An Efficient Codebook Based Radio Parameter Optimization Method for Mobile Networks

Linzhi Shen, and Shaowei Wang
School of Electronic Science and Engineering, Nanjing University, Nanjing 210023, China.
Email: 141180100@smail.nju.edu.cn, wangsw@nju.edu.cn

GLOBECOM 2022 - 2022 IEEE Global Communications Conference | 978-1-6654-3540-6/22/$31.00 ©2022 IEEE | DOI: 10.1109/GLOBECOM48099.2022.10001704

***Abstract*—With the increasing of the directional antennas in millimeter wave mobile communication systems, tuning the azimuths and downtilts of these antennas in an optimal manner is the core method to enhance the network performance, including the power coverage and the system throughput of the network. Since the formulated optimization problem is generally computational intractable, we introduce a codebook based optimization scheme to handle this difficult task efficiently. Specifically, a codebook with a small set of promising azimuths and downtilts is generated for each base station according to its surrounding radio propagation environment, based on which all the antennas are adjusted to serve the target area with strong enough signals while alleviating the interference to adjacent base stations. Numerical results show that the proposed scheme can yield promising antenna configurations with limited computing resources.**

***Index Terms*—Codebook, millimeter wave, Monte Carlo tree search, radio parameter optimization.**

I. Introduction

Radio parameter optimization (RPO) is a critical part of radio resource management in mobile communication systems, where radio frequency parameters of base stations (BSs) are adjusted to improve the power coverage of the network and the system throughput [1, 2]. Power coverage aims to maximize the area with strong enough signals, while system throughput, also referred to as capacity coverage, depends on the signal-to-interference-plus-noise ratio (SINR), which is related to the strength of not only the target signals but also the interference ones. More specifically, the power increase of transmission may improve the power coverage of a cell, but such a power increase also throws more interference to the adjacent cells and lower the capacity coverage. Therefore, a tradeoff between the two objectives should be achieved in practical mobile networks.

The 5G and beyond mobile communication systems introduce millimeter wave (mmWave) technology to get more bandwidth resources, which however results in further challenges for the RPO problem. Specifically, mmWave bands suffer from severe fading due to the shorter wavelength as compared to the sub-6GHz bands. In response to such kind of fading, directional antennas which concentrate energy on particular narrow directions have been installed in large numbers [3]. Compared to the ordinary omnidirectional antenna model, a directional antenna has a flexible radiation pattern that varies with its azimuth or downtilt. As a result, the target signals and the interference ones are closely related to the configuration of azimuths and downtilts, which is quite different from the sub-6GHz scenarios [4]. Indeed, recent studies have shown that tuning the azimuths and downtilts of directional antennas can improve the spectrum utilization efficiency of mmWave communication systems in the spatial domain [5, 6].

This work was supported in part by the National Natural Science Foundation of China under Grants 61931023 and U1936202.

Despite its great significance, searching for the optimal antenna configurations always yields a combinatorial optimization problem, which is intractable for even medium-scale antennas. For example, if we try to find out the optimal azimuth and downtilt of one antenna, it generates 72×181 combinations when the azimuth range $[0°, 360°)$ is discretized into intervals of $5°$ and the downtilt range $[-90°, 90°)$ is discretized into intervals of $1°$. In this way, the number of parameter combinations for 30 antennas is $(72 \times 181)^{30}$. Therefore, enumerating all feasible solutions is impossible even with huge computing resources.

Different kinds of algorithms are investigated to obtain promising antenna configurations in the literature, among which reinforcement learning algorithms attach widespread attention. Within the framework of reinforcement learning, the RPO problem can be treated as a Markov decision process and the policy of antenna configuration can be obtained via the interaction between the RPO engine and the network environment [7, 8]. In [7], the antennas are successively adjusted by evaluating the outcomes of all feasible configurations, where the remaining antennas are fixed when a given antenna is to be configured. The major limitation of this method is that it ignores the coupled dependence of the adjacent antennas. In [8], a Monte Carlo tree search method is proposed to overcome this limitation, where the impact of policies on the adjacent antennas is averaged via multiple Monte Carlo simulations. However, this method is costly due to the curse of dimensionality brought by the vast combinations of antenna configurations.

The main dilemma of conventional reinforcement learning schemes for the RPO comes from the conflict between modeling the dependence among antennas and controlling the size of the solution space. This conflict can be effectively eased if the candidate antenna configuration set of each antenna is compressed. One promising approach to implement such a compression is to design a codebook containing a small set of the most promising antenna settings for each BS. In this way, antennas can determine their configurations from the corresponding codebooks instead of considering all feasible antenna settings.

