Common Radio Resource Management Algorithms for Multimedia Heterogeneous Wireless Networks

Miguel López-Benítez, *Student Member, IEEE*, Javier Gozalvez, *Member, IEEE*

Abstract—Heterogeneous wireless systems are envisaged as the integration and joint cooperative management of diverse radio access networks and technologies through which network providers can satisfy the wide variety of user/service demands in a more efficient manner by exploiting their varying characteristics and properties. To achieve this objective, a key tool is common radio resource management techniques designed to jointly manage the radio resources from different radio access technologies. In this context, this work proposes and optimizes new common radio resource management techniques designed to efficiently distribute traffic among the available radio access technologies while providing adequate quality of service levels under heterogeneous traffic scenarios. The obtained results demonstrate the ability of the proposed solutions to provide high user/service satisfaction levels while adequately exploiting the overall system resources.

Index Terms—Beyond 3G heterogeneous wireless systems; common radio resource management; quality of service provision.

✦

1 INTRODUCTION

FUTURE heterogeneous wireless networks, also referred to as Beyond 3G (B3G) networks, are intended to provide an open and flexible architecture for the co-existence of a wide variety of Radio Access Technologies (RATs) with diverse features and capabilities that should be able to support different applications and services with various Quality of Service (QoS) demands. The B3G concept assumes that different radio access networks can be cooperating components of a heterogeneous wireless system through which network providers can satisfy the wide variety of demands in a more efficient manner. The complementary characteristics offered by different wireless technologies in terms of transmission rates, coverage, and cost/revenue, make it possible to exploit the trunking gain leading to a higher overall performance than the aggregated performances of the standalone networks. Nevertheless, the potential gains of B3G networks will only turn into reality by means of an appropriate and efficient management of the whole pool of available radio resources provided by each one of the individual RATs. In this context, the term Common Radio Resource Management (CRRM) is used to refer to the set of functions addressed to ensure an efficient and coordinated use of the available radio resources in a heterogeneous wireless system.

The CRRM concept embraces diverse functionalities. One of the most important functions within the CRRM concept is the RAT selection procedure, which is in charge of deciding the most suitable RAT to transmit the user information. It is important to highlight that diverse service types and user profiles with varying QoS requirements are usually present and not all the available RATs may be able to fulfill such requirements. Therefore, user-to-RAT assignments must be decided to satisfy the user QoS demands and optimally exploit the overall radio resources. While previous work on CRRM and RAT selection has mainly focused on the development of solutions aimed at maximizing the overall system capacity, the design of strategies from the QoS-provision viewpoint has received less attention. In this context, the work reported in this paper is focused on the development of new CRRM techniques designed to efficiently distribute heterogeneous traffic among the available RATs in order to provide appropriate user/service QoS levels while adequately exploiting the available radio resources of each RAT. Different novel criteria for deciding the most suitable user-to-RAT assignment are proposed in this work and evaluated by means of extensive system level simulations.

The rest of this work is organized as follows. First, Section 2 provides a brief overview of the related previous work. Then, the new RAT selection strategies proposed in this work are presented in Section 3. The SPHERE simulation platform, which has been used to assess the performance of the proposed algorithms, is described in Section 4. Section 5 presents and analyzes the obtained results, and Section 6 summarizes this research.

- M. López-Benítez was with the Ubiquitous Wireless Communications Research Laboratory (Uwicore), University Miguel Hernández, Elche, Spain. He currently is with the Department of Signal Theory and Communications, Universitat Politècnica de Catalunya, Barcelona, Spain.
  E-mail: miguel.lopez@tsc.upc.edu
- J. Gozalvez is with the Ubiquitous Wireless Communications Research Laboratory (Uwicore), University Miguel Hernández, Elche, Spain.
  E-mail: j.gozalvez@umh.es

2 RELATED WORK

The CRRM concept has been identified as an important aspect by the 3GPP, which defines some recommendations and architectures for CRRM operation [1], [2]. RAT selection algorithms, however, are not defined by the standardization bodies. Therefore, their development has become an important research topic attracting the interest of the research community during the last years, resulting in a number of solutions published in the literature. For instance, [3] proposes a general framework for the definition of policy-based initial RAT selection strategies. Some simple policies based on prestablished service-to-RAT assignments are also defined and evaluated in [3]. Reference [4] extends the work reported in [3] by considering the impact of multi-mode terminals, i.e. terminals able to operate in different RATs. It is possible to take advantage of the presence of multi-mode terminals by performing a Vertical Handover (VHO) of multi-mode terminals between RATs in order to free the capacity required to accept new calls from single-mode terminals. Several strategies to perform this traffic rearrangement are discussed in [5] and references therein.

The potential benefits of CRRM are evaluated in [6]. The load-balancing strategies employed in the study are aimed at achieving a uniform traffic distribution, which is pointed out to be desirable in order to maximize the trunking gain and minimize the probability of making unnecessary VHOs. For non real-time services, the load balancing is performed based on the measured buffer delay, while for real-time services a load balancing principle based on load thresholds is proposed. In [7], the authors present a refinement where these thresholds are dynamically modified according to the load of the alternative RATs in order to avoid unnecessary VHOs, thus reducing the associated signaling load. In [8], a comparison between load balancing principles and service-based CRRM policies is performed.

Several RAT selection principles based on the signal strength (coverage) and instantaneous load are suggested in [9]. VHOs are performed according to a given signal-to-noise ratio threshold, which is selected to obtain a certain load distribution.

In [10], a terminal-based strategy where users compete for the best RAT, and a network-based strategy where the network chooses the users to serve at any given time are compared. The work assesses the performance gain due to both multi-radio access and multi-user diversity with respect to the non-multi-radio case, where users are constrained to connect to the same RAT.

A theoretical study on the distribution of multiple bearer services onto different subsystems in multi-access wireless systems is shown in [11]. Based on subsystem's multi-service capacities, near-optimum subsystem service allocations that maximize combined multi-service capacity are derived through optimization methods.

As it can be derived from the previous revision, different solutions based on diverse approaches have been proposed in the literature for the RAT selection problem in heterogeneous wireless systems. However, the work in this field is still open and more efforts must be devoted to the development of adaptive and user/service efficient solutions. In this context, this work advances the state-of-the-art by proposing novel algorithms that are not basically based on system performance and requirements but that focus on providing adequate QoS levels under heterogeneous traffic scenarios while optimally exploiting the overall available radio resources.

3 PROPOSED ALGORITHMS

The solutions proposed in this work are characterized by the use of utility functions that numerically quantify the factors taken into account in the RAT selection decision and assign a utility value to each available RAT. These utility functions have been designed in order to provide an adequate trade-off between modeling accuracy (i.e., representation of the considered system aspects) and implementation complexity (i.e., ready availability of the required input information) in real systems. In the first proposed algorithm, the utility function quantifies, based on the perceived channel quality, the estimated throughput that may be expected from each candidate RAT. Therefore, this approach will be referred to as UBET (Utility Based on Expected Throughput) algorithm. In the second proposed algorithm, the transmission rate expected from each candidate RAT is also taken into account, but the utility that each user assigns to a given transmission rate depends on that required by the service in order to achieve an adequate QoS level. Hence, this second approach will be referred to as UBReQoS (Utility Based on Required Quality of Service) algorithm.

3.1 UBET Algorithm

In heterogeneous networks, diverse RATs with different capabilities are available. In principle, high QoS levels could be expected if users were assigned to the RAT with the best capabilities in terms of transmission rate. This policy, however, could lead to undesirable situations where such RAT would remain heavily loaded, even congested, while the resources of other alternative RATs would be underutilized. For a sufficiently high number of users, other alternative RATs experiencing lower loads might offer a suitable performance for the requested service despite their lower transmission rates. These observations suggest the need for more intelligent RAT selection policies in order to exploit the available resources in a more efficient manner. In this context, the UBET algorithm is here proposed. The underlying idea is to provide high QoS levels to users by preferably assigning them to the RAT with the best transmission capabilities until the load level supported by that RAT, which can be associated to the experienced interference levels and thus the measured channel quality, reaches a value such that a better performance is expected from other alternative RATs. To this end, each RAT is

assigned a utility value that represents an estimation of the throughput expected for the user if the RAT is chosen. Every time a RAT selection decision needs to be made for an incoming user, the RAT offering the highest utility value (i.e., expected throughput) is selected.

