Load balancing algorithm by vertical handover for integrated heterogeneous wireless networks

Mohammad Ali Pourmina[*] and Navid MirMotahhary

Abstract

In this paper, an adaptive resource management scheme for hybrid WWAN/WLAN is proposed. Based on proposed joint velocity and average received power (ARP) estimation algorithms, a novel vertical handoff (VHO) for efficient load balancing in multitier network is developed. Simulation results show that proposed scheme achieves significant improvements over conventional schemes.

Keywords: mobility aware vertical handover, load balancing heterogeneous networks, power estimation

1. Introduction

Future wireless communication systems can be visualized as the integration of different radio access network (RAN) technologies, to provide always best connected. Heterogeneous wireless networks (HWN) will give the service provider, a chance to provide sufficient capacity, needed to support the temporally and spatially fluctuating traffic demands generated by mobile users. A practical benefit is that users can be served at lower cost and with the better quality of service (QoS). To support freedom of movement between HWNs and seamless roaming, several VHO management architectures and decision-making algorithms have already been proposed [1]. These will allow full exploitation of flexible HWN infrastructure, resources, and services. Although individual radio resource management (RRM) schemes can be tuned to optimally perform within their respective RANs, they may not efficiently perform in an HWN if the different RRM schemes are not properly managed. Hence, a major issue is how to jointly utilize the resources of the different RANs in an efficient manner while simultaneously achieving the desired QoS and minimizing the service cost from both user and service provider perspectives. In an integrated heterogeneous wireless network (IHWN), a mobile terminal (MT) is equipped with heterogeneous network interface, which is called the *multimode* terminal. When a multimode MT generates or originates a new call in an IHWN, it can select connections among different types of the IHWN

based on network selection strategies. An active multi-mode MT can also change its connections among different types of IHWN. Such a process is called the *vertical hand-off* (VHO). Traditional handoff algorithms are based on link quality or estimate of ARP. However, this measure is not sufficient for VHO, and other factors like mobile user velocity, network condition, and user preferences should be considered. Also because of complex structure of next-generation networks, more precise and sophisticated method for link measurements is required. In mul-tipath channel, received signal strength (RSS) is consisted of three different phenomena (path loss, shadow fading, and fast fading). Because of MT mobility, multipath effect, and shadowing, the RSS has fluctuations, which make raw signal strength an unstable criterion for triggering vertical and horizontal handoffs (HHOs). Shadowing, large-scale variation in path loss, is caused by obstacles in the propagation path between the MT and the base station (BS). The small-scale variation is due to the Doppler shift along the different signal paths and the time dispersion caused by the multipath propagation delays. As one primary indicator of channel quality, the power of the slowly varying shadow component is important for handoff decisions and power control. Most existing handoff algorithms assume that multipath fluctuations can be adequately filtered and base their handoff decisions on local mean power estimates [2,3]. Although these variations bring back uncertainty in the act of VHO decision making, they can be utilized to extract precious information about propagation environment and mobility behavior of mobile user [4]. In order to mitigate these variations in RSS, efficient smoothing

* Correspondence: pourmina@srbiau.ac.ir
Department of Electrical Engineering, Science and Research Branch, Islamic Azad University, Tehran, Iran

techniques must be considered. If the averaging interval is too short, fluctuations may not be effectively removed, or if the interval is too long, it may cause delay in handoff procedure, or in non-line of site (NLOS) scenarios, it can average out useful information of corner's positions. To fully exploit the capacity of the wireless channel, and to overcome ping-pong effect, an efficient power estimation method is required [5,6]. The ping-pong effect occurs if factors for VHO decision are changing rapidly and an MT performs the handoff as soon as it detects a more suitable BS. Because of heterogeneity, PHY and MAC layer of different IHWN are different, so a unified approach must be taken into consideration for collection of specific measures from different networks. As a result, more sophisticated VHO algorithm is required to extend the throughput of multilayer network and to increase efficiency of resource management for next generation of HWNs. In this paper, utilizing an accurate joint velocity and ARP estimation algorithms, a novel VHO algorithm is proposed which can effectively be used for load balancing and internetwork ping-pong effect reduction in HWNs. Also based on Markov model, an analytical model for performance evaluation of the VHO algorithm is proposed. This paper is organized as follows. Section 2 reviews related work on VHO algorithms. In Section 3, propagation model for an IHWN is introduced. In Section 4, proposed load balancing algorithm is discussed. VHO algorithm analysis framework is described in Section 5. The performance of the proposed VHO algorithm is analyzed through a theoretical model and simulations based on probability of blocking and probability of false network layer assignment in Section 6. Finally, paper is concluded in Section 7.

2. Previous works

Early works on VHO considered multitier homogeneous networks and used the RSS as the main factor of the hand-off decision-making process [7]. However, the VHO needs to be triggered considering a few more factors [8]. In [9], a VHO algorithm is proposed based on a assumption that a data call is kept in the higher bandwidth network as long as possible and voice calls are vertically handed over as soon as possible to avoid handoff delay. In [10], a network selection strategy that only considers mobile users' power consumption is introduced. To maximize the battery life, the mobile user selects the uplink or downlink that has the lowest power consumption from all of the available networks. In [11], a policy-enabled network selection strategy is proposed, which combines several factors such as bandwidth provision, price, and power consumption. By setting different weights over different factors based on the user's preference, a mobile user can connect to the most desired network. Reference [12] presents a signaling protocol for the exchange of information between a network management system and intelligent

multimodal wireless terminals in a heterogeneous environment; some preliminary measurement works were done mainly between WLANs, general packet radio service (GPRS), and digital video broadcasting for terrestrials (DVB-T), but the described algorithms were very simple, and it did not present how such redistribution would be performed in detailed steps. Efforts on standardization of the VHO operation can be review on [13-15]. Although the above network selection strategies have their own advantages, they did not put much attention on system performance, such as the blocking probability of originating calls and the forced termination probabilities of horizontal and VHO calls. They fail to address any ARP and also velocity estimation method in order to perform an accurate and seamless VHO. Also these have not considered corner effect and effect of low SNR in cell boundaries. In a homogenous environment, the ping-pong effect is a phenomenon that rapidly repeats HHOs between two BSs and can be mitigated by means of *dwell timer* (DT) or hysteresis margin [16]. In a heterogeneous environment, the ping-pong effect occurs if factors for the VHO decision are changing rapidly and an MT performs handoff as soon as the MT detects the better BS [17]. The DT scheme has been used to avoid such ping-pong effects due to the fact that RSS from HWNs is not comparable to each other [5,6]. The ping-pong effect can also occur if the MT's speed is high or its moving direction is irregular. Therefore, the proposed VHO algorithm balances the traffic load in each network based on efficient mitigation of inter-network ping-pong effect and also based on MT mobility behavior.

3. Propagation and noise model

The propagation model discussed here takes into account correlated multipath fading, correlated lognormal shadowing, and a distance dependant trend [18]. A discrete model (with the sampling rate of $1/2BW$) for the received signal (RS), $\gamma[n]$, is given by

$$\gamma[n] = \sqrt{s[n]}.r[n] + \eta[n] \tag{1}$$

where $r[n]$ is the complex envelop due to multipath propagation and user mobility, which contains the mobile's Doppler amplitude information and $s[n]$ is ARP (local mean) at the MT and $\eta[n]$ is AWG noise with zero mean and variance of σ_n^2 which is added to the RS. $r[n]$, $s[n]$, and $\eta[n]$ are mutually independent. $r[n]$ is defined by [4]

$$r[n] = \frac{1}{\sqrt{k}} \sum_{i=1}^{K} a_i e^{j(2\pi f_d \cos(\theta_i)n + \phi_i)} \tag{2}$$

where f_d is the Doppler frequency, θ_i and φ_i are mutually independent random variables uniformly

distributed over $(-\pi, \pi, a_i$ is the gain of ith scatter and K is the number of independent scatters; usually $K = 20$ is sufficient to provide good approximation). The process $s[n]$ is a wide-sense stationary lognormal random process, which contains distance-dependent trend and lognormal shadowing. We denote its mean and variance by μ_s and σ_s^2, respectively. Shadow fading process is assumed to have the exponential correlation function proposed by [19] based on the measured autocovariance function of $s[n]$ in urban environments. Path loss, $(\mu_s(d))$ which is a mean of $s[n]$, decreases monotonically with distance from the BS. Let d_c be the distance between BS and intersection at which MT makes a turn. Following [4], dimensionless parameter x_0, the distance parameters x_c, y_0, y_c and exponents ζ, η, χ are introduced. Corner effect could cause ΔS dB signal drop, in y_0 meters. P_0 is a constant that accounts for transmitted power and antenna gain. Path loss for microcellular structure at position d is modeled by

$$\mu_s(d) = \begin{cases} P_0 - 20\log_{10}\left(\dfrac{d}{x_0}\right) \\ \quad -\dfrac{10}{x}\log_{10}\left(1 + \left(1 + \left(\dfrac{d}{x_c}\right)^{(\zeta-2)\chi}\right)\right), 0 < d_1 < d_c \\ P_L(d_1)10^{\dfrac{-\Delta S \cdot (d)}{10\gamma_0}} \qquad d_c < d < y_0 \\ P_L(d_2) - 20\log_{10}\left(\dfrac{'d}{\gamma_0}\right) \\ \quad -\dfrac{10}{x}\log_{10}\left(1 + \left(1 + \left(\dfrac{d}{\gamma_c}\right)^{(\eta-2)_\chi}\right)\right), d_2 < y_0 \end{cases} \qquad (3)$$

To suppress noise and interference terms, $\gamma[n]$ is passed through a unit-gain, square-root raised cosine lowpass filter with a bandwidth $BW > f_{\max}$, since we are only interested in the Doppler power spectrum, which is narrowband and variable $(0 - f_{\max}(Hz))$ in microcellular structure. $f_{\max}$ is maximum possible Doppler frequency of channel. An example of RSS in a microcellular environment is plotted in Figure 1, for variable mobile speed when long-term SNR is 20 dB. As it is seen, short-term SNR is high near base stations. Long-term SNR is plotted for 100 s observation.

4. Proposed traffic load balancing algorithm
Knowledge of MT's position and velocity plays an important role on offering efficient network controlling mechanisms and variety of offered services in IHWNs. Mainly in IHWNs structure, WLANs have less coverage than WWAN, so a reliable mobility tracking algorithm is desirable to reduce the number of handoffs and waste of bandwidth due to unnecessary signaling. Researches about MT speed estimation are divided into two different groups. First one uses statistics of RSS from different BSs

measured at an MT and second one uses corresponding propagation times called cell sojourn time (CST) in order to show if user speed is slow, medium, or fast. Both categories are subjected to strong irregular variations caused by Rayleigh fading and shadowing [17,20]. Many of fading distribution property (FDP)-based methods give an accurate estimate in noiseless environments, but in noisy environment, results are unreliable [17,20]. In urban area which is modeled as Manhattan structured microcell, speed estimation is more complicated, due to complexity introduced by severe fading and noise. On the other side, although the CST-based methods work well in noisy area [17], but these algorithms lose their accuracy when they are used on highly dense urban area with variety of building structures. Because CST-based methods calculate MT speed by comparing the CST with a predefined time threshold, a manoeuvring user with variable speed might have more (less) actual speed than what is estimated based on comparing CST with predefined threshold. High MT mobility in cell borders is another issue which can result in many cell border crossings. The sojourn timer resets every time MT crosses the cell border, thus a slow-moving MT that is repeatedly crosses cell border is considered as a fast-moving MT by CST-based schemes [17]. In addition, CST-based algorithms fail to estimate variable speed. When variable speed MT is circulating around the cell boundary, CST-based algorithms classify MT as a slow-moving MT. CST-based speed estimation algorithms require estimation and calculation of statistical properties of CST in coverage area [17]. In order to calculate probability and cumulative distribution functions (PDF and CDF) of CST, while taking into account presence of fast fading, shadowing, corner effects, and uncertainty region between neighboring BSs, it is required to perform simplifications and assumptions, which results in lack of generality. Hence, an FDP-based method for MT speed estimation is proposed in this paper. A simple type of window-based ARP estimators, namely weighted sample average estimators of local mean power, is currently deployed in many commercial communication systems, and various other window-based estimators (WBEs) have been proposed in [21]. These WBEs work well under the assumption that the shadowing is constant over the duration of the averaging window, and in this case, their performance improves as the window size increases. In practice, however, the shadow process varies with time (albeit slowly relative to the fast-fading process), and this variation must be considered since both analysis (developed herein) and experiment show that the mean square error (MSE) performance of these WBEs deteriorates severely when the window size increases beyond a certain value. For variable speed, the observation window must be adapted constantly, and the rate of adaptation depends not only on the MT speed but also on the sampling period and shadow fading

