Budgeted Cell Planning for Cellular Networks With Small Cells

Shaowei Wang, *Senior Member, IEEE*, Wentao Zhao, *NA , IEEE*, and Chonggang Wang, *Senior Member, IEEE*

***Abstract*—Heterogeneous networks (HetNets), where small cells are deployed within the coverage of macrocells, can increase capacity and enhance coverage of a cellular system. Compared with macrocells, small cells have superiority in terms of low installation and operation cost because of their small physical size and low transmission power. Hence, HetNets are deemed as a cost-effective way to address the everlasting radio spectrum crisis. On the other hand, HetNets also unfold a new paradigm from the viewpoint of cellular network planning. In this paper, we study the *budgeted cell planning problem* in HetNets, where our aim is to maximize the number of traffic demand nodes whose required rates are fully satisfied with a given budget. Our optimization task is challenging, and the formulated problem is hard to solve because of constraints in practical cellular systems, including power limitation, available bandwidth, and traffic requirements. We develop an approximation algorithm that yields an $(e-1)/2e$ fraction of the optimum, which not only provides quality-guaranteed solutions to the *budgeted cel planning problem* but sheds useful lights on how to plan a HetNet with limited capital expenditure as well. Preliminary numerical results show that small cells can improve the capacity of a cellular system significantly if they are properly planned.**

***Index Terms*—Approximation algorithm, budgeted cell planning, heterogeneous network (HetNet).**

I. Introduction

THE ever-growing service demand has been stimulating researchers in academia and industry to exploit advanced techniques to use the scarce radio spectrum as efficiently as possible for wireless communications, among which cell planning has been of crucial importance since the very commercialization of cellular mobile communications. Basically, planning a cellular network involves both capacity and coverage design. Conventional cell planning emphasizes site acquisition, frequency assignment, etc. There are usually two counter objectives for cell planning: minimizing the total deployment cost while satisfying the traffic requirements of users [also referred to as traffic demand points (TDPs)] and maximizing the number of fully satisfied TDPs with a given deployment budget.

Manuscript received May 13, 2014; revised September 5, 2014; accepted October 18, 2014. Date of publication November 3, 2014; date of current version October 13, 2015. This work was supported by the JiangsuSF (BK2011051). The review of this paper was coordinated by Prof. F. R. Yu.

S. Wang and W. Zhao are with the School of Electronic Science and Engineering, Nanjing University, Nanjing 210023, China (e-mail: wangsw@nju.edu.cn; zhaowt@smail.nju.edu.cn).

C. Wang is with InterDigital Communications, King of Prussia, PA 19406 USA (e-mail: cgwang@ieee.org).

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.2014.2366783

For the second-generation (2G) cellular network, e.g., the Global System for Mobile communications system that employs frequency-division multiple access with a time-division component, cell planning usually consists of two separate steps that correspond to coverage design and capacity planning. First, cell sites are selected from candidate locations and configured with a series of radio parameters, such as transmission power, antenna type, azimuth, tilt, and height, to provide strong enough signal strength for coverage of the service region [1]–[7]; second, radio frequency is grouped to provide a sufficient number of channels per cell to serve the users in the service region. More importantly, frequency grouping is also the key technique that suppresses the mutual interference among adjacent cells so that users can obtain high enough signal-to-interference-plus-noise ratio (SINR) [3]–[7].

Code-division multiple access (CDMA) is adopted in the third-generation (3G) cellular network, e.g., the Universal Mobile Telecommunications System. Frequency grouping is no longer necessary because each cell can use all radio spectra in CDMA systems. Thus, cell planning in such cellular systems always focuses on how to select a set of sites from candidate sites to install cells and configure their radio parameters to keep the SINRs of users above a predefined threshold [8]–[12]. Since the SINR of the user in the CDMA system is related to the relative power of its association cell and the neighboring cells, power control is important in this situation. Power control is the key technique to address the near-far problem in the CDMA system that can seriously degenerate the performance of the victim user.

For the fourth-generation (4G) cellular network and beyond, such as the Long-Term Evolution Advanced (LTE-A) system, orthogonal frequency-division multiplexing is deemed as an ideal air interface because of its potential to allocate radio resource flexibly, which cannot only equip cellular network with the ability of satisfying diverse users' demands but also exploit diversity gains among users to improve the system throughput [13], [14]. Furthermore, the LTE-A system employs a HetNet infrastructure, where small cells, namely, micro/picocells, femtocells, and relays, are introduced to underlay macrocells. These small cells can fill the coverage blind spots of macrocells and improve the capacity of the cellular system dramatically, making the HetNet a promising solution to the radio spectrum crisis [15]. Moreover, small cells can offer flexible site acquisitions due to their low transmission power and small physical size, resulting in the HetNet usually being more cost efficient than the conventional macro-only cellular system in the case of satisfying equal traffic demand.

However, the HetNet also brings challenges to cell planning and other interference-related issues for the cellular system. For either the second or the third cellular system, approximately equal transmission power of each cell is always preferred because it is the most effective and efficient way to avoid interference among adjacent cells. However, it is not the case for the HetNet, where macrocells and small cells are equipped with different radio parameters, particularly different transmission power. For example, the maximum transmission power of a macrocell is 46 dBm, whereas a small cell can transmit with a maximum power of only 30 dBm. Thus, conventional cell planning that can suppress interference between adjacent cells cannot work well in the HetNet scenario. As a result, coordinated multipoint (CoMP) transmission and reception, intercell interference coordination (ICIC), and other advanced signal processing techniques are required to address the interference management issues in HetNets and have been intensively investigated in both academia and industry [16]–[18] recently.

However, to the authors' knowledge, jointly planning the macrocells and small cells in the HetNet is not extensively investigated except for some preliminary results reported recently [19]–[21]. As discussed in [22] and [23], as long as a cellular network is interference limited and a mobile terminal connects to the cell that provided the strongest signals, statistically speaking, the coverage probability of the cellular system is independent of the number and the density of different types of cells and their relative power levels. That is to say, cell planning is also an efficient way to exploit the potential of the HetNet even if there exists heavy interference between macrocells and small cells. A well-planned HetNet can provide capacity as large as possible with limited available radio spectrum and, at the same time, relieve the pressure of exploiting complex signal processing techniques to address the interference between small cells and macrocells. Furthermore, the total number of cells can be reduced without deteriorating the performance of the cellular system. In summary, the HetNet can also be deployed in a cost-efficient way for given capital expenditure, which is always the key object of service providers and also the motivation of this research.

In this paper, we investigate the cell planning problem in the HetNet, where our optimization objective is to maximize the number of fully satisfied TDPs in a cellular system while keeping the total deployment cost below a given budget. Both macrocells and small cells are considered in our network model, as well as transmission power limitation, available bandwidth, and rate requirements of TDPs. We show that the formulated *budgeted cell planning problem* is NP-hard and address it with a divide-and-rule strategy. First, we work out the optimal bandwidth and power allocation for a given cell that can satisfy a set of TDPs. We show that it is a convex optimization problem and can be efficiently solved by using Karush–Kuhn–Tucker (KKT) conditions. Second, a $1/2$-approximation algorithm is developed to yield promising solutions to the problem of maximizing the number of fully satisfied TDPs for a given set of cells. Finally, we propose an $(e-1)/2e$-approximation algorithm for the *budgeted cell planning problem*.

The main contributions of this work are twofold: First, we develop an optimization framework for cell planning in the HetNet, which covers practical application scenarios in the next-generation cellular networks; second, we design an approximation algorithm to address the intractable cell planning problem, which is different from other heuristic methods that have no worse-case performance guarantees. Moreover, the approximation ratio of our proposed algorithm is also attractive, indicating that it could be used to plan a practical HetNet, which is confirmed by preliminary numerical results. In brief, our proposed cell planning scheme sheds some insights on how to deploy a heterogeneous cellular network at the lowest possible cost.

TABLE I
NOTATIONS

Symbol	Semantics
$\mathcal{N}_m$	The set of potential macro cells
$\mathcal{N}_s$	The set of potential small cells
$\mathcal{N}$	The set of potential cells
N	The number of all potential cells
c_n	The installation cost of BS n
$P_{n,max}$	The maximum transmission power of BS n
B	The total bandwidth
$\mathcal{K}$	The set of TDPs
K	The number of TDPs
$R_{k,min}$	The minimum traffic requirement of TDP k
$b_{k,n}$	The bandwidth of BS n allocated to TDP k
$p_{k,n}$	The power of BS n allocated to TDP k
$g_{k,n}$	The channel gain from BS n to TDP k
Γ	The SNR gap
N_0	The power spectral density of AWGN
R_k	The satisfied transmission rate of TDP k
C	The installation budget
$y_{k,n}$	The TDP assignment index
z_n	The selection variable of cell n
Y_k	The selection variable of TDP k

The remainder of this paper is organized as follows. In Section II, we present the network model and formulate the *budgeted cell planning problem*. Related works are also discussed in this section. In Section III, our proposed algorithms are detailed. To make the idea easy to follow, three subsections are organized in a gradual manner. In Section IV, fundamental numerical results are reported, as well as discussions about the interference management issue. Conclusions are drawn in Section V.