The data rate expected from a RAT strongly depends on the features of its transmission technology. In general, RATs offer several Transmission Modes (TMs), i.e. various combinations of transmission parameters such as modulation and coding scheme, that provide different trade-offs between resilience to propagation errors and transmission rates. In this case, the TM employed is usually changed dynamically according to the experienced instantaneous channel quality following different Link Adaptation (LA) or Adaptive Modulation and Coding (AMC) techniques [12], [13]. Considering these adaptive policies, the throughput, and hence the utility for each RAT, can be estimated using the BLock Error Rate (BLER) for each TM as a function of the channel quality. Assuming a set of T_m TMs in RAT m, $\mathcal{T}_m = \{\text{TM}_i^m\}_{i=1,2,\ldots,T_m}$, the throughput $\Gamma_{\text{TM}_i^m}(\gamma_m)$ for each transmission mode TM_i^m of RAT m can be expressed as a function of the channel quality γ_m experienced in such RAT as:

$$\Gamma_{\text{TM}_i^m}(\gamma_m) = R_{\text{TM}_i^m} \cdot \left[1 - \text{BLER}_{\text{TM}_i^m}(\gamma_m)\right] \qquad (1)$$

where $R_{\text{TM}_i^m} \in \mathbb{R}^+$ represents the data rate of transmission mode TM_i^m and $\text{BLER}_{\text{TM}_i^m}(\gamma_m) \in [0,1]$ represents the BLER measured for the experienced channel quality γ_m when TM_i^m is used. Assuming as usually done that the TM maximizing the throughput is always employed, the envelope of these curves can be used as an estimation of the throughput that may be experienced if RAT m is selected. Thus, this envelope constitutes the utility function of the UBET algorithm. The utility function for the m-th RAT, denoted as $U_1^m(\gamma_m)$, is thus computed as:

$$U_1^m(\gamma_m) = \max_i \Gamma_{\text{TM}_i^m}(\gamma_m), \qquad i = 1, 2, \ldots, T_m \quad (2)$$

By applying this general method to every available RAT, a set of curves representing an estimation of the expected throughput as a function of the channel quality is obtained for each candidate RAT. Figure 1 shows an example of $U_1^m(\gamma_m)$ for the HSDPA technology.

When computing $U_1^m(\gamma_m)$ it is implicitly assumed that the user will be assigned a channel every time it is requested, completely neglecting the effect of other users in the current cell that are connected to the same RAT, and therefore competing for the available channels. To account for this effect, a second utility function is introduced in the definition of the UBET algorithm:

$$U_2^m(N_{u,\text{eff}}^{m,n}, N_c^{m,n}) = \begin{cases} 1, & N_{u,\text{eff}}^{m,n} \leq N_c^{m,n} \\ \dfrac{N_c^{m,n}}{N_{u,\text{eff}}^{m,n}}, & N_{u,\text{eff}}^{m,n} > N_c^{m,n} \end{cases} \quad (3)$$

where $N_{u,\text{eff}}^{m,n} \in \mathbb{R}^+$ represents the effective number of active users currently connected to the same RAT m and the same cell n as the user that the RAT selection policy is currently addressing, and $N_c^{m,n} \in \mathbb{N}$ represents the
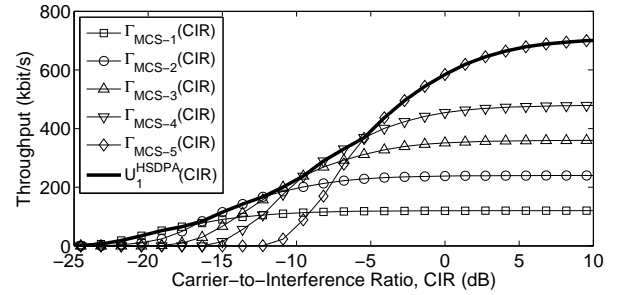

Fig. 1. UBET utility function $U_1^m(\gamma_m)$ for HSDPA.

number of channels available for RAT m in cell n. The concept of *effective number of active users* is introduced in order to account for the possibility offered by some RATs to simultaneously allocate several channels to a single user (i.e., multi-channel operation) and share a single channel among various users (i.e., user multiplexing). The effective number of active users is defined as:

$$N_{u,\text{eff}}^{m,n} = \sum_{l=1}^{N_u^{m,n}} \frac{\alpha_l^{m,n}}{\beta_l^{m,n}} \qquad (4)$$

where $N_u^{m,n} \in \mathbb{N}$ is the actual number of real users (i.e., mobile stations) connected to RAT m in cell n, $\alpha_l^{m,n} \in \mathbb{N}$ represents the number of channels simultaneously allocated to user l in RAT m and cell n, and $\beta_l^{m,n} \in \mathbb{N}$ denotes the total number of users (including user l) sharing the set of $\alpha_l^{m,n}$ channels. Notice that $N_{u,\text{eff}}^{m,n} = N_u^{m,n}$ if each user is assigned a single channel ($\alpha_l^{m,n} = \beta_l^{m,n} = 1$).

The final UBET utility value is computed as:

$$U^m = U_1^m(\gamma_m) \cdot U_2^m(N_{u,\text{eff}}^{m,n}, N_c^{m,n}) \qquad (5)$$

If $N_{u,\text{eff}}^{m,n} \leq N_c^{m,n}$, the allocation of a channel every time it is requested is guaranteed for all cell users. In this case $U_2^m(N_{u,\text{eff}}^{m,n}, N_c^{m,n}) = 1$, meaning that the utility value for RAT m is given by $U^m = U_1^m(\gamma_m)$. On the other hand, if $N_{u,\text{eff}}^{m,n} > N_c^{m,n}$, some users might not receive a channel whenever requested. Considering the worst case scenario in which all cell users simultaneously request a channel and assuming a relatively fair scheduling policy, it has been assumed that if $N_{u,\text{eff}}^{m,n} = r \cdot N_c^{m,n}$ ($r > 1$) in average each user would have access to a channel $U_2^m(N_{u,\text{eff}}^{m,n}, N_c^{m,n}) = 1/r$ times it is requested. Therefore, $U^m = U_1^m(\gamma_m)/r$, meaning that in such a case a lower throughput is expected in RAT m due to the scarcity of channels with respect to the number of active users. The UBET algorithm operation is summarized in Figure 2.

3.2 UBReQoS Algorithm

The UBET algorithm assigns users to the RAT with the highest expected throughput performance regardless of the specific user/service QoS requirements. If users of services with low QoS requirements, which could be satisfactorily served by several RATs, are assigned to the RAT with the best data rates, they could degrade the performance of other users requesting services with
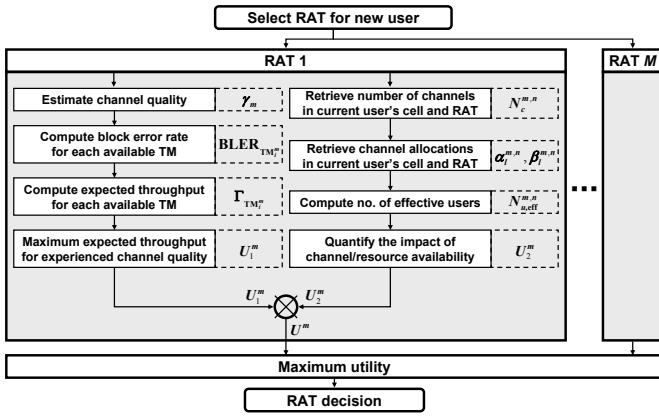
Fig. 2. UBET algorithm operation.

higher QoS requirements that can only be provided by the RAT with the highest capabilities. This situation can be avoided if users are allocated to a RAT offering the required QoS level but not necessarily higher. As a first approach, each service could be preassigned to a RAT according to the required QoS and the RAT capabilities. This simple solution could however be inefficient since a service demand increase may lead to congestion on the associated RAT. A more elaborated solution is therefore required to satisfy the requested QoS levels while exploiting the available resources more efficiently. To this end, the UBReQoS algorithm is here proposed.

The aim of the UBReQoS algorithm is to intelligently distribute the users of a heterogeneous traffic scenario among the available RATs according to the QoS level required by the user/service. To this end, the utility function defined for this algorithm assigns to each candidate RAT a utility value that depends not only on the data rates offered by each candidate RAT, but also on the specific user QoS requirements.

