

Deploying C-RAN in Cellular Radio Networks: An Efficient Way to Meet Future Traffic Demands

Qiu Shen, *Member, IEEE*, Zhan Ma, *Member, IEEE*,
and Shaowei Wang , *Senior Member, IEEE*

Abstract—Cloud radio access networks (C-RANs) have been proposed to provide high energy-spectral efficiency and high data rate at low cost. In this correspondence paper, we introduce a novel C-RAN deployment scheme that is complementary to traditional cellular radio networks. The key idea is to install the remote radio heads (RRHs) of the C-RAN in a load-balanced way to lighten the transmission burden of the backhauls between the RRHs and the centralized baseband unit pool, and to activate the RRHs on demand so as to reduce system-wide energy consumption. Numerical results show that our proposed method performs well in practical scenarios. It provides QoS-guaranteed performance with relatively low capital expenditure and operating expenditure.

Index Terms—Cloud radio access networks, heterogeneous network, load balancing, network planning.

I. INTRODUCTION

With the popularity of smartphones, the traffic demand of mobile network is increasing dramatically. It is reported that the global mobile data traffic will increase sevenfold between 2016 and 2021, which means a compound annual growth rate of 47% during this period [1]. Meanwhile, mobile applications become more diversified and complicated, especially for multimedia services such as video streaming transmission with 4K, 8K, and gigapixel resolution, as well as virtual reality and augmented reality. Future mobile networks have to deal with all these challenges effectively.

To enhance network capacity, a feasible way is to increase the number of base stations (BSs) or cells in the cellular radio networks. However, the capacity gain can be achieved in this way only if the BSs are sparsely distributed in the service area. Another solution to capacity expansion is heterogeneous network (HetNet) [2], where densely deployed low power nodes (LPNs), such as pico cells, femto cells and relays, are expected to meet the high data rate requirement of traffic hot zones. However, the HetNet faces severe cross-tier interference issue and the system performance even degrades if the density of LPNs is too high [3]. Besides, site acquisition in urban areas is always expensive. Deploying too many LPNs inevitably incurs unacceptable capital expenditure (CAPEX) and operating expenditure (OPEX)

Manuscript received September 13, 2017; revised December 26, 2017 and March 19, 2018; accepted May 12, 2018. Date of publication May 18, 2018; date of current version August 13, 2018. This work was supported in part by the National Natural Science Foundation of China under Grant 61671233, in part by the Jiangsu Science Foundation under Grant BK20151389, and in part by the Fundamental Research Funds for the Central Universities under Grant 021014380094. The review of this paper was coordinated by Dr. Y. Mao. (Corresponding author: Shaowei Wang.)

The authors are with the School of Electronic Science and Engineering, Nanjing University, Nanjing 210023, China (e-mail: shenqiu@nju.edu.cn; mazhan@nju.edu.cn; wangsw@nju.edu.cn).

Color versions of one or more of the figures in this paper are available online at <http://ieeexplore.ieee.org>.

Digital Object Identifier 10.1109/TVT.2018.2838133

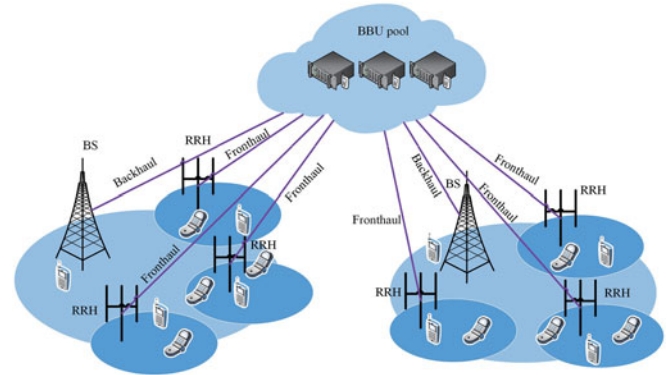


Fig. 1. Deploying C-RAN in an existing mobile network.

for service providers who always concern the average revenue per user.

