Ann. Telecommun.
DOI 10.1007/s12243-012-0288-9

An AHP-based resource management scheme for CRRM in heterogeneous wireless networks

Choongyong Shin · Jinsung Cho · Jeong Geun Kim · Ben Lee

Received: 29 April 2011 / Accepted: 20 January 2012

Abstract In a heterogeneous wireless environment, a variety of Radio Access Technologies (RATs) coexist. Since the number of RATs is anticipated to increase in the near future, it is desirable to have radio and network resources managed in a cooperative manner using the Common Radio Resource Management (CRRM) strategy. In order to make RAT-specific radio resources manageable in CRRM, this paper proposes the Analytical Hierarchy Process (AHP) based resource management scheme that efficiently allocates resources among heterogeneous wireless networks. The proposed AHP-based method is simple and flexible enough to be used in any network environment and can consider a multitude of decision factors. In addition, the proposed scheme uses a radio bandwidth model, which properly reflects transmission rates under given channel conditions, as the actual radio resources to be allocated. The model considers the AMC (Adaptive Modulation and Coding) scheme that is widely used in current broadband wireless access technologies, and thus, packet service characteristics, such as response time, can be analyzed. This is in contrast to existing work that focuses only on circuit service characteristics (e.g., blocking probability). The effectiveness and flexibility of the proposed method are demonstrated by implementing a number of existing methods and performing extensive simulation study on several different scenarios.

C. Shin · J. Cho (✉) · J. G. Kim
Kyung Hee University, Yongin-si 446-701, Korea
e-mail: chojs@khu.ac.kr

B. Lee
Oregon State University, Corvallis 97331, OR, USA

Keywords Heterogeneous network · Common radio resource management (CRRM) · Radio bandwidth model · AHP

1 Introduction

The next generation wireless communication services are envisioned to be supported by heterogeneous networks consisting of variety of wireless access technologies. In this environment, users with multiple interfaces can access various wireless networks according to their preferences, but *Radio Resource Management* (RRM) strategies are implemented separately for different networks. Therefore, heterogeneous networks require a more efficient RRM strategy to optimally coordinate radio resources among different Radio Access Technologies (RATs). The Common RRM (CRRM) strategy has been proposed for this purpose. If radio and network resources in heterogeneous network environment are managed in cooperative manner with CRRM strategy, users can receive better services and network operators can manage their resources more efficiently.

The CRRM concept is based on a two-tier RRM model shown in Fig. 1 [1]. The local RRM entity that manages and allocates RAT-specific radio resources is located in the lower tier. The CRRM entity resides in the upper tier of the model and is responsible for efficiently managing multiple RRM entities. A CRRM entity controls a number of RRM entities and communicates with other CRRM entities. Based on the degree of interaction between RRM and CRRM entities, the following functions can be performed by either RRM

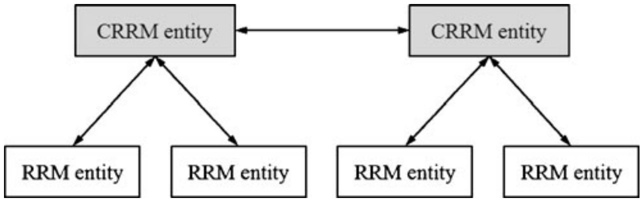

Fig. 1 CRRM interaction model

or CRRM: RAT selection, vertical handover, admission control, congestion control, horizontal handover, packet scheduling, and power control [1, 2]. A higher degree of interaction between RRM and CRRM entities can achieve more efficient resource management, but it requires more frequent interactions between these entities thus leading to higher amount of signaling. Most of the current research on resource management focus on the low or intermediate interaction degree models, which allow RAT selection, vertical handover, admission control, and congestion control functions to be performed by a CRRM entity.

This paper proposes a resource management scheme for CRRM in heterogeneous wireless networks. The proposed scheme manages and allocates the abstracted radio bandwidth as resource for CRRM, and at the same time considers various factors such as resource efficiency, network load or congestion, cost, user preferences, etc. For example, when users request their services and/or vertical handovers, the proposed CRRM scheme allocates the required radio bandwidth from all the available resources across heterogeneous network entities. Therefore, the proposed scheme can be used in RAT selection, vertical handover, admission control, and congestion control, which are the functions in CRRM entity.

Since there are many factors to consider in resource allocation, the proposed method exploits a multiple criteria decision-making method, called *Analytic Hierarchy Process* (AHP) [3], which is a structured technique for dealing with complex decisions. Rather than simply prescribing a correct decision, the AHP helps to find one that best suits the goal. It provides a comprehensive and rational framework for structuring a decision problem by representing and quantifying its elements, relating those elements to overall goals, and evaluating alternative solutions. By exploiting AHP, the proposed scheme is flexible enough to be used in heterogeneous network environments containing variety of networks and various number of factors. The proposed method is validated through extensive simulations of different network environments.

The rest of this paper is organized as follows: Section 2 discusses the related work and explains the basics of AHP as a background for this paper. Section 3 formally describes the problem and proposes the AHP-based resource management scheme for CRRM. Section 4 examines several scenarios and validates that our proposed CRRM scheme can be adapted to various heterogeneous network environments, and finally, Section 5 concludes the paper.

2 Background

2.1 Analytical hierarchy process (AHP)

The Analytical Hierarchy Process (AHP) [3] is a multiple criteria decision-making method that decomposes a complex problem into a hierarchy of simpler and more manageable sub-problems. These sub-problems are referred to as *decision factors* or *criteria* and each factor is given a *weight* according to its relative importance to the problem. Finally, their importance to the problem is synthesized to find the best solution. AHP consists of the following three main steps: *hierarchy structuring*, *local weights calculation*, and *weight synthesis for global weights*. These steps are explained below:

2.1.1 Hierarchy structuring

The first step in AHP is to structure a problem as a hierarchy of multiple criteria. Figure 2 depicts the general AHP hierarchy structure, where the top-level representes the goal of the decision problem (e.g., resources allocation), the mid-level consists of various factors (e.g., efficiency, network load, cost, user preferences, etc.), and the bottom-level consists of alternatives or candidates (e.g., heterogeneous networks).

