Received November 8, 2018, accepted November 28, 2018, date of publication December 6, 2018, date of current version January 4, 2019.

Digital Object Identifier 10.1109/ACCESS.2018.2885313

PCI Planning Based on Binary Quadratic Programming in LTE/LTE-A Networks

JIHONG GUI[1], ZHIPENG JIANG, AND SUIXIANG GAO
[1]School of Mathematical Sciences, University of Chinese Academy of Sciences, Beijing 100049, China
Corresponding author: Zhipeng Jiang (13611356256@139.com)

ABSTRACT In recent years, interference has played an increasingly significant part in bulkier and denser Long Term Evolution (LTE/LTE-Advanced) networks. Though intra-cell interference is successfully improved by Orthogonal Frequency Division Multiple Access (OFDMA), inter-cell interference (ICI) could cause a degradation of throughput and significantly impact Signal-to-Noise-Ratio (SINR) in the downlink (DL) network. Physical Cell ID (PCI) planning, an effective approach to eliminate ICI, is required to reduce collision, confusion and mod q interference, where $q = 3$ for Single-Input Single-Output (SISO) system, and $q = 6$ for Multiple-Input Multiple-Output (MIMO) system. In this study, a new definition of neighborhood relations was proposed based on the measurement report (MR) data in the actual network. Binary quadratic programming (BQP) model was built for PCI planning through a series of model deductions and mathematical proofs. Since BQP is known as NP-hard, a heuristic Greedy algorithm was proposed and its low complexity both in time and space can ensure large-scale computing. Finally, based on the raw data extracted from the actual SISO system network and the simulation calculation of MATLAB, the experimental results demonstrated that Greedy algorithm not only eliminates conflict and confusion completely, but also reduces the mod 3 interference of 26.213% more than the baseline scheme and far more than the improvement ratio of 4.436% given by the classical graph coloring algorithm.

INDEX TERMS LTE Network, ICI, PCI Planning, BQP Problem, Greedy algorithm

I. INTRODUCTION

A. MOTIVATION

As the base station (BS) and user equipment (UE) are booming, existing wireless communication network becomes bulkier and denser. However, the complexity of this network has made it difficult to manage and maintain by network operators and engineers [1], [2]. Third Generation Partnership Project (3GPP) has adopted Long Term Evolution (LTE/LTE-A) standard to improve network throughput and satisfy the data demands of end-user [3]–[5]. LTE network is based on the single carrier Frequency Division Multiple Access for uplink (UL) network and Orthogonal Frequency Division Multiple Access (OFDMA) for downlink (DL) network [6], [7].

Though OFDMA technique has been employed in downlink network to eliminate intra-cell interference including inter symbol interference and inter code interference [8], [9], frequency reuse factor leads to Inter-Cell Interference (ICI), which significantly affects Signal-Interference-plus-to-Noise-Ratio (SINR) of active UEs, especially UEs in cell-edge. As a result, the total throughput is significantly degraded [10], [11].

Physical Cell ID (PCI) planning is considered the most effective solution to reduce ICI and increase the throughput over Physical Downlink Control Channels (PDCCHs) under the constraint of non-orthogonal control regions between neighboring cells [12]. PCI, a significant parameter in the DL network, is the cell identifier in the physical layer. It is employed by the BSs for channel estimation and separation of cells in the handover (HO) measurements [13], [14]. The number of PCIs is limited to 504 due to the combined special construction of the Primary Synchronization Signal (PSS) and the Secondary Synchronization Signal (SSS) [15]. This helps UEs achieve frequency and time synchronization during the cell search stage. To ensure the proper cell search, the PCI planning should reduce the impacts of ICI including collision, confusion, and mod 3 interference. As a result, given the limited PCIs and massive cells, the proper automated configuration of PCIs in the whole network is necessary to reduce the manual operations.

B. RELATED WORK

A large number of recent studies on PCI planning are investigated as follows. Abdulkareem *et al.* [16] presented a robust algorithm to configure and assign PCIs with guaranteeing conflict free and minimum PCI reassignment. An approach in the field of automatic PCI assignment in LTE-A network was developed in [17]. In the study, an algorithm was established based on graph coloring problem (GCP) allowing for multi-target optimization in different network deployment scenarios to minimize the number of PCI conflicts and confusion. Liu *et al.* [18] proposed an automatic centralized PCI assignment mechanism using operation administration and maintenance as central server to collect cell information of network. Besides, an improved graph-coloring algorithm was presented, which can significantly reduce time complexity while keeping a high PCI utility ratio to create a collision and confusion free PCI for new cells. The PCI self-configuration problem was mapped as a well-known minimum spanning tree (MST) to optimize the PCI reuse distance and avoid collision and confusion in the whole network [19]. In the meantime, some studies built combinatorial optimization models for PCI planning and solved them with intelligence algorithm, e.g., genetic algorithms [20], [21]. Besides, based on the considerations of collision and confusion free between cells of different layers, some recent works have extended the analysis of PCI planning to heterogeneous LTE network [22]–[24].

However, due to the superposition of reference signal (RS), PCI mod 3 interference in downlink network will significantly impact the channel evaluation in the case of multiple antennas (e.g., dual antenna or eight antenna). This will degrade SINR estimates [25], which will be reported by UE and later used to select the Modulation and Code Scheme (MCS). This leads to wrong evaluation of channel quality indicator (CQI) and the delay of transmission in the downlink network [26]. Acedo-Hernndez *et al.* [27] presented a comprehensive performance analysis of the significant effects of PCI mod 3 interference on LTE downlink performance. Acedo-Hernndez *et al.* [28] presented a graph partition algorithm to reduce PCI collision, confusion and Cell-specific Reference Signal (CRS) collisions, e.g., PCI mod 3 interference.

