) as well as the value of ω can be defined beforehand for each application. For the signaling cost function in (6), the value of $K_{i,a}$ is provided by the network operators and can be included within one of the IE fields. The value of $K_{i,a}$ can be set according to the agreements between two network operators. Suppose the current network is i and the chosen network a is j where $i \neq j$. A small value of $K_{i,j}$ will give an incentive to switch from network i to network j, whereas a large value of $K_{i,j}$ can indicate that there is no agreement between these two networks. Finally, the probability function in (8) can be estimated by the network operator based on its own network traffic statistics.

Given the values of the reward function and distributions of the networks, the VIA can be used to determine the MDP optimal policy $\delta^*(\mathbf{s})$. The VIA operates by calculating successive approximation to the value function $\nu(\mathbf{s})$. Each iteration of the VIA is performed in $O(|A||S|^2)$, where A is the action set and S the state space. Several variants of the VIA have been proposed to enhance the speed of convergence [23]. The VIA is widely used due to its conceptual simplicity, and to its ease in coding and implementation. Once the optimal policy is calculated, it can be stored in a matrix format. Each entry of the matrix specifies the optimal network to be used given the current state (i.e., bandwidth, delay), and switching cost of all collocated networks. The calculation of the optimal policy is performed by the operator off-line, and is updated periodically whenever spare processing capacity is available at the network access controller.

VI. MODEL EXTENSIONS

In this section, we extend the MDP model to include additional QoS parameters and also explain how the proposed vertical handoff scheme can handle user's preferences and be integrated into a complete handoff management framework for heterogeneous wireless networks.

A. Extension to Additional QoS Parameters

Suppose that the jitter information needs to be included as one of the QoS parameters. From Section IV-A, the state space S can be extended as:

$$S = \{1, 2, \ldots, M\} \times N^1 \times N^2 \cdots \times N^M,$$

where N^m represents the set of QoS parameters offered from network m for the connection, and

$$N^m = B^m \times D^m \times J^m, \qquad m = 1, 2, \ldots, M,$$

where B^m and D^m are the same as introduced in Section IV-A, and J^m denotes the set of jitter information from network m in a multiple of units of jitter. That is,

$$J^m = \{0, 1, 2, \ldots, j^m_{max}\}, \qquad m = 1, 2, \ldots, M,$$

where j^m_{max} denotes the maximum jitter provided to a connection by network m. As an example, each unit of jitter in a WLAN and in a wireless cellular system can be 10 ms and 5 ms, respectively.

Given the current state $\mathbf{s}$ and the chosen action a, the link reward function $f(\mathbf{s}, a)$ in (3) can be extended as:

$$f(\mathbf{s}, a) = \omega_b \ f_b(\mathbf{s}, a) + \omega_d \ f_d(\mathbf{s}, a) + \omega_j \ f_j(\mathbf{s}, a),$$

where $\sum_{k=\{b,\ d,\ j\}} \omega_k = 1$, and $f_j(\mathbf{s}, a)$ is the jitter reward function. The weight of the QoS parameters ω_k can be defined beforehand according to the specific requirements of the application. In general, N^m can be extended to include other relevant QoS parameters.

B. Extension to User's Preferences and Horizontal Handoff

Here, we propose a framework to integrate both horizontal handoff and vertical handoff with the user's preferences. First, we classify B^m, D^m, and J^m from network m as *network-based* QoS parameters, while parameters such as access cost, power consumption, and security denoted by C^m, Ψ^m, and Φ^m, respectively, as *user-based* QoS parameters.

Let us now consider a mobile terminal, which is connected to network i. We assume that horizontal handoff has priority over vertical handoff. The mobile terminal uses a conventional horizontal handoff scheme based on the received signal strength (RSS) from network i (e.g., [26]). Hence, if the horizontal handoff algorithm detects that the RSS is below a certain threshold, then the horizontal handoff is executed. On the other hand, if the mobile terminal discovers that it is within a collocated coverage area because of the information provided by the IEEE 802.21 MIIS, then a *screening phase* is invoked. This phase is capable of filtering networks not suitable for performing vertical handoff based on user-based QoS parameters such as C^m, Ψ^m, and Φ^m. This set of values may be specified by the user or stored in a user profile beforehand in the mobile terminal. Only suitable candidate networks will be considered for the vertical handoff decision. Towards the end, the vertical handoff decision is based on the MDP optimal policy $\delta^*(\mathbf{s})$ considering the network-based QoS parameters B^m, D^m and J^m. Finally, the handoff management framework can be summarized in algorithmic form as follows.

Horizontal and Vertical Handoff Decision Algorithm:

```
1)  While connecting to network $i$
2)    If RSS of network $i$ is below a certain threshold,
3)      Perform horizontal handoff.
4)    If the mobile terminal is within the collocated coverage area,
5)      Invoke screening phase based on $C^m$, $\Psi^m$, and $\Phi^m$.
6)      Determine vertical handoff decision policy $\delta^*(\mathbf{s})$ based on $B^m$, $D^m$, $J^m$
7)      If action $a \neq i$,
8)        Perform vertical handoff to network $a$.
9)      else
10)       Remain connected to network $i$.
```

TABLE I

PARAMETERS OF THE WIRELESS NETWORKS.