Figure 1 Simulated RSS for variable mobile speed, long-term SNR = 20 dB for noisy RSS.
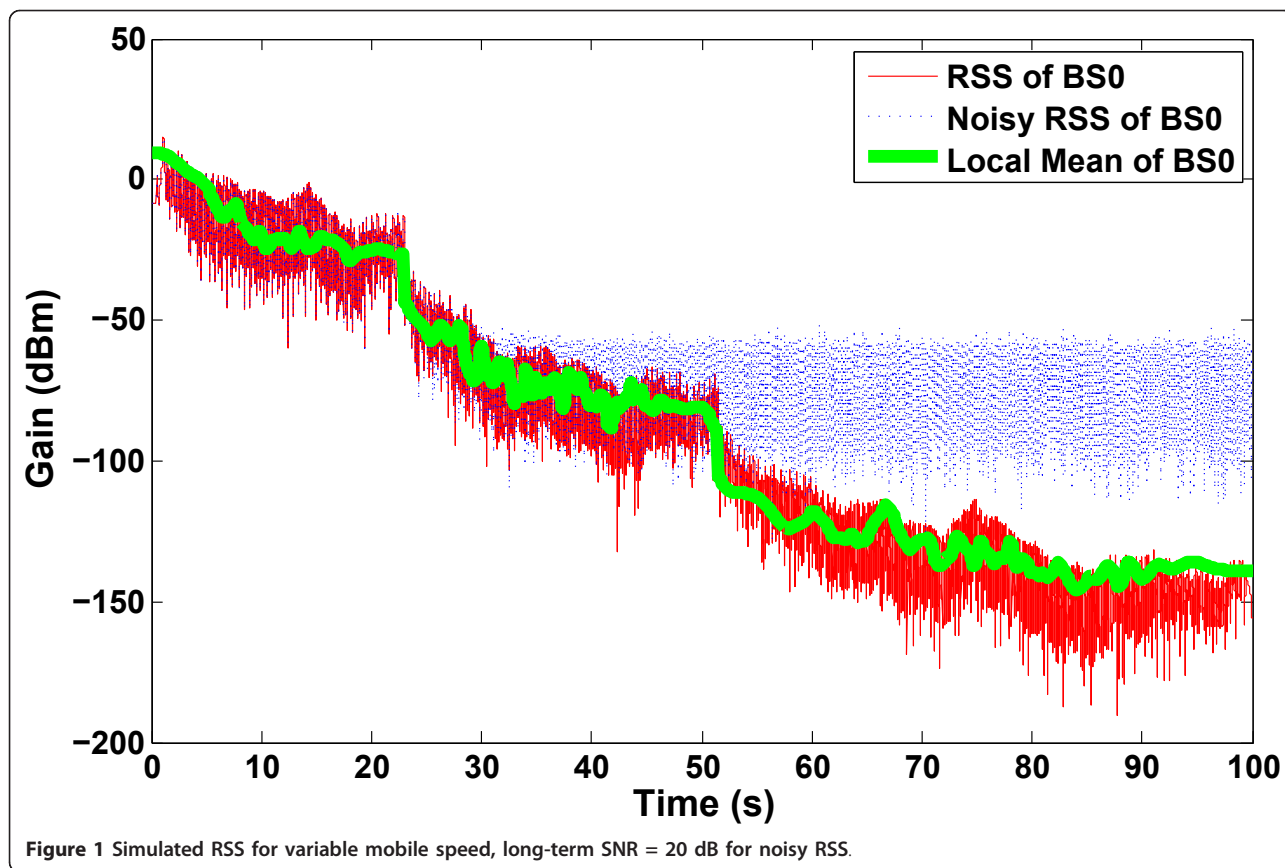
characteristics. In particular, errors in the estimates could propagate due to suboptimal observation windows. A joint velocity and power estimator are proposed in this paper to calculate mobility behavior of ma-noeuvering MT in dense urban area also to mitigate fluctuations of RSS in order to minimize number of VHOs and simultaneously assign different service requests and MTs efficiently to different networks in a IHWN. This in turn can minimize probability of blocking and probability of false network assignment.

4.1. Joint velocity and average received power estimation

In HWNs in order to reduce number of VHOs state of velocity plays an important role in classification of MTs, therefore our proposed method in [4] is modified in such a way that state of velocity is determined in real time for act of VHO decision making. Due to time-varying nature of communication channel, the signal properties include (amplitude, frequencies, and phases) will change with time. Thus, we utilized time-dependent Fourier transform also referred to as the short-time Fourier transform to estimate power spectral density (PSD) of RS by using DFT of finite-length segments of the RS, obtained by banks of rectangular filters such that, each filter has different duration. The N-DFT of segmented RS, $\gamma_i[n]$, is

$$V_i[k] = \sum_{n=0}^{N-1} w_i[n]\gamma[n]e^{-j\frac{2\pi n}{N}k} \tag{4}$$

where $w_i[n]$ is the ith window with the length L_i. An estimate of PSD (called periodogram) would be

$$\mathrm{PSD}_i(\omega_k) \approx \frac{1}{L_iF}\{|V_i[K]|^2\} \tag{5}$$

where $\omega_K = \frac{2\pi K}{N}$ for $K = 0, 1,..., N - 1$. It is well known that fast-moving MTs cause high Doppler spread while slow-moving MTs cause low Doppler spread. Therefore, the shift in local maxima of estimated periodogram (PSD) is related to maximum Doppler frequency, which is proportional to the mobile speed ($\hat{v} \propto \hat{f}_d.\lambda$) in which $\hat{v}_i$ is estimated speed under ith window and λ is wavelength [4].

$$\hat{v}_i \propto \arg\underset{\omega}{\mathrm{Max}}\{\mathrm{PSD}_i(\omega)\} \tag{6}$$

As it is seen in Figure 2, increase in velocity can be interpreted as a Doppler spread in frequency domain. Thus, the mobile velocity can be obtained from estimated Doppler spread. First, the estimated PSD is

**Figure 2 Integrated heterogeneous wireless network**.
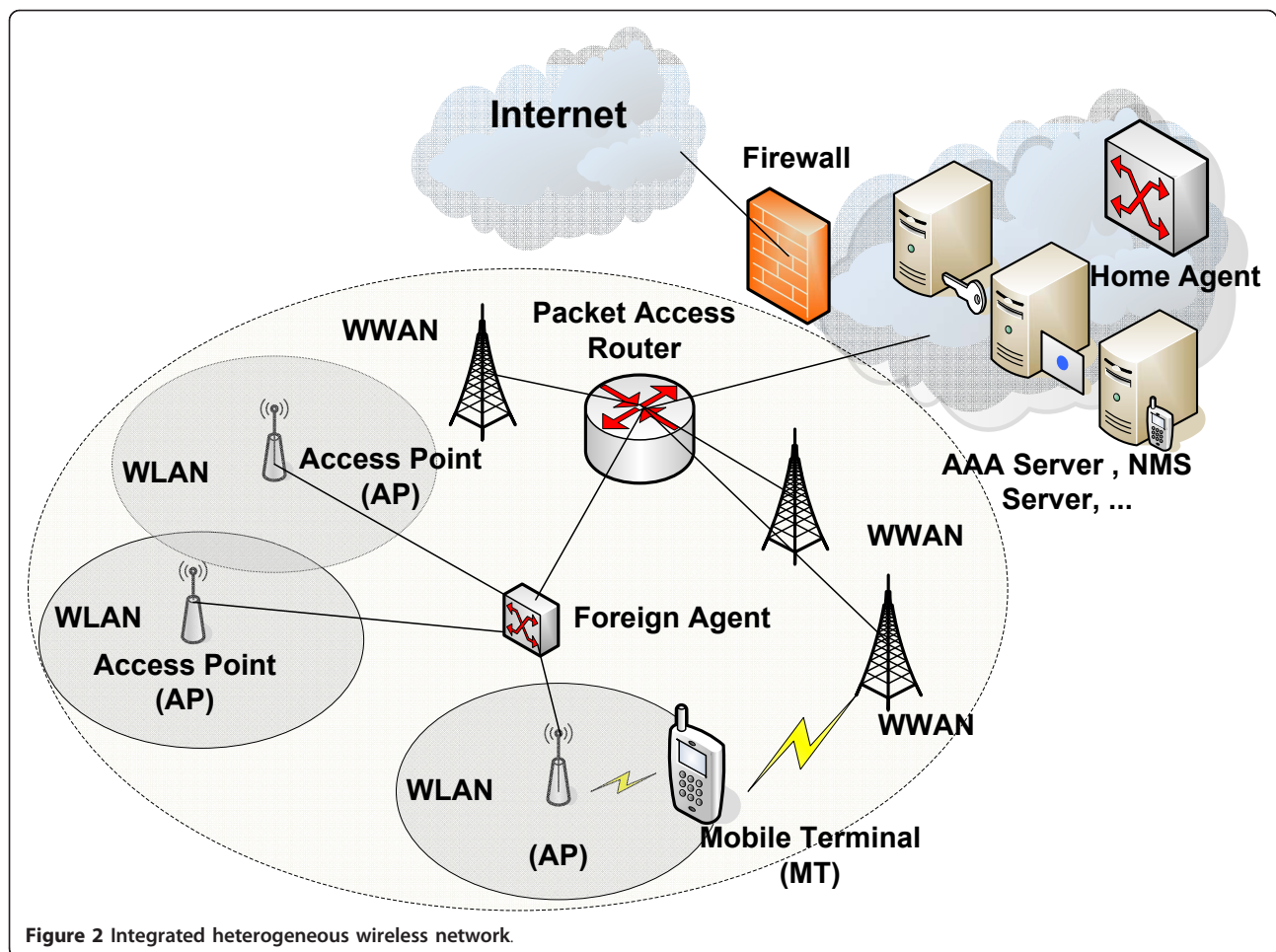
normalized to estimate $\hat{f}_d$ then due PSD symmetry, PSD is folded (folded PSD is plotted in Figure 2). Bandwidth of PSD is detected which is proportional to Doppler spread. In order to find the relative velocity of user, we classify user's mobility model into two classes, pedestrian and fast. As a result, only two segments (subsets) of frequency domain of estimated PSD are going to be considered as regions of interest, and $\hat{f}_d$ extraction is done only in these two segments. In proposed algorithm, maximum is searched only in pedestrian subset of estimated PSD of RS, which limits search space and as a result increases estimation speed. The ARP (RS local mean) is estimated based on a fact that slow fading and path loss have slow variations while MT is manoeuvering in coverage area; therefore, they are present only in DC component of the estimated PSD of the RS. For manoeuvering MT, the duration of observation window must be constantly adapted, and the rate of adaptation is critical on performance of speed and power estimators [18]. DC component of estimated PSD is adaptively extracted from different filters

$$\hat{S}_i \approx \frac{1}{L_i \Delta t F} |V_i(e^{j\omega})|^2 \mid_{\omega=0} \approx \frac{1}{L_i \Delta t F} \left| \sum_{n=0}^{L_i-1} w_i[n]P[n] \right|^2 \quad (7)$$

where $\hat{s}_i$ is the estimated ARP, L_i is window length, Δt is sampling interval, and F is normalization factor [4]. Active smoothing window is switched to another window in which duration is selected proportional to inverse of estimated velocity for any iteration. The schematic of joint ARP and velocity estimator is shown in Figure 3. Bias and variance of proposed local mean estimator power are given in Appendix A.