In this paper, we propose a codebook based Monte Carlo tree search (CB-MCTS) scheme for the RPO problem, which only explores limited antenna configurations while preserving the ability of modeling coupled dependence. The CB-MCTS method contains two stages: codebook generation and antenna configuration. In the codebook generation stage, a codebook is generated for each BS based on its surrounding propagation environment. The elements in the codebook are supposed to be potential for improving the power coverage and the total throughput in the surrounding area of the corresponding BS. We formulate this subtask as a binary quadratic programming problem and handle it with a branch and bound procedure followed by semi-definite relaxation algorithm. In the antenna configuration stage, the antennas are tuned based on the predefined codebooks to optimize the power coverage of the network, as well as the system throughput. This subtask is formulated as a Markov decision process and a Monte Carlo tree search method is introduced to generate a policy on antenna configuration. Numerical results show that our proposed scheme obtains promising antenna settings in real environments with reasonable computing load.

II. System Model and Problem Formulation

Consider an urban mobile communication network composed of several BSs, where each BS contains three directional antennas, as shown in Fig 1. The sets of BSs and antennas are represented by $\mathbb{N}$ and $\mathbb{A}$, respectively. For each antenna $a \in \mathbb{A}$, its azimuth θ_a ranges in $[0°, 360°)$ while its downtilt ϕ_a ranges in $[-90°, 90°]$. Denote $\boldsymbol{\theta} = \{\theta_a\}_{a\in\mathbb{A}}$ and $\boldsymbol{\phi} = \{\phi_a\}_{a\in\mathbb{A}}$ as the azimuth set and downtilt set of antennas, respectively. To illustrate the power coverage and the capacity coverage of the network explicitly, the target area is divided into 5m × 5m grids. The grid centers, also referred to as traffic demand points (TDPs), are used to measure the signals in the corresponding grids. The set of TDPs is represented by $\mathbb{U}$.

For each $a \in \mathbb{A}$ and $u \in \mathbb{U}$, the power of signal received by u from a, which varies with the azimuth and downtilt of a, is represented by $P_{a,u}(\theta_a, \phi_a)$. Note that each TDP is always associated with one antenna that provides the strongest signal, the strength of the target signal received by u is given by

$$P_u(\boldsymbol{\theta}, \boldsymbol{\phi}) = \max_{a\in\mathbb{A}} P_{a,u}(\theta_a, \phi_a), \tag{1}$$

and the SINR at u can be expressed as

$$\rho_u(\boldsymbol{\theta}, \boldsymbol{\phi}) = \frac{P_u(\boldsymbol{\theta}, \boldsymbol{\phi})}{\sum\limits_{a\in\mathbb{A}} P_{a,u}(\theta_a, \phi_a) - P_u(\boldsymbol{\theta}, \boldsymbol{\phi}) + \delta^2}, \tag{2}$$

where δ^2 denotes the power of additive white Gaussian noise. The power coverage, which refers to the ratio of the TDPs

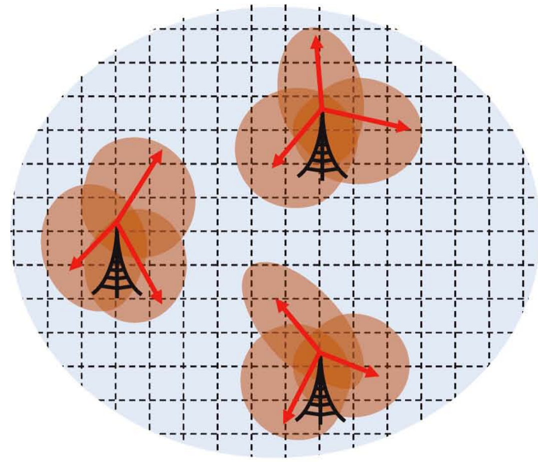


Fig. 1. Illustration of the network scenario.

whose maximum received signal strength exceeds a certain threshold T^p to the total TDPs $|\mathbb{A}|$, is represented by

$$\eta^p(\boldsymbol{\theta}, \boldsymbol{\phi}) = \frac{\sum_{u\in\mathbb{U}} \mathbb{1}(P_u(\boldsymbol{\theta}, \boldsymbol{\phi}) > T^p)}{|\mathbb{U}|}, \tag{3}$$

where $\mathbb{1}(x)$ is an indicator denoted by

$$\mathbb{1}(x) = \begin{cases} 1 & , \quad \text{if } x \textit{ is } \textbf{true}, \\ 0 & , \quad \text{otherwise.} \end{cases} \tag{4}$$

Similarly, the capacity coverage η^c, which refers to the ratio of the TDPs whose SINR exceeds a certain threshold T^c to $|\mathbb{U}|$, is expressed as

$$\eta^c(\boldsymbol{\theta}, \boldsymbol{\phi}) = \frac{\sum_{u\in\mathbb{U}} \mathbb{1}(\rho_u(\boldsymbol{\theta}, \boldsymbol{\phi}) > T^c)}{|\mathbb{U}|}. \tag{5}$$

