RESEARCH　　　　　　　　　　　　　　　　　　　　　　　**Open Access**

CrossMark

Energy efficiency optimization-oriented control plane and user plane adaptation with a frameless network architecture for 5G

Xiaodong Xu[1*], Xun Dai[1], Ya Liu[2], Rui Gao[3] and Xiaofeng Tao[1]

Abstract

In order to fully explore the merits of the multi-tier heterogeneous network (HetNet) and ultra-dense small cell deployments, the evolution of the cellular network architecture is on the way. Based on the frameless network architecture, this paper proposes the control plane (CP) and user plane (UP) adaptation strategy to improve the system energy efficiency (EE). In this paper, we give a three-step system EE optimization with constraints on the CP/UP adaptation. We optimize the system EE via CP and UP construction and adaptation while guaranteeing the user quality of service (QoS). In the first step, the CP construction and adaptation scheme is proposed through the Voronoi Diagram with EE improvements. In the second step, by jointly allocating the AE and subchannel resources, we construct the on-demand user-centric UP. In the last step, based on the UP construction from the second step, we adjust each UP to optimize the EE performance. This is achieved by employing game theory. Furthermore, we verify the existence and uniqueness of the Nash Equilibrium for the non-cooperative game model with pricing. Finally, the system-level simulation results show that, with constraints on the QoS of the users, the system EE performances are improved.

Keywords: CP and UP adaptation; Frameless network architecture; Energy efficiency; Joint resource allocation; Power adjustment; Genetic algorithm; Game theory; Nash Equilibrium

1 Introduction

Currently, the mobile Internet applications and versatile mobile services are affecting every aspect of our daily life. Specifically, the dramatic increase of the data traffic poses great challenge to the network capacity and forces the mobile operators to make revolutionary changes. Besides expanding the spectrum and improving the radio transmission, the mobile network architecture is considered as another potential way of further increasing the capacity of the 5th generation mobile systems (5G) [1, 2].

Heterogeneous network (HetNet) was proposed to meet the future demand by deploying multiple types of base stations (BSs) into the existing networks [3]. In addition, the small cell network (SCN) was introduced into the standard framework of the 3rd Generation Partnership Project (3GPP) to increase hot-spot capacity [4]. Considering the coordinated multi-point transmission/reception (CoMP) and relay techniques [5], the future mobile radio access network (RAN) architecture will have many distinct features comparing to current networks, such as:

- Small cell-based ultra-dense cellular deployments;
- Multi-tier heterogeneous network supporting;
- More BSs than users;
- Coordination among different types of BSs;
- Dense and irregular cellular network topology;
- Separation of the control plane (CP) and user plane (UP);
- Flexible cell association and dual connectivity for users.

With all these new features of the HetNet, the RAN architecture has to evolve dramatically. There are many aspects needed to be researched in the future, such as

* Correspondence: xuxiaodong@bupt.edu.cn
[1]National Engineering Laboratory for Mobile Network Security, Beijing University of Posts and Telecommunications, Beijing 100876, China
Full list of author information is available at the end of the article

Springer

Xu *et al. EURASIP Journal on Wireless Communications and Networking* (2015) 2015:159

Page 2 of 15

the cellular network modeling with dense and irregular deployments, the influence of separated CP&UP on system energy efficiency (EE), and so on.

Currently, there is significant research on the C-RAN, Open RAN, Soft Cell, and frameless network architecture (FNA) [6–10]. Based on the baseband pool, China Mobile Research Institute proposed C-RAN featuring a centralized base band unit, coordination, and cloud computing [6]. A software-defined RAN architecture is implemented through the virtualization in [7]. For the Soft Cell concept proposed in [8], the transparent sets of BSs are provided for users.

Research on the FNA has also been studied substantially. In [9, 10], the concept of the FNA was given to explore the capacity gain in the network architecture evolution. Several important metrics, e.g., the user-centric frameless coverage definition with coordinated serving set construction, centralized resource management, CP/UP separation and adaptation, were considered and found to support the FNA networking. In the FNA, the resource pooling and service slicing schemes are proposed for different user quality of service (QoS) guaranteed resource allocations. According to the booming and strong research conducted in this field, FNA is a promising solution for fulfilling the requirements of 5G. Specifically, the user-centric service and cloud computing features are important for managing the multi-tier HetNets.

The research topics for the FNA include the network topology modeling, on-demand user-centric serving set construction, CP/UP adaptation, mobility management, and so on. In this paper, we focus on the CP/UP adaptation schemes based on the FNA.

In mobile networks, 3GPP has already started discussions on the CP/UP separation. The RAN2 Study Item - Small Cell Enhancement studies the dual connectivity in the small cell scenarios. It allows users to connect to both small cells and macrocells at the same time, which can effectively reduce the frequency of handovers. Therefore, CP and UP are separated to achieve high efficiency, flexibility, and low cost. Furthermore, 3GPP has designed two alternative architectures for the CP and nine alternative architectures for the UP. In this scheme, the CP/UP may both be maintained cooperatively by the macro- and small cells.

However, the separation of the CP and UP introduces new relationship between their coverage areas, which could be modeled by constraints on the deployment of CP and UP. Considering the constraints, the system EE could be further improved by dynamically adjusting the separated CP and UP. This method is defined as the CP/UP adaptation. The CP/UP adaptation also raises challenges for the AE association on the aforementioned user connectivity, such as choosing the "best" AE for the specific user considering both the user QoS requirement and the system EE.

In this paper, we present the constraints and adaptation scheme for the CP and UP separation. A three-step optimization solution is given. The main contributions of this paper are highlighted as follows:

- CP/UP separation-based adaptation model is constructed with constraints.
- The CP construction and adaptation scheme is given using the Voronoi Diagram with EE optimization.
- Considering the user QoS requirements, on-demand user-centric UP construction with joint resource allocation (AE and subchannel) for establishing the serving set is studied. Equal power allocation for all the UPs are maintained by the AEs within the serving set. The coordination between the AEs is implemented for better supporting different user QoS requirements.
- Game theory is employed for power adjustment based UP adaptation to achieve the optimal EE performance. Additionally, we also verified the existence and uniqueness of the Nash Equilibrium for the proposed non-cooperative game model.

The structure of this paper is arranged as follows. In Section 2, we give the system model for the FNA and put constraints on the CP/UP. In Section 3, we propose a Voronoi Diagram-based CP construction and adaptation for EE optimization. Then, we study the joint allocation of the AE and subchannels for UP construction in Section 4 and game theory-based power adjustment for UP adaptation in Section 5. Finally, we give the EE performance evaluation and simulation of CP/UP adaptation in Section 6.

2 FNA deployment scenario for the CP/UP adaptation

The target of FNA is to give a set of principles to guide the development of user-centric network architecture for the 5G. The typical deployment scenario of FNA is shown as in Fig. 1 with the core network (CN), RAN deployments, and user equipment (UE).

The RAN of FNA consists of two network elements. By decomposing the traditional BS into centralized processing entity (CPE) and antenna element (AE), the RAN of FNA consists of two new network elements. The function of the CPE is to maintain the networking, implement the signal processing, handle the CP and UP, manage the radio resources including the connected AEs, and construct on-demand user-centric serving set for specific users. The CPE can be located at a macro BS or at another type of BS that has the required processing ability. The CPE mainly lies as a logical node. The AE is
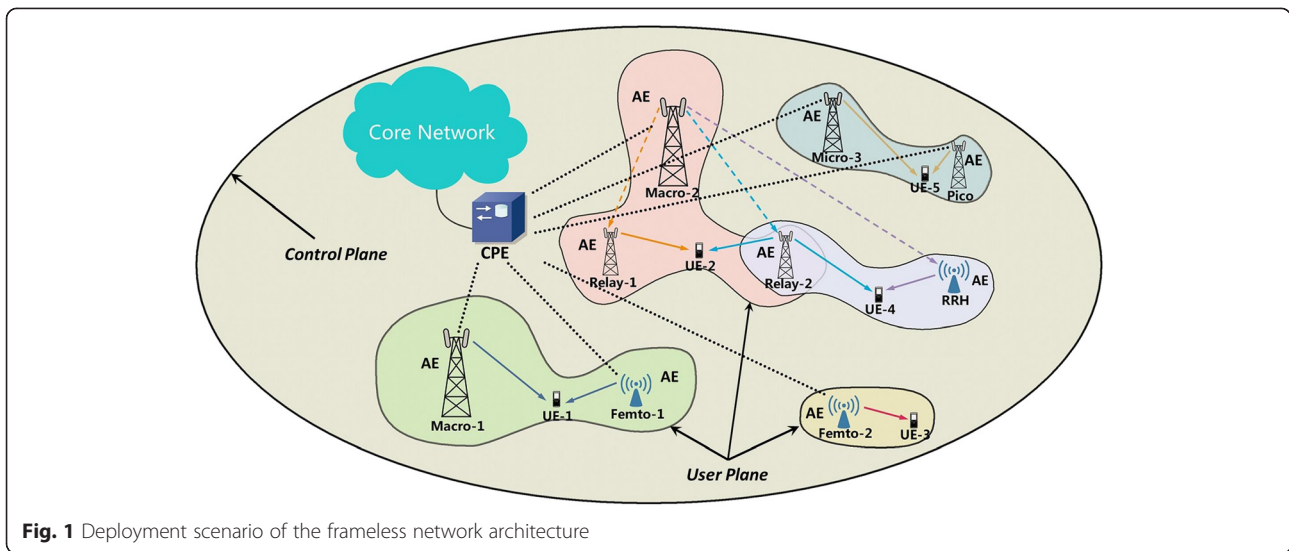
Xu *et al. EURASIP Journal on Wireless Communications and Networking* (2015) 2015:159

Page 3 of 15

Fig. 1 Deployment scenario of the frameless network architecture

responsible for the radio signals' transmission/reception. The backhaul links between the CPE and AE could be the fiber, wireless backhaul, wired connections, and so on. The capacity and latency abilities for different backhaul links are different, which also pose challenges to resource allocations. The backhaul issues will be focused on for future researches.