II. NETWORK MODEL AND PROBLEM FORMULATION

A. Network Model

Some frequently used notations are listed in Table I.

Consider an area served by a cellular system. The candidate sites to deploy macrocells and small cells are denoted as $\mathcal{N}_m$ and $\mathcal{N}_s$, respectively.[1] Denote $\mathcal{N} = \mathcal{N}_m \cup \mathcal{N}_s = \{1, 2, \ldots, N\}$, for each cell $n \in \mathcal{N}$, c_n is the installation cost, and $P_{n,\max}$ is the maximum transmission power. The maximum available bandwidth is B for both macrocells and small cells.

The set of TDPs is denoted as $\mathcal{K} = \{1, 2, \ldots, K\}$. Each TDP can be served by one cell. The rate requirement of TDP

[1]There are also small cells that are installed by users in an unplanned manner, which are usually referred to as femtocells. Such cells are not considered in this work.

$k \in \mathcal{K}$ is $R_{k,\min}$, where $R_{k,\min} > 0$. The channel gain between cell n and TDP k is $g_{k,n}$, which is determined by the path-loss propagation model, usually including antenna gain, path-loss exponent, and shadow fading. Denote the bandwidth and the power allocated to TDP k by cell n as $b_{k,n}$ and $p_{k,n}$, respectively, the achievable rate from cell n to TDP k can be calculated as

$$r_{k,n} = b_{k,n} \log_2 \left(1 + \frac{p_{k,n}|g_{k,n}|^2}{\Gamma N_0 b_{k,n}}\right) \tag{1}$$

where Γ is the SNR gap related to a given bit error rate (BER) for a specific modulation/demodulation scheme, e.g., $\Gamma = -\ln(5\text{BER})/1.6$ for an uncoded multilevel quadrature amplitude modulation system [24]. N_0 is the power spectral density of additive white Gaussian noise (AWGN). Denote $G_{k,n} = g_{k,n}^2/\Gamma N_0$, the rate of the TDP k is

$$R_K = \sum_{n\in\mathcal{N}} y_{k,n} b_{k,n} \log_2 \left(1 + \frac{p_{k,n}G_{k,n}}{b_{k,n}}\right) \tag{2}$$

where $y_{k,n}$ indicates whether the TDP k is served by cell n or not, that is

$$y_{k,n} = \begin{cases} 1, & \text{TDP } k \text{ is associated to cell } n \\ 0, & \text{otherwise.} \end{cases} \tag{3}$$

B. Problem Formulation

The *budgeted cell planning problem* is to select a subset from $\mathcal{N}$ to maximize the number of fully satisfied TDPs, while the total installation cost cannot exceed a given budget C. Define two index variables z_n and Y_k, then

$$z_n = \begin{cases} 1, & \text{cell } n \text{ is selected} \\ 0, & \text{otherwise} \end{cases} \tag{4}$$

$$Y_k = \begin{cases} 1, & \text{TDP } k \text{ is satisfied} \\ 0, & \text{otherwise.} \end{cases} \tag{5}$$

Our optimization task can be formulated as

$$\begin{aligned}
\max_{\substack{z_n, y_{k,n}, Y_k \\ b_{k,n}, p_{k,n}}} \quad & \sum_{k\in\mathcal{K}} Y_k \\
\text{s.t.} \quad C_1: \quad & \sum_{k\in\mathcal{K}} b_{k,n} \le z_n B \qquad \forall\, n \in \mathcal{N} \\
C_2: \quad & \sum_{k\in\mathcal{K}} p_{k,n} \le z_n P_{n,\max} \qquad \forall\, n \in \mathcal{N} \\
C_3: \quad & Y_k R_{k,\min} \le R_k \qquad \forall\, k \in \mathcal{K} \\
C_4: \quad & 0 \le b_{k,n} \le y_{k,n} B \qquad \forall\, k \in \mathcal{K}, n \in \mathcal{N} \\
C_5: \quad & 0 \le p_{k,n} \le y_{k,n} P_{n,\max} \qquad \forall\, k \in \mathcal{K}, n \in \mathcal{N} \\
C_6: \quad & z_n, y_{k,n}, Y_k \in \{0, 1\} \qquad \forall\, k \in \mathcal{K}, n \in \mathcal{N} \\
C_7: \quad & \sum_{n\in\mathcal{N}} c_n z_n \le C \\
C_8: \quad & \sum_{n\in\mathcal{N}} y_{k,n} = Y_k \qquad \forall\, k \in \mathcal{K}.
\end{aligned} \tag{6}$$

C_1 and C_2 indicate that the available bandwidth and power of cell n are limited to B and $P_{n,\max}$, respectively. C_3 is the rate requirement of the TDP k served by cell n. C_4 and C_5 are the bandwidth and power constraints of each TDP, that is, if $y_{k,n} = 1$, we have $b_{k,n} \le B$ and $p_{k,n} \le P_{n,\max}$; otherwise, $b_{k,n} = 0$, and $p_{k,n} = 0$. C_7 indicates that the total installation cost cannot exceed a given budget C. C_8 shows that each TDP can be served by only one cell.

C. Related Work

Cell planning in the 2G/3G cellular systems is usually formulated as a capacitated facility location problem. Since such kinds of problems are generally NP-hard, heuristic methods, such as tabu search, simulated annealing, and evolutionary algorithms, are usually attractive because they can produce feasible and even promising solutions with reasonable complexity. Many heuristic methods have been proposed to handle the cell planning problem for the 2G/3G cellular systems in the literature [5]–[12]. Although the proposed algorithms may work well for some practical cases as shown in these works, few of them can guarantee the performance of the proposed planning algorithms in the worst case. It is inconclusive to encourage service providers to apply them for a practical cell planning. In this paper, we try to develop approximation algorithms that hopefully give guaranteed performance in the worst case.

Equation (6) is a mixed-integer programming problem, which is NP-hard [25]. The general form of (6) is the *budgeted maximum coverage problem* [26]: Given a collection of sets $\mathcal{S} = \{S_1, S_2, \ldots, S_m\}$ with associated costs $c_i, i = 1, 2, \ldots, m$, which is defined over a domain of elements $X = \{x_1, x_2, \ldots, x_n\}$ with associated weights $w_j, j = 1, 2, \ldots, n$, the objective is to find a subset $\mathcal{S}_1$ of $\mathcal{S}$ to maximize the total weight of elements of the subset under the constraint that the total cost of $\mathcal{S}_1$ does not exceed a given budget. If each set has unit cost, it is also called a *weighted maximum coverage problem* [27]. Particularly, if $w_j = 1,\ \ j = 1, 2, \ldots, n$, it is also called a *max-k-cover problem* [28], [29]. Given a positive integer k, the *weighted maximum coverage problem* is to find a subset $\mathcal{S}_2$ with $|\mathcal{S}_2| = k$ to maximize the total weight of the elements, and the *max-k-cover problem* is to find a subset $\mathcal{S}_3$ with $|\mathcal{S}_3| = k$ to cover the largest number of elements.

The *weighted maximum coverage problem*, the *budgeted maximum coverage problem*, and the *max-k-cover problem* are all NP-hard. A greedy algorithm that achieves an approximation ratio of $1 - 1/e$ for the *max-k-cover problem* is proposed [28]. In [29], Feige proved that the *max-k-cover problem* cannot be approximated within a ratio of $1 - 1/e + \epsilon$ for any $\epsilon > 0$ unless NP $\subseteq$ DTIME($n^{O(\log\log n)}$), which means that $1 - 1/e$ is the best approximation ratio for such kind of problems. In [27], a $1 - (1 - 1/\rho)^\rho$ approximation algorithm for the *weighted maximum coverage problem* is proposed, where $\rho = \max_{i=1,2,\ldots,m} |S_i|$. For small values of ρ, this approximation ratio is better than $1 - 1/e$. In [26], Khuller *et al.* developed two greedy-based algorithms for the *budgeted maximum coverage problem*, one of which achieves an approximation ratio of $1 - 1/\sqrt{e}$, whereas the other achieves an approximation ratio of $1 - 1/e$. They also proved that no approximation algorithm with a ratio better than $1 - 1/e$ exists for the *budgeted maximum coverage problem* unless NP $\subseteq$ DTIME($n^{O(\log\log n)}$).

