QoS Provisioning in Beyond 3G Heterogeneous Wireless Systems through Common Radio Resource Management Algorithms

Miguel López-Benítez Javier Gozalvez
Signal Theory and Communications Division, University Miguel Hernández
Avda. de la Universidad, s/n 03202 Elche (Spain)
j.gozalvez@umh.es

ABSTRACT
Beyond 3G mobile communication systems are being defined as the integration of diverse Radio Access Technologies (RATs) into what is generally known as heterogeneous wireless systems. One of the main challenges that such systems must overcome is the ability to guarantee the interoperability and efficient management of the different RATs in order to provide the user with a suitable and consistent Quality of Service (QoS) level. To this end, one of the key elements that must be considered by the network provider is the Common Radio Resource Management (CRRM) of the different RATs. The work reported in this paper is focused on the development of new CRRM techniques designed to efficiently distribute traffic across the diverse RATs of a heterogeneous wireless network. The main objective of this work is to provide solutions that are able to fulfil user QoS requirements and to optimally exploit overall system resources.

Categories and Subject Descriptors
J.2 [**Physical Sciences and Engineering**]: Engineering.

General Terms
Algorithms, Management, Performance, Design.

Keywords
Heterogeneous wireless systems, quality of service, common radio resource management, radio access technology selection.

1. INTRODUCTION
While cellular communications are witnessing the emergence of current third generation (3G) systems, new proposals are being defined by the research community for Beyond 3G (B3G) or fourth generation (4G) systems. Some initiatives are developing new radio access technologies for B3G/4G systems. However, most of the research community envisages such systems as the integration and joint management of current cellular systems, Wireless Local Area Networks (WLAN), Digital Broadcasting Systems (DBS), and any potential new radio access technologies that might appear in the future. This concept assumes that different radio access networks can be cooperating components in a heterogeneous wireless infrastructure, through which network providers can more efficiently achieve the required QoS levels.

In this context, one of the main challenges that such systems must overcome is the ability to guarantee the interoperability and efficient management of the different Radio Access Technologies (RATs) in order to provide the user with a suitable Quality of Service (QoS) level. To this end, one of the key elements that must be considered by the network provider is the Common Radio Resource Management (CRRM) of the different RATs.

The CRRM concept embraces several techniques of diverse nature. One important function within the CRRM concept is the RAT selection function, which is in charge of deciding the RAT that must be selected for the transmission of the information of each user in the heterogeneous network. As a first approach, all users could always be assigned to the RAT with better transmission capabilities. However, this solution could lead to an undesirable situation where a RAT is highly loaded, even saturated, and the resources of other available RATs in the system are unused. This discussion suggests that users in a heterogeneous network may be distributed across the different RATs of the system in order to achieve a better utilisation of the radio resources available on the system. However, very distinct RATs with different capabilities and diverse service types with dissimilar QoS requirements are usually present in a heterogeneous system, and not all radio access alternatives are able to fulfil the QoS expectations of all users. Therefore, user assignments must be decided in such a way that not only overall radio resources in the system are optimally exploited but also diverse QoS requirements are satisfied.

The work reported in this paper is focused on the development of new CRRM techniques designed to efficiently distribute the traffic across the diverse RATs of a heterogeneous wireless network. Users in the system are assigned to the RAT that results optimum according to a certain criterion. Different criterions for deciding such an assignment are proposed and evaluated in this work.

2. RELATED WORK
Several approaches and algorithms for RAT selection have been proposed in the literature. For instance, in reference [9] a general framework for the definition of policy based initial RAT selection strategies is proposed and some specific examples based on pre-established assignments are defined and evaluated.

Q2SWinet'06, October 2, 2006, Torremolinos, Malaga, Spain.

The potential benefits of CRRM are evaluated in [10]. The RAT selection strategies employed in this study are aimed at achieving a uniform traffic distribution, which is pointed out to be desirable to maximise the trunking gain and to minimise the probability of making unnecessary handovers. For non real-time services, the load balancing is performed based on the measured buffer delay of each RAT. For real-time services the load balancing is based on load thresholds that trigger vertical handovers between RATs.

In [12] the authors propose several RAT selection principles that are based on the signal strength (coverage) and instantaneous load. Vertical handovers are performed according to a signal-to-noise ratio threshold, which is selected in order to obtain a certain load distribution. Furthermore, reference [12] also presents other RAT selection principles where the QoS that can be offered by each RAT is also taken into account through the expected transmission rate. The work reported in reference [11] is another example where RAT selection is also based on the expected QoS. A terminal-based strategy where users compete for the best RAT, and a network-based case 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 single radio access.

A theoretical study on the distribution of multiple bearer services onto different subsystems in multi-access wireless systems is performed in [2]. Based on the included subsystem's multi-service capacities, near-optimum subsystem service allocations that maximise combined multi-service capacity are derived through simple optimisation procedures.

As 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 relatively recent and more efforts must be devoted to the study of new solutions. In this context, this work proposes new algorithms aimed at satisfying user QoS expectations and optimally exploiting overall system resources.

3. PROPOSED ALGORITHMS

In this section, new CRRM algorithms designed to efficiently distribute the traffic across the diverse RATs of a heterogeneous wireless system are presented and described. Users in the system are assigned to the RAT that results optimum according to a given criterion. In this work, different criterions for deciding such an assignment are proposed and evaluated. All them are characterised by the use of a utility function, which numerically quantifies all the factors that are taken into account in the decision. The RAT with the highest utility is selected for transmission.

