Load Balance for Multi-Layer Reuse Scenarios on Mobile WiMAX System

Juliano J. Bazzo and André M. Cavalcante
Nokia Technology Institute (INdT)
Manaus-AM, Brazil
{juliano.bazzo,andre.cavalcante}@indt.org.br

Marco J. de Sousa
Federal University of Pará (UFPA)
Belém-AM, Brazil
marcojsousa@ufpa.br

Lauri Kuru and Jani Moilanen
Nokia Siemens Networks (NSN)
Espoo, Finland
{lauri.kuru,jani.moilanen}@nsn.com

***Abstract*—This paper intends to propose a novel handover algorithm to balance the load of the layers in a multi-reuse scenario. Each layer works independently from each other and it has its own scheduling process, coverage and mobile stations attached. The algorithm proposed balances the resources among layers by moving mobile stations from one to another layer according their QoS requirements and channel condition. Results show that it can increase the cell coverage and still keeping a satisfactory quality for the VoIP calls. Effective spectral efficiency also increased when comparing with single tri-sectorized reuse 1 scenario. The algorithm is also prepared to avoid ping-pong handovers and to work for any number of layers.**

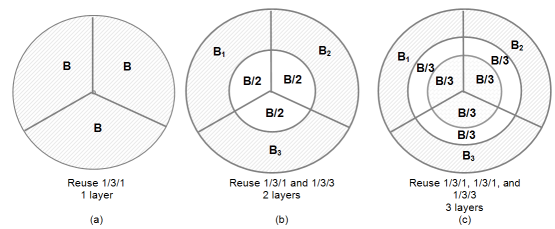

Fig. 1. BS scenarios: (a) Single layer, (b) 2-layers scenario and (c) 3-layers scenario with different reuses.

I. Introduction

ITU IMT-Advanced global standardization group has worked to meet the requirements of the next generation mobile networks [1]. One key target of these networks is to increase the cell throughput capacity without penalizing the quality of service (QoS) of the users. Several approaches have been studied in order to try to achieve it, like smart antennas, new scheduling algorithms and overlay techniques for different systems. The approach proposed in this paper is to exploit the characteristics of multi-layer reuse scenarios through a novel load balancing (LB) handover algorithm. The term layer used in this paper refers to the coverage area under a given frequency reuse. Figure 1 shows three examples of layered scenarios. Figure 1a is a single layered scenario with reuse 1/3/1 (1 base station with 3 sectors with reuse factor equal to 1) and total bandwidth of B. Figure 1b is a two layered scenario with reuse 1/3/1 and bandwidth B/2 for inner layer and reuse 1/3/3 and bandwidth B1+B2+B3=B/2 for outer layer. As can be noted, this scenario presents the same total bandwidth of single layer scenario. By inner layer we mean the smaller circle which corresponds to smaller coverage area. By outer layer we mean the larger circle and hence larger coverage. Figure 1c is a three layered scenario having the bandwidth per layer of one-third the total bandwidth of single layered scenario.

In a usual single layered scenario there is a classical trade-off between throughput capacity, coverage and QoS for Guaranteed Bit Rate (GBR) VoIP users that are much more sensible to delay and packet errors. Frequency reuse 1/3/1 is the general strategy to increase the throughput capacity. However it decreases the SINR (Signal to Interference plus Noise Ratio), which can lead to a worst VoIP user satisfaction for instance (criteria for VoIP calls outage can be found in [2]). In other hand, an increasing in the user satisfaction can be achieved by reuse to 1/3/3, but the throughput capacity decreases because the bandwidth per sector is one-third of the total. Additionally, concerning the power consumption, mobile stations (MSs) connected under reuse 1/3/3 consume less power than under reuse 1/3/1, if considered that the total power per carrier is kept the same for both reuses.

Multi-layer scenario can be used properly to combine efficiently different reuses in order to better exploit the trade-off already mentioned. In this context, a coordinated user distribution among the layers is the key problem to be solved and it is the main purpose of this work. It proposes a handover algorithm among layers to move users according their QoS requirements.

To the best of our knowledge, handover algorithms for multi-layer reuse scenario driven to increase the network throughput capacity, while guaranteeing user QoS requirements, have not been published. At least those are not known by the authors of this work. However, some related ideas have been already published:

Principle of partitioning reuse was first published by S. W. Halpern in [3]. Halpern demonstrated that the partitioning reuse can potentially increase the network capacity in terms of voice channel availability keeping the same SINR objective. However, access selection strategies were not evaluated.