Interestingly, the *generalized maximum coverage problem* is also akin to the *budgeted maximum coverage problem*. For the *generalized maximum coverage problem*, the weight and the cost of an element are dependent on which set covers it. The objective is also to select sets to maximize the weighted sum of elements in these sets, while the total cost cannot exceed a given budget. In [30], two approximation algorithms for solving the *generalized maximum coverage problem* are proposed, which achieve approximation ratios of $(e-1)/(2e-1)-\epsilon$ and $(e-1)/e-\epsilon$ for every $\epsilon > 0$.

It is worth noticing that in [31], the *budgeted cell planning problem* is formulated as a *capacitated maximum coverage problem*, which aims to find a subset of cells to cover the largest number of TDPs while the capacity of each cell is limited, as well as the total deployment cost. The formulated problem is NP-hard, and an $(e-1)/(3e-1)$-approximation algorithm is presented for a special case. However, link gains between TDPs and base stations are not considered, resulting in an unreasonable assumption that a TDP always consumes equal radio resources whichever cell serves it. As a matter of fact, a TDP should be served by the cell that can provide the highest SINR so as to consume radio resources as few as possible to provide the required transmission rate from the viewpoint of the system. Particularly for the LTE-A system, the transmission rate can be adjusted based on the link gain between the base station and the user served by it. In other words, only considering the link gain can the number of TDPs be maximized for the cell planning in HetNets.

III. $(e-1)/2e$-Approximation Algorithm

To solve the *budgeted cell planning problem*, we first answer the following two questions: Given a set of TDPs and a cell, is it possible for the cell to satisfy all the TDPs' rate requirements with its bandwidth and power budgets? Given a set of cells, how can we maximize the number of TDPs whose required rates can be fully satisfied? We call the former the *bandwidth and power allocation problem*, the optimal solution to which can be obtained by using KKT conditions and numerical algorithms. The latter is called the *TDP assignment problem*, and we develop a 1/2-approximation algorithm for this problem. Finally, we generalize the method proposed in [26] and develop an $(e-1)/2e$-approximation algorithm for the *budgeted cell planning problem*.

A. Bandwidth and Power Allocation Problem

Denote the bandwidth and power budgets of the cell n as B and $P_{n,\max}$, respectively, the set of TDPs cell n to be served by the cell n is $\mathcal{K}_n$, the *bandwidth and power allocation problem* can be directly described as follows:

$$\begin{aligned} &\text{find} \quad b_{k,n}, p_{k,n} \\ \text{s.t.} \quad &C_1: \quad \sum_{k\in\mathcal{K}_n} b_{k,n} \le B \\ &C_2: \quad \sum_{k\in\mathcal{K}_n} p_{k,n} \le P_{n,\max} \\ &C_3: \quad r_{k,n} = R_{k,\min}, \qquad \forall\, k \in \mathcal{K}_n \\ &C_4: \quad b_{k,n} \ge 0,\ p_{k,n} \ge 0, \qquad \forall\, k \in \mathcal{K}_n. \end{aligned} \tag{7}$$

If we can find a feasible solution to (7), we claim that all TDPs in $\mathcal{K}_n$ can be covered by cell n. However, (7) cannot be solved straightforward as can be seen from the objective function. We can give an alternate form of (7). Assume that all available bandwidth B is consumed to serve the TDPs in $\mathcal{K}_n$, if the required power to satisfy the TDPs' rate requirements does not exceed the power budget $P_{n,\max}$, cell n can certainly cover $\mathcal{K}_n$. Correspondingly, we define the following optimization problem to achieve the given ideas:

$$\begin{aligned} \min_{b_{k,n},p_{k,n}} \quad &\sum_{k\in\mathcal{K}_n} p_{k,n} \\ \text{s.t.} \quad &C_1: \quad \sum_{k\in\mathcal{K}_n} b_{k,n} = B \\ &C_2: \quad r_{k,n} = R_{k,\min} \qquad \forall\, k \in \mathcal{K}_n \\ &C_3: \quad b_{k,n} \ge 0, p_{k,n} \ge 0 \qquad \forall\, k \in \mathcal{K}_n. \end{aligned} \tag{8}$$

Based on C_2 in (8) and the definition of $r_{k,n}$, we have

$$p_{k,n} = \frac{b_{k,n}}{G_{k,n}} \cdot \left(2^{\frac{R_{k,min}}{b_{k,n}}} - 1\right). \tag{9}$$

Substituting (9) into (8), we have

$$\begin{aligned} \min_{b_{k,n}} \quad &\sum_{k\in\mathcal{K}_n} \frac{b_{k,n}}{G_{k,n}} \cdot \left(2^{\frac{R_{k,\min}}{b_{k,n}}} - 1\right) \\ \text{s.t.} \quad &C_1: \quad \sum_{k\in\mathcal{K}_n} b_{k,n} = B \\ &C_2: \quad b_{k,n} \ge 0 \qquad \forall\, k \in \mathcal{K}_n. \end{aligned} \tag{10}$$

Equation (10) defines a convex problem because the objective function is convex, and all the constraints are affine. Hence, it can be solved by standard convex optimization techniques. The Lagrangian of (10) is

$$J = \sum_{k\in\mathcal{K}_n} \frac{b_{k,n}}{G_{k,n}} \cdot \left(2^{\frac{R_{k,\min}}{b_{k,n}}} - 1\right) + \lambda\left(\sum_{k\in\mathcal{K}_n} b_{k,n} - B\right) - \sum_{k\in\mathcal{K}_n} \mu_{k,n} b_{k,n}$$

where λ and $\mu_{k,n}$ are the Lagrange multipliers. Let $b^*_{k,n}$ and λ^*, $\mu^*_{k,n}$ be the primal and dual optimal points with zero duality gap [32]. By using KKT conditions [32], we get the following equations:

$$\lambda^* = -\frac{1}{G_{k,n}}\left[\left(1 - \frac{R_{k,\min}\ln 2}{b^*_{k,n}}\right) 2^{\frac{R_{k,\min}}{b^*_{k,n}}} - 1\right] \tag{11}$$

$$\sum_{k\in\mathcal{K}_n} b^*_{k,n} = B \tag{12}$$

$$\mu^*_{k,n} = 0, b^*_{k,n} > 0. \tag{13}$$

Equation (13) is intuitive because the rate requirement of the TDP k cannot be satisfied with consuming no power. $b^*_{k,n}$ and λ^* can be solved via a bisection method. The outline of the algorithm (see Algorithm 1) for the *bandwidth and power allocation problem* is described in Table II, where ϵ is a tolerance, and Δ is a suitably large number. Let $P_n(\mathcal{K}_n) = \sum_{k\in\mathcal{K}_n} p^*_{k,n}$

TABLE II
ALGORITHM 1: BANDWIDTH AND POWER ALLOCATION

Initialize: $l = 0$, $\lambda^{(l)} = 0$, $\lambda_{min} = 0$, $\lambda_{max} = \Delta$;
repeat
    $l = l + 1$;
    $\lambda^{(l)} = (\lambda_{max} + \lambda_{min})/2$;
    for $k \in \mathcal{K}_n$
        Calculate $b_{k,n}$ that satisfies Eq.(10);
        $b_{k,n} = max\{0, b_{k,n}\}$;
    end for
    if $\sum_{k\in\mathcal{K}_n} b_{k,n} > B$;
        $\lambda_{min} = \lambda^{(l)}$;
    else
        $\lambda_{max} = \lambda^{(l)}$;
    end if
until $|\lambda^{(l)} - \lambda^{(l-1)}| \le \epsilon$
for $k \in \mathcal{K}_n$
    $b^*_{k,n} = b_{k,n}$;
    Calculate $p^*_{k,n}$ using Eq.(9);
end for
return $b^*_{k,n}, p^*_{k,n}, \sum_{k\in\mathcal{K}_n} p^*_{k,n}$.

be the optimal value of (8). If $P_n(\mathcal{K}_n)$ does not exceed $P_{n,\max}$, the rate requirements of all TDPs in $\mathcal{K}_n$ can be satisfied by cell n, and we claim that cell n can cover $\mathcal{K}_n$; otherwise, cell n cannot serve all TDPs in $\mathcal{K}_n$.

