

Monte Carlo Tree Search for Network Planning for Next Generation Mobile Communication 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

Abstract—In this paper, we investigate the network planning problem in mmWave mobile communication systems, where the narrow-beam antennas can adjust azimuths and downtilts of antennas so as to maximize the power coverage of the network, as well as the system throughput. Searching for the optimal configurations of antennas generally yields a combinatorial optimization problem, which cannot be addressed even for a medium scale antenna set case. We formulate this optimization task as a finite Markov decision process, and develop a multi-layer Monte Carlo tree search method to produce a promising solution with reasonable complexity, which evaluates the outcome of given azimuth and downtilt settings without acquiring all configurations of antennas. Experiments in a real urban environment show that our proposed scheme outperforms the state-of-the-art algorithms over 10% in terms of system throughput while guaranteeing high power coverage.

Index Terms—Antenna configuration, Monte Carlo Tree Search, mmWave, network planning.

I. INTRODUCTION

Mobile network operators are obliged to improve the power coverage and the system throughput, responding to the ever-increasing service demands. Power coverage tries 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 reference signals but also the interference ones. More specifically, increasing the transmission power may improve the power coverage of a cell, however, such a power increasing also throws more interference to the adjacent cells and lower the capacity coverage of the network. As a result, network planning should achieve a balance between the power coverage and the capacity coverage.

Network planning technique continues to evolve with the development of mobile communication systems. For the GSM system, which is the main technique in the 2G mobile network, network planning generally contains two steps [1, 2]: First, a set of base stations (BSs, also referred to as cells) is selected from a candidate site list, each configured with proper RF parameters to cover the service area with strong enough signals; second, adjacent cells use different frequency bands to avoid mutual interference among cells. In contrast to the GSM system, CDMA technology is commonly used in the 3G, where the interference between adjacent cells is alleviated by carefully power controlling [3]. LTE-A system in the 4G

adopts heterogeneous network. Small cells, e.g., picocells and femtocells, cover much smaller areas than macrocells and serve the coverage holes of macrocells. Network planning for heterogeneous networks focuses on coordinating interference between small cells and macrocells [4, 5].

Millimeter wave (mmWave) is introduced to the 5G and beyond mobile networks, which brings new challenges to network planning. Electromagnetic wave in mmWave band is more likely to reflect and diffract along the propagation path due to the shorter wavelength as compared to sub-6GHz spectrum utilized nowadays, causing more severe fading. To increase the propagation distance of mmWave, sector antennas with narrow beams are generally imported in BSs [6]. In contrast to omnidirectional antenna model, the beam direction of a narrow-beam antenna in mmWave systems varies under different azimuths and downtilts, and network planning focuses on optimizing power coverage and capacity coverage by adjusting the azimuths and downtilts of the antennas [7, 8]. Searching for the optimal configuration of antennas is always an intractable combinatorial optimization problem. For example, if we try to find out the optimal azimuth and downtilt for just one antenna, it generates $72 * 181$ combinations when the azimuth range of $[0^\circ, 360^\circ)$ is discretized with an interval of 5° and the downtilt range of $[-90^\circ, 90^\circ)$ is discretized with an interval of 1° . Thus the number of azimuths and downtilts combinations for an area with 30 antennas is $(72 * 181)^{30}$, implying that enumerating all feasible solutions is impossible even with huge computing resource.

Quite a few algorithms have been developed to work out promising solutions to the aforementioned network planning problem. Tabu search explores the solution space from a feasible configuration of antennas, and maintains a memory of previous exploration to avoid cycling around local optimal solutions [9]. Genetic algorithm (GA) simulates the biological evolution process in nature, generating a series of feasible azimuth and downtilt configurations through crossover and mutation operations, and also eliminating unattractive configurations based on fitness value [10]. Stochastic gradient descent (SGD) method minimizes the continuous objective function under the guidance of gradient iteratively, which can optimize the relaxed antenna configuration problem [11]. However, the gap between the solutions produced by these algorithms and the optimal ones is not known, i.e., the performance of these algorithms cannot be guaranteed. Approximation algorithms can produce feasible solutions with worst-case performance

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

guarantee for network planning with omnidirectional antennas as discussed in [12, 13], however, they cannot be promoted to mmWave scenarios straightforwardly since the gain of a narrow-beam antenna is anisotropic. Reinforcement learning (RL) methods discussed in [14, 15] tend to evaluate the outcome of all feasible antenna configurations, which is not suitable for the scenarios with massive combinations of azimuths and downtilts due to the computation burden. The authors of [16] try to overcome the curse of dimensionality through mean field theory, which however simplifies the coupled interference between antennas.