The target of the RPO is to provide both high η^p and high η^c by configuring $\boldsymbol{\theta}$ and $\boldsymbol{\phi}$. A new index called valid capacity coverage ratio is introduced to handle this multi-objective optimization problem, which is given by

$$\eta^e(\boldsymbol{\theta}, \boldsymbol{\phi}) = \frac{\sum_{u\in\mathbb{U}} \mathbb{1}(P_u(\boldsymbol{\theta}, \boldsymbol{\phi}) > T^p \wedge \rho_u(\boldsymbol{\theta}, \boldsymbol{\phi}) > T^c)}{|\mathbb{U}|}. \tag{6}$$

With the range of azimuths and donwtilts discretized into intervals of 5° and 1°, respectively, the RPO task can be formulated by

$$\begin{aligned} \max_{\boldsymbol{\theta}, \boldsymbol{\phi}} \quad & \eta^e(\boldsymbol{\theta}, \boldsymbol{\phi}), \\ \text{s.t.} \quad & \theta_a \in \{0°, 5°, 10°, ..., 355°\}, \qquad \forall a \in \mathbb{A}, \\ & \phi_a \in \{-90°, -89°, -88°, ..., 90°\}, \quad \forall a \in \mathbb{A}. \end{aligned} \tag{7}$$

III. Codebook based method for RPO

The optimization problem (7) is intractable due to the huge number of feasible combinations of antenna configurations. The proposed CB-MCTS method only explores limited antenna configurations instead of enumerating all the configuration combinations. It contains two stages: First, a site-specific codebook generation algorithm is designed for each BS to determine its candidate antenna configurations. Then, the antennas are successively adjusted based on the codebooks to optimize the valid capacity coverage ratio of the network.

A. Codebook Generation

The cardinality of candidate configuration set for each antenna is $D = 360 \times 181$, which is quite large but unnecessary due to the following facts: First, similar antenna configurations result in almost the same valid capacity coverage ratio. Moreover, some specific antenna settings for a given BS can hardly improve the valid capacity coverage ratio of the network in any state owing to the characteristics of the surrounding propagation environment. As a result, a set containing the most potential antenna configurations should be generated for each BS according to its surrounding propagation environment so as to improve the efficiency of exploring antenna configurations.

To this end, the metric for measuring the performance of candidate antenna configuration sets should first be determined. Without loss of generality, we consider an antenna configuration set $\{(\theta_a^i, \phi_a^i)\}_{i=1}^D$ for a specific BS, where a denotes the antenna deployed in this BS. The discrepancy between the i-th and j-th antenna setting is defined as

$$c_{i,j} \triangleq \frac{\sum_{u\in\mathbb{U}}(P_{a,u}(\theta_a^i, \phi_a^i) > T^p) \oplus (P_{a,u}(\theta_a^j, \phi_a^j) > T^c)}{|\mathbb{U}|}. \quad (8)$$

where $\oplus$ is the exclusive or sign. The optimization objective is defined as maximizing the discrepancies between antenna settings within the codebook. The task can be expressed as

$$\begin{aligned} \max_{\boldsymbol{x}} \quad & \boldsymbol{x}^T \boldsymbol{C} \boldsymbol{x} = \sum_{i=1}^D \sum_{j=1}^D c_{i,j} x_i x_j \\ \text{s.t.} \quad & \boldsymbol{e}^T \boldsymbol{x} = K \\ & \boldsymbol{x} \in \{0,1\}^D, \end{aligned} \quad (9)$$

where $\boldsymbol{x}$ is a D-dimensional 0-1 vector used to indicate whether each antenna setting is included in the codebook, $\boldsymbol{C}$ is a symmetric matrix, $\boldsymbol{e}$ is an all-one vector, and K is the size of codebook that is predefined. This is a linearly constraint binary quadratic programming (BQP) problem, which is known to be NP-hard [9].

Branch and bound (BnB) algorithm is typically used to find out the optimal solution of BQP problems. Within the framework of BnB, the set of feasible solutions is thought of as forming a rooted tree with the full set at the root, and the algorithm explores branches of this tree which represents subsets of the solution set. Each branch corresponds to an optimization problem that consists of the original optimization problem with some additional constraints. One key step of the BnB algorithm is to obtain a tight upper bound for one branch so that the algorithm can decide whether to continue exploring this branch or just prune it. However, the matrix $\boldsymbol{C}$ in (9) is not semi-definite since the diagonal elements of $\boldsymbol{C}$ equal to 0, which makes the objective function non-convex. As a result, the upper bounds of the branches are hard to obtain just by linear programming relaxation. Semi-definite relaxation (SDR) is an efficient approximation technique for host of non-convex optimization problems, which relaxes the original problem to a semi-definite programming (SDP) problem so that polynomial methods can be utilized to solve it [10]. In this paper, we embed the SDR within a simple BnB algorithm to solve the proposed codebook generation problem in an effective manner.

