Approximation Algorithms for Cellular Networks Planning with Relay Nodes

Shaowei Wang and Wentao Zhao
School of Electrical Science & Engineering
Nanjing University
Nanjing 210093, China
E-mail: wangsw@nju.edu.cn, zhangwt@smail.nju.edu.cn

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

***Abstract*—Relay nodes are introduced to the next generation cellular networks to enhance coverage and improve system capacity, leading to a new radio network planning paradigm. In this paper, we study two planning problems for cellular networks with relay nodes: Minimum cost cell planning and budgeted cell planning. The former is to minimize the total installation cost for opening base stations (BSs), including macro BSs and relay nodes, while satisfying all users' traffic demands. The latter is to maximize the number of users with predefined traffic demands under a given budget. Both of the problems are NP-hard. We present approximation algorithms to work out promising solutions to these problems. For the minimum cost cell planning, we develop an $O(\log W)$-approximation algorithm, where W is the maximum capacity of macro BSs. For the budgeted cell planning, we prove that the problem is NP-hard to approximate and give an $\frac{e-1}{3e-1}$-approximation algorithm for a special case of the problem, which is general enough to meet practical requirements.**

I. INTRODUCTION

Generally, traditional cell planning involves positioning base stations (BSs), determining the power level of each BS, and grouping radio frequency among antennas of BSs. The optimization objective is to minimize the total deployment cost or maximize the number of demand nodes with a given deployment budget. For the second generation (2G) cellular system, cells are usually planned in two stages. First, a set of sites is selected from a candidate list, and configured with a series of radio parameters. These BSs should satisfy predefined service requirements, such as converging the service area with strong enough signal strength, supplying system throughput as high as possible [1–4]. In some cases, the objective of planning is to minimize the total installation cost or maximize the number of users with QoS guarantee [5, 6]. Second, radio frequency is grouped to decrease the interference among different cells [2–6] to an acceptable level.

For the third generation (3G) cellular networks, no frequency grouping is required because all cells can use the same spectrum. The key purpose of planning is how to select a set of sites from the candidate ones to minimize the cost while keeping the signal-to-interference-plus-noise ratio (SINR) above a predefined threshold [7–11]. Because the SINR is mainly determined by the relative power of neighboring cells, power control/allocation is one the most important issues for cell planning in 3G systems. In [7–9], cell planning is formulated as the classical capacitated facility location problem, where the optimization objective is to minimize the installation cost and the transmission power. In [10, 11], the optimization objective is to maximize the number of users served by the cellular system with a given number of macro BSs.

In the standardization process of the next generation systems, such as 3GPP Long Term Evolution-Advanced (LTE-A), relay technology has been actively discussed [12–14]. A relay node can help a macro BS to expand its coverage or increase throughput for the spots where the direct link between the BS and a user suffers deep fading. Such spots can exist at a cell edge or in the shadow of large objects, such as the towers in downtown, underground constructions. Moreover, a relay node can also compensate the coverage hole. Compare with macro BSs, relay node has lower transmission power, smaller physical size and much cheaper cost. Thus it can offer flexible site acquisitions with relative low cost.

Most research for cell planning is concentrated on the deployment of macro BSs. However, in this paper, we also consider the deployment of relay nodes in cellular networks, where the relay node can expand the service coverage of a macro BS but cannot provide additional capacity. We formulate two general cell planning problems. The first is how to position the macro BSs and relay nodes with minimum cost while satisfying all traffic demands of users. The second one is how to maximize the number of demand nodes with a given installation budget. Both problems are NP-hard. We develop approximation algorithms to tackle them in an efficient way.

The remainder of this paper is organized as follows. In Section II, we formulate the cell planning problems. In Section III, approximation algorithms are presented in details. In Section IV, numerical results are given and discussed. Conclusion is drawn in Section V.

II. PROBLEM FORMULATION

Consider an area served by a cellular network. Denote the candidate sites for deploying macro BSs as $\mathcal{I} = \{1, 2, \ldots, I\}$ and the demand nodes as $\mathcal{J} = \{1, 2, \ldots, J\}$. Each macro BS $i \in \mathcal{I}$ has a capacity w_i, installation cost c_i^M, and each demand node $j \in \mathcal{J}$ requests a traffic demand d_j. The coverage of macro BS i is denoted as $S_i^M \subseteq \mathcal{J}$. A demand node can be served by multiple macro BSs.

