Optimal Load Balancing in Cloud Radio Access Networks

Chen Ran and Shaowei Wang
School of Electronic Science & Engineering
Nanjing University, Nanjing 210023, China
E-mail: wangsw@nju.edu.cn

Chonggang Wang
InterDigital Communications
781 Third Avenue, King of Prussia, PA 19406
Email: cgwang@ieee.org

***Abstract*—Cloud radio access network (CRAN) has been seen as an effective means to address the challenges faced by cellular radio networks, such as high capital expenditure and operating expense, high energy consumption and low spectral efficiency. Especially, CRAN has the potential to equip a cellular network with the load-balancing capability to cope with the non-uniformly distributed traffic in the service area. In this paper, we develop an optimal load balancing scheme for CRAN-based cellular systems by employing an infinite optimization technique. A fairness index is defined to measure the load balancing level of the cellular system and monitored by periodically inspecting the load distribution among all cells. When the fairness index is below a warning threshold, we divide the service zone into compact and connected subregions based on an infinite optimization formulation. Each subregion served by a cell has almost equal area and almost equal throughput requirement. To avoid yielding ill-shaped subregion that is difficult to be covered by a practical cell, a penalty term is introduced to the objective function. Then we update the cell association of each user so that the fairness index return to an acceptable level. Numerical experiments show that our proposal can provide performance-guaranteed load balancing for the cellular network with almost no additional operating expense.**

***Index Terms*—Cloud radio access networks, infinite optimization, load balancing.**

I. Introduction

The proliferation of lightweight hand-held devices, tablets, and other media-hungry devices, together with the significant benefit of anywhere anytime Internet access has spurred the deployment of wireless *hot-spot* networks. A key challenge for managing a cellular network is to use the available resources such as radio spectrum and energy to derive the best return on investment while satisfying users' service demands. In a cellular network, a terminal equipment generally associates itself with a cell that provides the strongest signals, while being ignorant of the load of the cell. As users are typically not evenly distributed in the service area of the cellular network, some cells tend to suffer from heavy load, while their adjacent cells may carry only light load. Such load imbalance among cells is undesirable as it hinders the network from fully utilizing its capacity and consuming energy as little as possible to provide users with high quality services, and degrades the stability of the cellular system.

In order to achieve high resource usage efficiency without significantly changing the infrastructure of the cellular system and the terminals of users, cloud radio access network (CRAN) [1–3] is introduced to address the challenges that mobile service providers are faced with, such as spectrum efficiency and energy reduction. CRAN-based cellular system can carry out baseband signal processing in a centralized manner which can greatly reduce the power consumption of the cellular system. And co-operative radio with distributed antenna equipped by remote radio head (RRH) provides higher spectrum efficiency. Furthermore, base station virtualization technique enables the CRAN-based cellular system the ability of processing aggregation and dynamic resource allocation, which can increase the infrastructure utilization efficiency significantly. In particular, CRAN can address the non-uniformly distributed traffic in cellular systems due to the load-balancing capability in the aggregated baseband unit (BBU) pool. Though the serving RRH changes dynamically according to the movement of users, the serving BBU can still be in the same BBU pool because the coverage of a BBU pool is larger than a traditional base station. However, how to implement the load balancing between the BBUs is not investigated in both academic and industry as far as the authors known.

Load balancing in cell networks has been discussed in the literature extensively. Cell breathing has been proposed [4] as a promising load balancing method and widely investigated in the second and third generation cellular systems. The core of cell breathing is to adjust the coverage area of each cell to cope with the traffic variations with time in the cellular system adaptively. Specifically, the heavily loaded cells shrink their coverage areas, whereas the lightly loaded cells expand their coverage areas to serve the users previously associated with the heavily loaded cells. Essentially, cell breathing is to balance the load of the cellular system so that the whole system can work in a more stabilized and more energy-efficient way. In [5–7], the authors proposed several load balancing schemes by controlling the size of the cell. A greedy algorithm is proposed in [5], which can provide capacity where and when it is needed to reduce the transmission power level of the congested access point (APs) until any of the congested APs reaches the minimal power level. But it could not achieve an overall load balance among multiple APs. Two algorithms are presented in [6], one of which can minimize the load of the most congested AP and the other can produce an optimal min-max load balancing solution. However, the furthest located users may be expelled randomly, making indistinctive load