In the first proposed algorithm, the utility function quantifies, based on the perceived channel quality, the estimated throughput that could be obtained from every RAT of the system. Therefore, this algorithm will be referred to as UBET (Utility Based on Expected Throughput) algorithm. In the second proposed algorithm, expected transmission rate is also taken into account, but the utility that a given user assigns to a given transmission rate is relative to that required by its own service to be satisfactorily provided. Thus, this algorithm will be referred to as UBReQoS (Utility Based on Required Quality of Service) algorithm.

The definition of these algorithms will be illustrated considering a heterogeneous wireless system where the GPRS (General Packet Radio Service), EDGE (Enhanced Data-rates for GSM/Global Evolution), and HSDPA (High Speed Downlink Packet Access) technologies are integrated. This scenario will be considered in section 5 to assess the performance of the proposed algorithms.

3.1 UBET Algorithm

In a heterogeneous wireless system, diverse RATs with different capabilities are available. High QoS levels could be achieved if users were assigned to the RAT with the best transmission capabilities. This is true while system load remains at low or moderated levels. However, if all users are directed to the same RAT and the load increases, resource availability decreases and higher interference levels are experienced. In this case the actual performance of the loaded RAT decreases, and other alternative and unused RATs, even with lower capabilities, might offer a better performance. The aim of this algorithm is to provide high QoS levels to users by preferably directing them to the RAT with the best capabilities, until that RAT is loaded and a better estimated performance is expected from other alternative RATs.

To this end, each RAT of the system is assigned a utility value that represents an estimation of the throughput that can be expected by the user if that RAT is selected. In order to obtain the utility value, a set of curves representing the BLock Error Rate (BLER) as a function of the channel quality, expressed in terms of the Carrier-to-Interference Ratio (CIR), is employed for each transmission mode (modulation and coding scheme) of each RAT. The throughput can be represented as a function of CIR as

$$Throughput = R \cdot [1 - BLER(CIR)] \quad (1)$$

where R is the bit-rate corresponding to the transmission mode considered. Assuming that the transmission mode maximising the throughput is always selected, the envelope of these curves can be utilised as an estimation of the throughput that could be achieved if a given RAT is selected for transmission. This envelope is therefore selected as utility function U_I. A utility function U_I is obtained for each RAT, which represents an estimation of the expected throughput as a function of the channel quality.

The estimated throughput will depend on the experienced channel quality. When a larger number of users are assigned to a certain RAT, interference levels increase and channel quality decreases, thus reducing the expected throughput for that RAT. In such a case, an alternative RAT experiencing lower loads might offer a better throughput performance and, hence, a higher utility. Some users would be assigned to the less loaded RAT where they would obtain a higher throughput, thus improving global QoS and preventing from saturation in the RAT with the best capabilities.

It is worth noting that the utility function U_I previously discussed offers an estimation of the throughput that could be expected from a given RAT. This estimation takes into account the effect of interfering users through the CIR value, which represents the experienced channel quality and determines the estimated throughput. However, when computing this value it is assumed that a mobile will be assigned a channel every time it is requested, completely neglecting the effect of users in the current cell that are connected to the same RAT, and therefore competing for the available channels. If the number of users in the current cell is large enough in comparison with the number of available channels for the considered RAT, not all users would be assigned a channel every time it is solicited, and this will affect the actual throughput performance. In order to include this effect in the estimated throughput, a second utility function U_2 is introduced. The final utility U for a RAT is then given by $U = U_1 \cdot U_2$. The U_2 utility function is defined as

$$U_2 = \begin{cases} 1 & N_u \le N_c \\ N_c / N_u & N_u > N_c \end{cases} \quad (2)$$

where N_u represents the number of users in the current cell that are connected to the considered RAT, and N_c represents the number of available channels for that RAT in the considered cell. If the number of users is lower than the number of available channels, the allocation of a channel every time it is requested is guaranteed for all users in the cell. Thus, for $N_u \le N_c$, $U_2 = 1$ and $U = U_1$ meaning that the estimated throughput is given by U_1. However, if the number of users is greater than the number of available channels, some users could not be assigned a channel when it is needed. For instance, if $N_u = 2 \cdot N_c$, it can be considered as a simple approximation that in average each user would have access to a channel half the times it is requested (actually, this could heavily depend on the scheduling policy). Therefore, for $N_u = 2 \cdot N_c$, $U_2 = 0.5$ and $U = 0.5 \cdot U1$, meaning that the expected throughput in such a case is assumed to be 0.5 times the estimation given by U_1.

