A Novel Metric for Context-Aware RAT Selection in Wireless Multi-Access Systems

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

***Abstract*—This paper presents a novel metric, named *fittingness factor*, for developing Common Radio Resource Management (CRRM) algorithms in heterogeneous wireless scenarios. It reflects the suitability of allocating a given Radio Access Technology (RAT) and cell to a user depending on the specific context in terms of requested service and on the terminal and network capabilities at each instant. Thus, the fittingness factor allows synthesizing a number of considerations supporting CRRM decisions into a single and generalized parameter. 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

Future heterogeneous wireless scenarios will be characterized by the coexistence of a large variety of wireless access technologies, with different protocol stacks and supporting a number of applications and services with different Quality of Service (QoS) demands to be provided to terminals with different degrees of multi-mode capabilities to access the available networks. Each mobile and wireless Radio Access Network (RAN) differs from the others by the specific air interface technology, cell-size, services supported, bit rate capabilities, coverage, mobility support, etc. Therefore, the heterogeneous characteristics offered by these networks make possible to exploit the trunking gain resulting from the joint consideration of all the networks as a whole. As a result, the additional dimensions introduced by the multiplicity of available radio access technologies (RATs) provide further flexibility in the way how radio resources can be managed and, consequently, overall improvements may follow with respect to the 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 they are referred as CRRM (Common Radio Resource Management) [1][2]. In particular, when a multi-RAN scenario is considered, a specific functionality arises, namely RAT selection (i.e. the functionality devoted to decide to which RAT a given service request should be allocated), which 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 conditions 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.

The scenario heterogeneity is also present from the customer side, because users may access the requested services with a variety of terminal capabilities (e.g. single or multi-mode) and different market segments can be identified (e.g. business or consumer users) with their corresponding QoS levels. Then, selecting the proper RAT and cell is a complex problem due to the number of variables involved in the decision-making process, as reflected in Figure 1 with some possible inputs. Furthermore, some of these variables may vary dynamically, making the process even more difficult to handle.

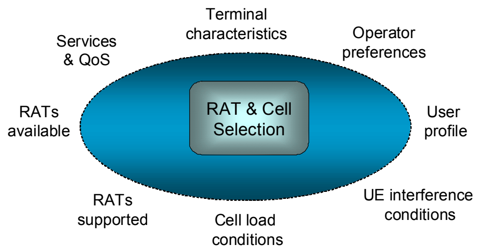


Figure 1. Factors influencing the RAT and Cell selection

In this context, CRRM in general and RAT selection mechanisms in particular have received a lot of attention in recent years, clearly acknowledging the key role that these strategies will have for a full realization of Beyond 3G (B3G) scenarios. Research efforts have been oriented either to propose and assess the performance of heuristic algorithms [3]-[6] or to identify architectural and functional aspects for CRRM support [1][7][8]. From an algorithmic point of view, in [3] and [4], mechanisms to balance the load in different RATs by means of vertical handover decisions are analyzed. However, the service-dimension is not captured in the problem because only real time services are considered. Similarly, Lincke discusses the CRRM problem from a more general perspective in e.g. [5] and references therein, comparing several substitution policies and including the multi-mode terminal dimension with speech and data services. Finally, in [6] the authors propose a RAT allocation methodology that considers the specific radio network features of a CDMA network to reduce the interference by allocating users to RATs depending on the total measured path loss and capturing also the service dimension but considering that all terminals support the available RATs and that the involved RATs support the same services.

With all the above, it would be prime important to devise a generic framework that takes all these diverse aspects into account and comes up with suitable RAT selection principles under any possible circumstance that may arise in a practical implementation. In this respect, this paper introduces a novel metric, named *fittingness factor*, which reflects the suitability of selecting each available RAT depending on the specific terminal and network context and capabilities. Then, by abstracting all these multiple aspects into a single and generalized metric, a more clear representation of the dynamic reality can be obtained.