Moreover, some relevant studies on binary quadratic programming (BQP) problems are to be presented here since BQP is the final model in our work. BQP problem is generally known as NP-hard [29] and has numerous important applications in many fields, including Max-Cut [30] and vertex coloring problem [31] in graph theory, capital budgeting and financial analysis problems [32]–[34] in economics, and image segmentation [35]–[37] in computer vision. Spectral relaxation and semi-definite relaxation are two classic methods broadly used for dealing with convex BQP problems. However, it is hard for spectral relaxation to ensure properties of solution in many cases because the loose bound [38]–[40] and semi-definite relaxation will cost large computational complexity for large-scale problems [41]. In the meantime, there are various heuristic algorithms for quickly finding a high quality solution in large-scale problems. For instance, in [42] and [43], Scatter Search and Tabu Search algorithms are used to solve unconstrained binary quadratic programming (UBQP) problems. Greedy and Random Greedy algorithms can be also used to solve UBQP problems [44]. For non-convex BQP problems, Branch and Bound method (BBM) is suitable for small-scale problems with exponential complexity [45], [46], which has been applied in existing optimization software.

C. OUR CONTRIBUTIONS

In this study, a new definition of neighborhood relations based on Electro-Magnetic Interference (EMI) of signals was proposed in the downlink network thereby giving a quantitative mathematical description for ICI in various scenarios. Finally, a general BQP optimization model was built, which synthetically considers collisions, confusion, and mod 3 interference through a series of mathematical transformations and equivalent proofs. BQP problem is known as NP-hard, so it is hard to find an exact solution in polynomial time unless $P = NP$, especially for the non-convex objective function with constraints. Though the algorithms in [44] can solve the UBQP well, it is invalid for this problem because it cannot even ensure the solution is feasible. Also, other existing algorithms, e.g., BBM will cost considerable time complexity and space complexity. Thus, a heuristic greedy algorithm is originally designed based on the special property of constraints. The time complexity of this algorithm is $O(m^2n^3)$ and space complexity is only $O(mn + n^2)$, where m and n denote the amount of PCI and cells respectively. The Greedy algorithm with low complexity can be applied for the large-scale problems easily. To evaluate the performance of the proposed algorithm, this algorithm is compared with the scheme implemented in the live network and classical graph coloring algorithm based on actual data in LTE downlink network provided by China Mobile Group Beijing Company Limited (CMBJ).

The rest of this study is organized as follows. Section 2 gives the definition of PCI planning problem. Section 3 builds a BQP optimization model and presents some basic analysis of PCI problem's hardness. Section 4 and 5 provide the Greedy algorithm and a series of performance analysis, respectively, based on multiple simulation experiments. Section 6 draws the conclusion.

II. THE PCI PLANNING PROBLEM

A. THE DEFINITION OF PCI

As described in the introduction, the number of PCIs is limited to 504 and each of them is indicated by SSS and PSS in the downlink network. SSS is valued from 0 to 167, and PSS is valued from 0 to 2. The definition of PCI is given as [47]:

$$PCI = 3 * SSS + PSS \tag{1}$$

where the minimum of PCI value is 0 when $SSS = PSS = 0$ and the maximum is 503 when $SSS = 167, PSS = 2$.

B. THE DEFINITION OF NEIGHBORHOOD RELATION

It is very necessary to clarify the neighborhood relations before allocating PCI to each cell. The traditional neighborhood relation of different cells depends on whether coverage areas of cells are conterminous or separable. In the meantime, the coverage area of each cell is often theoretically assumed as a regular hexagon of cellular network or an irregular figure generated by the emulation of field intensity distribution practically.

However, PCI planning should consider the EMI of signals in the downlink network since the coverage area of cell is over idealistic or based on propagation model. EMI occurs because UE receives signals from multiple cells but only access one cell. Let binary variable $\psi_{u,i}$ denote the access status between user u and cell i. Then $\psi_{u,i} = 1$ if user u accesses cell i and $\sum_i \psi_{u,i} = 1, \forall u$ because of the unique access of each user u. Moreover, let Ψ_i denote the set of users who access cell i, then

$$\Psi_i = \{u|\psi_{u,i} = 1, \forall u\} \tag{2}$$

In fact, UE u receives signals from multiple cells but only accesses cell i. Then cell i is the unique master cell. In the meantime, interference may result from the signal from cell j. Let $P_{u,i}$ and $P_{u,j}$ denote the received power of UE u from cell i and j, respectively. Moreover, the distance between cell i and j is expressed as $d_{i,j}$ and the neighborhood relation set of ordered pair (i, j) is denoted by Γ formed as

$$\Gamma = \{(i,j)|P_{u,i} - P_{u,j} \le \delta_s \text{ and } d_{i,j} \le \delta_d,\ \exists u \in \Psi_i\} \tag{3}$$

where δ_s and δ_d denote the threshold of signal strength and distance, respectively. Then, cell i is the master cell, and cell j is the neighbor cell if ordered pair $(i, j) \in \Gamma$ and they make up a couple of neighborhood relations.

Fig. 1 gives an example of neighborhood relation. The UE in Fig. 1 accesses to cell 1 as marked by red full line and cell 1 is the master cell. In the meantime, it receives signals from neighbor and other cells, illustrated by purple and white dotted line, respectively. To be specific, ordered pair set $\Gamma = \{(1, 2), (1, 4), (1, 7), (1, 9)\}$ as $P_{u,1} - P_{u,j} \le \delta_s$ and $d_{1,j} \le \delta_d$ are met simultaneously, $\forall j \in \{2, 4, 7, 9\}$. Accordingly, neighbor cells 2, 4, 7, 9 and master cell 1 make up 4 couples of neighborhood relations.

Furthermore, the co-frequency neighborhood relation set of ordered pair (i, j) is denoted by Γ_f, which is expressed as

$$\Gamma_f = \{(i, j) \in \Gamma | f_i = f_j\} \tag{4}$$

where f_i and f_j denote the frequency of cell i and j, respectively. Cell j is the co-frequency neighbor cell of the master cell i if ordered pair $(i, j) \in \Gamma_f$. They make up a couple of co-frequency neighborhood relations.