Notation	Parameter definition	Network $i = 1$	Network $i = 2$
b^i_{max}	Maximum available bandwidth in network i	25 units	10 units
d^i_{max}	Maximum delay in network i	8 units	8 units
$K_{1,2}$	Switching cost from network 1 to network 2	1	-
$K_{2,1}$	Switching cost from network 2 to network 1	-	1

TABLE II

PARAMETERS OF THE REWARD FUNCTIONS.

Notation	Parameter definition	CBR	FTP
L_B	Minimum available bandwidth required	2 units	2 units
U_B	Maximum available bandwidth required	4 units	16 units
L_D	Minimum delay required	2 units	7 units
U_D	Maximum delay required	7 units	8 units
ω	Weight factor	0.25	0.90

VII. NUMERICAL RESULTS AND DISCUSSIONS

In this section, we compare the performance between our proposed MDP-based vertical handoff decision algorithm, SAW [9], TOPSIS [9], and GRA [10]. We denote these policies as δ^*, δ^{SAW}, δ^{TOP}, and δ^{GRA}, respectively. In addition, two other heuristic policies are included. For the first heuristic, the network to be selected in each decision epoch is the one which has the highest available bandwidth. We denote this policy as δ^{BAN}. For the second heuristic, no vertical handoff is performed during the connection lifetime. We denote this policy as δ^{NEV}.

The performance metrics are the *expected total reward per connection*, and the *expected number of vertical handoffs per connection*. The expected total reward is defined in Section III. By following the notation introduced in Sections III and IV-A, let the state at time t be denoted as $\mathbf{s}_t = [i(t), b_1(t), d_1(t), \ldots, b_M(t), d_M(t)]$. The expected number of vertical handoffs per connection given policy δ with initial state $\mathbf{s}$ is given by:

$$\varsigma^{\delta}(\mathbf{s}) = E^{\delta}_{\mathbf{s}} \left\{ \sum_{t=1}^{\infty} \lambda^{t-1} \cdot \mathbf{1}[a_t \neq i(t)] \right\}, \tag{11}$$

where $\mathbf{1}[\cdot]$ denotes the indicator function (i.e., $\mathbf{1}[a_t \neq i(t)]$ is equal to 1 if a_t is not equal to $i(t)$ at time t, and is equal to 0 otherwise).

The average time between successive decision epochs is assumed to be 15 sec. We consider a scenario where there are two collocated networks (i.e., $M = 2$) as in the heterogeneous system depicted in Fig. 1. Network 1 is a WLAN and network 2 is a wireless cellular system. The parameters of the networks used in the numerical results are summarized in Table I. For simplicity, the switching costs in the signaling cost function in (6) are the same (i.e., $K_{1,2} = K_{2,1}$).

Two applications are considered. The first one is constant bit rate (CBR) voice traffic using the user datagram protocol (UDP) as the transport protocol. The second one is file transfer protocol (FTP) data traffic using the transmission control protocol (TCP). For these applications, the parameters of the link reward function in (3), as well as the bandwidth and delay reward functions in (4) and (5) are summarized in Table II.

The unit of bandwidth is equal to 16 $kbps$, and the unit of delay is equal to 60 ms. For the connections with CBR traffic, L_B and U_B are set to match voice coders and protocol overheads of IP multimedia services [27], and L_D and U_D are set to match the target delay lower than 150 ms. The connection is still acceptable if the delay is between 150 ms and 400 ms. The quality of the connection is not acceptable if the delay exceeds 400 ms. These ranges are defined according to ITU Recommendation G.114 [28]. For the connections with FTP traffic, L_B, U_B, L_D and U_D are set to match an elastic application without stringent delay requirements. Note that the importance factor ω in the link reward function in (3) is also set accordingly to each of the application's bandwidth requirements. Unless stated otherwise, the initial state at the beginning of the connections with CBR traffic is assumed to be 48 $kbps$ of bandwidth and 60 ms of delay, and 160 $kbps$ and 60 ms for the connections with FTP traffic.

For the state transition probability function of the wireless cellular system (i.e., network 2) in (8), we assume that the values of bandwidth and delay are guaranteed for the duration of the connection. Thus,

$$P[b_2^{'}, d_2^{'} \mid b_2, d_2] = \begin{cases} 1, & b_2^{'} = b_2,\, d_2^{'} = d_2, \\ 0, & \text{otherwise.} \end{cases} \tag{12}$$

For the state transition probability function of the WLAN, however, we cannot make the same assumption as in the wireless cellular system. Instead, we follow a simulation-based approach to estimate such probabilities. A typical IEEE 802.11b WLAN is simulated by using the ns-2 (v2.29) network simulator [29], where users arrive and depart from the network according to a Poisson process with an average rate of 0.2 users per second. The user's connections are either 64 $kbps$ CBR traffic using UDP, or FTP traffic using TCP. The basic rate and data rate of the WLAN are 1 $Mbps$ and 11 $Mbps$, respectively. The available bandwidth is calculated from the WLAN's capacity, which is approximated to be the achievable throughput by the WLAN under saturation, minus the aggregated traffic of the users. The values of the available bandwidth and delay are rounded according to the units defined in Section IV-A. The counting of transitions among states is performed to estimate the state transition probabilities.