978-1-4799-8406-0/15/$31.00 ©2015 IEEE

contribution as explained in [7], where a distributed algorithm is developed to allow APs to adjust their coverage range according to its own and the neighboring load conditions. It provides a better balance from the viewpoint of the system. However, it requires that every AP has known a complete list of the load information on its adjacent APs, which brings a lot of signalling overhead and may be impossible to implement in practical wireless systems. In [8], short- and long-term power control policies are presented for the OFDMA-based downlink of a single-cell system. Numerical results show that base station on-off power control is more effective for light-load cellular system while range adaptation becomes more effective for the heavy-load case.

In this paper, we develop a novel load balancing technique for the cellular network, especially for the CRAN-based one due to its cloud infrastructure that can implement centralized mass computation. The main idea of our proposal is as follows: Given a region served by a cellular system and the distribution of users (also referred to as traffic demand points, TDPs) in the region, we monitor the load of all cells periodically. When the load fairness index of the network exceeds a threshold, our proposed load balancing algorithm is triggered to redesign the service areas of all cells in an area- and traffic-balanced way. Based on the repartition results, the cell association of each TDP is updated and the cellular system returns to a relative load balanced state. The key challenge to implement our proposal is how to divide a connected polygonal region into subregions with almost equal areas and almost equal traffic demands, which will be explained in detail in Section III. The effectiveness of our proposed method is verified by experiment results. Notice that our proposed load balancing method are not only dedicated to the cellular system based on which we describe our proposal and derive results in an intuitive manner, it is also suitable for the CRAN-based system. In fact, CRAN infrastructure is the ideal platform to employ our proposal because it can address the potential computation-related challenge caused by the proposed algorithms.

The remainder of this paper is organized as follows. In Section II, we give our load balancing strategy in detail and point out the key technique required to implement our proposal. In Section III, an infinite optimization task is formulated, and efficient algorithms are developed to solve the involved problem. Numerical results are reported with discussions in Section IV. Conclusions are drawn in Section V.

II. Notational Conventions and Proposed Strategy

Consider a given region R made up of n districts $\;_i R_i = R$ is served by n cells $P = \{p_1, p_2, \ldots, p_n\}$, as shown in Fig.1. It is reasonable to assume that R is a connected, polygonal region with non-empty interior for practical cellular system. Without consideration of the offloading effect, we assume that each cell p_i serves a subregion R_i, and each subregion should not overlap with others. As it is our natural desire that all subregions should be connected, a penalty function denoted as $u_i(\cdot)$ is introduced to punish our objective function by

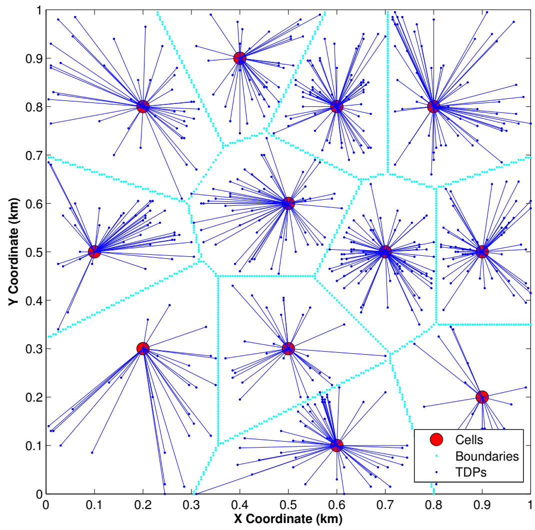

Fig. 1. Illustration of a cellular system.

preventing it from getting to its optimality when the subregions are far from connected or seriously ill-shaped. In order to measure the connectivity of the subregion served by a cell, we can define $u_i(\cdot)$ to be the Euclidean distance between the user node x and the cell i, which means $u_i(x) = \|x - p_i\|$. As a result, the overall penalty of the cell i is presented as the integral $\iint_{R_i} u_i(x) dA$. Besides, the density of demand nodes across R at the moment is statistically formulated as $f(x)$, where x is a bi-vector representing coordinates. Thus the integral $\iint_{R_i} f(x) u_i(x) dA$ would denote the overall penalty of the cell i. Finally, a parameter of fairness index to evaluate the cell designing scheme for load balancing among districts, denoted as ϕ is introduced.