On the other hand, cloud radio access networks (C-RANs) [4] have been proposed as a promising architecture to provide high transmission data rate with high energy efficiency, and most importantly, with lower CAPEX and OPEX as compared to the HetNet [5]. The C-RAN is mainly composed of three parts: distributed remote radio heads (RRHs), centralized baseband unit (BBU) pool, and high-data-rate and low-latency backhauls between the RRHs and the BBU pool. Since most of the signal processing procedures are performed in the virtual BBU pool, the operational functionalities of RRHs are much less than the LPNs in the HetNet. It is more cost-efficient by deploying RRHs in the C-RAN with the same density as the LPNs in the HetNet. However, the C-RAN also faces a challenging problem: The coverage area of an RRH is much smaller than that of a macro cell due to its lower transmission power. The number of required RRHs would be very large to serve an area covered by macro cells previously. It is costly to construct an exclusive C-RAN to provide users reliable wireless access for service providers concerning the CAPEX and the OPEX. Moreover, the backward compatibility issue with conventional cellular networks is also an obstacle if deploying the C-RAN subversively.

Motivated by the respective advantages of the HetNet and the C-RAN, we can consider to deploy the C-RAN selectively in traditional cellular radio networks to jointly provide users with high data rate and seamless coverage [6]. Moreover, it can efficiently address the problem of dynamic traffic because it is easy to switch on or off an RRH to adapt to the traffic fluctuation in the service area [7]. In such kind of network, the conventional macro BSs can provide seamless coverage for the whole service area and cooperate with the BBU pool of the C-RAN to control the RRHs [8]. The RRHs make decisions on their coverage via the signalling from the BBU pool or the macro BSs, and adjust their states (e.g., on or off) based on the traffic distribution change in the service area to reduce the system energy consumption. Fig. 1 shows the architecture of the proposed network model. The questions are how to deploy the RRHs in a cost-efficient way and how to adjust the states of RRHs according to the traffic variations.

TABLE I
FREQUENTLY USED NOTATIONS

| Symbol | Semantics |
|--------------|---|
| A | 2-dimension integration variable |
| f_k | Average of $f(x)$ on ξ_k |
| $f(x)$ | Traffic density distribution in R |
| $I_j(x)$ | Index function |
| p_i | The i th BS |
| Q | Set of RRHs in a subdistrict |
| q_j | The j th RRH |
| R | The whole service area |
| R_i | Service area of p_i |
| S_j | Service area of q_j |
| t | Maximum traffic demand in RRHs |
| $u_j(\cdot)$ | Euclidean distance between user and q_j |
| u_{jk} | Average of $u_j(x)$ on ξ_k |
| z_{jk} | Fraction of zone ξ_k served by q_j |
| μ | Penalty term |
| ξ_k | Grid zone k |
| ϵ | Area of grid zone k |
| Ω | Constant |

In this paper, we try to answer these two questions from the viewpoint of load balancing, where we allocate each RRH almost the same traffic via cutting-edge optimization technique so that the number of activated RRHs can be minimized. Our proposed method can also deal with the situation that the traffic demand varies greatly in the service area. The main contributions of this work are summarized as follows:

- We propose a backward-compatible C-RAN deploying scheme for mobile networks, which can satisfy the ever-increasing traffic demand at low CAPEX and OPEX by taking advantage of HetNet and C-RAN.
- We introduce an efficient algorithm to fulfill the intractable RRH planning task, which can balance the traffic loads among RRHs so as to reduce the number of active RRHs, as well as the CAPEX and the OPEX.
- Our proposed method is a quasi-dynamic scheme, which adapts to traffic distribution and provides a feasible solution to address the traffic fluctuation issue faced by practical networks.