Due to the importance of the service's data rate requirements for the definition of the UBReQoS algorithm, a detailed discussion of this aspect is firstly required. The nominal average data rate needed by the considered service to be satisfactorily provided, hereafter denoted as $R_{\mathrm{nom}} \in \mathbb{R}^+$, may be evident for some services, but not for some others. For real-time and delay-sensitive services, R_{nom} can be easily determined since these services are generally characterized by a nominal average source bitrate (e.g., video at 64 kbit/s). On the other hand, the identification of R_{nom} for many non-real-time services may not be immediate. To avoid ambiguities, a method for deriving R_{nom} for such services is proposed. The idea is to compute a representative packet size value for the service and divide such value by the time interval considered by the user as admissible for each packet transmission. Web browsing and email services are considered here as illustrative examples, although the proposed method could be extended to other services. Considering these two services, the first step is to compute $F_S(\cdot)$, the Cumulative Distribution Function

(CDF) of the web page or email sizes S. To this end, we have computed $F_S(\cdot)$ considering the web browsing and email traffic models mentioned in Section 4. Another option would be to capture the packet sizes on a real network and compute the empirical CDF. Once $F_S(\cdot)$ is obtained, a size value is then derived as $S_\kappa = F_S^{-1}(\kappa)$, where $F_S^{-1}(\cdot)$ is the inverse of $F_S(\cdot)$ and the value of $\kappa \in [0,1]$ depends on the target objectives. In this work, the median packet size value (obtained for $\kappa = 0.5$, i.e. the maximum size for 50% of the samples), is considered unless otherwise stated. Although the average packet size would be simpler to compute, it may not be statistically representative when packet sizes take very dissimilar magnitudes, since a few packets with very low/high sizes can induce appreciable changes in the average value of the whole sequence (this trend was indeed observed in our studies). In this case, the median value usually is a more representative value of the packet sizes present in a packet sequence. As a matter of fact, our studies showed a good match between the median packet size and the most frequently observed packet size for the traffic models cited in Section 4. Following the indications provided by the 3GPP in [14], a web page or email transmission is considered to be satisfactory if performed in less than $\delta_{\max} = 4$ seconds (other values may be defined for other services). For non-real-time services, R_{nom} is then computed as $R_{\mathrm{nom}} = S_\kappa / \delta_{\max}$.

Based on the user's service R_{nom}, a utility function $u_1(R)$ is defined in order to assign a utility value to the different data rates R offered by the candidate RATs:

$$u_1(R) = \begin{cases} \mu \cdot \dfrac{R}{R_{\mathrm{nom}}}, & R < \dfrac{R_{\mathrm{nom}}}{\mu} & \text{(6a)} \\[3mm] 1, & R \geq \dfrac{R_{\mathrm{nom}}}{\mu} & \text{(6b)} \end{cases}$$

where $\mu \in (0,1]$ is introduced for a fine tuning of the algorithm and will be discussed later on. Figure 3 depicts $u_1(R)$ for different values of R_{nom} and μ. Notice that different combinations with the same R_{nom}/μ ratio lead to the same utility function $u_1(R)$.

To understand the rationale of $u_1(R)$, assume $\mu = 1$ at this point. Equation (6a) then corresponds to data rates R that are not enough to satisfy R_{nom}. In this region, $R = R_{\mathrm{nom}}/r$ $(r > 1)$ is assigned a utility $1/r$ times the maximum possible value; this is a simple and straightforward manner to decide the utility of different transmission rates based on the service requirement R_{nom}. On the other hand, for $\mu = 1$ equation (6b) corresponds to data rates R capable to satisfy the required R_{nom}. In this case, the utility does not increase with R and for $R > R_{\mathrm{nom}}$ the user perceives the same utility as for $R = R_{\mathrm{nom}}$. Hence, a RAT providing a rate R_{nom} is sufficient to satisfy the user QoS requirements, and selecting other RATs offering higher transmission rates will not increase utility beyond one. As a result, the user would be assigned to a RAT capable of satisfying the data rate requirements, even if there are other RATs providing higher data rates. This process helps distributing traffic among RATs adequately
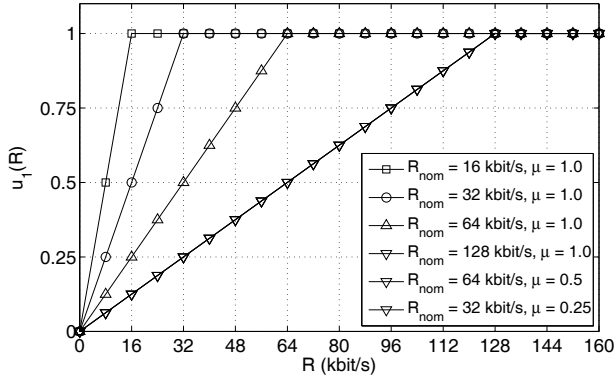
Fig. 3. UBReQoS utility function $u_1(R)$.

according to the real user's QoS needs.

The parameter μ in equation (6) is defined as the utility $u_1(R)$ perceived by the user when the transmission data rate R is equal to R_{nom}. This parameter controls the boundary between the regions of data rates that are regarded as sufficient or not in order to satisfy the user QoS requirements. Although the natural choice for this parameter would be $\mu = 1$, other values are eligible for a fine algorithm tuning. This parameter has been introduced in order to provide a higher degree of flexibility in the design and configuration of the algorithm and hence a more accurate control of the algorithm performance.

Based on equation (6), a utility value can be assigned to each candidate RAT based on the relation between the RAT transmission rate R and the service's R_{nom}. The difficulty in applying equation (6) comes from the fact that the final transmission rate R a user will experience in a certain RAT cannot be predicted beforehand, which depends on the set of available TMs, $\mathcal{T}_m = \{\text{TM}_i^m\}_{i=1,2,\ldots,T_m}$, the corresponding data rates $\mathcal{R}_m = \{R_{\text{TM}_i^m}\}_{i=1,2,\ldots,T_m}$, and how often each TM is used. Since the channel quality a user will experience if assigned a RAT cannot be predicted at the moment of making the RAT selection decision, it is not possible to predict the TMs the user will employ and therefore the resulting rate R and its corresponding utility $u_1(R)$. To address this issue, the decision problem can be characterized from a probabilistic point of view by considering the set $\mathcal{P}_m = \{p_{\text{TM}_i^m}\}_{i=1,2,\ldots,T_m}$, where each element $p_{\text{TM}_i^m}$ represents the probability to use each mode TM_i^m. The probability $p_{\text{TM}_i^m}$ can be computed as the quotient between the number of times that TM_i^m has been selected in previous transmissions and the number of TM selection decisions in RAT m. The probabilities $p_{\text{TM}_i^m}$ would then be updated upon each TM selection decision. Given the sets of transmission rates offered by each candidate RAT, $\mathcal{R}_m$, and their corresponding sets of probabilities, $\mathcal{P}_m$, the utility function $u_1(R)$ can be employed to assign a utility value to each candidate RAT by means of a von Neumann-Morgenstern expected utility function, which computes the expected utility of an arbitrary lottery (set of probabilities) as a linear combination of the individual utilities of its parts [15]. Thus, the utility for each candidate RAT is given by:

$$U_1^m(\mathcal{R}_m, \mathcal{P}_m) = \sum_{i=1}^{T_m} p_{\text{TM}_i^m} \cdot u_1(R_{\text{TM}_i^m}) \qquad (7)$$

The function $U_1^m(\mathcal{R}_m, \mathcal{P}_m)$ assigns to each candidate RAT a utility value that depends not only on the expectable RAT transmission rates but also on the specific user QoS requirements.

Similar to the UBET algorithm, the utility function $U_2^m(N_{u,\text{eff}}^{m,n}, N_c^{m,n})$ defined in equation (3) is considered to account for the negative effects deriving from user overload. Although the RAT selection process depends on the RATs available in a given cell and the cell user load, the interference resulting from neighboring cells can also impact the performance and should therefore be taken into account in the RAT selection process. To include the effect of interfering users, the function:

$$U_3^m(\mathcal{P}_m, \gamma_m) = 1 - \sum_{i=1}^{T_m} p_{\text{TM}_i^m} \cdot \text{BLER}_{\text{TM}_i^m}(\gamma_m) \qquad (8)$$

is introduced, where $\text{BLER}_{\text{TM}_i^m}(\gamma_m)$ represents the same parameter as in equation (1). Since $\text{BLER}_{\text{TM}_i^m}(\gamma_m)$ is a decreasing function of γ_m, $U_3^m(\mathcal{P}_m, \gamma_m)$ decreases as the channel quality γ_m degrades. This method is therefore an adequate way to quantify the impact of the channel quality and the expected usage of each TM on the utility value of the candidate RATs.

The final utility U^m for the m-th RAT is computed as:

$$U^m = U_1^m(\mathcal{R}_m, \mathcal{P}_m) \cdot U_2^m(N_{u,\text{eff}}^{m,n}, N_c^{m,n}) \cdot U_3^m(\mathcal{P}_m, \gamma_m) \quad (9)$$

The user is finally assigned to the RAT with the highest utility value as illustrated in Figure 4. In case the maximum utility value $U^m = 1$ is obtained for more than one RAT, other additional criteria may be employed to make the final decision. The solution adopted in this work is to select the RAT offering the lowest transmission capabilities out of all the RATs achieving $U^m = 1$. This approach has been chosen in order to select a RAT capable to satisfy the required QoS level and reserve resources in RATs offering higher performances for more QoS-demanding services.

