A Hybrid (N/M)CHO Soft/Hard Vertical Handover Technique for Heterogeneous Wireless Networks

Anna Maria Vegni, Enrico Natalizio

A Hybrid (N/M)CHO Soft/Hard Vertical Handover Technique for Heterogeneous Wireless Networks

Anna Maria Vegni*, and Enrico Natalizio†

* Department of Engineering, COMLAB Communication Lab., University of Roma Tre, Rome, Italy

† Heudiasyc UMR CNRS 7253, Université de Technologie de Compiègne, Compiègne, France

*annamaria.vegni@uniroma3.it, †enrico.natalizio@hds.utc.fr

Abstract

In this paper we investigate the potentiality and the benefits of a soft Vertical Handover (VHO) mechanism, compared with the traditional hard approach. More specifically, we present an analytical scheme for seamless service continuity in a heterogeneous network environment, modeled by means of a multi-dimension Markov chain. The call blocking probabilities, as well as the soft and hard vertical handover probabilities, are computed for specific networks (*i.e.*, UMTS and WLAN).

We propose a soft/hard VHO technique working either as (*i*) a Mobile Controlled Handover, on the basis of a reward and cost model, which consider the data rate and the bandwidth allocation, or (*ii*) as a simple Network Controlled Handover scheme, by assuming a probabilistic approach as the handover decision metric. Simulation results validate the benefits of the proposed handover algorithm when operating in soft mode, which outperforms the traditional hard approach in terms of network performance and limitation of unwanted and unnecessary handovers. Finally, the effectiveness of the proposed approach is proven with respect to other single and multi-parameter VHO techniques, by extensive simulations.

Index Terms

Heterogeneous wireless networks, mobility, handover, Markov chains.

I. INTRODUCTION

Heterogeneous Wireless Networks (HWNs) represent the new scenario of Next Generation Network (NGN) architecture, where different technologies, such as GSM (Global System for Mobile Communications), GPRS (General Packet Radio Service), UMTS (Universal Mobile Telecommunications System), WLAN (Wireless Local Area Network), WiMAX (Worldwide Interoperability for Microwave Access), LTE (Long Term Evolution) etc., co-exist and offer an overlapped wireless coverage [1].

Currently, wireless mobile networks and devices are becoming increasingly popular to provide user seamless Internet access, anytime and anywhere. This has led to the concept of *nomadic computing*, which involves portable devices (such as smartphones, laptop and handheld computers) providing Internet access to users connecting from their home or office networks. Furthermore, multimedia services requirements encompass not only large bandwidth communications, but also *on-the-move* facilities. NGN communication systems aim at providing seamless mobility support to access heterogeneous wired and wireless networks [2], [3].

NGNs are based on the cellular approach where the area is covered by cells that overlap each other. User services available in such scenario are several, such as entertainment, video on demands, and also emergency and safety assistance. In all these cases, connection quality provision between Mobile Terminals (MTs) is the main factor of network performance and represents a challenging issue [4]. In NGN scenarios, high-quality services can be guaranteed through the interoperability and interworking with NGN-compliant networks, by (*i*) managing heterogeneous wireless networks to service continuity support, and (*ii*) preserving the connectivity to mobile users moving across different radio access technologies with no noticeable performance degradation [4].

The use of reconfigurable devices, equipped with several network interfaces, is an open issue that has the objective of facilitating coordination among major co-existing mobile and wireless access systems. Currently, as no single wireless communication technology can provide a high performance to mobile users, networking in NGNs gives a solution to this limitation.

Seamless service continuity can be guaranteed by the *handover* process, which represents the switching of the connection from a Serving Network (SN) to a Candidate Network (CN), needed when a user moves [1], [5]. In NGNs,

both intra and inter-technology handovers take place [4]. The first one reflects the traditional *Horizontal Handover* (HHO) process, in which the MT hands-over between two neighbouring cells of the same access technology [6]. On the other hand, inter-technology handover, referred as *Vertical Handover* (VHO), occurs when the MT moves between different access technologies, *e.g.* from WLAN to UMTS, and vice versa [7]. Hence, the HHO is a symmetric process because the MT switches from a cell to another one of the same access technology [6], while VHO is asymmetric as the MT moves between different networks.

The benefits of vertical handover mechanisms are well-known in literature, and several surveys have investigated them [5], [8], [9], [10]: user connectivity is preserved by switching technology, and service quality is improved due to a reduced number of disconnections. However, vertical handovers are typically performed through *hard* mechanisms, where the connectivity between mobile users and the serving network is broken before the connection with a new network is established (namely, "*break-before-make*"). In this vision, the *soft* vertical handover mechanism (namely, "*make-before-break*") represents a viable solution to improve seamless connectivity [5], [8], [9].

Typically handover procedures fall into two main categories, (*i*) the Network Controlled Handover (NCHO) and (*ii*) the Mobile Controlled Handover (MCHO), depending on whether a handover is initiated and controlled by the network or by the MT, respectively [8], [11]. In horizontal handover management, MCHO is the most common case, especially for WLAN environments, while NCHO is generally the preferred choice for cellular networks where resource optimization and load management are centralized. To the best of our knowledge no scheme currently exists that is able to use both approaches.

In this paper, we propose a novel VHO algorithm, relaying on both MCHO and NCHO mode and making hard/soft vertical handover decisions. Our approach has been tested in WLAN and UMTS heterogeneous networks, in order to prove the benefits of Soft Vertical Handover (SVHO) as opposed to Hard Vertical Handover (HVHO). In this way, service continuity in NGNs can be assured by a soft network switching, while assuring high network performance as well as low handover occurrences.

The contribution of this paper can be summarized as follows:

1) We propose a mathematical analysis of both hard and soft vertical handover schemes based on a six-dimension (6-D) Markov chain model;
2) We define a *reward* and *cost* model based on the data rate and the bandwidth allocated to the incoming calls. The model can be used by mobile users working in MCHO mode to decide whether perform a hard/soft handover or not;
3) We define a *soft/hard handover decision technique*, which can work by using either the NCHO or the MCHO approach, depending on the decision metric used. Handover decisions made on the basis of hard/soft vertical handover probability are well suitable for NCHO, while a handover decision based on the *gain function* is a viable solution for MCHO.

We first compare network performance by evaluating (*i*) the handover gain, expressed in terms of user bandwidth, as well as (*ii*) the handover frequency, in two cases, such as with and without the possibility for the MT to perform a SVHO. Then, we simulate the performance results of the proposed handover scheme, which aims at maximizing the network performance (*i.e.*, throughput and cumulative received bits), and limiting the unnecessary and unwanted vertical handovers. Finally, a comparison to other techniques (*i.e.*, single and multi-parameter based vertical handover algorithms) is carried out, in order to highlight the benefits of our approach. The Markov chain model is exploited in our technique by using both the HVHO/SVHO occurrence probabilities as well as the gains associated to the execution of a HVHO/SVHO. Acting in both NCHO and MCHO modes, as well as in hard and soft approach, the proposed handover model results as a more efficient technique, with respect to traditional approaches, which perform only hard or soft handover.

The paper is organized as follows. In Section II we investigate related works on vertical handover schemes, specifically those techniques exploiting Markov chains. Section III describes the system model for an overlay network architecture. The proposed 6-D Markov chain is illustrated in Section IV; both hard and soft vertical handover cases are analysed by a probabilistic approach and an analytical model. We also describe how a soft and hard vertical handover decision problem can be modeled as a Markov decision process. In Section V the benefit and penalty functions are defined, in order to compare hard and soft VHO cases in terms of *bandwidth gain*. The proposed hard/soft vertical handover algorithm is described in Section VI, while in Section VII we define our simulation overlay heterogeneous network environment, and present the obtained results. As expected, from the comparison between hard and soft mode, we show that system that use soft handover perform better than systems using the hard approach. Finally, conclusions are drawn in Section VIII.

II. Related Work

In this section we first recall the handover concept and different types of handover decisions. Then, we classify recent research works on handover management, and position our technique in the existing literature.

Usually, in HWN environments continuous service is achieved by supporting VHO mechanism from one network to another. Handover is the process of switching channel (*i.e.* frequency, time slot, spreading code, and other medium access techniques) associated with the current connection, while a call is in progress [5]. Handover schemes are classified into two broad categories *i.e.*, (*i*) hard and (*ii*) soft handovers. In hard handover, current resources are released before new resources are used, while in soft handover, for a period of time the mobile terminal has at least two active interfaces for data transmission, such as both existing and new resources are used during the handover process.

There are numerous methods for performing handover, relaying on centralized or decentralized decision-making processes of handover (*i.e.*, the handover decision may be made at the MT or network). From the decision process point of view, we have (*i*) Network-Controlled Handover (NCHO), (*ii*) Mobile-Controlled Handover (MCHO) and (*iii*) Mobile-Assisted Handover (MAHO) [5], [8]. In a NCHO approach, the serving network makes a handover decision based on the measurements of the MTs. This approach is used in first-generation analog systems, like AMPS (Advanced Mobile Phone System) and TACS (Total Access Communication System). In a MCHO approach, each MT is completely in control of the handover process, through measurements of the signal strengths and interference levels from neighboring networks. Eventually, in a MAHO process, the MT takes the measurements and the network makes the decision. This approach is typically used in GSM networks.

The standardization process of the handover mechanism is driven by the need of having a media independent handover, working with different technologies. The IEEE 802.21 working group is involved in, [12]:

> *"Developing standards to enable handover and interoperability between heterogeneous network types including both 802 and non 802 networks."*

The IEEE 802.21 standard provides quick handover of data sessions with small switching delays and minimized latency. As a result, the handover process becomes more flexible and appropriate, seamless and media independent. It considers both wired and wireless technologies, not only belonging to the IEEE 802 family (such as 802.3, 802.11, 802.16, and so on), but also to cellular systems, such as 3GPP, and 3GPP2. The IEEE 802.21 introduces the Media Independent Handover (MIH) functional model, which enhances the handover process across heterogeneous media. It is intended to introduce a new protocol layer located between the Network Layer (Layer 3) and the interface-specific lower layers *i.e.*, Medium Access Control (MAC) and physical layers (PHY) in the case of IEEE interfaces, Radio Resource Control (RRC) and Link Access Control (LAC) in the case of 3GPP or 3GPP2 interfaces, respectively.

Many researchers have addressed solutions based on the IEEE 802.21 standard. For example, the use of Mobile IP (MIP) together with MIH has been largely explored, in the attempt to improve the handover management [13], [14], [15], [16], [17], [18]. Particularly, in [14], Benoubira *et al.* investigate the issue of simultaneous connections to be supported in Mobile IPv6 network scenarios. In the framework of IEEE 802.21, the authors introduce a novel entity that communicates with Layer 3 and the MIH, and implements MIP functionalities. In [15], Song *et al.* use the Proxy Mobile IPv6 into IEEE 802.21 environment, and analogously, in [16] Tamijetchelvy and Sivaradje present a fast vertical handover MIP, which works in heterogeneous WLAN/WiMAX networks. They showed that this approach provides an efficient solution for mobility, multihoming, and routing in integrating heterogeneous networks. Finally, in [18] a QoS-based vertical handover mechanism involving WiMAX and WiFi networks is illustrated. Based on the Signal-to-Interference and Noise Ratio (SINR), the proposed technique adopts the IEEE 802.21 standard in order to assist in the handover decisions.

Apart from the recent IEEE 802.21 standard, novel techniques for seamlessly switching from a serving network to a candidate network have been proposed, especially in order to minimize the handover latency and enhance network performance. For example, vertical handovers issues have been investigated through radio resource management techniques to minimize unnecessary (and unwanted) handovers [19], [20], [21]. At the same time, forced vertical handovers may be imposed on certain groups of users to free up resources in the network, so as to admit more suitable users. In [22], Farbod and Liang investigate networking issue in NGNs as a solution for Call Admission Control issue, and call dropping probabilities are considered as system performance parameters. The admission control for handover drops is a key resource management technique in wireless networks. Lee *et al.* in [23] consider a Markov chain scheme, based on intra- and inter-domain for the analysis of equilibrium states and transient states, in order to model a simple adaptive bandwidth reservation mechanism.

In [24], the authors focus on the factors that trigger the network to execute a hard vertical handover, and a new scheme that takes into account different normalized factors is presented and evaluated. Inzerilli *et al.* [25] propose a multiparameter algorithm for VHO that considers also power and location information, resulting in a remarkable network performance and in the limitation of unnecessary handovers. Basically, multi-parameter approaches are very effective with respect to single-parameter techniques, since the decision making process is optimized by combining networking information [26]. In [27], the authors use some services of the IEEE 802.21 standard, together with geolocation, map information, surround context information and route calculation, in order to improve the handover performance.

In [10], a recent survey on vertical handover decision algorithms in 4G HWNs classifies 12 different works in 4 groups, depending on the parameter the VHO decision is based on *i.e.*, power level, bandwidth, cost function, and on the combination of these factors. Even though this classification takes into account exclusively works on HVHO among heterogeneous networks, it is also useful to classify our scheme of SVHO. In fact, when the NCHO is running, by following the previous classification, our scheme would be considered as a bandwidth-based algorithm. Regarding the characterization of the VHO, in [28] Fang *et al.* deal with VHO issue in HWNs as a two-state Markov model. Also in [29], a Markov model is introduced in order to analyze vertical handovers in overlay networks.

The use of Markov chains has been also investigated in several issues, such as regarding how to select a candidate network to hand over. In [30], Ying *et al.* propose two novel weighted Markov chain approaches based on candidate networks rank aggregation, so that the favorite network is selected as top one of rank aggregation result, fused from multiple ranking lists based on several decision criteria.