A feasible scheme of configuration for antennas can be obtained by sequentially adjusting the azimuth and downtilt of each antenna. However, the power coverage and the system throughput for a given setting of antennas cannot be evaluated until all configurations of antenna parameters are determined since we have to take mutual interference in different cells into consideration. If we can measure the power coverage and the system throughput of a given antennas setting without acquiring all configurations of the antennas, the sequential antennas configuration process can be performed efficiently. In this paper, we formulate the antenna configuration task in mmWave mobile networks as a finite Markov Decision Process (MDP) and propose a multilayer Monte Carlo tree search (MCTS) algorithm to find a promising configuration of azimuths and downtilts. The MCTS attracts widespread attention recently due to its impressive capacity of dealing with huge scale of feasible parameter combinations [17]. For our considered network planning problem, antenna configurations with all antenna parameters determined are obtained by Monte Carlo methods from a given setting without acquiring all antenna configurations, and the power coverage and the system throughput are utilized to evaluate the performance of the given setting. Experiment results show that our proposed scheme yields a promising configuration of antennas in real environments with reasonable computation load, and outperforms the state-of-the-art schemes over 10% in terms of system throughput while guaranteeing high power coverage.

The remainder of this paper is organized as follows. We describe system model in Section II, as well as problem formulation. In Section III, we transform the network planning task to a finite MDP and present the MCTS algorithm. In Section IV, experiment results are given with discussions. Conclusions are drawn in Section V.

II. SYSTEM MODEL AND PROBLEM FORMULATION

Consider an urban area containing multiple BSs, each of which carrying three sector antennas, as shown in Fig. 1. Let B denote the set of BSs and A the set of antennas, where $|A| = 3|B|$. Divide the area into $5\text{m} \times 5\text{m}$ grids, the grid centers, also referred to as the traffic demand points (TDPs), are taken as reference points to measure the signals in the corresponding grids. Let U denote the set of TDPs, power coverage η^p refers to the ratio of the TDPs at which the reference signal power surpasses a certain threshold T^p to the total TDPs $|U|$, while capacity coverage η^c refers to the ratio

of the TDPs at which the SINR exceeds a certain threshold T^c to $|U|$. Our optimization task is to maximize η^p and η^c .

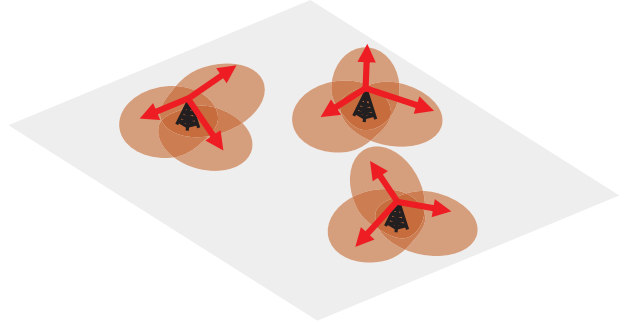


Fig. 1. Network layout.

Define a standard coordinate system ($o-xyz$), for antenna $a \in A$ where azimuth θ_a is the angle between the horizontal projection of the main lobe direction of a and x -axis, and downtilt ϕ_a is the angle between the main lobe direction of a and the horizontal plane, as shown in Fig. 2. For each TDP $u \in U$, the reference signal strength received from a is the function with regard to θ_a and ϕ_a , denoted by $P_{a,u}(\theta_a, \phi_a)$.

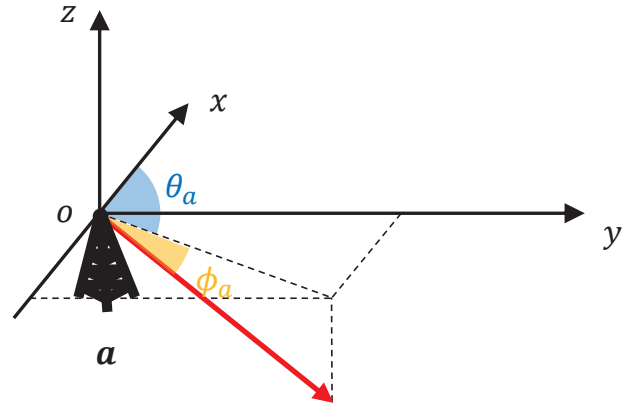


Fig. 2. Azimuth and downtilt of a sector antenna.