3.3 Channel Quality Estimation

Some of the utility functions defined in this work, in particular $U_1^m(\gamma_m)$ of the UBET algorithm and $U_3^m(\mathcal{P}_m, \gamma_m)$ of the UBReQoS algorithm, require the experienced channel quality γ_m to be measured during system operation. Although several methods have been proposed in the literature for this purpose [16], [17], they are usually complex, technology-dependent, and require an important computational cost. In order to increase the applicability of the algorithms proposed in this work and to simplify its practical implementation, we also propose an approximation that can be used to easily estimate the channel quality for each candidate RAT. The proposed
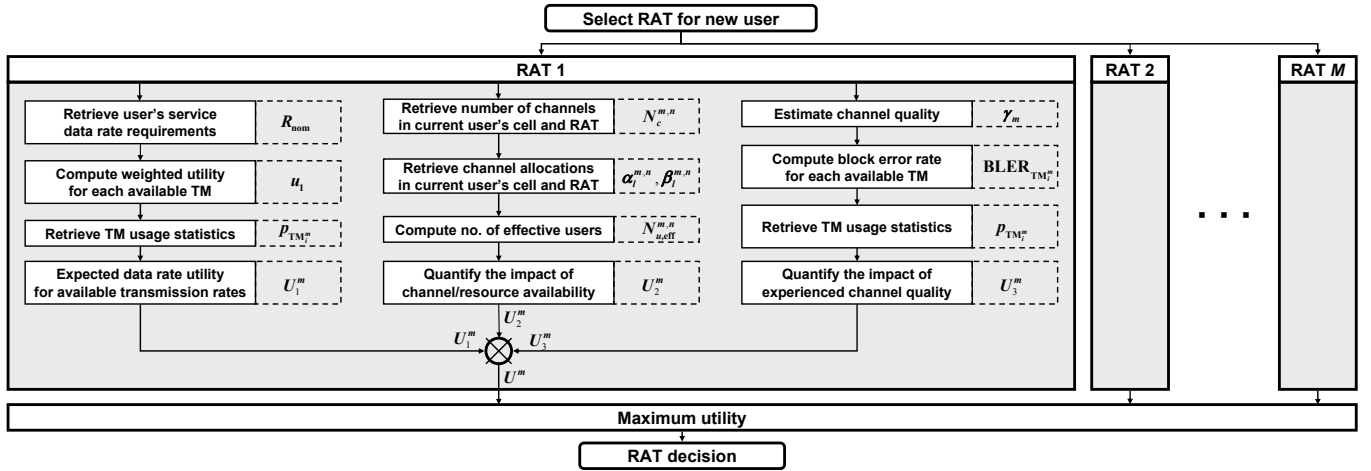
Fig. 4. UBReQoS algorithm operation.

method is not intended to provide an accurate value of the channel quality as in [16], [17] but rather an adequate estimation with a reduced computational cost that simplifies the implementation of the algorithms proposed in this work. To cope with the potential dimensioning effects of channel quality estimation inaccuracies, the approximation here described provides an estimate of the worst case channel quality conditions that might be experienced. Considering the worst case scenario is a common system and network engineering approach.

The suggested approximation is based on the CIR value predicted by a given propagation model according to the number of interfering users. The procedure is here illustrated considering the downlink, but could be used in the uplink as well. The implemented methodology estimates separately the carrier and interference levels. The carrier level is obtained assuming a worst-case situation in which the considered user is located at the border of the cell and the serving base station is in the center. The motivation for this approach is to guarantee that the actual user performance on the selected RAT will not be lower than that estimated based on the channel quality, even in the most unfavorable case. To estimate the interference level a distinction is made between Frequency/Time-Division Multiple Access (FDMA/TDMA) and Code-Division Multiple Access (CDMA) systems. In FDMA/TDMA-based systems, interference originates from co-channel cells separated from the interfered cell by a distance equal to the reuse distance D (considering only the first tier of interfering cells). Interfered and interfering users have been assumed to be situated in the center of their respective cells in order to eliminate mobility effects on the instantaneous channel quality. The CIR value for FDMA/TDMA systems is therefore approximated by:

$$\text{CIR}_{\text{FDMA/TDMA}} = \frac{\dfrac{P_j}{L_P(d_j)}}{\displaystyle\sum_{k\in\Omega} \dfrac{P_k}{L_P(D_k)} + N_0 \cdot W} \quad (10)$$

where P_j is the average transmission power of the desired signal in the reference cell (cell j), $L_P(d_j)$ is the path loss between the center and the border of cell j (i.e., over a distance equal to the cell radius d_j), Ω is the set of co-channel interfering cells, P_k is the average interfering power from the k-th interfering cell, $L_P(D_k)$ is the path loss over a distance equal to the reuse distance D_k from the center of the k-th interfering cell to the center of the considered reference cell, and $N_0 \cdot W$ represents the thermal noise at the receiver in the reference cell, with N_0 being the noise power spectral density and W the channel bandwidth. $L_P(\cdot)$ depends on the selected propagation model, not restricted by this method.

In CDMA-based systems, interference is generated in surrounding cells but also inside the considered cell since multipath fading decreases the orthogonality between the cell's channelization codes, leading to some intra-cell interference. Inter-cell interference is estimated as for FDMA/TDMA systems. For intra-cell interference estimation, the interfered user is supposed to be at the cell border while the transmitting base station is in the center. In this case, the interference power is attenuated by an orthogonality factor α [18, p. 58], with typical values $\alpha \in [0.4, 0.9]$ for multipath channels [19, p. 194]. Thus, for CDMA systems the CIR value is computed as indicated in equation (11), where P_{T_j} and P_{T_k} represent the average total transmitted power in the reference cell j and in the k-th interfering cell, respectively.

$$\text{CIR}_{\text{CDMA}} = \frac{\dfrac{P_j}{L_P(d_j)}}{\displaystyle\sum_{k\in\Omega} \dfrac{P_{T_k}}{L_P(D_k)} + \dfrac{(P_{T_j} - P_j)(1-\alpha)}{L_P(d_j)} + N_0 \cdot W} \quad (11)$$

The proposed procedure offers a simple way to estimate the experienced CIR for each present candidate RAT since it only needs to know the current average transmission power of the interfering cells and the distance to each one of them. The model can be further simplified by assuming that all the interfering cells are

located at a given equivalent distance (according to the actual radio planning) and transmit at a constant average power per interfering user. This model then establishes a direct relation between the number of interfering users and the CIR value, which can be stored in a table. Upon each RAT selection decision, the number of interfering users would be determined (the average number of active users should be exchanged periodically among sets of nearby base stations or core network controlling nodes) and the associated CIR value would be queried in the aforementioned table. With respect to the case in which the experienced CIR is constantly measured during system operation following previous proposals [16], [17], the approximation described in this section allows an easier implementation of the proposed algorithms and requires a significantly lower computational cost since the only requirement is to query a value in a table whenever a RAT selection decision is to be made.

4 SIMULATION PLATFORM

The performance of the proposed algorithms has been assessed by means of extensive system level simulations. These simulations have been carried out using SPHERE, an advanced Simulation Platform for HEterogeneous wiREless systems [20]. The SPHERE platform integrates three advanced system level simulators for GPRS, EDGE, and HSDPA technologies. The specifications of their radio interfaces are implemented in detail and radio transmissions are emulated at the packet level. This confers to the SPHERE platform the ability of accurately evaluating the final user perceived QoS.

Table 1 summarizes the simulation configuration. A cell layout of 27 omni-directional cells with a radius of 500 m is considered, with all GPRS, EDGE and HSDPA base stations co-located in the same site for the 27 cells. To avoid border effects, a wrap-around technique is applied. Once a service request is granted, free channels are assigned randomly. When no free channel is available, the requesting user is queued. For GPRS and EDGE users, a First-Come First-Served (FCFS) scheduling policy is applied, while a round robin policy is used for HSDPA users. Unless otherwise stated, each user is exclusively assigned a single channel ($\alpha_l^{m,n} = \beta_l^{m,n} = 1$). In case of multi-channel operation ($\alpha_l^{m,n} > 1$), allocated channels are not shared among users ($\beta_l^{m,n} = 1$).

The simulation platform implements all the TMs of the simulated RATs and models their adaptive utilization by means of a LA technique, referred to as AMC in HSDPA. LA/AMC selects the TM that, based on certain criterion, optimizes the performance for the experienced channel quality conditions. For non-real-time services, the TM maximizing the throughput in equation (1) is selected [12]. For real-time services, the algorithm proposed in [24] is used since it improves real-time performance.

This work considers three traffic types: web browsing, email, and H.263 real-time video. The implemented models are described in [25], [26] and [27], respectively.

Erroneously received data are retransmitted through an Automatic Repeat reQuest (ARQ) protocol. For GPRS and EDGE, the sliding-window selective-repeat ARQ protocol described in the 3GPP specifications [23] is implemented. For HSDPA, retransmissions are performed by an N-channel Stop-And-Wait (SAW) protocol as detailed by 3GPP in [28]. These retransmission protocols are activated for non-real-time services where transmission reliability is of key importance and some delay can be tolerated. Real-time services are characterized by tight delay constraints and retransmission protocols for these services have been deactivated to avoid excessive delays except for HSDPA, which allows for some retransmission attempts due to its high transmission rates.