II. RRHS DEPLOYING STRATEGY IN CELLULAR RADIO NETWORKS

Some frequently used notations is listed in Table I. Consider a given area R made up of n districts R_i , $\bigcup_i R_i = R$, and is served by n BSs $P = \{p_1, p_2, \dots, p_i, \dots, p_n\}$. Without loss of generality, we assume that R is a connected, polygonal region with non-empty interior. BS p_i serves district R_i . According to the traffic demand distribution in the service area, n_i RRHs, denoted as $Q = \{q_1, q_2, \dots, q_j, \dots, q_{n_i}\}$, are required to facilitate BS p_i to provide robust high speed data transmission in R_i . RRH q_j serves a small district S_j without overlap. Since the district served by an RRH should be connected, a penalty function $u_j(\cdot)$ is introduced to punish the objective function so that the districts are not far from connected. To measure the connectivity of the district served by an RRH, we define $u_j(\cdot)$ as Euclidean distance between a user and RRH q_j serving this user, which means $u_j(x) = \|x - q_j\|$. The density of traffic demand across R is statistically formulated as $f(x)$, where x is a bi-vector

representing coordinates. Thus the integral $\iint_{S_j} f(x)u_j(x)dA$ would denote the overall penalty of RRH q_j , where A is 2-dimension integration variable. Our RRHs deployment strategy is as follows:

- Estimate the number of RRHs, denoted as n_i , that should be deployed to provide high speed data transmission in each objective region R_i . Necessary capacity margin should be reserved to account for the continuous traffic variation in the coverage of an RRH. Candidate sites for deploying RRHs should be determined under the consideration of height, terrain, and density of population.
- Design the service area of S_j of each RRH q_j under four conditions. First, traffic load among RRHs $\iint_{S_j} f(x)dA$ should be as balanced as possible; second, the service subregions of RRHs $\iint_{S_j} dA$ should not differ too much from each other to avoid the case where some subregions are too large to be covered by only one RRH. Furthermore, it can also help avoid yielding ill-shaped subregions; third, there should not be any coverage holes in the service area that is thirsty for high traffic rate; fourth, the service subregions of RRHs should not overlap with each other for simplification.
- Relocate RRHs to minimize the system power consumption. Specifically, the RRH that yields the minimum power consumption in a given subregion should be activated to provide service in this subregion.

The load distribution among RRHs can be adjusted adaptively via our proposed method, for instance, we can define a fairness index to measure the load imbalance degree of the network and trigger the RRHs adjustment procedure when the index exceeds a threshold [9].

III. EFFICIENT ALGORITHMS FOR RRHS PLANNING AND OPTIMIZATION

A. Initialization

We firstly need to know the macro BSs that cannot provide users with expected data rate. In practical networks, it is not a difficult task since the traffic demand in the service area of a macro BS is usually recorded by service providers. Then the number of required RRHs could be estimated according to the traffic gap in the target area R_i . We need to collect not only the traffic distribution in R_i but also the candidate sites to install RRHs. Since the input of our RRHs planning method is mainly based on historical data, it can yield more effective planning scheme than conventional cellular radio planning schemes that are usually based on demographic geography [10]–[12].

Reserving capacity margin is important as it is difficult if not impossible to guarantee that all RRHs can yield the same capacity in wireless environment. Another advantage of capacity margin is to absorb the constant variation of traffic demand that is not violent. For a planned C-RAN, our proposed method can work as follows. We collect the candidate sites of all installed RRHs over the region R_i and the traffic demand therein. If the number of required RRHs is smaller than the active ones, we abandon the RRH with the least traffic demand from the current active RRHs. Otherwise, we select some RRHs randomly from the inactive ones until the traffic demand is satisfied.