978-1-4673-5939-9/13/$31.00 ©2013 IEEE

Denote the candidate sites for deploying relay nodes as $\mathcal{K} = \{1, 2, \ldots, K\}$. The relay node $k \in \mathcal{K}$ has a installation cost c_k^R and is connected to a macro BS. Similar to macro BS, relay node k has a coverage area $S_k^R \subseteq J$. Denote the set of relay nodes connected to the macro BS i as K_i, i.e., $\mathcal{K} = \bigcup_{i=1}^{I} K_i$ and $K_i \cap K_{i'} = \emptyset$, for all $i, i' \in \mathcal{I}, i \neq i'$. We assume that $|K_i|$ is limited by a constant number.

Define z_i^M as an index variable which indicates macro BS i is selected or not

$$z_i^M = \begin{cases} 1 & \text{macro BS } i \text{ is selected,} \\ 0 & \text{otherwise,} \end{cases} \quad \forall i \in \mathcal{I}.$$

Let z_k^R be the selection variable of relay node k

$$z_k^R = \begin{cases} 1 & \text{relay node } k \text{ is selected,} \\ 0 & \text{otherwise,} \end{cases} \quad \forall k \in \mathcal{K}.$$

A. Minimum Cost Cell Planning

The minimum cost cell planning problem is how to select a subset from $\mathcal{I}$ and a subset from $\mathcal{K}$ to satisfy all traffic demands with the minimum installation cost. The optimization problem can be formulated as follows,

$$\begin{aligned} & \min \sum_{i \in \mathcal{I}} c_i^M z_i^M + \sum_{k \in \mathcal{K}} c_k^R z_k^R \\ s.t. \quad & C_1 : \sum_{i \in \mathcal{I}} x_{ij} = d_j, \forall j \in \mathcal{J}, \\ & C_2 : \sum_{j \in \mathcal{J}} x_{ij} \leq w_i z_i^M, \forall i \in \mathcal{I}, \\ & C_3 : x_{ij} \geq 0, \forall i \in \mathcal{I}, j \in S_i^M, \\ & C_4 : \max\nolimits_{k: j \in S_k^R, k \in K_i} w_i z_k^R \geq x_{ij} \geq 0, \forall i \in \mathcal{I}, j \notin S_i^M, \\ & C_5 : z_i^M \in \{0, 1\}, \forall i \in \mathcal{I}, \\ & C_6 : z_k^R \in \{0, 1\}, \forall k \in \mathcal{K}, \\ & C_7 : z_i^M \geq z_k^R, \forall i \in \mathcal{I}, k \in K_i, \end{aligned} \tag{1}$$

where x_{ij} is the capacity of macro BS i allocated to demand node j, $\forall i \in \mathcal{I}, j \in \mathcal{J}$. C_1 means that all traffic demands should be satisfied. C_2 is the capacity constraint of macro BS. C_3 and C_4 are the coverage constraints. C_7 indicates that relay node $k \in K_i$ can be selected only when macro BS i has been selected.

B. Budgeted Cell Planning

Unlike the minimum cost cell planning problem, the budgeted cell planning problem is how to select a subset from $\mathcal{I}$ and a subset from $\mathcal{K}$ to maximize the number of satisfied demand nodes, while the total installation cost cannot exceed a given budget B. Let y_j be the selection variable of demand node j

$$y_j = \begin{cases} 1 & \text{demand node } j \text{ is fully satisfied,} \\ 0 & \text{otherwise,} \end{cases} \quad \forall j \in \mathcal{J}.$$

Mathematically, the optimization problem can be formulated as,

$$\begin{aligned} & \max \sum_{j \in \mathcal{J}} y_j \\ s.t. \quad & C_1 : \sum_{i \in \mathcal{I}} x_{ij} = d_j y_j, \forall j \in \mathcal{J}, \\ & C_2 : \sum_{j \in \mathcal{J}} x_{ij} \leq w_i z_i^M, \forall i \in \mathcal{I}, \\ & C_3 : x_{ij} \geq 0, \forall i \in \mathcal{I}, j \in S_i^M, \\ & C_4 : \max\nolimits_{k: j \in S_k^R, k \in K_i} w_i z_k^R \geq x_{ij} \geq 0, \forall i \in \mathcal{I}, j \notin S_i^M, \\ & C_5 : \sum_{i \in \mathcal{I}} c_i^M z_i^M + \sum_{k \in \mathcal{K}} c_k^R z_k^R \leq B, \\ & C_6 : z_i^M \in \{0, 1\}, \forall i \in \mathcal{I}, \\ & C_7 : z_k^R \in \{0, 1\}, \forall k \in \mathcal{K}, \\ & C_8 : z_i^M \geq z_k^R, \forall i \in \mathcal{I}, k \in K_i, \\ & C_9 : y_j \in \{0, 1\}, \forall j \in \mathcal{J}. \end{aligned} \tag{2}$$

C_5 indicates that the total installation cost cannot exceed B.