C. THE DEFINITION OF PCI PLANNING

The limited PCIs will be inevitably reused in the downlink network since the number of cells is far more than 504. However, the wrong allocation of PCI will significantly increase

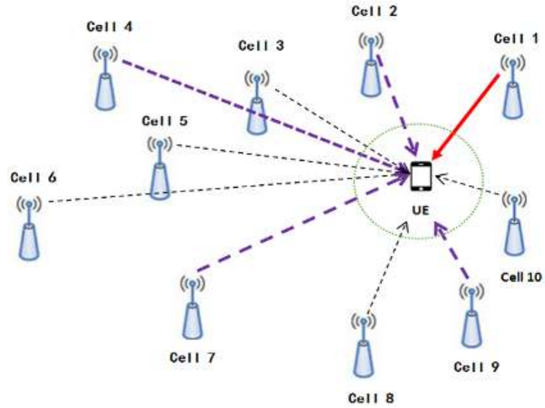


FIGURE 1. The diagram about neighborhood relation.

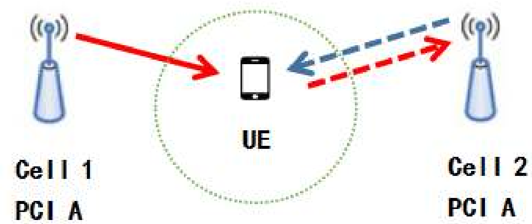


FIGURE 2. PCI collision.

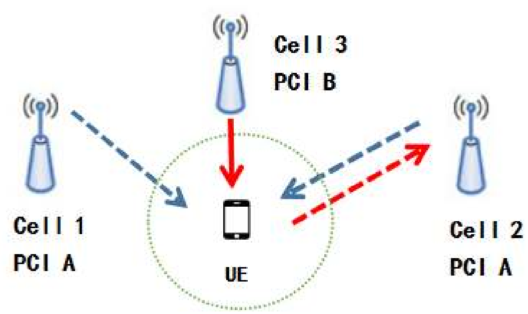


FIGURE 3. PCI confusion.

ICI and further effect the network quality. In order to reduce these ICIs, we should comprehensively consider and analyze the scenario of collision, confusion, and mod 3 interference occurring in neighboring cells at the same frequency.

Collision may occur in some neighboring cells with the same frequency and PCI, which is shown as master cell 1 and neighbor cell 2 with the same PCI A in Fig. 2. In the case of collision, UE may access cell 2 wrongly whereas cell 1 is the target. As it is difficult for the UE to judge these nearly signals effectively, the wrong access command will lead to a service interruption and cause the resource misconfiguration in the downlink network.

Confusion may occur in some master cells with two or more neighbor cells sharing the same frequency and PCI, which is shown as the master cell 3 allocated with PCI

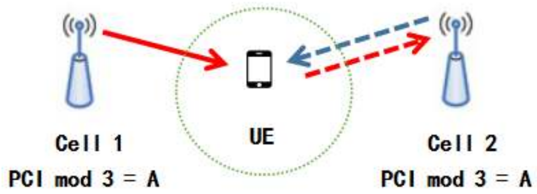


FIGURE 4. PCI mod 3 interference.

B and its neighbor cell 1 and 2 both allocated with PCI A in Fig. 3. In the case of confusion, a UE served by master cell 1 is to be handed over to one of its co-frequency neighbors and PCI will lead to distinguishing problems for the UE. For instance, the UE may be handed over to cell 2 wrongly whereas cell 1 is the target. This scenario may also result in a service interruption and resource misconfiguration for UEs in the downlink network.

Similar to the scenario of collision, mod 3 interference may occur in some neighboring cells with the same frequency and PSS, which is shown in Fig. 4 as master cell 1 and neighbor cell 2 with the same PSS A. In other words, the PCI mod 3 value of cell 1 and cell 2 are equal. In the case of mod 3 interference, there will be a significant decline of the received signal quality (RXQUAL) for the UE due to the superposition of RS from cell 1 and cell 2. This scenario will lead to wrong evaluation of CQI and transmission delay in the downlink network. Besides, since the mod 3 interference is for the Single-Input Single-Output (SISO) antennas in OFDMA system while Multiple-Input Multiple-Output (MIMO) antennas in the 5th generation mobile networks (5G), the planners are required to consider mod 6 interference [28] to eliminate the conflict of reference signals.

Thus, synthetic analysis and evaluation including collision, confusion, and mod 3 interference should be considered to eliminate and reduce ICIs in the downlink network.

III. OPTIMIZATION MODEL

A. MATHEMATICAL DEFINITION

The whole downlink network with n cells and m distributable PCIs can be considered as an unweighted digraph $G = (V, E)$ with n nodes. The node i in digraph G is the cell i in the network. If master cell i and its neighbor cell j share the same frequency, arc $e(i, j)$ exists. Let matrix $A = [a_{ij}]_{n\times n}$ denote adjacency matrix of G. If arc $e(i, j)$ exists, the binary variable $a_{ij} = 1$. A special case is the diagonal element of matrix A, $a_{ii} = 0$ for each node i since there are co-frequency neighborhood relations in different cells.

Then the element of matrix $C = A^T A = [c_{ij}]_{n\times n}$, $c_{ij} = \sum_k a_{ki}a_{kj}$ denotes the same master cells' amount of neighbor cell i and j. Moreover, a binary matrix $B = Sign(C - Diag(C)) = [b_{ij}]_{n\times n}$ is introduced for the subsequent definition of confusion, where $Sign(\bullet)$ denotes an indicator function and $Diag(\bullet)$ is a diagonal matrix. Also, the element of B, b_{ij} reflects whether neighbor cell i and j share the same master cell. For the same reason, the diagonal element of B, b_{ii} should be equal to 0 for each i, thus $C - Diag(C)$ is used here. Thereby matrix B is binary ultimately by means of function $Sign(\bullet)$.