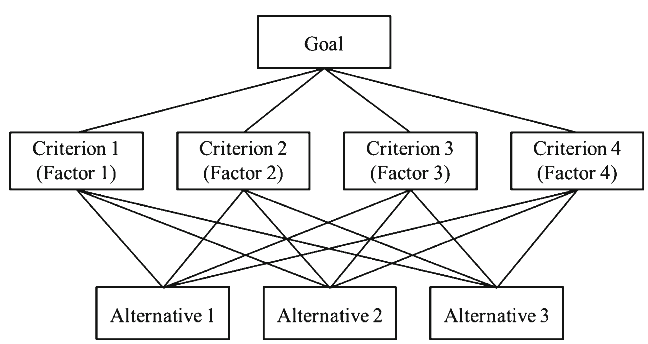

Fig. 2 AHP hierarchy structure

2.1.2 Calculating local weights

The second step involves determining the local weights, which represent the relative weights of the nodes within a group of siblings with respect to their parent. This is done by comparing each factor with all other factors within the same parent. The local weights consist of two parts: the weight of each decision factor to the goal and the weight of each candidate to each factor. Both are calculated using the same procedure consisting of pairwise comparison, calculation of weight vector, and consistency check. The following illustrates this process of determining the local weights of each factor relative to the goal.

A. Making pairwise comparison An evaluation matrix is developed by performing pairwise comparison of each decision factor based on the topmost goal. The comparison results are based upon user's expertise and experience by asking questions such as "Which is more important and by how much?" These initial values are captured in a square matrix A given by

$$A = (a_{ij})_{n\times n} = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{pmatrix}, \tag{1}$$

where a_{ij} denotes the ratio of the weight for i^{th} factor to the weight of the j^{th} factor, and n is the number of factors. The fundamental scale of 1 to 9 can be used to rank the judgments as shown in Table 1 [3]. The smaller weight in a pair is chosen as a unit and the larger weight is estimated as a multiple of that unit, and then a number is assigned based on the perceived importance. Similarly, the reciprocals of these numbers can be used to show the inverted comparison results. Thus, a reciprocal matrix can be obtained where the entries are symmetric with respect to the diagonal.

B. Calculating weight vector For the given matrix A in Eq. 1, its eigenvalue equation is defined as $AW = \lambda_{\max} W$, where W is a non-zero vector called the eigenvector, and $\lambda_{\max}$ is a scalar value called the eigenvalue.

Table 1 A fundamental scale of 1 to 9 (2, 4, 6, 8 indicate the medium value of pairwise comparison)

Number rating	Verbal judgment of preferences
1	Equally
3	Moderately
5	Strongly
7	Very
9	Extremely

W and $\lambda_{\max}$ appear as a pair and cannot be separated. After standardizing the eigenvector W, its elements are regarded as approximate local weights of decision factors denoted as

$$W = (w_1 \;\; w_2 \;\; \dots \;\; w_n)^T. \tag{2}$$

As a result, the weights of the decision factors can be obtained by calculating the eigenvector of an AHP matrix and its eigenvalue that is approximately equal to the number of assessed elements.

C. Checking for consistency If every element in Eq. 1 satisfies $a_{ij} = 1/a_{ji}$ and $a_{ik} \cdot a_{kj} = a_{ij}$, then the matrix A is a consistency matrix. However, the evaluation matrices are often not perfectly consistent due to users' random judgments. These judgment errors can be detected by a consistency ratio (CR), which is defined as the ratio of consistency index (CI) to random index (RI) given as follows:

$$CR = CI/RI, \tag{3}$$

where RI[1] is given in Table 2 [3], and CI is defined as

$$CI = (\lambda_{\max} - n) / (n - 1). \tag{4}$$

When $CR \leq 0.1$, the judgment errors are tolerable and the weight coefficients in W are the weights of decision factors for the topmost goal. Otherwise, the pairwise comparisons need to be adjusted until matrix A satisfies the consistency check, i.e., matrix A needs to be readjusted.

Similarly, the weights of candidates to each factor can also be calculated. For example, if there are m candidates, a $m \times m$ matrix B_k can be obtained for each factor k ($1 \leq k \leq n$) from Eq. 1 given as

$$B_k = (b_{ij})_{m\times m} = \begin{pmatrix} b_{11} & b_{12} & \dots & b_{1m} \\ b_{21} & b_{22} & \dots & b_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ b_{m1} & b_{m2} & \dots & b_{mm} \end{pmatrix}, \tag{5}$$

where b_{ij} denotes the ratio of the weight for i^{th} candidate to the weight of the j^{th} candidate.

Next, the eigenvector C_k is derived for each factor k, $1 \leq k \leq n$, using Eq. 2 as follows:

$$C_k = (c_{1k} \;\; c_{2k} \;\; \dots \;\; c_{mk})^T. \tag{6}$$

2.1.3 Weight synthesis for global weights

After the local weights of candidates are calculated, their global weights can be obtained by multiplying

[1] RI is the average CI of 500 randomly filled matrices and is compared with CI of A obtained in Eq. 1 for consistency check.

Table 2 Random index

n	1	2	3	4	5	6	7	8	9	10	11
RI	0	0	0.58	0.90	1.12	1.24	1.32	1.41	1.45	1.49	1.51

their local weights to the weight of its corresponding parent, i.e., decision factor. This is done by constructing a $m \times n$ matrix C from C_k, $1 \leq k \leq n$, as follows:

$$C = (c_{ij})_{m\times n} = \begin{pmatrix} c_{11} & c_{12} & \dots & c_{1n} \\ c_{21} & c_{22} & \dots & c_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ c_{m1} & c_{m2} & \dots & c_{mn} \end{pmatrix}. \tag{7}$$

Finally, the global weights for the candidates are calculated as

$$W_{\text{global}} = CW = (w_1 \;\; w_2 \;\; \dots \;\; w_m)^T. \tag{8}$$

The larger the global weight of a candidate, the higher the probability it will be selected.

2.2 Related work

As mentioned in Section 1, a CRRM entity performs RAT selection, vertical handover, admission control, and congestion control. Since the purposes of the RAT selection and vertical handover are usually associated with admission and congestion control, these control functions can be considered together. This section discusses the RAT selection and vertical handover methods with regard to CRRM functions.

Most of the existing vertical handover schemes focus on reducing the latency of handovers in order to provide seamless services to users [4–6]. There are few proposals where vertical handovers are performed for the purpose of admission/congestion controls regarding CRRM. Taha et al. proposed forced vertical handoffs as a powerful RRM scheme [7]. The basic idea is to balance the load across networks by forcing existing users to handover to the other supporting networks in order to make room for new users when the capacity of the current network is full. Gelabert et al. addressed the problem of congestion control in a scenario that considers GERAN and UTRAN, and proposed a vertical handover scheme jointly with a bit-rate reduction [8]. In their scheme, bit-rates of UTRAN users are reduced to avoid congestion and some GERAN users are handed over to UTRAN because the bit-rate reduction technique is not implemented in GERAN.

In contrast to vertical handovers, RAT selection methods are closely related to CRRM functions and can be used for both service initiation and vertical handover. Numerous RAT selection methods have been proposed and they can be classified into the following three categories: load-balancing based, service-type based, and RAT-type based.