The mentioned fittingness factor can be particularly useful when taking into account the foreseen evolution of wireless mobile networks towards IP-based architectures accompanied by more and more decentralized management concepts. In this respect, such a single metric can facilitate the implementation of cell-by-cell RRM strategies by reducing signaling exchanges. On the other hand, radio resource usage optimization in mobile terminal autonomic decision making processes could also be facilitated by the fittingness factor concept as long as the operator would communicate to the terminals a single metric that integrates a large number of variables that the operator would not be willing to communicate over the air separately. It is worth noting that e.g. the IEEE standard committee on next generation radio and spectrum management is progressing in the IEEE P1900.4 group [9] towards the specification of protocols carrying information, including context as well as resource selection constraints, between network resource managers and device resource managers enabling distributed decision making which will result in optimization of radio resource usage, in the context of heterogeneous radio access technologies.

The rest of the paper is organized 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 computing the fittingness factor under specific 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 summarizes the conclusions.

II. FITTINGNESS FACTOR DEFINITION

In order to cope with the multi-dimensional heterogeneity reflected in Figure 1, two main levels are identified in the RAT selection problem:

1) Capabilities. A user-to-RAT association may not be possible because of 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 the 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 path loss, intercell interference level, 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, measured interference, etc.).

All these concepts can be captured for each detected RAT in a new measure, the so-called *fittingness factor*, which reflects the degree of adequacy of a given RAT to a given user. The fittingness 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:

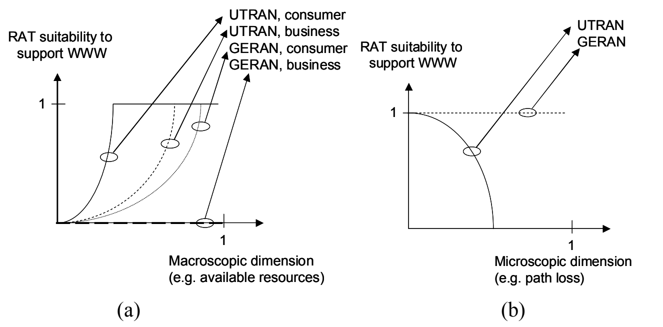
$$\psi_{i,p,s,j} = C_{i,p,s,j} \times Q_{i,p,s,j} \quad (1)$$

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

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

The term $T_{i,p,j}$ 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.

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. To define this suitability, one or both of the macro and microscopic views can be considered, as reflected in Figure 2. 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 if needed. Notice also that the fittingness factor should be defined at a cell-level considering the resources available in a given cell or the path-loss or interference with respect to this cell.



(a) (b)

Figure 2. Suitability as a function of the macroscopic (a) and microscopic (b) dimensions.

Concerning the specific shape and formulation of the function $Q_{i,p,s,j}$ for each service and RAT, several approaches can be used, ranging from empirical functions towards more detailed analytical-based expressions. In the context of this paper, the latter approach will be retained for some specific RATs, as it will be described in Section IV.

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 should be done separately for uplink and downlink, both measurements can be weighted to obtain a unique measurement:

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

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, thus giving the same importance to both links.

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 less loaded RAT).

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 denoted as "servingRAT" and the cell denoted as "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

$$\psi_{i,p,s,j}(k_j) > \psi_{i,p,s,servingRAT}(servingCell) + \Delta_{VHO} \quad (5)$$

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 THE FITTINGNESS FACTOR COMPUTATION

The above framework relays on the so-called suitability function $Q_{i,p,s,j}$ included in the fittingness factor definition in Section II. This function can be defined either empirically or analytically for the different situations. In particular, in this section, some analytical expressions for the computation of this function under specific considerations regarding service and RAT are presented. 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. However, similar expressions could be derived for other RATs (e.g. WLAN, WiMAX, etc.) and services.

A. GERAN

a) Voice users

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

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

where L_{max} is the maximum path loss according to sensitivity level S_{min} and the maximum transmitted 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:

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

This expression is general for the uplink and the downlink. *M* is the multislot capability in the uplink or downlink, $R_{MCS}(L_i)$ 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 $R_{bmax,s,p}$ 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 shared channels depending on the profile *p*. It is computed as the quotient between the number of slots allocated to the service profile *p* in a given frame with respect to the total number of users with the same service and profile that have currently an established TBF (Temporary Block Flow). Then, it basically is a measure of the average number of slots per frame allocated to the profile *p*.