Tripathi et. al. [4] proposed a handover algorithm based on Fuzzy logic system for overlaid network. In that work the partitioning reuse was not considered and the analysis for a single traffic model did not take into account the user satisfaction requirements.

K. Ivanov and G. Spring [2] evaluated a coordinated handover algorithm based on mobile speed in a mixed wireless network technology environment. Results presented there showed that the number of handovers and the call termination were reduced in comparison to uncoordinated handover case (i.e., without taking into account the mobile speed for instance). Multi-layer scenario, reuse diversity and satisfaction evaluation were not taken into account.

978-1-4244-3574-6/10/$25.00 ©2010 IEEE

V.A. de Sousa Jr. et. al. [5] proposed an anticipatory handover algorithm for multi-access WCDMA and WiFi networks. The strategy was to perform the handover from WiFi to WCDMA network of that user with lowest throughput without checking the WiFi or WCDMA system load which could lead to service degradation due to unbalance conditions and, delays and package losses due to excessive handovers. The performance results from [5] are concentrated on a limited simulation scenario without usual mobility constraints like loading variation due to call losses and handovers.

K. Kim, Y. Han and J. Lee [6] proposed a load sharing algorithm between heterogeneous wireless systems. Their proposal was to evaluate the user satisfaction level in terms of minimum required throughput and maximum accepted head of line (HOL) delay and then select those users to be moved to another system. The results were based on an OFDMA and 1xEVDO wireless system simulation which shown that the users might be better distributed between the systems to get the desired throughput. Packet HOL delay, despite mentioned in the beginning of the paper and crucial for VoIP applications, was not evaluated in the results. They mentioned that the parameters were not well dimensioned and deserves further improvements.

The algorithm proposed on this work can increase the coverage and the effective spectral efficiently while keeping the VoIP user satisfied according to the speech analysis performed by [7].

The paper is organized as follows: In section II, it presents the load balance handover algorithms. In Section III, it describes the simulation scenario, parameters and assumptions. Performance evaluation results are presented and discussed in section IV. Finally, conclusions are given in Section V.

II. Load Balance Handover Algorithm

Traditional handover algorithms are based on the MS received signal strength (RSS), SINR or MS speed. Apart from the concern about QoS requirements to move users from one to another cell, previous works [4] [2] also don't take into account either the MS radio conditions before performing the load balancing handover.

A general handover algorithm flow is presented on Figure 2 and it concerns on two parts, the default and load balance handover cycles. Checking period, i.e., when the algorithm is evaluated, for Default and Load Balancing handovers are not necessarily the same. By the Default Handover, we mean the movement based on radio conditions (RSS and SINR), i.e., MSs with radio conditions below a given threshold are moved to another base station (BS), layer or sector that can provide better signal quality than the serving BS. By the Load Balance Handover, we mean the movement based on layer loading conditions. Detailed algorithm flow is presented on Figure 3. The algorithm starts by measuring the layer loading, which can be based on, for example, the layer average throughput, number of occupied slots or even the number of terminals per layer. The adjacent layers are paired and ranked starting with the most unbalanced pair as depicted in step 2 from Figure 3. The most unbalanced layers are those that vary most in terms of loading measures.

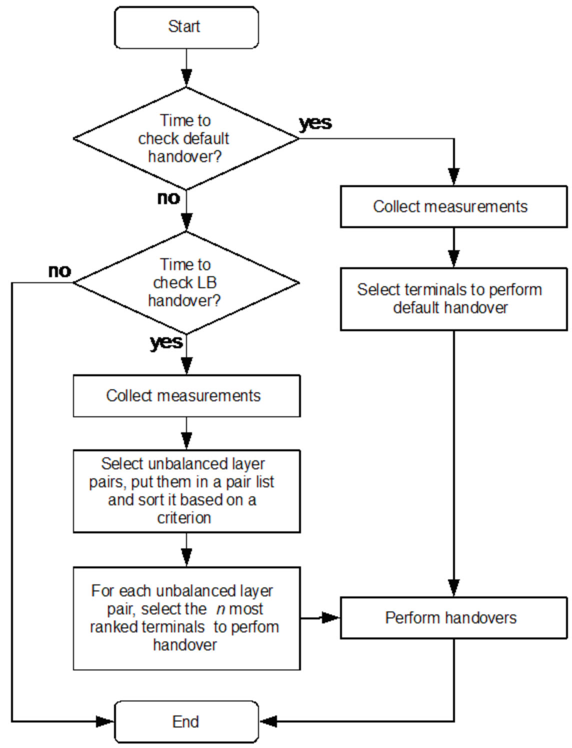

Fig. 2. Generic algorithm that evaluates the radio and balancing conditions.