Several transforms for (9) are formulated before SDR to make the bound tight. Define $\boldsymbol{z} = 2\boldsymbol{x} - \boldsymbol{e}$, the problem (9) can be reformulated by

$$\begin{aligned} \max_{\boldsymbol{z}} \quad & \frac{1}{4}(\boldsymbol{z}^T \boldsymbol{C} \boldsymbol{z} + 2\boldsymbol{z}^T \boldsymbol{C} \boldsymbol{e} + \boldsymbol{e}^T \boldsymbol{C} \boldsymbol{e}) \\ \text{s.t.} \quad & \boldsymbol{e}^T \boldsymbol{z} = 2K - D \\ & \boldsymbol{z} \in \{-1,1\}^D. \end{aligned} \quad (10)$$

Consider the following optimization problem

$$\begin{aligned} \max_{\tilde{\boldsymbol{z}}} \quad & \tilde{\boldsymbol{z}}^T \tilde{\boldsymbol{C}} \tilde{\boldsymbol{z}} \\ \text{s.t.} \quad & \tilde{\boldsymbol{z}}^T \tilde{\boldsymbol{C}}_0 \tilde{\boldsymbol{z}} = 4K - 2D \\ & \tilde{\boldsymbol{z}} \in \{-1,1\}^{D+1}, \end{aligned} \quad (11)$$

where $\tilde{\boldsymbol{z}} = \begin{pmatrix} z_0 \\ \boldsymbol{z} \end{pmatrix}$, and the symmetric matrices $\tilde{\boldsymbol{C}}$ and $\tilde{\boldsymbol{C}}_0$ are represented by

$$\tilde{\boldsymbol{C}} = \frac{1}{4}\begin{bmatrix} \boldsymbol{e}^T \boldsymbol{C} \boldsymbol{e} & \boldsymbol{e}^T \boldsymbol{C} \\ \boldsymbol{C} \boldsymbol{e} & \boldsymbol{C} \end{bmatrix}, \quad \tilde{\boldsymbol{C}}_0 = \begin{bmatrix} 0 & \boldsymbol{e}^T \\ \boldsymbol{e} & \boldsymbol{0} \end{bmatrix}.$$

The equivalence between (10) and (11) can be easily proved as follows: On the one hand, if $\boldsymbol{z}^*$ is optimal in (10), then $\begin{pmatrix} 1 \\ \boldsymbol{z}^* \end{pmatrix}$ is feasible in (11) with the same objective value. On the other hand, if $\begin{pmatrix} z_0 \\ \boldsymbol{z}^* \end{pmatrix}$ is optimal in (11), then $z_0 \boldsymbol{z}^*$ is feasible in (10) with the same objective value. Thus, (10) and (11) have the same optimal objective value, and the relationship between their optimal solutions is easy to obtain. □

A crucial step to derive an SDR is to observe that

$$\begin{aligned} \tilde{\boldsymbol{z}}^T \tilde{\boldsymbol{C}} \tilde{\boldsymbol{z}} &= \mathrm{Tr}(\tilde{\boldsymbol{z}}^{\mathrm{T}} \tilde{\boldsymbol{C}} \tilde{\boldsymbol{z}}) = \mathrm{Tr}(\tilde{\boldsymbol{C}} \tilde{\boldsymbol{z}} \tilde{\boldsymbol{z}}^{\mathrm{T}}), \\ \tilde{\boldsymbol{z}}^T \tilde{\boldsymbol{C}}_0 \tilde{\boldsymbol{z}} &= \mathrm{Tr}(\tilde{\boldsymbol{z}}^{\mathrm{T}} \tilde{\boldsymbol{C}}_0 \tilde{\boldsymbol{z}}) = \mathrm{Tr}(\tilde{\boldsymbol{C}}_0 \tilde{\boldsymbol{z}} \tilde{\boldsymbol{z}}^{\mathrm{T}}). \end{aligned}$$