Finally, it is worth noting that all the RATs are simultaneously emulated and VHOs are performed dynamically during simulations. This modeling approach, although time and computing resource consuming, is important to realistically and accurately evaluate the performance and implementation costs of CRRM techniques.

5 RESULTS

This section presents the results obtained by simulation and analyzes the performance of the proposed algorithms. Different traffic scenarios have been considered in the simulations. Table 2 shows the number of users per cell in each evaluated traffic scenario. To analyze the performance of the proposed algorithms, several performance metrics have been considered. Since the user perceived performance is strongly determined by the frequency of use of each RAT, the first evaluated aspect is the RAT selection percentage, defined as the quotient (expressed in percentage) between the number of times a RAT is selected and the number of RAT selection decisions. To objectively quantify the user perceived QoS, several traditional figures such as throughput, delay or error rate have been evaluated in the context of this work. Moreover, and to numerically quantify the subjective user QoS perception, another figure of merit, referred to as user satisfaction level, has also been considered. The definition of this parameter depends on the considered service. Following the indications in [14], a web page or email transmission is considered to be satisfactory if it is performed in less than 4 seconds. A real-time video frame transmission is assumed to be satisfactory if completed before the next video frame is generated. The user satisfaction level is then defined as the percentage of times the transmission of a web page, email or video frame is satisfactory according to the previously described service criteria. It is worth noting that, as expected, the objective and subjective measured figures of merit have been observed to be closely related, i.e. higher satisfaction levels are usually associated to higher throughput as well as lower delay and error rate values. As a result, and given length restrictions, our analysis will mainly focus on the user satisfaction level; the throughput, delay and error performances will not

TABLE 1
Configuration of the SPHERE simulation platform.

Parameter	GPRS	EDGE	HSDPA
Environment	Urban macro cellular		
User speed	3 km/h		
Simulated link	Downlink		
Number of cells	27 omni-directional		
Reuse factor	3	3	1
Cell radius	500 m		
Channels/cell	4	4	4
Channel allocation	Random		
Scheduling	FCFS	FCFS	Round Robin
Power/channel	30 dBm	30 dBm	30 dBm
Path loss model	Okumura-Hata COST 231 [21]		
	$f_c = 1.8$ GHz	$f_c = 1.8$ GHz	$f_c = 2.0$ GHz
Shadowing model	Log-normal with standard deviation of 6 dB. Shadowing correlation implemented using the Gudmundson model [22] with de-correlation distance of 20 m.		
Thermal noise	-121 dBm	-121 dBm	-107 dBm
ARQ protocol configuration	Windows size of 64 blocks. Reporting period of 16 blocks.	Window size according to multi-slot class [23]. Reporting period of 32 blocks.	4 SAW processes. Maximum 3 retransmissions.
LA/AMC period	60 ms	60 ms	2 ms

TABLE 2
Traffic scenarios (users/cell).

Service	Scenario		
	I	II	III
Web browsing	2	3	4
Electronic mail	2	3	4
Video 32 kbit/s	4	6	8
Video 64 kbit/s	3	4	6
Video 256 kbit/s	1	2	2
Total	12	18	24

be shown in detail for all scenarios. Finally, another interesting aspect evaluated here is the cost derived from the application of the proposed algorithms, quantified by means of the signaling load in terms of the average number of VHOs requested per minute and per user.

5.1 UBET Performance

Figure 5 shows the RAT usage percentage for the UBET algorithm in all the considered traffic scenarios. This figure shows the obtained performance values, averaged over all the simulated services according to Table 2. However, it is worth noting that the values individually obtained for each service are very similar to those shown in Figure 5. Each figure shows the results when allocating one ($\alpha_l^{m,n} = 1$, $\beta_l^{m,n} = 1$) or several ($\alpha_l^{m,n} > 1$, $\beta_l^{m,n} = 1$) channels per EDGE user. For EDGE multi-channel operation, the number of EDGE channels per cell is not increased but kept constant at the same value as for EDGE single-channel operation, as shown in Table 1. GPRS and HSDPA are always here operated in single-channel mode ($\alpha_l^{m,n} = \beta_l^{m,n} = 1$). As it can be appreciated, when all RATs are operated in single-channel mode, users are assigned to HSDPA

by the UBET algorithm since HSDPA data rates are considerably higher than those offered by GPRS and EDGE in single-channel mode. Such difference explains why even under high user loads (Figure 5c) UBET massively assigns users to HSDPA. If EDGE is operated in multi-channel mode, its offered data rates increase and UBET moves some users from HSDPA to EDGE. This trend accentuates as more channels are simultaneously allocated to an EDGE user and the cell load increases. This behavior indicates that when the candidate RATs provide very different transmission capabilities, UBET mostly selects the RAT offering the higher data rates, but when such RAT capabilities are comparable, UBET is capable to distribute users among RATs according to the offered RAT capabilities. To verify this trend, a network composed of GPRS and EDGE was simulated in single-channel mode. The obtained results indicated that GPRS and EDGE were selected by UBET 33.82% and 66.18% of the times, respectively. In this case, since GPRS and EDGE data rate capabilities are more similar, a more homogeneous traffic distribution is performed by UBET among the available RATs. This result confirms the previous observation about UBET traffic distribution.

To analyze how the UBET operation affects the user perceived QoS, Table 3 shows the user satisfaction level as well as the number of VHOs per minute and per user. First of all, it is worth noting that high satisfaction levels are generally observed with UBET. In fact, under low loads satisfaction values higher than 90% are observed for all configurations. As the user load increases, the satisfaction level decreases due to higher interference levels and lower resource availability. Nevertheless, even under high load conditions (scenarios II and III), the observed satisfaction levels are comprised between 80% and 90%, except for scenario III with EDGE operating
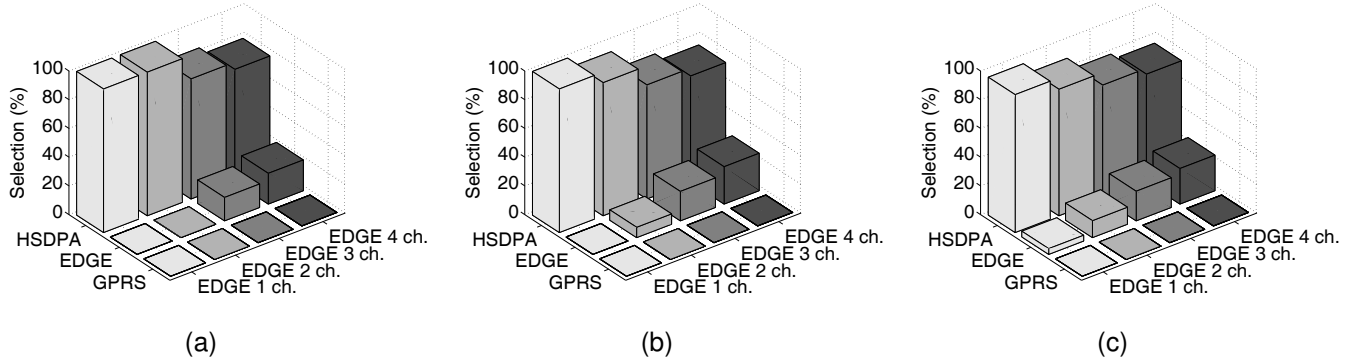
Fig. 5. UBET RAT selection percentage: (a) scenario I, (b) scenario II, and (c) scenario III.

TABLE 3
User satisfaction and average number of VHOs per
minute and per user for UBET.

Scenario	Parameter	EDGE channels per user			
		1	2	3	4
I	Satisfaction (%)	92.73	92.17	91.38	92.29
	VHOs/min/user	0.00	0.00	1.11	1.37
II	Satisfaction (%)	83.09	84.02	87.99	89.51
	VHOs/min/user	0.00	0.26	1.25	1.28
III	Satisfaction (%)	71.64	80.55	85.27	87.85
	VHOs/min/user	0.28	0.78	1.06	1.18

in single-channel mode (71.64%). The obtained results indicate that UBET operates adequately since it generally fulfills its objective, i.e. to provide high user QoS levels.