Finally, for the most unbalanced pairs MSs from the inner loaded layer are ranked and handed over in order to balance the load. The ranking can be done e.g. by using the cost function in the Eq.1. The values $w_1 ... w_n$ are the weights for each measures $M_1 ... M_n$. An integral part of this article is that the algorithm takes into account the traffic types (e.g. QoS class with an associated measure value) associated of each MS. Examples of other measures are received signal strength (RSS), SINR, throughput and etc.

$$C = \sum_{i=1}^{n} w_i . M_i \tag{1}$$

It should be also noted that the cost function from Eq.1 is only based on MS specific measurements while the unbalanced condition is based on layer specific measurements (such as frame load).

MSs with highest cost function value are moved to the target layer. However, in order to avoid ping-pongs handovers, a MS is only moved to the target layer if the radio conditions there, in terms of RSS and SINR, is a given margin above the minimum thresholds. With this strategy, it is not probable that the MS is moved to the target layer with Load Balance algorithm and then returns to the original layer due to Default handover in a short period of time. Additionally, different cost function weights can be assigned to each direction, i.e., weights from first to second layer handover can be different for the other way movement. After the MSs are moved in order to achieve the balancing, the layers that took part of the initial evaluation are removed from all other layer pairs that they take part as depicted in step 2 of Figure 3. All layer pairs are evaluated following the same flow explained previously.

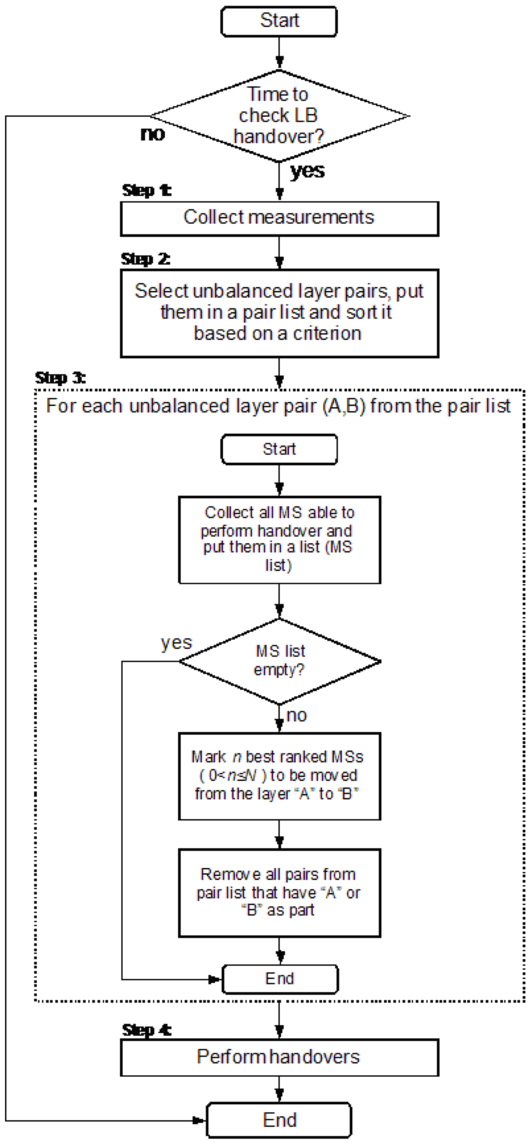

Fig. 3. Load balance handover algorithm in details.

III. Simulation Scenario

In order to evaluate the performance of the load balance handover algorithm described above, a dynamic system-level simulator was used. The simulator models a standard-compliant IEEE 802.16e WiMAX multi-cell and multi-user radio network including modeling of network layout, terminal distribution and movement, radio environment, PHY layer, MAC layer (RRM and ARQ algorithms, scheduling, access schemes) and traffic generation. As simulation scenario, a network of homogeneous hexagon cells (1.5km site-to-site distance), 27 Base Stations (BSs) and 75 antenna sectors was adopted. The BS cell arrangement is depicted in Figure 1b.

Table I shows the most important simulation parameters concerning the network topology and deployment. The propagation model applied at the simulation is presented in Table II.