III. Our Proposed Algorithms

In general, Eq.(1) and Eq.(2) cannot be solved efficiently since both of them are NP-hard. Furthermore, even if the number of macro BSs I is a constant, Eq.(2) is still intractable because maximizing the number of demand nodes severed by a given set of macro BSs is a multidimensional 0-1 knapsack problem which is also NP-hard.

A. $O(\log W)$-Approximation Algorithm for Minimum Cost Cell Planning

Let $f(I', K')$ be the maximum traffic demands supplied by a given set of macro BSs $I' \subseteq \mathcal{I}$ and relay nodes $K' \subseteq \mathcal{K}$. Denote $S_i \subseteq \mathcal{J}$ as a set of demand nodes which can be served by macro BS i or relay node k, where $k \in K_i \cap K'$. We can calculate the $f(I', K')$ by solving the following linear program problem:

$$\begin{aligned} & \max \sum_{j \in \mathcal{J}} \sum_{i \in I'} x_{ij} \\ s.t. \quad & C_1 : \sum_{i \in I'} x_{ij} \leq d_j, \forall j \in \mathcal{J}, \\ & C_2 : \sum_{j \in \mathcal{J}} x_{ij} \leq w_i, \forall i \in I', \\ & C_3 : x_{ij} \geq 0, \forall i \in I', j \in S_i, \\ & C_4 : x_{ij} = 0, \forall i \in I', j \notin S_i. \end{aligned} \tag{3}$$

Denote $c(I', K')$ as the total cost for such a deployment. For $\widetilde{I} \subseteq \mathcal{I}$ and $\widetilde{K} \subseteq \mathcal{K}$, define

$$F_{I', K'}(\widetilde{I}, \widetilde{K}) = \frac{f(I' \cup \widetilde{I}, K' \cup \widetilde{K}) - f(I', K')}{c(\widetilde{I}, \widetilde{K})}, \tag{4}$$

the outline of the approximation algorithm (Algorithm 1) is described in Table I, where Algorithm 2 is given in Table II and $X \backslash Y = \{x | x \in X, x \notin Y\}$.

Let (I'_l, K'_l) be a solution at the end of the lth iteration of Algorithm 1. Without loss of generality, let i_l and k_l be the

TABLE I
ALGORITHM 1

$I' = \emptyset$, $K' = \emptyset$, $I^* = \mathcal{I}$, $K^* = \mathcal{K}$, $l = 0$.
while $f(I', K') < \sum_{j \in \mathcal{J}} d_j$ **do**
$l = l + 1, F = 0$;
for all $i \in I^*$ **do**
find $G_i \subseteq K_i$ that maximizes $F_i = F_{I',K'}(\{i\}, G_i)$ and update F, i_l, k_l by using Algorithm 2;
end for
for all $i \in I'$ **do**
find $k'_i \in K_i \backslash K'$ that maximizes $F'_i = F_{I',K'}(\emptyset, \{k'_i\})$;
if $F_i > F$
$i_l = \emptyset$, $k_l = \{k'_i\}$, $F = F'_i$;
end if
end for
$I' = I' \cup i_l$, $I^* = I^* \backslash i_l$, $K' = K' \cup k_l$, $K^* = K^* \backslash k_l$;
end while
return (I', K')

TABLE II
ALGORITHM 2

$K_i^* = \{G_i^* | F_{I' \cup \{i\}, K'}(\emptyset, G_i^*) > \max\{F, F_{I',K'}(\{i\}, \emptyset)\}, |G_i^*| = 1\}$, $\gamma = 1$, $K'_i = K_i^*$;
if $K'_i = \emptyset$ and $F_{I',K'}(\{i\}, \emptyset) > F$
$i_l = \{i\}$, $k_l = \emptyset$, $F = F_{I',K'}(\{i\}, \emptyset)$;
end if
while $K'_i \neq \emptyset$ and $\gamma \leq |K_i|$
find $G_i \subseteq K'_i$ that maximizes $F_i = F_{I',K'}(\{i\}, G_i)$;
if $F_i > F$
$K_i^* = \{G_i^* | F_{I' \cup \{i\}, K'}(\emptyset, G_i^*) > F, |G_i^*| = 1\}$;
$i_l = \{i\}$, $k_l = G_i$, $F = F_i$;
end if
$K'_i = \{G'_i \cup G_i^* | F_{I' \cup \{i\}, K'}(\emptyset, G'_i) > F, G'_i \in K'_i, G_i^* \in K_i^*, G'_i \cap G_i^* = \emptyset\}$, $\gamma = \gamma + 1$;
end while
return i_l, k_l, F

set of selected macro BSs and the set of selected relay nodes added to I'_{l-1} and K'_{l-1} at the lth iteration, respectively.