The AEs are selected to construct a serving set for the specific user according to its QoS requirement. The serving set may contain one or several AEs. The AE can be equipped with a single antenna or an antenna array. According to different transmission power limitations, the AEs are classified into several types that have different coverage abilities, such as the Macro AE, Micro AE, Pico AE, Femto AE, RRH AE, and etc. Each AE has its own limitations of resource usage and power constraints. The AEs in the serving set can be of different types, which provide more flexible serving set constructions.

Based on the FNA, each user is always focused as being the coverage center of the serving AE set, which means the cell boundary or the traditional cellular structure will no longer exist. The user does not need to handover while moving through the network. Specifically, a user's serving set will be updated adaptively to fulfill its QoS requirement and ensure the user is always staying at the coverage center. Such mechanism provides a typical on-demand user-centric environment with FNA.

Therefore, the AE is actually a new dimension of radio resource, which can be dynamically allocated and scheduled by CPE. Each AE could be viewed as a type of coverage resource with different transmission power constraints. As shown in Fig. 1, Macro AE, Pico AE, Femto AE, Relay AE, and RRH AE are deployed as underlay. The UE-1 is served with a coordinated transmission mode with AEs Macro-1 and Femto-1 as the serving set with UE-1 QoS requirement. The UE-2 is served by AEs Macro-2, Relay-1, and Relay-2. The Femto-2 serves the UE-3 as the only corresponding AE. The AE Relay-2 in the serving set for UE-4 is the common node for the serving set of the UE-2. For the UE-5, the AEs Micro-3 and Pico AE make up its serving set. The coordinated transmission scheme can be a joint processing scheme based on CoMP or an enhanced coordinated transmission scheme with pre-coding techniques. The coverage area for the serving set of each UE will be amorphous due to the adaptive serving set constructing and updating.

In FNA, the resource pooling scheme is proposed to manage the multi-dimensional resources in a centralized manner. Multi-dimensional resource management is processed in the CPE with different optimization goals for different deployment scenarios. A uniform resource allocation strategy can take full advantage of centralized optimization that can improve the resource and energy efficiency, where [9] have already provided some results.

The power limitation of the AE is always the main constraint for the RAN deployments. Due to the centralized controlling capability of CPE, the transmission power of AE can be adjusted dynamically within its radio frequency (RF) power limitation. Then, one Macro AE can be changed to a Pico AE by adaptively decreasing its maximal transmission power. This adjustment can both reduce interferences and cope with the traffic tiding problem. In dense networks, the Pico AEs and Femto AEs are expected to handle more data traffic, while Macro AE mainly handles the large-coverage but low-rate services, such as the voice traffic. In order to fulfill user requirements, the FNA can adjust the type of AE by changing its power output limitations through the CP/UP adaption.

The CP/UP is separated in the FNA. Based on the centralized resource management, the CP and UP are both maintained by the CPE. The CP needs to cover all the deployment area for supporting all possible user access requests and other necessary signaling processes. In order to cover a relatively large area, the CP is usually handled by the Macro AEs with higher power limitations, which are logically named as the Controlling-AEs. The AEs are defined as Data-AEs, when they make up the UP as an on-demand user-centric serving set. The Data-AEs only need to cover the required area for the specific user. The beamforming or pre-coding techniques with multiple AEs can also be implemented for further improving the capacity and decreasing the interferences. The UPs are also constrained and controlled by the CPE, which could be adjusted adaptively for better system performances.

In this paper, we will give the CP/UP adaptation scheme for improving the system EE performance in the downlink scenario of the FNA networks. A three-step EE optimization process is designed as follows.

3 CP construction and adaptation with Voronoi Diagram

In the FNA, the coordinated transmission is managed by one CPE with an arbitrary deployment of AEs within the coverage area. The CP and UP are separated based on the FNA. The designated Controlling-AE implements the function of CPE, which is handling and maintaining the control plane. The Data-AEs maintain their own user plane under the control of Controlling-AEs.

The Data-AEs distributed within the coverage area of a Controlling-AE are supposed to be managed by the Controlling-AE through the CPE. To quantize the master–slave relationship between the Controlling-AE and Data-AE, we focus on a simplified scenario that includes a single Controlling-AE and multiple related Data-AEs. For notational simplicity, the Controlling-AE is denoted as AE *0* while Data-AEs are denoted as *AE i* ($i = 1, \cdots, N$). We denote P_i as the maximum transmission power of the *i*th AE, P_c, P_o as the allocated transmission power for CP and UP of the Controlling-AE, which satisfy the constraints that $p_c \leq p_o$, $p_o \leq p_o$. For the Data-AE, the constraints should be $p_i \leq p_i (i = 1, ..., N)$, where p_i denotes the allocated transmission power of the *i*th Data-AE for the UP. Since the CP transmits the necessary signaling for the UP, additional coverage constraints should be made for the UP. That is, the whole coverage of all the UPs constructed by the Data-AEs should not surpass the coverage of the CP. Then the above constraint can be transformed as the coverage radius of Data-AE:

$$d_i + r_i \leq r_0 \tag{1}$$

where d_i is the distance between the Controlling-AE and *i*th Data-AE. r_i is the coverage radius of the *i*th Data-AE, and r_o is the coverage radius of the Controlling-AE. The coverage radius of the *i*th Data-AE r_i is actually determined by its transmission power p_i, while the coverage radius of the Controlling-AE r_o is determined by its CP transmission power p_c. Then, Equation 1 can be further transformed to a power constraint of *i*th Data-AE:

$$p_i \leq p_c - P(d_i), i \in \{1, \cdots, N\} \tag{2}$$

Where $P(d_i)$ is the power attenuation from the *i*th Data-AE to the Controlling-AE. The above power constraint still guarantees that the coverage of UPs does not surpass the CP coverage, which will be used in the UP construction step of Section 4.

In order to formulate the above constraints, a basic signal propagation model capturing pathloss as well as shadowing is defined as [11]:

$$p_{rx} = K \left(\frac{r}{r_o} \right)^{-\alpha} \varphi P_{tx} \tag{3}$$

Where P_{rx}, P_{tx}, r, and α denote the receiving power, transmission power, propagation distance, and pathloss exponent, respectively. The random variable φ is used to model slow fading effects and commonly follows a lognormal distribution. K is set to the free-space path gain at distance r_o with assumption of omnidirectional antennas. Here, the coverage is defined as the maximum coverage range, which satisfies the UE's minimum required received power P_{min}. The effect of shadowing will be averaged out for the network planning of the CP construction and adaptation. The coverage radius can be expressed as $r_i = r_o \sqrt[\alpha]{KP/P_{min}}$.

In the designed three-step EE optimization for the CP/UP adaptation, the first step aims at constructing a seamless deployment of the CP with minimum transmission power. The Voronoi Diagram, a geometric structure in the computational geometry, divides the space into a number of regions consisting of all the points closer to a specific site than to any other. As energy consumption is proportional to distance, the Voronoi Diagram also defines regions where less energy consumption is required. In order to achieve better EE in the CP construction and adaptation, we construct Voronoi coverage area for the Controlling-AEs. The Data-AEs located within the Voronoi coverage area are controlled by the corresponding Controlling-AE.