TABLE I
Network topology and Deployment

Parameters	Value
Number of BSs	27
Number of sectors per BS	3
Total number of sectors	75
Number of layers	2
Inner cell reuse	1-3-1
Outer cell reuse	1-3-3
BS-BS distance	1.5 km
Center frequency	2.5 GHz
Inner cell bandwidth	5 MHz
Outer cell bandwidth	5 MHz
Outer/inner carrier power ratio	3dB
BS height	30 m
Number of Tx/Rx antennas	1/2 (Maximum Ratio Combing - MRC)
Tx antenna pattern	70°(-3dB) w/ 20dB front/back ratio
Tx antenna gain	15 dBi
MS height	1.5 m
Number of Rx antennas	2
Rx antenna pattern	Omnidirectional
MS Noise Figure	8 dB

TABLE II
Propagation Models

Parameters	Value
Path loss model	COST-231 Hata
Lognormal shadowing	μ = 0 dB, σ = 8 dB
Shadowing correlation	100% inter-sector, 50% inter-BS
Channel model	ITU Pedestrian A (3 km/h)

Figure 4 refers to the movements for radio conditions, i.e., Default, and Load Balance handovers that are enabled for the simulation scenario. The cell to cell and layer to layer radio conditions handover are triggered when the target cell has 3dB higher RSS than the serving cell. The inner to outer happens when the RSS goes below a threshold. Only users located on outer layer and with a RSS above a given threshold are eligible for load balance handover. These users are selected by applying the cost function described by Eq.1 and following the algorithm from Figure 3.

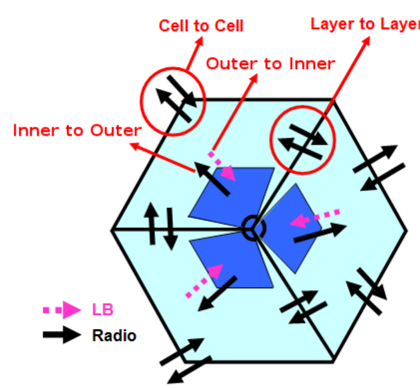

Fig. 4. The handover movements between dual layer scenario.

Table III presents the scheduler and HARQ (Hybrid Automatic Repeat Request) most important configuration for VoIP and best effort (BE) users. The load balance handover parameters are presented in Table IV.

TABLE III
SCHEDULER AND HARQ PARAMETERS

Parameters	Value
Number of simultaneous active calls per sector	20
DL/UL ratio in frame	24/23 [symbols/symbols]
CQI feedback period/Delay	1/0 [frames/frames]
Traffic Type	50% VoIP and 50% BE
Scheduling Algorithm	Weighted Round-Robin (WRR) with twice more scheduling opportunity for VoIP users
HARQ Type	Chase Combining
Maximum number of HARQ Retrans.	3
HARQ Retransmission Delay	1 frame

TABLE IV
LOAD BALANCE HANDOVER PARAMETERS

Parameters	Value
LB checking period	1.5 seconds
Handover delay	0.1 seconds
Enabled handovers due radio conditions (default handover)	cell to cell layer to layer inner to outer
Enabled handovers due balancing conditions (LB handover)	inner to outer
Unbalance Condition	The loading criterion is based on the frame usage, i.e. percentage of occupied slots per frame. Then, the condition is: Serving $\geq$ 80% AND Target $\leq$ 100% AND (Target - Serving) $\geq$ 10%
Number of MSs selected for HO at the time	3
Cost function HO prioritization	MSs with BE traffic type and highest RSS

IV. PERFORMANCE RESULTS

The simulation results were evaluated in terms of:

1) ***Call drop***: The percentage of call dropped.
2) ***HARQ packet retransmissions***: The percentage of HARQ packet retransmissions.
3) ***VoIP User Satisfaction***: The percentage of satisfied VoIP users. An user call is considered satisfied if:
   - the packet error rate (PER) is below 2%.
   - the air interface delay is below 50ms.
   - the call is not dropped.
4) ***Frame Load***: The WiMAX OFDMA frame utilization percentage.
5) ***Spectral Efficiency***: The system spectral efficiency from physical layer taking into account the total sector t-put and bandwidth.

As noted by the load balancing algorithm configuration, users that have their predominant traffic type as BE and with better radio conditions are selected to be handed over from outer to inner layer. The reuse 1/3/1 for inner layer can offer a better throughput condition for BE users than the outer layer. Users with highest RSS are potentially candidates for higher Modulation and Coding Schemes (MCSs) which might lead to a better spectral efficiency.