Another important factor, which has been analyzed through a Markov chain, is the user mobility pattern. Mobility significantly affects the performance of a vertical handover scheme, resulting in variations of handover occurrences. This topic has been discussed in [31], where the statistics of the user mobility pattern are analyzed to determine main parameters in a Markov chain for modeling the handover delay. Finally, the use of Markov chains is also applied for channel allocation mechanisms. In [32], Xiao and Kim propose a new channel allocation policy for satellite networks, and the use of a Markov chain with hysteresis control allows to calculate the new call blocking probability, as well as handover blocking probability and the resource utilization. However, this model is quite simple and is limited to two arrival calls in addition to the full state probabilities. Also, in [33] Wang *et al.* present a handover scheme for dynamically management of the channels reserved for handover calls, based on the current status of the handover queue. The proposed scheme consists of a three-dimensional Markov model in order to analyze network performance and investigate the desirable tradeoff.

Till this point, we have cited works that do not take into consideration the effects of the *soft* VHO in a heterogeneous networks environment. Actually, the benefits introduced by considering the SVHO are well known, but no work exists in literature that mathematically characterizes SVHO and HVHO in HWNs and evaluate the performance of the network in a general case. In fact, in [34] soft handover issue is just considered by traditional horizontal scheme, where a MT is connected to two or more WLAN cells. In [35], Li *et al.* propose an approach for SVHO based on location information and they compare their scheme with pre-existent Mobile IP technique, by handoff latency and packet loss. Finally, in [36] Ali and Pierre evaluate the impact of SVHO on the performance of a specific service (*i.e.*, the voice admission control) and propose a new algorithm based on their analytical model.

The main novelty of this work resides in the design and simulation of a vertical handover algorithm, which, in respect of existing techniques, is more:

- *Flexible*: it can be driven by either network requirements (NCHO mode) or user requirements (MCHO mode);
- *Complete*: it can perform both hard and soft vertical handover.

We characterize the SVHO and HVHO impact on blocking and handover probabilities for generic services by a 6-D Markov chain, and then we present new *reward* and *cost* factors in the SVHO-HVHO selection. A new scheme of soft-hard handover selection, based on both the mathematical model, and the reward and cost parameters, is presented. Finally, it is worth observing that our solution is expected to fit appropriately into the IEEE 802.21 framework. As shown in the next sections, the proposed technique aims at improving handover initiation and preparation, which is one of the main purposes of the IEEE 802.21 standard.

III. SYSTEM MODEL

In this section, we introduce the system model for a heterogeneous network environment, where we assume that both hard and soft handovers can occur.

Wireless and cellular communication systems are generally based on the assumption that mobile users can be served by the base station (or access point) that provides the best link quality [37]. At many locations *overlapping*
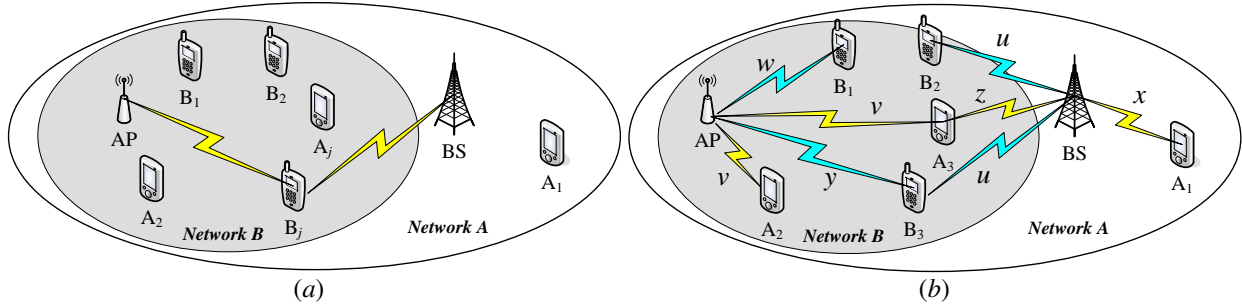
Fig. 1. Overlay heterogeneous network scenario consisting of network A (*e.g.*, UMTS) and B (*e.g.*, WLAN). (*a*) User B_j is in SVHO call with both the networks; (*b*) $Type_{(1,2)}$ hard and soft handover calls, as well as default calls. Yellow and light blue connections are for $Type_1$ and $Type_2$ service calls, respectively.

coverage usually occurs by nearby base stations. In such scenarios, a mobile user can establish a communication link of acceptable quality with more than one base station. A multiple connection decreases the blocking probability and the handover failure probability that the user can experience.

In this work, we consider an area where two networks *i.e.*, A and B, operate simultaneously, as in Fig. 1 (*a*). Without loss of generality, we consider A as a UMTS network and B as a WLAN. In the example in the figure, the two networks do not have the same coverage area, but MTs in the coverage area of B can be connected to A as well. Thus, the network consisting of the overlap of A and B is an overlay heterogeneous network. For sake of simplicity, we assume that A and B provide only two services, $Type_1$ and $Type_2$.

The system model is shown in Fig. 1 (*a*), where users A_i (*i.e.*, $i = 1, 2, \ldots, n$) are in the coverage area of network A, while users B_j (*i.e.*, $j = 1, 2, \ldots, k$ with $k < n$) are in the coverage area of network B and A as well. Obviously, the following scheme can be extended to more complicated scenarios, by increasing the number of networks and service types, as is typical in NGNs.

Both the networks have a limited number of resources, which can be frequency channels, time slots, scrambling codes or a combination of them to provide users with the requested service. By using the same terminology of [1], we will refer to the network that a call is initially connected to as the Serving Network (SN), and we will refer to the other network as the Candidate Network (CN). New calls are managed on a first-come-first-served policy by the network they attempt to connect to. When the number of calls increases and the traffic demand exceeds the capacity provided by the SN, the call blocking probability increases too. The excess of traffic may be served by the CN, if the user who generated the traffic is inside the coverage area of the CN and if there are any resources available in the CN. If no resources are available in the CN, a call block occurs. We assume that the different services can access the resources of the CN when it is needed. The need to switch an existing call from the SN to the CN can be justified by several reasons, such as signal loss, movement of the user, request for higher quality or a cheaper connection, etc. In all these cases, in our system model a VHO will occur from the SN to the CN. When the connection of a call with the CN is created and the link to the SN is broken, we talk of Hard Vertical Handover; instead, when the creation of the link with the CN does not cause the removal of the link with the SN, and the two links co-exist and provide the user with the requested service, we talk of Soft Vertical Handover (SVHO), as in the case illustrated by terminal B_j in Fig. 1 (*a*).

IV. MARKOV CHAINS MODEL

This section presents our proposed Markov chain, which describes how hard and soft vertical handovers occur in a heterogeneous network environment. This model will be exploited by the decision process of the proposed handover technique (see Section VI). More precisely, the Markov chains model provides information on (*i*) the probabilities of performing either a HVHO or a SVHO, as well as on (*ii*) the gains associated with the execution of a HVHO or a SVHO. The former information will be used by our technique when performing a NCHO, whereas the latter will be useful for a MCHO to check if the user requirements are satisfied.

In order to better understand the following explanation the reader can refer to Fig. 1. Let us assume networks A and B have N_A and N_B units of resources, respectively. U_1 and U_2 (with $U_2 > U_1$) denote the units of resources required by a $Type_1$ and $Type_2$ service, respectively. Without loss of generality, we assume that the main service provided by network A is $Type_1$, and the main service of network B is $Type_2$. Hence, calls of $Type_1$ ($Type_2$)

TABLE I
LIST OF PARAMETERS USED IN THE ANALYTICAL MODEL.

Parameter	Description
$N_{A,B}$	Network A and B
$Type_{1,2}$	Type of service 1 and 2
$U_{1,2}$	Units of resources for service Type 1 and 2
u	Number of $Type_2$ VHO calls to network A
v	Number of $Type_1$ VHO calls to network B
w	Number of $Type_2$ calls in network B
x	Number of $Type_1$ calls in network A
y	Number of $Type_2$ VHO calls to network B
z	Number of $Type_1$ VHO calls to network A
$\lambda_{1,2}$	Mean arrival rates of $Type_{1,2}$ calls
$\mu_{1,2}$	Mean service rates of $Type_{1,2}$ calls

will be first routed by network A (B) [1]. Furthermore, we assume that the same service consumes the same units of resources when provided by either of the networks [4], [38]. The possible states, denoted as (u, v, w, x, y, z), are modeled using a 6-D Markov chain. Specifically:

1) u is the number of $Type_2$ VHO calls to network A;
2) v is the number of $Type_1$ VHO calls to network B;
3) w is the number of $Type_2$ calls in network B;
4) x is the number of $Type_1$ calls in network A;
5) y is the number of $Type_2$ VHO calls to network B;
6) z is the number of $Type_1$ VHO calls to network A.

The variables u and v indicate the number of calls that are currently experiencing a traditional *hard vertical handover* [5]. In fact they are served by the CN, which is, respectively, network A for u and network B for v. The calls which are connected to the SN are split into two subsets (please, refer to Fig. 1 (*b*)):

(*i*) The calls currently connected to the SN only. These calls are indicated with w for network B, and x for network A;

(*ii*) The calls currently connected to the SN while they are also connected to the CN. These calls are indicated with y for network B and z for network A.

By using variables y and z we introduce the *soft vertical handover*, because these variables take into account the connection of mobile terminals to their SN while they are also connected to the CN. In our model, when we increase (decrease) variables y and z, we also increase (decrease) variables u and v, respectively [2]. This is well described in Fig. 1 (*b*), where the mobile terminal B_3 (A_3) is connected with both the networks in a soft VHO call, which is characterized by variable y (v) and u (z).

As we can see from the variables definition of our system, $Type_2$ ($Type_1$) service calls towards network A (B) are considered only by variable u (v). This means that a $Type_2$ ($Type_1$) call handled by network A (B) is considered in HVHO by variable u (v) when no other connection is established, while in SVHO it is considered by variables u (v) and y (z) when the call is also connected to network B (A). As an example, the state $(1, 2, 3, 3, 2, 1)$ represents one $Type_2$ HVHO call, three $Type_1$ calls and one $Type_1$ SVHO call in network A, as well as two $Type_1$ HVHO calls, three $Type_2$ calls and two $Type_2$ SVHO calls in network B. Fig. 1 (*b*) depicts the overlay heterogenous network scenario, where all previous calls are shown. As a reminder, all the previous parameters have been collected in TABLE I.

Now we can introduce the following definitions:

Definition 4-1. ***A block call*** *(for a particular service) occurs when none of the networks can provide resources to fulfill the request of an incoming call (of that service).*

---

[1]In case there are no resources available in network A (B), the call will be routed to network B (A) and considered by our system as a VHO call.

[2]This will be shown in TABLE II, Transition $T_5 - T_6$ for the increase, and $T_{15} - T_{16}$ for the decrease.

Definition 4-2. ***A vertical handover call*** *(for a particular service) occurs when a call (of that service) is switched from the SN to the CN with or without breaking the connection with the SN.*

We assume a memoryless mobility pattern for the users. The arrivals of $Type_1$ and $Type_2$ calls are assumed to follow a memoryless Poisson process with mean rates λ_1 and λ_2, respectively. The call holding time (service time) follows a negative exponential distribution with mean rates μ_1 and μ_2, resulting in an exponential channel holding times [39]. This is valid since we also assume the cell residence time is exponentially distributed, that typically occurs in NGNs, especially in WLAN. Moreover, we consider that the vertical handover execution requires a negligible delay and that all blocked calls are cleared.

At each state transition, only one type of service can be admitted or terminated. The served calls may increase or decrease according to twenty possible transitions, as enlisted in TABLE II.

TABLE II: List of transitions in the 6-D Markov chain.