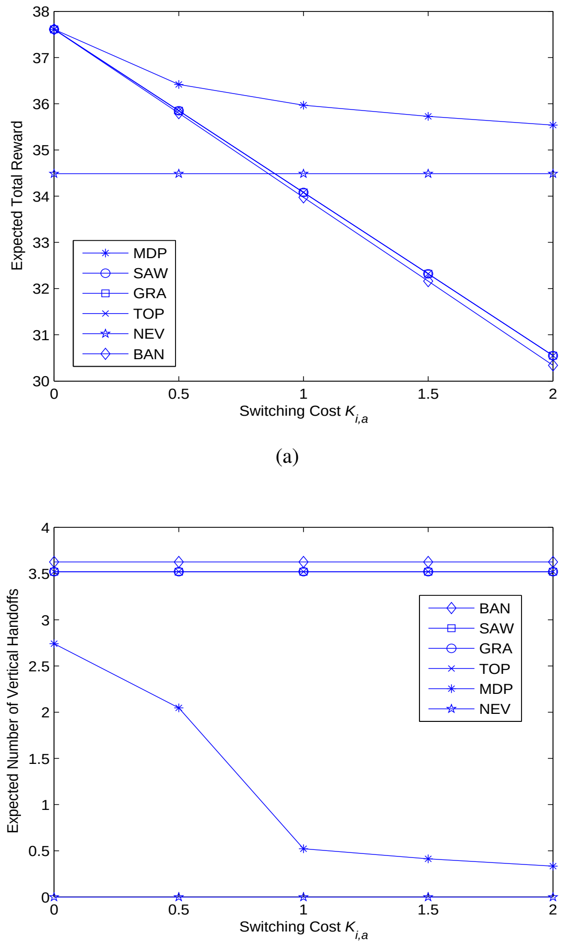
For the VIA, ε is chosen to be equal to 10^{-3}. From the MDP optimal policy, the VIA is used again to determine the expected number of vertical handoffs by solving (11). For the other vertical handoff decision algorithms SAW, TOPSIS and GRA, the ranking of each network is used to determine the policies δ^{SAW}, δ^{TOP} and δ^{GRA}. Given the reward functions and the state transition probabilities, the corresponding expected total reward and the expected number of vertical handoffs can be determined from (2) and (11), respectively. For the two heuristics policies δ^{BAN} and δ^{NEV}, the expected total reward and expected number of vertical handoffs can also be determined using the VIA.

A. Results for CBR Voice Traffic

Fig. 4 shows the expected total reward and the expected number of vertical handoffs versus the switching cost $K_{i,a}$ for connections with CBR traffic. The average connection duration is 10 min (i.e., $\lambda = 0.975$). Fig. 4(a) shows that when $K_{i,a}$ increases, the policy δ^* obtained from MDP algorithm gives the highest expected total reward per connection compared to the other algorithms. Fig. 4(b) shows that when $K_{i,a}$ increases, there is less incentive to perform vertical handoff and the MDP algorithm chooses not to switch more often. On the other hand, δ^{SAW}, δ^{TOP} and δ^{GRA} select the network based on the current available bandwidth and delay, and do not take the switching cost into consideration. δ^{BAN} only considers the available bandwidth and δ^{NEV} by definition never performs a vertical handoff. Thus, the expected number of vertical handoffs is constant for those algorithms. Recall that an increase in the number of vertical handoffs is directly related to an increase in the signaling load incurred in the networks to re-route the connections from one network to another.

The switching costs $K_{i,a}$ in the signaling cost function in (6) provide flexibility for the network operators. The values of $K_{i,a}$ can be selected to reflect the complexity of the re-routing operations, and the signaling load incurred on the network when vertical handoffs are performed at the network layer and above. As examples of such flexibility, small values of $K_{i,a}$ can be set among networks with roaming or interworking agreements that may facilitate or simplify the vertical handoffs, such as in the 3GPP/3GPP2-WLAN standardized architectures [4], [5]. Large values of $K_{i,a}$ can be set temporarily for overloaded networks as a load balancing technique in order to deter roaming users to connect to them and decrease the traffic load in the network.

Fig. 5 shows the expected total reward and the expected number of vertical handoffs versus the discount factor λ. Recall that in the discrete-time MDP model, we assume that the *time unit* (i.e., the time between successive decision epochs) is $1/4$ min. The average connection duration is equal to $1/(1-\lambda)$ time unit. When λ is varied from 0.9 to 0.983, it corresponds to the variation of the average connection duration from 2.5 min to 15 min. Fig. 5(a) shows that the policy δ^* from MDP gives the highest expected total reward per connection for all values of λ. As an example, when the average connection duration is 15 min, the MDP algorithm gives 4.4% more total expected reward than policies δ^{SAW}, δ^{TOP} and δ^{GRA}, 4.7% more than δ^{BAN} and 5% more than δ^{NEV}. Finally, when the average connection duration increases, the number of decision epochs and the difference in the total expected reward also increase. Thus, the expected number of vertical handoffs increases for all algorithms.