Our suggested load balancing strategy is as follows. We monitor a fairness index ϕ of the cellular system, which is measured by using the load distribution of all cells collected periodically. The fairness index ϕ can be defined as that in [9],

$$\phi = \frac{(\sum_{i=1}^{n} L_i)^2}{n(\sum_{i=1}^{n} L_i^2)}, \tag{1}$$

where L_i is the load of cell i and n is the number of cells. Obviously, the range of ϕ is between 0 and 1, with higher ϕ indicating more balanced load distribution among cells.

When ϕ is below a warning threshold for a certain period of time, we triggers our load balancing algorithm to redesign the subregion served by each cell. Firstly, we estimate the number of cells, denoted as n, that is required to serve all TDPs in the service region R by considering the sum rate of the TDPs at the moment and the capacity that can be provided by each cell. Secondly, we redesign the service subregion R_i of the cell i with the cell set P to make the traffic yielded by the TDPs in each cell almost the equal, as well as the coverage area. Last, each TDP is associated with the cell that covers it based on the repartition results.

III. Algorithms for Load Balancing

When the fairness index ϕ drops below a certain threshold for a certain period of time, actions will be taken to avoid unnecessary operations caused by abrupt emergence and disappearance of load imbalance.

A. Initialization

As the sum rate required by TDPs varies with time, before designing the service area R_i of each cell i, we should firstly determine the number of cells required to sever all TDPs according to the traffic demands at the moment. Firstly, we collect the total traffic demands of the area R, denoted as F_t and estimate the average capacity of a cell, denoted as C_B. The ideal number of cells required which exists in case when the traffic demands are perfectly balanced among subregions is $n_{ideal} = F_t/C_B$. If we define a capacity margin of m, then the number of cells we finally adopt can be expressed in the form $n = (1+m) * n_{ideal} = (1+m) * F_t/C_B$. Reserving capacity margin is important for a practical cellular system because it is impossible to guarantee that all cells can provide the same capacity with the same transmission parameters. Secondly, we collect all the candidate cells all over the region R. If the number of cells required at the moment is the same as before, the initial sites of cells for stating our load balancing procedure remain the same. When the number of cells required at the moment is larger than before, we select extra cells randomly from all the candidate sites for cells, and they, together with formal cell sites, make the starting condition for our scheme. In a similar way, if the number of cells required at the moment is smaller than before, we abandon cells with the least loads from formal cell sites.

B. Subregions Partition

The most critical step in this strategy is to construct the service subregion for each cell in a balanced fashion from both the area and the traffic demand aspects. One way to implement it is to minimize the maximum traffic demand of all cells with penalty phase while introducing constraints on the amount of $f(\cdot)$ that are served by them and the area served by each cell. The first step of our strategy can be written as [10, 11]

$$\begin{aligned} \min_{R} \quad & t + \mu \sum_{i=1}^{n} \iint_{R_i} f(x) u_i(x) dA \\ s.t. \quad C_1 : \quad & t \geq (1-\mu) \iint_{R_i} f(x) dA, \forall i, \\ C_2 : \quad & \iint_{R_i} dA \geq \Omega, \forall i, \\ C_3 : \quad & R_i \cap R_j = O, \forall i \neq j, \\ C_4 : \quad & {}_i R_i = R. \end{aligned} \tag{2}$$

Here, we introduce a variable μ to represent the penalty factor corresponding to the penalty phase, and a variable t to represent the maximum value of the traffic demands all over cells with the penalty factor, which is denoted as C_1. C_2 indicates that all cells should have an area larger than a constant denoted as Ω, which implies area balance. Areas of subregions should be as balanced as possible to avoid cases when some subregions are too large to be covered by only one cell per subregion or some subregions are so small as strong interference occurs. For simplification, we normalize the total area to constant one. So the constant Ω in C_2 can be set to $1/n$ so that the areas of cells are balanced as much as possible. C_3 presents the assumption that all cells should not overlap with each other, while C_4 denotes that there cannot have any coverage hole. The objective function in (2) is designed to minimize the maximum traffic capacity all over the cells with the value of the penalty function increasing when cells are not as connected as we expect to add obstructions to reaching the optimality.