Transition	Description
T_1	Increase in the number of $Type_1$ calls being served by network A i.e., $(0,0,w,x,0,0) \to (0,0,w,x+1,0,0)$ with rate λ_1 for $0 \le x \le x_{max} - 1$ and $0 \le w \le w_{max}$, where $x_{max} = N_A/U_1$, and $w_{max} = N_B/U_2$.
T_2	Increase in the number of $Type_2$ calls being served by network B i.e., $(0,0,w,x,0,0) \to (0,0,w+1,x,0,0)$ with rate λ_2 for $0 \le x \le x_{max}$ and $0 \le w \le w_{max} - 1$, where $w_{max} = N_B/U_2$ and $x_{max} = N_A/U_1$.
T_3	Increase in the number of $Type_1$ calls being served by network B i.e., $(0,v,w,x_{max},0,0) \to (0,v+1,w,x_{max},0,0)$ with rate λ_1 for $x = x_{max}$ and $(v+1)U_1 + wU_2 \le N_B$.
T_4	Increase in the number of $Type_2$ calls being served by network A i.e., $(u,0,w_{max},x,0,0) \to (u+1,0,w_{max},x,0,0)$ with rate λ_2 for $w = w_{max}$ and $(u+1)U_2 + xU_1 \le N_A$.
T_5	Increase in the number of $Type_1$ VHO calls being served by network A i.e., $(u,v,w,x,y,z) \to (u,v+1,w,x,y,z+1)$ with rate λ_1 for $(x+1)U_1 + wU_2 \le N_A$ and $(v+1)U_1 + wU_2 \le N_B$.
T_6	Increase in the number of $Type_2$ VHO calls being served by network B i.e., $(u,v,w,x,y,z) \to (u+1,v,w,x,y+1,z)$ with rate λ_2 for $(w+1)U_2 + xU_1 \le N_B$ and $xU_1 + (u+1)U_2 \le N_A$.
T_7	Decrease in the number of $Type_1$ VHO calls being served by network A i.e., $(u,v,w,x,y,z) \to (u,v,w,x,y,z-1)$ with rate $z\mu_1$ for $(x+1)U_1 + wU_2 \ge N_A$ and $(w+1)U_2 + vU_1 < N_B$.
T_8	Decrease in the number of $Type_2$ VHO calls being served by network B i.e., $(u,v,w,x,y,z) \to (u,v,w,x,y-1,z)$ with rate $y\mu_2$ for $vU_1 + (w+1)U_2 \ge N_B$ and $(x+1)U_1 + uU_2 < N_A$.
T_9	Increase in the number of $Type_1$ calls being served by network A i.e., $(u,v,w,x,y,z) \to (u,v,w,x+1,y,z)$ with rate λ_1 for $(x+1)U_1 + wU_2 \le N_A$ and $vU_1 + wU_2 \le N_B$.
T_{10}	Increase in the number of $Type_2$ calls being served by network B i.e., $(u,v,w,x,y,z) \to (u,v,w+1,x,y,z)$ with rate λ_2 for $(w+1)U_2 + vU_2 \le N_B$ and $zU_1 + uU_2 \le N_A$.
T_{11}	Decrease in the number of $Type_1$ calls being served by network A i.e., $(0,0,w,x,0,0) \to (0,0,w,x-1,0,0)$ with rate $x\mu_1$ for $1 \le x \le x_{max}$ and $0 \le w \le w_{max}$.
T_{12}	Decrease in the number of $Type_2$ calls being served by network B i.e., $(0,0,w,x,0,0) \to (0,0,w-1,x,0,0)$ with rate $w\mu_2$ for $1 \le w \le w_{max}$ and $0 \le x \le x_{max}$.
T_{13}	Decrease in the number of $Type_1$ VHO calls being served by network B i.e., $(0,v,w,x,0,0) \to (0,v-1,w,x,0,0)$ with rate $v\mu_1$ for $0 \le x \le x_{max}$ and $xU_1 + yU_2 \le N_B$ and $v \ge 1$.
T_{14}	Decrease in the number of $Type_2$ VHO calls being served by network A i.e., $(u,0,w,x,0,0) \to (u-1,0,w,x,0,0)$ with rate $u\mu_2$ for $0 \le w \le w_{max}$ and $(u+1)U_2 + zU_1 \le N_A$ and $u \ge 1$.

TABLE II: List of transitions in the 6-D Markov chain.

Transition	Description
T_{15}	Decrease in the number of $Type_1$ VHO calls being served by network A and B *i.e.*, $(u,v,w,x,y,z) \to (u,v-1,w,x,y,z-1)$ with rate $(v+z)\mu_1$ for $(x+1)U_1 + wU_2 > N_A$ and $(v+1)U_1 + yU_2 > N_B$.
T_{16}	Decrease in the number of $Type_2$ VHO calls being served by network A and B *i.e.*, $(u,v,w,x,y,z) \to (u-1,v,w,x,y-1,z)$ with rate $(u+y)\mu_2$ for $(w+1)U_2 + xU_1 > N_B$ and $zU_1 + (u+1)U_2 > N_A$.
T_{17}	Decrease in the number of $Type_1$ VHO calls being served by network B *i.e.*, $(u,v,w,x,y,z) \to (u,v-1,w,x,y,z)$ with rate $v\mu_1$ for $(w+1)U_2 + xU_1 > N_B$ and $(x+1)U_1 + wU_2 \le N_A$.
T_{18}	Decrease in the number of $Type_2$ VHO calls being served by network A *i.e.*, $(u,v,w,x,y,z) \to (u-1,v,w,x,y,z)$ with rate $u\mu_2$ for $(x+1)U_1 + wU_2 > N_A$ and $(w+1)U_2 + xU_1 \le N_B$.
T_{19}	Decrease in the number of $Type_1$ calls being served by network A *i.e.*, $(u,v,w,x,y,z) \to (u,v,w,x-1,y,z)$ with rate $x\mu_1$ for $(x+1)U_1 + wU_2 > N_A$ and $xU_1 + yU_2 \le N_B$.
T_{20}	Decrease in the number of $Type_2$ calls being served by network B *i.e.*, $(u,v,w,x,y,z) \to (u,v,w-1,x,y,z)$ with rate $w\mu_2$ for $(w+1)U_2 + xU_2 > N_B$ and $zU_1 + wU_2 \le N_A$.

Hard and soft VHO conditions are expressed by transitions T_3 and T_4, T_5 and T_6, respectively. SN connection down together with CN connection up is a condition expressed by transition T_7 and T_8, for network A and B as SN and CN, and vice versa respectively. The condition about SN and CN connection down is expressed by transitions T_{15} and T_{16}, for $Type_1$ and $Type_2$ vertical handover calls in network A and B, respectively. Then, transitions T_{17} and T_{18} represent respectively condition about CN connection down together with SN connection maintained, for $Type_1$ and $Type_2$ vertical handover calls. The missing transitions will be explained later in this section by means of an example. Finally, the sum of all steady-state probabilities must equal to one:

$$\sum_{h=0}^{u_{max}} \sum_{i=0}^{v_{max}} \sum_{j=0}^{w_{max}} \sum_{k=0}^{x_{max}} \sum_{l=0}^{y_{max}} \sum_{m=0}^{z_{max}} P(h,i,j,k,l,m) = 1, \tag{1}$$

where $w_{max} = \lfloor N_B/U_2 \rfloor$ and $x_{max} = \lfloor N_A/U_1 \rfloor$ are the largest integer value less than (or equals to) N_B/U_2 and N_A/U_1, such that $jU_2 + iU_1 \le N_B$ and $kU_1 + hU_2 \le N_A$, respectively. It follows $u_{max} = v_{max} = y_{max} = z_{max} = \min\{w_{max}, x_{max}\}$.

A block call for $Type_1$ service occurs when both network A and B cannot provide resources to allocate the new call, because this would cause the traffic to exceed the capacity of the networks. The $Type_1$ service blocking probability is:

$$P_{Block}^{Type_1} = \Pr\left[(uU_2 + (\min\{x+1,z+1\})U_1 > N_A) \cap (wU_2 + (v+1)U_1 > N_B)\right] | \lambda_1, \tag{2}$$

where the first condition considers the capacity of network A when the call is considered as either a new call (*i.e.*, $x \to x+1$) or a SVHO call (*i.e.*, $z \to z+1$), and the second condition considers the capacity of network B when the call is considered as a VHO call (*i.e.*, $v \to v+1$).

A hard vertical handover call for $Type_1$ service occurs when network A cannot provide resources to allocate the call, because this would cause the traffic to exceed its capacity, whereas network B can handle the call. The $Type_1$ HVHO probability is:

$$P_{HVHO}^{Type_1} = \Pr\left[(uU_2 + (\min\{x+1,z+1\})U_1 > N_A) \cap (wU_2 + (v+1)U_1 \le N_B)\right] | \lambda_1, \tag{3}$$

where the first condition considers the capacity of network A when the call is considered as either a new call (*i.e.*, $x \to x+1$) or a SVHO call (*i.e.*, $z \to z+1$), and the second condition considers the capacity of network B when the call is considered as a VHO call (*i.e.*, $v \to v+1$).

A $Type_1$ soft vertical handover call occurs when both the networks have the capacity to allocate the call. The

$Type_1$ SVHO probability is:

$$P_{SVHO}^{Type_1} = \Pr\left[(uU_2 + (z+1)U_1 \le N_A) \cap (wU_2 + (v+1)U_1 \le N_B)\right] | \lambda_1, \tag{4}$$

where the first condition considers the capacity of network A when the call is considered as a SVHO call (*i.e.*, $z \to z+1$), and the second condition considers the capacity of network B when the call is considered as a VHO call (*i.e.*, $v \to v+1$).

A block call for $Type_2$ service occurs when both network A and B cannot provide resources to allocate the new call, because this would cause the traffic to exceed the capacity of the networks. The $Type_2$ service blocking probability is:

$$P_{Block}^{Type_2} = \Pr\left[(xU_1 + (u+1)U_2 > N_A) \cap (vU_1 + (\min\{w+1, y+1\})U_2 > N_B)\right] | \lambda_2, \tag{5}$$

where the first condition considers the capacity of network A when the call is considered as a VHO call (*i.e.*, $u \to u+1$), and the second condition considers the capacity of network B when the call is considered as either a new call (*i.e.*, $w \to w+1$) or a SVHO call (*i.e.*, $y \to y+1$).

A hard vertical handover call for $Type_2$ service occurs when network B cannot provide resources to allocate the call, because this would cause the traffic to exceed its capacity, whereas network A can handle the call. The $Type_2$ HVHO probability is:

$$P_{HVHO}^{Type_2} = \Pr\left[(xU_1 + (u+1)U_2 \le N_A) \cap (vU_1 + (\min\{w+1, y+1\})U_2 > N_B)\right] | \lambda_2, \tag{6}$$

where the first condition considers the capacity of network A when the call is considered as a VHO call (*i.e.*, $u \to u+1$), and the second condition considers the capacity of network B when the call is considered as either a new call (*i.e.*, $w \to w+1$) or a SVHO call (*i.e.*, $y \to y+1$).

A $Type_2$ soft vertical handover call occurs when both the networks have the capacity to allocate the call. The $Type_2$ SVHO probability is:

$$P_{SVHO}^{Type_2} = \Pr\left[(xU_1 + (u+1)U_2 \le N_A) \cap (vU_1 + (y+1)U_2 \le N_B)\right] | \lambda_2, \tag{7}$$

where the first condition considers the capacity of network A when the call is considered as a VHO call (*i.e.*, $u \to u+1$), and the second condition considers the capacity of network B when the call is considered as a SVHO call (*i.e.*, $y \to y+1$). As an example, Fig. 2 shows a portion of the Markov chain diagram for variables w, x, and v, as they are strictly dependent. We assume $N_A = 4$, $U_1 = 1$, $N_B = 6$, $U_2 = 2$, $u = 0$, $w_{max} = 3$, $x_{max} = 4$, $v_{max} = 2$, and $y = z = 0$. Conditions $T_1 - T_4$, and $T_{11} - T_{14}$ are so characterized. As v increases with rate λ_1, and $x = x_{max}$, a hard handover can occur as $(v+1)U_1 + wU_2 \le N_B$, (*i.e.*, for $v = 0$, $7 \le N_B$, and for $v = 1$, $8 \le N_B$). On the other hand, transitions $T_5 - T_8$ and $T_9 - T_{15}$ are evinced in Fig. 3, where v and z are increasing at the same time, while z can decrease independently by v.

We will see in Section VI how the presented probabilities can be calculated by the network and used in a NCHO approach to control the handover process.

V. REWARD AND COST MODEL

After introducing the handover probabilities, in this section we deal with a constrained Markov decision process, and focus on *reward* and *cost functions* [40]. These concepts are investigated in this paper in order to evaluate the real benefits of performing a soft vertical handover as opposed to a hard one.

Our proposed vertical handover algorithm acts either as NCHO or MCHO approach. When the MCHO mode is selected the MT has to choose an action a based on its current state $\mathbf{s}$, at each decision epoch. The MT's transition to a new state depends on a probability function, as expressed in Section IV. This new state lasts for a period of time until the next decision epoch comes, and then the MT makes a new decision again.

For any action that the MT chooses at each state, there is a *reward* and a *cost* associated with it. The MT's goal is to maximize the expected total reward it can obtain during the connection lifetime, subject to the expected cost constraint.

Let us denote the state space of the MT by S. We only consider a finite number of states that an MT can possibly occupy [41]. The MT state contains information such as the SN identification, and the available bandwidth. Specifically, the state space can be expressed as:
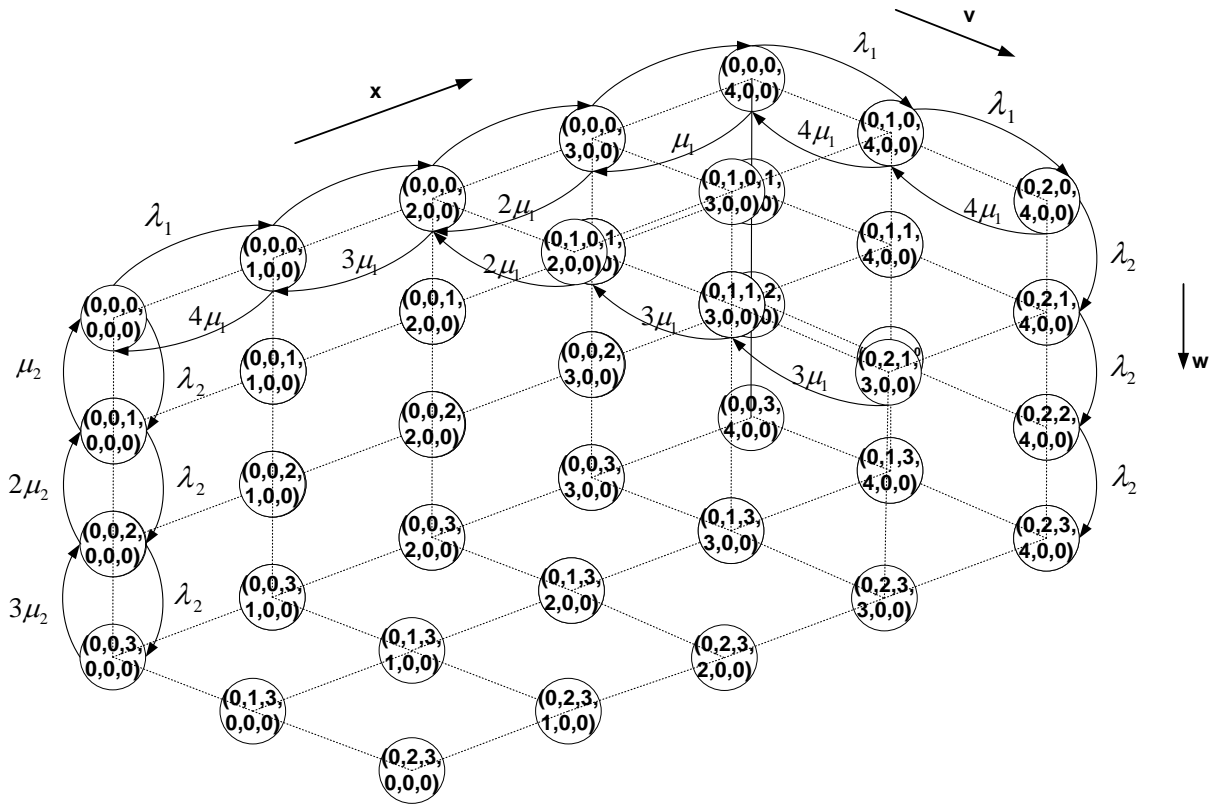
$$S = M \times B_1 \times ... \times B_M, \tag{8}$$