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, as follows:

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

with the maximum path loss L_{max} given as a function of the bit rate $R_{b,i}$, the *Eb/No* target, the maximum transmit power P_{Tmax}, the noise power P_N, the chip rate W=3.84 Mc/s and the measured uplink load factor η_{UL} as [2]:

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

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

$$Q^{DL}_{i,p,VOICE,UTRAN}=\begin{cases}1 & \text{if } P_{Ti}\le \Delta P_{\max,p,s}\\ 0 & \text{if } P_{Ti}>\Delta P_{\max,p,s}\end{cases} \tag{10}$$

with $\Delta P_{\max,p,s}=\min\left(P_{\max,p,s},P_{T\max}-P_T\right)$ 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_{Tmax} is the maximum power available in the downlink). In turn, the power requirement of the new user can be estimated as:

$$P_{Ti}=\frac{L_i RSSI_i-P_T(1-\rho)}{\frac{W}{\left(\frac{E_b}{N_o}\right)_i R_{b,i}}+\rho}=\frac{\frac{P_P}{\left(Ec/Io\right)_i}-P_T(1-\rho)}{\frac{W}{\left(\frac{E_b}{N_o}\right)_i R_{b,i}}+\rho} \tag{11}$$

where $RSSI_i$ is the total received power at the antenna input of the i-th user, which can be expressed as a function of the pilot power P_P and the Ec/Io of the pilot measured by this user. In turn, $R_{b,i}$ is the bit rate, Eb/No the target requirement, W=3.84 Mc/s the chip rate and ρ the orthogonality factor.

b) Interactive users

For the uplink direction the suitability function is given by:

$$Q^{UL}_{i,p,WWW,UTRAN}=\frac{f(R^*)}{R_{b\max,s,p}}\varphi_p \tag{12}$$

where R^* is the maximum bit rate that can be achieved depending on the path loss of the user, the noise power, the load factor and the maximum available power, given by:

$$R^*=\frac{W}{\left(\frac{E_b}{N_o}\right)_i\left(\frac{L_i}{1-\eta_{UL}}\frac{P_N}{P_{T\max}}-1\right)} \tag{13}$$

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

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).

For the downlink direction the suitability function is:

$$Q^{DL}_{i,p,WWW,UTRAN}=\frac{f(R^*)}{R_{b\max,s,p}}\varphi_p \tag{14}$$

where in this case the maximum bit rate R^* depends on the maximum power available for the service and profile $\Delta P_{\max,p,s}=\min\left(P_{\max,p,s},P_{T\max}-P_T\right)$ as follows:

$$R^*=\frac{\Delta P_{\max,p,s}W}{\left(\frac{E_b}{N_o}\right)_i\left(\frac{P_P}{\left(Ec/Io\right)_i}-(1-\rho)P_T-\rho\Delta P_{\max,p,s}\right)} \tag{15}$$

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 with cell radius 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 summarized 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(dB)=128.1+37.6log(d(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) making use of the transport channel type switching procedure. The considered RAB assumes a maximum bit rate of 64 kb/s in the uplink and 128 kb/s in the downlink 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 required sensitivity.

Concerning the fittingness factor evaluation for the vertical handover algorithm, the measurements are averaged in periods of T=1s. The hysteresis margin is Δ_{VHO} =0.1 and T_{VHO}=3s.

Furthermore, $\alpha_{p,s}$=0.5 for all services in order to give equal importance to the uplink and downlink directions in the fittingness factor computation.

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 1, the downlink transmitted power below 42dBm and there must be code sequences available. The active set size equals 1 with a cell replacement hysteresis of 3 dB and a time to trigger handover of 0.64s. The target Block Error Rate is 1% for voice and videophone services and 10% for interactive services, allowing retransmissions in the latter case. 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. For data slots, a maximum of 8 and 32 users can be simultaneously multiplexed in the uplink and downlink, respectively. A handover is triggered whenever the received power either at uplink or downlink is below -100 dBm during 3 consecutive samples.

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)

VI. RESULTS

In the following, some illustrative results regarding the behavior of the fittingness factor to capture the variations in the network are presented. With respect to the microscopic component, Figure 3 plots the time evolution of the average fittingness factor and the path loss corresponding to the central UTRAN cell in the scenario for a certain interactive consumer user. A total of 600 voice, 600 interactive and 100 videophone users are simulated. It can be observed that for high values of the path loss the fittingness factor takes low values, while when the path loss is low the fittingness factor increases, revealing the higher bit rates that can be obtained when the mobile is close to the base station.