We start by transforming the problem into the form of an infinite-dimensional integer program to solve (2). By introducing a $\{0,1\}$-valued function $I_i(x)$ to indicate whether the TDP x is served by cell i, our problem could be put as the equivalent formulation as follows,

$$\begin{aligned} \min_{I_1(\cdot),\ldots,I_n(\cdot)} \quad & t + \mu \sum_{i=1}^{n} \iint_{R} f(x) I_i(x) u_i(x) dA \\ s.t. \quad C_1 : \quad & t \geq (1-\mu) \iint_{R} f(x) I_i(x) dA, \forall i, \\ C_2 : \quad & \iint_{R} I_i(x) dA \geq \Omega, \forall i, \\ C_3 : \quad & \sum_{i=1}^{n} I_i(x) = 1, \forall x, \\ C_4 : \quad & I_i(x) \in \{0,1\}, \forall i, x. \end{aligned} \tag{3}$$

The most difficult part in solving (3) lies in the integer constraints. An intuitive way to cope with them is to relax the integer variables into continuous ones [12–14]. The linear programming relaxation of (3) is given by

$$\begin{aligned} \min_{I_1(\cdot),\ldots,I_n(\cdot)} \quad & t + \mu \sum_{i=1}^{n} \iint_{R} f(x) I_i(x) u_i(x) dA \\ s.t. \quad & C_1 \sim C_3 \text{ in (3)}, \\ & I_i(x) \geq 0, \forall i, x. \end{aligned} \tag{4}$$

Here, we can discretize (4) into N grid cells $\square_j$ of area ϵ, and f_j and u_{ij} are introduced to represent the average values of $f(x)$ and $u_i(x)$ on $\square_j$. We also use the symbol z_{ij} to denote the fraction of cell $\square_j$ served by base station i, and then the approximate formulation is listed below,

$$\begin{aligned} \min_{Z} \quad & t + \mu \epsilon \sum_{i=1}^{n} \sum_{j=1}^{N} f_j z_{ij} u_{ij} \\ s.t. \quad C_1 : \quad & t \geq (1-\mu) \epsilon \sum_{j=1}^{N} f_j z_{ij}, \forall i, \\ C_2 : \quad & \epsilon \sum_{j=1}^{N} z_{ij} \geq \Omega, \forall i, \\ C_3 : \quad & \sum_{i=1}^{n} z_{ij} = 1, \forall j, \\ C_4 : \quad & z_{ij} \geq 0, \forall i, j. \end{aligned} \tag{5}$$

By introducing Lagrange multiplier vectors $a \in R^n$,$b \in R^n$ and $d \in R^N$, we can obtain the dual problem to (5) as follows,

$$\begin{aligned} &\max_{a,b,d} \quad \sum_{i=1}^{n} \Omega b_i - \sum_{j=1}^{N} d_j \\ s.t.\ & C_1: \quad a_i \geq 0, \forall i, \\ & C_2: \quad \sum_{i=1}^{n} a_i = 1, \\ & C_3: \quad b_i \geq 0, \forall i, \\ & C_4: \quad \epsilon\mu f_j u_{ij} + (1-\mu)\epsilon a_i f_j - \epsilon b_i + d_j \geq 0, \forall i, j. \end{aligned} \tag{6}$$

In order to simplify the expression, we introduce new variables $\sigma_j = -d_j/\epsilon$, $\lambda_i = a_i$ and $\gamma_i = b_i$, and rewrite (6) as

$$\begin{aligned} &\max_{\lambda,\gamma,\sigma} \quad \sum_{i=1}^{n} \Omega \gamma_i + \epsilon \sum_{j=1}^{N} \sigma_j \\ s.t.\ & C_1: \quad \lambda_i \geq 0, \forall i, \\ & C_2: \quad \sum_{i=1}^{n} \lambda_i = 1, \\ & C_3: \quad \gamma_i \geq 0, \forall i, \\ & C_4: \quad \sigma_j \leq \mu f_j u_{ij} + (1-\mu)\lambda_i f_j - \gamma_i, \forall i, j, \end{aligned} \tag{7}$$