By introducing a new variable $\tilde{\boldsymbol{Z}} = \tilde{\boldsymbol{z}} \tilde{\boldsymbol{z}}^T$ and noticing that $\tilde{\boldsymbol{Z}}$ is a rank one symmetric positive semi-definite matrix whose diagonal elements are all 1, we obtain the following equivalent formulation of (11):

$$\begin{aligned} \max_{\tilde{Z}} \quad & \mathrm{Tr}(\tilde{\boldsymbol{C}} \tilde{\boldsymbol{Z}}) \\ \text{s.t.} \quad & \mathrm{Tr}(\tilde{\boldsymbol{C}}_0 \tilde{\boldsymbol{Z}}) = 4K - 2D \\ & \mathrm{Tr}(\boldsymbol{E}_i \tilde{\boldsymbol{Z}}) = 1, i \in \{1, 2, ..., D\} \\ & \tilde{\boldsymbol{Z}} = \tilde{\boldsymbol{Z}}^T \\ & \tilde{\boldsymbol{Z}} \geq \boldsymbol{0} \\ & \mathrm{Rank}(\tilde{\boldsymbol{Z}}) = 1, \end{aligned} \quad (12)$$

where $\boldsymbol{E}_i$ is the matrix with 0 entries except in position (i,i) where there is a 1. The only difficult constraint in (12) is the rank one constraint, which is non-convex. Thus, we may drop it to obtain a relaxed version:

$$\begin{aligned}\max_{\tilde{Z}} \quad & \mathrm{Tr}(\tilde{\boldsymbol{C}}\tilde{\boldsymbol{Z}}) \\ \text{s.t.} \quad & \mathrm{Tr}(\tilde{\boldsymbol{C}}_0\tilde{\boldsymbol{Z}}) = 4K - 2D \\ & \mathrm{Tr}(\boldsymbol{E}_i\tilde{\boldsymbol{Z}}) = 1, \forall i \in \{1,2,...,D\} \\ & \tilde{\boldsymbol{Z}} = \tilde{\boldsymbol{Z}}^T \\ & \tilde{\boldsymbol{Z}} \geq \boldsymbol{0}. \end{aligned} \tag{13}$$

The problem (13) is known as an SDR of (12).

For the optimization problem of a branch where some elements of $\boldsymbol{x}$ are fixed, the formulation can be written by

$$\begin{aligned}\max_{\boldsymbol{x}} \quad & \boldsymbol{x}^T\boldsymbol{C}\boldsymbol{x} \\ \text{s.t.} \quad & \boldsymbol{e}^T\boldsymbol{x} = K \\ & x_i = q_i, \forall i \in \mathbb{I} \\ & \boldsymbol{x} \in \{0,1\}^D, \end{aligned} \tag{14}$$

where $\mathbb{I}$ is the set of indexes of fixed element in $\boldsymbol{x}$, and q_i is a constant with a value of 0 or 1. Its corresponding SDR is expressed as

$$\begin{aligned}\max_{\tilde{Z}} \quad & \mathrm{Tr}(\tilde{\boldsymbol{C}}\tilde{\boldsymbol{Z}}) \\ \text{s.t.} \quad & \mathrm{Tr}(\tilde{\boldsymbol{C}}_0\tilde{\boldsymbol{Z}}) = 4K - 2D \\ & \mathrm{Tr}(\boldsymbol{E}_i\tilde{\boldsymbol{Z}}) = 1, \forall i \in \{1,2,...,D\} \\ & \tilde{z}_{i+1,j+1} = (2q_i - 1)(2q_j - 1), \forall i < j \in \mathbb{I} \\ & \tilde{\boldsymbol{Z}} = \tilde{\boldsymbol{Z}}^T \\ & \tilde{\boldsymbol{Z}} \geq \boldsymbol{0}. \end{aligned} \tag{15}$$

It can then be solved in MATLAB with the convex optimization toolbox CVX [11].

With the obtained bounds of branches, the BnB algorithm to solve the codebook generation task is designed as follows:

• **Initialization.** Good initial solution gives tight upper bound at the beginning of the BnB. We propose a simple heuristic algorithm to obtain a promising initial solution, which is illustrated in Algorithm 1.

Algorithm 1 Heuristic algorithm for initialization

1: Initialization: the target size of codebook K,
the codebook $\Psi = \varnothing$
2: $\Psi = \Psi \cup \arg\max_i \sum_{u \in \mathbb{U}} \mathbb{1}(P_{a,u}(\theta_a^i, \phi_a^i) > T^p)$
3: **for** $k = 2 : K$ **do**
4: $\Psi = \Psi \cup \arg\max_i \sum_{\psi \in \Psi} c_{i,\psi}$
5: **end for**
6: **return** Ψ

• **Branching strategy.** The branching scheme is based on a basic depth-first search. Specifically, the algorithm starts at the root node and explores as far as possible along each branch before backtracking. The backtracking is executed when a leaf node is reached or the SDR of the current branch is lower than the optimal value of solutions that have been obtained. The algorithm terminates when all branches have been explored or pruned.