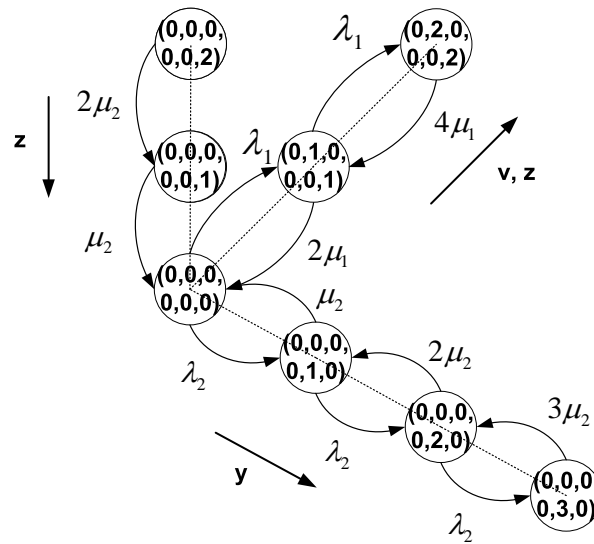Fig. 2. Markov chain diagram for $u = 0$, $w_{max} = 3$, $x_{max} = 4$, $v_{max} = 2$, $y = 0$, and $z = 0$.

Fig. 3. Markov chain diagram for y, v, and z.

1   where the symbol $\times$ denotes the Cartesian product, and B_M is the available bandwidth of M networks (*i.e.*,
2   $M = 2$, for network A and B). Basically, the bandwidths can be quantized into multiple of unit bandwidth (*i.e.*,
3   $B_A = \{1, 2, \ldots, b_{A,max}\}$, where $b_{A,max}$ is the maximum bandwidth available to a connection from network A).
4   For example, if network A is a WLAN, and network B corresponds to UMTS, we set B_A and B_B as 500 kbps
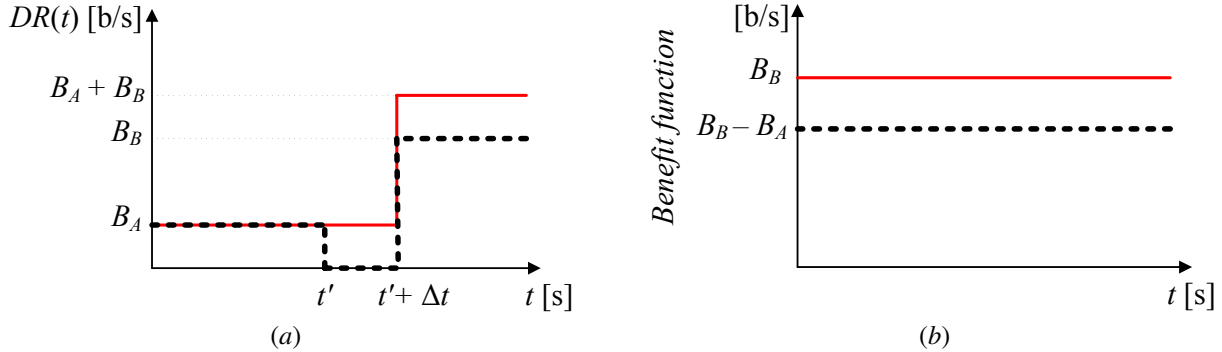5   and 5 Mbps, respectively [6].

Fig. 4. Comparison of (*a*) *data rate* behavior, and (*b*) *benefit function*, in SVHO (red solid line) and HVHO (black dotted line). $B_{A,B}$ [b/s] are the bandwidths in the SN and CN, respectively.

We assume a vector $\mathbf{s} = [i, b_A, b_B]$ as the MT's current state, where i is the SN identification, and $b_{A,B}$ the current bandwidths of network A and B, respectively. At each decision epoch, the MT on the current state $\mathbf{s}$ chooses an action $a \in A_{\mathbf{s}}$, where the action set $A_{\mathbf{s}}$ consists of the all available networks that the MT can potentially switch to (*i.e.*, $A_{\mathbf{s}} \subset 2$). When an MT chooses an action a in the state $\mathbf{s}$, it receives an immediate reward (*i.e.*, $r(\mathbf{s}, a)$).

We assume that the *reward function* depends on a *benefit function*, and a *penalty function* [41]. The *benefit function* is expressed in terms of data rate that a MT benefits from when she makes a hard/soft handover to a CN. The *penalty function* represents the network resources —in terms of bandwidth [b/s]— allocated when a MT makes a hard/soft handover to a CN.

From the contributions of both benefit and penalty functions we can define the *reward function* [3], which is mainly exploited for the selection of a Soft Vertical Handover, with respect to the Hard Vertical Handover.

A. Data Rate-based benefit function

The Data Rate (DR)-based *benefit function* represents the amount of data rate [b/s] that a MT can gain when it selects an action a in state $\mathbf{s}$, and makes a hard or soft vertical handover.

Let us assume a dual-mode MT is connected to a SN at a data rate B_A [b/s]. For any reason, at $t = t'$ the MT sends to the SN a (vertical) handover request message. The SN will check for available CNs. The (vertical) handover request will be approved and the MT will switch to the CN in a soft handover (SVHO) or hard handover (HVHO) way. We assume that the selected CN provides to the MT a data rate $B_B > B_A$ [b/s]. In this case, we expect that a vertical handover is necessary to improve the quality of service performance. The time interval from the handover request message sending and the switching to a CN is assumed to be Δt (*i.e.*, $\Delta t = 120$ ms, [1]).

For a SVHO request, during the Δt time interval, the MT will continue transmitting at B_A data rate, while for a HVHO the MT will interrupt the connection with the SN and no data transmission will be sent. Then, after Δt time interval, for a SVHO the MT will be connected both with the SN and the CN, and will be transmitting at $(B_A + B_B)$ [b/s] data rate, while for a HVHO, the MT will be connected only with the CN and the data rate will be B_B [b/s]. The data rate performance are depicted in Fig. 4 (*a*), both for SVHO and HVHO, respectively. More in general it can be expressed as

$$DR(t) = \begin{cases} b_i, & t < t' \\ b_i, & t' < t < \Delta t' \quad \text{in SVHO} \\ 0, & t' < t < \Delta t' \quad \text{in HVHO} \\ b_i + b_a, & t > \Delta t' \quad \text{in SVHO} \\ b_a, & t > \Delta t' \quad \text{in HVHO} \end{cases} \tag{9}$$

where $b_{i,a}$ [b/s] are respectively the bandwidths in the SN and CN [4].

We can now give the following definition:

---

[3] In this paper, the reward function is also called as *gain function*, indifferently.

[4] The subscripts i, a in $b_{i,a}$ [b/s] respectively mean the bandwidths in the *initial* (serving) and *accessed* (candidate) networks.

Definition 5-1. ***The benefit function*** *represents the difference between the data rate received by a MT when a (hard/soft) handover is executed, and the data rate before the handover will be executed.*

Generally, by assuming that the data rate improves when a soft (vertical) handover occurs, we consider the *benefit function* as:

$$f_b^{xVHO}(\mathbf{s}, a) = DR(t > t' + \Delta t) - DR(t < t'), \tag{10}$$

where we omitted the DR contribution during the transition time Δt. Notice that the *benefit function* represents the estimation of data rate gained when a handover is executed. It is constant versus time, as depicted in Fig. 4 (*b*).

By assuming the action a is a switching from the SN to a CN, two case studies can be described by the following scenarios:

- **Case 1**: The bandwidth in the SN is higher than that one in the CN (*i.e.*, $B_A > B_B$);
- **Case 2**: The bandwidth in the SN is lower than that one in the CN (*i.e.*, $B_A \leq B_B$).

Let us suppose a MT is in the state $\mathbf{s}$, and is moving in an heterogeneous environment comprised of two networks (*i.e.*, A and B). The SN has the maximum bandwidth b_i [b/s] (*i.e.*, $b_i = \max\{B_A, B_B\}$). This represents the **Case 1** from previous list. A hard VHO is a condition to avoid, because the data rate benefit from the CN would be null. On the other hand, a soft VHO can occur and the data rate benefit function will be represented by a ratio, in which (*i*) the numerator is the difference between available bandwidths in the CN and the SN, and (*ii*) the denominator is the maximum available bandwidth. This case for SVHO and HVHO is expressed respectively as:

$$f_b^{SVHO}(\mathbf{s}, a) = \frac{(b_i + b_a) - b_i}{\max\{B_A, B_B\}} = \frac{b_a}{b_i}, \quad \text{if} \quad b_i = \max\{B_A, B_B\}, a \neq i, \tag{11}$$

and

$$f_b^{HVHO}(\mathbf{s}, a) = \frac{b_a - b_i}{\max\{B_A, B_B\}} = \frac{b_a}{b_i} - 1 < 0, \quad \text{if} \quad b_i = \max\{B_A, B_B\}, a \neq i, \tag{12}$$

where b_a [b/s] is the bandwidth in the CN.

Now, let us assume the SN has $b_i \neq \max\{B_A, B_B\}$, and $b_a = \max\{B_A, B_B\}$ is the bandwidth in the CN. If the MT switches from the SN to the CN, a soft VHO can occur and the data rate benefit is represented by a ratio, in which (*i*) the numerator is the MT's actual increase of data rate due to handover, and (*ii*) the denominator is the MT's maximum available data rate. This is the **Case 2** from previous list. For SVHO and HVHO the data rate-based *benefit functions* are, respectively

$$f_b^{SVHO}(\mathbf{s}, a) = \frac{(b_i + b_a) - b_i}{\max\{B_A, B_B\}} = \frac{b_a}{b_a} = 1, \quad \text{if} \quad b_i \neq \max\{b_A, b_B\}, b_a > b_i, \tag{13}$$

and

$$f_b^{HVHO}(\mathbf{s}, a) = \frac{b_a - b_i}{\max\{B_A, B_B\}} = 1 - \frac{b_i}{b_a}, \quad \text{if} \quad b_i \neq \max\{B_A, B_B\}, b_a > b_i. \tag{14}$$

Equations (11) – (14) can be collected in the following expressions:

$$f_b^{SVHO}(\mathbf{s}, a) = \begin{cases} b_a/b_i, & \text{if} \quad b_i = \max\{B_A, B_B\}, a \neq i, \\ \\ 1, & \text{if} \quad b_i \neq \max\{B_A, B_B\}, b_a > b_i, \end{cases} \tag{15}$$

$$f_b^{HVHO}(\mathbf{s}, a) = \begin{cases} (b_a/b_i) - 1 < 0, & \text{if} \quad b_i = \max\{B_A, B_B\}, a \neq i, \\ \\ 1 - (b_i/b_a), & \text{if} \quad b_i \neq \max\{B_A, B_B\}, b_a > b_i. \end{cases} \tag{16}$$

B. Resource Allocation-based penalty function

When a MT requests a soft/hard handover initiation (*i.e.*, during a Δt time interval), the SN and the CN will be allocating bandwidth (network resources) for the MT. Particularly, the SN will continue to reserve bandwidth to the MT for data transmission; while the CN is responsible for authentication, authorization, and accounting load
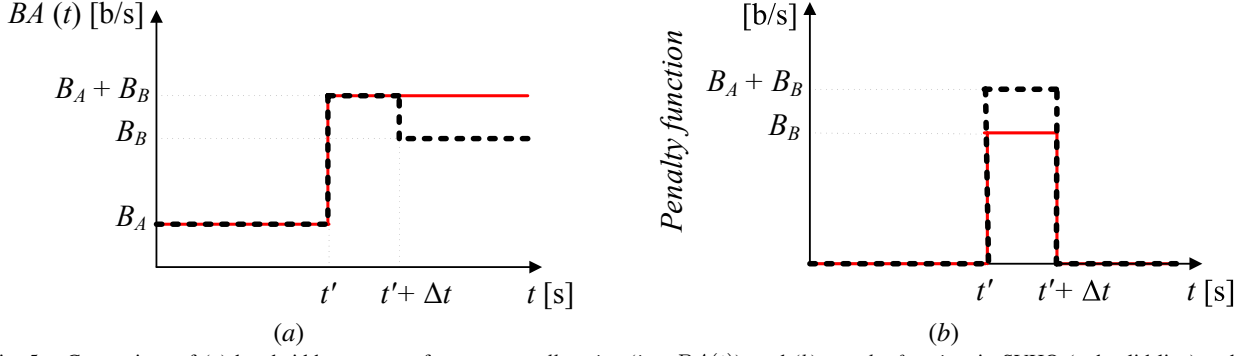
Fig. 5. Comparison of (*a*) bandwidth necessary for *resource allocation* (*i.e.*, $BA(t)$), and (*b*) *penalty function*, in SVHO (red solid line) and HVHO (black dotted line). $B_{A,B}$ [b/s] are the bandwidths in the SN and CN, respectively.

of the MT. After Δt time interval, in a SVHO both the SN and the CN will maintain resources for the MT, while in a HVHO, the SN will drop the allocated resources to the MT, and it will be connected only with the CN.