Load-balancing is the basic principle for congestion control in CRRM. Tolli et al. proposed a simple load-balancing based CRRM scheme and proved the benefits of CRRM [9]. If the load of a cell exceeds a predetermined threshold, new connections are directed to the least loaded RAT cell. The authors also improved the scheme by introducing an adaptive load threshold rather than a fixed one [10]. The dynamic load threshold of a cell is adjusted periodically according to the loads of its inter-RAT cells. Piqueras et al. proposed a decentralized RAT selection scheme [11], where the price of each RAT is adjusted periodically according to its load based on the assumption that users always choose the cheapest RAT. This results in a balanced load across multiple RAT networks.

Under the service-type based RAT selection policy, there are two types of services: real-time (e.g. voice) and non real-time (e.g. WWW) services. Romero et al. proposed VG and VU policies [12], where the former allocates voice users to GERAN and WWW users to UTRAN while the latter is the opposite. Their simulation results show that the VG policy performs better than the VU policy. Under WLAN and WWAN environment, Hasib et al. proposed that real-time users are allocated to WWAN and non real-time users are to WLAN [13].

RAT-type based selection policies take into account network-specific characteristics. Since CDMA capacity degrades for indoor user traffic, Romero et al. proposed that indoor users be allocated to FDMA/TDMA-based GERAN while outdoor users are allocated to CDMA-based UTRAN [12]. Similarly, users who are located at the cell edge and thus experience more interference are allocated to FDMA/TDMA cells [14, 15]. In contrast, users with low propagation loss are allocated to CDMA cells.

The aforementioned methods consider only one or two factors, such as network load, service-type, or RAT-specific characteristics. Considering multiple factors can improve the efficiency of CRRM. Pillekeit et al. proposed a force-based load balancing scheme under UMTS/GSM environment that considers four factors: load, QoS, migration attenuation (i.e., the time since the last vertical handover occurred), and handover (the signaling overhead of vertical handovers) [16]. The importance of each factor is multiplied by a weight and the third and fourth factors are applied to vertical handovers. The cell with the largest value is then selected as the target cell. Agusti et al. proposed a fuzzy-neural based approach [17, 18]. This complicated approach includes three main blocks: fuzzy-neural algorithm, reinforcement learning, and multiple decision making. The fuzzy-neural algorithm aims to allocate a numerical indication named Fuzzy Selected Decision (FSD) to each RAT. The value of a FSD is in the range from 0 to 1, which is determined by a set of linguistic variables based on technical measurements, such as signal strength, resource availability, and mobile speed. The reinforcement learning procedure is used to select and adjust the parameters used in the fuzzy-neural algorithm. Finally, the multiple decision making block decides the most suitable RAT based on technical related inputs (Received Signal Strength, Resource Availability, and Mobile Speed) coming from the fuzzy-neural block and techno-economic related inputs (Cost) such as user demand and operator preferences.

The related work discussed thus far deals with only specific factors and/or specific network characteristics. However, future heterogeneous wireless networks will have many different radio access technologies and factors to consider. Therefore, this paper aims to propose a simple and flexible scheme which is general enough to be used in any network environment with any number of factors. In Section 4, we will discuss how these related methods can be implemented using the proposed AHP-based scheme.

3 Proposed scheme

This section presents the proposed AHP-based resource management scheme that can be used in RAT selection, vertical handover, admission control, and congestion control, which are the functions in a CRRM entity. The goal is to efficiently allocate and manage radio resources among heterogeneous wireless networks by considering various factors, such as service-type, network load or congestion, wireless channel condition, resource efficiency, cost, user preferences, etc.

Before discussing the details of the proposed scheme, the meaning of *radio bandwidth* needs to be defined. In current broadband wireless access technologies, one or more basic *radio resource units* are allocated to each user. For example, Fig. 3 illustrates the radio resource unit in the Orthogonal Frequency-Division Multiplexing (OFDM) system. A unit in OFDM is represented in two dimensions (time and frequency domains) and is called 'radio unit' (RU) in 3GPP packet channel. The radio resource unit can also be represented in 1-dimension where the packet channel is shared in time domain only, which is the case for 'time slot' in 3GPP2 1xEV-DO.

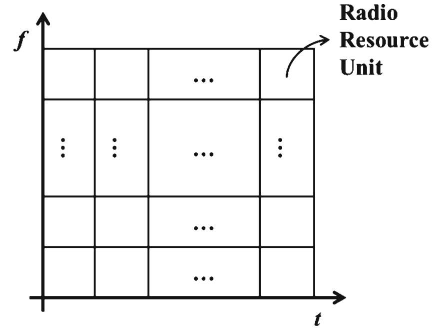

Fig. 3 Radio resource unit in OFDM system

For each radio resource unit, the Adaptive Modulation and Coding (AMC) scheme is applied according to each user's estimated channel condition. Thus, the amount of data may be different across all the radio resource units. For example in High-Speed Downlink Packet Access (HSDPA), a user u_1 in a bad channel may be allocated 1.8 Mbps with QPSK modulation, which is robust and tolerates higher levels of interference but provides low data rates. In contrast, 14.4 Mbps can be allocated to another user u_2 in a good channel with 64-QAM, which provides high data rate but is susceptible to interference signals. Therefore, if one radio resource unit is allocated to each of u_1 and u_2, u_2 can send 8 times more data than u_1. In other words, 8 times more radio resource units should be allocated to u_1, compared with u_2, to send the same amount of data. In order to make RAT-specific radio resources manageable in CRRM and to precisely model the current broadband packet channel, the *radio bandwidth* is abstracted as the actual radio resources allocated to support a transmission data rate under a given channel condition.

The process of allocating radio bandwidth is explained below:

– Initially, a base station has a radio bandwidth of $\Phi = \Phi_{max}$, where Φ_{max} is its maximum transmission data rate.
– If user u_i requests a real-time service (e.g., VoIP) with data rate of r_i and channel condition c_i, then a radio bandwidth of $\phi_i = f(r_i, c_i)$ is dedicated to u_i, where $f(r_i, c_i)$ represents a relationship between the required radio bandwidth and the actual transmission bandwidth with channel condition. After the base station allocates ϕ_i to u_i, its available radio bandwidth becomes $\Phi = \Phi - \phi_i$.

- If user u_j requests a non real-time service (e.g., WWW) with channel condition c_j, the available radio bandwidth of the base station Φ is shared among l users whom also requested non real-time services. Therefore, user u_j is serviced with the rate $r_j = g(\Phi/l, c_j)$, where $g(\phi_j, c_j)$ represents the transmission data rate when a user under channel condition c_j has radio bandwidth of ϕ. This concept is called *proportional fair scheduling* [19].