Besides, two new binary matrices $U = [u_{ij}]_{n\times n}$ and $V = [v_{ij}]_{n\times n}$ are introduced, where $u_{ij} = 1$ if node i and j are allocated the same PCI. Also, $v_{ij} = 1$ if these two nodes share the same mod value of PCI. For the interference of PCI mod 3 and PCI mod 6, there is just little difference between these two methods in mathematics. Theoretically, PCI mod q interference is defined here for the sake of simplicity, where q is equal to 3 for the mod 3 interference in OFDMA system and equal to 6 for the mod 6 interference in MIMO system.

B. ESTABLISH BASIC MODEL

According to the above, let $F_{1,i}$, $F_{2,i}$, $F_{3,i}$ denote the number of nodes which have conflict, confusion, and mod q interference with node i respectively. Subsequently, $F_{1,i} = \sum_{j=1}^{n} a_{ij}u_{ij}$ because $a_{ij}u_{ij} = 1$ if $a_{ij} = 1$ and $u_{ij} = 1$, i.e.,, the node i and j are co-frequency with the same PCI. Then the amount of total collision in the downlink network is $F_1 = \sum_{i=1}^{n} F_{1,i} = \sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij}u_{ij}$. Likewise, the amount of PCI confusion, and mod q interference in node i are expressed as $F_{2,i} = \sum_{j=1}^{n} b_{ij}u_{ij}$ and $F_{3,i} = \sum_{j=1}^{n} a_{ij}v_{ij}$ respectively. Moreover, the amount of confusion, and mod q interference value in the downlink network are also expressed as $F_2 = \sum_{i=1}^{n} F_{2,i} = \sum_{i=1}^{n}\sum_{j=1}^{n} b_{ij}u_{ij}$ and $F_3 = \sum_{i=1}^{n} F_{3,i} = \sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij}v_{ij}$, respectively.

A proper PCI plan should comprehensively consider the ICIs, which closely related to the collision, confusion, and mod q interference. Thus, the objective function F for the amount of ICIs is multi-level about these three criteria. Generally, the weighted summation is an effective way to solve the multi-objective optimization. Therefore, the total value $F = \omega_1 F_1 + \omega_2 F_2 + \omega_3 F_3 = \omega_1 \sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij}u_{ij} + \omega_2 \sum_{i=1}^{n}\sum_{j=1}^{n} b_{ij}u_{ij} + \omega_3 \sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij}v_{ij}$, where $\omega_1, \omega_2, \omega_3$ are the weight of collision, confusion, and mod q interference, respectively.

Since the objective function has been established, the constraints should be considered for the optimization model. Actually, we focus more on PCI allocation scheme than on the equality of PCI of certain nodes. For this purpose, we further define two 0-1 matrices $X = [x_{i,k}]_{n\times m}$ and $Y = [y_{i,h}]_{n\times q}$, where $x_{i,k} = 1$ if the PCI k will be allocated to the node i and $y_{i,h} = 1$ if the PSS h is allocated to node i. With these two matrices, the distribution of PCI and PSS can be easy to approach. Since each node should be allocated only one PCI and PSS value, equality constraints are written as $\sum_{k=1}^{m} x_{i,k} = 1$ and $\sum_{h=q}^{l} y_{i,h} = 1$.

Moreover, $y_{i,h}$ can be written as $x_{i,k}$ due to the corresponding relation between PCI and PSS, as reflected in equation (1). In other words, the value of $y_{i,h}$ for all h should be equal to 1 if $x_{i,q\times g+h} = 1$ for some g in each node i.

Furthermore, the implied relation between $x_{i,k}$ and u_{ij} can be established because they both reflect the distribution of PCI, despite from different angles. In other words, the value of $u_{i,j}$ should equal 1 if the allocated PCI in node i is same as j, i. e, $x_{i,k} = x_{j,k}$ for all PCI k. Likewise, the value of $v_{i,j}$ should be equal to 1 if the allocated PSS in node i is same as j, i. e, $y_{i,h} = y_{j,h}$ for all PSS h.

Finally, the basic optimization model (I) based on PCI planning problem can be established below, where the objective is to minimize ICI in the downlink network F. In this model, the constraints should be satisfied as described above. The decision variables in this combinatorial optimization model are $x_{i,k}$.

$$\begin{aligned}
\min_{i,j}\ F &= \omega_1 \sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij}u_{ij} + \omega_2 \sum_{i=1}^{n}\sum_{j=1}^{n} b_{ij}u_{ij} \\
&\quad + \omega_3 \sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij}v_{ij} \\
s.t.\ & \sum_{k=1}^{m} x_{i,k} = 1, \quad \forall i \qquad (1)\\
& \sum_{h=1}^{q} y_{i,h} = 1, \quad \forall i \qquad (2)\\
& u_{i,j} = \begin{cases} 1 & x_{i,k} = x_{j,k}, \forall k \\ 0 & \text{otherwise,} \end{cases} \quad \forall i \neq j \qquad (3)\\
& v_{i,j} = \begin{cases} 1 & y_{i,h} = y_{j,h}, \forall h \\ 0 & \text{otherwise,} \end{cases} \quad \forall i \neq j \qquad (4)\\
& y_{i,h} = \begin{cases} 1 & x_{i,q\times g+h} = 1, \forall h, \exists g \\ 0 & \text{otherwise,} \end{cases} \quad \forall i \qquad (5)\\
& x_{i,k} \in \{0, 1\}, \quad \forall i, k
\end{aligned} \qquad \text{(I)}$$

C. MODEL TRANSLATION

Considering the constraints (3)-(5) in model (I) are three knotty piecewise functions whereas the sum of $x_{i,k}$ and $y_{i,h}$ are both equal to 1. Thus the model (I) can be solved more conveniently with some mathematical deduction.