The value of the function U_2 can be easily obtained since it only depends on the number of users that are connected to the considered RAT in the current cell. However, the computation of the U_1 value requires the interference levels to be estimated during system operation. In order to simplify this process, we propose an approximation that can be used to easily determine a CIR value for each RAT. The suggested approximation has been elaborated considering the downlink and is based on the CIR value predicted by a given propagation model according to the number of interfering users. The carrier level is obtained assuming the worst case, i.e. the considered user is located at the border of the cell and the serving base station is in the centre. Interference levels are estimated considering only the first tier of interfering cells. A distinction is made between users in a FDMA/TDMA system and users in a CDMA system. In FDMA/TDMA systems such as GPRS and EDGE, interference proceeds from co-channel cells separated from the interfered cell by a distance equal to the reuse distance R (considering the first interfering tier). Interfered and interferers have been assumed to be situated in the centre of their respective cells in order to eliminate mobility effects. Therefore, for GPRS and EDGE, the CIR value is estimated as

$$CIR_{GPRS/EDGE} = \frac{\dfrac{P_i}{L_P(r)}}{N_j \cdot \dfrac{P_j}{L_P(R)} + N_0 \cdot W} \quad (3)$$

where P_i is the transmission power of the desired signal in the reference cell (cell i), $L_P(r)$ is the path loss between the centre and the border of the reference cell, i.e. over a distance equal to the radius r of the reference cell, N_j is the number of interfering transmitters in co-channel interfering cells, P_j is the transmission power of the interfering cells (assumed for simplicity to be the same in all interfering cells), $L_P(R)$ is the path loss over a distance equal to the reuse distance R, and $N_0 \cdot W$ represents the thermal noise at the receiver in the reference cell, with N_0 being the noise spectral density and W the channel bandwidth.

In CDMA systems such as HSDPA, interference proceeds from surrounding cells but also from inside the cell, since multipath fading decreases the orthogonality between channelisation codes and some intra-cell interference is therefore observed. Inter-cell interference is estimated assuming the same location of interfered and interferer users as for FDMA/TDMA systems. For intra-cell interference estimation, the interfered user is supposed to be at the border of the cell while interferers are assumed to be located at the centre of the cell. In this case interference power is attenuated by an orthogonality factor α [7]. Thus, for HSDPA the CIR value is computed as indicated by expression (4), where N_i represents the number of intra-cell active channels.

$$CIR_{HSDPA} = \frac{\dfrac{P_i}{L_P(r)}}{N_i \cdot \dfrac{P_i \cdot (1-\alpha)}{L_P(r)} + N_j \cdot \dfrac{P_j}{L_P(R)} + N_0 \cdot W} \quad (4)$$

The proposed procedure offers a simple way to establish a direct relation between the number of interfering users and the CIR value, which can be stored in tables and consulted every time it is needed, instead of estimating the CIR during system operation to compute the value of the utility function U_1.

To summarise, the operation of the UBET algorithm is as follows. In every RAT selection decision, the value of the utility function U is computed for each RAT. To obtain the value of the utility function U_1, the number of interfering users is determined and this value is then mapped to a utility value by means of a stored table whose values have been computed following the procedure described above. The number of users in the current cell is then mapped to a U_2 value as shown in expression (2) . Both values are multiplied in order to obtain a utility value U for each RAT. The user is assigned to the RAT with the highest utility value.

3.2 UBReQoS Algorithm

The criterion of the UBET algorithm described in the previous section is to assign users to the RAT with the highest expected throughput performance. This assignment is executed neglecting the possibility that users with low QoS requirements could be satisfied in other RATs with lower capabilities. Users with low QoS requirements who could be satisfied in diverse RATs but are occupying the RAT with the best capabilities could degrade the performance of users who require higher QoS levels and can only be satisfied by the RAT with the best capabilities. This situation could be avoided if users are assigned a RAT with capabilities in accordance with the required QoS level. As a first approach, each service type could be pre-assigned to a given RAT according to the required QoS. However, more elaborated principles are required since an increase on the demand of a certain service could saturate the associated RAT. The aim of the UBReQoS algorithm proposed in this section is to intelligently distribute users among the RATs of a heterogeneous system according to the required QoS level. To this end, the utility function defined for this algorithm assigns to each RAT a utility value that depends on the specific user QoS requirements.

We first introduce a utility function u_1 that is defined according to the bit rate R_{nom} required by the service to be adequately provided. Before describing this function, it is important to clarify the meaning of the R_{nom} parameter for different service types. Real-time services are clearly characterised by a required bit rate, which can be considered as R_{nom} for the application of the UBReQoS algorithm. However, the identification of this value for non real-time services is not immediate. We propose a method for deriving this value for non real-time services such as web browsing or email. The first step is to compute the cumulative distribution function (CDF) of the packet sizes. A packet corresponds to a web page or an email. We have used the size of the packets generated with the traffic models described in section 4. Once the CDF is computed, a packet size value is derived depending on the specific objectives. We have considered the value corresponding to a probability equal to 0.5 in the CDF, i.e. the maximum packet size for the 50 % of the samples. Assuming that a web or email transmission is assumed to be satisfactory if it is performed in a time interval lower than 4 seconds (as indicated in 3GPP 22.105), the packet size derived from the CDF is divided by 4 seconds. The value obtained with this procedure is then considered as the value of R_{nom} for web and email services.

In the definition of the u_1 function, a nominal utility coefficient μ is employed. This parameter is defined as the utility perceived by the user when the estimated bit rate is equal to the nominal mean bit rate R_{nom} of the service. This parameter therefore defines the utility u_1 as a function of an established relation between the estimated bit rate and the nominal bit rate R_{nom} required by the considered service. Figure 1 illustrates how this parameter and the utility function u_1 are related. As it can be appreciated, the function u_1 is a normalized function. For an estimated bit rate equal to zero, the utility is always equal to zero. As the estimated bit rate increases, the utility also increases until the maximum value is reached. The slope of the curve is determined by the μ parameter and the nominal bit rate R_{nom} of the service.