B. TDP Assignment Problem

The *TDP assignment problem* is to maximize the number of fully satisfied TDPs by a given set of cells $\mathcal{N}' \subseteq \mathcal{N}$ under their bandwidth and power budgets. The mathematical form is

$$\begin{aligned}
\max_{\substack{y_{k,n},Y_k\\ b_{k,n},p_{k,n}}} \quad & \sum_{k\in\mathcal{K}} Y_k \\
\text{s.t.} \quad C_1: \quad & \sum_{k\in\mathcal{K}} b_{k,n} \le B \qquad \forall n \in \mathcal{N}' \\
C_2: \quad & \sum_{k\in\mathcal{K}} p_{k,n} \le P_{n,\max} \qquad \forall n \in \mathcal{N}' \\
C_3: \quad & \sum_{n\in\mathcal{N}'} r_{k,n} = Y_k R_{k,\min} \\
C_4: \quad & y_{k,n} B \ge b_{k,n} \ge 0 \qquad \forall k \in \mathcal{K} \ \forall n \in \mathcal{N}' \\
C_5: \quad & y_{k,n} P_{n,\max} \ge p_{k,n} \ge 0 \qquad \forall k \in \mathcal{K} \ \forall n \in \mathcal{N}' \\
C_6: \quad & y_{k,n}, Y_k \in \{0,1\} \qquad \forall k \in \mathcal{K}, n \in \mathcal{N}' \\
C_7: \quad & \sum_{n\in\mathcal{N}'} y_{k,n} = Y_k \qquad \forall k \in \mathcal{K}.
\end{aligned} \tag{14}$$

Equation (14) is a mixed-binary-integer programming problem, which is NP-hard. We develop a 1/2-approximation algorithm to obtain promising solutions, which is described in Table III (see Algorithm 2). Denote $\mathcal{K}'_n$ as the set of TDPs assigned to cell n by Algorithm 2. First, we initialize $\mathcal{K}'_n = \emptyset$ for each cell $n \in \mathcal{N}'$. Moreover, we define a searching index matrix $I(k,n)$ for each $k \in \mathcal{K}$, $n \in \mathcal{N}'$, where $I(k,n) = 0$ indicates that the pair of TDP k and cell n has been checked; otherwise, $I(k,n) = 1$. Then, we calculate $P_n(\{k\})$ for each $k \in \mathcal{K}, n \in \mathcal{N}'$. $P_n(\{k\})$ can be simply calculated by substituting $b_{k,n} = B$ into (9). At each iteration, we check all entries of

TABLE III
ALGORITHM 2: TDP ASSIGNMENT

Initialize: $\mathcal{K}'_n = \emptyset$ for all $n \in \mathcal{N}'$, $I(k,n) = 1$ for each $k \in \mathcal{K}$, $n \in \mathcal{N}'$;
Calculate $P_n(\{k\})$ for each $k \in \mathcal{K}$, $n \in \mathcal{N}'$;
repeat
    $(k^*, n^*) = \arg\min_{(k,n):I(k,n)=1} P_n(\{k\})$;
    if $P_{n^*}(\mathcal{K}'_{n^*} \cup \{k^*\}) \le P_{n^*,max}$
        $\mathcal{K}'_{n^*} \leftarrow \mathcal{K}'_{n^*} \cup \{k^*\}$;
        set $I(k^*, n) = 0$ for each $n \in \mathcal{N}'$;
    else
        set $I(k, n^*) = 0$ for each $k \in \mathcal{K}$;
    end if
    set $I(k^*, n^*) = 0$;
until $I(k,n) = 0$ for all $k \in \mathcal{K}$, $n \in \mathcal{N}'$
return $\mathcal{K}'_n$

1's in matrix $I(k,n)$ and find out the $I(k^*, n^*)$ that corresponds to the minimum power consumption $P_{n^*}(\{k^*\})$. Moreover, $P_{n^*}(\mathcal{K}'_{n^*} \cup \{k^*\})$ can be worked out by using Algorithm 1. If the required power of cell n^* is less than its transmission power budget, we assign TDP k^* to cell n^* and set $I(k^*, n) = 0$ for each $n \in \mathcal{N}'$. Otherwise, cell n^* cannot satisfy any remaining TDP since k^* can use the same allocated bandwidth with less power consumption compared with all the remaining TDPs. Hence, we can safely set $I(k, n^*) = 0$ for each $k \in \mathcal{K}$. The check procedure terminates when all entries of the index matrix are zeros.

To prove the approximation ratio of Algorithm 2, we need the following fact.

Fact 1: Given positive numbers A, a_1, a_2 such that $a_1 - 1 \ge A \cdot (a_2 - 1)$, where $a_1 > 1, a_2 > 1$, and $A > 0$, then

$$a_1^n - 1 \ge A \cdot (a_2^n - 1) \qquad \forall n \ge 1, n \in \mathbf{R}.$$

Lemma 1: For cell n and two sets of TDPs M_1 and M_2, where $|M_1| = |M_2|$. If $P_n(\{k_1\}) \ge P_n(\{k_2\}) \ \forall k_1 \in M_1$, $k_2 \in M_2$, then $P_n(M_1) \ge P_n(M_2)$.

Proof: Let $p^*_{k,n}$ and $b^*_{k,n}$ be the optimal power and bandwidth allocation for $P_n(M_1)$. Since $P_n(\{k_1\}) \ge P_n(\{k_2\})$ $\forall k_1 \in M_1, k_2 \in M_2$, we have

$$\frac{1}{G_{k_1,n}} \cdot \left(2^{R_{k_1,\min}/B} - 1\right) \ge \frac{1}{G_{k_1,n}} \cdot \left(2^{R_{k_2,\min}/B} - 1\right).$$

The key idea of the proof is that the power consumption to serve the TDPs in M_2 is less than $P_n(M_1)$, even consuming the same bandwidth allocation $b^*_{k,n}$. For each $k_1 \in M_1$, $k_2 \in M_2$, we can obtain

$$\begin{aligned}
p^*_{k_2,n} &= \frac{b^*_{k_1,n}}{G_{k_1,n}} \cdot \left(2^{\frac{R_{k_1,\min}}{b^*_{k_1,n}}} - 1\right) \\
&= \frac{b^*_{k_1,n}}{G_{k_1,n}} \cdot \left(2^{\frac{R_{k_1,\min}}{B} \cdot \frac{B}{b^*_{k_1,n}}} - 1\right) \\
&\ge \frac{b^*_{k_1,n}}{G_{k_2,n}} \cdot \left(2^{\frac{R_{k_2,\min}}{B} \cdot \frac{B}{b^*_{k_1,n}}} - 1\right) \\
&= \frac{b^*_{k_1,n}}{G_{k_2,n}} \cdot \left(2^{\frac{R_{k_2,\min}}{b^*_{k_1,n}}} - 1\right)
\end{aligned}$$

where the inequality follows Fact 1 since $b^*_{k_1,n}$ is always less than B. We consider the same bandwidth allocation for the TDPs in M_2, that is, each TDP $k_2 \in M_2$ gets $b^*_{k_1,n}$ bandwidth. We have

$$\begin{aligned} P_n(M_1) &= \sum_{k_1 \in M_1} p^*_{k_1,n} \\ &\geq \sum_{k_2 \in M_2} \frac{b^*_{k_1,n}}{G_{k_2,n}} \cdot \left(2^{\frac{R_{k_2,\min}}{b^*_{k_1,n}}} - 1\right) \\ &\geq P_n(M_2). \end{aligned}$$

■

According to Lemma 1, we can also obtain the following corollary.

Corollary 1: For cell n and two sets of TDPs M_1 and M_2. If $P_n(\{k_1\}) \geq P_n(\{k_2\})\ \forall\, k_1 \in M_1,\ k_2 \in M_2$, and $P_n(M_1) < P_n(M_2)$, then $|M_1| < |M_2|$.

Proof: The conclusion is intuitive because if $|M_1| \geq |M_2|$, we consider a set $M' \subseteq M_1$ of TDPs, where $|M'| = |M_2|$, and then

$$P_n(M_1) \geq P_n(M') \geq P_n(M_2).$$

Thus, we always have $|M_1| < |M_2|$. ■

Now, we can give the approximation ratio of Algorithm 2.

Theorem 1: Algorithm 2 is a $1/2$-approximation for the TDP assignment problem.

Proof: Let $\mathcal{K}^{OPT}$ be the set of fully satisfied TDPs in the optimal solution of (14). For each cell $n \in \mathcal{N}'$, let $\mathcal{K}^{OPT}_n$ be the set of TDPs satisfied by cell n in the optimal solution. Let $\mathcal{K}'$ be the set of selected TDPs chosen by Algorithm 2 and $\mathcal{K}'_n$ be the set of TDPs assigned to cell n in $\mathcal{K}'$.