First, the piecewise function of constraint (3) in model (I) can transform into a quadratic function of $x_{i,k}$ based on the theorem 1 below.

Theorem 1: $u_{i,j} = \sum_{k=1}^{m} x_{i,k}x_{j,k}, \forall i \neq j$ if (1) (3) constraints in model (I) are simultaneously satisfied.

Proof: See Appendix A. ■

Likewise, a quadratic relationship between $v_{i,j}$ and $y_{i,h}$, $y_{j,h}$ can be approached as $v_{i,j} = \sum_{h=1}^{q} y_{i,h}y_{j,h}, \forall i \neq j$ based on the comprehensive consideration of the constraints (2)(4) in model (II). The last piecewise function of constraint (5) in model (I), we can adapt it as a linear function of $x_{i,k}$ based on the theorem 2 below.

Theorem 2: $y_{i,h} = \sum_{g=0}^{m/q-1} x_{i,q\times g+h}, \forall h \in \{1, 2, \cdots, q\}$ if (1) (5) constraints in model (I) are satisfied simultaneously.

Proof: See Appendix B. ■

According to the above theorem 1 and 2, model (II) with quadratic constraints can be transformed from model (I). The model (II) is expressed as

$$\begin{aligned}
\min_{i,j}\ F &= \omega_1 \sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij}u_{ij} + \omega_2 \sum_{i=1}^{n}\sum_{j=1}^{n} b_{ij}u_{ij} \\
&\quad + \omega_3 \sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij}v_{ij} \\
s.t.\ & \sum_{k=1}^{m} x_{i,k} = 1, \quad \forall i \qquad (1)\\
& \sum_{h=1}^{q} y_{i,h} = 1, \quad \forall i \qquad (2)\\
& u_{i,j} = \sum_{k=1}^{m} x_{i,k}x_{j,k}, \quad \forall i \neq j \qquad (3)\\
& v_{i,j} = \sum_{h=1}^{q} y_{i,h}y_{j,h}, \quad \forall i \neq j \qquad (4)\\
& y_{i,h} = \sum_{g=0}^{m/q-1} x_{i,q\times g+h}, \quad \forall i, h \qquad (5)\\
& x_{i,k} \in \{0, 1\}, \quad \forall i, k
\end{aligned} \qquad \text{(II)}$$

where the constraints (1) (5) in model (II) can easily demonstrate the establishment of constraint (2), which can be eliminated by ensuring the constraints (1) (5) to be satisfied.

Moreover, the equality relationships reflected by the constraints (3)-(5) in model (II) can replace variables in the objective function. To maintain the consistency of subscripts, the expression of constraint (5) will be converted using matrix $L = [l_{k,h}]_{m\times q}$. If $q = 3$, the form of L is

$$L = \begin{bmatrix} 1 & 0 & 0 & 1 & 0 & 0 & \cdots & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 & \cdots & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 & \cdots & 0 & 0 & 1 \end{bmatrix}^T$$

and if $q = 6$, the form of L is

$$L = [\underbrace{I_6, I_6, \cdots, I_6}_{m/6 \text{ items}}]^T$$

where I_6 is the 6 rows and 6 columns of unit matrix.

Subsequently, the constraint (5) can be expressed as $y_{i,h} = \sum_{k=1}^{m} x_{i,k} l_{k,h}$. Thus, model (II) can be further simplified and its equivalent conversion model, i.e, model (III) is expressed as

$$\begin{aligned}
\min_{i,j} F =\ & \omega_1 \sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij} \sum_{k=1}^{m} x_{i,k}x_{j,k} \\
& + \omega_2 \sum_{i=1}^{n}\sum_{j=1}^{n} b_{ij} \sum_{k=1}^{m} x_{i,k}x_{j,k} \\
& + \omega_3 \sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij} \sum_{h=1}^{q} (\sum_{k=1}^{m} x_{i,k} l_{k,h}) \cdot (\sum_{k=1}^{m} x_{j,k} l_{k,h}) \\
s.t.\ & \sum_{k=1}^{m} x_{i,k} = 1, \quad \forall i \\
& x_{i,k} \in \{0, 1\}, \quad \forall i, k
\end{aligned} \tag{III}$$

where the objective function is quadratic and the constraint is linear.

Therefore, the model based on PCI planning problem is finally converted into a binary quadratic programming (BQP) model. For the convenience of narrative, the model (III) is equivalently rewritten as matrix form based on matrix theory.

$$\begin{aligned}
\min_{X}\ & F = \omega_1 Tr(X^T AX) + \omega_2 Tr(X^T BX) \\
& \quad + \omega_3 Tr(L^T X^T AXL) \\
s.t.\ & X\mathbf{1} = \mathbf{1} \\
& X \in \{0, 1\}^{n\times m}
\end{aligned} \tag{IV}$$

where Tr denotes the trace of matrix and $\mathbf{1} = [1, 1, \cdots, 1]^T$ is an n dimensional column vector.

D. MODEL ANALYSIS

To further illustrate the difficulty of theoretically solving the model (IV), the convexity and concaveness of the objective function F should be determined. However, F is not always convex based on the deduction of theorem 3 as shown below.

Theorem 3: The convexity of F depends on the positive definiteness of matrix A and B.

Proof: See Appendix C. ■

Since the input matrix A and B are not always positive definite in actual network thereby F cannot hold the properties of convex. Since the general BQP without convexity is NP-hard, PCI planning problem is NP-hard.

Therefore, it is hard to approach the exact optimal solution for the PCI planning problem in polynomial time. Besides, large-scale inputs in practical cases will cost not only huge time complexity but also huge space complexity for existing classical algorithms, e.g., BBM.

IV. ALGORITHM

To ensure calculation speed and reduce storage resources, this section presents a heuristic Greedy algorithm with scalable weights solving the model (IV). The lower complexity in time and space of Greedy algorithm contributes to actual network optimization. Moreover, scalable weights provide multiple options for planning in different scenarios.