(a)

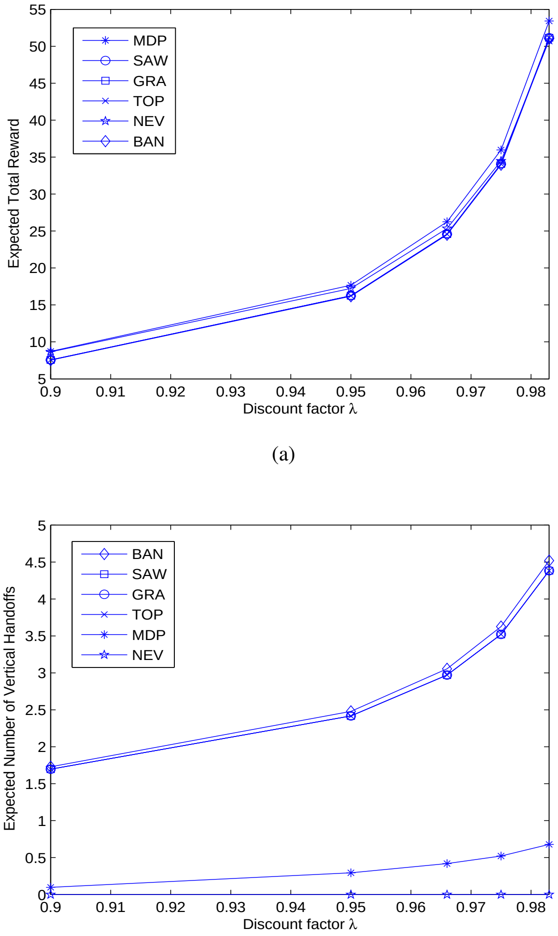
(b)

Fig. 4. Effect under different switching cost $K_{i,a}$ for CBR traffic. (a) Expected total reward. (b) Expected number of vertical handoffs. λ = 0.975 and ω = 0.25.

Fig. 6 shows the expected total reward and the expected number of vertical handoffs versus the weight factor ω for connections with CBR traffic. The policy δ^* from MDP gives the highest expected total reward per connection for all different values of ω. From (3), when ω increases, the bandwidth reward function (4) becomes more important than the delay reward function (5). As we can see, when $\omega \geq 0.25$, the total expected reward and the expected number of vertical handoffs of δ^{SAW}, δ^{TOP} and δ^{SAW} converge to δ^{BAN}. This is because more than 25% of the importance is placed on the available bandwidth.

B. Results for FTP Data Traffic

Fig. 7 shows the expected total reward and the expected number of vertical handoffs versus the switching cost $K_{i,a}$ for connections with FTP traffic. Fig. 7(a) shows that when



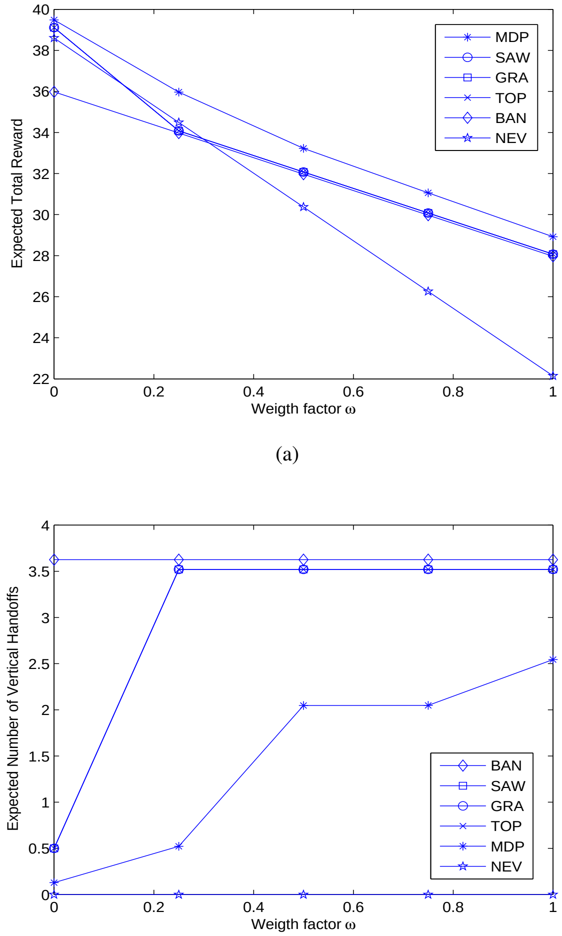
(a)

(b)

Fig. 5. Effect under different discount factor λ for CBR traffic. (a) Expected total reward. (b) Expected number of vertical handoffs. $K_{1,2}$ = $K_{2,1}$ = 1 and ω = 0.25.

$K_{i,a}$ increases, the policy δ^* obtained from the MDP algorithm gives the highest expected total reward per connection compared to the other algorithms. Note that in this case, the policy δ^* coincides with the heuristic policy δ^{BAN}. Recall that we set the parameters of the reward functions to match an elastic application, and this kind of application is suitable for the heuristic policy of selecting the network which offers the highest available bandwidth. A similar behavior of δ^* and δ^{BAN} is shown in Fig. 7(b) when $K_{i,a}$ increases.

Fig. 8 shows the expected total reward and the expected number of vertical handoffs versus the discount factor λ. Fig. 8(a) shows that the MDP policy δ^* gives the highest expected total reward per connection when the average connection duration is increased from from 2.5 min to 15 min, and also coincides with the heuristic policy δ^{BAN}. Finally, when the average connection duration increases, the number of decision epochs and the difference in the total expected reward increase as well. Note that the expected number of vertical handoffs increases slowly for all algorithms compared to the connections with CBR traffic. The reason is that the connections with FTP traffic tend to remain connected to network 1 whenever it is possible.



