Joint radio resource management for heterogeneous wireless systems

J. Gozalvez[a], M.C. Lucas-Estañ[a*], J. Sanchez-Soriano[b]

[a] Ubiquitous Wireless Communications Research (Uwicore) Laboratory
[b] Center of Operations Research (CIO)
University Miguel Hernandez of Elche, Avda. de la Universidad s/n, 03202, Elche, Spain.
j.gozalvez@umh.es, m.lucas@umh.es, joaquin@umh.es

Abstract— **Heterogeneous wireless systems are characterized by the physical coexistence of a variety of radio access technologies with different, but also complementary, technical characteristics and performance. A key aspect of heterogeneous systems is then the implementation of efficient joint radio resource management mechanisms. In this context, this paper presents and evaluates novel joint radio resource management techniques based on the CEA bankruptcy distribution rule. The proposed policies base their distribution decisions on the system conditions and the varying quality of service requirements present in multimedia scenarios. The obtained results demonstrate that the proposed policies can efficiently distribute the radio resources with a low computational cost.**

Index Terms— **Heterogeneous wireless systems, Joint Radio Resource Management, bankruptcy.**

1 Introduction

Future wireless networks will consist of a variety of Radio Access Technologies (RATs) physically coexisting. This framework will offer network operators the capability to manage in a coordinated way the distinct performance and technical characteristics offered by the different RATs. The management possibilities are also increased by the simultaneous coexistence of multimedia applications with varying Quality of Service (QoS) requirements. In this context, a key aspect of heterogeneous wireless systems is the design of advanced Joint Radio Resource Management (JRRM) policies aimed at guaranteeing the user QoS demands, and maximising the network performance, capacity and revenue. JRRM techniques are in charge of deciding for each incoming call, the RAT over which it will be conveyed (initial RAT selection), and the number of radio resources within the selected RAT (intra-RAT RRM) that will be necessary to satisfy the user/service QoS demands.

To date, most research efforts have been devoted to the design of initial RAT selection techniques. For example, [1] described the framework over which JRRM algorithms can be developed, and proposed some basic techniques to address the initial RAT selection dilemma. A key factor influencing the

* Corresponding author. Tel.:+34 965222031, fax: +34 966658903
E-mail address: m.lucas@umh.es (M. C. Lucas-Estañ)

performance of heterogeneous systems is the system load. In this context, [2] proposed load balancing strategies aimed at achieving a uniform traffic distribution in order to maximize the trunking gain and minimize the probability of making unnecessary vertical handovers. The proposals reported in [3] use vertical handovers of multi-mode terminals to free the capacity needed for incoming calls from single-mode terminals. It is also worthwhile highlighting the study reported in [4] that shows the importance of considering the different service QoS requirements and RAT capabilities in the design of a JRRM solution.

First proposals to jointly address the RAT selection and intra-RAT RRM dilemmas have been proposed in [5] and [6]. The JRRM algorithm proposed in [5] is based on neural networks and fuzzy logic, and considers factors such as the signal strength, resource availability and mobile speed. This algorithm decides the optimum RAT for the incoming call, and the bit rate that has to be granted to each user but does not determine the number of radio resources that must be assigned to each user to achieve the identified bit rate. The proposal is evaluated considering users with equal QoS demand, thereby overlooking the challenges and opportunities characteristic of multimedia scenarios. The proposal in [6] employs Hopfield neural networks and bases its decision criterion on user QoS constraints defined in terms of required bit rate or maximum delay. The reported technique simultaneously decides the optimum RAT for the incoming call, and the necessary radio resources at the assigned RAT. Although the authors mention the distinct nature of radio resources from different RATs, the reported study does not consider such nature when addressing the JRRM problem. A further evolution would benefit from considering in the resource assignment process, the diverse nature and characteristics of radio resources present in heterogeneous environments. In addition, only non-real time users are considered in the evaluation scenario, whereas the integration with real-time services characterised with more strict QoS demands introduces additional challenges in the resource management problem.

This work extends the current JRRM state of the art by proposing novel bankruptcy-based policies that take into account the radio resources' diversity, and that seek to efficiently distribute radio resources based on the system load and user/service QoS requirements. To this aim, the proposed JRRM techniques simultaneously determine the adequate combination of RAT and number of radio resources within such RAT that should be assigned to each active user in a multimedia environment. The use of bankruptcy theory in heterogeneous systems was first proposed in [7], where different RATs compete in coalitions for satisfying the bandwidth requirements of each user that requests access to the system ([7] allows the possibility of several RATs simultaneously serving a user). The bankruptcy problem modelled in [7] results in a cooperative game with transferable utility, where each RAT belonging to the coalition can offer any bandwidth so that the bandwidth offered by the coalition satisfies a user's demand. To solve the problem, [7] uses the Shapley value to determine how much bandwidth each RAT that forms the coalition must assign to a user. This approach cannot be applied in this paper since it is not envisaged the possibility that different RATs form a coalition to satisfy a single user's demand. In addition, while [7] focuses on individual user demands, this work investigates how to use the resources globally to efficiently satisfy the largest possible number of users in the system. In this context, solving the system bandwidth distribution problem using the Shapley value in a cooperative game would significantly increase the time needed to find a solution to the problem, and reduce the interest of the approach in practical systems. As a result, and differently from [7], this work proposes the use of bankruptcy distribution policies for managing, with a low computational cost, discrete radio resources in heterogeneous systems.

2 Joint radio resource management policies

This work focuses on the design of advanced JRRM policies that simultaneously determine the adequate combination of RAT and number of radio resources within such RAT that each active user would need to satisfy its service QoS requirements while maximizing the systems' efficiency and revenue. Determining the number of radio resources per RAT needed by each active user allows adequately dimensioning the number of users that can be assigned to each RAT while adequately satisfying their QoS demands. To achieve these objectives, this work proposes novel techniques that use the CEA (Constrained Equal Awards) bankruptcy rule [8] to distribute resources among active users.

2.1 Constrained equal awards policy

Bankruptcy theory addresses situations where a limited resource has to be divided among a number of agents with claims adding up to more than the value to divide. The bankruptcy dilemma can be formally

expressed as:

$$(c,E) \in R_+^N \times R_+ \text{ such that } \sum c_i \geq E \qquad (1)$$

where $E \in R_+$ is a continuous and infinitely divisible good that has to be divided among N agents, and $c \equiv (c_i)_{i \in N}$ is the requests vector, where $c_i \in R_+$ represents the amount of resources requested by each user $i \in N$. In this context, adequate rules must be defined to decide how scarce resources should be fairly distributed among the claiming creditors. One of the more widely used bankruptcy-based distribution rules is the CEA rule. The CEA rule tries to equally satisfy all agents by assigning them the same amount of resources without exceeding their individual demands [8]:

$$CEA_i(c,E) = \min\{c_i, \lambda\} \qquad (2)$$

where λ is chosen so that $\sum_i \min\{c_i, \lambda\} = E$.