4.2. Load balancing by VHO algorithm

As it is mentioned in previous section, due to heterogeneity of integrated networks, measurements of each network cannot be compared directly to other networks, so hysteresis method [10,11] cannot be utilized in VHO algorithm. Thus, thresholding technique is considered in this paper. In HWN structure shown in Figure 4, each network shall have its own thresholds. Also due to the fact that requirements of load balancing in each network

Figure 3 Folded PSD which shows Doppler spread for different velocities.
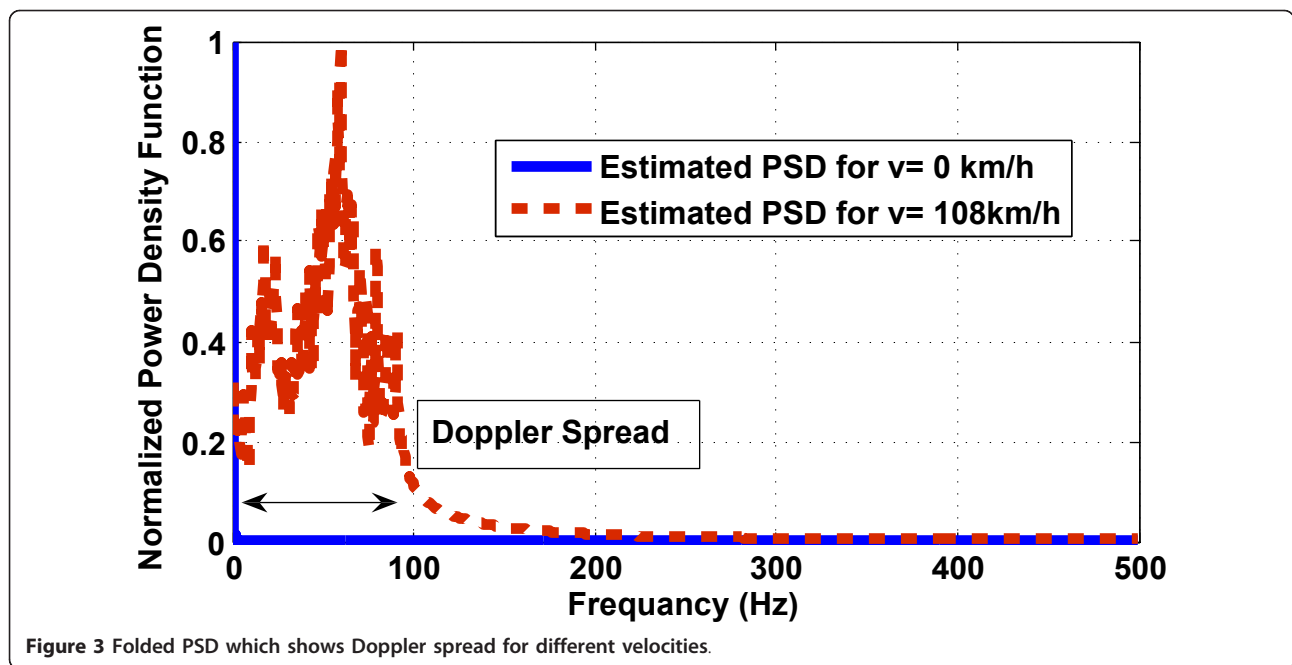
are different, so proposed algorithm is divided into 2 sections, (1) Downward VHO from WWAN to WLAN and (2) Upward VHO from WLAN to WWAN. In defined scenario, HHOs are only between BSs of WWAN network due to limitations of coverage that WLAN access points have. HHO algorithm utilizes joint velocity and ARP estimator in the act of HHO decision making. Proposed HHO is based on using hysteresis margin method [22] in addition to DT which is calculated based on inverse of velocity estimates. Block diagram of proposed VHO algorithm is shown in Figure 5. First, velocity and RS local mean are estimated from raw RSS. Values and thresholds used in the VHO decision making are also calculated and updated periodically. The proposed load balancing algorithm selects a network based on specifications of each network. Algorithm 1 illustrates VHO algorithm. According to Algorithm 1 in order to reduce the effect of velocity estimation error on accuracy of VHO decision making,

average of the all the velocity estimates during dwell time period is considered as a velocity estimate to classify MT as pedestrian or fast. The idea of utilizing 2 thresholds for adding MT to the list of upward VHO candidates solves the problem of lack of accuracy in scenarios that slow manoeuvering MT are bouncing inside and outside of cell boundary. Benefits gained from knowledge of user real velocity are to reduce ping-pong effect caused by mobility behavior of MTs, and idea of using adaptive DT $((t_{\mathrm{DT}} \propto 1/\hat{v}))$ velocity lack of ability to use hysteresis methods for HWNs and in addition to novel proposed ARP method mitigates ping-pong effect caused by fluctuations and RSS due to fading.

5. VHO algorithm analysis
Due to the fact that in VHO (HHO) algorithms, decision making on target network (cell) is based on present and previous step states, Markov model can be used for modeling and calculation of probability of transition in
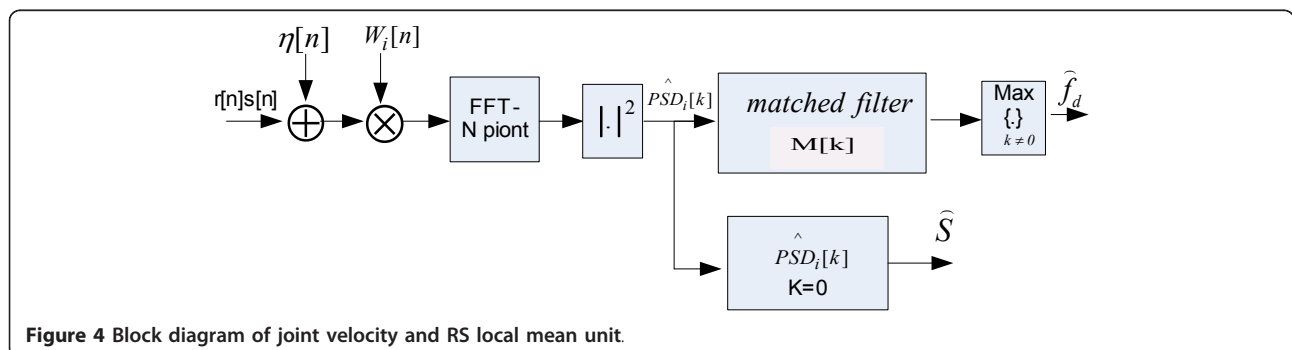

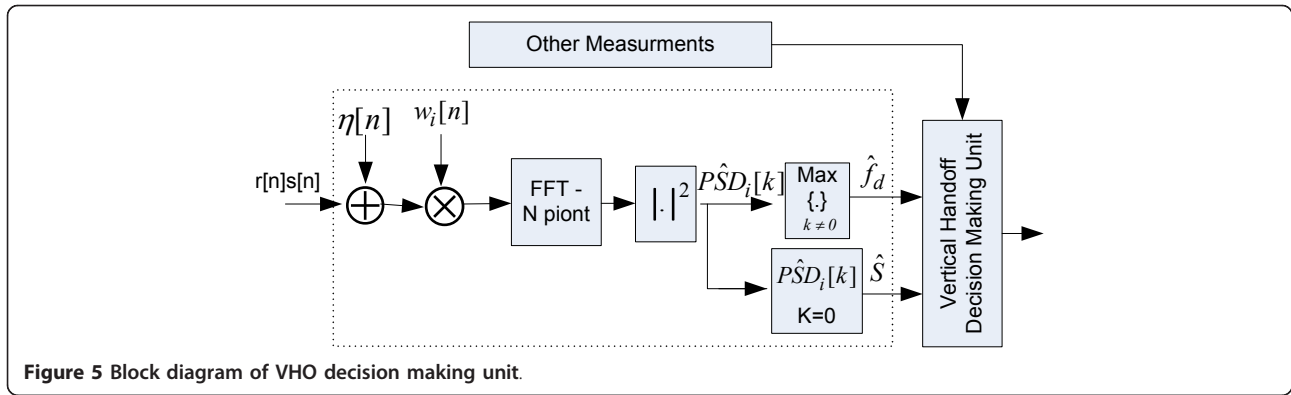**Figure 4 Block diagram of joint velocity and RS local mean unit**.

Figure 5 Block diagram of VHO decision making unit.

each state. For the given system model of Figure 4, a user in the system can be in one of the N_{states} defined as follows

$$N_{\text{state}} = 1 + \sum_{i=0}^{N_{\text{network}}-1} N_{\text{network}} - i \qquad (8)$$

In case of 2 integrated networks, we have

- State0) The user has no active session in progress and is not occupying any channel (i.e., the user is idle), independent of its location.
- State1) The user in the WLAN coverage area is occupying the WWAN resource.
- State2) The user in the WLAN coverage area is occupying the WLAN resource.
- State3) The user in the out of WLAN coverage area is occupying the WWAN resource.

Figure 5 shows the user state transition diagram for the 2 network system models, where $P_{si}[k]$ shows ith steady-state probability at kth moment and $P_{ij}[k]$ denotes the transition probability from state i to state j, $i, j \in 0, 1, 2,..., N_{\text{state}}$, and state transition probability matrix **P** is given by

$$\underbrace{\begin{bmatrix} P_{s_0}[k] \\ P_{s_1}[k] \\ P_{s_2}[k] \\ P_{s_3}[k] \end{bmatrix}}_{\pi[k]} = \underbrace{\begin{bmatrix} P_{00}[k] & P_{01}[k] & P_{02}[k] & P_{03}[k] \\ P_{10}[k] & P_{11}[k] & P_{12}[k] & P_{13}[k] \\ P_{20}[k] & P_{21}[k] & P_{22}[k] & P_{23}[k] \\ P_{30}[k] & P_{31}[k] & P_{32}[k] & P_{33}[k] \end{bmatrix}}_{P} \cdot \underbrace{\begin{bmatrix} P_{s_0}[k-1] \\ P_{s1}[k-1] \\ P_{s2}[k-1] \\ P_{s_3}[k-1] \end{bmatrix}}_{\pi[k-1]} \quad (9)$$

Based on Algorithm 1, a user can change state from state0 to any nonzero state when a new connection is made. If a session is completed while a user is currently in a nonzero state, the user changes to state0. The Markov chain depicted in Figure 5 is irreducible and aperiodic, and all the states are recurrent nonnull, so that the equilibrium state probabilities can be determined by solving the (9), subject to the normalization condition

$\sum_{i=0}^{N_{\text{state}}} P_{si} = 1$; however, the expressions for the transition probabilities remain to be determined; these form the topic of discussion in the succeeding sections.

5.1. Service model
The service model relies on four assumptions.

1. Due to wide coverage that WWAN has, we assume that probability of MT being inside the coverage area of WWAN, $P_{\text{WWAN}} = 1$ and as the coverage area of WLAN is subset of WWAN; it is clear that $P_{\text{WLAN}} < 1$.
2. New calls arrive in the macrocell and microcell according to a Poisson process with mean arrival rates λ_{WW} and λ_{WL}, respectively. A new call is randomly determined as the real time (RT) and non-RT (NRT) calls with probabilities P_{rt} and P_{nrt}, respectively. Similarly, a call is also independently determined as a low (high)-bandwidth application with probability $P_{\text{bwl}}(P_{\text{bwh}})$. Clearly, $P_{\text{rt}} + P_{\text{nrt}} = 1$ and $P_{\text{bwl}} + P_{\text{bwh}} = 1$

Algorithm 1 Proposed VHO algorithm
1: **loop**
2:    **if** $MS_i \in \{MS_i \mid Connected\ to\ WWAN\}$ **then**
3:      **if** $\hat{S}_{WLAN} \geq T_{r_{WLAN}}\ (dB)\}$ **then**
4:        $MS_i \in \{Downward\ VHO\ probability\ set\}$
5:        **if** $\hat{v} \leq V_{th}$**then**
6:          $MT_i \in \{Pedestrain\ Class\}\ and\ MT_i \in$ $\{VHO\ Candidate\ set\}$
7:          $Start(t_{DT})\{t_{DT} \propto 1/\hat{v}\}$
8:          **if** $\hat{v} \leq V_{th}$ untill the timer expires **then**
9:            at the end of

$$t_{DT}\hat{v} = \sum_{i=1}^{n} \frac{\hat{v}_i}{n}(t) \quad 0 \leq t \leq t_{DT}$$

10:          $MS_i \in \{VHO\ Active\ Set\}$
11:          VHO to target Network $VHO(targ_{network})$
12:        **else**

13:                     Reset Dwell Timer {Reset (t_{DT})}
14:             **end if**
15:         **else**
16:                 $MS_i \notin \{VHO\ Candidate\ set\} \rightarrow$ Unbeneficial

                VHO (MT Stays in Current Serving Network)
17:         **end if**
18:     **end if**
19: **else if** $MT_i \in \{MT\ is\ Connected\ to\ WLAN\}$ **then**
20:     **if** $\hat{S}_{WLAN} \leq T_{rWLAN_{Add}}(dB)$ and
        $\hat{S}_{WWAN} \geq T_{r_{WWAN}}(dB)$**then**
21:         $MT_i \in \{Upward\ VHO\ probability\ set\}$
22:         $Start(t_{DT})\{t_{DT} \propto 1/\hat{v}\}$
23:         **if** $\hat{v} \geq V_{th}$ untill the timer expires or
            $\hat{S}_{WLAN} \leq T_{r_{WLAN_{Drop}}}(dB)$**then**
24:             $MT_i \in \{VHO\ Active\ Set\}$
25:             VHO to target Network $VHO(targ_{network})$
26:         **else**
27:             Reset Dwell Timer {Reset (t_{DT})}
28:         **end if**
29:     **end if**
30:   **end if**
31: **end loop**

3. Call duration T_{call} is exponentially distributed with a mean of $1/\mu$, where μ is the average call completion rate. Hence, the call completion (termination) probability $P_{term} = P(T_{call} \leq T_{th})$, where T_{th} is the time unit for the user state transition diagram, as shown in Figure 5.