As shown in Fig. 2, the deployment of Controlling-AEs creates a Voronoi tessellation of the CP. Assuming that C represents the set of n Controlling-AEs. In 2D Euclidean space, $d_E(c_i, x)$ denotes the Euclidean
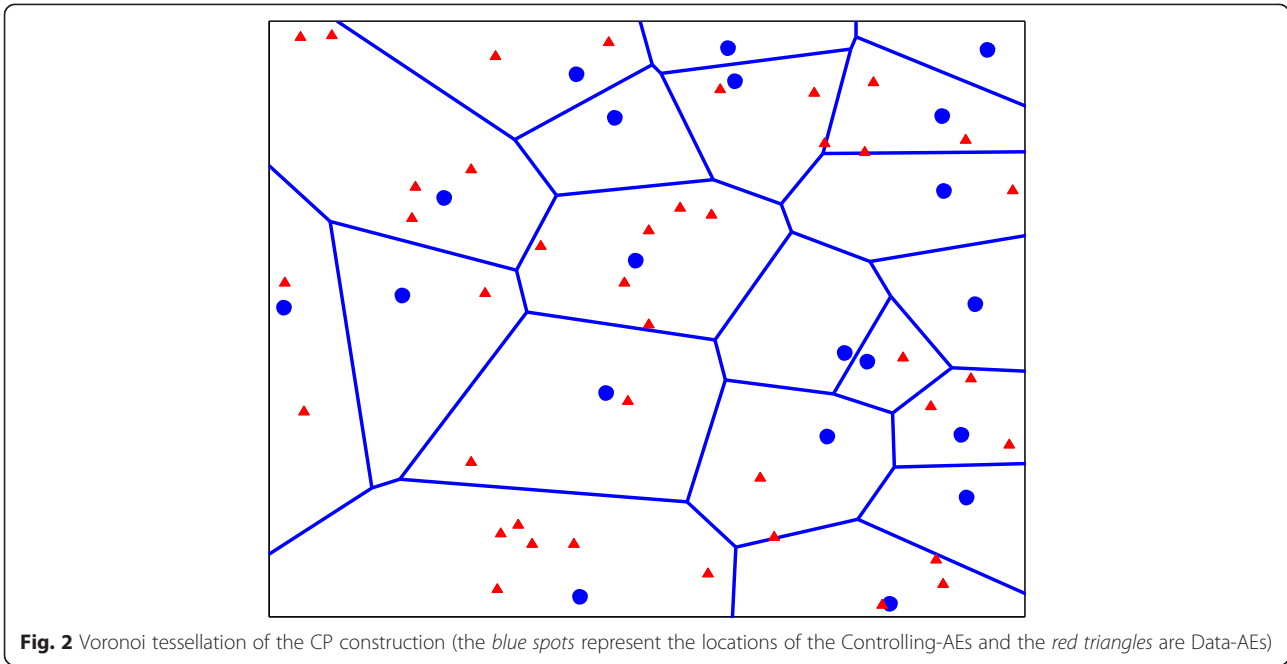
Xu et al. EURASIP Journal on Wireless Communications and Networking (2015) 2015:159

Page 5 of 15

Fig. 2 Voronoi tessellation of the CP construction (the *blue spots* represent the locations of the Controlling-AEs and the *red triangles* are Data-AEs)

distance between the ith Controlling-AE and a position x. Therefore, the Voronoi coverage of the ith Controlling-AE is defined as

$$Vor(c_i) = \left\{ x \in R^2 \big\| \forall j \neq i, d_E(c_i, x) < d_E(c_j, x) \right\} \quad (4)$$

In order to further adapt the transmission power of Controlling-AE with updated AE deployment or coverage area, the Voronoi coverage can be redefined based on the pathloss between the Controlling-AE and point x. Let $\alpha(c_i, x)_t$ be the pathloss between the ith Controlling-AE and the position x at time slot t, the Voronoi coverage will be revised as

$$Vor(c_i)_t = \left\{ x \in R^2 \big\| \forall j \neq i, \alpha(c_i, x)_t < \alpha(c_j, x)_t \right\} \quad (5)$$

Then, the whole CP can be formed into the expression as $\cup_{1 \leq i \leq n} Vor(c_i)$. According to this definition, any position in the Voronoi coverage area is closer to its Voronoi Controlling-AE than any others, which yields less power consumption. As a consequence, the required transmission power for the Controlling-AE is minimized.

The simulation evaluation of the proposed CP construction and adaptation could be found in Section 6.

4 User-centric UP construction with joint AE and subchannel allocation

For the second step of the designed optimization process, the initial deployment of the UP construction for each user should be set up with the available system resources. Based on FNA, AE is released as a new

dimensional radio resource for allocation and scheduling. By jointly allocating the AE and subchannel resources, the on-demand user-centric UP is constructed with the user's QoS requirement. The AEs' transmission powers are allocated equally in this step. Moreover, the transmission power will be further adjusted based on game theory in the third step.

4.1 Joint resource allocation model for the UP construction

Assume that there are N Data-AEs located in the network with two types of AEs, including the Macro AE and Small Cell AE. The bandwidth of each AE is the same and is divided into M subchannels. We set P_i as the maximum transmission power of the ith Data-AE. Meanwhile, K users are randomly distributed in the coverage area of the FNA, including K_1 users with guaranteed bit rate (GBR) service and K_2 users with non-GBR service. Each subchannel of an AE can only be allocated to one user. That is, we do not consider a space division multiple access (SDMA) scenario.

According to the concepts of FNA, in order to reduce the inter-AE interference and improve the capacity performance, users can use coordination techniques to receive signals from more than one AE in the same time slot. Let A_k^m represent the serving set of user k on subchannel m, while $\bar{A}_k^m$ is the complement set of A_k^m. The instantaneous received signal to interference and noise ratio (SINR) on subchannel m for user k is given by:

Xu *et al. EURASIP Journal on Wireless Communications and Networking* (2015) 2015:159

Page 6 of 15

$$\gamma_k^m = \frac{\sum_{i \in A_k^m} p_i \left| h_{i,k}^m \right|^2}{\sum_{j \in \overline{A_k^m}} p_j \left| h_{j,k}^m \right|^2 + n_k^m} \quad (6)$$

where $h_{i,k}^m$ denotes the complex channel response between AE i and user k on subchannel m, n_k^m denotes the additive white Gaussian noise with covariance σ^2 and p_i is the transmission power of AE i to the serving user. The transmission powers for the AEs are allocated to its serving users equally in this step for decreasing the complexity of the multi-dimensional joint resource allocations. The power adjustment for the UP adaptation will be conducted further in the last optimization step.

The data rate allocated to user k on subchannel m can be calculated as [12]:

$$R_k^m = B \log_2 \left(1 + \gamma_k^m \right) \quad (7)$$

The subchannel set allocated to user k is denoted as $S_k \subseteq \{0, 1, 2, \cdots, M\}, \forall\, k$, and we assume that at most N_S subchannels can be assigned to one user. Thus, it is obvious that $|S_k| \le N_s, \forall\, k$, where "$||$" denotes the cardinality of the set. We assume that the serving set for user k on all subchannels in S_k is selected form a common AE set. $A_k \subseteq \{0, 1, 2, \cdots, N + 1\}, \forall\, k$, i.e., $A_k^m \subseteq A_k, \forall m \in S_k, \forall k$. That is, A_k^m may contain all elements or partial elements in A_k. Assume that the maximal number of elements in A_k is N_a, i.e., $|A_k| \le N_a, \forall\, k$.

Hence, the instantaneous data rate for the user k can be given by:

$$R_k = \sum_{m \in S_k} R_k^m \quad (8)$$

In order to quantify the different QoS requirements of users, the utility theory in economics is introduced to describe the characteristics of service by mapping the data rate to the user satisfaction level [13]. According to the user service QoS constraints, the utility functions of the GBR and non-GBR service are verified as the S-shaped function and convex function correspondingly [13, 14]. To satisfy these two types of services simultaneously, the utility function $U(r)$ of all types of service is extracted from [15].

$$U(r) = \frac{E}{A + B e^{-C(r-d)}} + D \quad (9)$$

where r is the data rate allocated to user; R is the total resource of the system; C mainly influences the slope of the curve; A, B, D, E mainly effect the range of utility value; and d is the inflection point of the utility function, which indicates the user requirement of resource. By setting different parameter values, the utility function can present different characteristics,

both the S-shaped function and the convex function. The utility functions of the GBR service $U_{\text{real}}(r)$ and non-GBR service $U_{\text{non-real}}(r)$ are obtained from Equation 9 [15].

The system utility is defined as the linear weighted sum of all users' utility values. In Equation 10, λ represents the priority of GBR service and μ represent the priority of non-GBR service. These two weights are constrained by $\lambda, \mu \in [0, 1]$ and $\lambda + \mu = 1$.

$$U_{\text{system}} = \lambda \sum_{k=1}^{k_1} U_{\text{real}}(r_k) + \mu \sum_{k=K_1+1}^{k_1+k_2} U_{\text{non-real}}(r_k) \quad (10)$$

The system utility can be further extended to include more types of service. Since the utility value represents the satisfaction level of users, the system utility indeed represents all users' satisfaction level, which can provide a better reflection of system performance than throughput.

4.2 Generic algorithm-based centralized resource allocation

In the FNA, the UP construction process is jointly and simultaneously allocating two dimensions of resources (AEs and subchannels) to users with different QoS requirements. Such a multi-dimensional resource allocation problem can be solved by using the resource pooling-based centralized RRM scheme [9]. The above scheme is processed by the CPE to manage all of the available resources uniformly. Since the optimization problem of centralized resource allocation has a large and complex search space, genetic algorithm (GA) is implemented to obtain near-optimal solutions with a relatively fast convergence speed.