Considering the similarities between the bankruptcy and RRM problems, the authors proposed in [9] the DCEAM (Discrete CEA for Mobile radio resources distribution) policy, which applies the CEA rule to the design of efficient RRM solutions in single RAT scenarios. The CEA rule was chosen due to the similarity of its fairness approach and the objective to homogeneously satisfy all multimedia users in a system as long as sufficient radio resources were available. To this aim, the CEA rule was adapted in order to equally distribute user satisfaction levels among users rather than equally distributing resources. This adaptation was done because different services do not need the same number of resources to obtain an equal QoS level. In addition, an important novelty of the DCEAM proposal was the evolution of the CEA rule to the distribution of resources of discrete nature (e.g. timeslots or codes). Another characteristic of DCEAM is that it assigned resources using utility functions defined to quantify the QoS level that a user might experience based on the number of assigned radio resources and the requested traffic service.

2.2 Traffic service utility functions

The JRRM techniques proposed in this work are also based on service-dependent utility values. Such values identify the QoS level that a user would perceive for each combination of potentially assigned RAT and number of radio resources within such RAT. This work considers discrete utility values to account for the discrete nature of radio resources. The use of continuous utility values would certainly simplify the resource assignment problem, but might result in inappropriate assignment decisions given the discrete nature of radio resources in mobile networks.

Traffic service utility functions have been defined for web, email, and real-time H.263 video traffic (with different mean video bit rates). Web traffic is modeled following an ON/OFF pattern that represents the transmission of objects within a web page, and inactive periods between two consecutive object transmissions [10]. The model also accounts for the time a user will take to read downloaded information before initiating a new web request. The implemented email traffic model [11] assumes that incoming messages are stored at a dedicated email server from where the user downloads the emails it is interested in. The model accounts for the email size (with and without attachment), and the time needed to read an email before downloading the next one. Real-time H.263 video traffic is modeled following [12], which defines the statistical properties of the I, P and PB frame types included in the H.263 standard. For each frame, the model describes its size and duration given the time by which the next video frame will be generated. If the transmission of a video frame is not finished when the next frame is generated, its transmission is aborted.

The utility functions try to express the perceived user QoS as the transmission data rate varies. To establish the utility functions, the minimum, mean, and maximum QoS levels demanded by users are first defined per service class as illustrated in Fig. 1. For email and web services, utility values are expressed in terms of the user throughput. The minimum, mean and maximum QoS levels for web users have then been defined as the throughput needed to satisfactorily transmit 90%, 95% and 97.5% of web pages in less than 4 seconds as established by the 3GPP TS 22.105 recommendations. These high percentiles have been selected due the high transmission reliability requirements of non-real time data services. As previously mentioned, the email traffic model considers the transmission of emails with and without attachments. In this case, it might be difficult to successfully transmit emails with large attachments within the 4 seconds 3GPP recommendations. As a result, the email QoS thresholds have been established based on the throughput required to satisfactorily transmit 65%, 75% and 80% of the emails (with or without attachments). Once the QoS thresholds have been established for web and email services, the utility functions have been defined so that users perceive a null utility value if their minimum QoS demand is not satisfied. This condition avoids assigning radio resources to users that would experience very poor QoS levels. Web and email user satisfaction linearly
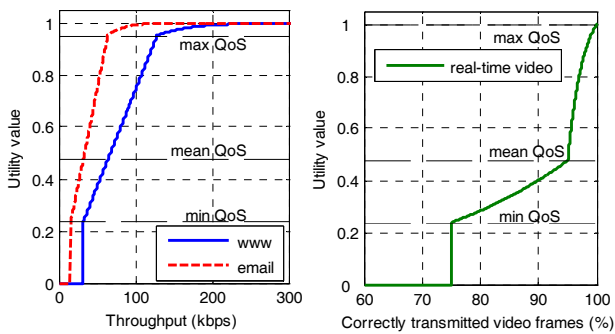
Fig. 1 Utility functions per traffic service

grows with the experienced throughput between the minimum and maximum QoS thresholds. Utility values equal to one have been avoided for web and email transmissions to account for the transmission reliability requirements of these services, and the dependence of the achievable throughput levels on the experienced channel quality conditions.

The real-time video QoS levels correspond to the percentage of correctly transmitted video frames (real-time video transmissions are considered satisfactory if video frames are completely transmitted before the next video frame is to be transmitted). As a result, real-time video utility functions are independent of the mean video bit rates. The studies reported in [13] show that a 25%, or even higher, dropping rate does not have a catastrophic effect on the QoS perceived by H.263 video users, and that dropping rates as high as 5% can be overcome if appropriate transmission techniques are invoked. Based on these observations, the minimum and mean real-time video QoS levels correspond to guaranteeing that 75% and 95% of video frames are transmitted before the next video frame needs to be transmitted. The maximum utility value for real-time video users has been set equal to one, and is achieved when all video frames are transmitted before the next video frame is to be transmitted. Although the 5% difference between the mean and maximum QoS levels might seem negligible, this 5% includes the H.263 I-frames. These frames include information of independently coded images in a video sequence, and are also used to code/decode other images exploiting temporal redundancy. As a result, I-frames have a significant impact on the user perceived QoS level, and require high transmission rates due to their potential large size. Real-time video users also perceive a null utility value below the minimum QoS threshold. Following the indications in [13] that establish that an

acceptable video quality requires a high percentage of correctly received video frames, the real-time video utility increases slowly with the percentage of transmitted frames until the mean QoS level is achieved, and then rapidly until the maximum QoS level.