The first step of the proposed method is to structure the problem as a hierarchy as explained in Section 2.1. Suppose there are n factors F_i $(1 \le i \le n)$ that need to be considered, and m candidate networks N_j $(1 \le j \le m)$ from which the radio resources can be allocated to a user request. Figure 4 depicts the AHP hierarchy structure, where the top-level is the goal of the resource allocation, the mid-level consists of various factors for resource allocation F_i $(1 \le i \le n)$, and the bottom-level consists of candidate networks N_j $(1 \le j \le m)$.

The second step is to calculate the relative local weights of decision factors to the goal. By comparing the importance of F_i $(1 \le i \le n)$ with each other, the scale α_i for F_i can be obtained from Table 1. If a_{ij} for $1 \le i, j \le n$ is given by α_i/α_j, then the matrix A and its eigenvector W can be obtained as defined by Eqs. 1 and 2, respectively. If $a_{ij} = \alpha_i/\alpha_j$, $a_{ij} = 1/a_{ji}$, and $a_{ik} \cdot a_{kj} = a_{ij}$ for $1 \le i, j \le n$ are satisfied, then this proves that the matrix A is consistent.

The third step is to calculate the relative local weights of each candidate network to each factor. The scale β_j for the candidate network N_j $(1 \le j \le m)$ is assigned from Table 1 by comparing the relative preference of N_j with each other from the perspective of each factor F_k $(1 \le k \le n)$. Then, c_{ij} for C_k is given by β_i/β_j for $1 \le i, j \le m$ in Eq. 7, which results in the matrices C_k and their eigenvectors X_k for $1 \le k \le n$.

The final step is to synthesize the above results to achieve the overall global weights as defined by Eqs. 7 and 8 and to choose the one with the largest weight. In case of a real-time service request, the radio bandwidth ϕ is allocated from the selected network. If the selected network has insufficient radio bandwidth for the request, then it can be allocated from the network with the second largest weight.

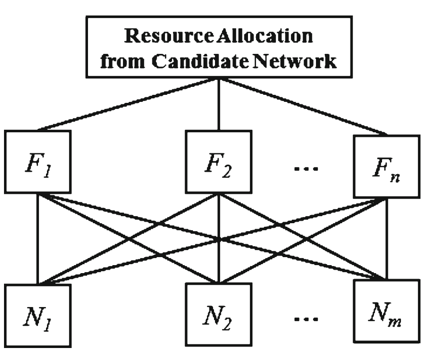

Fig. 4 CRRM hierarchical architecture

4 Evaluation

4.1 Validation of flexibility

This section applies the proposed method to the various related schemes described in Section 2.2 based on the four steps explained in Section 3. Our analysis validates that the proposed scheme is simple and flexible enough to be used in any network environment with any number of factors for CRRM. In addition, it also validates that the performance of our scheme is equal or superior to the related work because more factors can be considered.

The load-balancing based schemes discussed in [9–11] consider a simple goal, i.e., the least loaded candidate network has the highest priority. Implementing these schemes involve the following four steps:

- Construct a hierarchy as shown in Fig. 4, where there is only a single factor, F_1, reflecting the network load (i.e., $n = 1$). Since the authors of the prior work considered GSM and WCDMA as the candidate networks, N_1=GSM and N_2=WCDMA (i.e., $m = 2$).
- The second step is skipped since only one factor is considered.
- The scales β_1 and β_2 are used as local weights for the candidate networks relative to the network load. If the load of GSM is higher than that of UTRAN, $\beta_1 = 1$ and $\beta_2 = 2$; otherwise, $\beta_1 = 2$ and $\beta_2 = 1$. Then, the matrix C_1 is calculated from from Eqs. 5 and 6.
- Finally, W_{global} is obtained from Eq. 8 and the candidate network with the largest global weight is selected.

As mentioned in Section 2.2, the service-type based schemes [12, 13] give priorities to candidate networks according to their service-types, while the RAT-type based schemes [12, 14, 15] give priorities to candidate networks according to their RAT-types. The following steps implement the scheme presented in [12], since it considers both the service-type and the RAT-type:

- Construct a hierarchy as shown in Fig. 4, where F_1 indicates the service-type (i.e., real-time vs. non real-time services) and F_2 represents the user location (i.e., indoor vs. outdoor). The candidate networks are N_1 =GERAN and N_2 =UTRAN.

- For the local weights of two factors to the goal, if F_1 is considered more importantly than F_2, then $\alpha_1 = 2$ and $\alpha_2 = 1$; otherwise, it is vice versa. Romero et al. [12] showed that giving higher preference to F_1 results in better performance on high WWW load, while giving preference to F_2 is more desirable on high voice load. The local weights are obtained from Eq. 2.
- For the local weights of N_1 and N_2 for F_1, $\beta_1 = 2$ and $\beta_2 = 1$ for real-time services and $\beta_1 = 1$ and $\beta_2 = 2$ for non real-time services. For the local weights of N_1 and N_2 for F_2, $\beta_1 = 2$ and $\beta_2 = 1$ for indoor users and $\beta_1 = 1$ and $\beta_2 = 2$ for outdoor users. Then, the matrices C_1 and C_2 can be calculated from Eqs. 5 and 6.
- Finally, W_{global} is obtained from Eq. 8 and the candidate network with the highest global weight is chosen.

The fuzzy-neural based approach discussed in [17, 18] consider multiple factors, but their scheme is quite complicated as described in Section 2.2. Therefore, its approximate version is implemented as follows:

- Construct a hierarchy as shown in Fig. 4, where F_i's $(1 \le i \le 5)$ indicate signal strength, resource availability, mobile speed, user preference, and operator preference, respectively. The candidate networks are N_1 =UTRAN, N_2 =GERAN, and N_3 =WLAN.
- Their work defines signal strength, resource availability, and mobile speed as the *technical criterion*, and assumes that the user preference is three times more important than the technical criterion and the operator preference is two times more important than the technical criterion. Therefore, the local weights of these five factors to the goal are $\alpha_1 = 1$, $\alpha_2 = 1$, $\alpha_3 = 1$, $\alpha_4 = 3$, and $\alpha_5 = 2$. The local weights can then be obtained from Eq. 2.
- The local weights of N_1, N_2, and N_3 for each F_i's $(1 \le i \le 5)$ and β_j's $(1 \le j \le 3)$ are assigned according to the current status of mobile terminal or user/operator preference. For example, for F_1, if $s_1 < s_2 < s_3$, where s_j is the signal strength of N_j, $1 \le j \le 3$, then β_j's are assigned as $\beta_1 < \beta_2 < \beta_3$. Then, C_i's $(1 \le i \le 5)$ can be calculated from Eqs. 5 and 6.
- Finally, W_{global} is obtained from Eq. 8 and the candidate network with the highest global weight is chosen.