At the lth iteration of Algorithm 1, it always holds

$$F_{I'_{l-1},K'_{l-1}}(i_l, k_l) \geq F_{I'_{l-1},K'_{l-1}}(\{i\}, G), i \in \mathcal{I} \backslash I'_{l-1}, G \subseteq K_i, \quad (5)$$

Proof: At the beginning of proof, we need to prove that Algorithm 2 can always find $G_i \subseteq K_i$ which can maximize $F_{I'_{l-1},K'_{l-1}}(\{i\}, G_i)$ for each unselected macro BS $i \in \mathcal{I} \backslash I'_{l-1}$.

If $F_{I'_{l-1},K'_{l-1}}(\{i\}, \emptyset) \geq F_{I'_{l-1} \cup \{i\}, K'_{l-1}}(\emptyset, G), \forall G \subseteq K_i, |G| = 1$, it always holds

$$\begin{aligned} F_{I'_{l-1},K'_{l-1}}(\{i\}, \emptyset) \geq & \max_{G \subseteq K_i, |G|=1} F_{I'_{l-1} \cup \{i\}, K'_{l-1}}(\emptyset, G) \\ \geq & \frac{\sum_{G \subseteq G', |G|=1} f_{I'_{l-1} \cup \{i\}, K'_{l-1}}(\emptyset, G)}{\sum_{G \subseteq G', |G|=1} c(G)} \\ \geq & \frac{f_{I'_{l-1} \cup \{i\}, K'_{l-1}}(\emptyset, G')}{c(G')}, \forall G' \subseteq K_i, \end{aligned}$$

where follows by $\frac{a_0 + \ldots + a_n}{b_0 + \ldots + b_n} \leq \max_i \frac{a_i}{b_i}$, for $a_i, b_i > 0, \forall i$ and $f_{I'_{l-1},K'_{l-1}}(\emptyset, G') \leq \sum_{k \in G'} f_{I'_{l-1},K'_{l-1}}(\emptyset, \{k\})$. Hence, we have

$$F_{I'_{l-1},K'_{l-1}}(\{i\}, \emptyset) \geq F_{I'_{l-1},K'_{l-1}}(\{i\}, G'), \forall G' \subseteq K_i,$$

which follows $\frac{a_0}{b_0} \geq \frac{a_0 + a_1}{b_0 + b_1}$ when $\frac{a_0}{b_0} \geq \frac{a_1}{b_1}$, for $a_0, a_1, b_0, b_1 > 0$. For this case, the macro BS has the priority.

On the contrary, we ignore relay node k which satisfies $F_{I' \cup \{i\}, K'}(\emptyset, \{k\}) \leq F_{I',K'}(\{i\}, \emptyset)$ since we will get

$$F_{I'_{l-1},K'_{l-1}}(\{i\}, \emptyset) \geq F_{I'_{l-1},K'_{l-1}}(\{i\}, \{k\}).$$

Also, we also ignore the combination of relay nodes G' which meet the requirement of $F_{I' \cup \{i\}, K'}(\emptyset, G') \leq F$. When $K'_i = \emptyset$ happens, which also means $F_{I' \cup \{i\}, K'}(\emptyset, G') \leq F$, for all $G' \in K'_i$, we have

$$F_{I'_{l-1},K'_{l-1}}(\{i\}, G') \leq F,$$

which follows $\frac{a_0 + a_1}{b_0 + b_1} \geq \frac{a_0 + a_2}{b_0 + b_2}$ when $\frac{a_0 + a_1}{b_0 + b_1} \geq \frac{a_2}{b_2}$ and $\frac{a_1}{b_1} \geq \frac{a_0}{b_0}$, for $a_0, a_1, a_2, b_0, b_1, b_2 > 0$.

For unselected relay nodes of macro BS $i \in I'_{l-1}$, we ignore all combinations of relay nodes since

$$\begin{aligned} F_{I'_{l-1},K'_{l-1}}(\emptyset, G') \leq & \frac{\sum_{k \in G'} f_{I'_{l-1},K'_{l-1}}(\emptyset, \{k\})}{c(G')} \\ \leq & \max_{k \in G'} \frac{f_{I'_{l-1},K'_{l-1}}(\emptyset, \{k\})}{c(\{k\})} \\ = & \max_{k \in G'} F_{I'_{l-1},K'_{l-1}}(\emptyset, \{k\}), \end{aligned}$$

$\forall G' \subseteq K_i \backslash K'_{l-1}$. ∎

Algorithm 1 achieves an approximation of $O(\log W)$, where $W = max_{i \in \mathcal{I}} w_i$.