Once the utility functions depicted in Fig. 1 have been defined, the relation between the utility values and the possible radio resource assignments per RAT must be established. To this aim, it is necessary to quantify the throughput that could be achieved per radio resource in each RAT. It is important to note that the emulated RATs implement link adaptation techniques that result in varying data rates as the channel quality varies. To account for these variations, and considering the difficulty to predict the instantaneous throughput in adaptive radio interfaces, the relation between the utility values and radio resource assignments has been established considering the data rate of the transmission modes (modulation and coding schemes) providing a balance between high data rates and high error correction capabilities. Although this selection might seem conservative, it will increase the probability that JRRM assignment decisions instantaneously satisfy the user QoS requirements as the channel quality varies compared to more aggressive transmission mode selection strategies. This work considers a heterogeneous framework where the GPRS (General Packet Radio Service), EDGE (Enhanced Data rates for GSM Evolution), and HSDPA (High-Speed Downlink Packet Access) radio interfaces physically coexist. This evaluation scenario was chosen since it provides a mix of standards and multiple access technologies with varying performance capabilities (the proposed techniques could certainly be adapted to other scenarios and technologies). In this context, average throughput values of 13.4 kbps and 22.4 kbps per timeslot (TS) have been selected for GPRS and EDGE, respectively. These values correspond to the data rates of the coding scheme 2 (CS2) in GPRS, and the modulation and coding scheme 5 (MCS5) in EDGE. HSDPA offers a high number of transmission modes depending on the number of assigned codes. This work considers the transmission modes related to the 30 CQI (Channel Quality Indicator) values for User Equipment category 10 [14]. To achieve the sought balance between high data rate and high error correction capabilities, the selected transmission rate per number of assigned HSDPA codes is that achieved by the 'intermediate' mode out of all possible modes for a given number of codes.

Once an average data rate is identified for each possible combination of RAT and number of radio resources, the relation between utility values and RAT/resources combinations can be directly established for web and email users using the utility functions depicted in Fig. 1. For H.263 real-time video services, an additional step is necessary. Following the H.263 video model described in [12], the Cumulative Distribution Function (CDF) of the throughput needed to transmit each video frame before the next video frame is to be transmitted is derived for the mean video bit rates considered in this work (64, 128, and 256 kbps). Using these CDFs, the percentage of video frames reported in Fig. 1 can be related to the required throughput levels. Consequently, the relation between discrete utility values and RAT/resources combinations can also be derived for real-time video users. Table 1 shows an example of the utility values achieved per RAT/resources combinations for the web service (shown in increasing throughput order). In GPRS and EDGE, a user can receive up to 8 timeslots of the same carrier, and up to 15 codes in HSDPA when considering UE category 10. The RAT/resources combinations are denoted as xY, corresponding to x radio resources (timeslots for GPRS and EDGE, and codes for HSDPA) from RAT Y (GPRS is represented as G, EDGE as E, and HSDPA as H).

2.3 Bankruptcy-based JRRM proposals

This work considers a heterogeneous system where a set US of active users compete for the radio resources from different RATs that better satisfy their QoS demands. The set of all possible RAT/resources combinations is denoted as:

$$C = \left\{ c = xY \mid x = 1, 2, ..., r_Y \text{ and } Y = G, E, H \right\} \cup 0RS \quad (3)$$

Res./ RAT	Throughput (kbps)	Utility value	Res./ RAT	Throughput (kbps)	Utility value
1G	13.4	0.00	1H	116.5	0.87
1E	22.4	0.00	6E	134.4	0.96
2G	26.8	0.00	7E	156.8	0.98
3G	40.2	0.29	8E	179.2	0.99
2E	44.8	0.33	2H	396	0.99
4G	53.6	0.40	3H	741	0.99
5G	67	0.49	4H	1139.5	0.99
3E	67.2	0.50	5H	2332	0.99
6G	80.4	0.59	7H	4859.5	0.99
4E	89.6	0.67	8H	5709	0.99
7G	93.8	0.69	10H	7205.5	0.99
8G	107.2	0.79	12H	8618.5	0.99
5E	112	0.83	15H	11685	0.99

Table 1 Web utility values

In (3), r_Y represents the maximum number of radio resources available at each RAT Y. C also includes the '0RS' combination corresponding to users that do not receive any resources. Furthermore, the function $d: US \rightarrow C$ relates each user $i \in US$ with an element of C; $d_i = d(i)$ denotes the RAT/radio resources combination assigned to user i. In addition, each element of C has a corresponding data rate by means of the function $tr: C \rightarrow TR$, where TR is the set including all possible data rates achievable with the possible RAT/resources combinations. For each service class present in the system, a set of the utility values that can be achieved with the available RAT/resources combinations has also been defined: U^{email}, U^{www} and U^{hv}, where U^{hv} is the set of utility values related with an H.263 real-time video service with a mean bit rate of v kbps. These utility values are related with the elements of C by means of the following functions: $u^{email}: C \rightarrow U^{email}$, $u^{www}: C \rightarrow U^{www}$, and $u^{hv}: C \rightarrow U^{hv}$. Using these functions, the utility perceived by a user $i \in US$ with a RAT/resource assignment d_i can be denoted as $u_i(d_i)$:

$$u_i(d_i) = \begin{cases} u^{email}(d_i) & \text{if user } i \text{ demands an email service} \\ u^{www}(d_i) & \text{if user } i \text{ demands a www service} \\ u^{hv}(d_i) & \text{if user } i \text{ demands a video service} \\ & \text{with } v \text{ kbps mean bit rate} \end{cases} \quad (4)$$

To equally satisfy all users and maximize the satisfaction level experienced by all active users in the system, this work proposes JoDCEA (Joint Radio Resource Management based on Discrete CEA), a JRRM policy based on the bankruptcy CEA distribution rule. JoDCEA has been designed to simultaneously determine the RAT and number of resources within this RAT that active users would require to satisfy their QoS demands. The JoDCEA policy determines resources taking into account the user traffic service, the current system load, and the traffic distribution conditions. JoDCEA considers the set of RAT/resource combinations included in C as the discrete resources that have to be distributed among the active users in the system. As a result, each RAT/resource combination $c \in C$ is considered as an indivisible resource, and they are sequentially assigned to users one by one. To achieve its objective, JoDCEA assigns a RAT/resource combination to the user j that experiences the lowest utility value at each moment. The RAT/resource combination to be assigned to that user corresponds to that sequentially increasing its utility value. Considering the example in Table 1, if the user currently experiencing the lowest utility value had been previously assigned 2 GPRS timeslots (2G),

JoDCEA will assign that user 3 GPRS timeslots (3G). If this new assignment results in that this user is still experiencing the lowest utility value among active users, JoDCEA will then assign that user 2 EDGE timeslots (2E), and so on. Then, the user with the current lowest utility value will be the next one to be assigned radio resources. In case different users have the same utility value, radio resources are assigned to the highest priority user based on predefined service priorities.