Table 3 Simulation parameters

Symbol	Parameter	Value
V	VoIP data rate	19200 bps
T_s	VoIP service time	30 seconds
S	Web page size	3223 bytes
T_v	Interarrival time of VoIP user	180 seconds
T_w	Interarrival time of WWW user	30 seconds

4.2 Experimental results

This section analyzes the proposed AHP-based resource management scheme on several simulation scenarios. We first present traffic, user, radio bandwidth models that are independent of simulation scenarios. Then, scenario-dependent models are explained in each subsection.

As for the traffic model, VoIP (real-time) and WWW (non real-time) services are considered. It is assumed that the data rate of VoIP services is V and the service time is exponentially distributed with a mean of T_s. The web page size is assumed to follow an exponential distribution with a mean of S.

The inter-arrival times of VoIP and WWW users are assumed to be exponentially distributed with T_v and T_w, respectively, and 40% of users request VoIP while 60% of users request WWW services. Each user has a profile to indicate its preference: i.e., low cost, good service quality (e.g., service delay), or balanced cost and quality.

The radio bandwidth model described in Section 3 is assumed but the functions $\phi = f(r, c)$ and $r = g(\phi, c)$, which indicate the relationship between the actual transmission bandwidth r and the radio bandwidth ϕ with channel condition c, need to be defined. Therefore, four levels of channel conditions are assumed, and thus, four types of coding rates are used for a given channel condition. The channel condition c is given as 1, 1/2, 1/4, or 1/8 according to the distance of users from the base station. Based on this, ϕ and r are given as[2]:

$$\phi = f(r, c) = r/c \tag{9}$$

$$r = g(\phi, c) = \phi \cdot c \tag{10}$$

Finally, 40% of the radio bandwidth are reserved for VoIP users and the admission control is performed for real-time services. The simulation results are based on 3,000 to 15,000 users in steps of 3000. 15,000 users represent a threshold where the blocking probability

[2] In 3GPP HSDPA, the actual transmission bandwidth is 1.8, 3.6, 7.2, or 14.4 Mbps, which perfectly match our assumption.

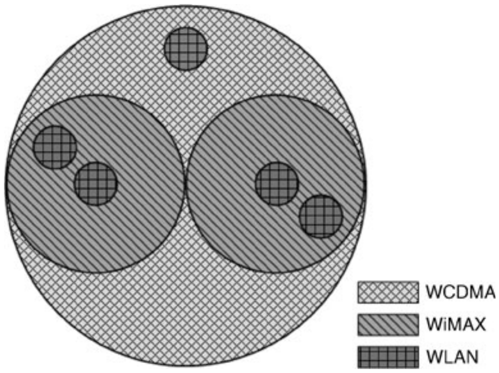

Fig. 5 Simulation topology for heterogeneous wireless networks

becomes over 10% and thus represents a heavily loaded environment.

Table 3 summarizes the above-mentioned simulation parameters used in the study, which are similar to ones used in [23, 24].

4.2.1 Scenario 1

In scenario 1, the proposed scheme is applied to a realistic heterogeneous wireless network consisting of WLAN, WiMAX, and WCDMA as shown in Fig. 5.

As mentioned earlier, three kinds of user groups exist according to their preferences: **Cost** (preference to low cost), **QoS** (preference to good service quality), and **Balanced** (balanced preference). Since users desire services with lost cost and/or good quality, these two factors are considered in our proposed scheme as follows:

– In Fig. 4, F_1, indicates the service cost and F_2 represents the available service bandwidth. The candidate networks are $N_1 = \text{WLAN}$, $N_2 = \text{WiMAX}$, and $N_3 = \text{WCDMA}$.
– For the local weights of two factors to the goal, $\alpha_1 = 2$ and $\alpha_2 = 1$ for the **Cost** user group; $\alpha_1 = 1$ and $\alpha_2 = 2$ for the **QoS** user group; and $\alpha_1 = 1$ and $\alpha_2 = 1$ for the **Balanced** user group.
– For the local weights of N_1 and N_2 for F_1, β_i's $(1 \leq i \leq 3)$ are given as 3, 2, and 1, respectively, based on the cost information provided in Table 4. For the local weights of N_1 and N_2 for F_2, β_i's $(1 \leq i \leq 3)$ are assigned according to the available service bandwidth of N_i's $(1 \leq i \leq 3)$, where the larger β_i the more the available bandwidth. A larger available service bandwidth indicates a lower blocking probability for VoIP services and a lower service delay for WWW services.
– Finally, W_{global} is obtained from Eq. 8 and the candidate network with the highest global weight is chosen.

Table 4 Network parameters [20–22]

Parameter	WLAN	WiMAX	WCDMA
Bandwidth	11Mbps	40Mbps	14Mbps
Cost	\$0.10/1MB	\$0.20/1MB	\$0.39/1MB
Radius	100*m*	500*m*	1000*m*

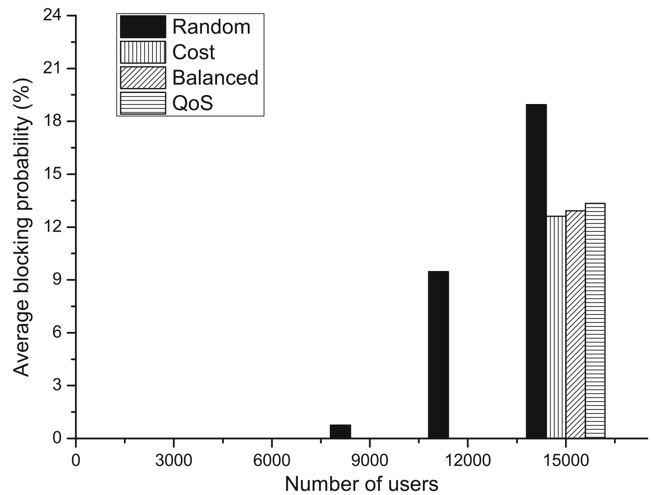

Fig. 6 Results of the blocking probability

The simulation was performed using the proposed scheme and a random selection scheme. For the proposed scheme, users are uniformly distributed across the three groups (**Cost**, **Balacned**, and **QoS**).