Another interesting observation from Table 3 is that for scenarios II and III the satisfaction level increases with the number of EDGE channels simultaneously assigned per user. This is due to the fact that UBET assigns a higher number of users to EDGE as the capabilities of EDGE and HSDPA become more comparable, thereby reducing the HSDPA load and increasing its throughput performance. For example, when 4 EDGE channels are allocated per user, the delay experienced by video users in HSDPA improves 39.6% for scenario II and 42.7% for scenario III with respect to the case in which EDGE is operated in single-channel mode. On the other hand, users assigned to EDGE operating in multi-channel mode experience a performance slightly lower than in HSDPA; for example, when 4 EDGE channels are allocated per user, the average throughput experienced for scenarios I, II and III decreseases, respectively, 6.63%, 9.72% and 9.61% with respect to the EDGE single-channel operation case. Although a lower overall throughput performance is observed for EDGE multi-channel operation, the UBET traffic distribution in this case finally results in an increased global user satisfaction level, as shown in Table 3. Therefore, as the maximum RAT capabilities become similar, the user distribution performed by the UBET algorithm results in a higher global user satisfaction level. This improvement

comes at the cost of an increase in the signaling load due to the higher frequency of VHOs between RATs. In the case of scenario I, an increase in the signaling load is also observed as the number of EDGE channels per user augments, but in this case the satisfaction level remains practically unchanged. This circumstance suggests that under low load conditions it may not be advisable to operate RATs with lower capabilities in multi-channel mode since the resulting load distribution among RATs leads to a higher cost in terms of signaling load with no improvements in the user perceived QoS. Under higher load situations, this strategy is able to increase the user satisfaction, as it can be appreciated for scenarios II and III, at the expense of an increased signaling level. In any case, it is worth highlighting that the obtained results suggest that when properly configured UBET is able to provide satisfaction levels greater than 90% for low loads, and close to 90% for moderated and high loads.

Despite the high satisfaction levels that UBET can achieve, it suffers from an inefficient resource utilization since the overall RAT usage is significantly unhomogeneous. Moreover, user QoS requirements are not directly taken into account in the RAT decision process, which could degrade the performance of users requesting services with high QoS requirements as it was pointed out in Section 3.2. To mitigate these potential disadvantages, the UBReQoS algorithm was proposed to try to efficiently distribute users among RATs based on RAT performance capabilities and user/service requested QoS.

5.2 UBReQoS Performance

Figure 6 shows the RAT selection percentage for the UBReQoS algorithm when all RATs are operated in single-channel mode under scenario I (for scenarios II and III the obtained percentages show similar trends). As it can be appreciated, a different RAT selection rate is observed for each service. Web browsing and email users (the least QoS demanding services) are mostly assigned to GPRS since it usually offers sufficient transmission capabilities to provide the services' required data rates (5 and 12 kbit/s for web browsing and email, respectively,
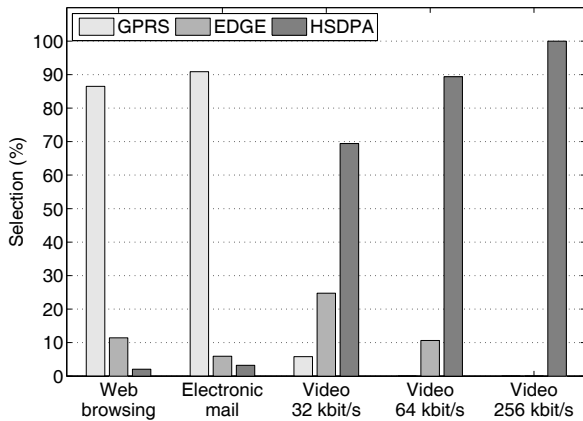
Fig. 6. UBReQoS RAT selection percentage (scenario I).

according to the procedure described in Section 3.2). Users for video at 256 kbit/s (the most QoS demanding service) are always assigned to HSDPA, the only RAT able to meet such QoS requirements. Finally, video users at 32 and 64 kbit/s are mostly distributed between EDGE and HSDPA. These RAT selection trends demonstrate that the UBReQoS algorithm is able to distribute users among the available RATs according to the RAT transmission capabilities and user/service QoS demands.

The first consequence of the UBReQoS user distribution that becomes apparent when comparing Figures 5a and 6 is a greater utilization of the overall available resources with respect to the UBET algorithm, and the fact that UBReQoS provides operators with a greater flexibility to exploit their deployed networks. Accordingly, many users that were usually assigned to HSDPA with the UBET algorithm, even if they did not fully require its transmission capabilities, are now directed to other RATs with lower transmission rates. As a result, the average throughput experienced by various services decreases, but it is important to highlight that the UBReQoS proposal has been designed to satisfy the user QoS demands while intelligently exploiting the radio resources, and not to maximize the system throughput. As it can be appreciated in Table 4, the highest throughput reduction is observed for web browsing and email users since the transmission rates available in GPRS (the most frequently RAT selected by UBReQoS for these services) are significantly lower than those provided by HSDPA (the preferred RAT with UBET). In the case of video users at 32 and 64 kbit/s, which are mostly allocated to HSDPA by UBReQoS, a notable throughput reduction with respect to UBET is also observed in Table 4 since these services are in some cases allocated by UBReQoS to EDGE, a RAT with lower data rates. On the other hand, video users at 256 kbit/s experience a throughput improvement when UBReQoS is employed. Due to the traffic distribution performed by UBReQoS, the load level supported by HSDPA decreases with respect to UBET, leading to lower interference levels and

increased availability of resources in HSDPA. As a result, video users at 256 kbit/s, which are always allocated to HSDPA with both UBET and UBReQoS, experience an increased throughput performance when UBReQoS is employed, with relative improvements of 6.37%, 14.44% and 28.01% for scenarios I, II and III, respectively.

Although the application of the UBReQoS algorithm results in an appreciable throughput reduction with respect to the UBET algorithm for web browsing, email and video at 32 and 64 kbit/s, the assignments performed by UBReQoS for such services are more appropriate in terms of provided QoS levels since the resulting experienced average throughput values shown in Table 4 for UBReQoS are more in accordance with the service required average nominal data rates. As a matter of fact, some questionable behaviors observed for UBET (e.g., average throughput values for web browsing higher than for video at 256 kbit/s), are not appreciated for UBReQoS. Moreover, it is worth noting that the average throughput experienced by the aforementioned services with UBReQoS is appreciably higher than the required average nominal data rate (notice that in general some packets may require peak data rates considerably higher than the average value). As opposed to the previous services, the average throughput performance obtained for video at 256 kbit/s is lower than the required average data rate. However, it is important to clarify here that this circumstance is due to the HSDPA performance upper bound caused by the particular configuration conditions considered in our simulations. Several factors such as the number of users per cell needed to force certain load conditions, the use of omni-directional antennas with low cell radius, even the low GPRS/EDGE reuse factor selected for computational cost reasons, result in important overall interference levels for all RATs. Such interference levels and the HSDPA single-channel operation considered in this research constrain the maximum throughput achieved in HSDPA, which explains the results shown in Table 4 for video users at 256 kbit/s. In terms of throughput performance, the results shown in Table 4 indicate that UBReQoS not only provides adequate QoS levels to web browsing, email and video at 32 and 64 kbit/s according to their nominal data rate requirements, but also is capable to enhance the QoS level provided to the most QoS demanding service (video at 256 kbit/s) with respect to the UBET algorithm.

To numerically quantify the impact of the UBReQoS RAT selection policy on the user perception, Table 5 shows the obtained user/service satisfaction levels. In general, user satisfaction statistics in Table 5 follow the same trend as throughput performance results in Table 4, with the exception of video users at 32 and 64 kbit/s in scenario III, for which the satisfaction level improves despite the lower experienced throughput. This circumstance, which was already observed in Section 5.1, reveals that throughput improvements are not necessarily related to higher satisfaction levels and vice versa. While significant satisfaction reductions are appreciated for

TABLE 4
Average throughput performance (kbit/s) for UBET and UBReQoS ($\kappa = 0.5$, $\mu = 1$).

Service	Scenario I		Scenario II		Scenario III	
	UBET	UBReQoS	UBET	UBReQoS	UBET	UBReQoS
Web browsing	291.49	18.69	246.15	14.29	206.44	13.15
Email	272.81	17.80	221.18	14.16	173.97	13.53
Video 32 kbit/s	237.28	104.34	203.21	78.21	173.48	81.92
Video 64 kbit/s	231.61	185.12	209.00	150.51	180.63	134.29
Video 256 kbit/s	233.27	248.14	210.45	240.83	185.60	237.59

TABLE 5
User satisfaction (%) for UBET and UBReQoS ($\kappa = 0.5$, $\mu = 1$).

Service	Scenario I		Scenario II		Scenario III	
	UBET	UBReQoS	UBET	UBReQoS	UBET	UBReQoS
Web browsing	97.70	59.64	96.98	44.12	94.72	34.16
Email	92.80	57.85	90.68	45.63	88.63	31.50
Video 32 kbit/s	99.53	90.96	95.20	92.38	84.60	93.78
Video 64 kbit/s	96.05	90.30	88.50	88.44	72.19	85.03
Video 256 kbit/s	52.49	55.89	42.58	54.62	27.80	49.72

web browsing and email services as a result of a more frequent use of GPRS, the final user satisfaction for video at 32 and 64 kbit/s is maintained around 90%, even under high load conditions, due to a more frequent assignment of these services to HSDPA. In the case of video at 256 kbit/s, the throughput obtained for this highly QoS demanding service cannot satisfy completely its users. Nevertheles, it is interesting to note that the UBReQoS traffic distribution is capable to increase the user satisfaction level for the most QoS demanding service with respect to UBET: while the improvement for scenario I is equal to 3.40%, better results were observed for other scenarios with higher load levels, in particular, 12.04% for scenario II and 21.92% for scenario III.