JoDCEA defines CD_j as a new set of RAT/resource combinations providing user j, the active user with the current lowest utility value, with higher utility values than the value it currently perceives with its assigned radio resources (d_j^{cur}):

$$CD_j = \left\{ c = xY \mid c \in C, u_j(c) > u_j(d_j^{cur}) \right\} \qquad (5)$$

After establishing CD_j, the RAT/resource combination that the user should be assigned (denoted as $c*$) must satisfy:

$$u_j(c*) = \min_{c \in CD_j} u_j(c) \qquad (6)$$

If several RAT/resource combinations satisfy this condition, the combination providing a lower data rate will be chosen. If ra_Y represents the number of radio resources available at RAT Y at a given moment, the set of RAT/resource combinations available in the system at each moment can be expressed as:

$$CA = \left\{ c = xY \mid x = 1, ..., ra_Y, Y = G, E, H \right\} \cup 0RS \qquad (7)$$

It is possible that the RAT/resource combination $c*$ that should be assigned to a given user following (6) is not available in the system at that moment, i.e. $c* \notin CA$. In this case, different variants of the JoDCEA policy have been defined. Before describing such variants, it is important to note that the JoDCEA technique is executed each time a new user requests access to the system or a user ends its transmission and frees its radio resources. In this case, all users competing for radio resources, except active real-time video users, begin each distribution round without any assigned radio resources. To account for the strict QoS requirements of real-time video services, each active video user (user k) is guaranteed at each distribution process to maintain at least the x_m radio resources of its previous assignment ($x_p Y_p$) that are needed to satisfy their minimum QoS demand ($u_{QoS_{min}}$)[1]. Active video users will then compete with other users for

---

[1] Vertical handovers can increase the user-perceived throughput, but also incur in an additional delay and overhead that must be carefully controlled, in particular for delay sensitive real-time services.

assignments improving their QoS demands (including the possibility of changing RATs):

$$u_k(c*) \geq u_k(x_m Y_p) \quad \text{with} \quad x_m = \min_{x \leq x_p} x \mid u_k(xY_p) \geq u_{QoS_{min}} \qquad (8)$$

The first JoDCEA variant (JoDCEAv1) determines for the user with the lowest utility value (user j) the RAT/resource combination currently available in the system ($c* \in CA$) that provides the user with the lowest utility value increase:

$$u_j(c*) = \min_{c \in CD_j \cap CA} u_j(c) \qquad (9)$$

Following (9), the first variant could provide a user with a radio resource assignment that over-satisfies its QoS requirements if the RAT/resource combination that the user needs to satisfy its QoS demand is not currently available in the system. This fact may also entail that these radio resources would not be available for users that actually require them to satisfy their QoS demands. To overcome this situation, JoDCEAv1 checks before assigning the $c*$ combination to user j, whether such combination might also improve the utility value perceived by a higher priority user that has already been assigned radio resources (user k)[2], i.e., $u_k(c*) > u_k(d_k^{cur})$. If this condition is satisfied, JoDCEAv1 checks whether user j also improves its perceived utility value if assigned the RAT/resources combination currently hold by user k ($u_j(d_k^{cur}) > u_j(d_j^{cur})$). If this is the case, JoDCEAv1 assigns $c*$ to the higher priority user (user k), and assigns to user j the minimum number (x_m) of radio resources freed by user k ($d_k^{cur} > x_k Y_k$) that satisfies its QoS requirements:

$$d_k = c* \quad \text{and} \quad d_j = x_m Y_k \quad \text{with} \quad x_m = \min_{x \leq x_k} x \mid u_j(xY_k) \geq u_j(d_j^{cur}) \quad (10)$$

JoDCEAv1 checks again whether the $x_m Y_k$ assignment could also improve the utility value perceived by another user with higher priority than user j. If this is the case, radio resources are exchanged following (10). In this context, it should be noted that since JoDCEAv1 assigns discrete resources from the set of currently available RAT/resources combinations, users can still receive resources that provide them with a higher QoS than initially demanded. To avoid a possible breach of the CEA principle, a second JoDCEA variant (JoDCEAv2) is proposed. JoDCEAv2 strictly determines for each user the RAT/resources combination sequentially increasing its utility value following (6), even if such combination might not be available at that moment ($c* \notin CA$). To account for the limited number of radio resources, and provide an end

---

[2] Users are checked from highest to lowest priority.

to the distribution process, JoDCEAv2 performs its first distribution process only limiting the radio resources from the RAT M that provides the highest data rates ($tr(c = xM) = \max_{c \subset C} tr(c)$). Users that receive radio resources from RAT M maintain them and do not participate in subsequent distribution rounds. The others users will compete again for the radio resources of the other RATs in a new and iterative distribution process. As it was previously the case, the process ends when the resources from the remaining RAT with the highest data rates are all assigned.

3 JRRM performance

The performance of the JoDCEA techniques is analyzed using a multi-RAT and multimedia C++ wireless platform. The implemented simulator is not aimed at accurately modeling radio transmissions, but at measuring the efficiency of the resource distribution proposals, and optimizing them according to each RAT's operational characteristics and specific system constraints. The platform models a heterogeneous wireless system where GPRS, EDGE, and HSDPA coexist, and emulates the distinct and discrete nature of radio resources for each of the simulated RATs, together with their main QoS characteristics. The simulated scenario considers that all three RATs provide the same radio coverage. The JRRM techniques have been implemented following a centralized architecture approach discussed in the 3GPP standards ([15] and [16]) where a common JRRM entity collects information of all available RATs[3]. Since the JoDCEA variants use utility functions to estimate the users' QoS demands, only updated information about each RAT's load must be transmitted to the JRRM entity. The 3GPP standards define the procedure to exchange cell traffic load measurements between different RATs ([17] and [18]), and thereby the proposed JRRM techniques could be easily accommodated in future networks using the current standards.

In this context, a multimedia scenario with email (lowest priority service), web and real-time H.263 video (highest priority) transmissions are simulated, with each service representing 50%, 30% and 20% of

the new service requests respectively in the S1 scenario, and 35%, 35% and 30% in the S2 scenario. Within the real-time H.263 service, new service requests are distributed as follows: 50% of 64kbps video users, 30% of 128kbps video users, and 20% of 256kbps video users. A cell with equal radio coverage for all RATs is modelled, and cell loads of 15, 20, and 25 simultaneous active users are simulated with two frequency carriers each (i.e. 16 timeslots) for GPRS and EDGE, and 14 HSDPA codes. The selected configurations are aimed at analysing the operation and performance of JRRM algorithms under varying user loads where the shortage of radio resources and user QoS demands prevent from guaranteeing maximum QoS levels to all users.

3.1 Reference JRRM techniques

The bankruptcy-based proposals are compared to some well established JRRM techniques reported in the literature:

- Service based RAT selection, SeRS [19]. This technique is based on pre-established service-to-RAT assignments. For each service, a prioritized list of RATs is maintained. When a new user requests access to the system, the system tries to allocate the user to the first RAT from its list with available capacity. SeRS has been implemented in this work considering the following prioritized list of RATs per service type: HSDPA-EDGE-GPRS for real-time users, EDGE-GPRS-HSDPA for web users, and GPRS-EDGE-HSDPA for email users.