Figure 5 presents the gains when comparing the coordinated user selection strategy mentioned previously and without any coordination, i.e., they are selected randomly to perform the handover. Particularly for this comparison, the Round-Robin (RR) scheduler was applied and only one MS is moved at load balancing handover time. For the following comparisons, the Weighted Round-Robin (WRR) was applied. The VoIP user satisfaction ("VoIP User Sat." in the figure) decreases in 5% because these users are more concentrated on outer layer for coordinated strategy while they are spread between both layers for the uncoordinated strategy. Therefore, the spreading improves the VoIP users however, it decreases the sector spectral efficiency ("Spec Eff." in the figure) in 9% and the inner layer throughput ("Inner L. Tput" in the figure) in 18%. The efficiency loss happens because the supposely closest BS users, with higher RSS, are not being handed over to inner layer. Additionaly, the uncoordinated handover moves much more VoIP users to inner layer and their contributions to increase the data rate there is small, about only 12.2 kbps. The outer layer is always overloaded due to its reduced bandwidth and because the admition control algorithm always select the outer layer as the first candidate.

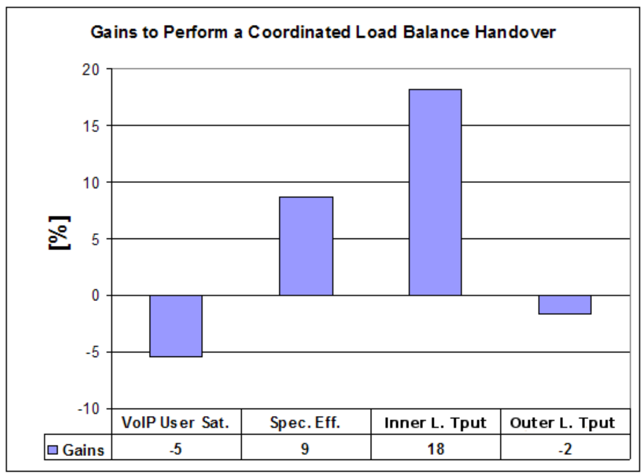

Fig. 5. Gains to perform the coordinated load balance handover.

Figure 6 refers to the gains to move three and five users at the load balance handover time in comparison with only one. The handover opportunity only happens in cycles of 1.5 seconds thus, it might be a long time to move only one user at the time particularly in a very active system, i.e., many BE users with short data transfer time, VoIP calls, handovers, call droppings and admissions. The inner layer only receives MSs through load balancing handovers and then it is important to identify an efficient user transfer rate in order to keep that layer well balanced. As shown in the graph, the VoIP satisfaction ("VoIP User Sat." in the figure) increases in 11% when moving three users at the time. The load transference from outer to inner can be noted by the inner layer throughput increasing and the reduction for the outer. Moving more than three users overload the inner layer which is the reason for the user satisfaction decreases.

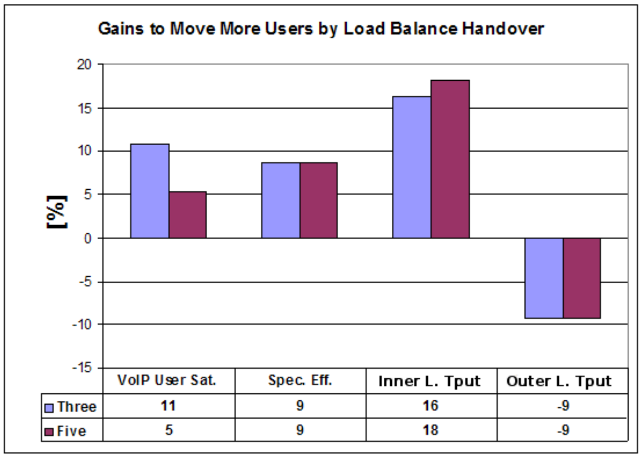

Fig. 6. Gains to move more users by load balancing handover.

Figure 7 shows the gains relative to single layer reuse 1/3/1 scenario for both layers and for each one independently. In a multi-reuse scenario, adding the outer layer with reuse 1/3/3 reduces the call dropp and the number of HARQ packet retransmissions ("HARQ Retrans" in the figure) in 76% and 16% respectively, due to the interference reduction for the whole system.

There isn't any gain in the user satisfaction for both layer and upper only, however the VoIP users moved to inner are experiencing much better satisfaction. The results are much better than it seems because the multi-reuse scenario has only 6.7 MHz of bandwidth than the single reuse scenario with 10MHz.