(a)

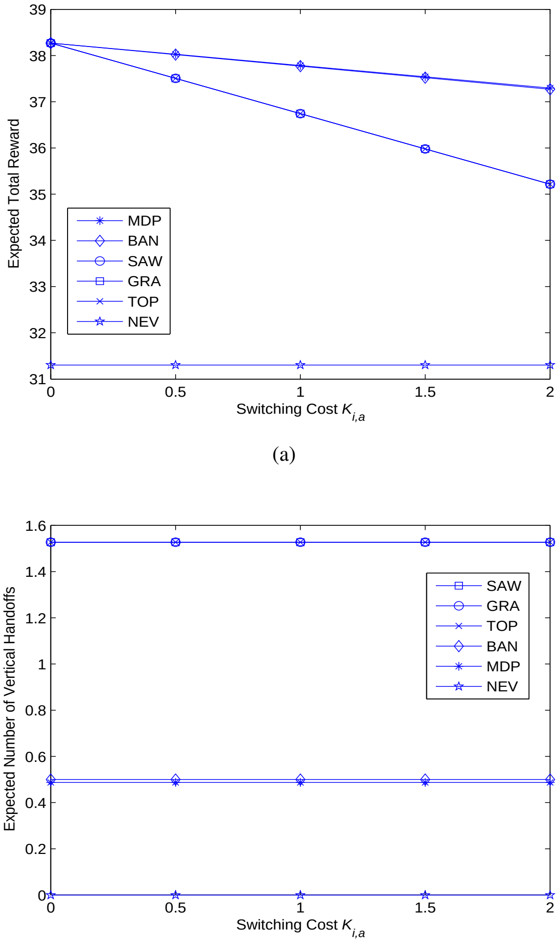
(b)

Fig. 6. Effect under different weight factor ω for CBR traffic. (a) Expected total reward. (b) Expected number of vertical handoffs. $K_{1,2}$ = $K_{2,1}$ = 1 and $\lambda = 0.975$.

C. Sensitivity Analysis

In order to calculate the maximum expected total reward of a connection, the MDP optimal policy δ^* needs to be determined. The optimal policy depends on different parameters. Although the parameters such as $K_{i,a}$ can be determined by the network, and parameters such as ω, L_B, L_D, U_B, U_D can be assigned by the user or set beforehand according to the application's requirements, the value of λ (i.e., average connection duration) may not always be estimated correctly by the mobile terminal during the connection setup. In that case, the policy may not indeed be the optimal one. In this section, we determine the percentage change of the expected total reward to the variation of the average connection duration. The procedures for the sensitivity analysis are as follows:

1) Given the actual average connection duration $c = (1-\lambda)^{-1}$ time units and other parameters, we first determine the maximum expected total reward, denoted as *Reward (optimal)*.
2) Let $\hat{c}$ denote the estimated average connection duration and Δ_c denote the percentage change of the average connection duration. These parameters are related by $\hat{c} = (1 + \Delta_c)c$. Based on the estimated average con-



(a)

(b)

Fig. 7. Effect under different switching cost $K_{i,a}$ for FTP traffic. (a) Expected total reward. (b) Expected number of vertical handoffs. λ = 0.975 and ω = 0.9.

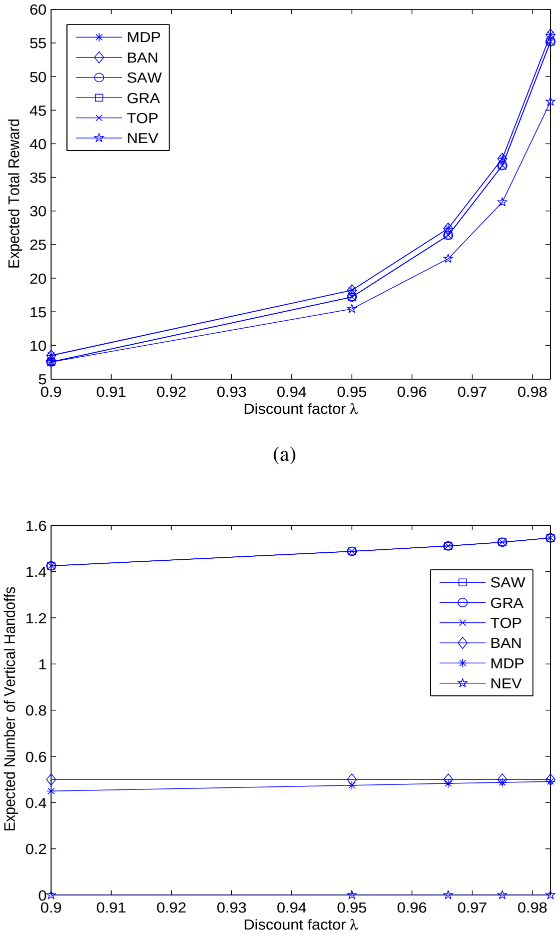
nection duration $\hat{c}$ and other parameters, the suboptimal policy is determined. The suboptimal expected total reward, denoted as *Reward (suboptimal)*, is calculated.

3) The change in the expected total reward with respect to the variation of the average connection duration is characterized by the reward ratio, which is defined as *Reward (suboptimal) / Reward (optimal)*.

