

A Generalized Framework for Multi-RAT Scenarios Characterisation

J. Pérez-Romero, O. Sallent, R. Agustí
Dept. of Signal Theory and Communications
Universitat Politècnica de Catalunya (UPC)
E-mail: [jorperez, sallent, ramon]@tsc.upc.edu

Abstract.- This paper proposes a novel approach for developing Common Radio Resource Management (CRRM) algorithms in heterogeneous wireless scenarios. It is based on a new indicator, denoted as *fittingness factor*, which reflects the suitability of allocating a given RAT and cell to a user depending on the service and on the terminal and network capabilities. Simulation results reveal that the proposed approach is able to capture the different terminal and service heterogeneities and obtains better performance than other approaches existing in the literature.

Keywords.- Common Radio Resource Management (CRRM), heterogeneous wireless networks, vertical handover, RAT selection.

I. INTRODUCTION

The heterogeneous radio access network (RAN) concept proposes a flexible and open architecture for a large variety of wireless access technologies, applications and services with different Quality of Service (QoS) demands as well as different protocol stacks. Mobile and wireless radio access networks differ from each other by air interface technology, cell-size, services supported, bit rate capabilities, coverage, mobility support, etc. The heterogeneous characteristics that these networks offer make possible to exploit the trunking gain resulting from the joint consideration of the different networks as a whole. That is, the additional dimensions introduced by the multiplicity of radio access technologies (RATs) available provide further flexibility in the way how radio resources can be managed and, consequently, overall improvements may follow. Thus, the interworking among heterogeneous RANs leads to a better overall performance than the accumulated performances of the stand-alone systems. This challenge calls for the introduction of new radio resource management (RRM) algorithms operating from a common perspective that take into account the overall amount of resources offered by the available RANs, and therefore are referred to as CRRM (Common Radio Resource Management) algorithms [1][2]. In particular, when a multi-RAN scenario is considered, a specific functionality arises, namely RAT selection devoted to decide the RAT that a given service request should be allocated to. It can be executed either at session initiation (i.e. the initial RAT selection procedure) or during an on-going session depending on how the network or the terminal position have changed since the session started. In this case, the RAT selection procedure may lead to a vertical or inter-system handover, changing the access network the mobile is currently connected to.

On the other hand, the scenario heterogeneity is also present from the customer side, because users may access the requested services with a variety of terminal's capabilities

(e.g. single or multi-mode terminals) and different market segments can be identified (e.g. business or consumer users) with their corresponding QoS levels. In this context, it is prime important to devise a generic framework to take all these diverse aspects into account and come up with suitable management principles under any possible circumstance. More specifically, in order to cope with this multi-dimensional heterogeneity, two main levels are identified in the RAT selection problem:

1) Capabilities. A user-to-RAT association may not be possible for limitations in e.g. the user terminal capabilities (single-mode terminals only able to be connected to a single RAT) or the type of services supported by the RAT (e.g. videophone is not supported in 2G networks).

2) Suitability. A user-to-RAT association may or may not be suitable depending on the matching between the user requirements in terms of QoS and capabilities offered by the RAT (e.g. a business user may require bit rate capabilities feasible on HSDPA and not on GPRS or these capabilities can be obtained in one technology or another depending on the RAT occupancy, etc.). In that respect, there are a number of considerations, which can be split at two different levels:

a) Macroscopic: Radio considerations at cell level such as load level or, equivalently, amount of radio resources available.

b) Microscopic: Radio considerations at user level such as measured path loss, intercell interference, etc. This component will be relevant for the user-to-RAT association when the amount of radio resources required for providing the user with the required QoS significantly depends on the local conditions where the user is located (e.g. power level required in WCDMA downlink).

In this framework, this paper proposes a new RAT selection algorithm that takes into account all these considerations by defining a new CRRM metric that reflects the suitability of selecting each available RAT depending on the specific terminal and system capabilities.