Let $\theta = \{\theta_a, a \in A\}$ and $\phi = \{\phi_a, a \in A\}$ denote the azimuth set and downtilt set of antennas, respectively. The maximum signal strength of u is

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

and the power coverage η^p is

$$\eta^p(\theta, \phi) = \frac{\sum_{u \in U} \mathbb{1}(P_u(\theta, \phi) > T^p)}{|U|}, \quad (2)$$

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

$$\mathbb{1}(x) = \begin{cases} 1 & , \text{ if } x \text{ is true,} \\ 0 & , \text{ otherwise.} \end{cases} \quad (3)$$

A TDP can only be served by one antenna in a transmission time interval. Let $\hat{a}_u = \arg \max_{a \in A} P_{a,u}(\theta_a, \phi_a)$ denote the

antenna that provides the strongest signal for TDP u , the SINR at u is

$$\rho_u(\boldsymbol{\theta}, \boldsymbol{\phi}) = \frac{P_u(\boldsymbol{\theta}, \boldsymbol{\phi})}{\delta^2 + \sum_{\substack{a \neq a_u, \\ a \in A}} P_{a,u}(\theta_a, \phi_a)}, \quad (4)$$

where δ^2 denotes the power of additive white Gaussian noise. The capacity coverage η^c can be expressed as

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

The effective coverage ratio η^e , referred to the ratio of the TDPs that meet the demand of both power coverage and capacity coverage to $|U|$, is given by

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

The value range of the downtilts can be reduced to $[0, 20]$ since antennas can hardly provide effective service for TDPs when the downtilts are too upward or downward [18]. With the range of azimuths discretized with an interval of 5° and the range of downtilts discretized with an interval of 1° , the formulated problem can be expressed as

$$\begin{aligned} \max_{\boldsymbol{\theta}, \boldsymbol{\phi}} \quad & \eta^e(\boldsymbol{\theta}, \boldsymbol{\phi}), \\ \text{s.t.} \quad & \theta_a \in \{0^\circ, 5^\circ, 10^\circ, \dots, 355^\circ\}, \forall a \in A, \\ & \phi_a \in \{0^\circ, 1^\circ, 2^\circ, \dots, 20^\circ\}, \quad \forall a \in A. \end{aligned} \quad (7)$$

III. MCTS BASED NETWORK PLANNING

Equation (7) defines a combinational optimization problem. The scale of all combinations of azimuths and downtilts is $(72 \times 21)^{|U|}$, which is impossible to enumerate even for medium $|U|$. It can be transformed to a finite MDP where antennas sequentially choose its azimuth and downtilt according to the configured antenna parameters. For antenna $a_i \in A$ at step i , action of $\theta_{a_i} \in \{0^\circ, 5^\circ, 10^\circ, \dots, 355^\circ\}$ and $\phi_{a_i} \in \{0^\circ, 1^\circ, 2^\circ, \dots, 20^\circ\}$ is selected according to the setting of antenna parameters (also referred to as state) S_i . The next state S_{i+1} appears with the configuration of $\{\theta_{a_i}, \phi_{a_i}\}$. The MDP thereby gives rise to a sequence as

$$\begin{aligned} S_1, \{\theta_{a_1}, \phi_{a_1}\}, S_2, \{\theta_{a_2}, \phi_{a_2}\}, \dots, \\ S_{|A|}, \{\theta_{a_{|A|}}, \phi_{a_{|A|}}\}, S_{fin}, Q_{fin}. \end{aligned} \quad (8)$$

where S_1 denotes the initial state, S_{fin} denotes the final state with all antennas configured, and Q_{fin} denotes the final effective coverage, equivalent to $\eta^e(\boldsymbol{\theta}, \boldsymbol{\phi})$. The azimuths of the three antennas in one BS differ from each other by 120 degrees to simplify computation. Let $\boldsymbol{\theta}_B$ denote the azimuth set of all BSs, the range of $\theta_b \in \boldsymbol{\theta}_B$ is reduced to $\{0^\circ, 5^\circ, 10^\circ, \dots, 115^\circ\}$ due to symmetry. The problem can be solved with BSs configuring their azimuths and downtilts sequentially, expressed as

$$\begin{aligned} S_1, \{\theta_{b_1}, \phi_{a_1^1}, \phi_{a_1^2}, \phi_{a_1^3}\}, S_2, \{\theta_{b_2}, \phi_{a_2^1}, \phi_{a_2^2}, \phi_{a_2^3}\}, \\ \dots, S_{|B|}, \{\theta_{b_{|B|}}, \phi_{a_{|B|}^1}, \phi_{a_{|B|}^2}, \phi_{a_{|B|}^3}\}, S_{fin}, Q_{fin}. \end{aligned} \quad (9)$$