The behavior of bandwidth for *resource allocation* during a soft and hard vertical handover is depicted in Fig. 5 (*a*). More in general it can be expressed as

$$BA\left(t\right) = \begin{cases} b_i, & t < t' \\ b_i + b_a, & t' < t < t' + \Delta t \\ b_i + b_a, & t > t' + \Delta t \quad \text{in SVHO} \\ b_a, & t > t' + \Delta t \quad \text{in HVHO} \end{cases} \tag{17}$$

We can now give the following definition:

Definition 5-2. **The penalty function** *represents the amount of bandwidth [b/s] that a MT gets when makes a hard or soft vertical handover, at the expense of bandwidth allocation.*

Formally, the *penalty function* is expressed as the difference between the bandwidth for resource allocation (*i.e.*, $BA(t)$), and that for data transmission (*i.e.*, $DR(t)$), such as

$$g_p^{xVHO}\left(\mathbf{s}, a\right) = BA\left(t\right) - DR\left(t\right), \tag{18}$$

which is zero for $t > t' + \Delta t$ and $t < t'$, as well as shown in Fig. 5 (*b*).

For the previous defined **Case 1**, the *penalty function* in hard and soft handover has the following expressions, respectively

$$g_p^{HVHO}\left(\mathbf{s}, a\right) = \frac{b_i + b_a}{\max\left\{B_A, B_B\right\}} = 1 + \frac{b_a}{b_i}, \quad \text{if } b_i = \max\left\{B_A, B_B\right\}, \, a \neq i \tag{19}$$

and

$$g_p^{SVHO}\left(\mathbf{s}, a\right) = \frac{b_a}{\max\left\{B_A, B_B\right\}} = \frac{b_a}{b_i} \quad \text{if } b_i = \max\left\{B_A, B_B\right\}, \, a \neq i \tag{20}$$

which are normalized with respect to the maximum available bandwidth.

Finally, for **Case 2** the *penalty function* in hard and soft handover is expressed as

$$g_p^{HVHO}\left(\mathbf{s}, a\right) = \frac{b_i + b_a}{\max\left\{B_A, B_B\right\}} = 1 + \frac{b_i}{b_a}, \quad \text{if } b_i \neq \max\left\{B_A, B_B\right\}, \, b_a > b_i \tag{21}$$

and

$$g_p^{SVHO}\left(\mathbf{s}, a\right) = \frac{b_a}{\max\left\{B_A, B_B\right\}} = 1, \quad \text{if } b_i \neq \max\left\{B_A, B_B\right\}, \, b_a > b_i \tag{22}$$
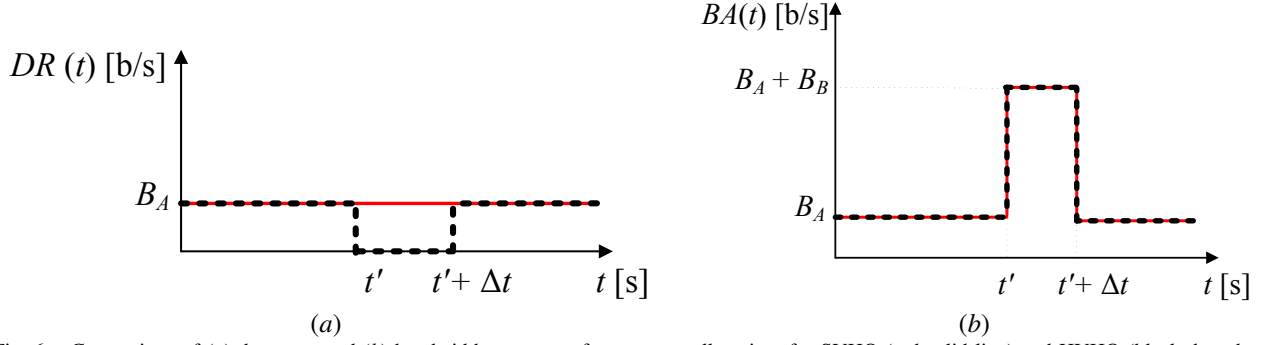
Fig. 6. Comparison of (*a*) data rate, and (*b*) bandwidth necessary for resource allocation, for SVHO (red solid line) and HVHO (black dotted line), in case of call rejection by the CN. $B_{A,B}$ [b/s] are the bandwidths in the SN and CN, respectively.

which are normalized with respect to the maximum available bandwidth, respectively. Equations (19) – (22) can be collected in the following expressions:

$$g_p^{HVHO}(\mathbf{s},a) = \begin{cases} 1 + (b_a/b_i), & \text{if } b_i = \max\{B_A, B_B\},\ a \neq i \\ 1 + (b_i/b_a), & \text{if } b_i \neq \max\{B_A, B_B\},\ b_a > b_i \end{cases} \tag{23}$$

$$g_p^{SVHO}(\mathbf{s},a) = \begin{cases} b_a/b_i, & \text{if } b_i = \max\{B_A, B_B\},\ a \neq i \\ 1, & \text{if } b_i \neq \max\{B_A, B_B\},\ b_a > b_i \end{cases} \tag{24}$$

The range $[t', t'+\Delta t]$ represents a VHO initiation time, necessary to authentication, authorization, and accounting information to MT's network access. We assume that Δt also includes the link allocation and acquisition delay. If a CN is not available to accept a new call, the time interval from the call rejection to the instant when the call is still served by the SN is again Δt, as assuming the call rejection time to be negligible.

Fig. 6 (*a*) and (*b*) respectively depict the data rate and bandwidth allocation performance, for SVHO and HVHO call rejection case, respectively. In case of call rejection, the *benefit function* is null, while the *penalty function* is again the same as (21) and (22), for HVHO and SVHO, respectively. As g_p^{HVHO} is greater than g_p^{SVHO}, this represents the benefit of a soft VHO respect to the hard one.

C. Reward function

After defining the *benefit* (*i.e.*, f_b^{xVHO}) and *penalty function* (*i.e.*, g_p^{xVHO}), we introduce the *reward function* experienced by a mobile terminal. It is expressed as:

$$G^{xVHO}(\mathbf{s},a) = f_b^{xVHO}(\mathbf{s},a) - r \cdot g_p^{xVHO}(\mathbf{s},a), \tag{25}$$

where r is the cost constraint for the *penalty function*.

Normally, the aim is a maximization of the *benefit function* (*i.e.* $f_b^{xVHO}(\mathbf{s},a)$), and a minimization of the *penalty function* (*i.e.* $g_p^{xVHO}(\mathbf{s},a)$). For this purpose, the ideal *reward function* —also called *gain function*— can be written as:

$$\max_{\mathbf{s}\in S}\left\{G^{xVHO}(\mathbf{s},a)\right\} = \max_{\mathbf{s}\in S}\left\{f_b^{xVHO}(\mathbf{s},a)\right\} - r \min_{\mathbf{s}\in S}\left\{g_p^{xVHO}(\mathbf{s},a)\right\}, \tag{26}$$

which occurs for SVHO and HVHO cases, respectively when

$$\max_{\mathbf{s}\in S}\left\{f_b^{SVHO}(\mathbf{s},a)\right\} = 1, \tag{27}$$

$$\max_{\mathbf{s}\in S}\left\{f_b^{HVHO}(\mathbf{s},a)\right\} = 1 - (b_i/b_a), \tag{28}$$
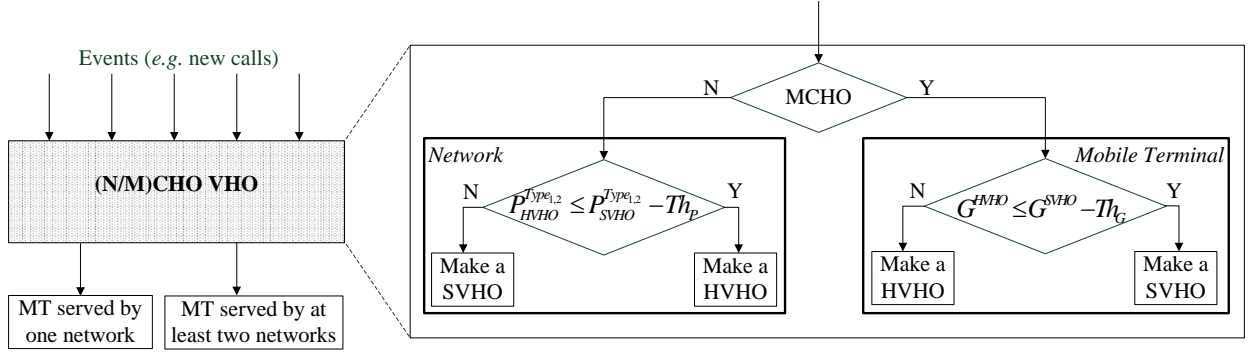
Fig. 7. Schematic of (N/M)CHO hard/soft VHO algorithm.

and

$$\min_{\mathbf{s} \in S} \left\{ g_p^{SVHO}\left(\mathbf{s}, a\right) \right\} = b_a/b_i, \tag{29}$$

$$\min_{\mathbf{s} \in S} \left\{ g_p^{HVHO}\left(\mathbf{s}, a\right) \right\} = 1 + (b_a/b_i). \tag{30}$$

By replacing (27) – (30), into (26), the ideal gain function for SVHO and HVHO is respectively

$$\begin{cases} \max_{\mathbf{s} \in S} \left\{ G^{SVHO}\left(\mathbf{s}, a\right) \right\} = 1 - r(b_a/b_i), \\ \\ \max_{\mathbf{s} \in S} \left\{ G^{HVHO}\left(\mathbf{s}, a\right) \right\} = -\left[b_i^2 + (r-1)b_a b_i + r b_a^2 \right]/(b_a b_i). \end{cases} \tag{31}$$

Notice that the choice between HVHO and SVHO strictly depends not only by the hard and soft vertical handover probability (*i.e.*, $P_{xVHO}^{Type_{1,2}}$), but also by the *gain function* (*i.e.*, G^{xVHO}). Basically, the probabilities of performing a hard or soft vertical handover take network congestions and traffic load into account, while the gain function mainly considers user and QoS requirements.

VI. HYBRID SOFT/HARD VHO ALGORITHM

In this section we present a simple algorithm that selects the proper vertical handover mechanism to maintain seamless connectivity. Our technique can work both as (*i*) NCHO by relaying on a probabilistic approach for the selection of hard/soft vertical handover, and as (*ii*) MCHO by exploiting the gain function as the handover decision criterion. A schematic representation of the proposed (N/M)CHO VHO technique is provided in Fig. 7. When an event occurs (*e.g.*, a new voice call originates in the network), the proposed algorithm first computes the (N/M)CHO decision, and then it evaluates the benefits of making a soft or a hard handover.

The handover scheme is then expressed by three phases, each of them associated to a time epoch as depicted in Fig. 8, and described as follows:

1) *VHO Initiation*: for $t < t'$, if the power level at the MT's receiver, expressed by the Received Signal Strength (RSS) from the SN, is lower than the receiver sensibility (S), the SN will check for available CNs. The power control is formulated as $RSS_{\text{SN}} \leq S$;

2) *VHO Preparation*: for $t' < t < t' + \Delta t$, immediately after the check for CNs availability, the decision on whether to perform an NCHO or an MCHO is made. Then, a soft or hard VHO is executed depending on the specific parameter used. In this phase, a VHO comparison is performed both between the NCHO and MCHO modalities, and between the soft and hard handover. Our scheme allows the MT to use either the NCHO or the MCHO approach, we are interested in using both and showing their different performance, therefore we will not propose a specific way to choose the approach. We simply assume that the MCHO approach is used when the MT possesses all the information to calculate the reward and cost function, presented in Section V, and the NCHO otherwise.