B. Designing Service Areas of RRHs

As discussed above, if each active RRH serves almost the same number of traffic demand points (TDPs) or users, the activated RRHs can be minimized for an area with given traffic distribution. We can achieve this goal by minimizing the maximum traffic load of the RRHs with penalty phase by introducing constraints on $f(\cdot)$ [13], [14] and the optimization task is formulated as follows:

$$\begin{aligned}
\min_{R_i} \quad & t + \mu \sum_{j=1}^{n_i} \iint_{S_j} f(x)u_j(x)dA \\
\text{s.t. } C_1 : \quad & t \geq (1 - \mu) \iint_{S_j} f(x)dA, \forall j, \\
C_2 : \quad & \iint_{S_j} dA \geq \Omega, \forall j, \\
C_3 : \quad & S_j \cap S_k = O, \forall j \neq k, \\
C_4 : \quad & \bigcup_j S_j = R_i,
\end{aligned} \tag{1}$$

where μ is a penalty term and t is the maximum traffic demand of the RRHs with penalty factor. C_2 means that the area of an RRH should be larger than a constant Ω , which guarantees that the service region of each RRH does not differ too much from each other to avoid generating any region too large to be served by only one RRH, or ill-shaped regions. If we denote the area of the service region as 1, Ω in C_2 can be normalized to a positive that is a little smaller than $1/n_i$. Put aside handover issue, C_3 indicates that the subregions do not overlap with each other. C_4 means that there is no coverage hole in the service area.

We first transform (1) into an infinite-dimensional integer program by introducing a $\{0, 1\}$ -valued function $I_j(x)$ which indicates whether TDP x is served by RRH q_j or not as follows:

$$\begin{aligned}
\min_{I_1(\cdot), \dots, I_{n_i}(\cdot)} \quad & t + \mu \sum_{j=1}^{n_i} \iint_{R_i} f(x)I_j(x)u_j(x)dA \\
\text{s.t. } C_1 : \quad & t \geq (1 - \mu) \iint_{R_i} f(x)I_j(x)dA, \forall j, \\
C_2 : \quad & \iint_{R_i} I_j(x)dA \geq \Omega, \forall j, \\
C_3 : \quad & \sum_{j=1}^{n_i} I_j(x) = 1, \forall x, \\
C_4 : \quad & I_j(x) \in \{0, 1\}, \forall j, x.
\end{aligned} \tag{2}$$

Technically, (2) is also hard to handle because of the integer constraints $I_j(x)$. We relax $I_j(x)$ into continuous ones as introduced in [15]. The linear programming relaxation form of (2) is as follows:

$$\begin{aligned}
\min_{I_1(\cdot), \dots, I_{n_i}(\cdot)} \quad & t + \mu \sum_{j=1}^{n_i} \iint_{R_i} f(x)I_j(x)u_j(x)dA \\
\text{s.t.} \quad & C_1 \sim C_3 \text{ in (2)}, \\
& I_j(x) \geq 0, \forall j, x.
\end{aligned} \tag{3}$$

Then we discretize (3) into N grid zones ξ_k of area ϵ . Denote f_k and u_{jk} as the average of $f(x)$ and $u_j(x)$ on ξ_k , respectively. z_{jk} is the fraction of zone ξ_k served by RRH q_j . An approximation formulation of (3) can be written as follows:

$$\begin{aligned}
\min_Z \quad & t + \mu \epsilon \sum_{j=1}^{n_i} \sum_{k=1}^N f_k z_{jk} u_{jk} \\
\text{s.t. } C_1 : \quad & t \geq (1 - \mu) \epsilon \sum_{k=1}^N f_k z_{jk}, \forall j, \\
C_2 : \quad & \epsilon \sum_{k=1}^N z_{jk} \geq \Omega, \forall j, \\
C_3 : \quad & \sum_{j=1}^{n_i} z_{jk} = 1, \forall k, \\
C_4 : \quad & z_{jk} \geq 0, \forall j, k.
\end{aligned} \tag{4}$$