The CRRM problem has been covered in the recent years in the literature, where different algorithms have been proposed. In [3] the effect of tuning the load-based handover (HO) thresholds depending on the load of inter-system/inter-layer/inter-frequency cells aiming at balancing the load in different RATs is studied. In [4], a force-based load balancing approach is proposed for initial RAT selection and vertical HO decision making. In turn, in [5] the authors compare the

load balancing principles with respect to service-based CRRM policies. Similarly, Lincke discusses the CRRM problem from a more general perspective in e.g. [6] and references therein, comparing several substitution policies. In [7] a framework for defining initial RAT selection policies combining service-based criteria and radio network considerations is presented. Finally, in [8][9] the authors propose a RAT allocation methodology that reduces the interference in the CDMA network by executing the RAT allocation depending on the total path loss measured by the user. This paper differs from the previous works in that a general methodology based on a new metric able to capture different types of network, service and terminal heterogeneities, and therefore applicable in multiple scenarios, is presented.

The rest of the paper is organised as follows. Section II defines the proposed new metric, denoted as *fittingness factor*. Section III presents the RAT selection strategy derived from the proposed methodology. Section IV provides some examples for evaluating the fittingness factor under different services and RATs. The proposed methodology is evaluated by means of a detailed system level simulator described in Section V and results are presented in Section VI. Finally, Section VII summarises the conclusions.

II. FITTINGNESS FACTOR DEFINITION

The concepts presented in Section I regarding capabilities and suitability can be captured for each RAT in a new measure, the so-called *fittingness factor*, which reflects the degree of adequacy of a given RAT to a given user. This factor is defined with respect to the j -th RAT for the i -th user, who belongs to the p -th customer profile requesting the s -th service, as follows:

$$\psi_{i,p,s,j} = C_{i,p,s,j} \times Q_{i,p,s,j} \quad (1)$$

The first term reflects the hard constraints posed by the capabilities of either the terminal or the technology, and therefore is defined as:

$$C_{i,p,s,j} = T_{i,p,j} \times S_{s,j} \quad (2)$$

$T_{i,p,j}$ reflects the terminal capabilities and equals 0 if the terminal of the i -th user belonging to the p -th profile does not support the j -th RAT and 1 if the RAT is supported. Similarly, the term $S_{s,j}$ reflects the RAT capabilities and equals 0 if the s -th service is not supported by the j -th RAT or 1 otherwise.

On the other hand, the term $Q_{i,p,s,j}$ reflects the suitability of the j -th RAT to support the s -th service requested by the i -th user with the p -th customer profile. In order to define this suitability, one or both of the macro and microscopic views described in Section I can be considered, as reflected in Figure 1. Notice that the x-axis in these functions (e.g. available resources, path loss, interference measurements, etc.) can be dynamically updated through measurements, so that the fittingness factor of the current and alternative RATs can be monitored and vertical handover procedures triggered in case. Some examples for the computation of $Q_{i,p,s,j}$ following an analytical approach will be presented in Section IV.

Notice that the fittingness factor should be defined at a

cell-level considering the specific resources available in a given cell or the path-loss or interference with respect to this cell.

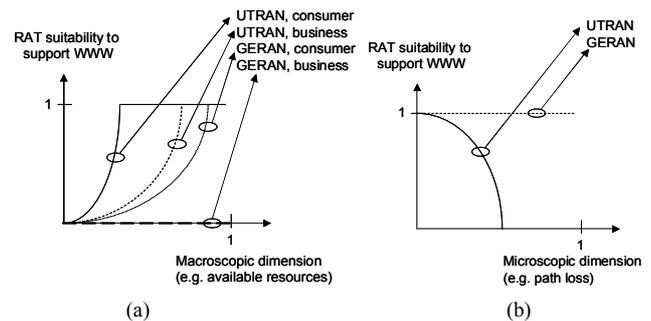


Figure 1 Suitability as a function of the macroscopic (a) and microscopic (b) dimensions.