B. Antenna Configuration

The size of the candidate exploration space for the RPO task is reduced from $D^{|\mathbb{A}|}$ to $W^{|\mathbb{A}|}$ with codebook size set to W, which is still too large to enumerate. Fortunately, this task can be formulated as an Markov decision process (MDP). In a general MDP, an agent observes the state of the environment before an action and obtains a reward after the action [12]. For the RPO task, the agent manages the configuration of all BSs, and the environment represents the radio communication network. The action of the agent is the combination of configurations for all antennas which are selected from the codebooks. The state of the environment is defined by the distributions of maximum signal power and SINR in the area, which is measured by all TDPs. The reward is the valid capacity coverage ratio η^e of the network. This MDP can be formed as a process where BSs are sequentially configured, as shown in Fig 2. Denote $n_j \in \mathbb{N}$ as the BS which is to be configured at step j, and s_j the state of the environment with $\{n_1, n_2, ..., n_{j-1}\}$ configured. For each step j, an action m_j is selected by agent n_j according to the current state s_j. It should be mentioned that m_j involves the configuration of the three antennas in n_j. One step later, a new state s_{j+1} arrives as a consequence of m_j. Finally, the reward η^e is worked out with the trajectory $[m_1, m_2, ..., m_{|\mathbb{N}|}]$. In this way, the antenna configuration task is converted to a tree search problem aiming at finding out the trajectory with the largest reward.

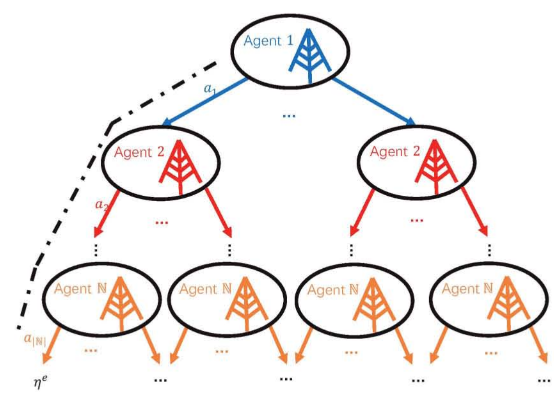


Fig. 2. Tree Search for RPO.

Monte Carlo tree search (MCTS) can effectively handle tree search problems, which explores the search tree by generating paths from the tree root iteratively until reaching a predefined computational budge, such as a time, memory or iteration constrained, at which point the search is halted and the action with best performance is returned [13]. In this paper, we adopt MCTS to explore the search tree derived from the RPO task. Denote $\{Q(s,m), V(s,m)\}$ as the statistic set of the branch

(s, m), where $Q(s, m)$ is the evaluation reward and $V(s, m)$ is the visit count. In the selection step, a path begins at the root node s_1, and finishes when reaching a node s_L in layer L which contains unexplored branches. At each layer $l < L$, an action m_l is selected from the candidate action set M_l, expressed as

$$m_l = \arg\max_{m \in M(s_l)} (Q(s_l, m) + U(s_l, m)), \tag{16}$$

where $U(s_l, m)$ denotes the upper confidence bound applied to trees [14], given by

$$U(s_l, m) = c \frac{\sqrt{\sum_{m' \in M(s_l)} V(s_l, m')}}{1 + V(s_l, m)}, \tag{17}$$

where c is a parameter determining the preference on exploration. In the exploration step, s_L is expanded by a random feasible action m_K with the statistic set of branch (s_L, m_L) initializing to

$$\{Q(s_L, m_L) = 0, V(s_L, m_L) = 0\}.$$

In the simulation step, the possible final reward through (s_L, m_L) is obtained by Monte Carlo method. Specifically, simulated configurations with all antenna parameters determined are generated from (s_L, m_L), and the reward R of (s_L, m_L) is estimated as the maximal η^e of these simulated configurations. In the backup step, R is fed back to (s_l, m_l) in each layer $h \le l \le L$ and updates the corresponding branches, written by

$$V(s_l, m_l) = V(s_l, m_l) + 1, \tag{18}$$

$$Q(s_l, m_l) = \max(Q(s_l, m_l), R). \tag{19}$$

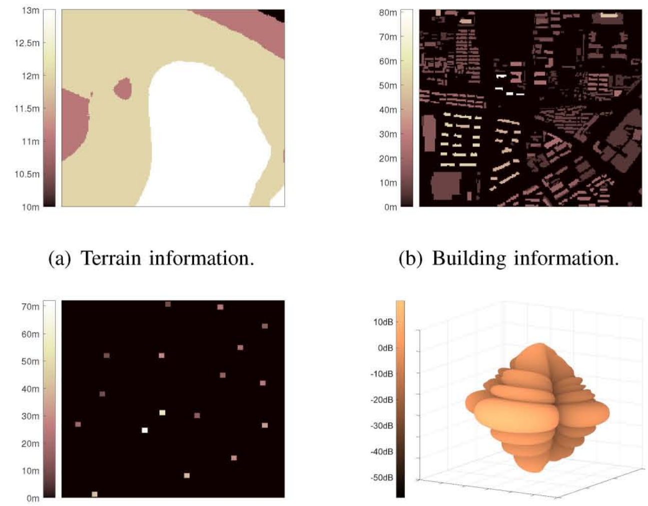
With a predefined maximal iteration G, the procedure of exploring the tree terminates, and the policy $\pi(m_h|s_h)$ is generated utilizing the statistic sets stored in all branches (s_h, m_h) where $m \in M(s_h)$, expressed as