The dual problem of (4) is

$$\begin{aligned}
\max_{\lambda, \gamma, \sigma} \quad & \sum_{j=1}^{n_i} \Omega \gamma_j + \epsilon \sum_{k=1}^N \sigma_k \\
\text{s.t. } C_1 : \quad & \lambda_j \geq 0, \forall j, \\
C_2 : \quad & \sum_{j=1}^{n_i} \lambda_j = 1, \\
C_3 : \quad & \gamma_j \geq 0, \forall j, \\
C_4 : \quad & \sigma_k \leq \mu f_k u_{jk} + (1 - \mu) \lambda_j f_k - \gamma_j, \forall j, k,
\end{aligned} \tag{5}$$

where $\lambda_j \in R^{n_i}, \gamma_j \in R^{n_i}$ and $\sigma_k \in R^N$ are Lagrange multiplier vectors. (5) is the discretization form of the following optimization problem:

$$\begin{aligned}
\max_{\lambda, \gamma, \sigma(\cdot)} \quad & \Omega \sum_{j=1}^{n_i} \gamma_j + \iint_{R_i} \sigma(x)dA \\
\text{s.t.} \quad & \sigma(x) \leq \mu f(x)u_j(x) + (1 - \mu) \lambda_j f(x) - \gamma_j, \forall j, x, \\
& C_1 \sim C_3 \text{ in (5)}.
\end{aligned} \tag{6}$$

Moreover, a simpler form of (6) can be rewritten as follows:

$$\begin{aligned}
\max_{\lambda, \gamma} \quad & \iint_{R_i} \min_j (\mu f(x)u_j(x) + (1 - \mu) \lambda_j f(x) - \gamma_j) dA \\
& + \Omega \sum_{j=1}^{n_i} \gamma_j \\
\text{s.t.} \quad & C_1 \sim C_3 \text{ in (5)}.
\end{aligned} \tag{7}$$

Equation (7) defines a convex, $2n_i$ -dimensional optimization task which can be solved by standard optimization techniques.

When the solution to (7) is obtained, the dual variables λ and γ corresponding to the optimal solution to (1) can be obtained. The term $\mu f(x)u_j(x) + (1 - \mu) \lambda_j f(x) - \gamma_j$ among $j \in [1, n_i]$ is minimized. Moreover, each user/TDP is associated with an RRH

q_j and the coverage boundary of each RRH can be produced by solving the following equation:

$$\{x \mid x \in R_i, \mu f(x)(u_j(x) - u_k(x)) + (1 - \mu)f(x)(\lambda_j - \lambda_k) = \gamma_j - \gamma_k\}. \quad (8)$$

For any point $x \in R_i$ and the optimal solution to (7), assume that \bar{j} is the index such that $\mu f(x)u_j(x) + (1 - \mu)\lambda_j f(x) - \gamma_j$ is minimal and unique. It is shown that the complementary slackness conditions of problem (6) stipulate that $I_j^*(x) = 0$ for all indices j other than \bar{j} , and consequently that $I_{\bar{j}}^*(x) = 1$ based on linear programming theory. As a result, the optimal solution to (1) remains valid despite relaxation [16].

C. Decreasing System Power Consumption

The traffic load of each RRH can be distributed in a balanced way with the proposed procedure. However, we do not take the system power consumption into consideration in the problem formulation. We can fine-tune the locations of RRHs based on the candidate sites to decrease the power consumption at the planning stage. If multiple candidate RRHs exist in a subregion, it is intuitive that we should choose the one with the minimum power requirement to serve the users in this region. Recall this procedure does not change the division of subregions. Generally speaking, the number of candidate RRHs in each region is limited in practical network so we can employ straightforward exhaustive search to work out the most promising site from all candidate ones. Since the design of the service area is based on the initial coordinates of RRHs that do not relate to the distribution of traffic demands or power consumption, it cannot be optimal from the viewpoint of the system. We can decrease power consumption further by repeating the RRHs planning and the fine-tuning procedures until the total power consumption cannot decrease any more.