- Load balancing based RAT selection, LBRS [20]. The LBRS mechanism assigns each user requesting access to the system to the RAT currently experiencing the lowest load. The load metric is computed as the ratio of utilized capacity to the total available capacity in each RAT. The utilized/available capacity is measured in terms of the number of timeslots in GPRS and EDGE, and in terms of the number of codes in HSDPA.

- Satisfaction based RAT selection, SaRS [21]. Each time a new user requests access to the system, this technique evaluates the number of satisfied users in each RAT, and assigns the new user to the RAT with a higher percentage of satisfied users. To this aim, a user is considered satisfied if its minimum QoS demand is fulfilled.

JoDCEA's performance and efficiency is also compared against that achieved with the MAXILOU (MAXImise Lowest Utility) policy [22]. This policy uses linear programming and optimization techniques

---

[3] In a centralized architecture, the JRRM entity might be placed in a stand-alone node or in an existing node of the radio access network of any RAT. The implemented JRRM techniques could also operate under the 3GPP JRRM distributed architecture given the limited amount of information they require to be exchanged among different RATs (number of active users and their requested services).

to provide similar, and highest possible, satisfaction levels to all users. When this target is not possible due to a shortage of resources, service priorities are applied. MAXILOU also makes use of the described utility functions to represent the user QoS that can be achieved with a given RAT/resource combination. MAXILOU seeks to maximize the lowest utility value assigned to any user in a distribution round. To apply linear programming tools, the MAXILOU objective function is expressed as:

$$\max z, \text{ with } z \leq u_j(d_j), \forall j \in US \tag{11}$$

where z is a real variable equal to the smallest utility value assigned to a user. The utility can only take specific values within a finite set due to the discrete nature of radio resources, and each user can only receive one RAT/resource combination d_j. As a result, $u_j(d_j)$ is expressed as:

$$u_j(d_j) = \sum_{c \in C} u_j(c) \cdot y_j^c \tag{12}$$

where y_j^c is a binary variable equal to one if user j is assigned the c RAT/resource assignment, and equal to 0 if not. In this context, MAXILOU must decide for each user which y_j^c variable is equal to one. The MAXILOU objective function is also subject to some system and service constraints that must be expressed as linear functions. MAXILOU also applies service priorities under resource shortage conditions, and seeks to guarantee the minimum QoS level for active video users. The MAXILOU distribution dilemma corresponds to a Mixed Integer Linear Programming (MIP) problem, which has been solved using the Branch and Cut method, and the Simplex Linear Programming (LP) mechanism. The interested reader is referred to [22] for additional details on the MAXILOU technique and the employed linear programming tools.

3.2 Performance comparison

Fig. 2 and Fig. 3 represent the percentage of users per service class that achieved the utility values corresponding to the minimum, mean and maximum QoS levels in each distribution round. The depicted results highlight that JoDCEAv1 seeks maximising the percentage of users obtaining the same and highest possible QoS level. As a result, JoDCEAv1 tries to first maximise the percentage of users obtaining their minimum and mean QoS demands before satisfying maximum QoS requests. Under cell loads of 15 and 20 users, JoDCEAv1 is capable to provide all users with their minimum and mean QoS demand. When the user load increases, there is a shortage of radio resources and the mean QoS level cannot be guaranteed to all users. In this context, JoDCEAv1 reduces the percentage of users achieving their mean QoS demand homogeneously for all service types, except for web. This is due to the resources requested by web users, and the order in which resources are being assigned. Given the established service priorities, web users receive resources after video users have been served. In this context, the results obtained showed that the radio resources that are better adjusted to the web minimum and mean QoS levels could be assigned to higher priority users. Consequently, web users can receive discrete resources over-satisfying their QoS demand, which explains why a higher percentage of web users obtained their mean QoS demand with 25 users per cell. This effect is also at the origin of why a higher percentage of email users than web users obtain their maximum QoS demand. Despite some users achieving higher QoS levels than users with higher priority, the information reported in Fig. 4 and Table 2 shows that this is not due to an inefficient assignment of resources; in fact, JoDCEAv1 never assigns to web users radio resources with higher data rates than to video users. The observed behaviour is instead due to the fact that JoDCEAv1 only assigns resources available in the system ($c* \in CA$), and that the assignment decision is based on the QoS level currently experienced by users and not on the QoS level that users will obtain with their new resource assignments. It is important to note that distributing resources based on the QoS levels that users will achieve with the assigned resources requires the design of JRRM techniques searching for optimum solutions (e.g. MAXILOU). As it will later be shown, such techniques usually result in higher computational costs. In addition, Fig. 2 and Fig. 3 show that the percentage of users reaching their maximum QoS demand is mainly reduced when the load increases for users with lower priority. This behaviour shows that JoDCEAv1 complies with the established service priorities when the available radio resources are not enough to homogeneously satisfy all users and service types.
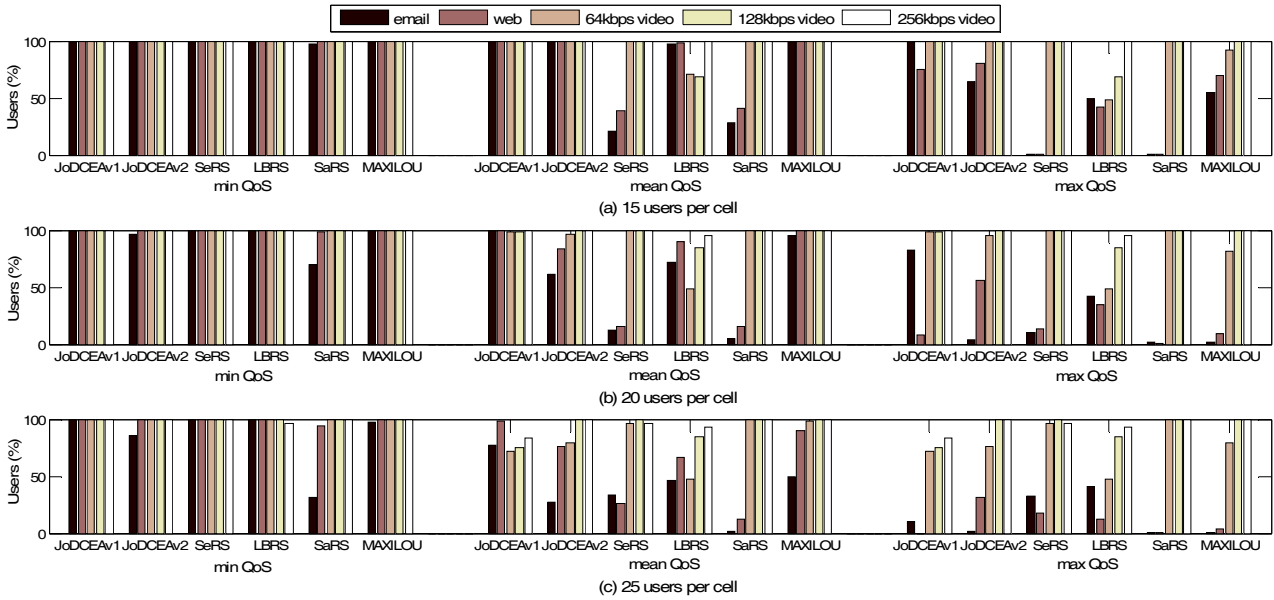
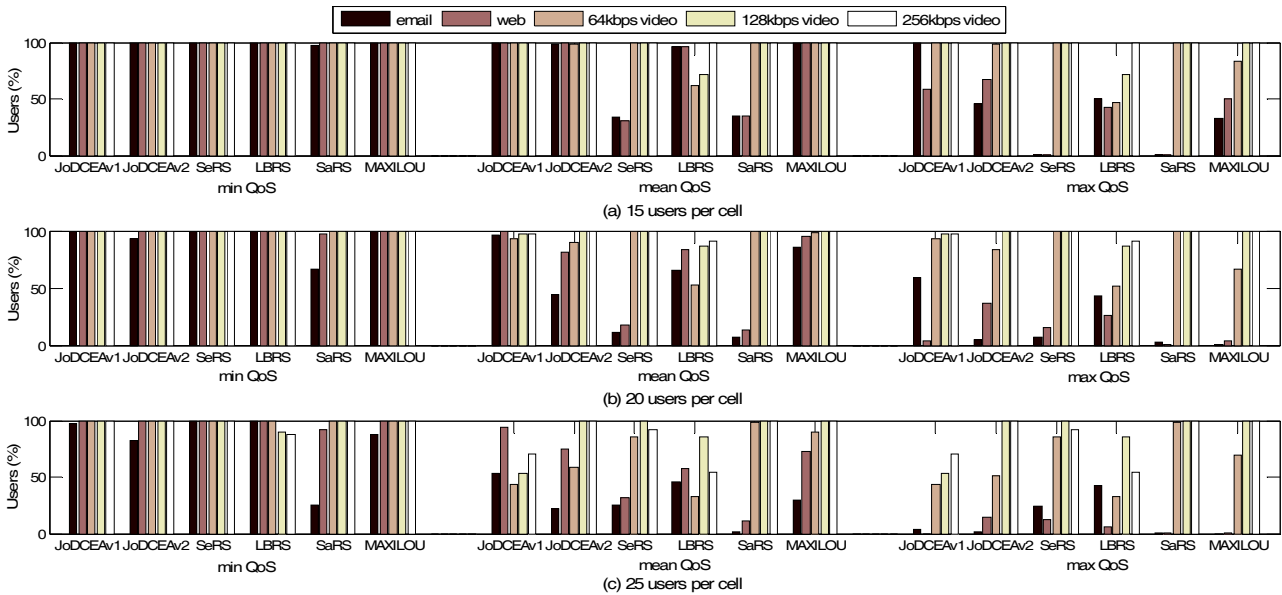**Fig. 2** QoS performance under S1 scenario