The results for different λ are shown in Fig. 9 for CBR and FTP traffic. Within the $(-95, -40)$ percentage range, there is a small decrease in the reward ratio for both applications, being more noticeable for the FTP traffic with shorter connections. To overcome this decrease, the results imply that if there is uncertainty in the estimation of the connection duration, it may be better to overestimate the connection duration in order to maintain the reward ratio to be close to one.

D. Structure of the Optimal Policy

The MDP optimal policy $\delta^*(\mathbf{s})$ is numerically computed by implementing the VIA. Here, we provide relevant examples of the impact of the switching costs $K_{i,a}$ of the signaling cost function in (6) on the structure of the policy. We focus on the MDP policy $\delta^*(\mathbf{s})$ for connections with CBR traffic because the impact of the $K_{i,a}$ is more noticeable.



(a)

(b)

Fig. 8. Effect under different discount factor λ for FTP traffic. (a) Expected total reward. (b) Expected number of vertical handoffs. $K_{1,2}$ = $K_{2,1}$ = 1 and ω = 0.9.

Figs. 10(a) and (b) show the structure of the MDP optimal policy $\delta^*(\mathbf{s})$ for the subset of states $\mathbf{s} = [i = 2, b_1, d_1, b_2, d_2 = 4]$ with switching costs $K_{i,a} = 1$ and $K_{i,a} = 0$, respectively. Note that in this numerical example, the state space S is five-dimensional. In order to plot the MDP policy, two parameters are fixed to a specific value. We choose the current network in use (i) to be network 2 and the current delay in network 2 (d_2) to be 4 units. The average connection duration is 10 min (e.g., $\lambda = 0.975$), and the time between successive decision epoch is 15 sec. The cubes (bars) represent action $a = 1$ (i.e., perform vertical handoff to network 1) while the absence of cubes (bars) represents action $a = 2$ (i.e., remain using network 2). Note that when the switching costs change from 0 to 1, there are fewer states in which the action is to perform the vertical handoff. This decrease in the number of states is expected since, as shown in Fig. 4(b), the expected number of vertical handoffs per connection is reduced from 2.75 to 0.5 when the switching cost $K_{i,a}$ increases from 0 to 1.

E. Number of Iterations

The performance metrics, which are the expected total reward and the expected number of vertical handoffs per

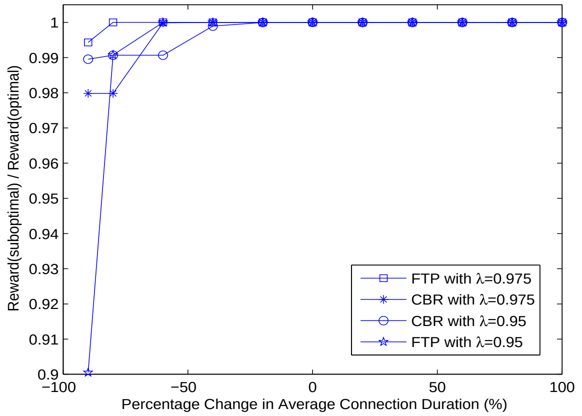


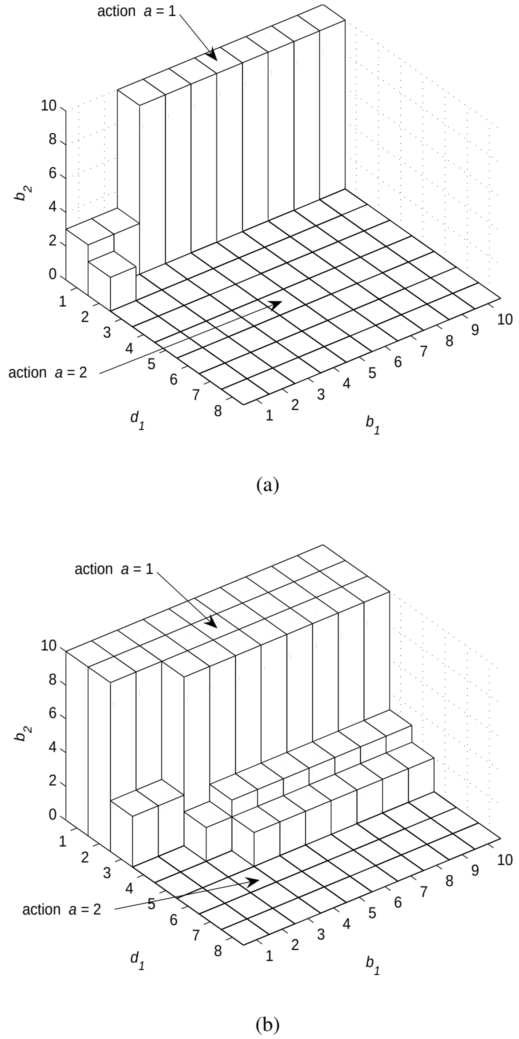
Fig. 9. Variation of the average connection duration with $K_{1,2} = K_{2,1} = 1$, $\omega = 0.25$ for CBR traffic, and $\omega = 0.90$ for FTP traffic.

connection, and the MDP policy δ^* presented in previous subsections were computed by using the VIA. The VIA proves to be a very efficient and stable iteration algorithm. The number of iterations required to converge from point to point is predictable. In general, the number of iterations neither depends on the switching costs $K_{i,a}$ of the signaling cost function (6), nor on the weight factor ω in the link reward function (3), but depends on the value of the discount factor λ (e.g., average connection duration). As an example, in Fig. 5(a), the VIA requires 45 iterations to converge when the average connection duration is 2.5 min (e.g., $\lambda = 0.9$), but it requires 183 iterations when the average connection duration is 10 min (e.g., $\lambda = 0.975$). As mentioned in Section V, variants of the VIA have been proposed to enhance convergence [23].