Consider cell $n \in \mathcal{N}'$. According to the TDP assignment procedure shown in Algorithm 2, we always have

$$P_n\{k_2\} \geq P_n\{k_1\} \qquad \forall\, k_1 \in \mathcal{K}'_n, k_2 \in \mathcal{K}^{OPT}_n \backslash \mathcal{K}'. \tag{15}$$

Moreover, we have

$$P_n\left(\mathcal{K}'_n \cup \{k_2\}\right) > P_{n,\max} \tag{16}$$

for each $k_2 \in \mathcal{K}^{OPT}_n \backslash \mathcal{K}'$; otherwise, TDP k_2 would be assigned to cell n according to Algorithm 2. At the same time, since the TDPs in $\mathcal{K}^{OPT}_n$ are assigned to cell n in the optimal solution, we also have

$$P_{n,\max} \geq P_n\left(\mathcal{K}^{OPT}_n \backslash \mathcal{K}'\right). \tag{17}$$

According to (16) and (17), we can obtain

$$P_n\left(\mathcal{K}'_n \cup \{k_2\}\right) > P_n\left(\mathcal{K}^{OPT}_n \backslash \mathcal{K}'\right). \tag{18}$$

Based on (15), (18), and Corollary 1, we have

$$\left|\mathcal{K}'_n \cup \{k_2\}\right| > \left|\mathcal{K}^{OPT}_n \backslash \mathcal{K}'\right|.$$

Thus

$$\left|\mathcal{K}'_n\right| \geq \left|\mathcal{K}^{OPT}_n \backslash \mathcal{K}'\right|.$$

TABLE IV
ALGORITHM 3: CELL SELECTION

```
Initialize: N' = ∅, K' = ∅.
for all H ⊆ N, |H| < 3, ∑_{n∈H} c_n ≤ C
    Calculate K'_H using Algorithm 2;
    if |K'_H| > |K'|
        N' = H; K' = K'_H;
    endif
endfor
for all H ⊆ N, |H| = 3, ∑_{n∈H} c_n ≤ C
    G = N\H;
    repeat
        n = arg min_{n∈G}(c_n/|K'_{H∪{n}}| − |K'_H|);
        if c_n + ∑_{n'∈H} c_{n'} ≤ C
            H = H ∪ {n};
        endif
        G = G\{n};
    until G = ∅
    if |K'_H| > |K'|
        N' = H; K' = K'_H;
    endif
endfor
return N', K'
```

Moreover

$$\begin{aligned} 2 \cdot |\mathcal{K}'| &= |\mathcal{K}'| + \sum_{n \in \mathcal{N}'} |\mathcal{K}'_n| \\ &\geq |\mathcal{K}^{OPT} \cap \mathcal{K}'| + \sum_{n \in \mathcal{N}'} \left|\mathcal{K}^{OPT}_n \backslash \mathcal{K}'\right| \\ &= |\mathcal{K}^{OPT} \cap \mathcal{K}'| + |\mathcal{K}^{OPT} \backslash \mathcal{K}'|. \end{aligned}$$

Since $|\mathcal{K}^{OPT} \cap \mathcal{K}'| + |\mathcal{K}^{OPT} \backslash \mathcal{K}'| = |\mathcal{K}^{OPT}|$, we can get

$$|\mathcal{K}'| \geq \frac{1}{2}|\mathcal{K}^{OPT}|.$$

■

C. Cell Selection

Based on the discussions of the *bandwidth and power allocation problem* and the *TDP assignment problem*, we can develop an approximation algorithm to solve the *budgeted cell planning problem* defined by (6), which is described in Table IV, where $\mathcal{K}'_H$ denoted the set of TDPs selected by Algorithm 2 for a given set of cells $H \subseteq \mathcal{N}$.

The key idea of our proposed algorithm is: The cell with the least average cost to serve newly added TDPs has the priority to be selected unless such an addition violates the budget constraint C7 in (6). At the beginning of the cell selection procedure, all possible sets, which include less than three cells, are tested, based on which we can test all possible combinations of three cells. For each combination, the inner loop in Table IV is performed to select the cell with the least average cost to serve newly added TDPs if the total cost does not exceed the budget C until all cells are checked. Both the set of satisfied TDPs and the corresponding set of cells that serve these TDPs are updated after each outer loop.

To prove the approximation ratio of Algorithm 3, we need the following mathematical facts:

Fact 2: Given positive numbers $a_1, \ldots, a_n$ and $b_1, \ldots, b_n$, then

$$\max_{i=1,\ldots,n} \frac{a_i}{b_i} \geq \frac{\sum_{i=1}^{n} a_i}{\sum_{i=1}^{n} b_i}.$$

Fact 3: Given positive numbers $a_1, \ldots, a_n, A$, if $\sum_{i=1}^{n} a_i = A$, then

$$\prod_{i=1}^{n} \left(1 - \frac{a_i}{A}\right) \leq \left(1 - \frac{1}{n}\right)^n.$$

For an instance of cell planning, let $z_n^*,\ y_{k,n}^*,\ Y_k^*,\ b_{k,n}^*,\ p_{k,n}^*$ be the optimal solution of (6). Denote $\mathcal{N}^*$ as the corresponding set of selected cells and $\mathcal{K}^*$ as the set of TDPs served by $\mathcal{N}^*$, respectively. Let $\mathcal{K}'_{H,n}$ be the set of TDPs assigned to cell n, $n \in H$. Since Algorithm 2 achieves a 1/2-approximation ratio, we have

$$|\mathcal{K}'_{\mathcal{N}^*}| \geq \frac{1}{2}|\mathcal{K}^*|.$$

Without loss of generality, we assume that $|\mathcal{N}^*| > 3$ (otherwise, Algorithm 3 must find the optimal solution), and it will cover at least $(1/2)|\mathcal{K}^*|$ TDPs. Order the cells in $\mathcal{N}^*$ by selecting the cells in $\mathcal{N}^*$ at each step that covers the maximum number of newly covered TDPs. Since all possible combinations of selecting three cells are tested by Algorithm 3, we consider the first three cells in this order as shown in line 8 of Algorithm 3. r is the number of iterations executed by Algorithm 3 when the first cell from $\mathcal{N}^*$ is checked but still not added to $\mathcal{N}'$ since such an addition would excess the budget C. L is the number of cells selected in the first r iterations, and n_l is the lth selected cell, $l = 1, 2, \ldots, L$, where n_1, n_2, and n_3 are the first three cells that are checked in line 8 of Algorithm 3. n_{L+1} denotes the cell checked at the rth iteration. l denotes the iteration index when cell n_l is checked. $\mathcal{N}'_l = \bigcup_{j=1}^{l} \{n_j\}$ is the set of all selected cells for $l = 1, 2, \ldots, L+1$, and $C' = C - \sum_{j=1}^{3} c_{n_j}$.

Lemma 2: After the lth iteration, $l = 4, \ldots, L+1$, the following equation holds:

$$\sum_{n \in \mathcal{N}^* \setminus \mathcal{N}'_l} \left(\left|\mathcal{K}'_{\mathcal{N}'_l \cup \{n\}}\right| - \left|\mathcal{K}'_{\mathcal{N}'_l}\right| \right) \geq |\mathcal{K}'_{\mathcal{N}^*}| - \left|\mathcal{K}'_{\mathcal{N}'_l}\right|.$$

Proof: Since each cell $n \in \mathcal{N}^*$ can cover $|\mathcal{K}'_{\mathcal{N}^*,n}|$ TDPs in the optimal solution, $|\mathcal{K}'_{\mathcal{N}'_l \cup \{n\}}| - |\mathcal{K}'_{\mathcal{N}'_l}|$ is no less than the number of the TDPs covered by cell n in $\mathcal{N}'$ but not covered by $\mathcal{N}'_l$. Thus, we always have

$$\left|\mathcal{K}'_{\mathcal{N}'_l \cup \{n\}}\right| - \left|\mathcal{K}'_{\mathcal{N}'_l}\right| \geq \left|\mathcal{K}'_{\mathcal{N}^*,n} \setminus \mathcal{K}'_{\mathcal{N}'_l}\right|$$

for each cell $n \in \mathcal{N}^* \setminus \mathcal{N}'_l$. We can obtain

$$\sum_{n \in \mathcal{N}^* \setminus \mathcal{N}'_l} \left|\mathcal{K}'_{\mathcal{N}'_l \cup \{n\}}\right| - \left|\mathcal{K}'_{\mathcal{N}'_l}\right| \geq \sum_{n \in \mathcal{N}^* \setminus \mathcal{N}'_l} \left|\mathcal{K}'_{\mathcal{N}^*,n} \setminus \mathcal{K}'_{\mathcal{N}'_l}\right| \geq |\mathcal{K}'_{\mathcal{N}^*}| - \left|\mathcal{K}'_{\mathcal{N}'_l}\right|$$

where the inequality follows from the fact that $|\mathcal{K}'_{\mathcal{N}^*}| - |\mathcal{K}'_{\mathcal{N}'_l}|$ is no more than the number of the TDPs covered by $\mathcal{N}^*$ but not covered by $\mathcal{N}'_l$. ■

The following two lemmas are generalized from the weighted budgeted maximum coverage problem [26].