Based on GA, chromosome, which is a two-dimensional integer matrix, is used to represent the potential resource allocation solution. Each row of the matrix represents the resource allocation strategy for the specific user. Moreover, each row can be further divided into several parts. Each part lists the allocated elements of a particular dimension of resource. In particular, the chromosome G in the following GA process is given by:

$$G = \begin{cases} a_{1,1}, a_{1,2}, \cdots, a_{1,N_a}; b_{1,1}, b_{1,2}, \cdots, b_{1,N_s}; \\ \quad\quad \cdots\cdots \\ a_{K_1,1}, a_{K_1,2}, \cdots, a_{K_1,N_a}; b_{K_1,1}, b_{K_1,2}, \cdots, b_{K_1,N_s}; \\ a_{K_1+1,1}, a_{K_1+1,2}, \cdots, a_{K_1+1,N_a}; b_{K_1+1,1}, b_{K_1+1,2}, \cdots, b_{K_1+1,N_s}; \\ \quad\quad \cdots\cdots \\ a_{K_1+K_2,1}, a_{K_1+K_2,2}, \cdots, a_{K_1+K_2,N_a}; b_{K_1,K_2,1}, b_{K_1,K_2,2}, \cdots, b_{K_1+K_2,N_s}; \end{cases}$$

$$(11)$$

where the first K_1 rows represent the resource allocation strategies of the K_1 users with GBR service, and the remaining K_2 rows represent the resource allocation strategies of the users with the non-GBR services. Each row is further divided into two parts. The first part

Xu *et al. EURASIP Journal on Wireless Communications and Networking* (2015) 2015:159

Page 7 of 15

containing N_a integers indicates the allocated AEs, while the second part containing N_s integers lists the allocated subchannels. The initial population which includes N_p chromosomes is generated by a random process.

Fitness function is constructed by the system utility function, which is the sum of all users' utility values. It can be used to evaluate the chromosomes. The larger the fitness value is, the better the solution is. Thus, the optimized objective is to maximize the fitness value that is to maximize the system utility. Since the utility value represents the satisfaction level of user, the proposed algorithm tends to meet the requirements of two types of services simultaneously under the three constraints in the user-centric UP constructions process. Specifically, we assume that at most N_a out of N AEs and N_s out of M subchannels can be allocated to user k. In addition to AE and subchannel limitations, we also apply the constraint derived from Equation 2 into the power limitation where we choose the minimum value between the two power constraints P_i (maximum transmission power limitation) and p_c-$P(d_i)$. By using such constraint, we are able to guarantee that the coverage of UPs will not surpass the CP coverage required in Section 3.

$$\max F = \max U_{\text{system}}$$
$$= \max \left[\lambda \sum_{k=1}^{K_1} U_{\text{real}}(r_k) + \mu \sum_{k=K_1+1}^{K_1+K_2} U_{\text{non-real}}(r_k) \right]$$
(12)

s.t.

$$|A_k| \le N_a, \ A_k \subseteq \{0, 1, 2, \cdots, N\}, \forall k$$
$$|S_k| \le N_s, \ S_k \subseteq \{0, 1, 2, \cdots, M\}, \forall k$$
$$p_i \le \min \{P_i, P_c - P(d_i)\}, i \in \{1, 2, \cdots, N\}$$

The chromosomes will be passed to the next generation through a four-step breeding process including Selection, Crossover, Mutation, and Modification.

First of all, a pair of parent individuals is selected based on the so-called "Roulette Wheel Selection", such that the higher the fitness, the greater the opportunity for the individual to be selected. The possibility of chromosome G_i being selected is:

$$p(G_i) = \frac{F(G_i)}{\sum_{k=1}^{N_p} F(G_k)}$$
(13)

Where $F(G_i)$ is the value of the fitness function of chromosome G_i. Note that the selected chromosomes are still in the population, and as a result, it is entirely likely that a chromosome is selected more than once.

Then, two children are generated by combining their parent's genes. In particular, a crossover point is first chosen randomly at a certain column of the two given chromosomes. Next, in order to form the first offspring, all the row

vectors before the crossover point of the first matrix will combine with the row vectors after the crossover point of the second matrix. The second offspring is generated in the opposite way. Herein, the crossover process is illustrated.

Assume two selected parent individuals A and B as follows:

$$A = \begin{cases} a_{1,1}, a_{1,2}, \cdots, a_{1,N_a}; b_{1,1}, b_{1,2}, \cdots, b_{1,N_s} \\ \cdots\cdots \\ a_{K_1,1}, a_{K_1,2}, \cdots, a_{K_1,N_a}; b_{K_1,1}, b_{K_1,2}, \cdots, b_{K_1,N_s}; \\ a_{K_1+1,1}, a_{K_1+1,2}, \cdots, a_{K_1+1,N_a}; b_{K_1+1,1}; b_{K_1+1,2}, \cdots, b_{K_1+1,N_s}; \\ \cdots\cdots \\ a_{K_1+K_2}, a_{K_1+K_2,2}, \cdots, a_{K_1+K_2,N_a}; b_{K_1+K_2,1}, b_{K_1+K_2,2}, \cdots, b_{K_1+K_2,N_s}; \end{cases}$$
(14)

$$B = \begin{cases} a'_{1,1}, a'_{1,2}, \cdots, a'_{1,N_a}; b'_{1,1}, b'_{1,2}, \cdots, b'_{1,N_s}; \\ \cdots\cdots \\ a'_{K_1,1}, a'_{K_1,2}, \cdots, a'_{K_1,N_a}; b'_{K_1,1}, b'_{K_1,2}, \cdots, b'_{K_1,N_s}; \\ a'_{K_1,1}, a'_{K_1,2}, \cdots, a'_{K_1,N_a}; b'_{K_1+1,1}, b'_{K_1+1,2}, \cdots, b'_{K_1+1,N_s}; \\ \cdots\cdots \\ a'_{K_1+K_2,1}, a'_{K_1+K_2,2}, \cdots, a'_{K_1+K_2,N_a}; b'_{K_1+K_2,1}, b'_{K_1+K_2,2}, \cdots, b'_{K_1+K_2,N_s}; \end{cases}$$
(15)

Suppose that the crossover point is located between $b_{1,1}$ and $b_{1,2}$. Then the two children chromosomes C and D can be expressed as:

$$C = \begin{cases} a_{1,1}, a_{1,2}, \cdots, a_{1,N_a}; b_{1,1}, b'_{1,2}, \cdots, b'_{1,N_s}; \\ \cdots\cdots \\ a_{K_1+1}, a_{K_1,2}, \cdots, a_{K_1,N_a}; b_{K_1,1}, b'_{K_1,2}, \cdots, b'_{K_1,N_s}; \\ a_{K_1+1,1}, a_{K_1+1,2}, \cdots, a_{K_1+1,N_a}; b_{K_1+1,1}, b'_{K_1+1,2}, \cdots, b'_{K_1+1,N_s}; \\ \cdots\cdots \\ a_{K_1+K_2,1}, a_{K_1+K_2,2} \cdots, a_{K_1+K_2,N_a}; b_{K_1+K_2,1}, b'_{K_1+K_2,2}, \cdots, b'_{K_1+K_2,N_s}; \end{cases}$$
(16)

$$D = \begin{cases} a'_{1,1}, a'_{1,2}, \cdots, a'_{1,N_a}; b'_{1,1}, b_{1,2}, \cdots, b_{1,N_s}; \\ \cdots\cdots \\ a'_{K_1,1}, a'_{K_1,2}, \cdots, a'_{K_1,N_a}; b'_{K_1,1}, b_{K_1,2}, \cdots, b_{K_1,N_s}; \\ a'_{K_1+1,1}, a'_{K_1+1,2}, \cdots, a'_{K_1,N_a}; b'_{K_1,1}, b_{K_1,2}, \cdots, b_{K_1,N_s}; \\ \cdots\cdots \\ a'_{K_1+K_2,1}, a'_{K_1+K_2,2}, \cdots, a'_{K_1+K_2,N_a}; b'_{K_1+K_2,1}, b_{K_1+K_2,2}, \cdots, b_{K_1+K_2,N_s}; \end{cases}$$
(17)

In this way, the offsprings are expected to provide better chromosomes with their parents' partial characteristics.

After the above processes, all of the children will go through the mutation operation to avoid converging to a local optimized solution. Since the individuals generated by crossover and mutation may no longer satisfy system constraints, some modification should be made.

Finally, in order to prevent good solutions from being lost in the breeding process and to ensure the convergence of the algorithm, we take the two best solutions (named "Elites") from the parent generation and direct them into the child generation. Meanwhile, all other parents will be replaced by the offspring generation. The new generation will replace the original generation and the above procedures are repeated for a total of N_g

Xu *et al. EURASIP Journal on Wireless Communications and Networking* (2015) 2015:159

Page 8 of 15

times. The whole population will evolve from generation to generation and gradually converge to the optimized solution. When the algorithm is terminated, the centralized resource allocation solution is based on the best individual G_{best} t among the current population.

Using the joint resource allocation strategy of AEs and subchannels in the G_{best}, the UP is constructed with equal transmission power allocation mentioned before, where the transmission power of each serving set Data-AE is P_i, $i \in \{1, 2, \cdots, N\}$. After that, based on the UP construction parameters, the UP adaptation via a power adjustment scheme can be implemented to further optimize the EE performance.

5 UP adaptation with game theory-based power adjustment

Based on above CP and UP construction, game theory could be further implemented as a power adjustment strategy for the UP adaptation in order to maximize the system EE. The non-cooperative game model with the pricing function is used to achieve the optimized system EE for the Data-AEs. Here, penalty is defined as the excessive power consumption from Macro Data-AEs, which yields severe interferences. The existence and uniqueness of the Nash Equilibrium for the proposed game model will also be verified.