$$\pi(m_h|s_h) = \frac{V(s_h, m_h)^\tau}{\sum_{m'_h \in M(s_h)} V(s_h, m'_h)^\tau}, \tag{20}$$

where τ denotes a temperature parameter that controls the level of exploration. The configuration m_h^* is then decided under the guidance of $\pi(m_h|s_h)$, after which the root node is updated with (s_h, m_h^*). The algorithm starts at the initial root state s_1, and terminates when reaching a leaf node where all BSs are configured.

IV. Numerical Experiment

Consider an urban area of 1km×1km, where the terrain information and the building information are presented in Fig. 3(a) and Fig. 3(b), respectively. The network in this area is composed of 17 BSs, whose locations and heights are shown in Fig. 3(c). Each antenna provides a three-dimensional directional transmit gain ranging from -58dB to 18dB, and the radiation pattern is shown in Fig. 3(d). The transmit power is 15.4dBm, and the path loss model refers to COST 231 Final [15]. The variance of noise δ^2 is -120dBm. The power threshold T^p is -95dBm, and the SINR threshold T^c is 0dB.



(a) Terrain information. (b) Building information.

(c) Site information. (d) Antenna radiation pattern.

Fig. 3. Map information and antenna radiation pattern in numerical experiment.

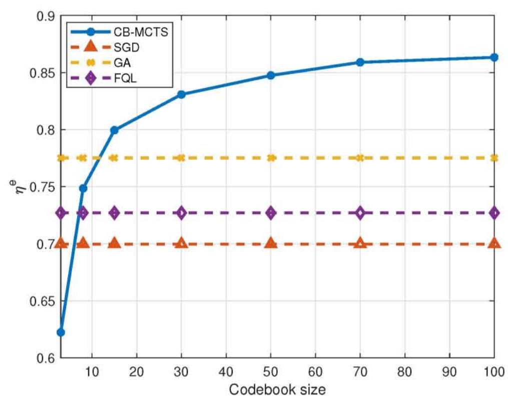
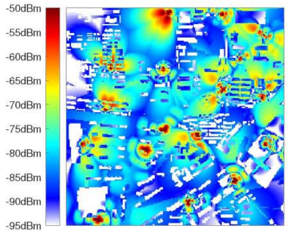
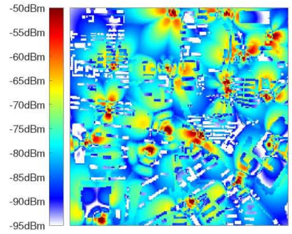


Fig. 4. Valid capacity coverage ratio under different methods.

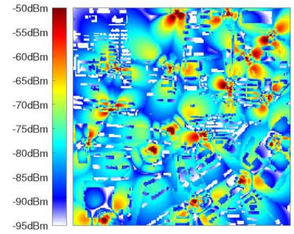
Fig. 4 shows the valid capacity coverage ratio of the network under our proposed CB-MCTS and the conventional methods, including fuzzy Q-learning (FQL) [7], stochastic gradient (S-GD) [16] and genetic algorithm (GA) [17]. The dashed curves represent the performance of conventional methods, which do not change with codebook size. The blue solid curve represents the performance of the CB-MCTS. When the codebook size is small, it is often difficult for a cell to find a proper antenna configuration to alleviate the interference to the adjacent cells, which makes the probability of interference between antennas high and results in low valid capacity coverage ratio. As the codebook size increases, each BS can better alleviate the interference to adjacent BSs, and the valid capacity coverage ratio increases. When the codebook size exceeds 70, the increase of codebook size can hardly bring performance gain since the codebooks have already contained almost all the promising



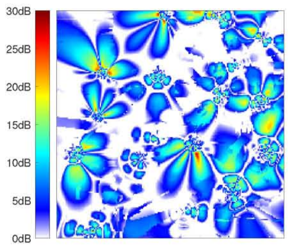
(a) The distribution of maximum signal strength with the codebook size set to 3.



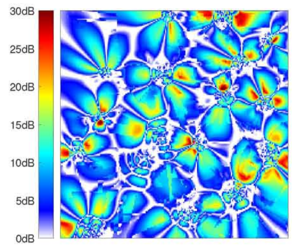
(b) The distribution of maximum signal strength with the codebook size set to 30.



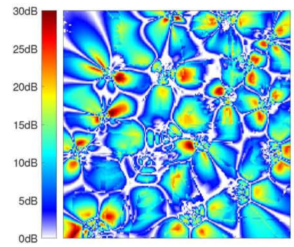
(c) The distribution of maximum signal strength with the codebook size set to 100.