Lemma 3: After the lth iteration, $l = 4, \ldots, L+1$, the following equation holds:

$$\left|\mathcal{K}'_{\mathcal{N}_l}\right| - \left|\mathcal{K}'_{\mathcal{N}'_{l-1}}\right| \geq \frac{c_{n_l}}{C'} \cdot \left(|\mathcal{K}'_{\mathcal{N}^*}| - \left|\mathcal{K}'_{\mathcal{N}'_{l-1}}\right| \right).$$

Proof: For each iteration l, $l \leq r$, the cell n_l satisfies

$$\begin{aligned} &\left|\mathcal{K}'_{\mathcal{N}'_l}\right| - \left|\mathcal{K}'_{\mathcal{N}'_{l-1}}\right| \\ &\geq \max_{n \in \mathcal{N}^* \setminus \mathcal{N}'_l} \frac{c_{n_l}}{c_n} \cdot \left(\left|\mathcal{K}'_{\mathcal{N}'_l \cup \{n\}}\right| - \left|\mathcal{K}'_{\mathcal{N}'_l}\right| \right) \\ &\geq \frac{c_{n_l}}{\sum_{n \in \mathcal{N}^* \setminus \mathcal{N}'} c_n} \cdot \sum_{n \in \mathcal{N}^* \setminus \mathcal{N}'_l} \left(\left|\mathcal{K}'_{\mathcal{N}'_l \cup \{n\}}\right| - \left|\mathcal{K}'_{\mathcal{N}'_l}\right| \right) \\ &\geq \frac{c_{n_l}}{C'} \cdot \left(|\mathcal{K}'_{\mathcal{N}^*}| - \left|\mathcal{K}'_{\mathcal{N}'_{l-1}}\right| \right). \end{aligned}$$

In the proof, the first inequality can be intuitively deduced from Algorithm 3, and the second inequality follows by Fact 2. The last inequality follows because $\sum_{n \in \mathcal{N}^* \setminus \mathcal{N}'} c_n$ is bounded by C'. ■

Lemma 4: After the lth iteration, $l = 4, \ldots, L+1$, the following equation holds:

$$\left|\mathcal{K}'_{\mathcal{N}'_l}\right| - \left|\mathcal{K}'_{\mathcal{N}'_2}\right| \geq \left[1 - \prod_{j=4}^{l} \left(1 - \frac{c_{n_j}}{C'}\right)\right] \cdot \left(|\mathcal{K}'_{\mathcal{N}^*}| - \left|\mathcal{K}'_{\mathcal{N}'_2}\right| \right).$$

Proof: Consider the 4th iteration, based on Lemma 3, we have

$$\left|\mathcal{K}'_{\mathcal{N}'_4}\right| - \left|\mathcal{K}'_{\mathcal{N}'_3}\right| \geq \frac{c_{n_4}}{C'} \cdot \left(|\mathcal{K}'_{\mathcal{N}^*}| - \left|\mathcal{K}'_{\mathcal{N}'_3}\right| \right).$$

Moreover, we can obtain

$$\begin{aligned} &\left|\mathcal{K}'_{\mathcal{N}'_l}\right| - \left|\mathcal{K}'_{\mathcal{N}'_3}\right| \\ &= \left|\mathcal{K}'_{\mathcal{N}'_{l-1}}\right| - \left|\mathcal{K}'_{\mathcal{N}'_3}\right| + \left(\left|\mathcal{K}'_{\mathcal{N}'_l}\right| - \left|\mathcal{K}'_{\mathcal{N}'_{l-1}}\right| \right) \\ &\geq \left|\mathcal{K}'_{\mathcal{N}'_{l-1}}\right| - \left|\mathcal{K}'_{\mathcal{N}'_3}\right| + \frac{c_{n_l}}{C'} \cdot \left(|\mathcal{K}'_{\mathcal{N}^*}| - \left|\mathcal{K}'_{\mathcal{N}'_{l-1}}\right| \right) \\ &= \left(1 - \frac{c_{n_l}}{C'}\right) \cdot \left(\left|\mathcal{K}'_{\mathcal{N}'_{l-1}}\right| - \left|\mathcal{K}'_{\mathcal{N}'_3}\right| \right) \\ &\quad + \frac{c_{n_l}}{C'} \cdot \left(|\mathcal{K}'_{\mathcal{N}^*}| - \left|\mathcal{K}'_{\mathcal{N}'_2}\right| \right) \\ &\geq \left(1 - \frac{c_{n_l}}{C'}\right) \cdot \left[1 - \prod_{j=4}^{l-1} \left(1 - \frac{c_{n_j}}{C'}\right)\right] \\ &\quad \cdot \left(|\mathcal{K}'_{\mathcal{N}^*}| - \left|\mathcal{K}'_{\mathcal{N}'_3}\right| \right) + \frac{c_{n_l}}{C'} \cdot \left(|\mathcal{K}'_{\mathcal{N}^*}| - \left|\mathcal{K}'_{\mathcal{N}'_3}\right| \right) \\ &= \left[1 - \prod_{j=4}^{l} \left(1 - \frac{c_{n_j}}{C'}\right)\right] \cdot \left(|\mathcal{K}'_{\mathcal{N}^*}| - \left|\mathcal{K}'_{\mathcal{N}'_3}\right| \right). \end{aligned}$$

■

Theorem 2: Algorithm 3 achieves an approximation factor of $(e-1)/2e$ for the cell planning problem.

Proof: According to Fact 3 and Lemma 4, at the $(L+1)$th iteration, we can obtain

$$\begin{aligned}
&\left|\mathcal{K}'_{\mathcal{N}'_{L+1}}\right| - \left|\mathcal{K}'_{\mathcal{N}'_3}\right| \\
&\geq \left[1 - \prod_{j=4}^{L+1}\left(1 - \frac{c_{n_j}}{C'}\right)\right] \cdot \left(\left|\mathcal{K}'_{\mathcal{N}^*}\right| - \left|\mathcal{K}'_{\mathcal{N}'_3}\right|\right) \\
&\geq \left[1 - \prod_{j=4}^{L+1}\left(1 - \frac{c_{n_j}}{\sum_{j=4}^{L+1} c_{n_j}}\right)\right] \cdot \left(\left|\mathcal{K}'_{\mathcal{N}^*}\right| - \left|\mathcal{K}'_{\mathcal{N}'_3}\right|\right) \\
&\geq \left[1 - \left(1 - \frac{1}{L+1}\right)^{L+1}\right] \cdot \left(\left|\mathcal{K}'_{\mathcal{N}^*}\right| - \left|\mathcal{K}'_{\mathcal{N}'_3}\right|\right) \\
&\geq \left(1 - \frac{1}{e}\right) \cdot \left(\left|\mathcal{K}'_{\mathcal{N}^*}\right| - \left|\mathcal{K}'_{\mathcal{N}'_3}\right|\right).
\end{aligned}$$

Then

$$\begin{aligned}
\left|\mathcal{K}'_{\mathcal{N}'_L}\right| &= \left|\mathcal{K}'_{\mathcal{N}'_{L+1}}\right| - \left(\left|\mathcal{K}'_{\mathcal{N}'_{L+1}}\right| - \left|\mathcal{K}'_{\mathcal{N}'_L}\right|\right) \\
&\geq \left(1 - \frac{1}{e}\right) \cdot \left(\left|\mathcal{K}'_{\mathcal{N}^*}\right| - \left|\mathcal{K}'_{\mathcal{N}'_2}\right|\right) + \left|\mathcal{K}'_{\mathcal{N}'_2}\right| - \frac{1}{3}\left|\mathcal{K}'_{\mathcal{N}'_2}\right| \\
&\geq \left(1 - \frac{1}{e}\right) \cdot \left(\left|\mathcal{K}'_{\mathcal{N}^*}\right| - \left|\mathcal{K}'_{\mathcal{N}'_2}\right| + \left|\mathcal{K}'_{\mathcal{N}'_2}\right|\right). \\
&= \left(1 - \frac{1}{e}\right) \cdot \left|\mathcal{K}'_{\mathcal{N}^*}\right| \\
&\geq \frac{e-1}{2e} \cdot |\mathcal{K}^*|.
\end{aligned}$$

Notice that the number of TDPs served by $\mathcal{K}'_{\mathcal{N}'_L}$ is no more than the number of TDPs selected by Algorithm 3. Hence, Algorithm 3 can achieve an approximation factor of $(e-1)/2e$ for the *budgeted cell planning problem*. ■

IV. Numerical Results and Discussions

A. Numerical Results

We give some numerical results to evaluate our cell planning algorithms. We compare the HetNet and the deployment scheme without small cells for the same celluar system. Simulation parameters, such as path-loss models, maximum transmission power, total available bandwidth, etc., are based on the specifications proposed in [33]. The service area is 3×3 km^2. There are 30 candidate macrocells and 50 candidate small cells in the area, which are randomly generated. The maximum transmission power of a macrocell is 46 dBm, and the installation cost varies within an interval of (8, 12). The transmission power of each small cell is limited to 30 dBm, and the installation cost is uniformly distributed in the interval $(8 \cdot t, 12 \cdot t)$, where t is the ratio of the average installation cost of a small cell to that of a macrocell. The total available bandwidth is 20 MHz. Each TDP is uniformly distributed in the deployment area with a rate requirement of 3 Mb/s. The path loss (in decibels) from the BS of the macrocell to TDP is $128.1 + 37.6 \log_{10}(R)$, and the path loss from the BS of the small cell to TDP is $140.7 + 36.7 \log_{10}(R)$ for distance R in kilometers. The standard deviation of lognormal shadowing is 10 dB. The noise power spectral density is -180 dBm/Hz, and Γ is set to 7.6288 (BER $= 10^{-6}$).