5.1 Game theory model for the UP adaptation

As we assumed above, the FNA network consists of N_1 Macro Data-AEs and N_2 Small Cell Data-AEs. One Data-AE is able to serve several users within its coverage area. It is assumed that there is only one scheduled active user in each serving set during each signaling slot. Let $k \in \{1, 2, \cdots, N\}$ denote the scheduled user k. According to the results of G_{best} and p_i, $i \in \{1, 2, \cdots, N\}$ derived in the UP construction step, the received SINR on subchannel m for the scheduled user k can be expressed as:

$$\gamma_k^m = \frac{\sum_{i \in A_k^m} p_i \left| h_{i,k}^m \right|^2}{\sum_{j \in \overline{A_k^m}} p_j \left| h_{j,k}^m \right|^2 + n_k^m} \quad (18)$$

where $\gamma_k = \gamma_k^m$ and $p_k = \sum_{i \in A_k^m} p_i$ are the aggregate SINR and transmission power of the scheduled user k, respectively. Considering the user k's QoS requirement, its received SINR has the constraint as $\gamma_k \geq \gamma_k^{threshold}$. But the threshold $\gamma_k^{threshold}$ for GBR service and non-GBR service are different.

5.2 UP adaption based on game theory

The UP Adaption focuses on the transmission power of serving set Data-AEs. Based on the game theory, we propose a power adjustment scheme for the system EE

optimization from the perspective of Data-AEs. Each Data-AE within the serving set will choose a reasonable transmission power to maximize its own utility of EE performance, which is a typical non-cooperative N-player game problem. Let $G = [N, \{p_k\}, \{u_k(p_k, \gamma_k | \mathbf{P}_{-k})\}]$ denote the Non-cooperative Power adjustment Game with Pricing (NPGP), where

- $N = \{1, 2, \ldots, N\}$ is the index of the serving set Data-AE.
- $\{p_k\} = \{p_k | p_k \in [0, p_{max}]\}$ are the transmission power of serving set Data-AEs for user k and $p_{max} > 0$ is the maximum power constraint of the corresponding Data-AEs.

Let $\{u_k(p_k, \gamma_k | \mathbf{p}_{-k})$ denote the utility of the scheduled user k, where γ_k and p_k is the aggregated SINR and transmission power of the user k respectively. $\mathbf{p}_{-k}$ is the vector of transmission power of all serving set Data-AEs other than the serving set Data-AEs for user k.

Considering the energy efficiency measured in bit/Joule [16], the utility function of the serving set Data-AEs for user k is defined as:

$$u_k(p_k, \gamma_k | \mathbf{P}_{-k}) = a_k \frac{f(\gamma_k)}{p_k} - b_k p_k \quad (19)$$

where $a_k \frac{f(\gamma_k)}{p_k}$ denotes the EE of the user k and $b_k p_k$ denotes the linear pricing for the user k. Both the a_k and b_k are positive factors.

$f(\gamma_k)$ is defined as:

$$f(\gamma_k) = 1 - e^{-\frac{\gamma_k}{2}} \quad (20)$$

which means that the revenue of the user k will increase slowly as γ_k increases [17].

The pricing $b_k p_k$ will ensure that the serving set Data-AEs for the user k could be penalized when they cause serious interference to other users with more transmission powers.

Finally, the EE optimization problem for the UP adaption is formulated as the game model:

$$\max_{0 \leq p_k \leq pmax} u_k(p_k, \gamma_k | \mathbf{p}_{-k}), \forall k = 1, 2, \ldots, N \quad (21)$$

5.3 Nash Equilibrium for the power adjustment game

The Nash Equilibrium is a steady state that offers a predictable outcome of a game, where Data-AEs compete with selfish actions through self-optimization and converge to a point that no Data-AEs wish to deviate unilaterally. For the proposed game model (21), the Nash Equilibrium is defined as:

Definition 1:

Xu *et al. EURASIP Journal on Wireless Communications and Networking* (2015) 2015:159

Page 9 of 15

Suppose $p_k^*, \forall k = 1, 2, ..., N$ is a solution for (21). Hence, the point $\mathbf{p}^*$ is a Nash Equilibrium for the proposed non-cooperative game if for any $\mathbf{p}$, the following conditions are satisfied:

$$u_k(p_k^*, \gamma_k^* | \mathbf{p}_{-k}^*) \geq u_k(p_k, \gamma_k | \mathbf{p}_{-k}), \forall k = 1, 2, ..., N \tag{22}$$

The proposed game model with pricing in this paper is a supermodular game [18]. The existence and uniqueness of the Nash Equilibrium will be verified as follows.

(1) Existence of Nash Equilibrium

Theorem 1:

The set of Nash Equilibria of a supermodular game is nonempty. Furthermore, the Nash set has a largest element and a smallest element.

A proof of the theorem can be found in [19]. Let E denote the set of Nash Equilibria. Let $\mathbf{p}_S$ and $\mathbf{p}_L$ denote the smallest and the largest elements of E respectively. The theorem states that all the equilibria $\mathbf{p} \in E$ are located such that $\mathbf{p}_S < \mathbf{p} < \mathbf{p}_L$.

We introduce a totally asynchronous algorithm that generates a sequence of powers that converges to the smallest Nash Equilibrium $\mathbf{p}_S$. Suppose that the serving set Data-AEs for user k update their power at time instances given by the set $T_k = \{t_{k1}, t_{k2}, t_{k3}, ...\}$ where $t_{kl} < t_{k(l+1)}$ and for $t_{k0} = 0$ all k. Define $T = \{\tau_1, \tau_2, \tau_3, ...\}$ as the set of update instances $T_1 \cup T_2 \cup ... \cup T_N$ sorted in increasing order. Assume that there are no two time instances in set T that are exactly the same. The algorithm for finding the Nash Equilibrium is designed as follows.

Algorithm 1:

Let us consider the proposed non-cooperative power adjustment game with pricing as given in (21). We will generate a sequence of the transmission powers:

1) Set the initial power vector at time $t = 0$: $\mathbf{p} = \mathbf{p}(0)$. Also let $l = 1$.
2) For all l such that $\tau_l \in T$
   For all serving set Data-AEs for the user k such that $\tau_l \in T_k$
   i. Given $\mathbf{p}(\tau_{l-1})$, compute $p_k^* = \arg\max u_k(p_k, y_k | \mathbf{p}_{-k})$
   ii. If $y_k^* \geq y_k^{threshold}$
   Then $p_k(t_l) = \min(p_k^*, p_{max})$.
   else
   Remove user k in this iteration and continue the algorithm in next iteration [20].
   end For.
   end For.

Theorem 2:

The proposed Algorithm 1 converges to a Nash Equilibrium of the NPGP. Furthermore, it is the smallest equilibrium $\mathbf{p}_S$, in the set of Nash Equilibria.

The proof can be found in [17], which implies that the Nash Equilibrium in the proposed NPGP exists and can be reached from either the top or the bottom of the strategy space with Algorithm 1. Since we do not know if there is a unique equilibrium, we compare the equilibrium in the Nash set E to determine if there exists a single equilibrium that dominates all other equilibria. Indeed, we can show that $\mathbf{p}_S$ is the best equilibrium in the set E.

(2) Uniqueness of Nash Equilibrium

Theorem 3:

If $\boldsymbol{x}, \boldsymbol{y} \in E$ are two Nash Equilibria in NPGP and $\boldsymbol{x} \geq \boldsymbol{y}$, then $u_k(\boldsymbol{x}) \leq u_k(\boldsymbol{y})$ for all k.

Proof:

Notice that, for fixed p_k, the utility $u_k = a_k \frac{f(\gamma_k)}{p_k} - b_k p_k$ decreases with increasing $\mathbf{p}_{-k}$ for all k. Therefore, since $\boldsymbol{x}_{-k} \geq \boldsymbol{y}_{-k}$, we have:

$$u_k(x_k, \boldsymbol{x}_{-k}) \leq (x_k, \boldsymbol{y}_{-k}) \tag{23}$$

Also, by definition of Nash Equilibrium and since $\boldsymbol{y}$ is a Nash Equilibrium of NPGP, we have:

$$u_k(x_k, \boldsymbol{y}_{-k}) \leq u_k(y_k, \boldsymbol{y}_{-k}) \tag{24}$$

By the above equations,

$$u_k(\boldsymbol{x}) \leq u_k(\boldsymbol{y}) \tag{25}$$

According to the Theorem 3, we know that smaller Nash Equilibrium leads to higher utilities for all users. Since $\mathbf{p}_S < \mathbf{p}$ for all $\mathbf{p} \in E$, we conclude that for all $\mathbf{p} \in E$,

$$u_k(\mathbf{p}_s) \geq u_k(\mathbf{p}) \text{ for all } k \tag{26}$$

This result implies that, in case the NPGP has Nash Equilibria, the one that yields highest utilities is the Nash Equilibrium with the minimum total transmission powers.

In conclusion, the existence and the uniqueness of the Nash Equilibrium for the proposed NPGP have been proved. It means that the transmission power of the serving set Data-AEs at the Nash Equilibrium are regarded as a reasonable solution for the EE improvement of the UP adaption scheme.