Figures 6 and 7 show the results for VoIP services. As shown in Fig. 6, the average blocking probability of the proposed scheme is much lower than that of the

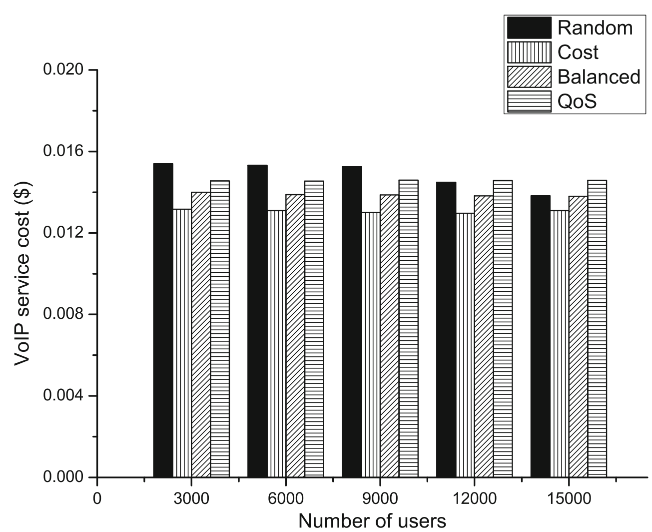

Fig. 7 Results of the voice service cost for VoIP users

random selection. This is expected because the proposed scheme considers the available service bandwidth to balance the load of networks for real-time services. Figure 7 shows that the **Cost** group has the lowest service cost, which makes sense since the users in this group chose cost as the highest priority. Moreover, the average costs for all three groups decrease slightly as the load (i.e., the number of users) increases. This is because WCDMA, which covers a wide area but provides low bandwidth, saturates early and as a result users are serviced on cheaper WLAN or WiMAX leading to lower cost. In case of the random selection, since WCDMA saturates earlier than the proposed scheme, the cost of the random selection decreases far more than that of the proposed scheme. In fact the cost for the random selection is approximately the same as the **Balanced** group when the network load is the highest.

Figures 8 and 9 show the results for WWW services. As expected, users that are located in the area covered only by WCDMA have no alternative choice. Users in the **Cost** group who are located in the other area tend to choose WLAN and users in the **QoS** group select WiMAX according to their preference. This can be seen from Figs. 8 and 9 where the **Cost** group has the lowest cost and the **QoS** group has the shortest service time. The service time increases as the number of users increases because the radio bandwidth is shared among users who are serviced concurrently. Similar to Fig. 7, the service cost decreases as the network load increases and the reason for this is the same as the case of VoIP services as mentioned above.

On the hand, users in the **Balanced** group consider both service cost and available bandwidth. On a lightly loaded network (i.e., when the number of users is small), users tend to choose WiMAX rather than WLAN because the former has the highest available bandwidth and medium cost while the latter has the lowest cost but the lowest available bandwidth. As the offered load increases, the available bandwidth of WiMAX decreases to the level provided by WLAN or WCDMA, and thus, the users in the **Balanced** group no longer select WiMAX. This can be seen from Figs. 8 and 9, where the service delay and cost of the **Balanced** group are close to those of the **QoS** group when the load is light load but becomes closer to those of the **Cost** group as the offered load becomes heavier.

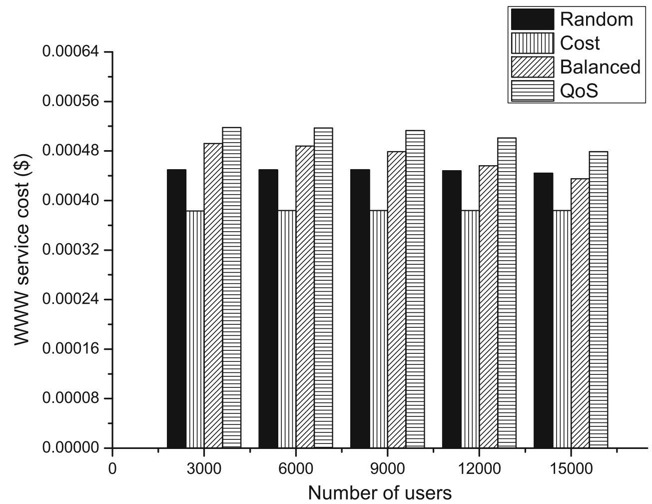

Fig. 9 Results of the service cost for WWW users

4.2.2 Scenario 2

The study of Scenario 1 showed that our proposed scheme properly reflects user's preference. However,

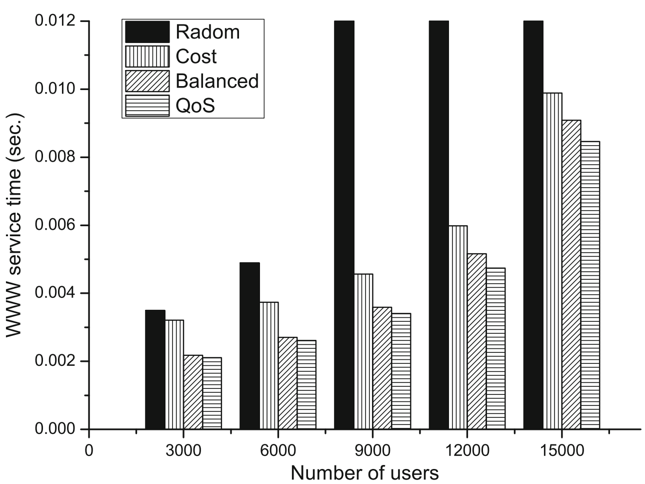

Fig. 8 Results of the service time for WWW users

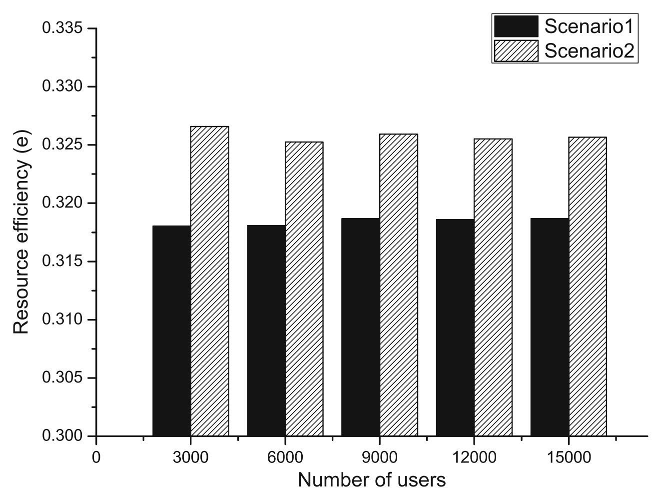

Fig. 10 Comparison of the resource efficiency for VoIP users between Scenarios 1 and 2