III. RAT SELECTION

A. Initial RAT selection algorithm

Based on the above framework, the proposed RAT selection algorithm for the i -th user of the p -th profile requesting a given service s would be as follows:

Step 1.- Measure the fittingness factor for each candidate cell k_j of the j -th detected RAT. Since the measurement is done separately for uplink and downlink, both measurements can be weighted to obtain a unique indicator:

$$\psi_{i,p,s,j}(k_j) = \alpha_{p,s} \psi_{i,p,s,j}^{UL}(k_j) + (1 - \alpha_{p,s}) \psi_{i,p,s,j}^{DL}(k_j) \quad (3)$$

Here, the weight factor is $\alpha_{p,s}$, depending in general of the specific service and profile, in the sense that for very asymmetric services $\alpha_{p,s}$ should be close to 0 so that the downlink is basically considered in the computation of the total fittingness factor (alternatively close to 1 if the uplink is the most important link). In turn, for symmetric services a proper setting could be $\alpha_{p,s}=0.5$.

Step 2.- Select the RAT J having the cell with the highest fittingness factor among all the candidate cells:

$$J = \arg \max_j \left(\max_{k_j} \psi_{i,p,s,j}(k_j) \right) \quad (4)$$

Step 3.- Try admission in the RAT J .

Step 4.- If admission is not possible, try with the next RAT in decreasing order of fittingness factor, provided that its fittingness factor is higher than 0. If no other RATs with fittingness factor higher than 0 exist, block the call.

In case that two or more RATs have the same value of the fittingness factor, then a decision can be taken based on other criteria (e.g. select the RAT with less load).

B. Vertical Handover algorithm

Similarly, the proposed criterion to execute a vertical handover algorithm based on the fittingness factor would be as follows, assuming that the terminal is connected to the RAT and cell denoted as "servingRAT" and "servingCell".

Step 1.- For each candidate cell and RAT, monitor the corresponding fittingness factor $\psi_{i,p,s,j}(k_j)$. Measures should be

averaged during a period T .

Step 2.- If the condition

$$\psi_{i,p,s,j}(k_j) > \psi_{i,p,s, \text{servicingRAT}}(\text{servicingCell}) + \Delta_{\text{VHO}} \quad (5)$$

holds during a period T_{VHO} then a vertical handover to RAT j and cell k_j should be triggered, provided that there are available resources for the user in this RAT and cell.

IV. EXAMPLES OF FITTINGNESS FACTOR COMPUTATION

This section provides some expressions for computing the suitability function $Q_{i,p,s,j}$ included in the fittingness factor definition in Section II, under specific considerations regarding service and RAT, and following an analytical approach. In particular, the GERAN (GSM/EDGE Radio Access Network) and UTRAN (UMTS Terrestrial Radio Access Network) RATs will be considered with three different service types, namely voice, videophone and interactive service. Similar expressions could be obtained for other RATs (e.g. WLAN, WiMAX, etc.).

A. GERAN

a) Voice users

For voice users, the suitability function in GERAN would basically depend on the measured path loss L_i , and will be defined both for uplink and downlink as:

$$Q_{i,p, \text{VOICE}, \text{GERAN}} = \begin{cases} 1 & \text{if } L_i \leq L_{\max} \\ 0 & \text{if } L_i > L_{\max} \end{cases} \quad (6)$$

where L_{\max} is the maximum path loss according to sensitivity level S_{\min} and maximum available power (i.e. $L_{\max} = P_{T\max} - S_{\min}$).

b) Interactive users

For data users working on Packet Data Channels (PDCHs), the suitability function can be defined taking into account both the occupancy and the measured path loss, leading to:

$$Q_{i,p, \text{INTERACTIVE}, \text{GERAN}} = \frac{R_{\text{MCS}}(L_i)}{R_{b\max,s,p}} \min(\varphi_p, M) \quad (7)$$

This expression is general for uplink and downlink. M is the multislot capability in the uplink or downlink, $R_{\text{MCS}}(L_i)$ is a function that provides the maximum bit rate among all the possible Modulation and Coding Schemes (MCS) available depending on the path loss L_i and the link adaptation mechanism, and $R_{b\max,s,p}$ is the maximum theoretical bit rate that the service could achieve among all the RATs.

The factor φ_p is the multiplexing factor and reflects how the users are multiplexed over the channels depending on the specific profile p . It is computed as the quotient between the number of slots allocated to the service profile p in a frame with respect to the total number of users with the same service and profile that have currently an established TBF (Temporary Block Flow).