6 Simulation evaluations

In this section, the system-level simulations are conducted to evaluate the proposed CP/UP adaptation scheme. The simulation environment is introduced at first with the simulation parameters. Then, the EE performance of the Voronoi Diagram-based CP construction and adaptation is given. Based on the CP construction, the simulation results of the user-centric UP construction with joint AE

Xu *et al. EURASIP Journal on Wireless Communications and Networking* (2015) 2015:159

Page 10 of 15

and subchannel allocation are provided. After obtaining the UP construction results, the non-cooperative game model-based power adjustment for the UP adaptation are evaluated with the EE performances. This is the last step for the whole CP/UP adaptation solutions on the FNA featured networks.

6.1 Simulation environment

Based on the deployment scenario of the FNA, the heterogeneous coverage environment is set up, which contains different types of AEs connected with one centralized CPE. There are 9 Macro AEs and 72 Small Cell AEs generated in a 2 × 2-km-square coverage area. In the coverage of each Macro AE, Small Cell AEs are randomly distributed as several clusters. The CPE constructs and maintains the CP by choosing Controlling-AEs to guarantee the signaling requirements in the coverage area. In addition, the CPE also constructs and updates the user-centric UP by allocating Data-AEs to make up the coordinated serving set to fulfill the user QoS requirements.

The simulations use parameters that are based on the guidelines for 3GPP channel model [21] and that are carried out through the Matlab platform. There are total 20 subchannels that can be fully reused by all the AEs. The channel conditions take into account pathloss, shadow fading, and Rayleigh fading. The maximum number of AEs in the coordinated serving set is limited to 3. Furthermore, in order to guarantee fairness among users, the maximum number of subchannels that can be allocated to each user is also limited to 3. More detailed simulation settings including the system and GA parameters for the joint resource allocation are listed in Table 1.

6.2 Simulation results and analyses

(1) The CP construction and adaptation with Voronoi Diagram.

According to the proposed Voronoi Diagram-based CP construction and adaptation scheme, the Controlling-AEs are selected to construct the CP for serving the whole coverage area. That is, the selected Controlling-AEs need to cover the farthest Voronoi vertices of the coverage area, while the transmission power is adapted with the actual network deployments.

Traditionally, the CP is constructed based on network planning, which is mainly uniform deployed within the coverage area. Furthermore, the transmission power of CP is set as the maximum transmission power of the BS. In this paper, the EE performance of the proposed CP construction and adaptation scheme is compared with the traditional CP construction scheme as the baseline. The simulation results are collected in the generated coverage area with choosing different numbers of Controlling-AEs.

Table 1 Simulation setting

System parameter	
Number of Macro AEs	9
Number of Small Cell AEs	72 (each Macro AE has two clusters with four Small AEs)
Number of subchannels	20
Maximum power of Macro AE	46 dBm
Maximum power of Small Cell AE	30 dBm
Carrier frequency	2 GHz
Bandwidth	10 MHz
Pathloss model	$PL = 128.1 + 37.6 \log_{10}d$, $d(km)$
Shadowing standard deviation	8 dB
Shadowing correlation distance	50 m
Fast fading	Rayleigh fading
Noise density	−174 dBm/Hz
Maximal size of the serving set N_a	3
Maximal size of the subchannel set N_s	3
GA parameters	
Population size N_p	500
Number of generations N_g	200
Mutation rate p_m	0.001

Figure 3 shows the EE performance versus different numbers of Controlling-AE selected for the CP construction. As expected, the Voronoi Diagram-based CP construction and adaptation scheme outperforms the traditional CP construction scheme. As the number of Controlling-AEs in the generated coverage area increases, the EE performances of both schemes experience a decline. This is because the decrease in the inter-site distance between Controlling-AEs actually causes unnecessary energy consumption. For the coverage area generated in the simulation, less Controlling-AE are enough to handle all the signaling coverage.

(2) The user-centric UP construction with GA-based joint AE and subchannel allocation.

The proposed GA-based resource allocation scheme is abbreviated to "GA" in the following simulation results. The performance metrics of the system utility, user access succeed rate, and system throughput are evaluated. In detail, the system utility is the sum of all users' utility values, which reflects the satisfaction level of users. The access succeed rate is defined as the proportion that GBR users' QoS requirement is satisfied. That is, the ratio between the number of GBR users (whose required data rate are reached) and the total number of GBR users. The number of users is increasing from 200 to 800, while the number of users with GBR service and non-GBR service remains the same in this step.
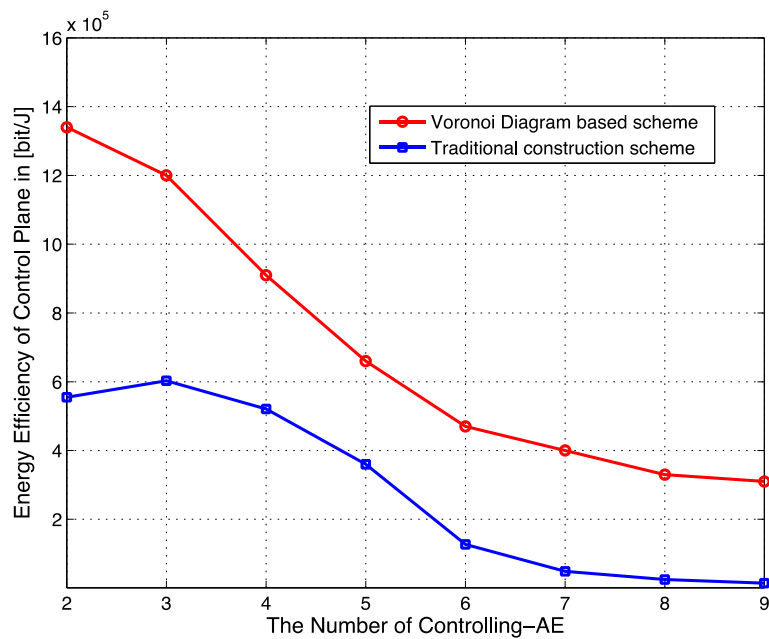
Fig. 3 Energy efficiency of the CP construction and adaptation

During the simulations, two algorithms are used as the comparisons. One is the pathloss-based AE selection for the user-centric serving set construction and randomized subchannel set allocation (abbreviate to "PL-Random"). Each user will select N_a best AEs with the lowest pathloss between the AEs and the user. Then the user will be randomly allocated the common available subchannels of all the AEs in the serving set.

The other compared algorithm is the pathloss-based user-centric serving set construction and maximum SINR subchannel set allocation (abbreviate to "PL-Max-SINR"). Each user will select N_a best AEs with the lowest pathloss between the AEs and the user. Then the user will be allocated the common available subchannels of the selected AEs, which tend to provide the highest SINR.

The simulation results of the system utility, the user access succeed rate, and the system throughput are plotted as Figs. 4, 5 and 6, respectively.

As shown in Fig. 4, the performance of the proposed UP construction scheme is shown in terms of the system utility value versus the number of users. It is observed that the GA-based scheme achieves the highest system utility value, PL-MaxSINR ranks second and PL-Random has the lowest performance. Note that the performance gap between GA and other algorithms becomes larger as the number of users increases. This is mainly because when the resources are not sufficient, the optimized resource management can allocate limited resources more effectively based on different user QoS requirements, which improves the system utility further.

For GBR users, the satisfaction level can be measured by access succeed rate, the larger the better. For instance, the access succeed rate 0.8 means that 80 % of GBR users' prescribed QoS requirements are fulfilled. As shown in Fig. 5, the proposed GA-based algorithm can achieve larger access succeed rate than the other solutions (PL-MaxSINR, PL-Random), which means the GA-based scheme tends to guarantee that a higher number of GBR users will meet the required data rate. This is because the S-shaped utility function curve of the GBR service rises rapidly at the required bit rate, and allocating enough resources to the GBR users will have more contribution to the overall system utility, compared to allocating more resources to the non-real-time services.

The system throughput performances of these three algorithms against the number of users were also simulated. In Fig. 6, we can observe that the PL-MaxSINR algorithm, which pursues maximal throughput, ranks first as expected. The GA-based scheme ranks second and also achieves relative high system throughput. This confirms that the GA-based scheme can achieve a better balance between user satisfaction and system throughput. In other words, it can provide better resource utilization, while just slightly worse system throughput.