where $b_i \in B$ denotes the selected BS at step i and a_i^j denotes the selected antenna j in b_i . This MDP can be solved by computing the optimal value in a search tree with a breadth of 24×21^3 and a depth of $|B|$. The initial state S_1 refers to the root of the search tree, while all possible S_i denote the nodes of layer i in the search tree. The pairs of S_i and $\{\theta_{b_i}, \phi_{a_i^1}, \phi_{a_i^2}, \phi_{a_i^3}\}$ form the edges of the search tree. Assume that each configuration of azimuth or downtilt obtains value, the MDP can be rewritten as

$$\begin{aligned} S_1^\theta, Q_1^\theta, \theta_{b_1}, S_1^{\phi_1}, Q_1^{\phi_1}, \phi_{a_1^1}, S_1^{\phi_2}, Q_1^{\phi_2}, \phi_{a_1^2}, S_1^{\phi_3}, Q_1^{\phi_3}, \\ \phi_{a_1^3}, S_2^\theta, Q_2^\theta, \dots, S_{|B|}^\theta, \theta_{b_{|B|}}, S_{|B|}^{\phi_1}, Q_{|B|}^{\phi_1}, \phi_{a_{|B|}^1}, \\ S_{|B|}^{\phi_2}, Q_{|B|}^{\phi_2}, \phi_{a_{|B|}^2}, S_{|B|}^{\phi_3}, Q_{|B|}^{\phi_3}, \phi_{a_{|B|}^3}, S_{fin}, Q_{fin}. \end{aligned} \quad (10)$$

where S_i^θ and Q_i^θ denote the state and value before BS b_i configures azimuth, $S_i^{\phi_j}$ and $Q_i^{\phi_j}$ denote the state and value before antenna a_i^j configures downtilt. Thus, the breadth of search tree turns into 24, while the depth turns into $4|B|$.

Enumerating all branches of the search tree is infeasible, but the exploration of the search tree can be more effective if there exists a reasonable policy $\pi(m|S)$ where branches corresponding to the antenna configuration with high effective coverage are more likely to be explored. To obtain such a policy, the contribution of action m at different states S should be evaluated accurately. However, the effective coverage cannot be evaluated until all configurations of antennas are acquired. Monte Carlo simulation is a promising approach to acquire configurations of all antenna parameters from the setting without all configurations acquired, and the effective coverage ratio achieved by these simulated configurations can be used to evaluate the value of the setting, which is widely used in MCTS.

A. MCTS Method

In the MCTS procedure, each state S in the search tree contains edges (S, m) for all feasible actions $m \in M(S)$, and each edge stores a set of data,

$$\{Q(S, m), N(S, m)\},$$

where $Q(S, m)$ denotes the evaluation value, and $N(S, m)$ denotes the visit count. The search tree only contains root node at the beginning, and expands in the MCTS process. A round of MCTS consists of four steps:

- **Selection.** The selection procedure begins at the root node of the search tree and finishes when reaching a leaf node S_L at step L . At step $i < L$, an action m_i is selected according to the data set at all (S_i, m) in layer i , $m \in M(S_i)$, i.e.,

$$m_i = \arg \max_{m \in M(S_i)} (Q(S_i, m) + U(S_i, m)), \quad (11)$$

where $U(S_i, m)$ denotes the upper confidence bound based on the estimated value and the visit count according to [19]:

$$U(S_i, m) = c \frac{\sqrt{\sum_{l \in M(S_i)} N(S_i, l)}}{1 + N(S_i, m)}, \quad (12)$$