A. BASIC IDEA

The process of non-convex $F(X)$ is inspired by the algorithm proposed by [44], while the Greedy algorithm proposed in this study is original. Moreover, there are also essential differences mathematically between the unconstrained BQP problem and BQP problem with constraint. The algorithms in [44] can solve the former well but does not work for the latter because it cannot ensure the solution to be feasible. Other existing algorithms such as BBM will lead to huge time and space complexity. The key point of the proposed algorithm is to use constraints' special property that the sum of each row in matrix X is 1. This ensures the solution is absolutely feasible and the time and space complexity of the Greedy algorithm are both low.

The Greedy algorithm starts from the initial relaxation solution relaxing the binary limits and gradually iterates it with maximum gain in each iteration. To be specific, all elements of the column vector of the matrix X can change to 0 or 1 in each iteration. If an element x_{ik} of X attempts to change, the improvement of F is defined as its gain. As the maximum improvement of all changed elements will fastest decrease the value of objective function, the corresponding element is really adjusted in this iteration. Finally, the termination condition of the algorithm is that all elements x_{ik} are adjusted to 0 or 1. For the convenience of description, the element is defined as changeable if $0 < x_{ik} < 1$ and unchangeable if $x_{ik} = 0$ or $x_{ik} = 1$.

The detail of the Greedy algorithm is presented as follows. First, the initialized relaxation solution matrix $\bar{X}$ is formed as

$$\bar{X} = \begin{bmatrix} \frac{1}{m} & \frac{1}{m} & \cdots & \frac{1}{m} \\ \frac{1}{m} & \frac{1}{m} & \cdots & \frac{1}{m} \\ \vdots & \vdots & \vdots & \vdots \\ \frac{1}{m} & \frac{1}{m} & \cdots & \frac{1}{m} \end{bmatrix}_{n\times m} \tag{5}$$

which satisfies the constraints but fails to meet the binary limits.

For this purpose, the element x_{ik} should change to 0 or 1 and meet $\sum_{k}^{m} x_{ik} = 1, \forall i$. Therefore, other elements $x_{ik'}, k' \neq k$ in the same row i will be changed passively when x_{ik} changes yet other row elements $x_{jk}, j \neq i$ are all unaffected. For this reason, only the one row vector elements X_i will change if changeable element x_{ik} is adjusted in each iteration.

If there are also l changeable elements and x_{ik} is adjusted to 0, then other changeable elements $x_{ik'}, \forall k' \neq k$ in the same row should be passively adjusted to $\frac{1}{l-1}$ to keep $\sum_{k}^{m} x_{ik} = 1$

satisfied. Subsequently, the row vector X_i is adjusted to $z^0_{i,k}$ as

$$z^0_{i,k} = [\underbrace{0,\cdots,0}_{\text{unchangeable elements}},\cdots,\underbrace{\frac{1}{l-1},\cdots,\frac{1}{l-1}}_{l-1 \text{ passivle changed elements } x_{ik'}},\cdots,\underbrace{0}_{x_{ik}},\cdots]$$

On the contrary, if x_{ik} is changed to 1, other changeable elements $x_{ik'}, \forall k' \neq k$ should be changed to 0 immediately for holding $\sum_k^m x_{ik} = 1$. Subsequently, the row vector X_i is adjusted to $z^1_{i,k}$ as

$$z^1_{i,k} = [\underbrace{0,\cdots,0}_{\text{unchangeable elements}},\cdots,\underbrace{0,\cdots,0}_{l-1 \text{ passivle changed elements } x_{ik'}},\cdots,\underbrace{1}_{x_{ik}},\cdots,0]$$

Obviously, the unchangeable elements are all 0 in each iteration as the row vector X_i will loss the chance to change if an element of X_i has been changed to 1 in the previous iteration.

Additionally, the change direction of element x_{ik} is determined by the improvement of objective function $F(X)$. For matrix X, the change of element x_{ik} leads to the adjustment of column vector X_i. Subsequently, when X_i changes to $z^0_{i,k}$, the improvement of $F(X)$ is denoted by $\Delta F(z^0_{i,k})$. When X_i changes to $z^1_{i,k}$, the improvement of $F(X)$ is denoted by $\Delta F(z^1_{i,k})$. Moreover, let $\tilde{X}$ denote the matrix differing from X just in the i-th row elements and the difference of $F(X)$ and $F(\tilde{X})$ is denoted by $\Delta F(z_{i,k})$ as

$$\begin{aligned}\Delta F(z_{i,k}) &= F(\tilde{X}) - F(X)\\ &= \omega_1 Tr(\tilde{X}^T A\tilde{X} - X^T AX)\\ &\quad + \omega_2 Tr(\tilde{X}^T B\tilde{X} - X^T BX)\\ &\quad + \omega_3 Tr(L^T\tilde{X}^T A\tilde{X}L - L^T X^T AXL)\end{aligned}$$

Though the $O(mn^2)$ time complexity of above matrices' multiplication for each change of element z_{ik} seems to be inessential, lots of repeated calculation for all elements in each iteration will cost unbearable time. Thus, the method of directly calculating trace is improved by vector multiplication as expressed below, and the derivation can be obtained by simple knowledge of matrix theory.

$$\begin{aligned}\Delta F(z_{i,k}) = &\ \omega_1(W^1_i(z_{i,k} - X_i)^T + A_{ii}(z^T_{i,k}z_{i,k} - X^T_i X_i))\\ &+ \omega_2(W^2_i(z_{i,k} - X_i)^T + B_{ii}(z^T_{i,k}z_{i,k} - X^T_i X_i))\\ &+ \omega_3(W^3_i(z_{i,k} - X_i)^T + A_{ii}((z_{i,k}L)^T(z_{i,k}L)\\ &- (XL)^T_i(XL)_i))\end{aligned}$$

and

$$\begin{aligned}W^1 &= (A + A^T - 2diag(A))X\\ W^2 &= (B + B^T - 2diag(B))X\\ W^3 &= (A + A^T - 2diag(A))(XL)\end{aligned}$$

where $W^1_i, W^2_i, W^3_i, (XL)_i$ are the i-th row of matrix W^1, W^2, W^3, XL, respectively, and $diag(A), diag(B)$ are the diagonal matrix of A and B, respectively. In each iteration, these matrices W^1, W^2, W^3, XL can be calculated in advance and the calculation of $F(z_{i,k})$ for all changeable elements x_{ik} just needs to call their i-th vector. The remaining multiplication calculation of vectors is just cost $O(n)$ time complexity.