The effect of the macroscopic component in the proposed methodology is shown in Figure 4 and Figure 5. They plot the average fittingness factor for voice, videophone and interactive consumer users with different load levels in the scenario. To that end, the number of voice users in the scenario has been varied from 50 to 800 while keeping a total of 600 interactive users (50% consumer and 50% business). The figures plot the cases without videophone users in the scenario (denoted as 0v) and with 100 videophone users (denoted as 100v). Figure 4 shows that the fittingness factor of voice users reaches high values in the two RATs but its reduction when increasing the load (either by increasing the number of voice or videophone users) is more significant in UTRAN than in GERAN. The reason is the cell breathing effect in UTRAN, which reduces the effective coverage when the load is increased.

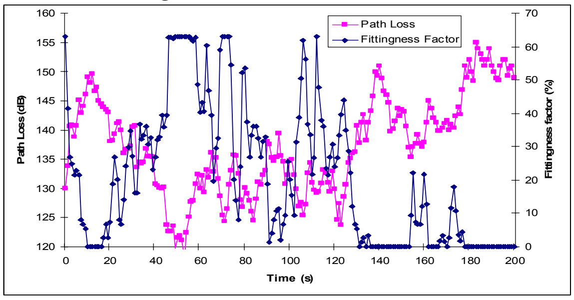


Figure 3. Example of time evolution of the path loss and the fittingness factor for a consumer interactive user in the central cell of UTRAN

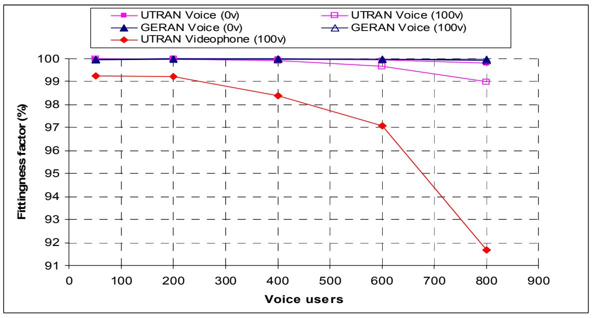


Figure 4. Fittingness factor of voice and videophone users for different loads

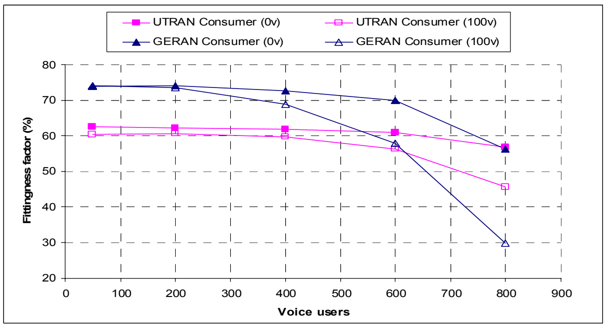


Figure 5. Fittingness factor of consumer interactive users for different loads

Concerning the interactive consumer users in Figure 5, for low load levels, GERAN is in general more attractive because of the higher potential bit rate that can be achieved in both uplink and downlink. In that sense, notice that the maximum potential bit rate in GERAN would be 177.6 kb/s in downlink and 118.4 kb/s in uplink, corresponding to the MCS-9 (i.e. 59.2 kb/s per slot) with 3 and 2 slots, respectively, while in UTRAN the maximum bit rate for this profile is 64 kb/s in uplink and 128 kb/s in downlink. In practice, simulations reveal that for the case with 50 voice and 100 videophone users, an interactive consumer user is allocated on average to 1.28 and 2.88 slots per frame in the UL and DL, respectively in GERAN. However, when increasing the load in the scenario, several transmissions must be multiplexed over the same time slot and as a result the achievable bit rate is significantly reduced. For example, in the case with 800 voice and 100 videophone users, an average of

0.58 and 0.78 slots is allocated to each interactive consumer, which turns into a lower bit rate than the achievable one in UTRAN. Consequently, the fittingness factor in UTRAN becomes higher than in GERAN, as shown in Figure 5 for high numbers of voice users. Concerning business users, although the results are not plot here for the sake of brevity, the fittingness factor is always higher in UTRAN than in GERAN because of the higher downlink bit rate (384 kb/s) that can be allocated to this user profile in UTRAN.