Fig. 3 QoS performance under S2 scenario

As depicted in Fig. 2 and Fig. 3, JoDCEAv2 exhibits a different behaviour and performance when there is a high shortage of radio resources to satisfy all users. In particular, JoDCEAv2 increases the QoS level perceived by the most demanding and priorized services at the expense of the lowest priority ones. With 25 users per cell, JoDCEAv2 increases the percentage of video users that satisfy their mean and maximum QoS demand compared to JoDCEAv1, but reduces the percentage of web and emails users that satisfy their minimum and mean QoS demand. The results depicted in Fig. 4 show that JoDCEAv2's performance is obtained by adjusting the resource

assignments to the users' QoS demands. For example, while JoDCEAv1 assigned HSDPA resources to email users in 9% of the distribution rounds, JoDCEAv2 never assigns HSDPA resources to email users under high cell loads. These results confirm that different QoS strategies can be designed by adapting the original bankruptcy-based JoDCEA policy.

The obtained results also show that both JoDCEA variants outperform the three reference techniques in all the simulated scenarios. Only in the scenarios with 20 and 25 users per cell, LBRS achieves higher QoS levels for the lowest priority users, but this is done at the expense of not satisfying the mean QoS demand
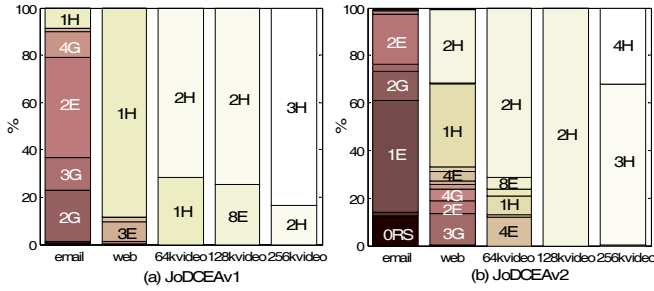
Fig. 4 Assigned resources per service class (%) - S1 and 25 users per cell

	email	web	64k video	128k video	256k video
Min QoS level	1E	3G	4E	8E	2H
Mean QoS level	3G	5G	7E	2H	3H
Max QoS level	5G	6E	8E	2H	4H

Table 2 Assignments needed to satisfy the services QoS levels

for a large percentage of real-time video users. In addition, the JoDCEA variants outperform SeRS and SaRS in low load scenarios, where a more efficient resource management results in higher percentages of users achieving mean and maximum QoS levels. Under higher loads, SeRS and SaRS provides the maximum QoS level to a higher percentage of video users compared to the JoDCEA variants. However, this is done at the expense of not even satisfying the mean QoS demand, and even the minimum QoS demand in the case of SaRS to email and web users. These results emphasize that the JoDCEA variants improve user fairness thanks to their resource allocations, and hence can provide higher and more homogeneous QoS levels for all service types compared to the simulated reference techniques.