The interest of such definition of the utility function u_1 is as follows. Consider one user of a certain service who perceives the maximum utility when the estimated bit rate equals R_{nom}, i.e. μ = 1.0. For bit rates greater than R_{nom} the user perceives the same utility as for R_{nom}, meaning that a RAT capable of providing a bit rate of at least R_{nom} kbps will result sufficient to satisfy the user. Higher bit rates will not increase the utility perceived by the user, so there is no need to assign the user to other RAT with higher transmission capabilities. Therefore, the user would be assigned to a RAT capable of satisfying its bit rate requirements, although another RAT with higher bit rates is available in the system. On the other hand, users of services characterised by a high R_{nom} value will be assigned to RATs with better capabilities, since RATs that offer bit rates under R_{nom} cannot satisfy such users. The utility function u_1 is therefore intended to distribute the traffic among the RATs of the system according to the QoS required by each service, expressed in terms of the bit rate.

Considering u_1 to determine the utility of the RATs on the system, an estimated bit rate must be determined for every RAT. To this end, the transmission modes (TMs) of each RAT must be taken into account. Different TMs offering a certain range of bit rates are usually available for radio transmission, and are dynamically changed according to the experienced channel quality. The actual performance of a RAT is strongly determined by the set of available TMs and the frequency of use of each one. For a RAT with T available TMs, a utility value U_1 is computed as

$$U_1 = \sum_{i=1}^{T} p(TM_i) \cdot u_1(TM_i) \quad (5)$$

where $p(TM_i)$ denotes the probability (percentage) of utilisation of the transmission mode TM_i, and $u_1(TM_i)$ represents the utility u_1 associated to the bit rate of the transmission mode TM_i.

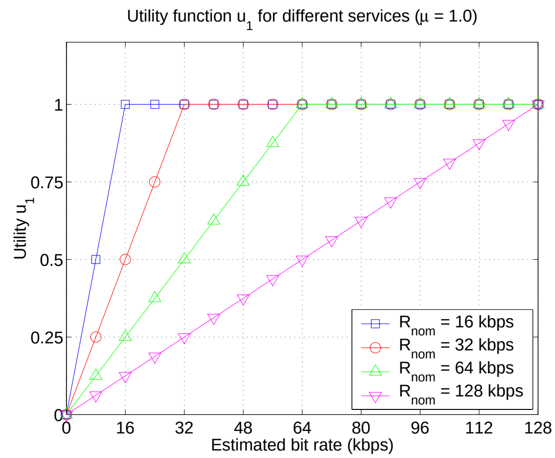


Figure 1. Utility function u_1 for different services (μ = 1.0).

As for the UBET algorithm, the utility function U_2 defined in expression (2) is considered to account for the negative effects on the performance that occurs when having to serve a population of users larger than the number of available channels. Moreover, users producing interference degrade the experienced channel quality, and therefore affect the actual performance. In order to include the effect of interfering users, a third utility function U_3 is introduced, which is defined as

$$U_3 = 1 - \sum_{i=1}^{T} p(TM_i) \cdot BLER(TM_i) \quad (6)$$

where $p(TM_i)$ denotes the probability (percentage) of utilisation of the transmission mode TM_i, and $BLER(TM_i)$ represents the BLER of the transmission mode TM_i for the experienced channel quality. This last term can be derived from curves relating the BLER performance to the experienced CIR. The procedure described in section 3.1 for mapping the number of interfering users to a CIR value can also be employed. Note that the value returned by the utility function U_3 decreases as the BLER value increases, i.e. as the channel quality degrades. The RAT offering the highest utility value $U = U_1 \cdot U_2 \cdot U_3$ is selected for transmission.

4. SIMULATION PLATFORM

The performance of the algorithms proposed in this work is assessed by means of system level simulations. These simulations have been carried out using SPHERE, an advanced Simulation Platform for HEterogeneous wiREless systems. The SPHERE platform integrates three advanced system level simulators of the GPRS, EDGE, and HSDPA technologies. The specifications of their radio interfaces are implemented in detail. Moreover, radio transmissions are emulated at the packet level, which allows an accurate evaluation of the final user perceived QoS.

Table 1 summarises the configuration used for the simulations carried out in this work. A cell layout of 27 omni-directional cells with a radius of 500m is considered. In order to avoid border effects, a wrap-around technique has been applied. Coverage from all RATs is provided in each cell. Users are assigned free channels randomly. When no free channel is available, the requesting user is then placed in a queue. For GPRS and EDGE queued users, a First Come First Served (FCFS) scheduling policy is applied. Users in the HSDPA queue are served in a round robin fashion.

Table 1. Configuration of the simulation platform.