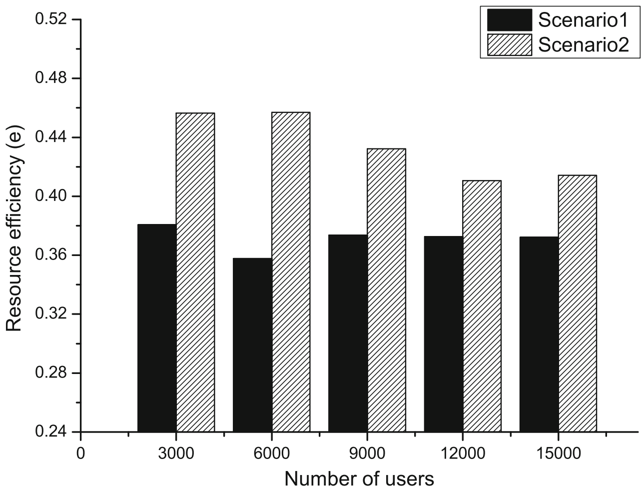

Fig. 11 Comparison of the resource efficiency for WWW users between Scenarios 1 and 2

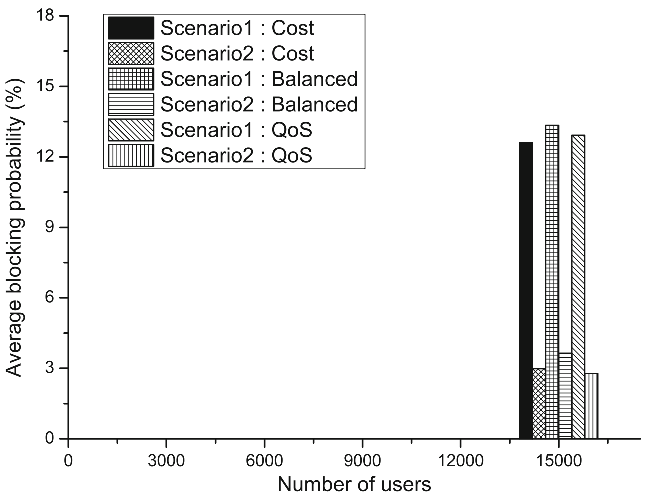

Fig. 12 Comparison of the blocking probability between Scenarios 1 and 2

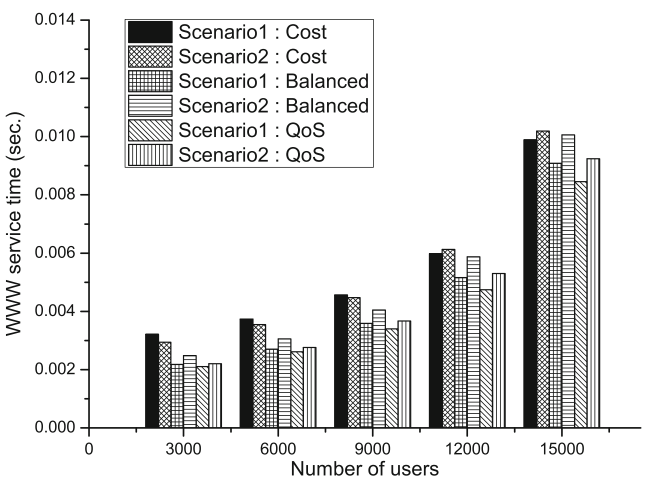

Fig. 13 Comparison of the WWW service time between Scenarios 1 and 2

Table 5 Energy comparison using WLAN, GSM and WCDMA. [25]

GSM	WLAN & WiMAX	WCDMA
0.10 Joule/KB	0.15 Joule/KB	0.25 Joule/KB

Scenario 1 does not consider operator's preferences. Therefore, Scenario 2 considers the resource efficiency for network operators. The resource efficiency of the radio bandwidth e is given as

$$e = r/\phi, \tag{11}$$

where the radio bandwidth ϕ and the actual transmission bandwidth r were defined in Eqs. 9 and 10, respectively.

In order to increase e, the channel condition of users needs to be considered. That is, high resource efficiency can be obtained if users choose a network with a good channel condition. Therefore, a factor F_3 that indicates the channel condition is added to the AHP-based resource management algorithm described in Scenario 1.

Figures 10 and 11 show the increase in resource efficiency compared with Scenario 1. Figure 12 shows that the blocking probability is significantly lower when the number of VoIP services is high. However, Fig. 13 shows that the service time is not affected by increase in WWW services. In certain cases, the service delays of **Balanced** and **QoS** groups in Scenario 2 are equal to or a little lower than that of Scenario 1 because users may select a network with a good channel condition but the

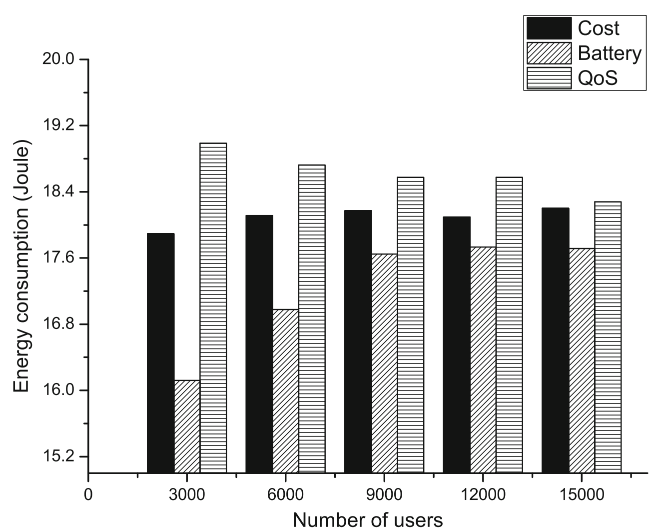

Fig. 14 Results of the energy consumption for VoIP services

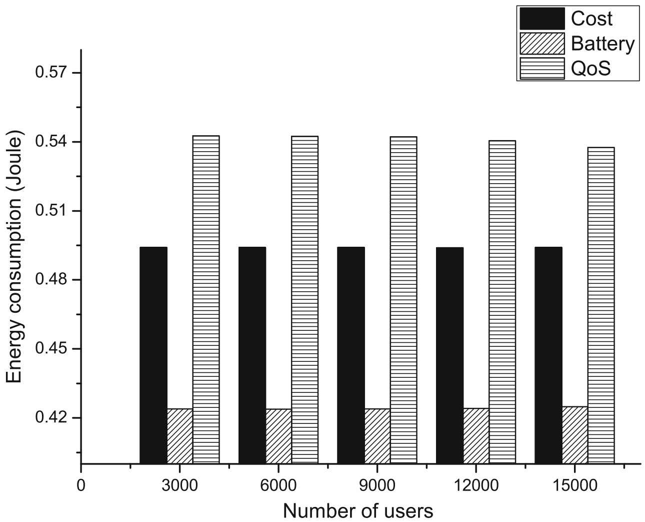

Fig. 15 Results of the energy consumption for WWW services

available bandwidth is low. In contrast, the service time for the **Cost** group is smaller compared with Scenario 1 when the load is light because the good channel condition increases the actual service bandwidth. However, the differences are very small and can be neglected. The gain in resource efficiency for WWW services, as shown in Fig. 11, can be used to provide more real-time services for network operators' benefit.