which is a discretization of the following optimization problem

$$\begin{aligned} &\max_{\lambda,\gamma,\sigma(\cdot)} \quad \Omega \sum_{i=1}^{n} \gamma_i + \iint_R \sigma(x) dA \\ &s.t. \quad \sigma(x) \leq \mu f(x) u_i(x) + (1-\mu)\lambda_i f(x) - \gamma_i, \forall i, x, \\ &\qquad C_1 \sim C_3 \text{ in (7)}. \end{aligned} \tag{8}$$

(8) can be rewritten in a simpler form as follows,

$$\begin{aligned} &\max_{\lambda,\gamma} \quad \iint_R \min_i (\mu f(x) u_i(x) + (1-\mu)\lambda_i f(x) - \gamma_i) dA \\ &\qquad + \Omega \sum_{i=1}^{n} \gamma_i \\ &s.t. \quad C_1 \sim C_3 \text{ in (7)}. \end{aligned} \tag{9}$$

Up to now, a convex, $2n$-dimensional dual problem is obtained [15]. (9) can be efficiently solved with standard convex optimization techniques. In this paper, we adopt a well-developed tool CVX, which is a modeling system for constructing and solving disciplined convex programs, to solve (9) [16].

After solving the dual problem (9), the dual variables λ and γ corresponding to the optimal solution to the original problem (2) are obtained. For any user node, it will be served by cell p_i which minimizes $\mu f(x) u_i(x) + (1-\mu)\lambda_i f(x) - \gamma_i$ among $i \in [1, n]$. In this case, the boundaries between the optimal cells to problem (2) are curves of the form

$$\partial(R_i^*) \cap \partial(R_j^*) \subseteq \Big\{ x \mid \quad x \in R, \mu f(x)(u_i(x) - u_j(x)) + (1-\mu) f(x)(\lambda_i - \lambda_j) = \gamma_i - \gamma_j \Big\} \tag{10}$$

Last but not the least, it still remains to be shown that the solution to (2) can be recovered from the optimal solution to (9). Consider any point $x \in R$ and the optimal solution to (9). Suppose $\bar{i}$ is the index such that $\mu f(x) u_i(x) + (1-\mu)\lambda_i f(x) - \gamma_i$ is minimal (assuming such an index is unique). From basic linear programming theory, we know that the complementary slackness conditions of problem (8) stipulate that $I_i^*(x) = 0$ for all indices i other than $\bar{i}$ [17], and consequently that $I_{\bar{i}}^*(x) = 1$. In a conclusion, despite relaxation, the optimal solution to (2) remains valid as proved in [18].

C. Cells Relocation

After the division of subregions, the traffic demands together with areas among different cells have been balanced. However, with the variation of the distribution of TDPs, the current cells that are active may not be as appropriate as before as the shapes of subregions have changed after the subregion partition. The power consumption of each cell may be reduced if the TDPs in a subregion are served by another available cell. Since there may be multiple candidate cells in each subregion, it is logical to select the one with the minimum power consumption. Considering the fact that the number of candidate cells in each subregion is very limited in practical cellular systems, we can adopt a quite straightforward method, exhaustive search, to obtain the best coordinates from all candidate cells. To be more specific, for the subregion i, there are g_i available cells. Denote W_{ij} as the power consumed in the subregion i when we select the jth candidate cell to serve the TDPs, we have a set

$$W_i = \{W_{i1}, W_{i2}, \ldots, W_{ig_i}\} \tag{11}$$

which indicates all possible power consumptions in subregion $i (i = 1, \ldots, n)$. Here,

$$W_{ij} = \iint_{R_i} B_{node} * (2^{C_{node}/B_{node}} - 1)/G \; dA, \tag{12}$$

where B_{node} and C_{node} denote the bandwidth and the traffic demand required by the TDPs respectively, and G is the signal-to-noise ratio (SNR) with unit power when selecting the jth candidate cell in subregion i. By searching for the minimum value in W_i, we set the candidate cell corresponding to the minimum value to be the revised coordinates for subregion i, through which we decrease the power consumption as much as possible.