4. From $\sum_{i=0}^{N_{state}} P_{si} = 1$, it is clear that for scenario depicted in Figure 5. $P_{WWAN} = \sum_{i=0}^{N_{state}} P_{si}$ and $P_{WLAN} = P_{s0} + P_{s1} + P_{s2} = 1 - P_{s3}$.

based on the service model, the transition probabilities are $P_{10} = P_{20} = P_{30} = P_{term}$. For exponentially distributed interarrival time T_{ar} with a mean of $1/\lambda_n (\lambda_n = \lambda_{WW} + \lambda_{WL})$, the new call arrival probability within the next time unit is $P_{new} = P(T_{ar} \leq T_{th})$. Also, it is clear that $P_{10} = P_{new\ arrival}$. The probability of a new call arrival in the WLAN area $P_{new\ WL}$ and the probability of a new call arrival in the WWAN area $P_{new\ WW}$ can be written as follows

$$P_{new\ WL} = P_{new} \times P_{WLAN} \tag{10}$$

and

$$P_{new\ WW} = P_{new} \times P_{WWAN} \tag{11}$$

5.2. State transition probabilities for VHO

As it is shown in Figure 5, VHO probability can be determined by state probabilities crossed by VHO index line

$$P_{VHO}[k] = P_{s_1}[k-1]P_{12}[k] + P_{s_2}[k-1]P_{21}[k]$$
$$+ P_{s_2}[k-1]P_{23}[k] + P_{s_3}[k-1]P_{32}[k] \tag{12}$$

According to list of states, each state consists of set of events (S_i)(e.g., each network has $N_{cell i}$ BSs, so probability of being in any state can be determined based on probability of being connected to any of cells inside network coverage area). Hence,

$$P_{S_i}[k] = \sum_{l=1}^{N_{Cell i}} P_{s_{il}}[k] \tag{13}$$

where $S_i = \{S_{i1}, ..., S_{iN_{Cell_i}}\} \{\forall i, l | i \in \{1,...N_{state}\}$ and $l \in \{1,..., N_{cell i}\}\}$, and $P_{sil}[k]$ is probability of occupying resources of tth BS from ith network at the kth moment. State transition probabilities can be defined as follows

$$p_{ij}[k] = P\{S_i[k]|S_j[k-1]\} = \sum_{l=1}^{N_{cell}} P\{S_{il}[k]|S_{jl}[k-1]\} \tag{14}$$

each $P\{S_{il}[k]|S_{jl}[k-1]\}$ will be determined based on predefined VHO algorithm. Before further discussion on how to determine conditional probability of transition between events in HWN, it is worth mentioning that based on scenario, some simplifications are possible in transition matrix (**P**) determination. For instance, state transition from state3 to state2 is not feasible directly, because if a user out of WLAN coverage area intends to change state to state2 which is occupying resources of WLAN, first it must change state to state1 which is entering into coverage area of WLAN and then it would be feasible to switch to state2 so for any k, we have

$$P_{32}[k] = P\{S_{21}[k]|S_{31}[k-1]\}$$
$$= P\{S_{22}[k]|S_{32}[k-1]\} = P\{S_{23}[k]|S_{33}[k-1]\} = 0 \tag{15}$$

because in proposed VHO algorithm, upward and downward VHO are different; thus for each scenario, equations must be calculated separately. VHO criteria (estimated velocity and ARP) are prepared according to Figure 6, and clearly, velocity and ARP estimates are independent. Event of presence of user in state1 and downward VHO to WLAN at the kth moment is

$$\{S_2[k]S_1[k-1]\}$$
$$= \{\hat{S}_{WLAN}[k] > T_{WLAN}, V[k] \leq V_{th}|S_1[k-1]\}$$
$$= \{\hat{S}_{WLAN}[k] > T_{WLAN}, |S_1[k-1]\} \times \tag{16}$$
$$\{V[k] \leq V_{th}|S_1[k-1]\}$$
$$= \{\hat{S}_{WLAN}[k] > T_{WLAN}|S_1[k-1]\} \times \{V[k] \leq V_{th}\}$$
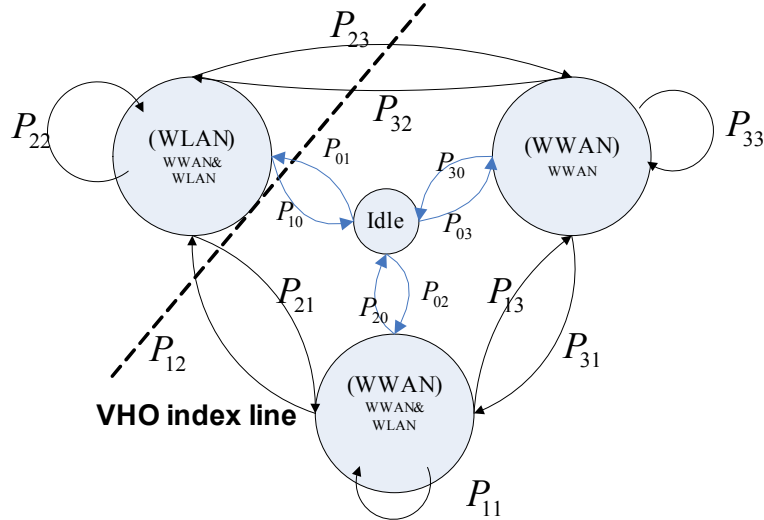
Figure 6 State transition diagram for $N_{\text{network}} = 2$.

and corresponding state transition probability is as follows

$$P_{12}[k] = P\left\{S_2[k]|S_l[k-1]\right\}$$
$$= P\left\{\hat{S}_{\text{WLAN}}[k] > T_{\text{WLAN}}|S_l[k-1]\right\} \times P\{V[k] \leq V_{\text{th}}\}$$
$$= \frac{P\left\{\hat{S}_{\text{WLAN}}[k] > T_{\text{WLAN}}|S_l[k-1]\right\}}{P\{S_l[k-1]\}} \times P\{V[k] \leq V_{\text{th}}\}$$
(17)

where $P\{S_l[k-1]\}$ is given by

$$P\{S_l[k-1]\} =$$
$$P\left\{\hat{S}_{\text{WWAN}}[k-1] > Tr_{\text{WWAN}}, \hat{S}_{\text{WLAN}}[k-1] \leq Tr_{\text{WLAN}}\right\}$$
$$\times P\{V[k-1] > V_{\text{th}}\}$$
(18)

It should be noted that because RSS from different networks cannot be compared to each other, they can be considered as independent variables. Hence,

$$P_{12}[k] = P\{V[k] \leq V_{\text{th}}\}$$
$$\times \frac{P\left\{\hat{S}_{\text{WLAN}}[k] > T_{\text{WLAN}}, \hat{S}_{\text{WLAN}}[k-1] \leq T_{\text{WLAN}}\right\}}{P\{\hat{S}_{\text{WLAN}}[k-1] \leq T_{\text{WLAN}}\}}$$
(19)

After a few simplifications, upward VHO to WWAN at the kth moment is given by

$$P_{23}[k] = P\left\{\hat{S}_{\text{WWAN}}[k] \geq T_{\text{WWAN}}\right\}$$
$$\left[\frac{P\left\{\hat{S}_{\text{WLAN}}[k] \leq T_{\text{Add,WLAN}}, \hat{S}_{\text{WLAN}}[k-1] \geq T_{\text{WLAN}}\right\}}{P\left\{\hat{S}_{\text{WLAN}}[k-1] \geq T_{\text{WLAN}}\right\}}\right.$$
$$\times P\{V[k] > V_{\text{th}}\}$$
$$+ \frac{P\left\{\hat{S}_{\text{WLAN}}[k] \leq T_{\text{Drop,WLAN}}, \hat{S}_{\text{WLAN}}[k-1] \geq T_{\text{WLAN}}\right\}}{P\left\{\hat{S}_{\text{WLAN}}[k-1] \geq T_{\text{WLAN}}\right\}}\right]$$
$$\times P\{V[k] \leq V_{\text{th}}\}$$
(20)

Other transition probabilities can be determined likewise. To calculate a close form for transition probabilities, it is required to calculate joint and conditional PDF of RSS which is given in Appendix B.