Parameter	GPRS	EDGE	HSDPA
Environment	Urban macro cellular		
Simulated link	Downlink		
No. of cells	27		
Reuse factor	3	3	1
Cell radius	500 m		
Channels/cell	4	4	4
Ch. allocation	Random		
Scheduling	FCFS	FCFS	Round Robin
Power/channel	30 dBm	30 dBm	30 dBm
Path loss Model	Okumura-Hata COST 231		
	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 and decorrelation distance of 20 m		
Thermal noise	−121 dBm	−121 dBm	−107 dBm
ARQ protocol configuration	Window size of 64 blocks. Reporting period of 16 blocks.	Window size according to multislot class. Reporting period of 32 blocks.	4 SAW processes. Maximum 4 transmissions.
LA/AMC upd. period	60 ms	60 ms	2 ms

The simulation platform implements the transmission modes of all considered RATs and models its adaptive utilisation for all RATs by means of a Link Adaptation (LA) technique, which is referred to as Adaptive Modulation and Coding (AMC) for HSDPA. This technique periodically selects the transmission mode that results optimum for the experienced channel quality according to a given criterion. For non real-time services, the transmission mode that maximises the throughput is selected. For real-time services, the algorithm proposed in [3] has been used since it outperforms the former in several key aspects for real-time services.

Erroneously received data is retransmitted by an Automatic Repeat reQuest (ARQ) protocol. For GPRS and EDGE, a selective ARQ protocol is implemented as described in 3GPP specifications. For HSDPA, retransmission of erroneous data is performed by a N-channel Stop-And-Wait (SAW) protocol. These retransmission protocols are activated for non real-time services since for such services the transmission reliability is an aspect of paramount importance and some delay can be tolerated. However, real-time services are characterised by tight delay constraints. Retransmission protocols are deactivated for real-time services in order to avoid excessive delays. Erroneous data is therefore discarded when received in error. However, some retransmission attempts are allowed in HSDPA for real-time services before a data block is discarded, since HSDPA transmission modes are characterised by high transmission bit rates.

Three traffic types have been considered in the context of this work and modelled in the employed simulation platform, namely web browsing [1], email [5], and H.263 real-time video [8].

Finally, it is worth noting that all the RATs of the heterogeneous wireless system are simultaneously emulated and RAT changes are performed dynamically during simulations. It is important for a heterogeneous wireless system simulation platform to be able of modelling this type of situations in order to allow for a realistic and accurate evaluation of the techniques under study.

5. PERFORMANCE EVALUATION

This section presents the results obtained through simulations and analyses the performance of the proposed algorithms. Different traffic scenarios have been considered in the simulations. Table 2 shows the number of users in each evaluated traffic scenario.

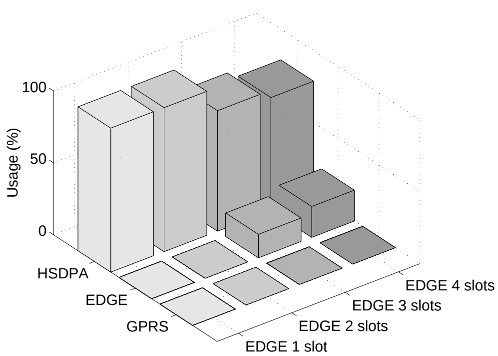
Table 2. Number of users in each evaluated traffic scenario.

Scenario	Web	Email	H.263 32 kbps	H.263 64 kbps	H.263 256 kbps
I	2	2	4	3	1
II	3	3	6	4	2

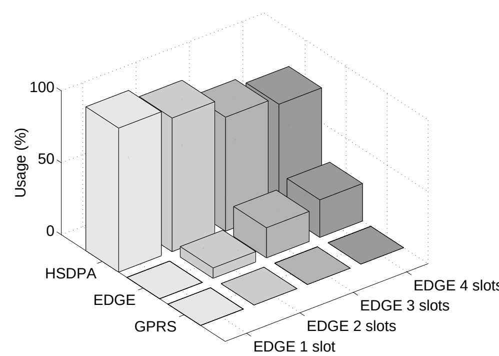
5.1 UBET Performance

Figure 2 shows the usage percentage of each RAT for the UBET algorithm considering scenarios I (Figure 2a) and II (Figure 2b). The usage percentage of a RAT is defined as the quotient between the number of times the RAT is selected and the number of RAT selection decisions. Each figure shows the results when EDGE is operated in single-channel and multi-channel mode. GPRS and HSDPA are always operated in single-channel mode. As shown in Figure 2a, when all RATs are operated in single-channel mode, users are always directed to HSDPA by the UBET algorithm. This is due to the considerable difference between the transmission capabilities of HSDPA and the rest of RATs, which leads to a significantly higher U_1 value for HSDPA. Even if the number of HSDPA users (12 in scenario I) is three times the number of available channels, i.e. U_2 = 1/3, HSDPA is still preferable for the UBET algorithm since HSDPA bit-rates are significantly higher to those offered by the other RATs. If EDGE is operated in multi-channel mode, the RAT selection percentage of EDGE increases. This trend accentuates as more channels are simultaneously allocated to each EDGE user. This behaviour indicates that when the RATs of the system offer very different transmission capabilities, the one with the best performance is massively selected. When the transmission capabilities of the RATs are comparable, users are distributed among them. In order to verify this statement, a heterogeneous wireless system composed by GPRS and EDGE, both operating in single-channel mode, was simulated. In this case, GPRS and EDGE were respectively selected for 33.82% and 66.18% of the RAT selection decisions, which confirms the previous assumption.