The power consumption of each subregion has been minimized till now based on the subregion partition we obtain before that. However, the cell partition is based on formally selected cells which have not given the distribution of TDPs at the moment enough attention, so we can decrease power consumptions further by do subregion partitions and cells relocation in a loop until the total power consumptions can not be decreased any more. Our proposed load balancing strategy is summarized in Table I.

IV. Experimental Results and Discussions

In this section, we present some preliminary experiment results to show the effectiveness of our proposed load balancing method. Firstly, we show the loads of cells before and after

TABLE I
PROPOSED LOAD BALANCING SCHEME

Initialize: $recycle = 1$, monitor the fairness index of load balancing ϕ, the alarm threshold of ϕ is set to T_{fair} and collect all the candidate sites for cells across R.
while $(\phi < T_{fair})$
  Determine the number of cells required n and set initial sites for cells $P^{(recycle)}$;
  repeat
    Calculate $\lambda^{(recycle)}$, $\gamma^{(recycle)}$ that satisfies Eq.(9);
    Recover $R_i^{(recycle)}(i = 1, \ldots, n)$ from $\lambda^{(recycle)}$ and $\gamma^{(recycle)}$;
    Relocate $P^{(recycle)}$ to $P^{(recycle+1)}$ to minimize power consumptions based on $R_i^{(recycle)}(i = 1, \ldots, n)$;
    Calculate total power consumptions $K_T^{(recycle)}$;
    $recycle = recycle + 1$;
  until Mean square error of $K_T^{(s)}(s = (recycle - 9), (recycle - 8), \ldots, recycle) \le \varepsilon$
end while
return $R_i^*(i = 1, \ldots, n)$, K_T^*, P^*.

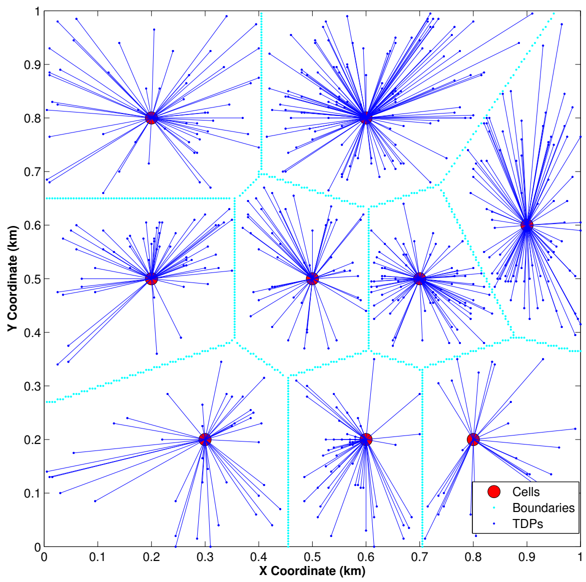

Fig. 2. Distributions of TDPs and cell associations before load balancing.

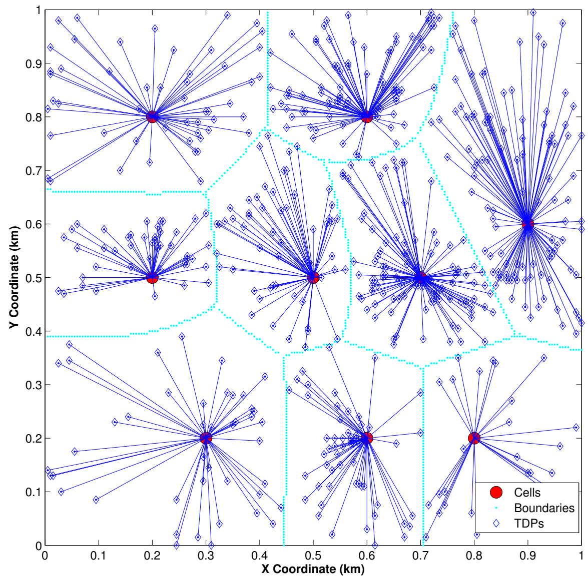

Fig. 3. Distributions of TDPs and cell associations after load balancing.

load balancing for a cellular network. Then we compare the fairness index of cells between the situations before and after load balancing, along with the number of TDPs increasing.