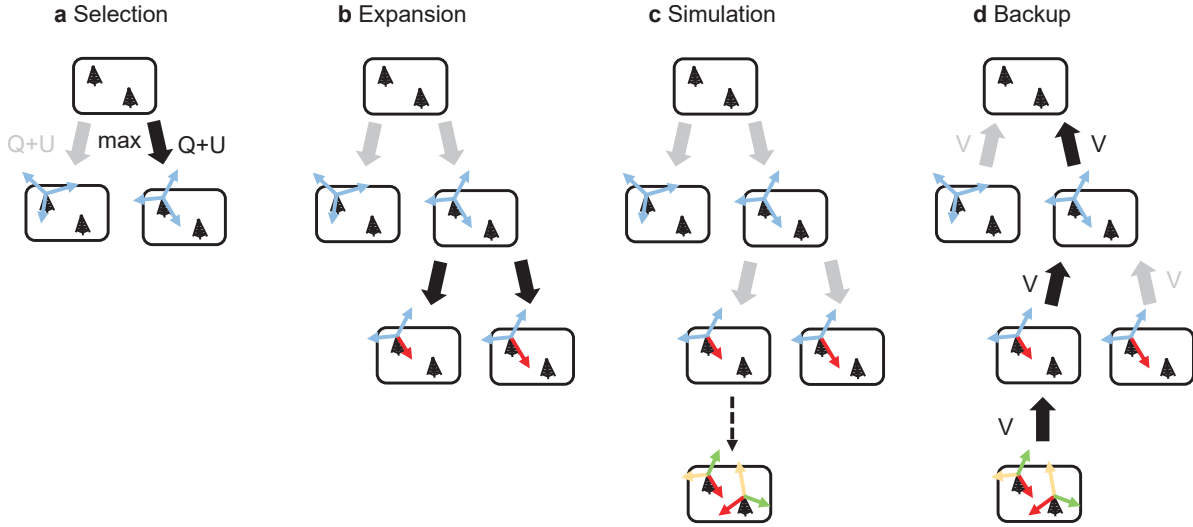


Fig. 3. S-MCTS for configuration of an azimuth or downtilt.

where c denotes the preference parameter of exploration. This selection strategy prefers action m with low visit count $N(S_i, m)$ at the early stage, and tends to select action m with high value $Q(S_i, m)$ with more evaluations obtained.

- **Expansion.** The search tree is expanded from S_L via all possible actions $m \in M(S_L)$ with the data sets initializing to

$$\{Q(S_L, m) = 0, N(S_L, m) = 0\}.$$

- **Simulation.** From a feasible action m_L of S_L , configurations are obtained by Monte Carlo simulation according to (S_L, m_L) , and interact with the environment to get the estimation value v .
- **Backup.** The simulated estimation value v is backed up through each layer $i \leq L$, with the data set of (S_i, m_i) updated as follows:

$$N(S_i, m_i) = N(S_i, m_i) + 1, \quad (13)$$

$$Q(S_i, m_i) = Q(S_i, m_i) + \frac{v - Q(S_i, m_i)}{N(S_i, m_i)}. \quad (14)$$

B. MCTS for Antenna Configuration

We propose a multi-layer MCTS framework to sequentially configure antenna parameters of each BS. Algorithm 1 (S-MCTS) illustrates the process of configuring the azimuth or downtilt of an antenna. First, we initialize a search tree with a root state S_r which stands for the existing setting of antenna parameters. Actions of azimuths or downtilts are selected according to (11) from S_r to a leaf node S_{cur} which has not be explored as shown in Fig. 3a. S_{cur} is then expanded with $\{Q(S_{cur}, m) = 0, N(S_{cur}, m) = 0\}$ for all feasible actions $m \in M(S_{cur})$ as depicted in Fig. 3b). Then multiple configurations of all antenna parameters are obtained from a

Algorithm 1 S-MCTS Algorithm

- 1: Initialization: S_r
 - 2: **for** $d = 1 : D$ **do**
 - 3: $S_{cur} \leftarrow S_r$
 - 4: **while** S_{cur} is not a leaf node **do**
 - 5: Select m_{cur} from all feasible (S_{cur}, m) using (11)
 - 6: Update $S_{cur} \leftarrow (S_{cur}, m_{cur})$
 - 7: **end while**
 - 8: Expand S_{cur} via all feasible $m \in M(S_{cur})$ with the data sets initializing to $\{Q(S_{cur}, m) = 0, N(S_{cur}, m) = 0\}$
 - 9: Randomly choose a feasible m_{cur}
 - 10: $v \leftarrow \text{simulate}(S_{cur}, m_{cur})$ using (15)
 - 11: $S_{old} \leftarrow \emptyset$
 - 12: **do**
 - 13: Backup v to (S_{cur}, m_{cur}) using (13), (14)
 - 14: Update $S_{old} \leftarrow S_{cur}$
 - 15: Update $S_{cur} \leftarrow \text{parent node of } S_{cur}$
 - 16: Update $m_{cur} \leftarrow m$ that connects S_{cur} and S_{old}
 - 17: **while** $S_{cur} \neq S_r$
 - 18: **end for**
 - 19: Generate policy $\pi(m|S_r)$ using (16)
 - 20: $m^* \leftarrow \pi(m|S_r)$
 - 21: **return** m^*
-

random edge of S_{cur} by Monte Carlo simulations, and the evaluation value v of this edge (S_{cur}, m_{cur}) is calculated as