B. UTRAN

a) Voice and Videophone users

For the voice and videophone services in the uplink direction, the suitability function can be defined depending on the maximum path loss allowable according to the current

uplink load factor and the service requirements. Then, it is given by:

$$Q_{i,p, \text{VOICE}, \text{UTRAN}}^{\text{UL}} = \begin{cases} 1 & \text{if } L_i \leq L_{\max} \\ 0 & \text{if } L_i > L_{\max} \end{cases} \quad (8)$$

with the maximum path loss L_{\max} given as a function of the bit rate $R_{b,i}$, the E_b/N_0 target, the maximum transmit power $P_{T\max}$, the noise power P_N , the chip rate $W=3.84$ Mc/s and the measured uplink load factor η_{UL} as [2]:

$$L_{\max} = \frac{P_{T\max}}{P_N} \left(\frac{W}{\left(\frac{E_b}{N_0}\right)_i R_{b,i}} + 1 \right) (1 - \eta_{\text{UL}}) \quad (9)$$

In turn, for the downlink, the suitability can be given as a function of the estimated required power with respect to the maximum available power for that service and user:

$$Q_{i,p, \text{VOICE}, \text{UTRAN}}^{\text{DL}} = \begin{cases} 1 & \text{if } P_{\text{Ti}} \leq \Delta P_{\max,p,s} \\ 0 & \text{if } P_{\text{Ti}} > \Delta P_{\max,p,s} \end{cases} \quad (10)$$

with $\Delta P_{\max,p,s} = \min(P_{\max,p,s}, P_{T\max} - P_T)$ is the maximum power available for the i -th user ($P_{\max,p,s}$ is the maximum power that can be allocated to a user with profile p and service s , P_T is the current transmitted power and $P_{T\max}$ is the maximum power available in the downlink). In turn, the power requirement of the new user can be estimated as:

$$P_{Ti} = \frac{L_i \text{RSSI}_i - P_T(1-\rho)}{\frac{W}{\left(\frac{E_b}{N_0}\right)_i R_{b,i}} + \rho} = \frac{\frac{P_p}{\left(\frac{Ec}{I_0}\right)_i} - P_T(1-\rho)}{\frac{W}{\left(\frac{E_b}{N_0}\right)_i R_{b,i}} + \rho} \quad (11)$$

where RSSI_i is the total received power at the antenna input by the i -th user, which can be expressed as a function of the pilot power P_p and the Ec/I_0 of the pilot measured by this user. In turn, $R_{b,i}$ is the bit rate, E_b/N_0 the target requirement, $W=3.84$ Mc/s the chip rate and ρ the orthogonality factor.

b) Interactive users

In the uplink direction the suitability function is defined as:

$$Q_{i,p, \text{WWW}, \text{UTRAN}}^{\text{UL}} = \frac{f(R_i^*)}{R_{b\max,s,p}} \varphi_p \quad (12)$$

where R_i^* is the maximum bit rate that can be achieved by the i -th user depending on its path loss, the noise power, the load factor and the maximum available power, given by:

$$R_i^* = \frac{W}{\left(\frac{E_b}{N_0}\right)_i \left(\frac{L_i}{1 - \eta_{\text{UL}}} \frac{P_N}{P_{T\max}} - 1 \right)} \quad (13)$$

where $f()$ is a function that adjusts the bit rate R_i^* to that of the closest transport format not exceeding R_i^* .

The factor φ_p , like in the GERAN case, is the multiplexing factor and reflects the average number of users in cellDCH state (i.e. with a dedicated channel DCH allocated) with respect to the total number of users of service profile p with data in their buffers (i.e. including those in cellDCH and those

in RACH_FACH, who are waiting for a DCH to be free).