Since the computational complexity of the solution to NP-Hard problems is an important factor for evaluating its performance, the complexities of the proposed GA-based algorithm and comparison schemes will be analyzed as follows.
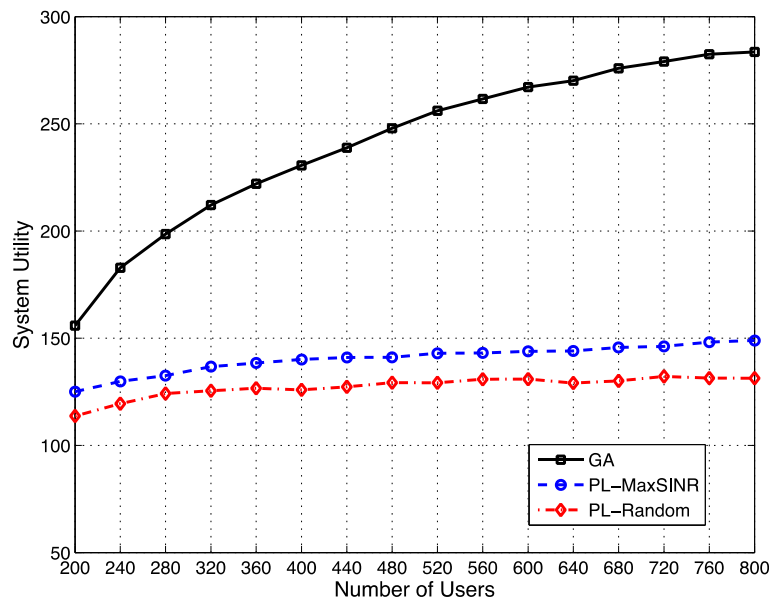
Xu et al. EURASIP Journal on Wireless Communications and Networking (2015) 2015:159

Page 12 of 15

Fig. 4 System utility of the user-centric UP construction

The heuristics solution, GA, is implemented here for handling the optimization problem of the two-dimensional resource allocation, which is proved as a NP-Hard problem. The merits of the GA are finding the near-optimal solution with lower computational complexity and fast converging speed. For the NP-Hard problem, the exhausted search (ES) is a way to solve the problem by searching for the optimal solution.

As assumed before, there are total N AEs and the bandwidth of each AE is divided into M subchannels,
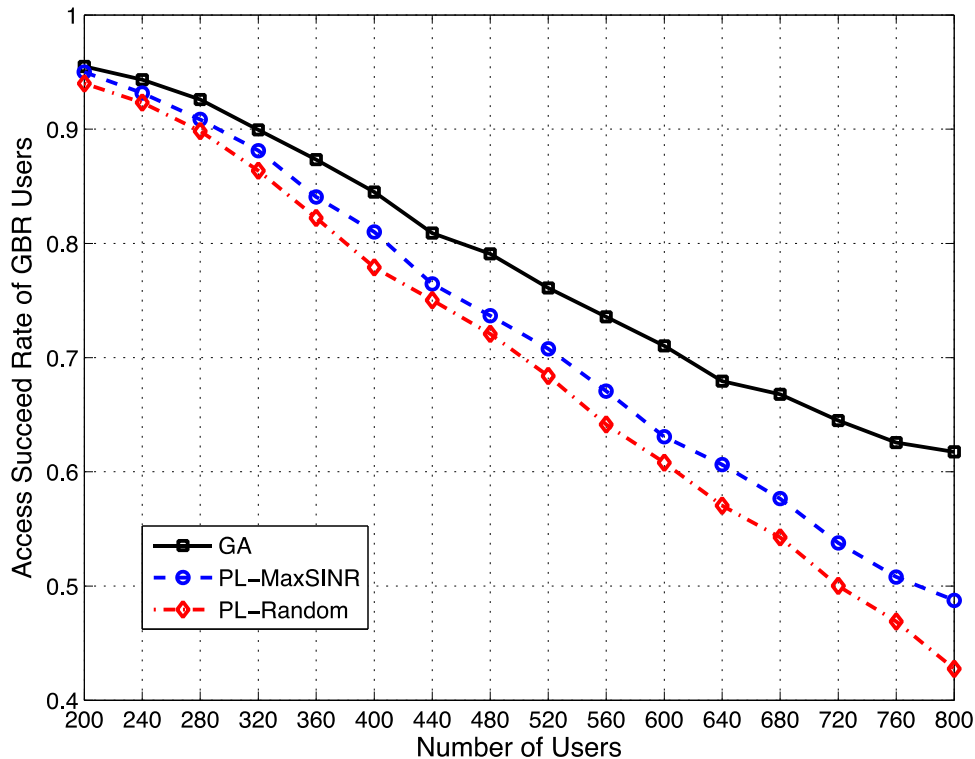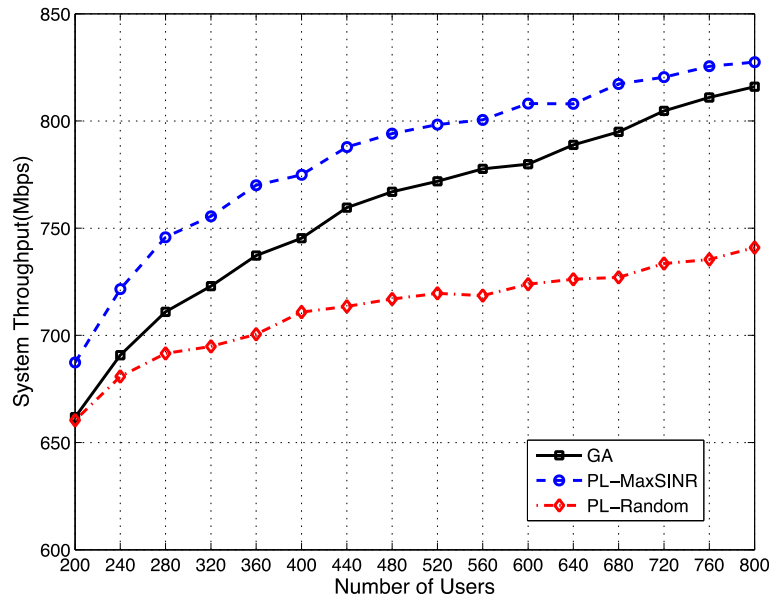

Fig. 5 Access succeed rate of GBR users

Xu *et al. EURASIP Journal on Wireless Communications and Networking* (2015) 2015:159

Page 13 of 15

Fig. 6 System throughput of the user-centric UP construction

the computational complexity of the K user ES will be $\left(KN^{KN_a}M^{KN_s}N_aN_s\right)$, where N_a denotes the maximal AE number in a serving set and N_s denotes the maximal subchannel number in a subchannel set. We can infer from the above equation that ES is too complex to be implemented in our evaluation scenario.

For the GA-based scheme, the computational complexity mainly comes from the fitness evaluation process, which can be measured with the number of SINR calculations. For other comparison schemes that we considered, namely the PL-Random and PL-MaxSINR, their computational complexities are measured with the number of SINR calculations in the resource allocation and criteria measurement procedure.

The following Table 2 lists the computational complexity of four different schemes, where the complexity is calculated for the worst case. In GA, N_p and N_g denote the population size and the iterations, respectively. As we can see, the complexity of PL-Random is the lowest. Using the order of magnitude provided in the system parameter (Table 1), it can be concluded that GA has an equal complexity to PL-MaxSINR but with superior performances as shown above. Thus, compared to the optimal ES, the complexity of the GA is much lower, while the performance is near-optimal.

(3) The UP adaptation with game theory-based power adjustment.

Based on the CP and UP construction results in the above steps, the user-centric UP will be further adapted by the proposed non-cooperative game model-based power adjustment. As assumed before, there are total K users randomly distributed in the system, including K_1 GBR service users and K_2 non-GBR service users. Let $\eta = K_1/K_2$ denote the ratio of users with GBR service to users with non-GBR service. The comparison algorithms include the equal transmission power allocation (EPA) derived in the second step with GA-based UP construction and another algorithm named as "efficient power control (EPC) scheme" [22]. The EPA scheme is compared as the performance baseline. The EPC scheme is proposed as the power control mechanism for the Small Cell HetNet scenarios. The EPC scheme also employs the non-cooperative game model, but the utility and pricing function design are different than those used in our proposed scheme. The system EE performance and average transmission power of the Data-AEs at the Nash Equilibrium solution (abbreviate to "NPGP" scheme) versus the ratio η are evaluated respectively with Figs. 7 and 8.

As shown in the Fig. 7, we can observe that the average achieved system EE performance of NPGP is better

Table 2 Computational complexity

Scheme	ES	GA	PL-MaxSINR	PL-Random
Complexity	$\mathcal{O}\left(KN^{KN_a}M^{KN_s}N_aN_s\right)$	$\mathcal{O}\left(KN_gN_pN_aN_s\right)$	$\mathcal{O}\left(KM^{N_s} + KN_aN_s\right)$	$\mathcal{O}(KN_aN_s)$

Moreover, we can achieve lower complexity by further improving the proposed GA algorithm. For example, by adopting the PL-Random algorithm in the initial process to produce the first population of GA, the number of generations can be further reduced
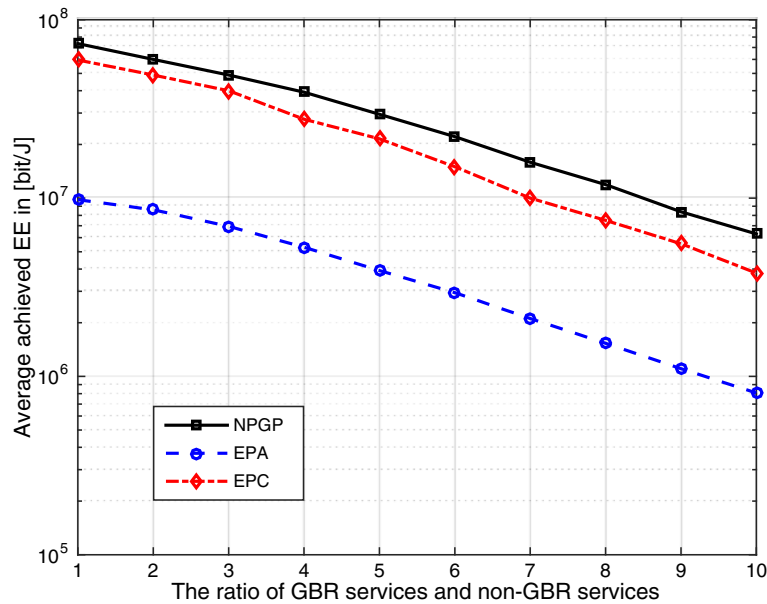
Xu *et al. EURASIP Journal on Wireless Communications and Networking* (2015) 2015:159

Page 14 of 15

Fig. 7 Average achieved system EE at the Nash Equilibrium

than that for the EPA and EPC schemes. The system EE at the Nash Equilibrium solution of the proposed NPGP is much better than that of EPA at each value of η. The utility function and pricing function design with the NPGP game model also show the gain of the EE performance over the EPC scheme.