6. Performance evaluation

This section presents numerical results that demonstrate the network-level performance of an HWN comprising mobile WIMAX and IEEE 802.11 WLAN RANs. Probability of blockage in WLANs and WWANs and probability of false load balancing based on velocity and ARP estimates are considered as performance criteria. Users are distributed uniformly in simulation area (Manhattan area). Mobility behavior of MTs is considered according to mobility profile proposed in [4]. Performance of proposed ARP estimation algorithm is compared to WBEs, like sample average (SA), uniformly minimum variance unbiased (UMVU) estimator, and Kalman-based method discussed in [21]. As it is seen in Figure 7, proposed local mean power estimator shows less MSE than window-based methods, although Kalman-based method shows better performance but requires large number of transmitted messages and heavy calculations so that signal strength variations can be filtered out. In Figure 8, estimated local mean power based on proposed method is plotted. Performance of proposed load balancing algorithm has been investigated in the case that total call arrival rate and the ratio of the fast to the slow traffic vary. For example, assume that the total call arrival rate to a WWAN and its embedded WLANs is 1,500 calls/h, and the ratio of the fast traffic to the slow traffic is 2:1. Then, the call arrival rate of fast traffic is 1,000 calls/h and the slow traffic is 500calls/h. The traffic load on the
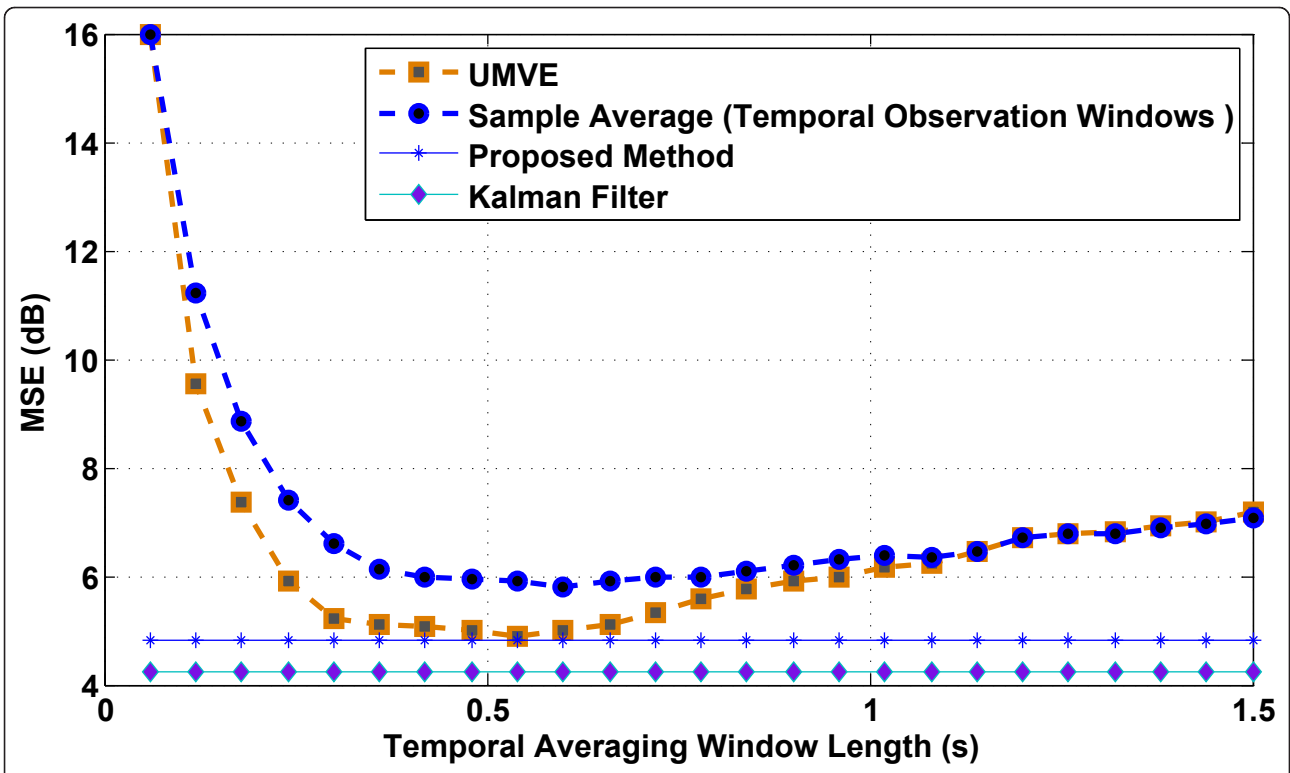
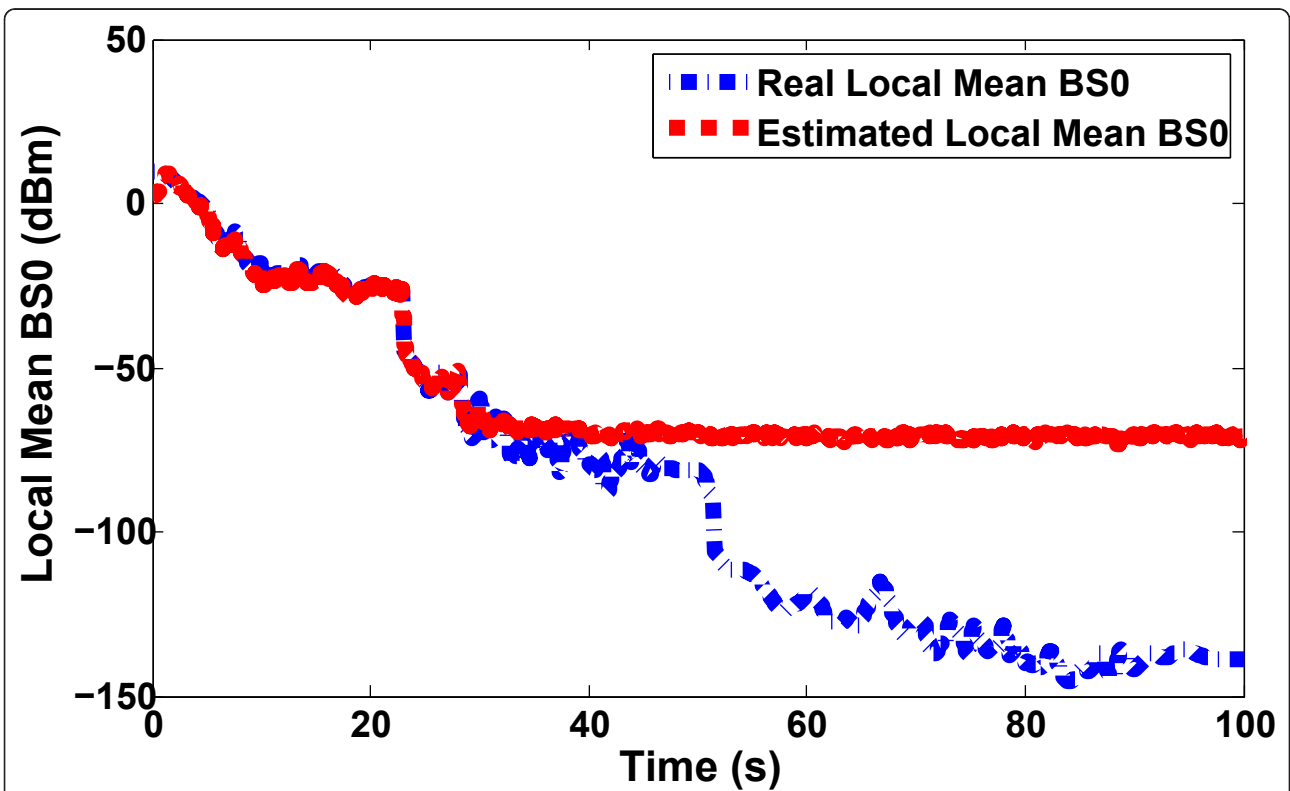**Figure 7** Local mean power estimation results.

Figure 8 MSE of local mean estimation versus averaging window length, v = 20 km h^{-1}, X_c = 10 m, shadowing variance $\sigma_s^2 = 6$ dB, LOS scenario.

WWAN is 140 calls/h, and the load on each WLAN is 136 calls/h, which consists of 86 calls/h for fast and 50 calls/h for slow. From these assumptions, the actual call arrival rates to the WWAN and each WLAN are given by

$$\lambda_{WW} = 140 \left(1 - P_{eWL} + 860 P_{eWW}\right)$$

and

$$\lambda_{WL} = 364 P_{eWL} + 86 \left(1 - P_{eWW}\right) + 50 \tag{21}$$

where P_{eWW} is the probability that slow user is allocated to WWAN, which gives arise to shortage in macrocell capacity, and P_{eWL} is probability that the fast user is allocated to WLAN, which causes increase in number of hand-offs. Speed threshold is considered V_{th} = 5 m/s. As it is demonstrated in Figure 9 and 10, in both LOS and NLOS scenarios, due to efficient load balancing based on mobility behavior of MTs and network conditions, probability of false VHO to WWAN and WLAN (i.e., two consecutive VHOs within a short duration for a roaming call in the hot-spot area or a very short WLAN session before the user moves out of the hot-spot area) is efficiently reduced. Effect of low SNR on proposed method is plotted in Figure 11 and 12 for upward and downward VHO, respectively. Because of robustness of proposed method even in low SNR scenarios, load balancing is done effectively. In Figure 13 and 14, blocking probability is plotted versus different call arrival rate (from 500 calls/h to 2,000 calls/h) for each layer with λ_n = 0.4 and $1/\lambda$ = 120s. Figure 13 shows that blocking probability is effectively decreased in WWANs due to more efficient load balancing, and as a result, blocking probability is increased in WLAN (Figure 14), which shows more WLAN resource utilization due to more efficient load balancing. For better demonstration, blocking probability in WLAN and WWAN is plotted in Figure 15 and 16 versus specific arrival traffic with variable ratio of slow traffic, respectively. As it is shown in Figure 15, when number of slow MTs is increased, blocking probability is decreased in WWAN and as a result for WLAN as it is plotted in Figure 16, there is a small increase in blocking probability.