Fig.2 shows the distributions of TDPs and their associated cells at t when the fairness index is below predefined threshold and the proposed load balancing algorithm is triggered. It can be observed that TDPs mainly gather at some central business districts and the traffic loads among cells at t are severely lopsided. By contrast, the traffic loads of cells after carrying out our proposed load balancing scheme look more balanced as shown intuitively in Fig.3. The load comparison of each cell before and after load balancing is shown in Fig.4. It can be seen that before load balancing, load is unevenly distributed among n cells, where cell 1 and cell 4 are seriously overloaded, while cell 3 and cell 9 are comparatively light-loaded, with 0.6492 as the fairness index. After carrying out our proposed load balancing strategy, the load balance is much better represented as can be seen from Fig.4, with 0.9423 as the fairness index.

Fig.5 is under the simulation that new TDPs join at a certain interval. The locations of TDPs are randomly set, as a result of which the situation of load imbalance of cells varies as new TDPs join. It can be seen in Fig.5 that with the number of TDPs increasing, the fairness of load balancing among cells is always better after applying our proposed strategy. When TDPs are rather unevenly distributed among cells which causes great load imbalance, for instance the number of TDPs at 100

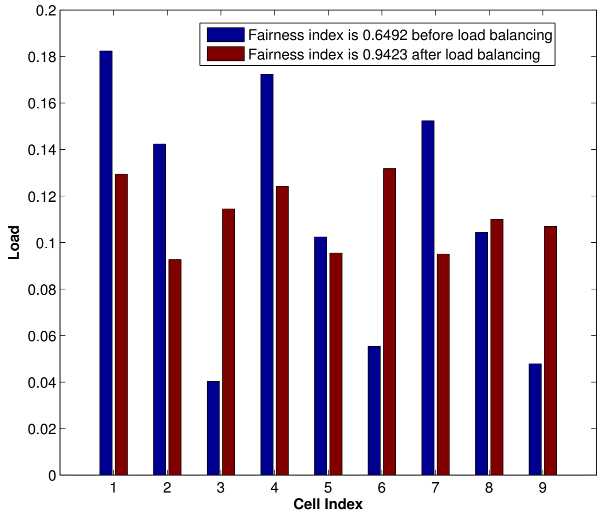

Fig. 4. Load of each cell before and after load balancing.

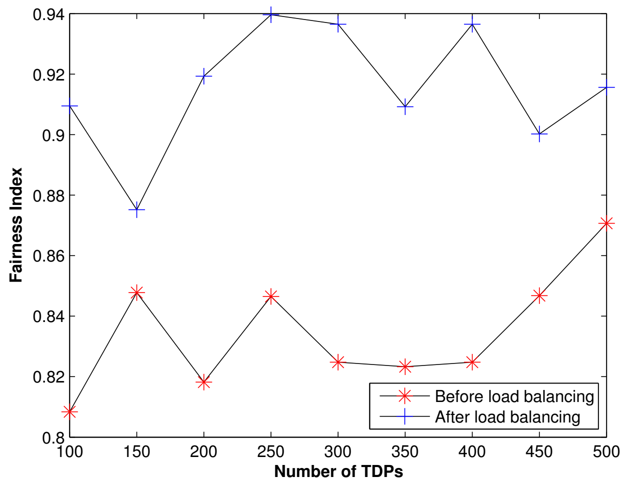

Fig. 5. Fairness index before and after load balancing with the number of TDPs increasing.

and 200, our strategy can improve the fairness significantly.

V. Conclusion

In this paper we presented a novel load balancing method for CRAN-based cellular networks. Our proposed method can balance the traffic demands and areas among all cells and decrease power consumption of the cellular system as much as possible, which can achieve an almost optimal solution from the viewpoint of load balance. The load balancing algorithm is triggered when the fairness index measured by monitoring the loads of all active cells periodically is below a threshold. We estimate the number of cells needed to serve all users and redivide the service zone into equal number of subregions corresponding to the cells while keeping the traffic demand and area of each subregion as equal as possible. Then we select the cell in each subregion that can minimize power consumption. The dividing and selecting procedures are run in an iterative manner until the overall consumed power of the cellular system can not be smaller any more. Experimental results verified the effectiveness and the efficiency of our proposed scheme. Only some preliminary results are reported in this paper because of space limitation. For future work, extensive experiments to evaluate our proposed algorithms should be conducted, such as removing or reactivating cells. The shapes of cells should also be further improved to meet the requirements of practical cellular systems. Moreover, the tradeoff between load balance and power consumption should also be paid more attention to to both improve the performance of users and decrease the energy required.