For all changeable elements in each iteration, the smallest $\Delta F(z_{i,k})$ determines the change direction of $\tilde{X}_i$, namely z^0_{ik} or z^1_{ik}. In particular, it will select a $z_{i,k}$ randomly if there are multiple possibilities. There will be a row vector $\tilde{X}_i$ adjusted in each iteration, which has no chance to re-change if x_{ik} intends to be 1. Even when there are only two changeable elements, the last one will change to $\frac{1}{l-1} = \frac{1}{2-1} = 1$ if the $n-1$th changes to 0. Finally, all elements x_{ik} become 0 or 1. At this time, the Greedy algorithm terminates. The final X is absolutely feasible and $x_{ik} = 1$ decides the finally assignment scheme, namely, the PCI k is allocated to cell i.

The pseudo code of the Greedy algorithm is shown as follows.

Algorithm 1 Greedy for PCI Planning Problem

Input: Cell amount n and PCI amount m, matrix A, weight $\omega_1, \omega_2, \omega_3$
Output: z

```
Φ = {(1, 1), (1, 2), ··· , (m, n)}
B = Sign(A^T A − Diag(A^T A))
while Φ ≠ ∅ do
    Calculate (i*, k*) with ΔF(z^0_{i*k*}) = min_{(i,k)} ΔF(z^0_{ik}) for
all x_ik and (i, k) ∈ Φ.
    Calculate (i**, k**) with ΔF(z^1_{i**k**}) = min_i ΔF(z^1_{ik})
for all x_ik and (i, k) ∈ Φ.
    if ΔF(z^0_{i*k*}) ≤ ΔF(z^1_{i**k**}) then
        X_{i*} := z^0_{i*k*}
        Φ := Φ\(i*, k*)
    else
        X_{i**} := z^1_{i*k*}
        Φ := Φ\(i**, k), ∀k
return X
```

B. COMPLEXITY ANALYSIS

The space complexity analysis of Greedy is rather simple and primarily cost in the storage of matrix A, B, X, L, which cost at most $O(mn) + O(mn) + O(n^2) + O(n \times 3) = O(nm + n^2)$ space complexity. However, the time complexity mainly depends on the calculation of each iteration and the amount of iterations.

In each iteration, the calculation consists of two parts, calculating the prepared matrix XL, W^1, W^2, W^3 and $\Delta F(z_{i,k})$ for each changeable element x_{ik}.

On the one hand, the calculation of prepared matrix XL will cost $O(mn)$ time complexity and W^1, W^2 cost $O(mn^2)$ time complexity. Subsequently, the calculation of W^3 uses matrix of $A + A^T - 2diag(A)$ and XL just cost $O(n^3)$. Accordingly, all prepared matrix calculation will cost at most $O(mn + mn^2 + n^2) = O(mn^2)$ time complexity.

On the other hand, the vector multiplication of W_i^1, W_i^2, W_i^3 and $(z_{i,k} - X_i)^T$ and the calculation of $z_{i,k}^T z_{i,k}, X_i^T X_i, z_{i,k}L, (XL)_i^T (XL)_i$ cost only at most $O(n)$ time complexity for each changeable element x_{ik}. However, the amount of all x_{ik} is not more than mn, so the calculation of $\Delta F(z_{i,k})$ for all changeable elements will cost not more than $O(mn^2)$ time complexity.

Hence, the total time complexity in each iteration is $O(mn^2 + mn^2) = O(mn^2)$.

Moreover, the most iteration amount is $O(mn)$ in the worst case that only one changeable element in each iteration is updated. Overall, the total time complexity of Greedy algorithm is at most $O(mn^2 \times mn) = O(m^2n^3)$.

V. PERFORMANCE EVALUATION

To evaluate the effectiveness and performance of Greedy algorithm described as above, multiple groups of comparative experiments will be simulated in this section. First, system simulation environments are illustrated, including the source of raw data and some important parameters in the actual downlink network. The raw data is taken from the actual SISO system network and the simulation calculation is implemented by MATLAB. Subsequently, the comparative results of these experiments are presented and the comparison algorithms include the classical graph color algorithm and the baseline PCI assign scheme implemented in the live network. Next, multiple experiments with different weights are simulated to investigate the effects of weight $\omega_1, \omega_2, \omega_3$ on the Greedy algorithm performance. Finally, the detail PCI assign scheme and PSS result based on the Greedy algorithm with weight $\omega_1 = 1, \omega_2 = 1, \omega_3 = 1$ will be presented as an example.

A. SIMULATION ENVIRONMENT

These experiments are all tested with real data provided by China Mobile Group Beijing Company Limited (CMBJ). The antennas are all deployed in the SISO system. Thus, only mod 3 interference is considered in the simulation. Moreover, the raw data is extracted from the live downlink network, which includes 1131 cells at the same frequency and covering a dense urban geographical area of 53.0859 km^2. The cell distribution in the live network is shown in Fig. 5. Moreover, the raw data known as measurement report (MR) data is based on the information regularly measured by UEs, which reports the status and performance of network. MR data is used to generate the elements of matrix A. Besides, after the preprocessing of MR data, the information of master cell and its co-frequency neighbor cells can be obtained. The element of matrix A, $a_{ij} = 1$ if master cell i has the co-frequency neighbor cell j. Fig. 6 shows a sample of MR data segment.