Finally, it is important to highlight the similar distribution and performance trend obtained with JoDCEAv2 and MAXILOU. MAXILOU was designed to achieve the highest possible QoS level for the user with the lowest utility value. With this approach, MAXILOU aims to obtain homogeneous QoS levels for all users. When the load increases, the possibility of achieving homogeneous QoS levels is reduced, and the effect of the service priorities is most notable. Despite its similarities, several differences can be observed between JoDCEAv2 and MAXILOU.

MAXILOU satisfies the mean QoS demands to a higher percentage of users compared to JoDCEAv2, but reduces the percentage of lower priority users that achieve their maximum QoS level. It is important to highlight that while both techniques satisfy the mean QoS demands of all users under the lowest load scenarios, JoDCEAv2 increases the percentage of users that see their maximum QoS demand satisfied. This is due to the fact that when MAXILOU satisfies its objective function (11), users stop competing for additional radio resources that could further improve their QoS. As an example, while MAXILOU did not assign all available radio resources in 30.8% of the distribution rounds (S1 scenario with 15 active users per cell), this percentage is reduced to only 0.8% with JoDCEAv2.

3.3 Implementation cost

An important characteristic of heterogeneous wireless systems is the possibility to conduct vertical handovers between RATs. Although such handovers can increase the final performance, they also incur in an additional delay and overhead that must be carefully controlled, in particular for delay sensitive real-time services. To this aim, the JoDCEA variants guarantee that in each distribution round, active real-time users will maintain at least their minimum QoS level using resources from the RAT they were previously assigned (8). These users will only change RATs if they can obtain higher QoS levels using resources available from other RATs. This approach has been adopted to achieve a balance between QoS and cost of switching RATs. For non real-time services, vertical handovers are permitted without any restrictions. Fig. 5 shows the percentage of transmissions that didn't perform a vertical handover in the S1 scenario. This figure confirms that the two JoDCEA variants limit the number of vertical handovers for real-time services, and that such handovers mainly occur when applying JoDCEAv2 under high cell loads. In this case, the vertical handovers are justified by the QoS improvements obtained by real-time video users when applying JoDCEAv2 (Fig. 2).
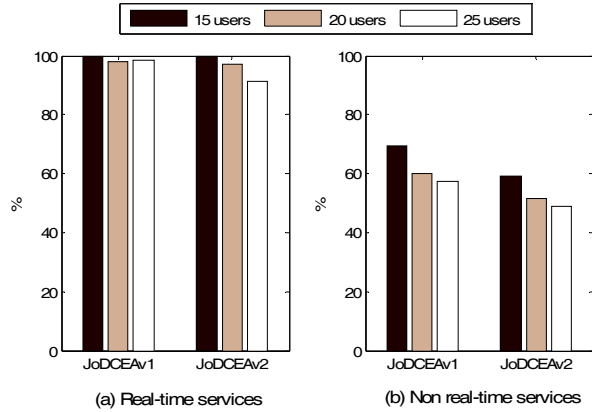
Fig. 5 Percentage of transmissions that ended up without switching RATs

Fig. 6 Average CPU time (milliseconds) needed to distribute radio resources

JRRM decisions are based on an increasing number of variables and data. As a result, the JRRM processing time might become an important factor that can compromise the implementation feasibility of advanced JRRM policies. The bankruptcy-based JRRM techniques proposed in this work are characterised by a polynomial computational complexity $O(n^2)$, where n is the number of users participating in a radio resource distribution. This order of growth has been estimated considering the worst case scenario where enough radio resources are available to provide all users with their maximum QoS demand. In this case, the two JoDCEA variants do not stop a JRRM distribution process until all users receive radio resources satisfying their maximum QoS level. When a limited number of radio resources is available, the computational complexity of the two JoDCEA variants is considerably reduced. Fig. 6 represents the average time that each JoDCEA variant needs to solve a JRRM process in the S1 scenario; the tests were conducted using a 2.6 GHz AMD Opteron processor with 1 MB of cache and 3 GB of RAM. The figure shows that the time required to solve a JRRM process is very low when the number of available resources is limited. In this case, the time needed to find a JRRM solution does not only depend on the number of active users, but also on the number of users that could be served with the available radio resources, and the QoS levels that could be achieved with such resources. This explains why the time JoDCEAv1 needed to distribute radio resources was higher with 20 users per cell than with 25 users per cell (Fig. 6). The results depicted in Fig. 6 show that the time needed to distribute radio resources increases with the user load for JoDCEAv2. This is due to the fact that this technique only limits during each distribution process the resources
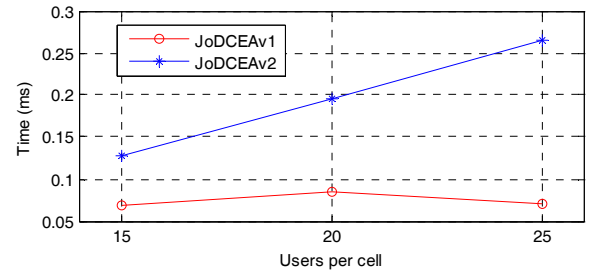
available from the RAT providing the highest data rates. As a result, the time needed to distribute resources is higher than for JoDCEAv1, and strongly dependent with the number of active users. However, it is important to note that such time can stop increasing when the number of users requesting resources is higher than the number of users that can be served with the available radio resources. In this case, it is possible to identify and eliminate from the distribution process those users that would not be served considering the available resources, the users' distribution per service class, and their minimum QoS demand. For example, it could be feasible to only consider per service class the maximum number of users that could be served if there were no active users from the other service classes. With this simple approach, it is possible to limit the increase in the time needed to distribute resources with the cell load in the case of JoDCEAv2 without modifying the final result.

Despite the higher computational cost of the second JoDCEA variant, the computational cost of both variants is low. In addition, the obtained results have shown that there is only a small difference between both variants when analysing the number of executed vertical handovers. In this context, both JoDCEA variants are considered to be equally suitable to be implemented in heterogeneous wireless networks. Since both JoDCEA variants provide different QoS strategies, the decision on the most adequate variant to implement should be based on the operator's QoS preferences and objectives.

The previous section showed that the bankruptcy-based JRRM policies can achieve a performance close to that obtained with JRRM techniques that seek an optimum resources distribution through the use of linear programming and optimization methods. The implementation feasibility of the MAXILOU technique was evaluated in [22], where it was highlighted that reasonable average computational costs could be achieved. However, the time needed to find an optimum JRRM solution with MAXILOU

considerably increased with the cell load when analysing the 95-percentile. In this case, suboptimum solutions, higher performance platforms or more advanced linear programming tools would be needed to reduce the MAXILOU computational cost to levels demanded by real-time systems. In this context, and given the similar performance trends exhibited by MAXILOU and JoDCEAv2, it is worthwhile comparing their computational cost. The average CPU time required by MAXILOU to solve a JRRM resource assignment problem was equal to 700 ms and 1239 ms when 20 and 25 active users demanded resources in the S1 scenario. Such values were reduced to 0.20 ms and 0.27 ms respectively when applying JoDCEAv2 (a reduction of three/four orders of magnitude). These results highlight the potential of bankruptcy-based JRRM techniques to achieve performance results close to those exhibited by techniques seeking to optimally distribute resources, but with a significantly lower computational cost. In this context, bankruptcy-based JRRM techniques provide a suitable trade-off between performance and computational cost in heterogeneous wireless systems.