In the downlink, the suitability function is given by:

$$Q_{i,p,WWW,UTRAN}^{DL} = \frac{f(R_i^*)}{R_{b,max,p}} \varphi_p \quad (14)$$

where in this case the maximum bit rate R_i^* depends on the maximum power available for the service and profile

$\Delta P_{max,p,s} = \min(P_{max,p,s}, P_{Tmax} - P_T)$ as follows:

$$R_i^* = \frac{\Delta P_{max,p,s} W}{\left(\frac{E_b}{N_o}\right)_i \left(\frac{P_p}{(Ec/I_o)_i} - (1-\rho)P_T - \rho\Delta P_{max,p,s}\right)} \quad (15)$$

V. SIMULATION MODEL

The considered approaches have been evaluated by means of system level simulations in a scenario that considers 7 omnidirectional cells for GERAN and 7 for UTRAN. The cells of both RANs are collocated. The cell radius is 1 km. In case of GERAN, it is assumed that the 7 cells represent a cluster so that all of them operate with different carrier frequencies. The main parameters of the User Equipment (UE) and the Base Station (BS) are summarised in Table I. It is assumed that all terminals have multi-mode capabilities, i.e. they can be connected either to UTRAN or to GERAN. The urban macrocell propagation model in [10] is considered for both systems, corresponding to $L(\text{dB})=128.1+37.6\log(d(\text{km}))$ with an additional shadowing with standard deviation 10 dB. The mobility model in [11] is considered with speed 3 km/h.

Voice and videophone calls are generated according to a Poisson process with an average call rate of 10 calls/h/user and exponentially distributed call duration with an average of 180 s. In UTRAN, the Radio Access Bearer (RAB) for voice users is the 12.2 kb/s speech defined in [12], while for videophone users the bit rate is 64 kb/s. In turn, GERAN does not support the videophone service and voice users are allocated to a TCH-FS (traffic channel full-rate speech), i.e. one time slot in each frame. Interactive users follow the www browsing model given in [11], with 5 pages per session and an average reading time between pages of 20s. In the uplink, there is an average of 25 packets per page, an interarrival packet time 0.05s and an average packet size of 366 bytes. In turn, in the downlink there are 50 packets per page on average, the interarrival packet time is 0.01s and the average packet size is 392 bytes. The average time between user sessions is 30s. It is assumed that half of the interactive users belong to the consumer profile and half to the business profile. WWW browsing service is provided in UTRAN by means of dedicated channels (DCH) using the transport channel type switching procedure. The considered RAB assumes an uplink maximum bit rate of 64 kb/s and a downlink maximum bit rate of 128 kb/s for consumer users and 384 kb/s for business users [12]. In turn, in GERAN, the www service is provided through a PDCH (Packet Data Channel) with a round robin scheduling algorithm to allocate transmissions to users sharing the same time slot. The algorithm allocates three times more resources

to business users than to consumer users in order to have the same bit rate relation than in UTRAN. On the other hand, a link adaptation mechanism operating in periods of 1s is used to select, for each user, the highest modulation and coding scheme (MCS) that ensures the specific sensitivity.

Concerning the fittingness factor evaluation for the vertical handover algorithm, the measurements are averaged in periods of $T=1\text{s}$. The hysteresis margin is $\Delta_{VHO}=0.1$ and $T_{VHO}=3\text{s}$. Furthermore, $\alpha_{p,s}=0.5$ for all services giving equal importance to uplink and downlink in the fittingness factor computation.

A summary of the main RRM parameters in UTRAN and GERAN is given in Table II. With respect to the admission control procedure in UTRAN, three conditions are checked [2], namely the uplink load factor after user acceptance should be below the threshold η_{max} , the downlink transmitted power below P_{max} and there must be code sequences available. With respect to GERAN, voice users are accepted provided that there are available time slots, while interactive users are always accepted at session initiation in idle state. Voice users have precedence over www users, so that slots occupied by www users are allocated to incoming voice users when there are not other free slots. All slots are reversible except the slot 0 of the carrier transmitting the broadcast channel.