Proof: Let (OPT_I, OPT_K) be the optimal solution. For each macro BS $i \in OPT_I$, define $H_i = OPT_K \cap K_i$. Let $\{\widetilde{x_{ij}}\}$ be a solution to the Eq.(3) with input (OPT_I, OPT_K). Macro BS $i \in OPT_I$ supplys $\sum_{j \in J} \widetilde{x_{ij}}$ traffic demands with a cost of $c(OPT_I, OPT_K)$.

Next, we inductively define $a_l(i)$ as the traffic demand which is not satisfied after l iterations for each macro BS $i \in OPT_I$. If add i to I'_l and H_i to K'_l, the remaining traffic demands $a_l(i)$ can be satisfied and we always have

$$a_l(i) \leq f(I'_l \cup \{i\}, K'_l \cup H_i) - f(I'_l, K'_l). \quad (6)$$

Consider macro BS $i \in OPT_I \backslash I'_l$. According to the Eq.(5) and the Eq.(6), we have

$$\begin{aligned} F_{I'_{l-1},K'_{l-1}}(i_l, k_l) & \geq F_{I'_{l-1},K'_{l-1}}(\{i\}, H_i) \\ & \geq a_{l-1}(i) / c(\{i\}, H_i). \end{aligned} \quad (7)$$

In the case that adding $i \in OPT_I$ to I', but not adding H_i to K' at the lth iteration, we also have

$$\begin{aligned} F_{I'_{l-1},K'_{l-1}}(i_l, k_l) & \geq F_{I'_{l-1},K'_{l-1}}(\emptyset, H_i \backslash K'_{l-1}) \\ & \geq a_{l-1}(i) / c(\emptyset, H_i \backslash K'_{l-1}) \\ & \geq a_{l-1}(i) / c(\{i\}, H_i). \end{aligned} \quad (8)$$

Now, we calculate the extra cost of the deployment determined by Algorithm 1 compared with the optimal solution. If we add $i \in OPT_I$ to I'_{l-1} and H_i to K'_{l-1} at the lth iteration, there is no extra cost, since OPT_I and OPT_K also include such items. Otherwise, the extra cost can be computed as follows,

$$c_l(i) = (a_{l-1}(i) - a_l(i)) \cdot F_{I'_{l-1},K'_{l-1}}(i_l, k_l). \quad (9)$$

Therefore, we have

$$\begin{aligned}\sum_{l=1}^{L} c_l(i) &= \sum_{l=1}^{L}(a_{l-1}(i) - a_l(i))F_{I'_{l-1},K'_{l-1}}(i_l, k_l) \\ &\leq c(\{i\}, H_i)\sum_{l=1}^{L}\tfrac{a_{l-1}(i)-a_l(i)}{a_{l-1}(i)} \\ &= c(\{i\}, H_i)\cdot H(\textstyle\sum_{j\in J}\widetilde{x_{ij}}) \\ &= c(\{i\}, H_i)\cdot O(\log(\textstyle\sum_{j\in J}\widetilde{x_{ij}})) \\ &= c(\{i\}, H_i)\cdot O(\log w_i),\end{aligned} \quad (10)$$

where $H(r)$ is the rth harmonic number, L is the number of iterations. So $c(I', K')$ is bounded by $O(\log W) \cdot c(OPT_I, OPT_K)$. ■

B. Inapproximability of Budgeted Cell Planning

The budgeted cell planning problem is strong NP-hard because finding a feasible solution to it is also NP-hard.

Proof: Given an instance of the budgeted cell planning problem with $\mathcal{I} = \{1, 2, \ldots, I\}$, $\mathcal{K} = \emptyset$, $|\mathcal{J}| = 1$, $B = d_1$ and $w_i = c_i^M$, $i \in \mathcal{I}$. The demand node is satisfied if and only if there exists $I' \subseteq \mathcal{I}$ with

$$\sum_{i\in I'} w_i = \sum_{i\in I'} c_i^M = d_1 = B. \quad (11)$$

The problem of finding a subset $I' \subseteq \mathcal{I}$ that satisfies the Eq.(11) follows the subset sum problem which is NP-hard [15]. Thus the budgeted cell planning problem is hard to approximate. ■