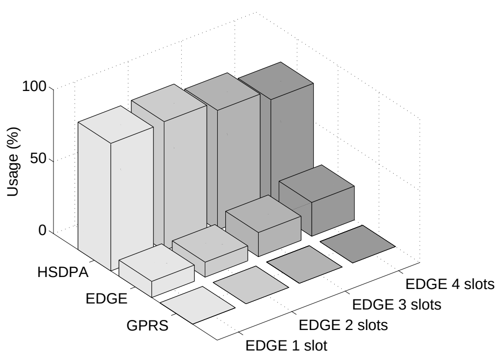
In order to analyse the behaviour of the UBET algorithm when system load increases, scenario II was simulated. The results are shown in Figure 2b. Comparing Figures 2a and 2b, a higher usage of EDGE is appreciated when system load increases. At low or moderated loads, the RAT with better transmission rates (HSDPA in this case) is expected to offer the best performance. However, as the number of users directed to the same RAT increases, interference levels within the RAT also increases and channel quality degrades, thus reducing the performance that could be expected. For instance, a reduction of the HSDPA experienced throughput of 11.4% was observed for scenario II with respect to scenario I. When the previous situation occurs, other RATs with



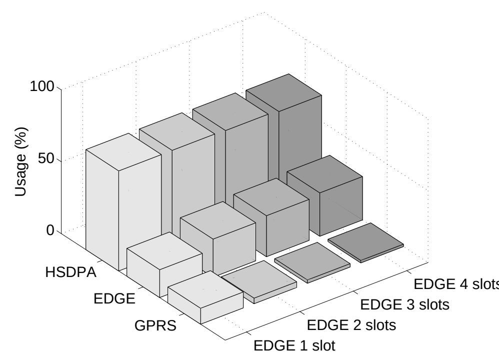
(a) UBET algorithm, scenario I



(b) UBET algorithm, scenario II



(c) Modified UBET algorithm, scenario I



(d) Modified UBET algorithm, scenario II

Figure 2. Usage percentage of each RAT for the UBET and modified UBET algorithms.

lower capabilities but experiencing lower load could offer a better performance, which is the case observed when comparing Figures 2a and 2b, especially for EDGE multi-channel operation. More users are therefore moved from HSDPA to EDGE as system load increases, which prevents HSDPA from saturating. However, this trend is timidly appreciated, even null, for the case where EDGE operated in single-channel mode. This means that when all RATs are operated in single-channel mode, HSDPA continues accepting new users generated by a load rise. As a result, the availability of resources decreases and the HSDPA queue occupation increases, as shown in Figure 3. This figure shows the average number of users in the HSDPA queue per second (for different EDGE multi-channel operation cases). To compute this parameter, the number of queued users in HSDPA is measured every second, and the values are averaged at the end of the simulation. As can be seen, there exists a relevant difference between EDGE 1/2 slots and EDGE 3/4 slots. This is observed for both scenarios. The reason for this is that for EDGE 1/2 slots, the number of users moved from HSDPA to EDGE when the load increases is very low, even null for EDGE 1 slot (see Figures 2a and 2b). As a result, the averaged number of queued users in scenario II considerably increases with respect to scenario I.

Users waiting in the HSDPA queue for a free resource could be served with acceptable QoS levels if they were directed to an alternative RAT with lower capabilities but with free resources. This circumstance suggests the design of a variant of the initial UBET algorithm where the availability of free resources in the selected RAT is checked before the assignment is performed. If no free channels are available, the next RAT with the highest utility value is then selected, and so on. The results for this modified version of the UBET algorithm are shown in Figures 2c and 2d. As it can be appreciated, the usage of alternative RATs increases with respect to the initial UBET algorithm, especially at higher loads (compare Figures 2b and 2d). The difference is more relevant for EDGE 1/2 slots. As a result, an important reduction on the number of queued users is observed in Figure 3 for EDGE 1/2 slots when the modified version of the UBET algorithm is applied. It was observed that the number of queued users in the system is reduced at the expense of a slight reduction of EDGE active users performance, since some new users are incorporated to EDGE. However, the HSDPA load reduction resulted in a throughput improvement, as shown in Table 3.

Table 3. Relative improvement of HSDPA throughput performance for modified UBET with respect to initial UBET.

EDGE slots	Scenario I	Scenario II
EDGE 1 slot	0.08 %	2.49 %
EDGE 2 slots	6.44 %	5.56 %
EDGE 3 slots	6.56 %	2.40 %
EDGE 4 slots	3.41 %	0.87 %

Finally, user satisfaction is measured as follows. A web or email user is assumed to be satisfied when a web page or email is received in a time interval shorter than 4 seconds (3GPP 22.105). H.263 real-time video users are assumed to be satisfied when a video frame is transmitted before the next video frame is generated. The user satisfaction is then defined as the percentage of times the user is satisfied. User satisfaction values around 90% were observed in most of the cases shown in Figure 2. This result indicates that the strategy of the UBET algorithm, which was designed to assign users to the RAT offering the best estimated performance, achieves high global satisfaction values.

5.2 UBReQoS Performance

The performance of this algorithm is compared against the performance of the algorithm based on assigning each service type always to the same RAT, which is considered as reference for evaluation purposes. For the application of the UBReQoS algorithm, $\mu = 1$ and the R_{nom} values assumed for web and email are derived from the 50% point of the corresponding CDFs as described in section 3.2. For the application of the reference algorithm, web and email users are assigned to GPRS, H.263 video users with R_{nom} equal to 32 kbps are assigned to EDGE, and H.263 video users with R_{nom} equal to 64 and 256 kbps are assigned to HSDPA. All RATs on the system are operated in single-channel mode.