4   Conclusion

Heterogeneous wireless networks require the design and implementation of advanced JRRM policies to efficiently manage the radio resources from the different RATs that could physically co-exist. In this context, this work has presented and evaluated novel bankruptcy-based JRRM techniques designed to assign incoming calls the most appropriate RAT and number of radio resources within such RAT needed to guarantee the user/service QoS demands. The proposed techniques have been defined considering the discrete and distinct nature of available radio resources in heterogeneous networks. Following an initial bankruptcy-based distribution policy, two techniques with varying QoS strategies and objectives have been introduced, and their performance has been compared with other reference JRRM techniques. The obtained results highlight the potential of bankruptcy theory to address the challenges of radio resource management in heterogeneous systems. In addition, the conducted study has shown the capacity of the proposed techniques to profit from the service QoS differentiation present in multimedia environments, and to adapt the resource assignments to the specific system conditions. Finally, the study has also shown that the proposed bankruptcy-based solutions can rapidly solve a JRRM process to distribute radio resources among active users, thereby improving its

implementation perspectives in real systems.

References

[1] Sallent, O. (2006). A perspective on Radio Resource Management in B3G. In *Proceedings of the 3rd International Symposium on Wireless Communication Systems (ISWCS)*, pp. 30-34, Valencia (Spain).

[2] Tolli, A., Hakalin, P., Holma, H. (2002). Performance evaluation of common radio resource management (CRRM). In *Proceedings of IEEE International Conference on Communications (ICC)*, vol. 5, pp. 3429–3433, New York (USA).

[3] Lincke, S. J. (2005) Vertical handover policies for common radio resource management. *International Journal of Communication Systems*, 18, 527–543.

[4] Furuskar, A., Zander, J., Multiservice allocation for multi-access wireless systems. *IEEE Transactions on Wireless Communications*, 4(1), 174–184.

[5] Giupponi, L., Agusti, R., Perez-Romero, J., Sallent, O. (2008). A novel approach for joint radio resource management based on fuzzy neural methodology. *IEEE Transactions on Vehicular Technology*, 57(3), 1789-1805.

[6] Calabuig, D., Monserrat, J.F., Martín-Sacristán, D., Cardona, N. (2010). Joint dynamic resource allocation for QoS provisioning in multi-access and multi-service Wireless Systems. *Mobile Networks and Applications,* 15(5), 627-638.

[7] Niyato, D., Hossain, E. (2006). A cooperative game framework for bandwidth allocation in 4G heterogeneous wireless networks. In *Proceedings of the IEEE International Conference on Communications (ICC)*, pp. 4357-4362, Istanbul (Turkey).

[8] Thomson, W. (2003). Axiomatic and game-theoretic analysis of bankruptcy and taxation problems: a survey. *Mathematical Social Sciences*, 45, 249-297.

[9] Lucas-Estañ, M. C., Gozalvez, J., Sanchez-Soriano, J. (2007). Multi-channel radio resource distribution policies in heterogeneous traffic scenarios. In *Proceedings of the 66th IEEE Vehicular Technology Conference (VTC2007-Fall)*, pp. 1674-1678, Baltimore (USA).

[10] Barford, P., Crovella, M. (1998). Generating representative web workloads for network and server performance evaluation. In *Proceedings. International*

Conference Measurement and Modeling of Computer Systems, pp. 151–160.

[11] Ho, J., Zhu, Y., Madhavapeddy, S. (1999). Throughput and buffer analysis for GSM general packet radio service (GPRS). In *Proceedings of the IEEE Wireless Communications and Networking Conference (WCNC)*, pp. 1427-1431.

[12] Lazaro, O., Girma, D., Dunlop, J. (2005). H.263 video traffic modelling for low bit rate wireless communications. In *Proceedings of the IEEE Personal, Indoor and Mobile Radio Communications Symposium (PIMRC)*, vol. 3, pp. 2124-2128, Barcelona (Spain).

[13] Hanzo, L., Cherriman, P., Streit, J., *Wireless video communications: second to third generation systems and beyond*, IEEE Press, 2001.

[14] 3GPP, Technical Specification Group Radio Access Network; Physical layer procedures (FDD), 3GPP TS 25.214, version 7.3.0, 2006.

[15] 3GPP, Improvement of Radio Resource Management (RRM) across RNS and RNS/BSS, 3GPP TR 25.881, version 5.0.0, 2002.

[16] 3GPP, Improvement of Radio Resource Management (RRM) across RNS and RNS/BSS post-Rel-5, 3GPP TR 25.891, version 0.3.0, 2003.

[17] 3GPP, Mobile Switching Centre - Base Station System (MSC-BSS) interface; Layer 3 specification, 3GPP TS 48.008, version 9.4.0, 2010.

[18] 3GPP, UTRAN Iu interface Radio Access Network Application Part (RANAP) signalling, 3GPP TS 25.413, version 9.5.0, 2010.

[19] Perez-Romero, J., Sallent, O., Agusti, R. (2005). Policy-based initial RAT selection algorithms in heterogeneous networks. In *Proceedings of 7th IFIP International Conference on Mobile and Wireless Communication Networks (MWCN')*, Marrakech (Marocco).

[20] Gelabert, X., Perez-Romero, J., Sallent, O., Agusti, R. (2005) On the suitability of load balancing principles in heterogeneous wireless access networks. In *Proceedings of the International. Symposium on Wireless Personal Multimedia Communications (WPMC'05)*, Aalborg (Denmark).

[21] da Silva, A.P., Cavalcanti, F.R.P., de O. Neto, R.A. (2007). VoIP capacity analysis of wireless multi-access networks using access selection schemes. In *Proceedings of the IEEE Personal, Indoor and Mobile Radio Communications (PIMRC)*, pp. 1-5, Athens (Greece).

[22] Lucas-Estañ, M. C., Gozalvez, J. (2010). Real-Time computational performance of advanced JRRM policies in B3G heterogeneous wireless systems. In *Proceedings of the 16th European Wireless Conference*, pp. 564-571, Lucca (Italy).