7. Conclusion

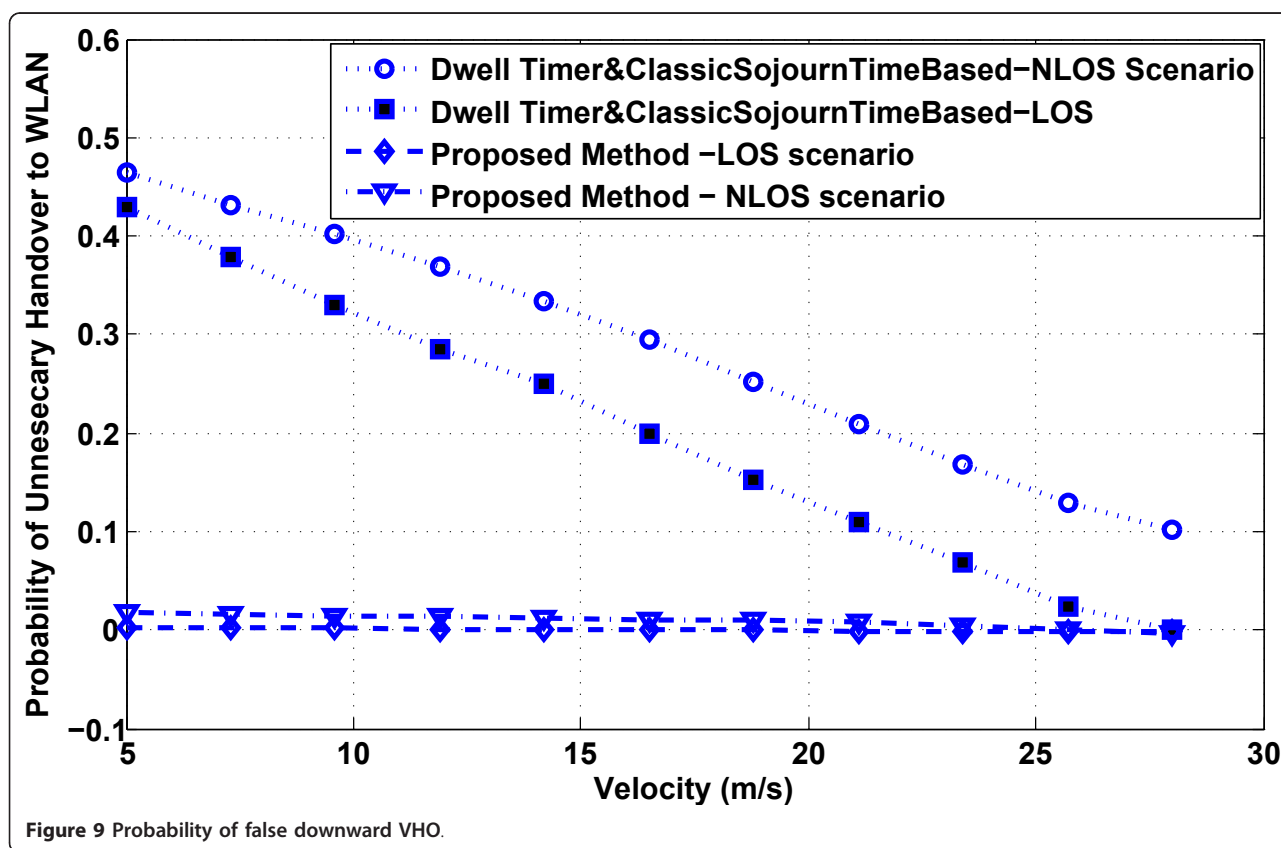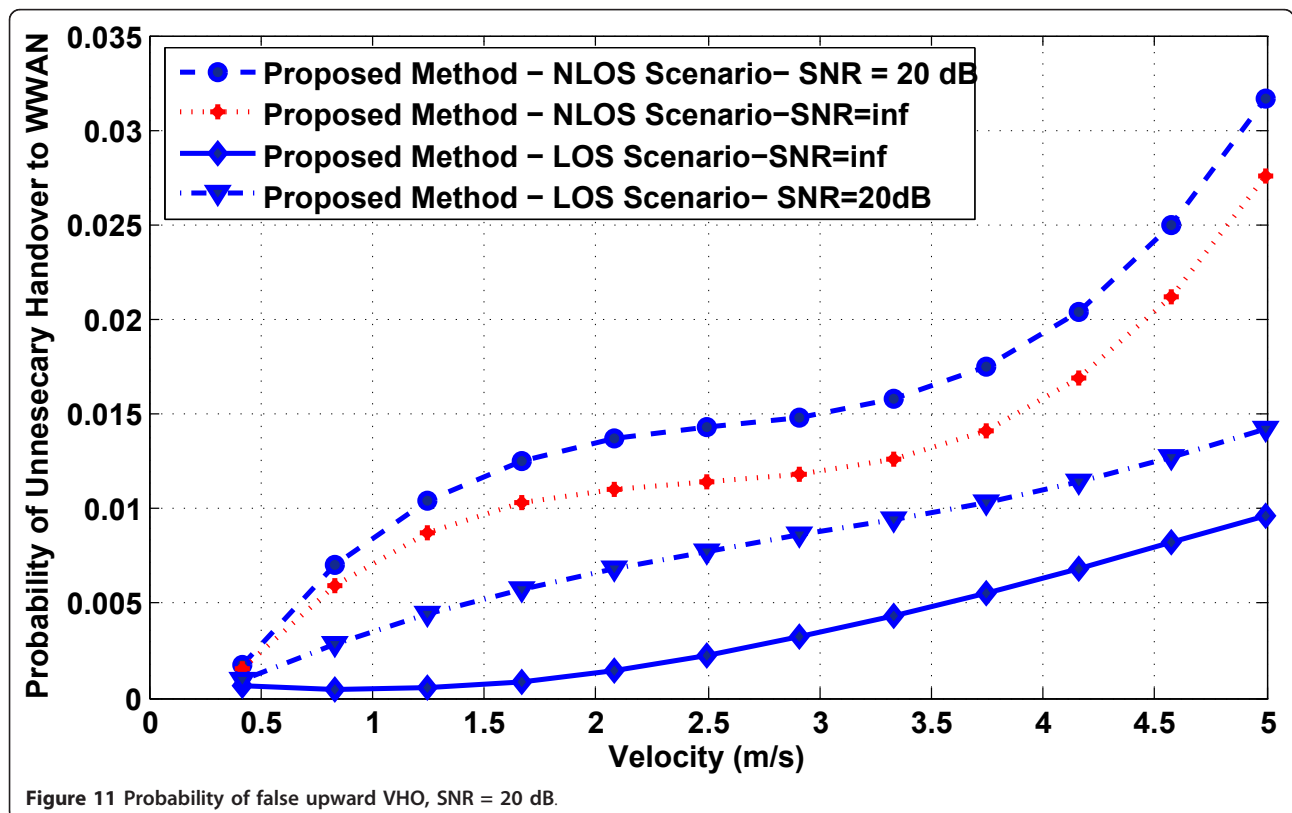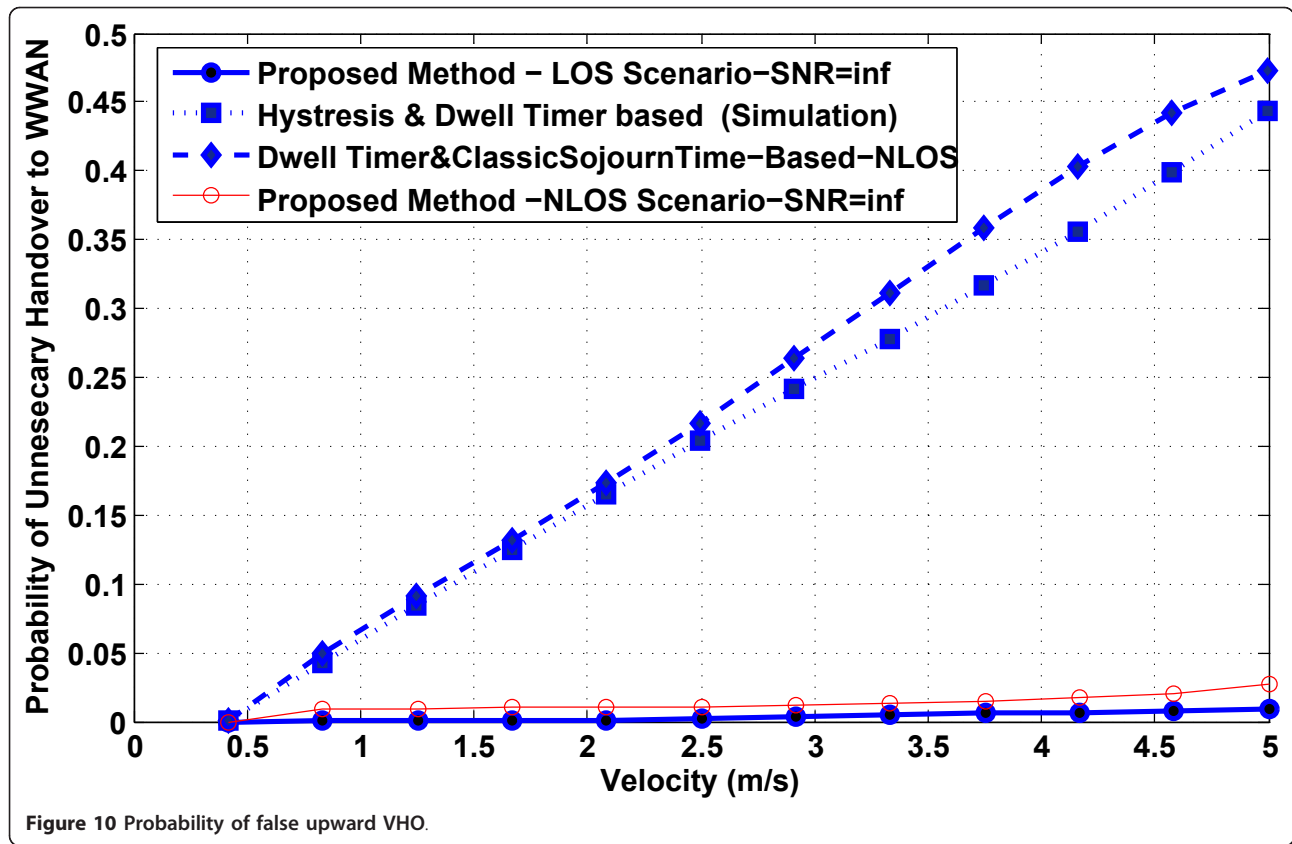This paper proposed a load balancing scheme that minimizes the VHO rate while achieving the desired service
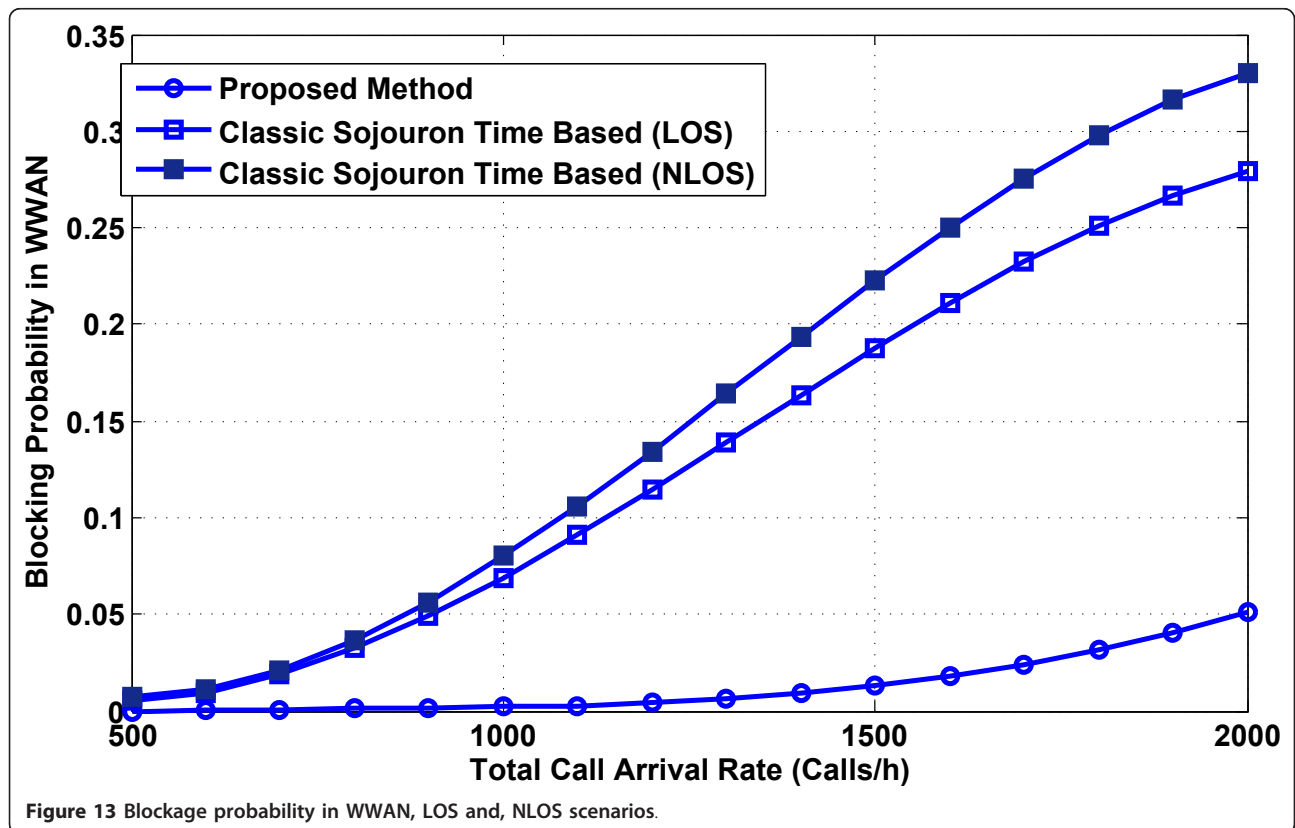

Figure 9 Probability of false downward VHO.

Figure 10 Probability of false upward VHO.

Figure 11 Probability of false upward VHO, SNR = 20 dB.

Figure 12 Probability of false downward VHO, SNR = 20 dB.

Figure 13 Blockage probability in WWAN, LOS and, NLOS scenarios.

Figure 14 Blockage probability in WLAN, LOS and NLOS scenarios.

Figure 15 Blockage probability in WWAN for different number of pedestrian users in LOS and NLOS scenarios.