TABLE I UTRAN BS AND UE PARAMETERS

BS parameters	UTRAN	GERAN
Maximum transmitted power	43 dBm	43 dBm
Thermal noise	-104 dBm	-117 dBm
Common Control Channels Power	33 dBm	43 dBm
Maximum DL power per user	41 dBm	N/A
Number of carriers	1	3
UE parameters	UTRAN	GERAN
Maximum transmitted power	21 dBm	33 dBm
Minimum transmitted power	-44 dBm	0 dBm
Thermal noise	-100 dBm	-113 dBm
DL Orthogonality factor	0.4	N/A
Multislot class (UL, DL, UL+DL)	N/A	2,3,4

TABLE II RRM PARAMETERS

UTRAN	
UL admission threshold (η_{max})	1.0
DL admission threshold (P_{max})	42 dBm
Active Set size	1
Replacement hysteresis	3 dB
Time to trigger handover	0.64 s
BLER target voice and videophone	1%
BLER target interactive	10%
GERAN	
Maximum number of TBFs per slot	UL: 8, DL:32
Minimum power to trigger handover	-100 dBm
Samples to trigger handover	3

VI. RESULTS

In the following, the results comparing the performance of the proposed fittingness factor-based algorithm against other schemes are presented. The considered schemes for comparison are:

- Load Balancing: It allocates the RAT having the lowest load.
- Service-based: This corresponds to the algorithm explained in [7] in which voice traffic is served in GERAN and interactive traffic in UTRAN.
- NCCB (Network Controlled Cell Breathing) algorithm,

presented in [8], applied to voice and data, and based on allocating low path loss users to UTRAN and high path loss users to GERAN, regardless of the service.

- NCCBvoice. This corresponds to the NCCB algorithm applied only to voice users, while www users are allocated to UTRAN, as explained in [9].

In all the cases videophone users are only allocated in UTRAN.

Figure 2 and Figure 3 plot the downlink average packet delay of consumer and business users, respectively, for the different strategies, as a function of the number of voice users in a scenario without conversational users and with a total of 600 interactive data users (50% consumer and 50% business). The proposed strategy based on the fittingness factor algorithm achieves the best behaviour among the considered ones for both user profiles. In this case, the fittingness factor algorithm tends to allocate the business users whenever possible in UTRAN, while the consumer users are distributed between UTRAN and GERAN.

Figure 4 plots the total aggregated throughput in the scenario (i.e. considering all the services). Not significant differences exist between the proposed fittingness algorithm, the load balancing approach and the NCCBvoice schemes, although the throughput with the fittingness factor strategy is slightly higher. In turn, the performance of the NCCB and the service-based strategies is poorer because of the degradation suffered by data users served mainly through GERAN.

In a scenario with videophone users the performance in terms of average delay for data users follows similar trends, as shown in Figure 5, being again the fittingness factor algorithm the one that achieves the best behaviour.

certain service profile. It allows then capturing several types of heterogeneities regarding terminal and network capabilities. By using this metric, a RAT selection algorithm has been presented and evaluated by means of simulations, revealing a better overall performance from both the user and the system point of view than that obtained with other approaches.

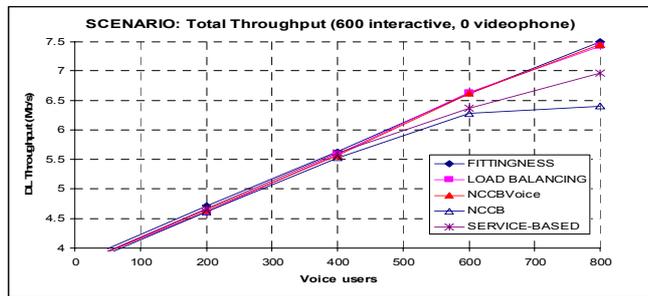


Figure 4 Total DL aggregated throughput in the scenario

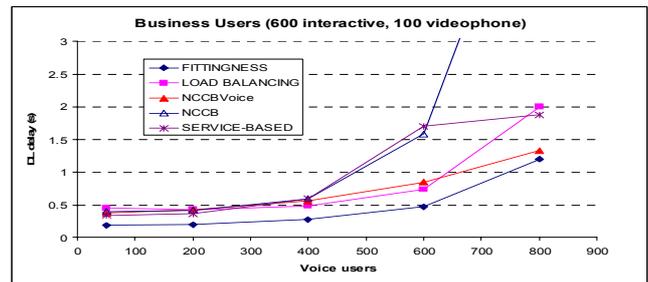


Figure 5 DL packet delay of business users with 100 videophone users

ACKNOWLEDGMENTS

This work has been performed in the framework of the IST AROMA project, which is partly funded by the European Community and by the Spanish Research Council (CICYT) under COSMOS grant TEC2004-00518.