Figure 8 compares serving set Data-AEs' average transmission power among the Nash Equilibrium solution of the NPGP, EPA, and EPC versus the user ratio η. As

observed from the simulations results, the serving set Small Cell Data-AEs' average transmission powers at Nash Equilibrium are much lower than that of the EPA and EPC with η increasing, which means that the transmission power of the Small Cell AEs are saved with the proposed NPGP scheme. For the average transmission power of Macro Data-AEs in the serving set, the NPGP at the Nash Equilibrium solution has larger transmission power requirements than that of the EPA and EPC
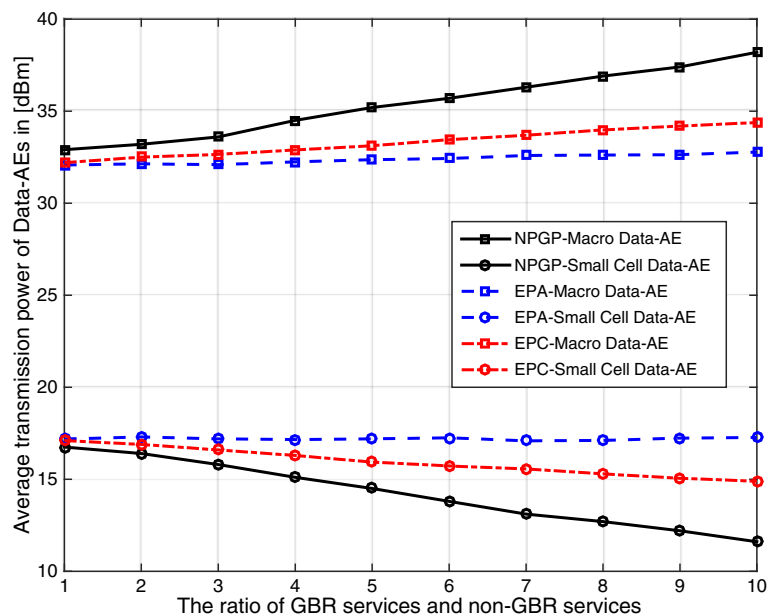

Fig. 8 Average power of the serving set Data-AEs

Xu *et al. EURASIP Journal on Wireless Communications and Networking* (2015) 2015:159

Page 15 of 15

scheme. This is due to the fact that the Macro Data-AE needs to cover the Macro cell edge users to guarantee their QoS requirements, especially with the increasing of GBR services. The users are randomly distributed in the whole coverage area, which means there are several GBR users located in the Macro coverage edge. The larger transmission power limitation lets the Macro Data-AE have the ability to handle the GBR requirements even at the edge of coverage. There is only about 7 % average increase of the Macro Data-AE transmission power compared with the NPGP scheme, which is still under the power limitation of the Macro Data-AE. Furthermore, since the number of Small Cell Data-AEs is much more than the Macro Data-AEs in the system, it is obvious that the proposed NPGP-based UP adaptation solution is helpful to save the transmission power and improve the system EE.

7 Conclusions

Based on the frameless network architecture, we propose a control plane/user plane adaptation strategy for system energy efficiency improvement. The constraints for the CP/UP separation-based adaptation are presented and the system EE performance optimization problem is designed as a three-step scheme. Firstly, the CP is constructed with the Voronoi Diagram to achieve better system EE. Secondly, we constructed the user-centric UP for each specific user and considered user-centric serving set construction by jointly allocating resources. The AEs and subchannels are jointly allocated by the GA algorithm with equal power allocations for different QoS services, while the power adjustments are further handled by the third step with game theory solutions for the system EE optimization. The non-cooperative power adjustment game model with pricing is given and the corresponding Nash Equilibrium is solved with the verifications of its existence and uniqueness. Finally, the performance evaluations are conducted with the system-level simulations. According to the simulation results, the CP/UP adaptation-based system EE performances are improved. Moreover, the user-centric UP construction scheme achieves a better user satisfaction level and access succeed user rate. Hence, considering the constraints on signaling and data transmission, the CP/UP adaptation is an innovative method for further improving the system energy efficiency.

Author details
[1]National Engineering Laboratory for Mobile Network Security, Beijing University of Posts and Telecommunications, Beijing 100876, China. [2]Division of Terminal Technology, China Mobile Research Institute, Beijing 100053, China. [3]Intellectual Property Department, Tendyron Corporation, Beijing 100083, China.

References
1. JG Andrews, Seven ways that HetNets are a cellular paradigm shift. Commun. Mag. IEEE **51**(3), 136–144 (2013)
2. L Hanzo, H Haas, S Imre, D O'Brien, M Rupp, L Gyongyosi, Wireless myths, realities, and futures: from 3G/4G to optical and quantum wireless. Proc. IEEE **100**, 1853–1888 (2012)
3. 3GPP, TR 36.839 v11.1.0, Mobility enhancements in heterogeneous networks (3GPP, Sophia-Antipolis, 2012)
4. 3GPP, TR 36.842 v0.4.1, Study on small cell enhancements for E-UTRA and E-UTRAN - higher layer aspects (3GPP, Sophia-Antipolis, 2013)
5. 3GPP, TR 36.814 v11.1.0, Further advancements for E-UTRA physical layer aspects (3GPP, Sophia-Antipolis, 2013)
6. China Mobile Research Institute, C-RAN, The road towards green RAN, White Paper, v3.0 (2013)
7. M Yang, Y Li, D Jin, OpenRAN: a software-defined RAN architecture via virtualization. ACM SIGCOMM **43**(4), 549–550 (2013)
8. X Jin, L E Li, L Vanbever, J Rexford, Softcell: Scalable and flexible cellular core network architecture. ACM conference on Emerging networking experiments and technologies, Santa Barbara, 2013.
9. X Xu, D Wang, X Tao, T Svensson, Resource pooling for frameless network architecture with adaptive resource allocation. Sci. China Inf. Sci **56**(12), 83–94 (2013)
10. X Xu, H Zhang, X Dai, Y Hou, X Tao, P Zhang, SDN based next generation mobile network with service slicing and trials. China Commun. **11**(2), 65–77 (2014)
11. A Goldsmith, *Wireless Communications* (Cambridge University press, Cambridge, 2005)
12. D Wang, X Xu, X Tao, *Joint Scheduling and Resource Allocation Based on Genetic Algorithm for Coordinated Multi-Point Transmission Using Adaptive Modulation* (IEEE PIMRC, Sydney, 2012)
13. C Liu, L Shi, B Liu, *Utility-based Bandwidth Allocation for Triple-Play Services* (ECUMN, Toulouse, 2007)
14. Z Niu, L Wang, X Duan, Utility-based radio re-source optimization for multimedia DS-CDMA systems. ACTA ELECTRONICA SINICA. **32**(010), 1594–1599 (2004)
15. L Chen, W Chen, Utility based resource allocation in wireless networks. Journal of Beijing University of Posts and Telecommunications **3**(6), 58–63 (2010)
16. YS Soh, TQS Quek, M Kountouris, Energy efficient heterogeneous cellular networks. IEEE J Selected Areas Commun. **31**(5), 840–850 (2013)
17. CU Saraydar, NB Mandayam, DJ Goodman, Efficient power control via pricing in wireless data networks. IEEE Trans. Commun. **50**(2), 291–303 (2002)
18. DM Topkis, Equilibrium points in nonzero Sum n-person submodular games. SIAM J. Control Optimization **17**(6), 773–787 (1979)
19. DM Topkis, *Supermodularity and Complementarity* (Princeton Univ Press, Princeton, 1998)
20. M Andersin, Z Rosberg, J Zander, Gradual removals in cellular PCS with constrained power control and noise. Wirel. Netw **2**(1), 27–43 (1996)
21. 3GPP, TR25.996 v.11.0.0, Spatial channel model for multiple input multiple output (MIMO) simulations (3GPP, Sophia-Antipolis, 2012)
22. Y Ma, T Lv, Y Lu, Efficient Power Control in Heterogeneous Femto-Macro Cell Networks, (IEEE WCNC, Shanghai, 2013)