Once the NCHO-MCHO choice has been made, the algorithm applies the following reasoning:

- In the **NCHO case**, the SN initiates and then executes a soft/hard vertical handover decision according to the SVHO/HVHO probabilities. We assume that a hard vertical handover will be performed whenever the
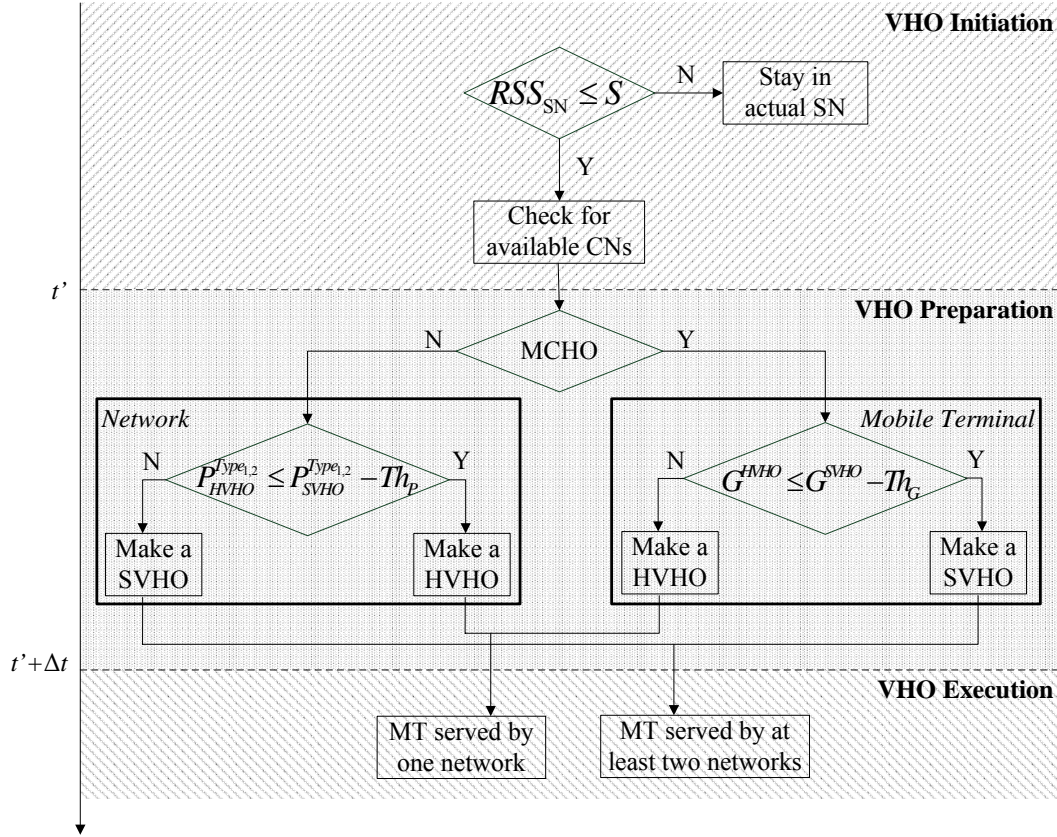
Fig. 8. Detailed MCHO/NCHO hard/soft vertical handover algorithm scheme.

HVHO probability (*i.e.*, $P_{HVHO}^{Type_{1,2}}$) is greater than the SVHO probability (*i.e.*, $P_{SVHO}^{Type_{1,2}}$) by the vertical handover threshold (*i.e.*, Th_P), such as

$$P_{HVHO}^{Type_{1,2}} > P_{SVHO}^{Type_{1,2}} + Th_P, \tag{32}$$

otherwise, a SVHO will occur;

- In the **MCHO case**, the MT makes a soft/hard vertical handover decision according to the SVHO/HVHO gain, as expressed in (25). We assume that a hard vertical handover is made whenever the HVHO gain (*i.e*, G_{HVHO}) is greater than SVHO gain (*i.e*, G_{SVHO}) by the gain threshold (*i.e*, Th_G), such as

$$G_{HVHO} > G_{SVHO} + Th_G, \tag{33}$$

otherwise, a SVHO will occur.

3) *VHO Execution*: for $t > t' + \Delta t$, the MT will be served only by the CN (that will become the new SN) if a HVHO occurs; otherwise, if a SVHO occurs, the MT will be served by both the CN and the SN.

VII. SIMULATION RESULTS

In this section, we report the simulation results that validate the proposed soft/hard vertical handover technique. The 6-D Markov chain scheme allows us to compute the probabilities of the probabilistic approach in order to evaluate the handover gain, for both soft and hard VHO cases. We expect high values of gain for the SVHO, along with a low number of handover occurrences. This trade-off can assure a maximization of the throughput and a limitation of unnecessary vertical handovers.

This section is organized as follows: we first describe the simulation setup in Subsection VII-A, and then in Subsection VII-B we present simulation results of our handover algorithm, when it works in NCHO or in MCHO mode.
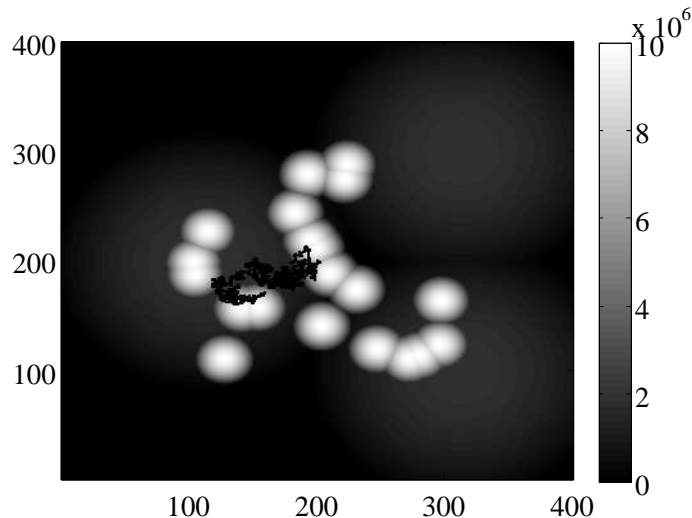
Fig. 9. Map of a generated outdoor scenario, composed of 20 WLAN access points (*white small cells*), and 3 UMTS base stations (*grey large cells*). The MT moves at low speed inside the map, forming a path (*black line*). The map is comprised of 400×400 zones of 25 m^2 each, characterized by the distribution of UMTS/WLAN data rate (in Mb/s).

A. Simulation setup

The simulation results are computed by using an event-driven simulator, which has been developed in Matlab 7.11 and has been extensively used in several works dealing with seamless connectivity solutions in heterogeneous networks [21], [25], [42], [43]. It simulates users that move with velocity between 0.5 and 30 m/s in a heterogeneous network environment, comprised of a variable set of network cells and base stations, respectively offering WLAN and UMTS connectivity. Simulation results have been averaged over a set of 100 Monte Carlo runs. Each of them simulates a network scenario (*i.e.*, a map), as depicted in Fig. 9.

In the simulation environment, we consider a heterogeneous scenario consisting of 20 WLAN Access Points (APs), and 3 UMTS Base Stations (BSs). Simulations generate a MT that moves in such scenario and receives packets from APs or BSs when it crosses the coverage area of WLAN or UMTS, respectively. Basically, the MT moves in a pedestrian environment (*i.e.*, at a speed of 0.5 m/s), making a journey composed of 2500 steps. It is worth noting that we simulate a pedestrian environment because our proposal arises from the need to provide seamless connection to users in densely crowded cellular scenarios, where they can use their multi-interface mobile devices while walking. The simulation scenario can be easily extended to include higher speed users (*e.g.*, in a vehicular environment). In this case, we would expect that an increase of the speed would reduce the network performance [42]. The heterogeneous scenario is modeled as a region of 4 km^2, split into 400×400 zones of 25m^2. For each simulated scenario, the displacement of BSs and APs is randomly generated, in order to approximate real outdoor environments. The APs and BSs positions are unknown *a priori*, and the MT moves within this area according to a Random Waypoint model. In Fig. 9 we show the simulated network environment and an example of path travelled by a MT. In the figure, the clearer the area, the higher the data rate available for the MTs.

For the network setup, we consider typical network parameters [7], [25]: the nominal UMTS/WLAN capacity at the centre of the cell (*i.e.*, $B_{UMTS} = 2$ Mb/s, and $B_{WLAN} = 10$ Mb/s), the maximum UMTS/WLAN transmitting power (*i.e.*, $P_{UMTS} = 43$ dBm, and $P_{WLAN} = 30$ dBm), the UMTS/WLAN carrier frequency (*i.e.*, $f_{UMTS} = 2.2$ GHz, and $f_{WLAN} = 2.4$ GHz), the UMTS/WLAN signal power attenuation (*i.e.*, $Att_{UMTS} = 20$ dB, and $Att_{WLAN} = 20$ dB), the UMTS/WLAN cell radius (*i.e.*, $r_{UMTS} = 600$ m, and $r_{WLAN} = 120$ m), and the UMTS/WLAN sensitivity at the mobile terminal receiver side (*i.e.*, $S = -100$ dBm). Finally, the signal power attenuation has been modeled according to the Okomura-Hata model, together with an AWG (Additive White Gaussian) channel model. In TABLE III (*a*) we have collected the main parameters used in the simulation scenario.

The proposed Markov chain provides the computation of the probabilities that a (H/S)VHO occurs as well as the gain obtained when a (H/S)VHO is initiated. We implemented equations (6), (7) and (31) into our Matlab simulator in order to have an integrated modeling tool. However, the probability that a MT performs a handover from UMTS to WLAN, and vice versa, is strictly dependent on the network setup. Hence, the MT makes the soft/hard VHO decision according to the probabilistic approach.

TABLE III
LIST OF PARAMETERS USED IN (*a*) THE SIMULATION SCENARIO SETUP, AND (*b*) THE MARKOV CHAIN MODEL.

Parameter	Value	Parameter	Value	Parameter	Value
Region size	4 km^2	B_{UMTS}	2 Mb/s	N_{UMTS}	6
AP/BS positions	Randomly generated	B_{WLAN}	10 Mb/s	N_{WLAN}	4
WLAN cells	20	P_{WLAN}	30 dBm	U_{UMTS}	2
UMTS cells	3	P_{UMTS}	43 dBm	U_{WLAN}	4
User mobility	Random WayPoint	Att_{UMTS}	20 dB	λ_1	0.4
User speed	0.5 m/s	Att_{WLAN}	20 dB	λ_2	0.6
S	-100 dBm	f_{UMTS}	2.2 GHz	Th_P	0.2
r_{UMTS}	600 m	f_{WLAN}	2.4 GHz	Th_G	0.2
r_{WLAN}	120 m	Channel	AWG	r	0.2
(*a*)				(*b*)	

B. Performance analysis

Simulation results have been obtained for our technique working either in NCHO or MCHO modality [5]. For each simulated scenario, different results can be collected. However, one hundred simulations are run for each scenario in order to get the averaged results in terms of [6]: (*i*) average throughput during the MT's journey [b/s], (*ii*) Cumulative Received Bits (CRBs) by the MT [b/s], (*iii*) bandwidth saving percentage, and (*iv*) number of handover occurrences during the MT's journey. Furthermore, in order to assess the effectiveness of the proposed technique, we show a comparison of our approach to other vertical handover solutions.

Let us assume a MT is a dual-mode device with UMTS and WLAN network interface cards; when the MT is connected to a WLAN AP, it uses a $Type_1$ service. Notice that, in our simulator, each service is defined by the amount of radio resources requested. A particular service request (*e.g.*, a new phone call) is served if the required radio resources are available; otherwise the request is switched to another network. If the new network is also unable to serve the request, the call will be blocked and cleared.

We assume that WLAN and UMTS networks have N_{WLAN} and N_{UMTS} units of resources, respectively; U_{WLAN} and U_{UMTS} (*i.e.*, $U_{WLAN} > U_{UMTS}$) are the units of resources required by a $Type_1$ and $Type_2$ service, respectively. We set $N_{WLAN} = 4$, $U_{WLAN} = 4$, $N_{UMTS} = 6$, $U_{UMTS} = 2$, and the average arrival rates $\lambda_1 = 0.4$ and $\lambda_2 = 0.6$, for $Type_1$ and $Type_2$ services, respectively. Eventually, the cost constraint has been set to $r = 0.2$, as shown in TABLE III (*b*).

NCHO-mode. When the RSS of the active connection drops below a threshold, the MT performs a soft or hard VHO decision according to SVHO and HVHO probabilities, as in (32), where the VHO threshold Th_P is set equal to 0.2 (see TABLE III (*b*)). In Fig. 10 (*a*) we show the performance in terms of throughput experienced by the MT during its journey when either a soft or hard VHO occurs. We observe that the performance of the SVHO shows average values which are greater than those of the HVHO. Particularly, by examining the last part of the MT's path (*i.e.*, at 2500-th step), we can see the gap between average Cumulative Received Bits (CRB) for SVHO and HVHO, (Fig. 10 (*b*)). This demonstrates that a SVHO provides better performance than a HVHO. The CRB performance is evaluated for different values of the *waiting time* parameter (*i.e.*, T_{wait} in the range $[30, 300]$ seconds). The *waiting time* parameter is employed to limit the handover frequency, as well as to save MT's battery life. It represents the time interval for a MT to be in idle mode and inhibit the use of any handover algorithm [11], [25]. Simulation results show that the CRB does not depend on the variation of the *waiting time*; only little fluctuations are experienced for HVHO case. This is caused by he fact that CRB performance only depends on handover probability comparison.

In Fig. 11 (*a*) is depicted the average MT's gain —in term of bandwidth percentage— for HVHO and SVHO cases versus the waiting time. The cost constraint in the penalty function is assumed equal to 0.2, that means that a VHO introduces a penalty of 20%. Maximum average gain is 1.62% in SVHO, and 1.04% in HVHO, for $T_{wait} = 30$ s,

---

[5]The selection of the algorithm working modality is out of the scope of this paper.