VIII. CONCLUSIONS

In this paper, we proposed a decision algorithm for the vertical handoff decision phase in heterogeneous wireless networks. The algorithm is based on MDP formulation with the objective of maximizing the expected total reward of a connection. A link reward function is used to model the QoS of the mobile connection. A signaling cost function is used to model the switching and re-routing operations when a vertical handoff occurs. Our work aims to illustrate the tradeoff between these two important aspects in the vertical handoff decision. A stationary deterministic policy is obtained when the connection termination time is geometrically distributed. We described the model extensions and the implementation guidelines, which are compatible with the IEEE 802.21 standard. For performance evaluation, we considered CBR voice traffic and FTP data traffic. Results show that our proposed MDP algorithm gives a higher expected total reward and lower expected number of vertical handoffs per connection than SAW, TOPSIS, GRA, and two heuristic policies under a wide range of conditions.

For future work, we plan to extend the proposed model to a continuous-time model formulation and consider general connection duration. In the discrete-time formulation, the connection duration is considered to be geometrically distributed. In addition, the current MDP model assumes that the vertical handoff decision is performed periodically. A more realistic model is to allow the decision to be performed whenever there is a state change in the networks. Other techniques to solve MDPs such as reinforcement learning will also be considered.



(a)

(b)

Fig. 10. Structure of the MDP optimal policy ($d_2 = 4$ units and the user is currently connected to network 2). (a) $K_{1,2} = K_{2,1} = 1$. (b) $K_{1,2} = K_{2,1} = 0$.

APPENDIX

In this section, we derive the expression $v^\pi(\mathbf{s})$, the expected total reward of the connection given a policy π and initial state $\mathbf{s}$. Assume the random variable N follows a geometric distribution with mean $1/(1-\lambda)$. That is,

$$P(N = n) = \lambda^{n-1}\,(1-\lambda), \qquad n = 1, 2, 3, \cdots$$

In this case, (1) can be written as:

$$v^\pi(\mathbf{s}) = E_{\mathbf{s}}^{\pi}\left\{\sum_{n=1}^{\infty}\sum_{t=1}^{n} r(X_t, Y_t)\,\lambda^{n-1}\,(1-\lambda)\right\}.$$

Since $\sum_{n=1}^{\infty}\sum_{t=1}^{n} = \sum_{t=1}^{\infty}\sum_{n=t}^{\infty}$, by interchanging the order of the summation, we have

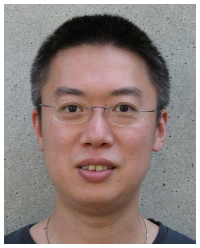
$$\begin{aligned} v^{\pi}(\mathbf{s}) &= E_{\mathbf{s}}^{\pi}\left\{\sum_{t=1}^{\infty}\sum_{n=t}^{\infty} r(X_t,Y_t)\,\lambda^{n-1}\,(1-\lambda)\right\} \\ &= E_{\mathbf{s}}^{\pi}\left\{\sum_{t=1}^{\infty}\lambda^{t-1}\,r(X_t,Y_t)\right\}. \qquad \square \end{aligned}$$