$$v = \max_{k \in \{1, 2, \dots, K\}} \eta^e(\theta_k, \phi_k), \quad (15)$$

where K is the number of Monte Carlo simulations, and θ_k, ϕ_k are the set of azimuths and downtilts in the k -th simulation, respectively. The procedure of simulation is shown in Fig. 3c. Finally, v is back propagated from (m_{cur}, S_{cur}) to (m_r, S_r) as described in Fig. 3d. With D rounds of MCTS,

the policy $\pi(m|S_r)$ is generated utilizing the data set stored in each edge (m, S_r) :

$$\pi(m|S_r) = \frac{N(S_r, m)^\tau}{\sum_l N(S_r, l)^\tau}, \quad (16)$$

where τ denotes a temperature parameter that controls the level of exploration. As the generation of $\pi(m|S_r)$ terminates, the action m^* is worked out under the guidance of $\pi(m|S_r)$ and output.

Algorithm 2 M-MCTS Algorithm

- 1: Initialization: $S_r \leftarrow S_1^\theta$
 - 2: **for** $i = 1 : |B|$ **do**
 - 3: Randomly select a BS $b_i \in B$
 - 4: $\theta_{b_i} \leftarrow \text{S-MCTS}(S_r)$
 - 5: Update $S_r \leftarrow (S_r, \theta_{b_i})$
 - 6: $\phi_{a_i^1} \leftarrow \text{S-MCTS}(S_r)$
 - 7: Update $S_r \leftarrow (S_r, \phi_{a_i^1})$
 - 8: $\phi_{a_i^2} \leftarrow \text{S-MCTS}(S_r)$
 - 9: Update $S_r \leftarrow (S_r, \phi_{a_i^2})$
 - 10: $\phi_{a_i^3} \leftarrow \text{S-MCTS}(S_r)$
 - 11: Update $S_r \leftarrow (S_r, \phi_{a_i^3})$
 - 12: **end for**
 - 13: **return** $\theta_B = \{\theta_{b_1}, \theta_{b_2}, \dots, \theta_{b_{|B|}}\}$,
 $\phi = \{\phi_{a_1^1}, \phi_{a_1^2}, \phi_{a_1^3}, \dots, \phi_{a_{|B|}^1}, \phi_{a_{|B|}^2}, \phi_{a_{|B|}^3}\}$
-

The configuration for all antennas can be produced by a process where each BS sequentially adjusts its azimuth and downtilts to cover the nearby area while restricting mutual interference, and the procedure is shown in Algorithm 2. First, we initialize the root state $S_r = S_1^\theta$ with no azimuth or downtilt decided. At each turn $i = 1$ to $|B|$, BS b_i is randomly selected since network performance is not affected by the sequence of configuration statistically. Then the azimuth and downtilts of b_i are individually decided by the S-MCTS algorithm. Finally, we collect all the azimuths and downtilts to get the result.

IV. EXPERIMENT RESULTS

We compare the proposed M-MCTS algorithm with GA [10] and SGD [11] in realistic network scenarios. Consider a service area of 1km * 1km with 12 BSs in a city where buildings are densely deployed. The map information of the area includes terrain information, building information, and BS information, as shown in Fig. 4(a-c). The detailed gain distribution of a narrow-beam antenna is shown in Fig. 4(d), ranging from -58dB to 18dB . The propagation model refers to the model in COST 231 Final [20]. The variance of noise δ^2 is -120dBm . The power threshold T_P is -95dBm , and the SINR threshold T_C is 3dB . The exploration preference parameter c is 0.05 , and the number of MCTS rounds D is 1000 . The number of MCTS simulations per round K is 10 .

Fig. 5 shows the complementary cumulative distribution function (CCDF) of the maximum signal strength received at TDPs. As we can see, over 85% of the TDPs receive strong enough signals that exceed power threshold for all the

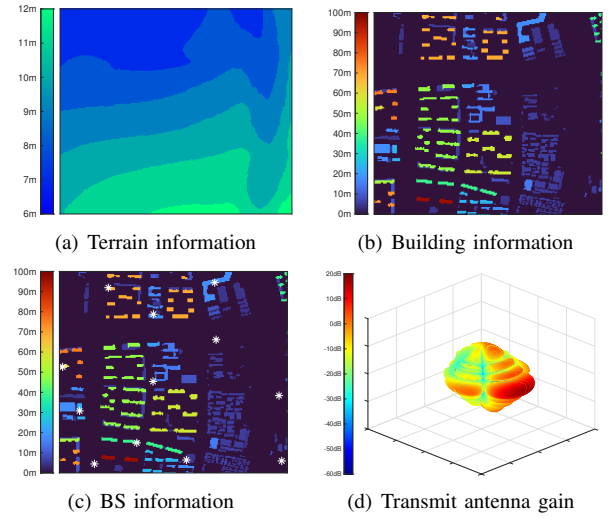


Fig. 4. Map and antenna information in numerical experiment.

schemes. It indicates that most of TDPs can obtain strong strength reference signal in network scenarios where BSs are densely deployed.