REFERENCES

- [1] 3GPP TR 25.881 v5.0.0 "Improvement of RRM across RNS and RNS/BSS"
- [2] J. Pérez-Romero, O.Sallent, R.Agustí, M. Díaz-Guerra, *Radio Resource Management Strategies in UMTS*, John Wiley & Sons, 2005.
- [3] A. Tölli, P. Hakalin, "Adaptive load balancing between multiple cell layers", *IEEE VTC Fall*, Vol. 3, Sept. 2002, pp.1691 – 1695.
- [4] A. Pillekeit, F. Derakhshan, E. Jugl, A. Mitschele-Thiel, "Force-based load balancing in co-located UMTS/GSM networks", *VTC 2004-Fall*. 2004 IEEE 60th Vol. 6, 26-29 Sept. 2004 pp. 4402 – 4406.
- [5] X. Gelabert, J. Pérez-Romero, O. Sallent, R. Agustí, "On the suitability of Load Balancing Principles in Heterogeneous Wireless Access Networks", *Wireless Personal Multimedia Communications Symposium (WPMC'05)*, Aalborg, Denmark, September, 2005.
- [6] S. Lincke-Salecker, "The Benefits of Load Sharing when Dimensioning Networks", *Proceedings of the 37th Annual Simulation Symposium (ANSS'04)*, April, 2004.
- [7] J. Pérez-Romero, O. Sallent, R. Agustí "Policy-based Initial RAT Selection algorithms in Heterogeneous Networks", *7th MWCN Conference*, Marrakesh, 2005.
- [8] J. Pérez-Romero, O.Sallent, R.Agustí, L.Wang, H.Aghvami "A Novel Algorithm for Radio Access Technology Selection in Heterogeneous B3G networks", *IEEE VTC 2006-Spring*, Melbourne, May, 2006.
- [9] J. Pérez-Romero, O.Sallent, R.Agustí, "Network Controlled Cell Breathing in Multi-service heterogeneous CDMA/TDMA scenarios", *IEEE VTC 2006-Fall*, Montreal, September, 2006.
- [10] 3GPP TR 25.942 "Radio Frequency (RF) system scenarios"
- [11] UMTS 30.03 v3.2.0 TR 101 112 "Selection procedures for the choice of radio transmission technologies of the UMTS", ETSI, April, 1998.
- [12] 3GPP TS 34.108 "Common Test Environments for User Equipment (UE); conformance testing"

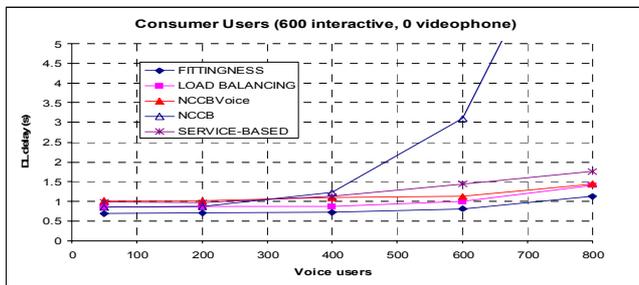


Figure 2 DL packet delay of consumer data users

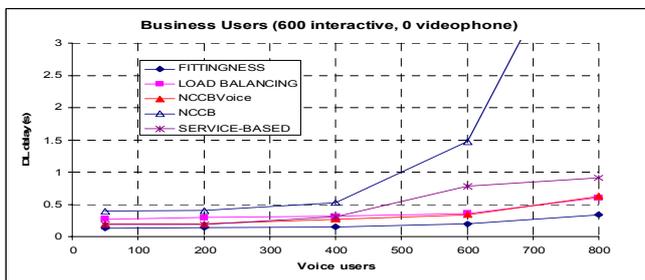


Figure 3 DL packet delay of business data users

VII. CONCLUSIONS

This paper has presented a new metric to develop CRRM strategies, denoted as fittingness factor. It reflects the suitability of allocating a given RAT to a given user of a