As a result of the previous behavior, Figure 6 plots the percentage of traffic that is served through UTRAN for the different services and profiles as a function of the number of voice and videophone users in the scenario. The inclusion of videophone users, which can only be allocated in UTRAN, has an impact over the allocation of the rest of services. Specifically, when there are no videophone users, around 60% of the voice traffic is served through UTRAN. On the contrary, in the case with 100 videophone users, the higher interference existing in UTRAN will tend to redirect the traffic towards GERAN, and less than 10% of voice traffic is served through UTRAN for the case with 50 voice users. When increasing the number of voice users, GERAN becomes more saturated and as a result more voice users are transferred to UTRAN. A similar effect is observed for interactive consumer users, which are mostly served through GERAN for low voice load levels. In turn, when increasing the voice load, the percentage of consumer traffic served by UTRAN increases because of the reduction in the fittingness factor of GERAN, as explained in Figure 5. Concerning interactive business users, they are mainly served through UTRAN, but the load increase in the scenario, particularly when videophone users are introduced, turns into a higher interference which causes some of the business users to be served through GERAN.

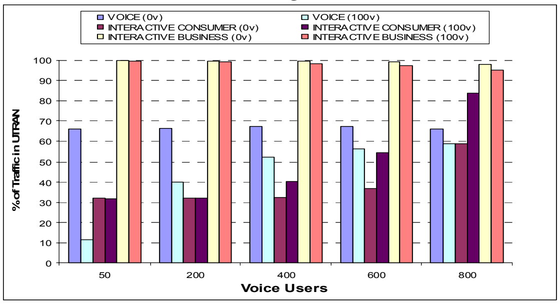


Figure 6. Percentage of traffic allocated in UTRAN for each service

In terms of performance, as an illustrative result, Figure 7 plots the average downlink packet delay for business users obtained with the proposed strategy (denoted as *fittingness* in the figure) for different load levels. For comparison purposes, a classical load balancing allocation scheme, which always selects the RAT having the lowest load, a service-based policy [2], in which voice traffic is served in GERAN and interactive and videophone in UTRAN, and the NCCBvoice strategy presented in [6], which combines a service-based policy with the allocation of low path loss voice users to UTRAN and high path loss voice users to GERAN, are also considered. The proposed strategy based on the fittingness factor algorithm achieves the best behavior among the considered ones.

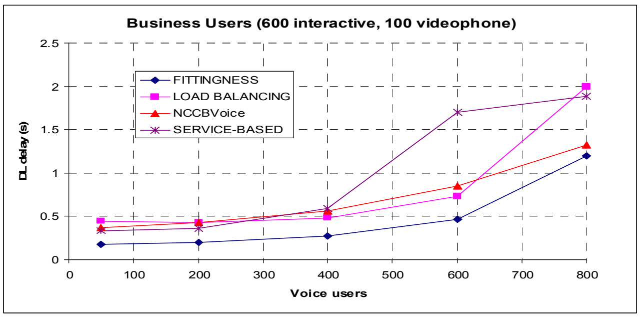


Figure 7. DL packet delay of business users.

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 certain service profile. It allows then capturing several types of heterogeneities regarding terminal and network capabilities. By making use of this metric, a RAT selection algorithm has been presented and evaluated by means of simulations, revealing a good ability to capture the scenario context from both the macroscopic and the microscopic dimensions and a better performance than other approaches in the literature.

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 grant TEC2006-26873-E, and of the COSMOS project 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] S. Lincke-Salecker, “The Benefits of Load Sharing when Dimensioning Networks”, Proceedings of the 37th Annual Simulation Symposium (ANSS’04), April, 2004.

[6] 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.

[7] W. Zhuang, Y. Gan, K. Loh, K. Chua, “Policy-Based QoS Management Architecture in an Integrated UMTS and WLAN Environment”, IEEE Communications Magazine, November, 2003, pp. 118-125.

[8] J. Pérez-Romero et al. “Common Radio Resource Management: Functional Models and Implementation Requirements”, IEEE PIMRC Conference, Berlin, September, 2005.

[9] IEEE P1900 Home Page, http://www.ieeep1900.org

[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”