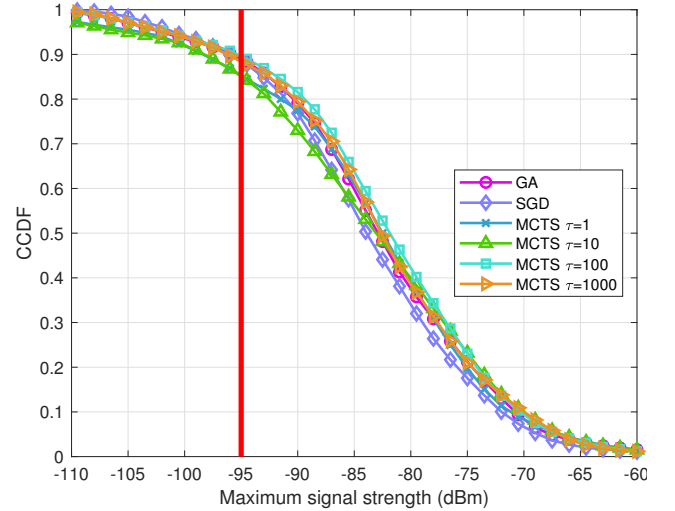


Fig. 5. Complementary cumulative distribution of maximum signal strength.

Fig. 6 shows the CCDF of SINR at TDPs. As depicted in this figure, 64.71% of the TDPs can obtain target SINR with the M-MCTS algorithm when the temperature parameter $\tau = 100$, while the ratio of TDPs reaching target SINR with GA and SGD are 56.59% and 50.40%, respectively. This implies that more TDPs can achieve high transmission rate with the proposed M-MCTS planning scheme.

Fig. 7 shows the effective coverage ratio as a function of SINR threshold, where the power threshold is fixed at -95dBm . When the SINR threshold is 3dB , the M-MCTS scheme achieves an effective coverage ratio of 60.69% with $\tau = 100$, while GA is 54.56% and SGD only reaches 48.33%. The proposed M-MCTS scheme outperforms the others by

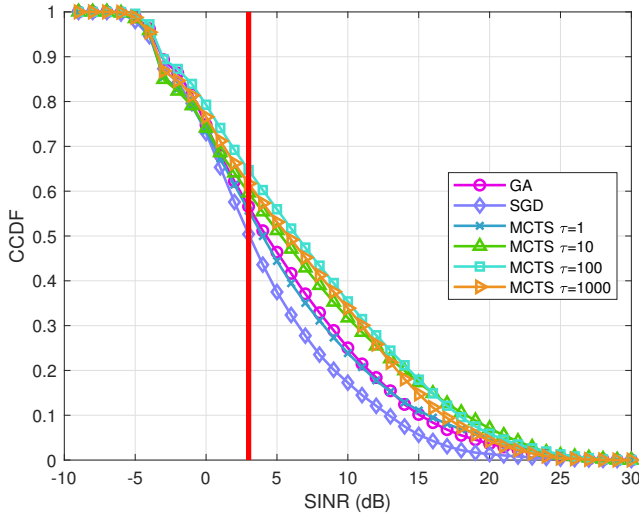


Fig. 6. Complementary cumulative distribution of received SINR.

more than 10% in terms of system throughput. More specifically, the effective coverage ratio only depends on the SINR distribution when the power threshold is fixed, which results in the consistency of the trend of the corresponding curves in Fig. 6 and Fig. 7.