Table 4 shows the RAT usage percentage for the UBReQoS algorithm. As it can be seen, each service type is often directed to a given RAT, but not constrained to always transmitting in that RAT as it is the case of the reference algorithm. The UBReQoS algorithm offers more flexibility in this sense since it dynamically adapts traffic distribution to changing load and transmission conditions. As a result, user satisfaction is significantly improved, which can be appreciated in Table 5. All services experience a higher satisfaction with the UBReQoS algorithm, except 64 kbps H.263 users. However, the satisfaction for this service is still arround 90 % with the UBReQoS algorithm. The satisfaction of H.263 users is also assessed by means of the percentage of H.263 video frames transmitted with a BLER value lower than 5%, which was pointed out in [4] to be required in order to obtain an acceptable image quality. The percentage of video frames with BLER lower than 5% and transmitted within the required deadline is also analysed (see Table 6). Although the values are not greater than 57% in the best case, it is important to note how the considered simulator configuration has influenced these results. On one hand, a 3-cell reuse factor was selected for GPRS and EDGE (see Table 1). This decision was necessary in order to reduce the computational cost. The low reuse factor employed in simulations favours the existence of important interference levels in GPRS and EDGE. Furthermore, common BLER values in HSDPA for the first radio transmission on macro-cellular environments are between 30% and 60% [6] (this also justifies the decision of allowing some retransmissions for H.263 users in HSDPA). As a result, the system configuration considered in this work imposes limits to the maximum percentage of frames that can be transmitted with a BLER lower than 5%.

As it can be seen in Table 6, while the value of these parameters for 64 kbps and 256 kbps video users remains approximately constant, for 32 kbps video users experiences an enhancement of 38.29% and 35.58% respectively, which represent an important improvement on the QoS perceived by such users.

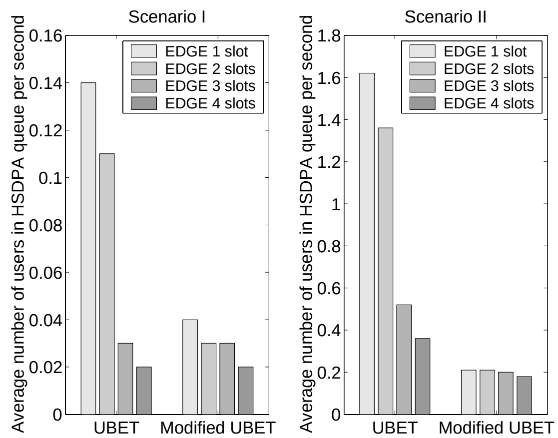


Figure 3. Average number of users in HSDPA queue per second.

Table 4. Usage percentage of each RAT for the UBReQoS algorithm (scenario I, single-channel operation).

Service	GPRS	EDGE	HSDPA
Web	86.51	11.41	2.08
Email	90.85	5.94	3.21
32 kbps	5.80	24.77	69.42
64 kbps	0.00	10.64	89.36
256 kbps	0.00	0.00	100.00

Table 5. User satisfaction (%) for UBReQoS and reference algorithms (scenario I, single-channel operation).

Service	Reference	UBReQoS	Improvement
Web	49.79	59.64	+ 9.85 %
Email	54.32	57.85	+ 3.53 %
32 kbps	88.23	90.96	+ 2.73 %
64 kbps	92.90	90.30	– 2.60 %
256 kbps	51.86	55.89	+ 4.03 %
Global	85.15	85.92	+ 0.77 %

Table 6. Percentage of video frames transmitted with BLER ≤ 5% (a), and with BLER ≤ 5% and without delay (b).

	Service	Reference	UBReQoS	Improvement
(a)	32 kbps	18.58	56.87	+ 38.29 %
	64 kbps	42.30	42.95	+ 0.65 %
	256 kbps	20.82	20.76	– 0.06 %
(b)	32 kbps	16.48	52.06	+ 35.58 %
	64 kbps	42.27	40.26	– 2.01 %
	256 kbps	19.02	18.82	– 0.20 %

Table 7. Throughput performance (kbps) for UBReQoS and reference algorithms (scenario I, single-channel operation).

Service/RAT	Reference	UBReQoS	Improvement
Web	15.64	18.69	+ 19.50 %
Email	15.41	17.80	+ 15.51 %
32 kbps	27.31	104.34	+ 282.06 %
64 kbps	227.91	185.12	– 18.77 %
256 kbps	230.29	248.14	+ 7.75 %
GPRS	15.57	15.66	+ 0.58 %
EDGE	27.46	41.30	+ 50.40 %
HSDPA	261.71	277.65	+ 6.09 %

The better performance of the UBReQoS algorithm is also shown in Table 7, where the throughput performance is presented. The performance of this parameter for 32 kbps H.263 video users is considerably enhanced at the expense of reducing the throughput experienced by 64 kbps H.263 video users. However, the final throughput performance of both services is great enough to obtain high satisfaction values, as manifested in Table 5. Furthermore, the throughput values measured in each RAT indicate that the dynamic adaptation of the UBReQoS algorithm leads to a better interference distribution, and hence to the better observed results. In addition, BLER improvements up to 18% were observed for the UBReQoS algorithm with respect to the reference algorithm.