Another interesting trend observed in Table 5 is the evolution of user/service satisfaction as system load increases. In general, higher traffic loads result in reduced user/service satisfaction levels. However, different trends are observed for UBET and UBReQoS. When UBET is employed, a load increase severely degrades the satisfaction level for the most QoS demanding services (i.e., video at 256 kbit/s), while satisfaction values for less QoS demanding services (especially web browsing and email) remain practically unaffected. This trend is due to the fact that UBET RAT selection decisions are solely based on the expected RAT transmission performances, completely neglecting the user QoS requirements. On the other hand, the QoS-aware RAT selection criterion of UBReQoS is capable to handle the system load growth in a more adequate manner, thereby preserving the satisfaction level for the most QoS demanding services at the expense of a satisfaction reduction for the least QoS demanding, best-effort services.

As it has been observed, the current aggressive UBReQoS configuration is able to improve the satisfaction of the most QoS demanding services, but reduces sig-

nificantly the QoS level for the least demanding ones. However, it is important to note that UBReQoS offers the possibility to optimize its performance through an adequate tuning and optimization of its configuration parameters κ and μ, which provides operators with a simple and flexible mean to tune the algorithm according to their particular interests. In the following, we discuss the optimization procedure that we have carried out in order to minimize the impact on the least QoS demanding services under different traffic load conditions.

The first UBReQoS configuration parameter, κ, determines the R_{nom} value for non-real-time services (web browsing and email in this research), and thus affects the algorithm behavior only for such services. The second UBReQoS configuration parameter, μ, affects the algorithm behavior for all services by correcting the R_{nom} value for all of them. Based on these observations, a reasonable approach is to first determine an adequate value of κ. In the previous results $\kappa = 0.5$ was assumed, i.e. the R_{nom} value considered for web browsing and email according to the method described in Section 3.2 corresponded to the highest packet size S_{κ} for 50% of the cases, which could explain the low satisfaction levels achieved for these services in Table 5. Therefore, an adequate value of κ should be found with progressive increments until representative packet size S_{κ} and R_{nom} values are obtained. Since the new R_{nom} values for web browsing and email modify the QoS requirements present in the system, a fine tuning of the algorithm behavior through the value of μ should then be performed until the desired performance is obtained.

Based on the previous procedure, the UBReQoS configuration was optimized for scenarios I ($\kappa = 0.9$, $\mu = 0.25$) and II ($\kappa = 0.7$, $\mu = 0.25$) with single-channel operation (see Table 6). When comparing the optimized configurations against the reference configu-

TABLE 6
User satisfaction (%) for UBET and UBReQoS (without and with optimization).

Service	Scenario I			Scenario II		
	UBET	UBReQoS $\kappa = 0.5$, $\mu = 1$	UBReQoS $\kappa = 0.9$, $\mu = 0.25$	UBET	UBReQoS $\kappa = 0.5$, $\mu = 1$	UBReQoS $\kappa = 0.7$, $\mu = 0.25$
Web browsing	97.70	59.64	94.01	96.98	44.12	62.74
Email	92.80	57.85	93.51	90.68	45.63	85.12
Video 32 kbit/s	99.53	90.96	97.27	95.20	92.38	96.86
Video 64 kbit/s	96.05	90.30	94.96	88.50	88.44	95.16
Video 256 kbit/s	52.49	55.89	55.66	42.58	54.62	53.93

ration previously considered ($\kappa = 0.5$, $\mu = 1$), significant performance improvements are observed for the least QoS demanding services. In particular, the satisfaction improvements achieved by the optimized configuration for web browsing and email services are 34.37% and 35.66% for scenario I, and 18.62% and 39.49% for scenario II, respectively. It is interesting to note that this satisfaction improvements are not obtained at the expense of the most QoS demanding ones, since the satisfaction levels for video users at 256 kbit/s are preserved with the optimized configuration. These results indicate that, by means of an appropriate configuration of the UBReQoS algorithm, it is possible to maintain the satisfaction improvement achieved for the most QoS demanding services with respect to the UBET algorithm while minimizing the impact on the rest of services, clearly highlighting the potential of UBReQoS to efficiently organize transmissions in a heterogeneous framework. In fact, the optimization of the UBReQoS algorithm shown in this work clearly indicates that it is possible to adequately serve multimedia traffic services through an adequate and intelligent management of the radio resources.

The UBReQoS algorithm was also compared with a reference case in which each service is always delivered by the same predefined RAT, as in [3]. In this case, the RAT for each service has been selected following the RAT capabilities and the service QoS requirements, resulting in the following assignments: web browsing and email services are always provided by GPRS, video at 32 kbit/s by EDGE, while video at 64 and 256 kbit/s are delivered by HSDPA. Despite also looking to assign services to RATs based on QoS requirements, UBReQoS offers a higher flexibility and is capable to dynamically adapt the user distribution to the instantaneously experienced load and demand conditions. The results shown in Table 7 confirm the potential of UBReQoS to further improve QoS and user satisfaction by dynamically and intelligently managing the available radio resources.

It is important to remark that the intelligent QoS-aware user distribution performed by UBReQoS is capable to dynamically adapt the user distribution to the instantaneously experienced load and demand conditions, offering adequate QoS levels under heterogeneous multimedia traffic conditions while providing operators with a high flexibility to efficiently exploit the overall radio resources of their deployed networks.

5.3 Channel Quality Estimation Impact on System Performance

This section analyzes the impact of simplifying the channel quality estimation (as described in Section 3.3) on the proposed algorithms system performance. To this end, the UBReQoS performance was analyzed under two different settings. In the first one, the utility for each candidate RAT was computed based on the value of γ_m (expressed in terms of CIR) provided by the approximations in equations (10) and (11). The second evaluated setting considered the actual experienced γ_m obtained by averaging CIR samples measured during the simulation, which is equivalent to the use in a real system of methods such as those proposed in [16], [17].

Table 8 shows the results obtained for the two settings under scenario I when UBReQoS is configured with $\kappa = 0.5$ and $\mu = 1$ (other configurations and traffic scenarios show analogous trends). As it can be appreciated, higher satisfaction levels are generally obtained when basing RAT selection decisions on the approximations of equations (10) and (11), with an average satisfaction improvement of 5.1%. Bearing in mind that the channel quality approximation is a more conservative approach, these results can qualitatively be explained as follows. The carrier level is computed assuming a worst-case situation in which the user is located at the cell border. To some extent, this approach provides some guarantees that the RAT regarded by UBReQoS as adequate to satisfy the user QoS requirements will also be adequate even under the most unfavorable conditions. On the other hand, if RAT utility values are computed based on the current experienced channel quality, RAT selection decisions might be based on channel quality conditions that might not be maintained for the remaining duration of the user connection, which can sometimes result in optimistic RAT selection decisions. This situation may occur for users that have been moving along positions near to the base station (the latest CIR measurements would indicate a good channel quality) but then move away (the channel quality degrades). In this case, the RAT initially regarded by the algorithm as an adequate option to satisfy the user QoS requirements may become insufficient to meet the user needs at the new location, resulting in the general satisfaction decrease observed in Table 8. In this context, approximating the channel

TABLE 7
User satisfaction (%) for fixed-assignments strategy and UBReQoS (optimized).

	Scenario I		Scenario II	
Service	Strategy of fixed assignments	UBReQoS $\kappa = 0.9$, $\mu = 0.25$	Strategy of fixed assignments	UBReQoS $\kappa = 0.7$, $\mu = 0.25$
Web browsing	49.79	94.01	42.17	62.74
Email	54.32	93.51	44.98	85.12
Video 32 kbit/s	88.23	97.27	79.67	96.86
Video 64 kbit/s	92.90	94.96	95.40	95.16
Video 256 kbit/s	51.86	55.66	45.79	53.93

TABLE 8
Channel quality estimation impact on user satisfaction (%) for UBReQoS ($\kappa = 0.5$, $\mu = 1$) under scenario I.

Service	Actual CIR	Eqs. (10) & (11)	Difference
Web browsing	52.11	59.64	+7.53
Email	57.40	57.85	+0.45
Video 32 kbit/s	93.43	90.96	-2.47
Video 64 kbit/s	76.79	90.30	+13.51
Video 256 kbit/s	49.14	55.89	+6.75

quality conditions not only enables an easier implementation of the solutions proposed in this work and requires a significantly lower computational cost, but can also provide higher satisfaction levels than accurately measuring the current channel quality conditions.