In summary, our proposal includes two parts: First, we avoid the case that some RRHs suffer from heavy loads while others are slightly loaded based on the traffic demand distribution in the service area during the planning stage, which can improve the utilization efficiency of RRHs so as to reduce the network CAPEX; second, we adjust the on/off states of RRHs to adapt to the traffic distribution change and activate the RRHs as few as possible to reduce the energy consumption and the OPEX during the optimization stage. On the other hand, the pressure of the transport network of the C-RAN can be alleviated in this way. Consider that the function of an RRH is greatly simplified compared to a conventional BS, it is more reasonable to implement the aforementioned planning and optimization procedures in the C-RAN than those in conventional cellular networks [9]. Notice that these operations unavoidably yield relatively heavy computation load on the core network. However, the centralized BBU pool in the C-RAN is generally equipped with powerful computing capacity which can support these operations effortlessly. Overall, we fully exploit the computing potentials of the C-RAN to enhance the system performance with lower cost.

IV. EXPERIMENTAL RESULTS AND DISCUSSIONS

Our proposed RRHs deployment scheme is tested in a special area of a real city, where 9 BSs provide wireless coverage but

TABLE II
TRAFFIC LOAD DISTRIBUTION IN THE SERVICE AREA

| BS | Load | RRHs | StD-V | StD-O | Mean |
|----|--------|------|--------|--------|--------|
| 1 | 0.1702 | 14 | 0.0525 | 0.0413 | 0.0714 |
| 2 | 0.0414 | 4 | 0.2412 | 0.2270 | 0.2500 |
| 3 | 0.1141 | 9 | 0.1024 | 0.0619 | 0.1111 |
| 4 | 0.1225 | 11 | 0.0997 | 0.0714 | 0.0909 |
| 5 | 0.1396 | 12 | 0.0737 | 0.0506 | 0.0833 |
| 6 | 0.1062 | 9 | 0.0988 | 0.0876 | 0.1111 |
| 7 | 0.1154 | 10 | 0.0863 | 0.0533 | 0.1000 |
| 8 | 0.1304 | 11 | 0.0783 | 0.0354 | 0.0909 |
| 9 | 0.0602 | 5 | 0.1945 | 0.1414 | 0.2000 |

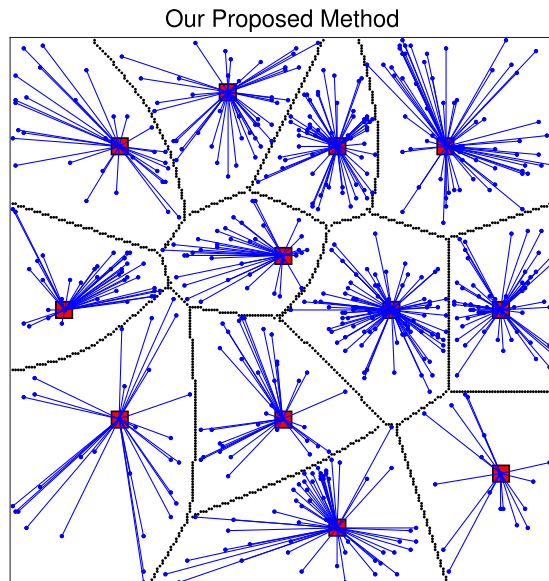


Fig. 2. RRHs deployment result of our proposed method.

suffer from capacity deficit. The traffic distribution is based on historical data recorded in three months. The number of required RRHs in each BS is shown in column 3 of Table II. Column 2 is the traffic load of each cell in proportion to the total traffic in the service area. We can see that more RRHs should be deployed in the cell with higher traffic load. To further evaluate the performance of our proposed method, we compare it with Voronoi diagram, which always associates a user with the nearest RRH. Recall that common user association schemes in cellular networks, such as maximum received signal power or signal-to-interference-plus-noise ratio, can be seen as the special cases of Voronoi diagram. Figs. 2 and 3 show the deployment results for a BS with 12 RRHs, where the red squares denote the RRHs and the blue points are the TDPs. The TDPs are generated from historical data as follows: First, we normalized the traffic of diverse services in each BS; second, we divide the coverage of each BS into uniform grids with $20m$ by $20m$; third, for the user with recorded location, the traffic yielded by the user is allocated to the corresponding grid, otherwise, the traffic of the user is randomly distributed in a grid.