(d) The distribution of SINR with the codebook size set to 3.



(e) The distribution of SINR with the codebook size set to 30.



(f) The distribution of SINR with the codebook size set to 100.

Fig. 5. The distribution of maximum signal strength and SINR of the CB-MCTS scheme under different hyperparameters.

antenna settings. The proposed CB-MCTS outperforms conventional methods by over 10% in terms of valid capacity coverage when codebook size is 100.

Fig. 5 shows the distribution of maximum signal strength and SINR of the CB-MCTS scheme under different hyperparameters. It can be seen from Fig. 5(a)-(c) that the codebook size hardly affects power coverage since the demand on signal strength is easy to meet as long as the BSs are dense enough. On the contrary, Fig. 5(d)-(f) show that the codebook size affects the capacity coverage especially when the size is small. The reason is that the small number of candidate configurations restricts the flexibility of patterns for each antenna, which leads to a high probability of interference to adjacent cells.

V. Conclusion

In this paper, we proposed a codebook based radio parameter optimization method for millimeter wave mobile communication networks, which can work out antenna configurations with promising power coverage and system throughput. The proposed method eliminates massive candidate antenna settings of each base station considering its surrounding propagation environment so as to avoid exploring the whole solution space. Numerical results show that this method outperforms the state-of-the-art ones by over 10% in terms of valid capacity coverage.

References

[1] W. Zhang *et al.*, "Toward intelligent network optimization in wireless networking: An auto-learning framework," *IEEE Wireless Commun.*, vol. 26, pp. 76–82, Jun. 2019.

[2] W. Zhao *et al.*, "Approximation algorithms for cell planning in heterogeneous networks," *IEEE Trans. Veh. Technol.*, vol. 66, pp. 1561–1572, Feb. 2017.

[3] S. Razavizadeh, M. Ahn, and I. Lee, "Three-dimensional beamforming: A new enabling technology for 5G wireless networks," *IEEE Signal Process. Mag.*, vol. 31, pp. 94–101, Oct. 2014.

[4] S. Wang, W. Zhao, and C. Wang, "Budgeted cell planning for cellular networks with small cells," *IEEE Trans. Veh. Technol.*, vol. 64, pp. 4797–4806, Oct. 2015.

[5] M. Rebato *et al.*, "Stochastic geometric coverage analysis in mmwave cellular networks with realistic channel and antenna radiation models," *IEEE Trans. Commun.*, vol. 67, pp. 3736–3752, May 2019.

[6] W. He *et al.*, "Intelligent optimization of base station array orientations via scenario-specific modeling," *IEEE Trans. Commun.*, vol. 70, pp. 2117–2130, Mar. 2022.

[7] V. Buenestado *et al.*, "Self-tuning of remote electrical tilts based on call traces for coverage and capacity optimization in LTE," *IEEE Trans. Veh. Technol.*, vol. 66, pp. 4315–4326, May 2017.

[8] L. Shen and S. Wang, "Monte Carlo Tree Search for Network Planning for Next Generation Mobile Communication Networks," in *Proc. IEEE GLOBECOM'21*, Madrid, Spain, Dec. 2021.

[9] M. Garey and D. Johnson, *Computers and intractabilityy: A Guide to the Theory of NP-Completeness*. freeman San Francisco, 1979.

[10] Z. Luo *et al.*, "Semidefinite relaxation of quadratic optimization problems," *IEEE Signal Process. Mag.*, vol. 27, pp. 20–34, May 2010.

[11] M. Grant, S. Boyd, and Y. Ye, "CVX: Matlab software for disciplined convex programming," 2008, available: https://web.stanford.edu/~boyd/software.html.

[12] R. Sutton and A. Barto, *Reinforcement learning: An introduction*. MIT press, 2018.

[13] C. Browne *et al.*, "A survey of monte carlo tree search methods," *IEEE Trans. Comput. Intell. AI Games*, vol. 4, pp. 1–43, Mar. 2012.

[14] C. Rosin, "Multi-armed bandits with episode context," *Ann. Math. Artif. Intell.*, vol. 61, pp. 203–230, Aug. 2011.

[15] E. Damosso and L. Correia, "COST action 231: Digital mobile radio towards future generation systems: Final report," 1999.

[16] Y. Liu *et al.*, "An efficient stochastic gradient descent algorithm to maximize the coverage of cellular networks," *IEEE Trans. Wireless Commun.*, vol. 18, pp. 3424–3436, Jul. 2019.

[17] Y. Qin *et al.*, "Boosting the cellular network coverage optimization in accordance with the metric structure of antenna variables," *IEEE Trans. Wireless Commun.*, vol. 19, pp. 8303–8314, Dec. 2020.