C. $\frac{e-1}{3e-1}$-Approximation algorithm for a Spacial Case of the Budgeted Cell Planning Problem

Consider a restrictive version of the budgeted cell planning problem which is general practical for cellular networks. Assume that a set of k macro BSs can fully satisfy at least k demand nodes, for a positive integer k, such a budgeted cell planning problem is still NP-hard but not hard to approximate. The algorithm for this spacial case generalizes the results of the [16]. First, consider the Eq.(2) with an additional constraint:

$$x_{ij} \in \{0, d_j\}, i \in \mathcal{I}, j \in \mathcal{J}, \quad (12)$$

which means that a demand node can be served by at most one macro BS. With such an additional constraint, we can achieve an approximation ratio of $(1 - 1/e)$ which is the best by now [17]. The outline of the approximation algorithm (Algorithm 3) for the Eq.(12) is described in Table III, where $N_{I',K'}(\widetilde{I}, \widetilde{K})$ can be calculated as follows,

$$N_{I',K'}(\widetilde{I}, \widetilde{K}) = \frac{N_A(I' \cup \widetilde{I}, K' \cup \widetilde{K}) - N_A(I', H')}{c(\widetilde{I}, \widetilde{K})}, \quad (13)$$

where $N_A(I', K')$ is the number of demand nodes satisfied by a given set of macro BSs I' and a given set of relay nodes K'. Computing $N_A(I', K')$ is also NP-hard. We give an approximation algorithm (Algorithm 4) as described in Table IV.

TABLE III
ALGORITHM 3

for all $I' \subseteq \mathcal{I}$, $K' \subseteq \mathcal{K}$, $c(I', K') \leq B$ and $|I'| + |K'| < 3$, compute $N_A(I', K')$. Let (I^*, K^*) be the solution with the highest value of N_A computed.
for all $I' \subseteq \mathcal{I}$, $K' \subseteq K$, $|I'| + |K'| = 3$ **do**
  for all $k \in K'$ **do**
    $I' = I' \cup \{i|k \in K_i\}$;
  end for
  if $c(I', K') \leq B$
    $U = \mathcal{I}\backslash I'$, $V = \mathcal{K}\backslash K'$, $K'_i = \emptyset, \forall i$;
    repeat
      find $i^* \in U$, $G_{i^*} \subseteq K_{i^*}$, $G_{i^*} \notin K'_{i^*}$ that maximize $a = N_{I',K'}(\{i^*\}, G_{i^*})$;
      find $k^* \in V$ that maximizes $b = N_{I',K'}(\emptyset, \{k^*\})$;
      if $a \geq b$
        if $c(I', K') + c_i^M + c(G_i) \leq B$
          $I' = I' \cup \{i^*\}$, $K' = K' \cup G_{i^*}$;
          $U = I\backslash\{i^*\}$, $V = V\backslash G_{i^*}$;
        else
          $K'_{i^*} = K'_{i^*} \cup \{G_{i^*}\}$;
        end if
      else
        if $c(I', K') + c_k^R \leq B$
          $K' = K' \cup \{k^*\}$;
        end if
        $V = V\backslash\{k^*\}$;
      end if
    until $U = \emptyset$ and $V = \emptyset$
  end if
  if $N_A(I', K') > N_A(I^*, K^*)$
    $(I^*, K^*) = (I', K')$;
  end if
end for
return (I^*, K^*)

TABLE IV
ALGORITHM 4

$J^* = \mathcal{J}$, $N_A(I', K') = 0$;
repeat
  $j^* = \arg\min_{j\in J^*} d_j$;
  for $i \in I'$ **do**
    if $j^* \in S_i$ and $w_i - d_{j^*} \geq 0$
      $N_A(I', K') = N_A(I', K') + 1$, $w_i = w_i - d_{j^*}$;
    end if
  end for
  $J^* = J^*\backslash\{j^*\}$;
until $J^* = \emptyset$
return $N_A(I', K')$

The approximation ratio of Algorithm 4 is $\frac{1}{2}$.

Proof: Let OPT be the optimal solution. For a macro BS i, let J_{OPT} be the set of demand nodes satisfied by i in OPT, but not be satisfied by Algorithm 4. J_{OPT} is the non-decreasing order of the traffic demand which is denoted as $j_1, j_2, \ldots, j_h$. The demand nodes connected to macro BS i by Algorithm 4 is $j'_1, j'_2, \ldots, j'_l$, which has also a non-decreasing order. We always have

$$d_{j'_l} \leq d_{j_1} \quad (14)$$

TABLE V
ALGORITHM 5

```
Let (I'_L, K'_L) be the output of Algorithm 3;
Find I' ⊆ I that maximizes |I'| and c(I', ∅) ≤ B;
if N_A(I'_L, K'_L) < |I'|
    Output I' and a set of demand nodes that can be covered by I';
else
    Output (I'_L, K'_L) and the demand nodes chosen by Algorithm 4;
end if
```

and

$$d_{j_1} + \sum_{t=1}^{l} d_{j'_t} > w_i. \tag{15}$$

Otherwise, j_1 would be conected to macro BS i by Algorithm 4. Since $j_1, j_2, \ldots, j_h$ are conected to macro BS i in the OPT, we have

$$\sum_{t=1}^{h} d_{j_t} \leq w_i. \tag{16}$$

Based on the Eq.(15) and the Eq.(16), we get

$$\sum_{t=1}^{l} d_{j'_t} > \sum_{t=1}^{h} d_{j_t} - d_{j_1} = \sum_{t=2}^{h} d_{j_t}. \tag{17}$$