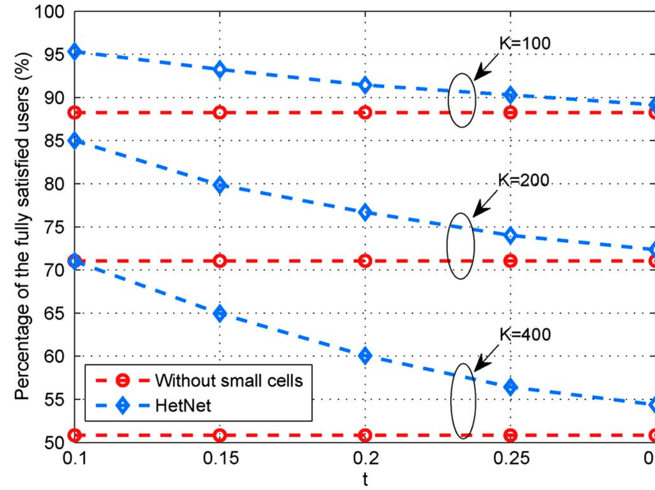


Fig. 1. Average percentage of the fully satisfied TDPs as a function of t, $K = 100, 200, 400$, and $C = 40$.

First, we investigate the average percentage of the fully satisfied TDPs versus parameter t. The number of TDPs varies from 100 to 400, as shown in Fig. 1. The budget C is set to 40. In Fig. 1, we can see that the percentage of the fully satisfied TDPs decreases as t increases. The reason is intuitive. When t is small, more small cells can be selected with given budget C. Thus, the area spectrum efficiency of the considered cellular system is improved. The performance of the HetNet is much better than the macrocell-only scheme, even if the installation cost of the small cell is not attractive, e.g., for the case of $t = 0.3$. On the other hand, for a given t, as the number of TDPs increases, the satisfied TDPs also increase for both the HetNet and the macrocell-only network. However, the average percentage of the satisfied TDPs decreases for both schemes as the number of TDPs increases. It can be explained as follows. As the density of TDPs increases, a small cell cannot serve all TDPs it covers. It is the same case for a macrocell.

Second, we investigate the average percentage of the fully satisfied TDPs as a function of deployment budget. The number of TDPs is set to $K = 200$. As expected, we can see in Fig. 2 that the average percentage of the fully satisfied TDPs increases as the budget increases because more cells can be installed for a big budget. Again, we can find that the HetNet can serve more TDPs than the macrocell-only scheme. However, the gap between the HetNet and the macrocell-only network becomes smaller as t increases. Particularly for the case of $t = 0.3$, the gap is slight for any given budget, indicating that the optimal number of small cells underlaying a macrocell is dependent on the cost ratio of the small cell to the macrocell.

B. Discussions of Interference Issues

In a practical HetNet, inter-layer and cross-layer interference are the major factors that influence the system capacity. For the LTE-A network, macrocells are interconnected with each other by X2 interface, which facilitates that CoMP and ICIC technologies can be employed to reduce intercell interference [17], [18] and significantly improve cell-edge user throughput.

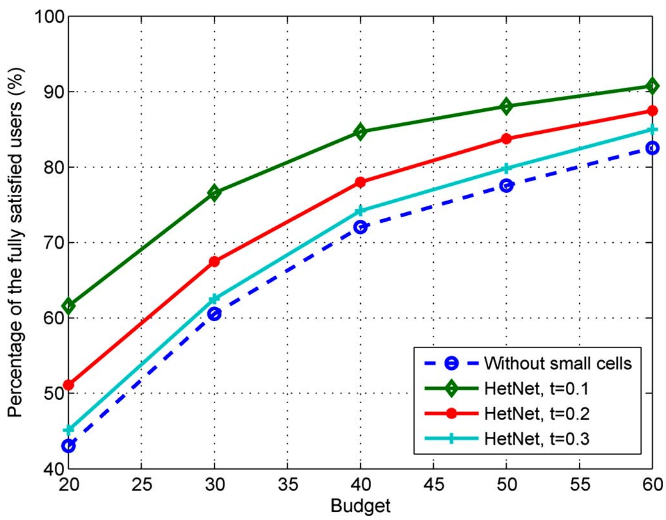


Fig. 2. Average percentage of the fully satisfied TDPs as a function of the budget; $K = 200$, and $t = 0.1,\ 0.2,$ and 0.3.

Additionally, due to the low transmission power of small cells, the interlayer interference between two small cells is usually slight if they are far apart. Therefore, we can safely put aside the interlayer interference in the cell planning phase for the HetNet. On the other hand, for the case of densely deployed small cells, fractional frequency reuse [34] for interference management can be employed. Our proposed cell planning strategy can also be adopted with necessary modifications.

Since small cells usually locate within the coverage of macrocells, cross-layer interference management becomes a critical issue for the HetNet. For instance, users served by a macrocell might suffer heavy interference from multiple small cells within the macrocell, and *vice versa*. Generally, the cross-layer interference is very hard to tackle for cell planning in HetNet because the infrastructure of the HetNet is totally different from conventional cellular systems where only homogeneous cells exist. Fortunately, as discussed in [22] and [23], for a k-tier heterogeneous cellular network, analysis shows that the number and density of different types of cells and their relative power levels do not change the coverage probability of the cellular system under some reasonable and practical assumptions. By carefully designing a cell association strategy, cell planning can be independent of interference management.

Although cell association and interference coordination are not discussed in this work, it is noteworthy that our proposed planning algorithm can work for any kind of cellular networks even if considering the interference between different cells, except that the approximation ratio may not be proved since the optimal solution cannot be described in these circumstances. In brief, our proposed algorithm provides a performance-guaranteed cell planning scheme for cellular networks that employ HetNet infrastructure. Thus, we can conservatively conclude that our proposal sheds insights on how to deploy a HetNet in a cost-efficient way.

V. Conclusion

In this paper, we have studied the *budgeted cell planning problem* in HetNets. We try to select a subset of candidate cells to maximize the number of fully satisfied TDPs while keeping the total deployment cost below a given budget. In addition to involving both small cells and macrocells, our network model also takes practical limitations in cellular systems into consideration, including transmission power of different types of cells, spectrum limitations, and traffic demands. In particular, link gains between the base station and users are considered in our problem formulation, which characterizes the distinctiveness of the HetNets. By minimizing the power consumption of a cell with given bandwidth, we design a $1/2$-approximation algorithm to assign TDPs to a given set of cells, and based on this, an $(e-1)/2e$-approximation algorithm is developed to solve the target *budgeted cell planning problem*. Preliminary numerical results and discussions of interference issues are also given. For future work, relay nodes should be investigated. Relays offer additional flexibility where wireline backhaul is unavailable or not economical. The potential sites of relays are constrained by the channel gain between the relays and the donor cell. Such a problem is important but hard to address because relays can only be selected with a given donor cell. Moreover, the interference issue should also be investigated in more detail to design a more practical cell planning scheme for HetNets.

Acknowledgment

The authors would like to thank the editors and the anonymous reviewers whose invaluable comments helped substantially improve the presentation of this paper.

References

[1] C. Lee and H. Kang, "Cell planning with capacity expansion in mobile communications: A tabu search approach," *IEEE Trans. Veh. Technol.*, vol. 49, no. 5, pp. 1678–1691, Sep. 2000.

[2] K. Tutschku, "Demand-based radio network planning of cellular mobile communication systems," in *Proc. IEEE INFOCOM*, Mar. 1998, pp. 1054–1061.

[3] S. Hurley, "Planning effective cellular mobile radio networks," *IEEE Trans. Veh. Technol.*, vol. 51, no. 2, pp. 243–253, Mar. 2002.

[4] J. Kalvenes, J. Kennington, and E. Olinick, "Hierarchical cellular network design with channel allocation," *Eur. J. Oper. Res.*, vol. 160, no. 1, pp. 3–18, Jan. 2005.