4.2.3 Scenario 3

In order to verify the flexibility and effectiveness of the proposed scheme, a factor F_4, which indicates the remaining battery power, and N_4 = GSM network are added to Scenario 2. We assume that one GSM network covers the same area as WCDMA in Fig. 5 and users in the **Battery** group desire low battery usages. Table 5 summarizes the energy consumption of networks used in the simulation [25]. Figures 14 and 15 show that users in the **Battery** group consume less energy than users in the other groups, while the results of **Cost** and **QoS** groups show a similar behavior with Scenario 2. This demonstrates the flexibility and effectiveness of our AHP-based resource management scheme. In particular, the proposed scheme reflects well the users' preference when the offered load is light as shown in Fig. 14.

5 Conclusion

In the near future, various wireless services could be increased in heterogeneous wireless networks. In this heterogeneous wireless network environment, radio and/or network resources need to be managed in a cooperative manner with the CRRM strategy. The CRRM strategy enables that users can receive better services and network operators can manage their resources more efficiently. The existing works on current CRRM strategies have limitations for application to various wireless networks with various considerations. This paper proposed an AHP-based resource management scheme for CRRM. AHP allows the proposed scheme to be simple and flexible enough to be used in any network environment with consideration of any decision factors. Our scheme allocates and manages the radio bandwidth and models the current broadband wireless access technologies precisely. The flexibility and effectiveness of the proposed scheme have been validated through extensive simulations of various scenarios.

Acknowledgements This research was supported by the MKE(The Ministry of Knowledge Economy), Korea, under the ITRC(Information Technology Research Center) support program supervised by the NIPA(National IT Industry Promotion Agency) (NIPA-2012-(C1090-1121-0003))

References

1. Wu L, Sandrasegaran K (2007) A survey on common radio resource management. In: Proc. of wireless broadband and ultra wideband communications, pp 60–66
2. Sallent O (2006) A perspective on radio resource management in B3G. In: Proc. of wireless communication systems, pp 30–34
3. Saaty TL (2000) Fundamentals of decision making and priority theory with the analytic hierarchy process. RWS Publications, USA
4. Stevens-Navarro E, Wong VWS, Lin Y (2007) A vertical handoff Decision algorithm for heterogeneous wireless networks. In: Proc. of wireless communications and networking conference, pp. 3199–3204
5. Wang LC, Chen A, Chen HH (2007) Network selection with joint vertical and horizontal handoff in heterogeneous WLAN and mobile WiMAX systems. In: Proc. of vehicular technology conference. Spring, pp 794–798
6. Stevens-Navarro E, Wong VWS (2006) Comparison between vertical handoff decision algorithms for heterogeneous wireless networks. In: Proc. of vehicular technology conference. Spring, pp 947–951
7. Taha AEM, Hassanein HS, Mouftah HT (2006) Exploiting vertical handoffs in next generation radio resource management. In: Proc. of communications, pp 2083–2088
8. Gelabert X, Perez-Romero J, Sallent O, Agusti R (2007) On managing multiple radio access congestion events in B3G scenarios. In: Proc. of vehicular technology conference. Spring, pp 1137–1141
9. Tolli A, Hakalin P, Holma H (2002) Performance evaluation of common radio resource management (CRRM). In: Proc. of communications, pp 3429–3433

10. Tolli A, Hakalin P (2002) Adaptive load balancing between multiple cell layers. In: Proc. of vehicular technology conference, Fall, pp 1691–1695
11. Piqueras R, Perez-Romero J, Salient O, Agusti R (2006) Dynamic pricing for decentralised RAT selection in heterogeneous scenarios. In: Proc. of personal, indoor and mobile radio communications, pp 1–5
12. Perez-Romero J, Salient O, Agusti R (2005) Policy-based initial RAT selection algorithms in heterogeneous networks. In: Proc of 7th mobile wireless communication networks, pp 1–5
13. Hasib A, Fapojuwo AO (2007) Performance analysis of common radio resource management scheme in multi-service heterogeneous wireless networks. In: Proc. of wireless communications and networking conference, pp 3296–3300
14. Perez-Romero J, Salient O, Agusti R, Garcia N, Wang L, Aghvami H (2006) Network-controlled cell-breathing for capacity improvement in heterogeneous CDMA/TDMA scenarios. In: Proc. of wireless communications and networking conference, pp 36–41
15. Perez-Romero J, Ferrus R, Salient O, Olmos J (2007) RAT selection in 3GPP-based cellular heterogeneous networks: from theory to practical implementation. In: Proc. of wireless communications and networking conference, pp 2115–2120
16. Pillekeit A, Derakhshan F, Jugl E, Mitschele-Thiel A (2004) Force-based load balancing in co-Located UMTS/GSMnetworks. In: Proc. of vehicular technology conference, Fall, pp 4402–4406
17. Agusti R, Salient O, Perez-Romero J, Giupponi L (2004) A fuzzy-neural based approach for joint radio resource management in a beyond 3G framework. In: Proc. of quality of service in heterogeneous wired/wireless networks, pp 216–224
18. Giupponi L, Agusti R, Perez-Romero J, Sallent O (2005) A novel joint radio resource management approach with reinforcement learning mechanisms. In: Proc. of performance, computing, and communications conference, pp 621–626
19. Bender P, Black P, Grob M, Padovani R, Sindhushyana N, Viterbi A (2000) CDMA/HDR: a bandwidth efficient high speed wireless data service for nomadic users. In: IEEE communications magazine, pp 70–77
20. http://www.wireless.att.com/
21. http://www.sprint.com/
22. http://www.verizonwireless.com/
23. Das SK, Lee E, Basu K, Kakani N, Sen SK (2002) Performance optimization of VoIP calls over wireless links using H.323 protocol. In: Proc. of INFOCOM, pp 1386–1394
24. Timmins PJ, McCormick S, Agu E, Wills CE (2006) Characteristics of mobile web content. In: Proc. of hot topics in web systems and technologies, pp 1–10
25. Niranjan B, Aruna B, Arun V (2009) Energy consumption in mobile phones: a measurement study and implications for network applications. In: Proc. of internet measurement conference, pp 280–293