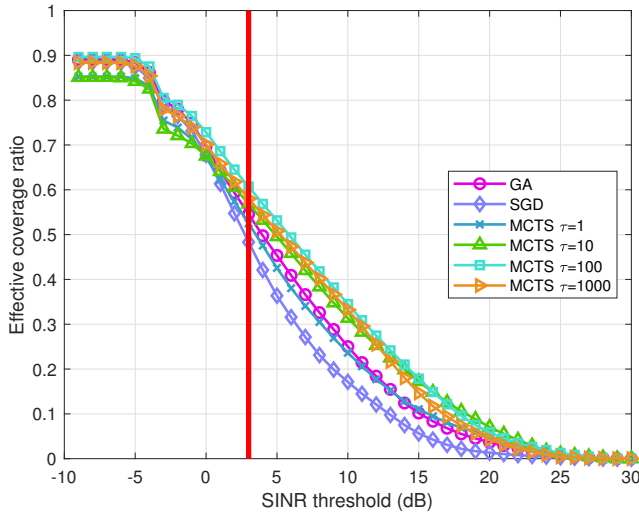


Fig. 7. Effective coverage as a function of SINR threshold.

V. CONCLUSION

In this paper, we converted the network planning problem in mmWave mobile communication systems to a finite Markov Decision Process, and developed an Monte Carlo tree search method to produce promising azimuths and downtilts. Experiment results show that our proposed algorithm outperforms the state-of-the-art ones by more than 10% in terms of system throughput while high power coverage is guaranteed, indicating that MCTS can explore the high-dimensional solution space of antenna configurations effectively. For future work,

the convergence speed analysis and the performance bound are interesting topics.

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, pp. 1678–1691, Sept. 2000.
- [2] K. Tutschku and P. Tran-Gia, "Spatial traffic estimation and characterization for mobile communication network design," *IEEE J. Sel. Areas Commun.*, vol. 16, pp. 804–811, Jun. 1998.
- [3] 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, pp. 939–952, Sept. 2003.
- [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] W. Zhao, S. Wang, C. Wang, and X. Wu, "Approximation algorithms for cell planning in heterogeneous networks," *IEEE Trans. Veh. Technol.*, vol. 66, pp. 1561–1572, Feb. 2017.
- [6] T. Bai and R. Heath, "Coverage and rate analysis for millimeter-wave cellular networks," *IEEE Trans. Wireless Commun.*, vol. 14, pp. 1100–1114, Oct. 2015.
- [7] T. Rappaport *et al.*, "Broadband millimeter-wave propagation measurements and models using adaptive-beam antennas for outdoor urban cellular communications," *IEEE Trans. Antennas Propag.*, vol. 61, pp. 1850–1859, Dec. 2013.
- [8] 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.
- [9] S. Hurley, S. Allen, D. Ryan, and R. Taplin, "Modelling and planning fixed wireless networks," *Wireless Netw.*, vol. 16, pp. 577–592, Apr. 2010.
- [10] Y. Yoon and Y. Kim, "An efficient genetic algorithm for maximum coverage deployment in wireless sensor networks," *IEEE Trans. Cybern.*, vol. 43, pp. 1473–1483, Oct. 2013.
- [11] Y. Liu, W. Huangfu, H. Zhang, and K. Long, "An efficient stochastic gradient descent algorithm to maximize the coverage of cellular networks," *IEEE Trans. Wireless Commun.*, vol. 18, pp. 3424–3436, Jul. 2019.
- [12] W. Zhao and S. Wang, "Cell planning for heterogeneous cellular networks," in *Proc. IEEE WCNC*, pp. 1032–1037, Apr. 2013.
- [13] S. Wang and X. Huang, "Aggregation points planning for software-defined network based smart grid communications," in *Proc. IEEE INFOCOM*, Apr. 2016.
- [14] R. Razavi, S. Klein, and H. Claussen, "A fuzzy reinforcement learning approach for self-optimization of coverage in LTE networks," *Bell Labs Tech. J.*, vol. 15, pp. 153–175, Nov. 2010.
- [15] M. Islam and A. Mitschele-Thiel, "Reinforcement learning strategies for self-organized coverage and capacity optimization," in *Proc. IEEE WCNC*, pp. 2818–2823, Apr. 2012.
- [16] E. Balevi and J. Andrews, "Online antenna tuning in heterogeneous cellular networks with deep reinforcement learning," *IEEE Trans. Cogn. Commun. Netw.*, vol. 5, pp. 1113–1124, Dec. 2019.
- [17] D. Silver *et al.*, "Mastering the game of go without human knowledge," *Nature*, vol. 550, pp. 354–359, Oct. 2017.
- [18] 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, no. 5, pp. 4315–4326, 2017.
- [19] L. Kocsis and C. Szepesvári, "Bandit based Monte-Carlo planning," in *Proc. ECML*, pp. 282–293, Sept. 2006.
- [20] E. Damosso and L. Correia, "COST action 231: Digital mobile radio towards future generation systems: Final report," *Brussels: European Union Publications*, pp. 190–207, 1999.