REFERENCES

[1] 3rd Generation Partnership Project (3GPP), http://www.3gpp.org/.
[2] 3rd Generation Partnership Project 2 (3GPP2), http://www.3gpp2.org/.
[3] IEEE 802.21 Media Independent Handover Working Group, http://www.ieee802.org/21/.
[4] 3GPP, "Requirements on 3GPP system to Wireless Local Area Network (WLAN) interworking," TS 22.234 (v8.0.0), March 2007.
[5] 3GPP2, "3GPP2-WLAN interworking," S.R0087-A (v1.0), February 2006.
[6] I. Akyildiz, J. Xie, and S. Mohanty, "A Survey of Mobility Management in Next-Generation All-IP-Based Wireless Systems," *IEEE Wireless Communications*, vol. 11, no. 4, pp. 16–28, August 2004.
[7] J. McNair and F. Zhu, "Vertical Handoffs in Fourth-generation Multi-network Environments," *IEEE Wireless Communications*, vol. 11, no. 3, pp. 8–15, June 2004.
[8] W. Chen, J. Liu, and H. Huang, "An Adaptive Scheme for Vertical Handoff in Wireless Overlay Networks," in *Proc. of ICPADS'04*, Newport Beach, CA, July 2004.
[9] W. Zhang, "Handover Decision Using Fuzzy MADM in Heterogeneous Networks," in *Proc. of IEEE WCNC'04*, Atlanta, GA, March 2004.
[10] Q. Song and A. Jamalipour, "A Network Selection Mechanism for Next Generation Networks," in *Proc. of IEEE ICC'05*, Seoul, Korea, May 2005.
[11] H. Wang, R. Katz, and J. Giese, "Policy-Enabled Handoffs Across Heterogeneous Wireless Networks," in *Proc. of IEEE Workshop on Mobile Computing Systems and Applications (WMCSA'99)*, New Orleans, LA, February 1999.
[12] K. Yoon and C. Hwang, *Multiple Attribute Decision Making: An Introduction*. Sage Publications, 1995.
[13] M. R. Kibria, A. Jamalipour, and V. Mirchandani, "A Location Aware Three-Step Vertical Handoff Scheme for 4G/B3G Networks," in *Proc. of IEEE GLOBECOM'05*, St. Louis, MO, November 2005.
[14] E. Stevens-Navarro and V. Wong, "Comparison Between Vertical Handoff Decision Algorithms for Heterogeneous Wireless Networks," in *Proc. of IEEE VTC'06-Spring*, Melbourne, Australia, May 2006.
[15] F. Zhu and J. MacNair, "Optimizations for Vertical Handoff Decision Algorithms," in *Proc. of IEEE WCNC'04*, Atlanta, GA, March 2004.
[16] W. Chen and Y. Shu, "Active Application Oriented Vertical Handoff in Next Generation Wireless Networks," in *Proc. of IEEE WCNC'05*, New Orleans, LA, March 2005.
[17] A. Zahran and B. Liang, "Performance Evaluation Framework for Vertical Handoff Algorithms in Heterogeneous Networks," in *Proc. of IEEE ICC'05*, Seoul, Korea, May 2005.
[18] A. Sur and D. Sicker, "Multi Layer Rules based Framework for Vertical Handoff," in *Proc. of BROADNETS'05*, Boston, MA, October 2005.
[19] 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, June 2004.
[20] O. Ormond, J. Murphy, and G. Muntean, "Utility-based Intelligent Network Selection in Beyond 3G Systems," in *Proc. of IEEE ICC'06*, Istanbul, Turkey, June 2006.
[21] A. Hassawa, N. Nasser, and H. Hassanein, "Tramcar: A Context-aware Cross-Layer Architecture for Next Generation Heterogeneous Wireless Networks," in *Proc. of IEEE ICC'06*, Istanbul, Turkey, June 2006.
[22] J. Zhang, H. C. Chan, and V. Leung, "A Location-based Vertical Handoff Decision Algorithm for Heterogeneous Mobile Networks," in *Proc. of IEEE GLOBECOM'06*, San Francisco, CA, November 2006.
[23] M. Puterman, *Markov Decision Processes: Discrete Stochastic Dynamic Programming*. John Wiley and Sons, 1994.
[24] IEEE, "Draft IEEE Standard for Local and Metropolitan Area Networks: Media Independent Handover Services," P802.21/D04.0, February 2007.
[25] IETF, "IP Performance Metrics (IPPM) Working Group," http://www.ietf.org/html.charters/ippm-charter.html.
[26] G. Pollini, "Trends in Handover Design," *IEEE Communications Magazine*, vol. 34, no. 3, pp. 82–90, March 1996.
[27] G. Camarillo and M. Garcia-Martin, *The 3G IP Multimedia Subsystem*. John Wiley and Sons, 2004.
[28] ITU-T, "One-way Transmision Time," Rec. G.114, May 2003.
[29] The Network Simulator - ns-2, http://www.isi.edu/nsnam/ns.

Enrique Stevens-Navarro received the B.Sc. and M.Sc. degrees in electrical engineering from Universidad Autonoma de San Luis Potosi (UASLP), San Luis Potosi, Mexico in 2000, and from Instituto Tecnologico y de Estudios Superiores de Monterrey (ITESM), Monterrey, Mexico in 2002, respectively. From 2002 - 2003, he was a project manager at Q-Voz IVR outsourcing in Monterrey, Mexico. In 2003, he joined the UASLP as a Lecturer with the Faculty of Science. He is currently working towards the Ph.D. degree in the Department of Electrical and Computer Engineering at the University of British Columbia, Vancouver, BC, Canada. His research interests are in mobility and resource management for heterogeneous wireless networks.

Yuxia Lin received his B.Sc. and M.Sc. degrees in applied physics and electrical engineering from Tsinghua University, Beijing, China, in 2000 and 2003, respectively. He is currently a Ph.D. candidate in the Department of Electrical and Computer Engineering at the University of British Columbia, Vancouver, BC. He has been granted the UBC Graduate Fellowship and the NSERC Canada Graduate Scholarship. His research interests are in the area of wireless networking with an emphasis on medium access control, cooperative communications and quality-of-service provisioning.



Vincent W.S. Wong received the B.Sc. degree from the University of Manitoba, Winnipeg, MB, Canada, in 1994, the M.A.Sc. degree from the University of Waterloo, Waterloo, ON, Canada, in 1996, and the Ph.D. degree from the University of British Columbia (UBC), Vancouver, BC, Canada, in 2000. From 2000 to 2001, he worked as a systems engineer at PMC-Sierra Inc. He is currently an Associate Professor in the Department of Electrical and Computer Engineering at UBC. His research interests are in resource and mobility management for wireless mesh networks, wireless sensor networks, and heterogeneous wireless networks. Dr. Wong is an associate editor of the IEEE Transactions on Vehicular Technology. He serves as TPC member in various conferences, including the IEEE International Conference on Communications (ICC) and Globecom. He is a senior member of the IEEE and a member of the ACM.