References

[1] China Mobile Res. Inst., "C-RAN: The road towards green RAN," White Paper, ver. 2.5, 2011.

[2] Z. Zhu, P. Gupta, and et al., "Virtual base station pool: towards a wireless network cloud for radio access networks." in *Proc. of the 8th ACM International Conference on Computing Frontiers*, 2010.

[3] Intel Cor., "Solution brief intel heterogeneous network solution brief," Solution brief, 2011.

[4] Z. Niu, Y. Wu, J. Gong, and Z. Yang, "Cell zooming for cost-efficient green cellular networks," *IEEE Commun. Mag.*, vol. 48, no. 11, pp. 74–79, November 2010.

[5] P. Bahl, M. T. Hajiaghayi, K. Jain, S. V. Mirrokni, L. Qiu, and A. Saberi, "Cell breathing in wireless LANs: Algorithms and evaluation," *IEEE Trans. Mobile Comput.*, vol. 6, no. 2, pp. 164–178, Feb. 2007.

[6] Y. Bejerano and S.-J. Han, "Cell breathing techniques for load balancing in wireless LANs," *IEEE Trans. Mobile Comput.*, vol. 8, no. 6, pp. 735–749, June 2009.

[7] E. Garcia, R. Vidal, and J. Paradells, "Cooperative load balancing in IEEE 802.11 networks with cell breathing," in *Proc. IEEE ISCC'08*, 2008, pp. 1133–1140.

[8] S. Luo, R. Zhang, and T. J. Lim, "Optimal power and range adaptation for green broadcasting," *IEEE Trans. Wireless Commun.*, vol. 12, no. 9, pp. 4592–4603, Sep. 2013.

[9] H. Velayos, V. Aleo, and G. Karlsson, "Load balancing in overlapping wireless LAN cells," in *Proc. IEEE ICC'04*, 2004, pp. 3833–3836.

[10] J. G. Carlsson and R. Devulapalli, "Shadow prices in territory division," *University of Minnesota, Available at: http://menet. umn. edu/~jgc/shadow-prices-rev2. pdf*, 2013.

[11] J. G. Carlsson, "Dividing a territory among several vehicles," *INFORMS J. Comput.*, vol. 24, no. 4, pp. 565–577, Aug. 2012.

[12] M. Ge and S. Wang, "Fast optimal resource allocation is possible for multiuser OFDM-based cognitive radio networks with heterogeneous services," *IEEE Trans. Wireless Commun.*, vol. 11, no. 4, pp. 1500–1509, Apr. 2012.

[13] S. Wang, Z.-H. Zhou, M. Ge, and C. Wang, "Resource allocation for heterogeneous cognitive radio networks with imperfect spectrum sensing," *IEEE J. Sel. Areas Commun.*, vol. 31, no. 3, pp. 464–475, Mar. 2013.

[14] S. Wang, M. Ge, and W. Zhao, "Energy-efficient resource allocation for OFDM-Based cognitive radio networks," *IEEE Trans. Commun.*, vol. 61, no. 8, pp. 3181–3191, Aug. 2013.

[15] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge University Press: New York, 2004.

[16] M. Grant, S. Boyd, and Y. Ye, "cvx users' guide," Technical Report Build 711, Citeseer. Available at: http://citeseerx. ist. psu. edu/viewdoc/download, Tech. Rep., 2009.

[17] D. G. Luenberger and Y. Ye, *Linear and nonlinear programming*. Springer, 2008.

[18] J. Carlsson, E. Carlsson, and R. Devulapalli, "Balancing workloads of service vehicles over a geographic territory," in *Proc. IEEE IROS'13*, Nov. 2013, pp. 209–216.