[6]The simulated results have been obtained for both soft and hard VHO cases.
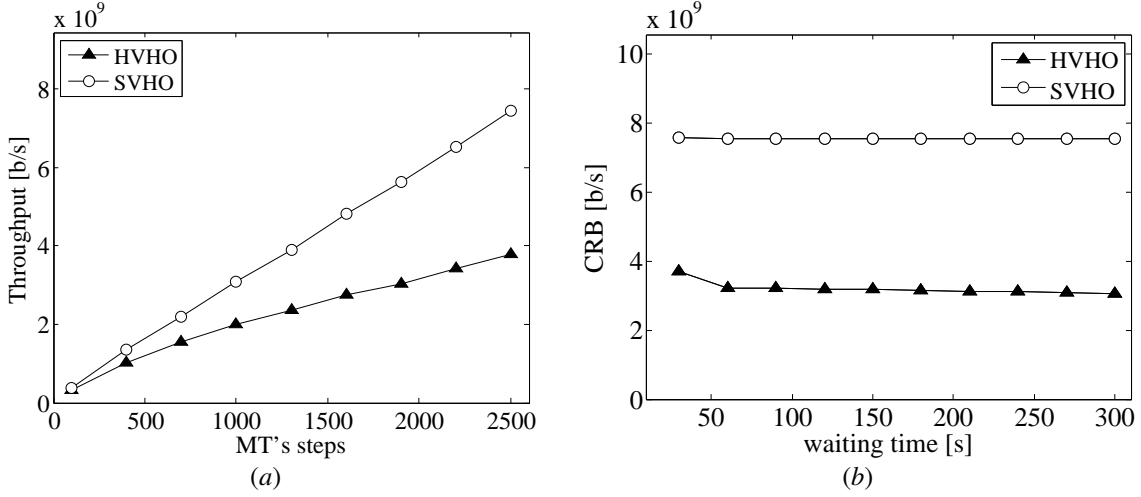
Fig. 10. Performance of hybrid soft/hard VHO in NCHO mode: (*a*) average throughput during the MT's journey in HVHO and SVHO approach; (*b*) average CRB in HVHO and SVHO vs. the waiting time parameter.
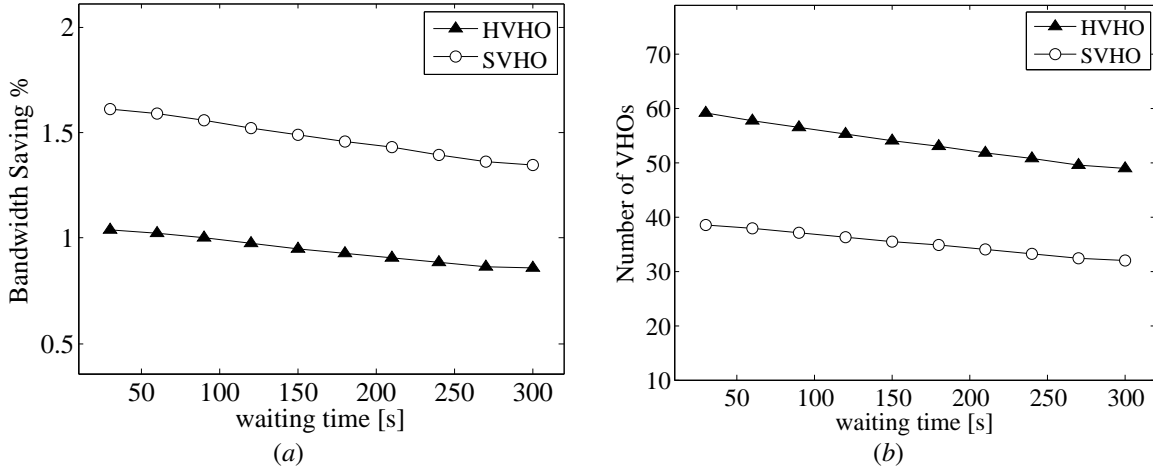

Fig. 11. Performance of hybrid soft/hard VHO in NCHO mode: (*a*) average bandwidth saving percentage for HVHO and SVHO, and (*b*) average number of hard and soft handovers during the MT's journey, vs. the waiting time.

while minimum values are 1.35% and 0.86% for SVHO and HVHO, respectively, when $T_{wait} = 300$ s. We observe that the waiting time parameter affects the average gain.

Finally, Fig. 11 (*b*) depicts the average VHO frequency *i.e.*, the number of handovers initiated and concluded between a SN and a CN, during the MT's journey. The soft VHO occurrences are lower than those in HVHO case. This represents a good result to avoid the well-known *ping-pong* effect, that is a cause of excessive network resource consumption and also affects the MT's performance [25]. By defining the *handover frequency bands* as the difference between the minimum and maximum number of handover occurrences, for hard and soft cases, we notice the HVHO frequency band is about 191, while it is 121 for SVHO case. As a conclusion, we claim that the waiting time parameter affects in a good way the SVHO frequency.

MCHO mode. In this case, the MT initiates and then executes a soft/hard vertical handover decision, according to SVHO/HVHO gain, as defined in (33), where the gain threshold Th_G is set equal to 0.2 (see TABLE III (*b*)). Fig. 12 (*a*) and (*b*) show the throughput and the CRB experienced by the MT during the journey, respectively. Results are in line with previous results for NCHO mode (see Fig. 10), although a small performance decrease is noticeable. This is expected since the algorithm working in NCHO allows MT to make handovers according to decisions taken directly by the SN, and no QoS or user requirements are considered.
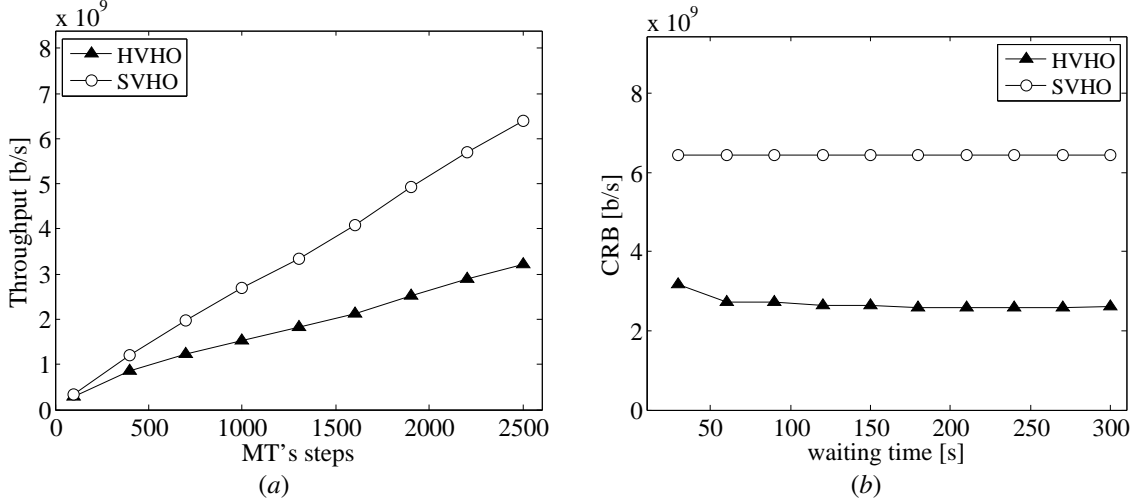
Fig. 12. Performance of hybrid soft/hard VHO in MCHO mode: (*a*) average throughput during the MT's journey in HVHO and SVHO approach; (*b*) average CRB in HVHO and SVHO vs. the waiting time parameter.
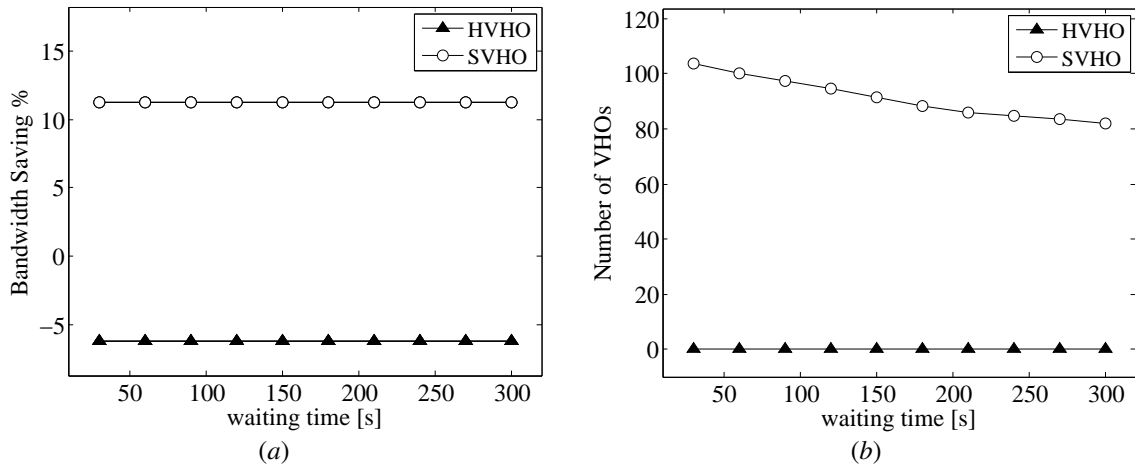

Fig. 13. Performance of hybrid soft/hard VHO in MCHO mode: (*a*) average bandwidth saving percentage for HVHO and SVHO, and (*b*) average number of hard and soft handovers during the MT's journey, vs. the waiting time.

Fig. 13 (*a*) depicts the average MT's gain in term of bandwidth saving percentage, for HVHO and SVHO cases respectively. In this case, the bandwidth saving does not dependent on the waiting time. The maximum value of gain is 11.20% in SVHO, while a negative value is obtained in HVHO (*i.e.*, −6.20%). This result shows that in MCHO, the SVHO is a preferred choice, and the HVHO shows low performance, like in NCHO mode. We can notice a strong increase of bandwidth gain in MCHO, with respect to NCHO mode. This aspect makes the MCHO the preferred choice for improving QoS requirements.

Finally, Fig. 13 (*b*) depicts the average vertical handover frequency in soft and hard cases. Simulated results show a higher number of soft handovers with respect to the hard case, that is null. This represents how soft handovers are the preferred choice in NCHO, whereas hard handovers are avoided due to the negative gain. As a consequence, the HVHO frequency band is 0, while it is 21 for the SVHO case.

In order to assess the performance of the proposed technique, we will compare our approach with other vertical handover solutions. Particularly, we have implemented three solutions based on different handover decision criteria:

- A multi-parameter vertical handover exploiting data rate requirements and information on interference level, *i.e.*, DRI, Data Rate and Interference-based vertical handover;
- Two single-parameter vertical handover approaches: (*i*) a traditional Power-Based approach (PB) and (*ii*) a Location-Based approach (LB). In order to initiate a handover, they use power measurements and location information, respectively.

All the implemented techniques are intended to limit the number of handover occurrences. In contrast, DRI is a combined vertical handover decision algorithm aiming to optimize user requirements (goodput), but at the same time takes into account the interference and noise level in a candidate network. Basically, in the DRI approach, the RSS measurement is used to drive a VHO during the initiation phase, while the data rate estimation from Signal-to-Noise-and-Interference Ratio (SINR) factor guides the handover execution phase.

In Fig. 14 (*a*) and (*b*) we respectively report the average number of handovers and CRB, for all the approaches, versus the waiting time. Notice that the benefit of a handover technique can be highlighted through a tradeoff between the number of handover occurrences, and the CRB. The optimal approach provides a high CRB with a low number of handovers, which means achieving performance maximization while limiting the energy consumption. In Fig. 14 (*a*), we can see that the SVHO shows the best performance in terms of CRB, for both NCHO and MCHO modalities. These results highlight the benefit of performing soft vertical handover in respect of the hard approaches. As a matter of fact, hard solutions, both in NCHO and MCHO, have a behavior as typical as that of other multi and single parameter schemes. However, considering Fig. 14 (*b*), we can notice that the SVHO in MCHO needs a high number of handovers, whereas lower values are shown for both hard and soft approaches in NCHO. Single-parameter approaches present a very low number of handovers, due to their purpose of limiting unwanted handovers.
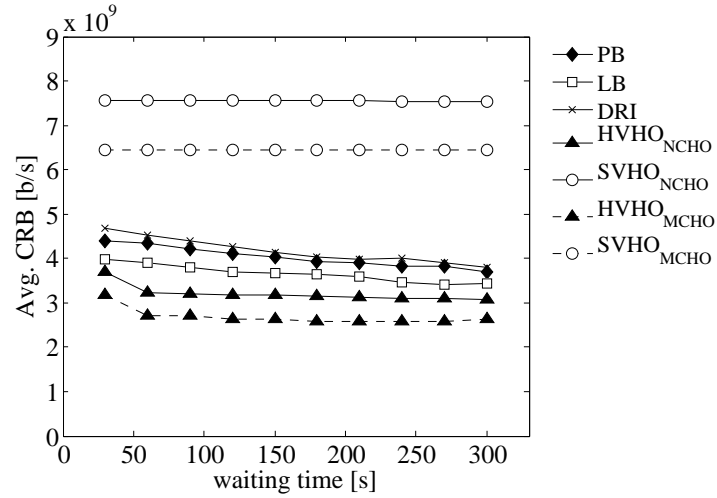
(*a*)
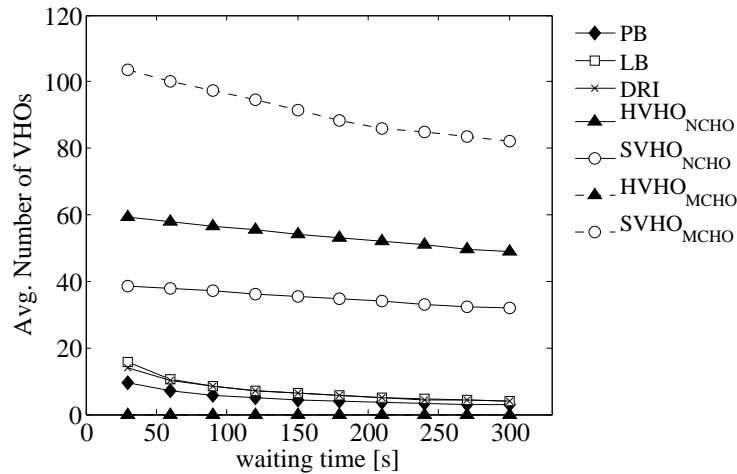
(*b*)

Fig. 14.   Comparison of the simulation results for different approaches. Average values of (*a*) the CRB, and (*b*) the number of handover occurrences, vs. the waiting time.