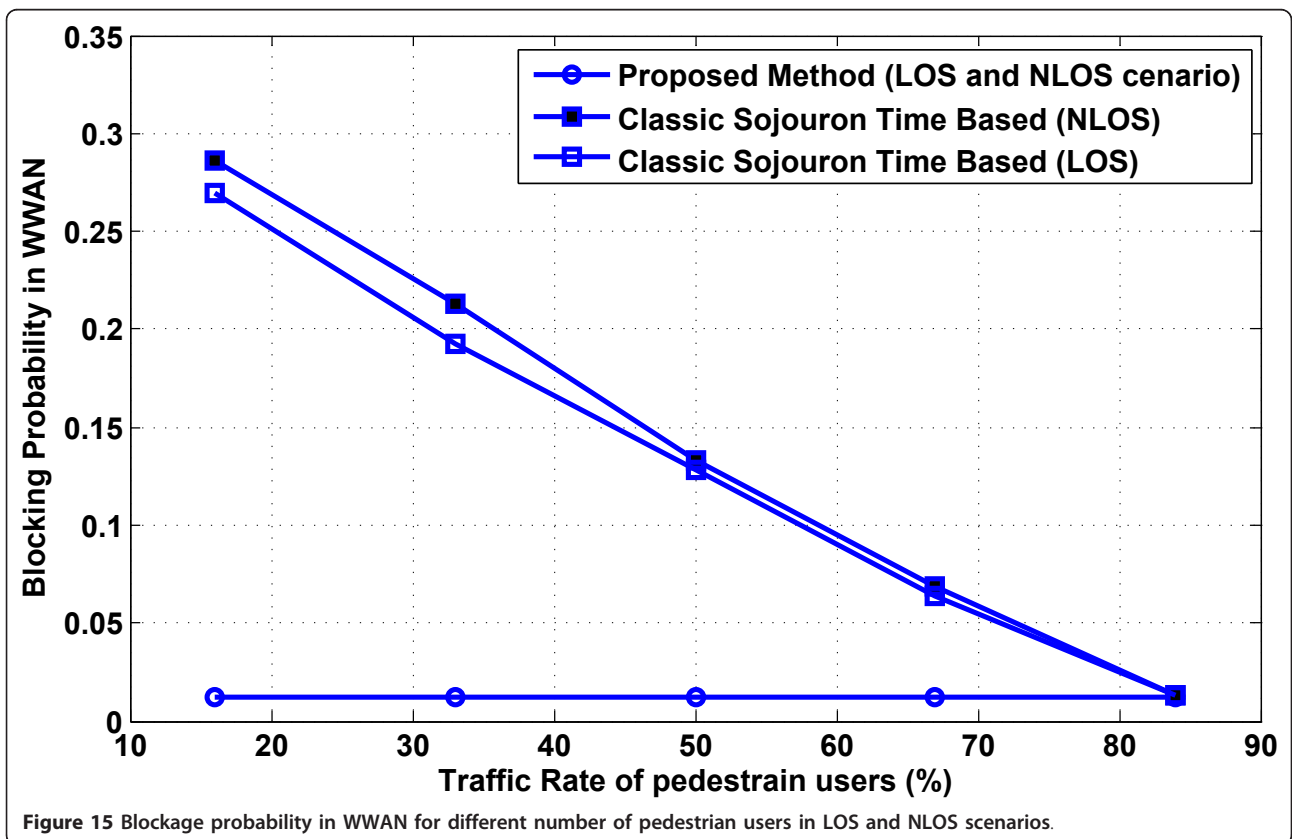**Figure 16 Blockage probability in WLAN for different number of pedestrian users in LOS and NLOS scenarios**.
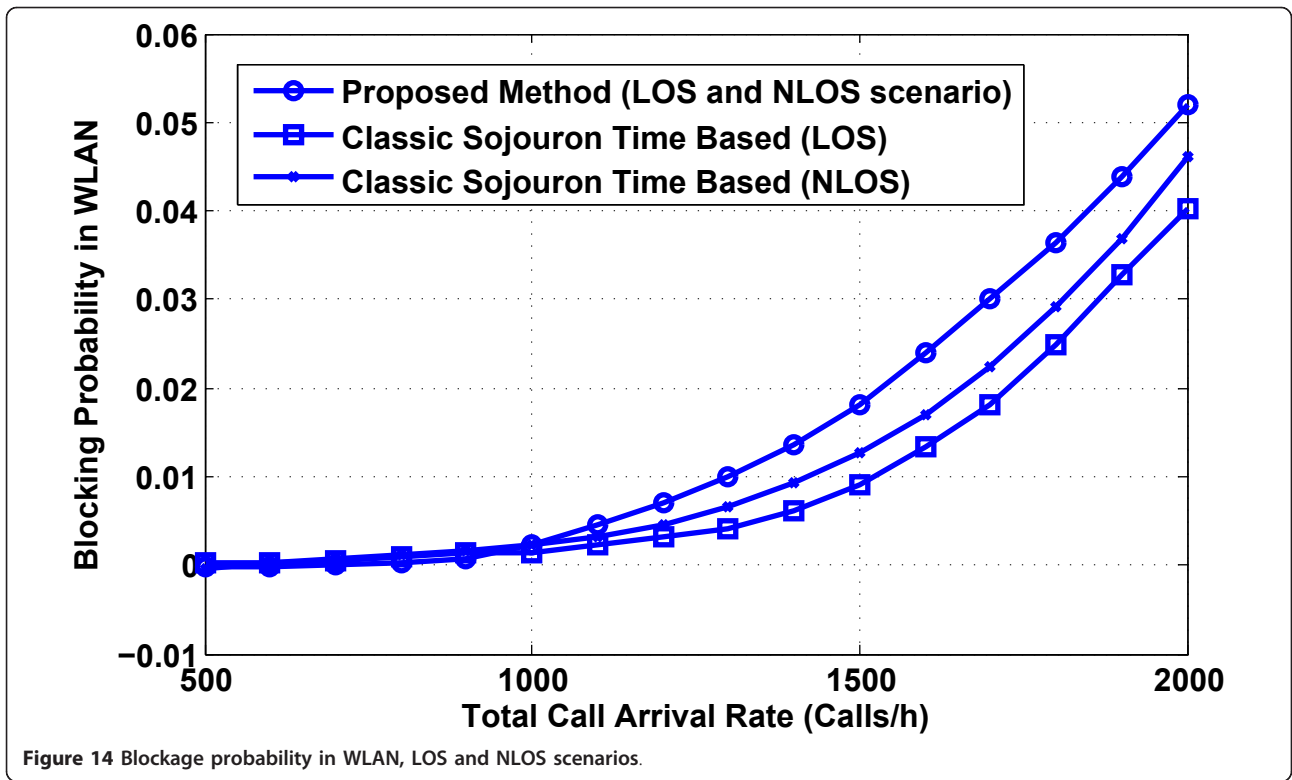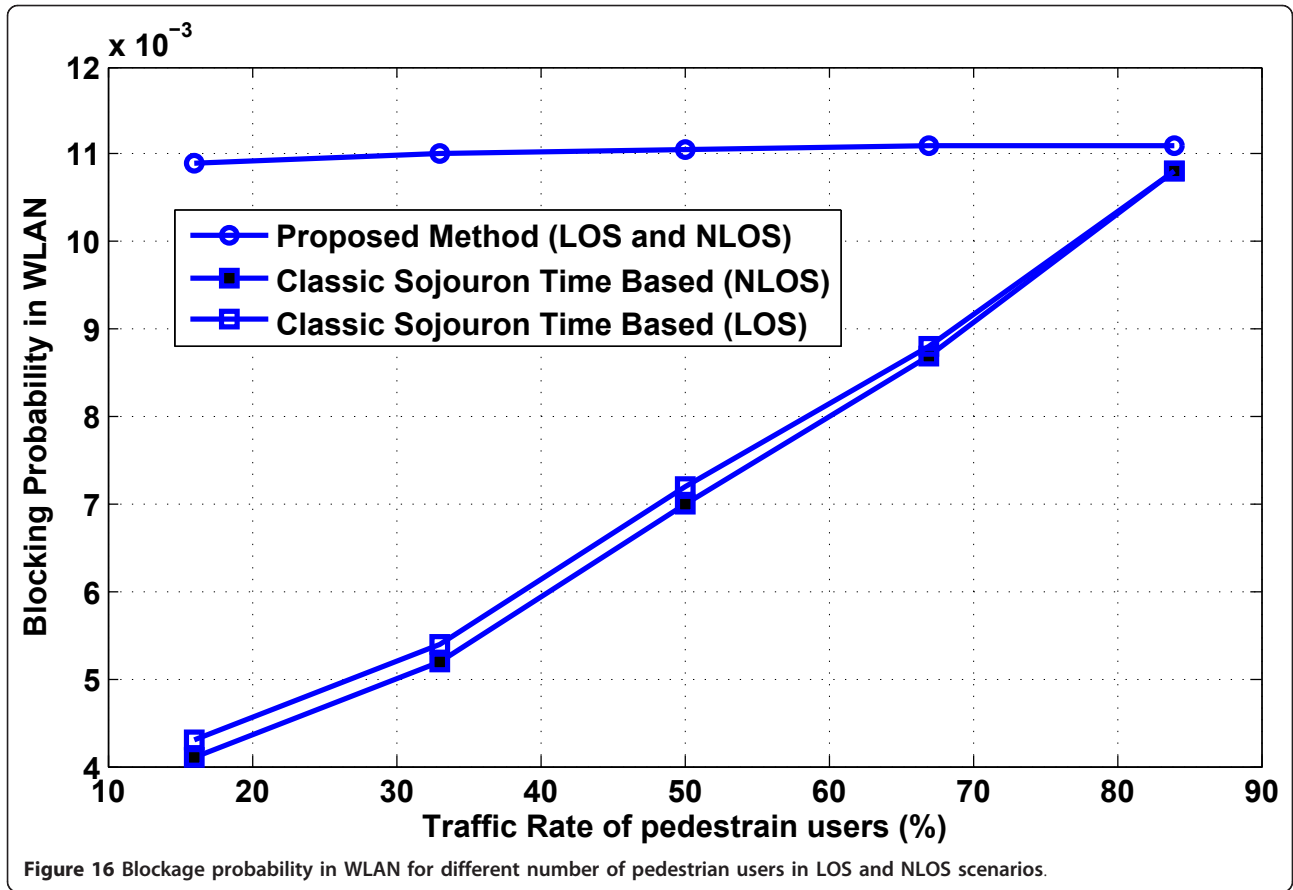
quality (i.e., low call blocking probabilities) for highly dense urban area. The performance of the proposed scheme was analyzed via user level Markov chains. Numerical results show that the proposed scheme achieves low VHO rate and low call blocking probability in comparison with the currently existing service-based and sojourn time-based load balancing schemes. Results show that, under different mobility conditions, the proposed scheme exhibits more recourse utilization in integrated networks while achieving low blocking probability in highly dense urban area. Based on these results, it is concluded that the proposed scheme exhibits a good service quality and, hence, serves as a viable alternative for practical IHWN deployment.

Appendix A. Derivation of bias and MSE for proposed ARP estimator

If shadowing is assumed to be constant over the length of temporal windows (LTW) and also if shadowing and multipath fading are considered to be independent, from (4)

$$E[\hat{S}_i] = \lim_{L_i \to \infty} \frac{E\left\{V_i\left(e^{j\omega}\right) \cdot V_i^*\left(e^{j\omega}\right)\right\}}{L_i \Delta tF}\Bigg|_{\omega=0} \tag{22}$$

$$E[\hat{S}_i] = \frac{1}{L_i \Delta tF} \sum_{n=0}^{L_i-1} \sum_{n=0}^{L_i-1} We^{-j\omega(n-m)} E\{s[n]s*[m]\}$$

$$\times\ E\left\{\sum_{i=0}^{k-1}\sum_{z=0}^{k-1} e^{j(\omega_d \cos\theta_i n + \phi_i)} e^{-j(\omega_d \cos\theta_z m + \phi_z)}\right\}\Bigg|_{\omega=0}$$

Considering the fact that shadowing has small variations over the LTW ($W = w[n]w^*[m]$ and $E_s\{s[n]s^*[m]\} = \sigma_s$), then

$$E[\hat{S}_i] = \frac{\sigma_s}{L_i \Delta tF} \lim_{L_i \to \infty} \sum_{n=0}^{L_i-1} \sum_{n=0}^{L_i-1} J_0\left(\omega_d(n-m)\right)$$

$$\times\ We^{-j\omega(n-m)}\Bigg|_{\omega=0} \tag{23}$$

$$= \lim_{L_i \to \infty} \left(G_i(\omega) * |w_i(\omega)|^2\right)\Bigg|_{\omega=0}$$

where $G_i(0)$ and $W_i(0)$ are DC component of RSS PSD and rectangular window, respectively, and * refers to convolution. Equation (23) shows that if LTW approaches to infinitum in time domain, in frequency domain it approaches to Kronecker delta function $\delta[\cdot]$, and as a result, the periodogram estimation error approaches to zero. Assume that window magnitude is

normalized ($F = 1$) and let $\psi = s[n]s^*[m]$ then from (7), we find a closed form for bias and variance of proposed ARP estimator, as follows S

$$\hat{S}_i = \frac{1}{L_i \Delta t} \sum_{i=0}^{k-1} \sum_{z=0}^{k-1} \sum_{n=0}^{k-1} \sum_{m=0}^{k-1}$$

$$W \psi e^{j(\omega_d \cos\theta_i n + \phi_i)} e^{-j(\omega_d \cos\theta_z m + \phi_z)}$$

$$= \frac{1}{L_i \Delta t} \sum_{i=0}^{k-1} \sum_{z=0}^{k-1} e^{j(\phi_i + \phi_z)}$$

$$\times \underbrace{\sum_{n=0}^{L_i-1} \sum_{m=0}^{L_i-1} W \Psi e^{j(\omega_d \cos\theta_i n)} e^{-j(\omega_d \cos\theta_z m)}}_{A} \quad (24)$$

if $n - m = q$ then $\begin{cases} \text{beginning of period, } 0 < q < L_i - 1 \\ \text{end of period, } -(L_i - 1) < q < 0 \end{cases}$

thus $-m \leq q \leq L_i - m - 1$ and therefore we have

$$A = \sum_{m=0}^{L_i-1} \sum_{q=-(L-1)}^{L_i-1} w[m+q]w*[m] \times$$

$$e^{j\omega_d(\cos\theta_i - \cos\theta_z)m} e^{j\omega_d \cos\theta_i q}$$

$$E[\hat{S}_i] = \frac{\sigma}{L_i \Delta_s} \sum_{m=0}^{L_i-1} \sum_{q=-(L-1)}^{L_i-1} w[m+q]w*[m]J_0(q\Delta t\omega_d) \quad (25)$$

$$= \frac{\sigma_s}{L_i \Delta t} \sum_{q=-(L-1)}^{L_i-1} J_0(q\Delta t\omega_d) \underbrace{\sum_{m=0}^{L_i-1} w[m+q]w*[m]}_{C_{ww}[q]}$$

where $C_{WW}[m]$ is autocorrelation function (ACF) of rectangular window. Thus, window ACF can be determined as

$$C_{ww}[q] = \begin{cases} L - |q|, & |q| \leq (L-1) \\ 0, & \text{elsewhere} \end{cases}$$

$$\text{then } C_{ww}(e^{j\omega}) = \left(\frac{\sin(\omega L/2)}{\sin(\omega/2)}\right)^2 \quad (26)$$

After substitution in (25), bias can be calculated as follows.

$$E[\hat{S}_i] = \sigma_s \sum_{q=-(L_i-1)}^{L_i-1} \left(\frac{L_i - |q|}{L_i}\right) J_0(q\Delta t\omega_d)$$

$$= \sigma_s \left[\frac{1}{L_i} + 2 \sum_{q=1}^{L_i-1} \left(\frac{L_i - |q|}{L_o}\right) J_0(q\Delta t\omega_d)\right] \quad (27)$$