6 CONCLUSIONS

This work has proposed new CRRM techniques designed to efficiently distribute multimedia traffic in heterogeneous wireless networks among the available RATs while providing adequate QoS levels under heterogeneous traffic scenarios. Different novel criteria for deciding adequate user-to-RAT assignments have been proposed and evaluated by means of extensive system level simulations. The first proposed solution is aimed at providing high global QoS levels by selecting the RAT with the highest estimated performance. The obtained results show that this scheme is able to provide high satisfaction levels but does not fully exploit the capabilities of all the deployed radio networks. On the other hand, the second proposed solution is designed to efficiently distribute users according to RAT transmission capabilities and user/service QoS requirements. The obtained results demonstrate that this second approach is capable to dynamically adapt the user distribution to the instantaneously experienced load and demand conditions, providing adequate QoS levels under heterogeneous multimedia traffic conditions while efficiently exploiting the overall available radio resources. Additionally, in order to increase the applicability of the proposed algorithms and simplify their implementation, this work has also considered a simple approximation that can be used to easily estimate the channel quality conditions. The suggested approximation provides a conservative channel quality estimate that was shown to generally improve the performance of the proposed policies.

REFERENCES

[1] 3GPP, *Improvement of Radio Resource Management (RRM) across RNS and RNS/BSS, TR 25.881, v5.0.0*, Jan. 2002.
[2] ——, *Improvement of Radio Resource Management (RRM) across RNS and RNS/BSS post-Rel-5, TR 25.891, v0.3.0*, Jun. 2003.
[3] J. Pérez-Romero, O. Sallent, and R. Agustí, "Policy-based initial RAT selection algorithms in heterogeneous networks," in *Proc. 7th IFIP Int'l Conf. Mobile and Wireless Comm. Networks (MWCN 2005)*, Sep. 2005, pp. 1–5.
[4] X. Gelabert, J. Pérez-Romero, O. Sallent, and R. Agustí, "On the impact of multimode terminals in heterogeneous wireless access networks," in *Proc. 2nd Int'l Symp. Wireless Comm. Systems (ISWCS 2005)*, Sep. 2005, pp. 39–43.
[5] S. J. Lincke, "Vertical handover policies for common radio resource management," *Int'l Journal of Comm. Systems*, vol. 18, pp. 527–543, Mar. 2005.
[6] A. Tolli, P. Hakalin, and H. Holma, "Performance evaluation of common radio resource management (CRRM)," in *Proc. IEEE Int'l Conf. on Comms. (ICC 2002)*, vol. 5, Apr. 2002, pp. 3429–3433.
[7] A. Tolli and P. Hakalin, "Adaptive load balancing between multiple cell layers," in *Proc. IEEE 56th Vehicular Technology Conf. (VTC 2002)*, vol. 3, Sep. 2002, pp. 1691–1695.
[8] X. Gelabert, J. Pérez-Romero, O. Sallent, and R. Agustí, "On the suitability of load balancing principles in heterogeneous wireless access networks," in *Proc. Wireless Personal Multimedia Comms. Symp. (WPMC 2005)*, Sep. 2005.
[9] O. Yilmaz, A. Furuskär, J. Pettersson, and A. Simonsson, "Access selection in WCDMA and WLAN multi-access networks," in *Proc. IEEE 61st Vehicular Technology Conf. (VTC 2005)*, vol. 4, Apr. 2005, pp. 2220–2224.
[10] R. Veronesi, "Multiuser scheduling with multi radio access selection," in *Proc. 2nd Int'l Symp. Wireless Comm. Systems (ISWCS 2005)*, Sep. 2005, pp. 455–459.
[11] A. Furuskär and J. Zander, "Multiservice allocation for multiaccess wireless systems," *IEEE Trans. Wireless Comms.*, vol. 4, no. 1, pp. 174–184, Jan. 2005.
[12] ETSI-SMG, *EDGE Feasibility study, Work item 184; Improved data rates through optimized modulation, Tdoc 97-331*, Dec. 1997.
[13] M. López-Benítez and J. Gozalvez, "Link adaptation algorithms for improved delivery of delay- and error-sensitive packet-data services over wireless networks," *Wireless Networks*, Dec. 2008.
[14] 3GPP, *Services and service capabilities, TS 22.105, v8.4.0*, Jun. 2007.
[15] J. von Neumann and O. Morgenstern, *Theory of games and economic behavior*. Princeton University Press, 1944.
[16] M. Turkboylari and G. L. Stüber, "An efficient algorithm for estimating the signal-to-interference ratio in TDMA cellular systems," *IEEE Trans. Comms.*, vol. 46, no. 6, pp. 728–731, Jun. 1998.

[17] D. Ramakrishna, N. B. Mandayam, and R. D. Yates, "Subspace based estimation of the signal to interference ratio for CDMA cellular systems," in *Proc. IEEE 47th Vehicular Technology Conf. (VTC 1997)*, vol. 2, May 1997, pp. 735–739.

[18] J. Laiho, A. Wacker, and T. Novosad, *Radio network planning and optimisation for UMTS*. Wiley, 2002.

[19] H. Holma and A. Toskala, *WCDMA for UMTS: Radio access for third generation mobile communications*, 3rd ed. Wiley, 2004.

[20] M. López-Benítez, M. C. Lucas-Estañ, and J. Gozalvez, "A dynamic radio simulation platform for the study of radio resource management techniques in heterogeneous wireless systems," in *Proc. 9th Int'l Symp. Modeling, Analysis and Simulation of Wireless and Mobile Systems (MSWiM 2006)*, Oct. 2006, pp. 143–171.

[21] COST 231, *Urban transmission loss models for mobile radio in the 900 and 1800 MHz bands, revision 2, TD(91)73*, Sep. 1991.

[22] M. Gudmundson, "Correlation model for shadow fading in mobile radio systems," *Electronics Letters*, vol. 27, no. 23, pp. 2145–2146, Nov. 1991.

[23] 3GPP, *Radio Link Control / Medium Access Control (RLC/MAC) protocol, TS 44.060, v8.3.0*, Dec. 2008.

[24] J. Gozalvez, M. López-Benítez, and O. Lázaro, "Link adaptation algorithm for improved wireless transmission of delay-sensitive packet data services," *Electronics Letters*, vol. 41, no. 14, pp. 813–815, Jul. 2005.

[25] P. Barford and M. Crovella, "Generating representative web workloads for network and server performance evaluation," in *Proc. Int'l Conf. Measurement and Modeling of Computer Systems (SIGMETRICS/PERFORMANCE 1998)*, Jun. 1998, pp. 151–160.

[26] J. Ho, Y. Zhu, and S. Madhavapeddy, "Throughput and buffer analysis for GSM general packet radio service (GPRS)," in *Proc. IEEE Wireless Comms. and Networking Conf. (WCNC 1999)*, Sep. 1999, pp. 1427–1431.

[27] O. Lázaro, D. Girma, and J. Dunlop, "H.263 video traffic modelling for low bit rate wireless communications," in *Proc. IEEE Int'l Symp. Personal, Indoor, and Mobile Radio Comms. (PIMRC 2004)*, Sep. 2004, pp. 2124–2128.

[28] 3GPP, *Physical layer aspects of UTRA High Speed Downlink Packet Access, TR 25.848, v4.0.0*, Mar. 2001.

Miguel López-Benítez received a Telecommunications Engineering degree (first-class honors) in 2006 from University Miguel Hernández (UMH), Elche, Spain, and was the recipient of the 2006 University Education National Award, a distinction from the Spanish Ministry of Education and Science to the best national academic records. His degree thesis was awarded a national research prize from the France Telecom Foundation. He currently is pursuing a Ph.D. degree at Universitat Politècnica de Catalunya (UPC), Barcelona, Spain, with the support of a FPU grant from the Spanish government. He has been and is actively involved in European- and Spanish-funded projects. Visit www.lopezbenitez.es for more details.

Javier Gozalvez received an electronics engineering degree from the French Engineering School ENSEIRB, a DEA in electronics from the Université de Bordeaux I, France, and a Ph.D. in mobile communications from the University of Strathclyde, Glasgow, U.K. Since October 2002, he has been with the Signal Theory and Communications Division of the University Miguel Hernández, Elche, Spain, where he is currently an Associate Professor and Director of the Uwicore Research Laboratory. At Uwicore, he is leading research activities in the areas of wireless vehicular communications, radio resource management, heterogeneous wireless systems, and wireless system design and optimization. He currently serves as Mobile Radio Senior Editor of IEEE Vehicular Technology Magazine, and previously served as Associate Editor of IEEE Communication Letters, and on the editorial board of the Journal of Communications and the Journal of Networks (Academic Publishers). He was TPC Co-Chair of the 2009 IEEE 69th Vehicular Technology Conference Spring, and General Co-Chair of the 3rd International Symposium on Wireless Communications Systems (ISWCS'2006). He is also the founder and General Co-Chair of the IEEE International Symposium on Wireless Vehicular Communications (WiVeC) in its 2007, 2008 and 2010 editions.