TABLE IV
RANKING OF SELECTED HANDOVER TECHNIQUES BASED ON THE CRB AND THE NUMBER OF VHO OCCURRENCES.

VHO technique	CRB	Number of VHO
SVHO$_{(NCHO)}$	*high*	*medium*
HVHO$_{(NCHO)}$	*medium*	*medium*
SVHO$_{(MCHO)}$	*high*	*high*
HVHO$_{(MCHO)}$	*medium*	*low*
PB	*medium*	*low*
LB	*medium*	*low*
DRI	*medium*	*low*

To summarize, we can state that SVHO in both NCHO and MCHO modalities shows the best tradeoff between the CRB and the average number of handovers. Through a classification of CRB values, we rank values in the range (*i*) $[0, 3.0]$ Gb as *low* values, (*ii*) $[3.0, 6.0]$ Gb as *medium* values, and (*iii*) $[6.0, 8.0]$ Gb as *high* values. In the same way, we classify values of the number of VHOs in the range (*i*) $[0, 30]$ as *low* values, (*ii*) $[30, 60]$ as *medium* values, and (*iii*) $[60, 100]$ as *high* values. Obviously, the higher the CRB the better the performance, and the higher the average number of handovers the worse the performance. According to such considerations, in TABLE IV we rank the handover approaches in order to highlight the effectiveness of the compared techniques. We notice that the soft approaches outperform other techniques concerning the CRB, whereas hard approaches show a decrease of the CRB as well as a limitation of the average number of handover occurrences. Thus, depending on the criterion we want to consider, a given handover technique can be selected accordingly.

VIII. CONCLUSIONS

In this paper we have proposed a novel vertical handover algorithm which exploits the potentiality of both hard and soft handovers, in a dual-mode configuration (*i.e.*, NCHO and MCHO). Our algorithm aims to maintain seamless connectivity to users moving in a heterogeneous networks environment (*i.e.*, comprised of WLAN hot-spots and UMTS base stations), while still guaranteeing user QoS requirements. We have designed a system model through a multi-dimension Markov chain, in order to define soft and hard vertical handover probability, as well as block probabilities. Simulation results have shown how soft vertical handover is preferred, compared to the traditional hard approach, since bandwidth gain and limitation of number of handovers are maximized. Moreover, the effectiveness of the proposed approach with respect to other single and multi-parameter VHO techniques has been proven by extensive simulation results.

Future work will investigate the adoption of our technique in the IEEE 802.21 framework. In this work, we highlighted the correspondence of MIH events and commands with the proposed technique, where an event corresponds to a new incoming call and a command to a (N/M)CHO decision.

REFERENCES

[1] K. Knightson, N. Morita, and T. Towle, "NGN architecture: generic principles, functional architecture, and implementation," *IEEE Communications Magazine*, vol. 43, no. 10, pp. 49–56, 2005.

[2] M. Lin, D. T. H. Choi, and T. L. Porta, "Network Integration in 3G and 4G Wireless Networks," in *Proc. of 19th Intl. Conf. on Computer Communications and Networks (ICCCN)*, pp. 1–8, August 2010.

[3] I. Papapanagiotou, D. Toumpakaris, J. Lee, and M. Devetsikiotis, "A Survey on Next Generation Mobile WiMAX Networks: objectives, features and technical challenges," *Communications Surveys & Tutorials, IEEE*, vol. 11, pp. 3–18, Fourth Quarter 2009.

[4] J. L. Salina and P. Salina, *"Next Generation Networks: perspectives and potentials"*. John Wiley & Sons, 2008.

[5] Q. A. Zeng and D. P. Agrawal, "Handoff in Wireless Mobile Networks," in *Handbook of Wireless Networks and Mobile Computing*, ch. 1, pp. 1–25, John Wiley & Sons, Inc., 2002.

[6] H. Holma and A. Toskala, *"WCDMA for UMTS: Radio Access for Third Generation Mobile Communications"*. 3rd Ed. John Wiley & Sons, Ltd., 2004.

[7] T. Ahmed, K. Kyamakya, and M. Ludwig, "A Context-Aware Vertical Handover Decision Algorithm for Multimode Mobile Terminals and its Performance," in *Proc. of IEEE/ACM Euro American Conference on Telematics and Information Systems (EATIS'06)*, pp. 19–28, 2006.

[8] A. Sgora and D. Vergados, "Handoff Prioritization and Decision Schemes in Wireless Cellular Networks: a survey," *Communications Surveys & Tutorials, IEEE*, vol. 11, no. 4, pp. 57–77, 2009.

[9] J. Márquez-Barja, C. T. Calafate, J.-C. Cano, and P. Manzoni, "An Overview of Vertical Handover Techniques: algorithms, protocols and tools," *Computer Communications*, vol. 34, pp. 985–997, June 2011.

[10] X. Yan, Y. A. Şekercioğlu, and S. Narayanan, "A Survey of Vertical Handover Decision algorithms in Fourth Generation Heterogeneous Wireless Networks," *Computer Networks*, vol. 54, no. 11, pp. 1848–1863, 2010.

[11] H. Cho, J. Park, W. Ko, K. Lim, and W. Kim, "A study on the MCHO method in Hard Handover and Soft Handover between WLAN and CDMA," in *Proc. of the International Conference on Consumer Electronics (ICCE'05)*, pp. 391–392, January 2005.

[12] "IEEE 802.21 Media Independent Handover Services – Media Independent Handover, http://www.ieee802.org/21/."

[13] C. Gu, J.-Y. Wang, and M. Song, "An Intelligent Seamless Handover Mechanism based on IEEE 802.21," in *Proc. of International Conference on MultiMedia and Information Technology, (MMIT '08)*, pp. 558–561, December 30–31, 2008.

[14] S. Benoubira, M. Frikha, S. Tabbane, and K. Ayadi, "Vertical Handover based on IEEE802.21 and Mobile IPv6 in UMTS/WLAN networks," in *Proc. of First International Conference on Communications and Networking, (ComNet 2009)*, pp. 1–6, November 3–6, 2009.

[15] H. Song, J. Kim, J. Lee, and H.-S. Lee, "Analysis of vertical handover latency for IEEE 802.21-enabled Proxy Mobile IPv6," in *Proc. of 13th International Conference on Advanced Communication Technology (ICACT)*, pp. 1059–1063, February 13–16, 2011.

[16] R. Tamijetchelvy and G. Sivaradje, "An Optimized Fast Vertical Handover Strategy for Heterogeneous Wireless Access Networks based on IEEE 802.21 Media Independent Handover standard," in *Proc. of Fourth International Conference on Advanced Computing (ICoAC)*, pp. 1–7, December 13–15, 2012.

[17] M. Zekri, B. Jouaber, and D. Zeghlache, "An Enhanced Media Independent Handover framework for Vertical Handover Decision Making based on Networks' Reputation," in *Proc. of IEEE 37th Conference on Local Computer Networks Workshops (LCN Workshops)*, pp. 673–678, October 22–25, 2012.

[18] A. A. Bathich, M. D. Baba, and M. Ibrahim, "IEEE 802.21 based Vertical Handover in WiFi and WiMAX Networks," in *Proc. of IEEE Symposium on Computers & Informatics (ISCI)*, pp. 140–144, March 18–20, 2012.

[19] I. Smaoui, F. Zarai, M. Banat, and L. Kamoun, "Heterogeneous Wireless Networks: Configuration and Vertical Handoff Management," *Wireless Personal Communications*, vol. 54, no. 3, pp. 417–445, 2010.

[20] T. Inzerilli, A. M. Vegni, A. Neri, and R. Cusani, "A Location-based Vertical Handover algorithm for limitation of the ping-pong effect," in *Proc. of IEEE International Conference on Wireless and Mobile Computing, Networking and Communications (WiMob'08)*, (Avignon (France)), October 2008.

[21] A. M. Vegni, G. Tamea, T. Inzerilli, and R. Cusani, "A Combined Vertical Handover Decision Metric for QoS Enhancement in Next Generation Networks," in *Proc. of 5th IEEE International Conference on Wireless and Mobile Computing, Networking and Communications (WiMob'09)*, (Marrakech, Morocco), pp. 233–238, October 12–14, 2009.

[22] A. Farbod and B. Liang, "Efficient structured policies for Admission Control in Heterogeneous Wireless Networks," *ACM Mobile Networks and Applications*, vol. 12, no. 5, 2007.

[23] D. Lee, J.-G. Choi, C. Joo, and S. Bahk, "Hierarchical Markov chain analysis of an adaptive bandwidth reservation algorithm in wireless communication systems," *Performance Evaluation*, vol. 65, p. 13, July 2008.

[24] J.-Y. Baek, W.-J. Kim, D.-W. Kim, Y.-J. Suh, S. Kang, and K. C. Shin, "An Efficient Multi-dimensional Handover scheme for Next Generation Networks," in *Proc. of the 3nd ACM workshop on Performance Monitoring and Measurement of Heterogeneous Wireless and Wired Networks (PM2HW2N'08)*, pp. 151–156, 2008.

[25] T. Inzerilli, A. M. Vegni, A. Neri, and R. Cusani, "A Cross-layer Location-based approach for Mobile-Controlled Connectivity," *International Journal of Digital Multimedia Broadcasting*, vol. 2010, p. 13, 2010.

[26] G. Mahardhika, M. Ismail, and K. Mat, "Multi-criteria Vertical Handover Decision in Heterogeneous Networks," in *Proc. of IEEE Symposium on Wireless Technology and Applications (ISWTA)*, pp. 1–4, September 23–26, 2012.

[27] J. Marquez-Barja, C. Calafate, J.-C. Cano, and P. Manzoni, "MACHU: A novel vertical handover algorithm for vehicular environments," in *Wireless Telecommunications Symposium (WTS)*, pp. 1–8, April 18–20, 2012.

[28] Y. Fang, I. Chlamtac, and Y. B. Lin, "Channel Occupancy Times and Handoff Rate for Mobile Computing and PCS Networks," *IEEE Trans. on Computer*, vol. 47, no. 6, 1998.

[29] A. H. Zahran, B. Liang, and A. Saleh, "Signal Threshold Adaptation for Vertical Handoff in Heterogeneous Wireless Networks," *ACM Mobile Networks and Applications*, vol. 11, no. 4, pp. 625–640, 2006.

[30] W. Ying, Y. Jun, Z. Yun, L. Gen, and Z. Ping, "Vertical Handover Decision in an Enhanced Media Independent Handover Framework," in *Proc. of IEEE Wireless Communications and Networking Conference (WCNC'08)*, pp. 2693–2698, March 31–April 3, 2008.

[31] P.-H. Ho, Y. Wang, F. Hou, and S. Shen, "A Study on Vertical Handoff for integrated WLAN and WWAN with Micro-Mobility Prediction," in *Proc. of 3rd International Conference on Broadband Communications, Networks and Systems (BROADNETS'06)*, pp. 1–11, October 2006.

[32] Y. Xiao and Y.-C. Kim, "Channel Allocation Scheme for Handoff Call based on Markov Chain with Hysteresis," in *Proc. of 1st International Conference on Information Science and Engineering (ICISE'09)*, pp. 2578–2581, December 2009.

[33] L. Wang, G. Min, and D. Kouvatsos, "Performance Analysis of a Dynamic Handoff scheme in Wireless Networks with Heterogeneous Call Arrival Processes," *Telecommunication Systems*, vol. 39, pp. 157–167, 2008.

[34] G. Cunningham, P. Perry, and L. Murphy, "Soft, Vertical Handover of Streamed Video," in *Proc. of 5th IEE International Conference on 3G Mobile Communication Technologies*, pp. 432–436, 2004.

[35] Y. Li, Y. Z. Cho, and W. Y. Jung, "Soft Vertical Handoff Scheme in Heterogeneous Networks," in *Proc. of 16th Joint Conference On Communication & Information*, 2006.

[36] R. B. Ali and S. Pierre, "On the Impact of Soft Vertical Handoff on Optimal Voice Admission Control in PCF-based WLANs loosely coupled to 3G networks," *IEEE Trans. Wireless. Comm.*, vol. 8, no. 3, pp. 1356–1365, 2009.

[37] S. G. Glisic, *"Advanced Wireless Networks – 4G Technologies"*. Wiley, 2006.

[38] P. K. Tang, Y. H. Chew, and C. O. M. Ling, "Improvement in the Grade-of-Service in a Cooperative Overlay Heterogeneous Network," in *Proc. of 6th International Conference on Information, Communications & Signal Processing*, pp. 1–5, December 2007.

[39] Y. Fang and Y. Zhang, "Call Admission Control schemes and Performance Analysis in Wireless Mobile Networks," *IEEE Trans. on Vehicular Technology*, vol. 51, no. 2, pp. 371–382, 2002.

[40] M. Puterman, *Markov Decision Processes: Discrete Stochastic Dynamic Programming*. New York, NY, USA: John Wiley & Sons, Inc., 1994.

[41] E. Stevens-Navarro, L. Yuxia, and V. W. S. Wong, "An MDP-based Vertical Handoff Decision Algorithm for Heterogeneous Wireless Networks," *IEEE Trans. on Vehicular Technology*, vol. 57, no. 2, 2008.

[42] F. Esposito, A.M. Vegni, I. Matta, and A. Neri, "On Modeling Speed-based Vertical Handovers in Vehicular Networks —Dad, slow down, I am watching the movie—," in *Proc. of IEEE Globecom 2010 Workshop on Seamless Wireless Mobility (SWM 2010)*, (Miami FL, USA), December 6–10, 2010.

[43] A. M. Vegni, *Multimedia Mobile Communications in Heterogeneous Wireless Networks*. PhD thesis, Department of Applied Electronics, Roma Tre University, March 8, 2010.