Clear difference of user association between the two methods can be found in Figs. 2 and 3, which leads to significant load difference among RRHs. Column 4 and 5 of Table II show standard derivation of normalized traffic load among RRHs in each BS obtained by Voronoi diagram (StD-V) and our proposed method

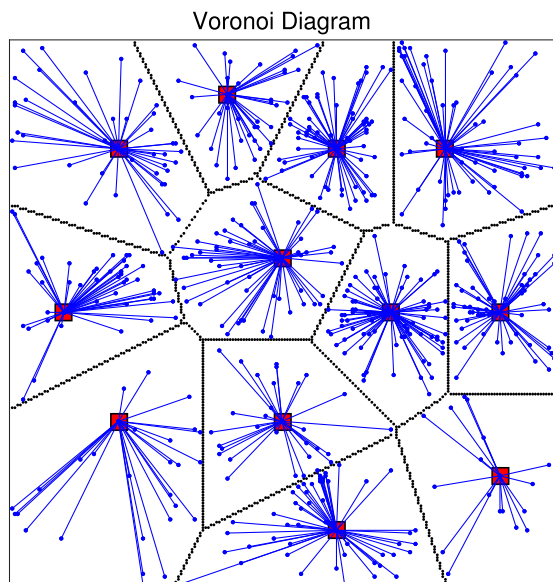


Fig. 3. RRHs deployment result of Voronoi diagram.

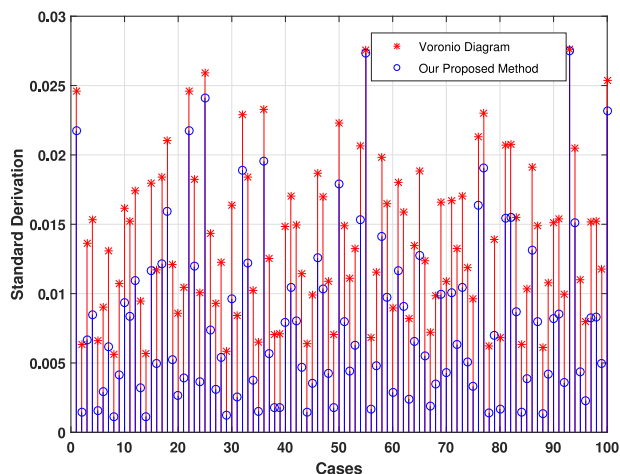


Fig. 4. Standard deviation of traffic demands among RRHs in 100 cases.

(StD-O), respectively. Our proposed method yields more balanced traffic distribution among RRHs as compared to Voronoi diagram. As can be found in the first column of Table II, the traffic distribution is heavily uneven in the considered mobile network, e.g., the maximum traffic load is more than 4 times than the minimum one among BSs based on the recorded historical data. In contrast the heaviest load RRH is less than 2 times as compared with others even in Cell 2 that has the largest standard deviation of normalized traffic load.

Our proposed algorithm is not sensitive to initialization. We test 100 cases that the initial sites of RRHs are random distributed in a given area, where the TDPs are generated randomly and remain unchanged in all 100 cases. The number of RRHs is 10. The standard deviation of the normalized traffic load among RRHs is shown in Fig. 4. Again, our strategy performs much better than Voronoi diagram in aspect of load balance since almost in all cases the standard deviation of our method is smaller

than that of Voronoi diagram. Moreover, in about 95% cases, the standard deviation of our proposed method is less than 0.02. We can conclude that our strategy is robust even though we cannot obtain enough prior knowledge about the initial locations of RRHs, which is usually the case in the RRHs planning stage.

V. CONCLUSION

In this paper, we presented a novel method for deploying C-RAN in traditional mobile networks while considering the backward compatibility and sustainable evolution. The core idea is to allocate almost equal traffic load to each access point to minimize the number of activated RRHs and alleviate the burden of backhauls in the C-RAN. Our proposed method can also address the issue of traffic fluctuation in mobile networks efficiently. Numerical results indicate that our proposal provides a cost-efficient way to enhance the performance of mobile networks. For future work, the RRHs on and off issue should be considered further since it can lead to many handovers, which would occur unbearable signalling overhead to the system.