Base on the Eq.(14) and the Eq.(17), we can conclude that $l > h-1$, which means the number of demand nodes satisfied by Algorithm 4 is always greater than $\frac{1}{2}|OPT|$. ■

Thus, Algorithm 3 can achieve an $\frac{e-1}{2e}$ approximation ratio for the Eq.(12). Then we develop an approximation algorithm (Algorithm 5) for Eq.(2) which is shown in Table V.

Algorithm 5 achieves an $\frac{e-1}{3e-1}$ approximation ratio.

Proof: Let n' be the number of satisfied demand nodes obtained by Algorithm 5, and n^* is the number of satisfied demand nodes obtained by the optimal solution. Let n_1^* and n_2^* be the number of demand nodes satisfied by a single macro BS and multiple macro BS in the optimal solution, respectively. Denote I^* as the set of macro BSs selected in the optimal solution to satisfy these $n^* = n_1^* + n_2^*$ demand nodes. Denote $N_A(OPT)$ as the optimal solution with considering the Eq.(12). Obviously $N_A(OPT) \geq n_1^*$.

Note that the number of demand nodes satisfied by more than one macro BS is at most the number of selected macro BSs for our considered special case. We have

$$n' \geq N_A(I'_L, K'_L) \geq (\frac{e-1}{2e})N_A(OPT) \geq (\frac{e-1}{2e})n_1^*. \tag{18}$$

Then we obtain

$$\begin{aligned}(\tfrac{3e-1}{2e})n' &= n' + (\tfrac{e-1}{2e})n' \\ &\geq n' + (\tfrac{e-1}{2e})|I^*| \\ &\geq (\tfrac{e-1}{2e})n_1^* + (\tfrac{e-1}{2e})n_2^* \\ &\geq (\tfrac{e-1}{2e})n^*.\end{aligned} \tag{19}$$

And we have

$$n' \geq (\frac{e-1}{3e-1})n^*. \tag{20}$$

■

TABLE VI
NUMERICAL RESULTS 1: THE AVERAGE COST AS A FUNCTION OF THE NUMBER OF DEMAND NODES

N	60	120	180	240
Our Method	592.92	731.81	815.79	915.04
Unplanned Manner	739.27	911.71	1025.73	1134.99

TABLE VII
NUMERICAL RESULTS 2: THE AVERAGE COST AS A FUNCTION OF THE COVERAGE RADIUS OF RELAY NODE

Coverage Radii(m)	300	450	600	750
Our Method	1023.26	970.25	913.58	856.70
Unplanned Manner	1061.88	1060.37	1060.56	1051.62

IV. NUMERICAL RESULTS

In this Section, we give the numerical results of the proposed algorithm. The service area is 6×10 km^2. There are 60 possible sites of macro BSs and 240 possible sites of relay nodes in the deployment area, which are randomly generated. For macro i, its capacity w_i is distributed uniformly within $(600, 900)$, and we set $c_i^M = w_i/10$. The coverage radius of each macro BS is 1.5km. Each macro BS can connect to four relay nodes. The relay nodes connected to each macro BS is uniformly distributed at the edge of the BS. The installation cost of each relay node is calculated as follows,

$$c_k^R = t\sum_{i=1}^{I} c_i^M / I.$$

where t is the ratio of the relay cost to the cost of a macro BS. Demand nodes are uniformly distributed in the deployment area. For demand node j, its traffic demand d_j is distributed uniformly within $(20, 30)$.

First, we show the average cost with different number of demand nodes. For comparison, the method without relay nodes is also shown in Table VI. The method is proposed in [1], which define the same problem as this work except for not considering relay nodes. The number of demand nodes varies from 60 to 240. The coverage radius of each relay node is 0.5 km and $t = 0.1$. As can be seen in Table VI, the average cost increases as the number of demand nodes increases, and our proposed algorithm performs better than the method without relays.