In order to calculate variance of local mean estimates, $4th$ moment of RSS is calculated. When LTW approaches infinity, the $4th$ moment of most of stochastic processes will approach to Gaussian process, and for a Gaussian process, it can be calculated based on

autocorrelation function [23]. Let X be square of RSS with ACF equals to $J_0^2(\omega_d K \Delta t)$[16]. Thus for RSS which is a Gaussian process, it can be shown that $E(X^2) = \sigma_x^2 + \overline{X}^2$ for a Gaussian process $E(X^2) \geq 2\overline{X}^2$ therefore $E[(\hat{S} - E[\hat{S}])^2] \geq E[\hat{S}_i]^2$ thus

$$E\left[\left(\hat{S} - E[\hat{S}]\right)^2\right] = \frac{\sigma_s^2}{L_i \Delta t} \times$$

$$\underbrace{E\left[\left(\sum_{n=q}^{L-1} \sum_{i=0}^{k-1} \sum_{z=0}^{k-1} e^{j\omega_d n(\cos\theta_i - \cos\theta_z)} e^{j(\phi_i - \phi_z)} e^{j\omega_d \cos\theta_z q}\right)^2\right]}_{A}$$

$$\underbrace{-\left(\sigma_s\left[\frac{1}{L_i} + 2\sum_{q=1}^{L_i-1}\left(\frac{L_i-|q|}{L_i}\right)J_0(q\Delta t\omega_d)\right]\right)^2}_{B}$$

$$A = \frac{\sigma_s^2}{L_i^2 \Delta t^2} E\left\{|p[n]|^2 |p[n-q]|^2\right\} \quad (28)$$

$$= \frac{R_{|p|^2|p|^2}(0)}{L_i^2 \Delta t^2} + \frac{2\sigma_s^2}{L_i^2 \Delta t^2} \sum_{q=1}^{L_i-1}(L_i - |q|)^2 J_0^2(\omega_d q\Delta t)$$

$$B = 4\sigma_s^2 \sum_{q=0}^{L_i-1} J_0^2(q\Delta t\omega_d) \underbrace{\sum_{q=0}^{L_i-1}\left(\frac{L_i^2 + q^2 - 2L_i|q|}{L_i^2}\right)}_{C}$$

$$C = \sum_{q=0}^{L_i-1} 1 + \frac{1}{L_i^2}\sum_{q=0}^{L_i-1}q^2 - \frac{2}{L_i}\sum_{q=0}^{L_i-1}|q| = \frac{(2L_i^2 - 3L_i + 1)}{6L_i}$$

$$B = 4\sigma_s^2 \sum_{q=0}^{L_i-1}\frac{(2L_i^2 - 3L_i + 1)}{6L_i}J_0^2(q\Delta t\omega_d)$$

A lower bound for variance of proposed local mean power estimator can be determined as follows (from (28))

$$\text{Var}(\hat{S}) \leq \sigma_s^2 \left[\left(\frac{R_{|p|^2|p|^2}(0) - 1}{L_i^2 \Delta t^2}\right) + \sum_{q=1}^{L_i-1}J_0^2(\omega_d q\Delta t)\right]$$

$$\left(\left(\frac{L_i - |q|}{L_i}\right)^2 - \frac{2(2L_i^2 - 3L_i + 1)}{3L_i}\right) \quad (29)$$

$$- 4\sum_{q=0}^{L_i-1}\left(\frac{L_i - |q|}{L_i^2}\right)J_0(\omega_d q\Delta t)$$

Appendix B. Derivation of joint and conditional PDF of received signal estimates

Due to the fact that RS estimates $\hat{S}$ dB have Gaussian distribution, joint $f_{\hat{S}_{k-1},\hat{S}_k}(\hat{S}_{k-1},\hat{S}_k)$ and $f_{\hat{S}_k|\hat{S}_{k-1}}(\hat{S}_k|\hat{S}_{k-1})$ PDF of each event will be Gaussian. Let $E\{\hat{S}[k-1]\} = \mu_{\hat{s}}[k-1]$, $\text{Var}\{\hat{S}[k-1]\} = \sigma_{\hat{S}[k-1]}$, $E\{\hat{S}[k]\} = \mu_{\hat{S}[k]}$, and $\text{Var}\{\hat{S}[k]\} = \sigma_{\hat{S}[k]}$

$$P\left\{\hat{S}_{\text{WLAN}}[k-1] \le T_{\text{WLAN}}\right\} = Q\left(\frac{T_{\text{WLAN}} - \mu_{\hat{S}_{WL}[k-1]}}{\sigma_{\hat{S}_{WL}}[k-1]}\right) \quad (30)$$

$$f_{\hat{S}_{k-1}, \hat{S}_k}\left(\hat{S}_{k-1}, \hat{S}_k\right) = \frac{1}{\sqrt{2\pi}\sigma_{\hat{S}_{k-1}}} \exp\left(\frac{\hat{S}_{k-1} - \mu_{k-1}}{2\sigma_{\hat{S}_{k-1}}^2}\right)^2$$

$$\times \frac{1}{\sqrt{2\pi}\sigma_{\hat{S}_k}\sqrt{1 - \rho_{\hat{S}\hat{S}}^2}}$$

$$\times \exp\left(\frac{\left(\hat{S}_k - \mu_k - \rho_{\hat{S}\hat{S}}\frac{\sigma_{\hat{S}_k}}{\sigma_{\hat{S}_{k-1}}}\left(\hat{S}_{k-1} - \mu_{k-1}\right)\right)^2}{2\sigma_{\hat{S}_k}^2\left(1 - \rho_{\hat{S}\hat{S}}^2\right)}\right) \quad (31)$$

$$f_{\hat{S}_k|\hat{S}_{k-1}}\left(\hat{S}_k \mid \hat{S}_{k-1}\right) = \frac{1}{\sqrt{2\pi}\sigma_{\hat{S}_k|\hat{S}_{k-1}}} \exp\left(-\frac{\hat{S}_k - \mu_{k|k-1}}{2\sigma_{\hat{S}_k|\hat{S}_{k-1}}^2}\right) \quad (32)$$

$$\mu_{k|k-1} = \mu_k + \rho_{\hat{S}\hat{S}} + \frac{\sigma_{\hat{S}_k}}{\sigma_{\hat{S}_{k-1}}}\left(\hat{S}_{k-1} - \mu_{k-1}\right) \quad (33)$$

$$\sigma_{\hat{S}_k|\hat{S}_{k-1}}^2 = \sigma_{\hat{S}_k}^2\left(1 - \rho_{\hat{S}\hat{S}}^2\right) \quad (34)$$

$$\rho_{\hat{S}\hat{S}} = \frac{1}{x_{\text{av}} + x_0}\left[x_0 \exp\left(\frac{-|v.t_S|}{x_0}\right) - x_{\text{av}} \exp\left(\frac{-|v.t_S|}{x_{\text{av}}}\right)\right] \quad (35)$$

where x_0 is correlation distance of local mean power and x_{av} is smoothing window length. From above, transition probabilities can be determined, for instance $p_{23}[k]$ is

$$P_{23}[k] = Q\left(\frac{T_{\text{WWAN}} - \mu_{\text{WW}[k]}}{\sigma_{\text{WW}[k]}}\right) \Big/ Q\left(\frac{T_{\text{WLAN}} - \mu_{\text{WL}[k-1]}}{\sigma_{\text{WL}[k-1]}}\right)$$

$$\times \left[\int_{T_{\text{WLAN}}}^{\infty} f_{\hat{S}_{k-1}}(\hat{S})\left[1 - Q\left(\frac{T_{\text{Add, WLAN}} - \mu_{k|k-1}}{\sigma_{\hat{S}_k}|\hat{S}_{k-1}}\right)\right] \text{d}\hat{S}\right]$$

$$\times P\{V[k] > V_{\text{th}}\}$$

Competing interests
The authors declare that they have no competing interests.

References
1. IF Akyildiz, J Xie, S Mohanty, A survey of mobility management in next-generation all-IP-based wireless systems. IEEE Trans Wirel Commun. 11(4), 16–28 (2004). doi:10.1109/MWC.2004.1325888
2. SN Enrique, YX Lin, WSW Vincent, An MDP based vertical handoff decision algorithm for heterogeneous wireless networks. IEEE Trans Veh Technol. 57(2), 1243–1254 (2008)
3. W Shen, QA Zeng, Cost-function-based network selection strategy in integrated wireless and mobile networks. IEEE Trans Veh Technol. 57(6), 3778–3788 (2008)
4. N Mirmotahhary, A Kohansal, H Zamiri-Jafarian, M Mirsalehi, Discrete mobile user tracking algorithm via velocity estimation for microcellular urban environment. in *IEEE Vehicular Technology Conference, Singapore* 2631–2635 (May 2008)
5. M Ylianttila, M Pande, J Makela, P Mahonen, Optimization scheme for mobile users performing vertical handoffs between IEEE 802.11 and GPRS/EDGE networks, in *IEEE Global Telecommunication Conference*, San Antonio, TX, (December 2001)
6. R Hsieh, ZG Zhou, A Senevirante, S-MIP: A seamless handoff architecture for mobile IP, in *IEEE International Conference on Computer Communication* San Francisco, CA, (April 2003)
7. GP Pollini, Trends in handover design. IEEE Commun Mag. 34(3), 82–90 (1996). doi:10.1109/35.486807
8. N Mirmotahhary, Y Mafinejad, F Atbaei, A Kouzani, An adaptive policy-based vertical handoff algorithm for heterogeneous wireless networks, in *IEEE 8th International Conference on Computer and Information Technology*, Sidney, pp. 188–193 (July 2008)
9. Q Zhang, C Guo, Z Guo, W Zhu, Efficient mobility management for vertical handoff between WWAN and WLAN. IEEE Commun Mag. 41(11), 102–108 (2003). doi:10.1109/MCOM.2003.1244929
10. N Nam, N Choi, Y Seok, Y Choi, WISE: energy-efficient interface selection on vertical handoff between 3G networks and WLANs. IEEE Int Symp Pers Indoor Mobile Radio Commun 692–698 (July 2004)
11. H Wang, R Katz, J Giese, Policy-enabled handoffs across heterogeneous wireless networks. in *IEEE Workshop on Mobile Computing Systems and Applications* 51–60 (1999)
12. Xiaodong Yang, Thomas OwensJ, Intersystem soft handover for Converged DVB-H and UMTS Networks. IEEE Trans Veh Technol. 57, 1887–1898 (2008)
13. G Lampropoulos, C Skianis, P Neves, Optimized Fusion of Heterogeneous Wireless Networks based on Media Independent Handover Operations. IEEE Wirel Commun Mag. 17(4), 78–87 (2010)
14. N Dimitriou, L Sarakis, D Loukatos, G Kormentzas, C Skianis, Vertical handover framework for future collaborative wireless networks, Int J Netw Manag, (Wiley)
15. WF El-Sadek, MN Mikhail, Universal mobility with global identity (UMGI) architecture. IEEE International Conference on Wireless Networks and Information Systems, WNIS 09 Los Alamitos389–394
16. N Zhang, J Holtzman, Analysis of handoff algorithms using both absolute and relative measurements. in *44th IEEE Vehicular Technology Conference* 82–86 (September 1994)
17. Y Chung, DJ Lee, DH Cho, BC Shin, Macrocell/microcell selection schemes based on a new velocity estimation in multi-tier cellular system. IEEE Trans Veh Technol. 51(5), 893–903 (2002). doi:10.1109/TVT.2002.801764
18. R Narasimhan, DC Cox, Estimation of mobile speed and average received power in wireless systems using best basis methods. IEEE Trans Commun. 49(12), 2172–2183 (2001). doi:10.1109/26.974264
19. D Wong, DC Cox, Optimal local mean signal power level estimator for rayleigh fading environments. in *IEEE International Conference on Information, Communications Signal Processing* 1701–1704 (September 1997)
20. KDA Appiah, On generalized covariance-based velocity estimation IEEE Trans. Veh Technol. 48, 1546–1557 (1999). doi:10.1109/25.790529
21. T Jaing, ND Sidiropulos, G Giannakis, Kalman filtering for power estimation in mobile communications. IEEE Trans Wirel Commun. 2(1), 151–161 (2003). doi:10.1109/TWC.2002.806386
22. AE Leu, BL Mark, An efficient timer-based hard handoff algorithm for cellular networks. in *IEEE Wireless Communication and Networking Conference* 1969–1974 (March 2003)
23. A Papoulis, SU Pillai, *Probability, Random Variables and Stochastic Processes*, 4th edn. (McGraw-Hill, New York, 2002)