The spectral efficiency for multi-reuse scenario is small due to reduced sector bandwidth. On the other hand, comparing the single and inner layer only, the spectral efficiency was the same however the frame usage was 7% smaller. Therefore, the inner layer has potentially better spectral efficiency than single layer.

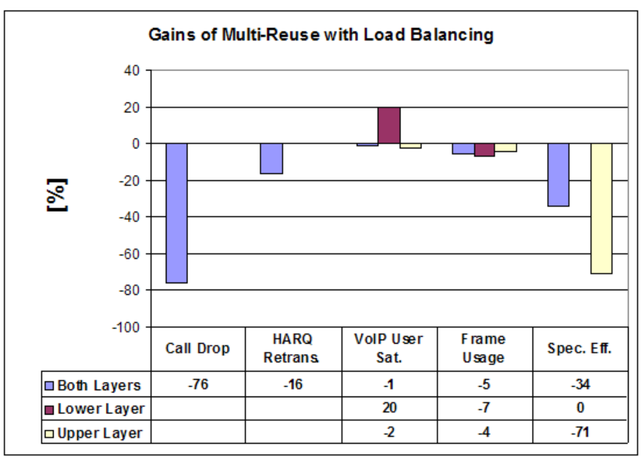

Fig. 7. Gains of multi-reuse with load balancing.

The Figure 8 shows the frame loading progress for inner and outer cells. The spikes show clearly the load balance handovers in action.

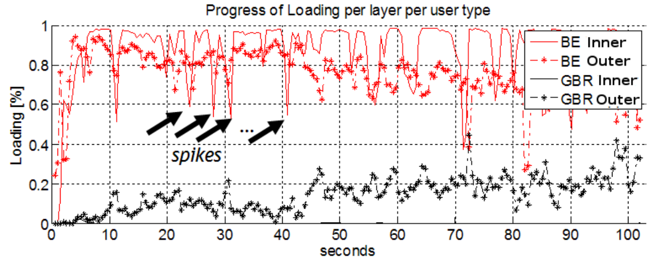

Fig. 8. Inner and outer cell loading distribution by flow type.

V. Conclusions

In this work, we presented a novel load balancing handover algorithm for multi-reuse scenario that takes into account the flow type and channel condition to move MSs from one to another layer. With multi-reuse and the load balancing algorithm presented, the VoIP calls satisfaction per sector remained the same when comparing with single 1/3/1 reuse scenario. The call drops and layer one HARQ retransmission reduced significantly with our proposal. The sector throughput was smaller because the effective bandwidth is only 67% of single 1/3/1 reuse however, comparing the inner layer with single scenario having both the same reuse, the potentially spectral efficiency for full frame is higher for inner layer.

Moving more user at the load balancing handover trigger is a good strategy to increase the VoIP user satisfaction and the overall throughput.

Further works will concentrate in create one algorithm to manage VoIP calls and another for best effort because their QoS requirements are very different. Additionally, performance influence due to handover delay and an analysis for wide bandwidth scenario, higher than 20MHz per sector, and middle load will be essential.

References

[1] ITU, "The ITU global standard for international mobile telecommunications (imt-advanced)," [Online] Available: www.itu.int/home/imt.html, May 2010.

[2] K. Ivanov and G. Spring, "Mobile speed sensitive handover in a mixed cell environment," *IEEE Veh. Technol. Conf.*, pp. 892–896, 1995.

[3] S. W. Halpern, "Reuse partitioning in cellular systems," *IEEE Veh. Technol. Conf.*, pp. 322–327, 1983.

[4] N. D. Tripathi, J. H. Reed, and H. F. VanLandingham, "Adaptive handoff algorithms for cellular overlay systems using fuzzy logic," *IEEE Veh. Technol. Conf.*, 1999.

[5] V. A. de Sousa Jr., R. A. de O. Neto, F. de S. Chaves, L. S. Cardoso, J. F. Pimentel, and F. R. P. Cavalcanti, "Access selection with anticipatory vertical handover for multi-access networks," *Personal Indoor and Mobile Radio Communications (PIMRC)*, 2006.

[6] K. Kim, Y. Han, and J.-R. Lee, "Load sharing in heterogeneous wireless systems using dynamic load measures," *IEEE Veh. Technol. Conf.*, 2006.

[7] E.-C. Popovici and T. Radulescu, "Coding scheme impact on the ip-qos network utilization and voice quality," *8th COST290 Management Committee and Technical Meeting*, Feb 2007.