Then we investigate the average cost with different coverage radius of relay nodes. In general, the transmission power of a macro BS (e.g., 46dBm) is normally much higher than that of a relay node (e.g., 30dBm). However, one related work in relay-based networks, range expansion, increases the footprint of low-power BSs such as relay nodes by adding a positive

TABLE VIII
NUMERICAL RESULTS 3: THE AVERAGE COST AS A FUNCTION OF INSTALLATION COST OF RELAY NODE

t	0.05	0.2	0.35	0.5
Our Method	810.44	914.06	967.23	994.50
Unplanned Manner	1056.05	1050.01	1061.09	1056.13

bias to their measured signal strengths during cell association [18]. Thus a user still may be served by one relay node even if this user is not within the coverage area of the relay node. In Table VII, there are 200 demand nodes in the deployment area and $t = 0.1$. The ratio of the coverage radius of relay node to that of macro BS varies from 0.2 to 0.5. When the coverage radius of relay node is small, the average cost by our method is close to the method without relays because the number of demand nodes covered by a relay node is far less than the number of demand nodes covered by a macro BS, which does not meet the practical scenarios. The average cost of our proposed method decreases as the ratio increases, which still performs better than the method without relays.

Finally, we give the average cost with different installation cost of relay node which is shown in Table VIII. The simulation parameters are same as Table VII except that the coverage radius of relay node is 0.5 km and t varies from 0.05 to 0.5. Again, the average cost of our proposed method is lower than the case without relays even though the installation cost of relay node is not attractive.

V. Conclusion

In this paper we studied cell planning for cellular networks with relay nodes. We formulate two different problems which are general for practical deployment cases. For the minimum cost cell planning, where the optimization objective is to minimize the total deployment cost while satisfying all traffic demands, we developed an $O(\log W)$-approximation algorithm. For the budgeted cell planning, where the objective is to maximize the number of fully satisfied demand nodes with a given budget, we proposed an $\frac{e-1}{3e-1}$-approximation algorithm for a practical case of the problem. Though some constraints are not considered in this work, we think our proposal throw some insights for cellular networks planning, especially for the network with relays.

Acknowledgement

This work was partially supported by the JiangsuSF (BK2011051), Huawei Technologies Co., Ltd (YB2012120195,YBWL2012SKL09) and the Fundamental Research Funds for the Central Universities.

References

[1] C. Lee and H. Kang, "Cell planning with capacity expansion in mobile communications: a tabu search approach," *IEEE Transactions on Vehicular Technology*, vol. 49, no. 5, pp. 1678–1691, Sep. 2000.

[2] S. Hurley, "Planning effective cellular mobile radio networks," *IEEE Transactions on Vehicular Technology*, vol. 51, no. 2, pp. 243–253, Mar. 2002.

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

[4] A. Dutta and V. Hsu, "Cellular network design site selection and frequency planning," *Annals of Operations Research*, vol. 106, no. 1, pp. 287–306, Sep. 2001.

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

[6] R. Mathar and M. Schmeink, "Integrated optimal cell site selection and frequency allocation for cellular radio networks," *Telecommunication Systems*, vol. 21, no. 2, pp. 339–347, Dec. 2002.

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

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

[9] A. Kamar, S. Nawaz, M. Patwary, M. Abdel-Maguid, and S.-U.-R. Qureshi, "Optimized algorithm for cellular network planning based on terrain and demand analysis," in *Proc. ICCTD 2010*, Cairo, Egypt, Nov. 2010, pp. 359–364.

[10] R. Mathar and M. Schmeink, "Optimal base station positioning and channel assignment for 3G mobile networks by integer programming," *Annals of Operations Research*, vol. 107, no. 1, pp. 225–236, Oct. 2001.

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

[12] J. Sydir and R. Taori, "An evolved cellular system architecture incorporating relay stations," *IEEE Communications Magazine*, vol. 47, no. 6, pp. 115–121, June 2009.

[13] Y. Yang, H. Hu, J. Xu, and G. Mao, "Relay technologies for WiMax and LTE-advanced mobile systems," *IEEE Communications Magazine*, vol. 47, no. 10, pp. 100–105, Oct. 2009.

[14] W. Guo and T. O'Farrell, "Relay deployment in cellular networks: Planning and optimization," *IEEE Journal on Selected Areas in Communications*, vol. PP, no. 99, pp. 1–10, 2012.

[15] M. R. Garey and D. S. Johnson, *Computers and Intractability: A Guide to the Theory of NP-Completeness*. San Francisco: W. H. Freeman, 1979.

[16] D. Amzallag, J. Naor, and D. Raz, "Coping with interference: From maximum coverage to planning cellular networks," in *Proc. WAOA 2006*, Zurich, Switzerland, Sep. 2006, pp. 29–42.

[17] S. Khuller, A. Moss, and J. Naor, "The budgeted maximum coverage problem," *Information Processing Letters*, vol. 70, no. 1, pp. 39–45, Apr. 1999.

[18] I. Guvenc, "Capacity and fairness analysis of heterogeneous networks with range expansion and interference coordination," *IEEE Communications Letters*, vol. 15, no. 10, pp. 1084–1087, Oct. 2011.