Comparing UBET and UBReQoS algorithms, it was observed that the UBET algorithm at the expense of degrading the satisfaction level of the most demanding users, i.e. 256 kbps video users, obtains high global satisfaction values. However, the obtained simulation results shown that the UBReQoS algorithm offers similar global user satisfaction values than UBET and modified UBET algorithms, even with slight enhancements at high loads. Moreover, higher satisfaction values were observed for 256 kbps H.263 video users when the UBReQoS algorithm was employed, with improvements up to 7% and 18% with respect to the UBET and modified UBET algorithms, respectively. This enhancement is due to the fact that the UBReQoS algorithm intelligently distributes the traffic among the RATs of the system according to the required user QoS and the capabilities offered by each available RAT on the heterogeneous system, instead of directing users to the RAT with the highest expected performance (usually HSDPA for case considered in this work). Since users of less demanding services are directed to other alternative RATs, they do not degrade the performance of 256 kbps H.263 video users, who imperatively need to be assigned to HSDPA in order to be satisfactorily served. It is important to highlight that satisfaction values shown in Table 5 256 kbps H.263 users are limited by the configuration considered for the simulations carried out in this work. While similar global user satisfaction values are obtained for both UBET and UBReQoS algorithms, the QoS level experienced by the users of the most demanding service is significantly improved when the UBReQoS algorithm is applied.

6. CONCLUSIONS

Beyond 3G heterogeneous wireless systems are envisaged as the integration and joint management of diverse radio technologies. This concept assumes that different radio access networks can be cooperating components in a heterogeneous wireless infrastructure, through which network providers can more efficiently achieve the required QoS levels. To this end, one of the key elements that must be considered by the network provider is the Common Radio Resource Management (CRRM) of the different radio access technologies. In this context, this work has proposed new CRRM algorithms designed to efficiently distribute the traffic across the diverse RATs of a heterogeneous wireless system. The first one is aimed at achieving high global QoS levels by assigning users to the RAT with the highest estimated performance. The obtained results show that high satisfaction values are obtained for all the services considered in this work. The second proposed algorithm was designed to distribute traffic among the RATs of the system according to the user required QoS and the capabilities of each RAT. This solution provides adequate QoS levels and leads to a high overall system resource utilisation.

7. ACKNOWLEDGMENTS

This work has been supported by the Ministry of Education and Science (Spain) and FEDER funds under the project TEC2005-08211-C02-02, and by the Generalitat Valenciana (Spain) under the project GV05/189.

8. REFERENCES

[1] Barford, P. and Crovella, M. Generating representative web workloads for network and server performance evaluation. In *Proceedings of the 1998 ACM SIGMETRICS Joint International Conference on Measurement and Modeling of Computer Systems (SIGMETRICS'98/PERFORMANCE'98)* (Madison, United States, June 24-26, 1998). ACM Press, New York, NY, USA, 151–160.

[2] Furuskär, A. and Zander, J. Multiservice allocation for multiaccess wireless systems. *IEEE Transactions on Wireless Communications*, (Jan. 2005), 4(1):174-184.

[3] Gozálvez, J., López-Benítez, M. and Lázaro, O. Link adaptation algorithm for improved wireless transmission of delay-sensitive packet data services. *Electronic Letters*, (Jul. 2005), 813-815.

[4] Hanzo, L., Cherriman, P. and Streit, J. *Wireless Video Communications: Second to Third Generation Systems and Beyond*. IEEE Press, 2001.

[5] Ho, J., Zhu, Y. and Madhavapeddy, S. Throughput and buffer analysis for GSM general packet radio service (GPRS). In *Proceedings of IEEE Wireless Communications and Networking Conference (WCNC'99)* (New Orleans, United States, September 21-24, 1999), 1427–1431.

[6] Holma, H. and Toskala. A *WCDMA for UMTS*. Wiley, 2002.

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

[8] Lázaro, O., Girma, D. and Dunlop, J. H.263 video traffic modelling for low bit rate wireless communications. In *Proceedings of IEEE International Symposium on Personal, Indoor, and Mobile Radio Communications (PIMRC'04)* (Barcelona, Spain, September 5-8, 2004), 2124–2128.

[9] Pérez-Romero, J., Sallent, O. and Agustí, R. Policy-based initial RAT selection algorithms in heterogeneous networks. In *Proceedings of Mobile and Wireless Communication Networks (MWCN'05)* (Marrakech, Morocco, September 19-21, 2005), 1–5.

[10] Tolli, A., Hakalin, P. and Holma, H. Performance evaluation of common radio resource management (CRRM). In *Proceedings of IEEE International Conference Communications (ICC'02)* (New York, United States, 28 April-2 May, 2002). 5:3429-3433.

[11] Veronesi, R. Multiuser scheduling with multi radio access selection. In *Proceedings of 2nd International Symposium on Wireless Communication Systems (ISWCS'05),* (Siena, Italy, September 5-7, 2005), 455-459.

[12] Yilmaz, O., Furuskär, A., Pettersson, J. and Simonsson, A. Access selection in WCDMA and WLAN multi-access networks. In *Proceedings of IEEE 61st Vehicular Technology Conference (VTC'05)* (Stockholm, Sweeden, 30 April – 1 May, 2005), 4:2220-2224.