REFERENCES

- [1] Cisco, "Cisco visual networking index: Global mobile data traffic forecast update," White Paper, 2017.
- [2] J. G. Andrews, S. Singh, Q. Ye, X. Lin, and H. S. Dhillon, "An overview of load balancing in hetnets: Old myths and open problems," *IEEE Wireless Commun.*, vol. 21, no. 2, pp. 18–25, Apr. 2014.
- [3] M. Peng, K. Zhang, J. Jiang, J. Wang, and W. Wang, "Energy-efficient resource assignment and power allocation in heterogeneous cloud radio access networks," *IEEE Trans. Veh. Technol.*, vol. 64, no. 11, pp. 5275–5287, Nov. 2015.
- [4] China Mobile Res. Inst., "C-RAN: The road towards green RAN," White Paper, 2011.
- [5] Y. Cai, F. R. Yu, and S. Bu, "Dynamic operations of cloud radio access networks (C-RAN) for mobile cloud computing systems," *IEEE Trans. Veh. Technol.*, vol. 65, no. 3, pp. 1536–1548, Mar. 2016.
- [6] Y. Zhong, T. Q. S. Quek, and W. Zhang, "Complementary networking for C-RAN: Spectrum efficiency, delay and system cost," *IEEE Trans. Wireless Commun.*, vol. 16, no. 7, pp. 4639–4653, Jul. 2017.
- [7] Y. Zhong, T. Q. S. Quek, and X. Ge, "Heterogeneous cellular networks with spatio-temporal traffic: Delay analysis and scheduling," *IEEE J. Sel. Areas Commun.*, vol. 35, no. 6, pp. 1373–1386, Jun. 2017.
- [8] V. N. Ha, L. B. Le, and N. D. Dao, "Coordinated multipoint transmission design for Cloud-RANs with limited fronthaul capacity constraints," *IEEE Trans. Veh. Technol.*, vol. 65, no. 9, pp. 7432–7447, Sep. 2016.
- [9] S. Wang and C. Ran, "Rethinking cellular network planning and optimization," *IEEE Wireless Commun.*, vol. 23, no. 2, pp. 118–125, Apr. 2016.
- [10] K. Tutschku and P. Tran-Gia, "Spatial traffic estimation and characterization for mobile communication network design," *IEEE J. Sel. Areas Commun.*, vol. 16, no. 5, pp. 804–811, Jun. 1998.
- [11] W. Zhao, S. Wang, C. Wang, and X. Wu, "Approximation algorithms for cell planning in heterogeneous networks," *IEEE Trans. Veh. Technol.*, vol. 66, no. 2, pp. 1561–1572, Feb. 2017.
- [12] S. Wang, W. Zhao, and C. Wang, "Budgeted cell planning for cellular networks with small cells," *IEEE Trans. Veh. Technol.*, vol. 64, no. 10, pp. 4797–4806, Oct. 2015.
- [13] J. G. Carlsson and R. Devulapalli, "Shadow prices in territory division," *Netw. Spatial Econ.*, vol. 16, no. 3, pp. 893–931, Sep. 2016.
- [14] C. Ran, S. Wang, and C. Wang, "Cellular networks planning: A workload balancing perspective," *Comput. Netw.*, vol. 84, no. 19, pp. 64–75, Jun. 2015.
- [15] S. Wang, M. Ge, and C. Wang, "Efficient resource allocation for cognitive radio networks with cooperative relays," *IEEE J. Sel. Areas Commun.*, vol. 31, no. 11, pp. 2432–2441, Nov. 2013.
- [16] J. G. Carlsson, "Dividing a territory among several vehicles," *INFORMS J. Comput.*, vol. 24, no. 4, pp. 565–577, Aug. 2012.