[5] N. Weicker, G. Szabo, K. Weicker, and P. Widmayer, "Evolutionary multiobjective optimization for base station transmitter placement with frequency assignment," *IEEE Trans. Evol. Comput.*, vol. 7, no. 2, pp. 189–203, Apr. 2003.

[6] A. Dutta and V. Hsu, "Cellular network design site selection and frequency planning," *Ann. Oper. Res.*, vol. 106, no. 1, pp. 287–306, Sep. 2001.

[7] R. Mathar and T. Niessen, "Optimum positioning of base stations for cellular radio networks," *Wireless Netw.*, vol. 6, no. 6, pp. 421–428, Dec. 2000.

[8] R. Mathar and M. Schmeink, "Optimal base station positioning and channel assignment for 3G mobile networks by integer programming," *Ann. Oper. Res.*, vol. 107, no. 1, pp. 225–236, Oct. 2001.

[9] J. Yang, M. Aydin, J. Zhang, and C. Maple, "UMTS base station location planning: A mathematical model and heuristic optimisation algorithms," *IET Commun.*, vol. 1, no. 5, pp. 1007–1014, Oct. 2007.

[10] E. Amaldi, A. Capone, and F. Malucelli, "Planning UMTS base station location: Optimization models with power control and algorithms," *IEEE Trans. Wireless Commun.*, vol. 2, no. 5, pp. 939–952, Sep. 2003.

[11] E. Amaldi and A. Capone, "Radio planning and coverage optimization of 3G cellular networks," *Wireless Netw.*, vol. 14, no. 4, pp. 435–447, Aug. 2008.

[12] H.-Y. Zhang, Y.-G. Xi, and H.-Y. Gu, "A rolling window optimization method for large-scale WCDMA base stations planning problem," *Eur. J. Oper. Res.*, vol. 183, no. 1, pp. 370–383, Nov. 2007.
[13] Y. Zhou, J. Wang, and M. Sawahashi, "Downlink transmission of broadband OFCDM systems—Part I: Hybrid detection," *IEEE Trans. Commun.*, vol. 53, no. 4, pp. 718–729, Apr. 2005.
[14] Y. Zhou and J. Wang, "Downlink transmission of broadband OFCDM systems—Part IV: Soft decision," *IEEE J. Sel. Areas Commun.*, vol. 24, no. 6, pp. 1208–1220, Jun. 2006.
[15] S.-P. Yeh, S. Talwar, G. Wu, N. Himayat, and K. Johnsson, "Capacity and coverage enhancement in heterogeneous networks," *IEEE Wireless Commun.*, vol. 18, no. 3, pp. 32–38, Jun. 2011.
[16] A. Ghosh *et al.*, "Heterogeneous cellular networks: From theory to practice," *IEEE Commun. Mag.*, vol. 50, no. 6, pp. 54–64, Jun. 2012.
[17] R. Irmer *et al.*, "Coordinated multipoint: Concepts, performance, and field trial results," *IEEE Commun. Mag.*, vol. 49, no. 2, pp. 102–111, Feb. 2011.
[18] G. Boudreau *et al.*, "Interference coordination and cancellation for 4G networks," *IEEE Commun. Mag.*, vol. 47, no. 4, pp. 74–81, Apr. 2009.
[19] W. Zhao, S. Wang, C. Wang, and X. Wu, "Cell planning for heterogeneous networks: An approximation algorithm," in *Proc. IEEE INFOCOM*, Apr. 2014, pp. 1087–1095.
[20] W. Zhao and S. Wang, "Cell planning for heterogeneous cellular networks," in *Proc. IEEE WCNC*, Apr. 2013, pp. 1032–1037.
[21] S. Wang, W. Zhao, and C. Wang, "Approximation algorithms for cellular networks planning with relay nodes," in *Proc. IEEE WCNC*, Apr. 2013, pp. 3230–3235.
[22] H. Dhillon, R. Ganti, F. Baccelli, and J. Andrews, "Modeling and analysis of k-tier downlink heterogeneous cellular networks," *IEEE J. Sel. Areas Commun.*, vol. 30, no. 3, pp. 550–560, Apr. 2012.
[23] H.-S. Jo, Y. Sang, P. Xia, and J. Andrews, "Heterogeneous cellular networks with flexible cell association: A comprehensive downlink SINR analysis," *IEEE Trans. Wireless Commun.*, vol. 11, no. 10, pp. 3484–3495, Oct. 2012.
[24] A. Goldsmith and S.-G. Chua, "Variable-rate variable-power MQAM for fading channels," *IEEE Trans. Commun.*, vol. 45, no. 10, pp. 1218–1230, Oct. 1997.
[25] M. R. Garey and D. S. Johnson, *Computers and Intractability: A Guide to the Theory of NP-Completeness*. San Francisco, CA, USA: Freeman, 1979.
[26] S. Khuller, A. Moss, and J. Naor, "The budgeted maximum coverage problem," *Info. Process. Lett.*, vol. 70, no. 1, pp. 39–45, Apr. 1999.
[27] A. Ageev and M. Sviridenko, "Approximation algorithms for maximum coverage and max cut with given sizes of parts," in *Proc. IPCO*, 1999, pp. 17–30.
[28] D. S. Hochbaum, *Approximation Algorithms for NP-Hard Problems*. Boston, MA, USA: PWS, 1997.
[29] U. Feige, "A threshold of $\ln n$ for approximating set cover," *J. ACM*, vol. 45, no. 4, pp. 634–652, Jul. 1998.
[30] R. Cohen and L. Katzir, "The generalized maximum coverage problem," *Inf. Process. Lett.*, vol. 108, no. 1, pp. 15–22, Sep. 2008.
[31] D. Amzallag, J. Naor, and D. Raz, "Coping with interference: From maximum coverage to planning cellular networks," in *Proc. WAOA*, Sep. 2006, pp. 29–42.
[32] S. Boyd and L. Vandenberghe, *Convex Optimization*. New York, NY, USA: Cambridge Univ. Press, 2004.
[33] 3rd Generation Partnership Project (3GPP), " Evolved universal terrestrial radio access (E-UTRA); further advancements for E-UTRA physical layer aspects (TR 36.814)," Mar. 2010.
[34] A. Stolyar and H. Viswanathan, "Self-organizing dynamic fractional frequency reuse in OFDMA systems," in *Proc. IEEE INFOCOM*, Apr. 2008, pp. 1364–1372.

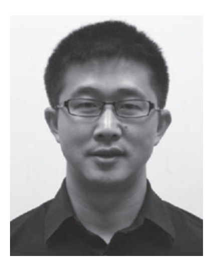
Shaowei Wang (S'06–M'07–SM'13) received the B.S., M.S., and Ph.D. degrees in electronic engineering from Wuhan University, Wuhan, China, in 1997, 2003, and 2006, respectively.

From 1997 to 2001, he was an R&D Scientist with China Telecom. Since 2006, he has been with the School of Electronic Science and Engineering, Nanjing University, Nanjing, China, where he is currently an Associate Professor. From 2012 to 2013, he was a Visiting Scholar/Professor with Stanford University, Stanford, CA, USA, and The University of British Columbia, Vancouver, BC, Canada. He has published more than 60 papers in leading journals and conference proceedings in his areas of interest. He organized the Special Issue on Enhancing Spectral Efficiency for LTE-Advanced and Beyond Cellular Networks for *IEEE Wireless Communications* and the Feature Topic on Energy-Efficient Cognitive Radio Networks for the *IEEE Communications Magazine*. His research focuses on wireless communications and networking.

Dr. Wang is on the Editorial Board of the *IEEE Communications Magazine* and serves/served on the Technical Committee or the Executive Committee of many reputable conferences, including the IEEE Conference on Computer Communications, the IEEE International Conference on Communications, the IEEE Global Communications Conference, the IEEE Wireless Communications and Networking Conference, etc.

Wentao Zhao (S'13) received the M.S. degree from Nanjing University, Nanjing, China, in 2013, where he is currently working toward the Ph.D. degree with the School of Electronic Science and Engineering.

His current research focuses on resource allocation in wireless networks, cellular networks, planning, and optimization.



Chonggang Wang (SM'09) received the Ph.D. degree from Beijing University of Posts and Telecommunications, Beijing, China, in 2002.

He is currently a Member of Technical Staff of InterDigital Communications with focuses on Internet of Things (IoT) R&D activities, including technology development and standardization. He has (co)authored more than 100 journal/conference articles and book chapters. His current research interests include IoT, mobile communications and computing, and big data management and analytics.

Dr. Wang is the founding Editor-in-Chief of the *IEEE Internet of Things Journal* and is on the Editorial Board for several journals, including *IEEE Access*.