The geographical distribution and performance configurations of cells constitutes the nodes of topological graph G. The edges of G are derived from the average statistics of MR data and two cells are considered as adjacent only if

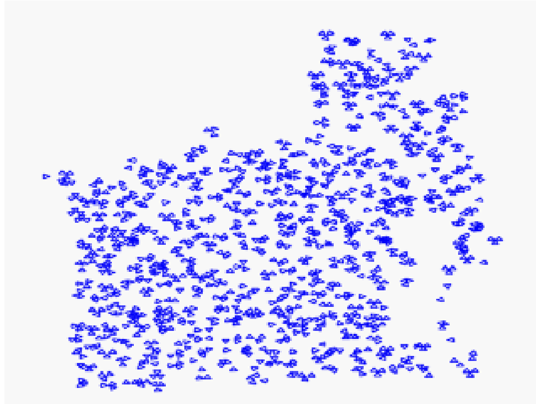
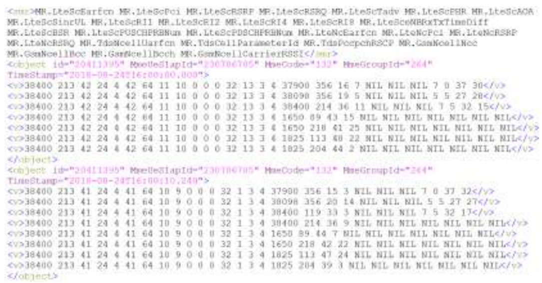


FIGURE 5. The geographical distribution of cells in the live network.




```
<mr>MR.LteScEarfcn MR.LteScPci MR.LteScRSRP MR.LteScRSRQ MR.LteScTadv MR.LteScPHR MR.LteScAOA
MR.LteScSinrUL MR.LteScRI1 MR.LteScRI2 MR.LteScRI4 MR.LteScRI8 MR.LteSceNBRxTxTimeDiff
MR.LteScBSR MR.LteScPUSCHPRBNum MR.LteScPDSCHPRBNum MR.LteNcEarfcn MR.LteNcPci MR.LteNcRSRP
MR.LteNcRSRQ MR.TdsNcellUarfcn MR.TdsCellParameterId MR.TdsPccpchRSCP MR.GsmNcellNcc
MR.GsmNcellBcc MR.GsmNcellBcch MR.GsmNcellCarrierRSSI</mr>
<object id="20411395" MmeUeS1apId="230706705" MmeCode="132" MmeGroupId="264"
TimeStamp="2018-08-24T16:00:00.000">
<v>38400 213 42 24 4 42 64 11 10 0 0 0 32 13 3 4 37900 356 16 7 NIL NIL NIL 7 0 37 30</v>
<v>38400 213 42 24 4 42 64 11 10 0 0 0 32 13 3 4 38098 356 19 5 NIL NIL NIL 5 5 27 28</v>
<v>38400 213 42 24 4 42 64 11 10 0 0 0 32 13 3 4 38400 214 36 11 NIL NIL NIL 7 5 32 15</v>
<v>38400 213 42 24 4 42 64 11 10 0 0 0 32 13 3 4 1650 89 43 15 NIL NIL NIL NIL NIL NIL NIL</v>
<v>38400 213 42 24 4 42 64 11 10 0 0 0 32 13 3 4 1650 218 41 25 NIL NIL NIL NIL NIL NIL NIL</v>
<v>38400 213 42 24 4 42 64 11 10 0 0 0 32 13 3 4 1825 113 48 22 NIL NIL NIL NIL NIL NIL NIL</v>
<v>38400 213 42 24 4 42 64 11 10 0 0 0 32 13 3 4 1825 204 44 2 NIL NIL NIL NIL NIL NIL NIL</v>
</object>
<object id="20411395" MmeUeS1apId="230706705" MmeCode="132" MmeGroupId="264"
TimeStamp="2018-08-24T16:00:10.240">
<v>38400 213 41 24 4 41 64 10 9 0 0 0 32 1 3 4 37900 356 15 3 NIL NIL NIL 7 0 37 32</v>
<v>38400 213 41 24 4 41 64 10 9 0 0 0 32 1 3 4 38098 356 20 14 NIL NIL NIL 5 5 27 27</v>
<v>38400 213 41 24 4 41 64 10 9 0 0 0 32 1 3 4 38400 119 33 3 NIL NIL NIL 7 5 32 17</v>
<v>38400 213 41 24 4 41 64 10 9 0 0 0 32 1 3 4 38400 214 36 9 NIL NIL NIL NIL NIL NIL NIL</v>
<v>38400 213 41 24 4 41 64 10 9 0 0 0 32 1 3 4 1650 89 44 7 NIL NIL NIL NIL NIL NIL NIL</v>
<v>38400 213 41 24 4 41 64 10 9 0 0 0 32 1 3 4 1650 218 42 22 NIL NIL NIL NIL NIL NIL NIL</v>
<v>38400 213 41 24 4 41 64 10 9 0 0 0 32 1 3 4 1825 113 47 24 NIL NIL NIL NIL NIL NIL NIL</v>
<v>38400 213 41 24 4 41 64 10 9 0 0 0 32 1 3 4 1825 204 39 3 NIL NIL NIL NIL NIL NIL NIL</v>
</object>
```


FIGURE 6. The sample of MR data.

handover occurs between the master cells and neighbors at the same frequency and $\delta_s \leq 1.5\ km$, $\delta_d \leq 6\ dB$. Based on the statistics and preprocessing of raw MR data, the network topological graph G is constructed and the input adjacent matrix A is approached. Moreover, the hardware of simulation environment is a Windows-based computer with 2.4 GHz of CPU and 64 GB of RAM.

B. ALGORITHMS SIMULATION

First, seven different PCI planning schemes are all tested in the aforementioned simulation environment with PCI amount $m = 504$. The first five of them assign PCI to the cells from multiple perspectives and are approached by the Greedy algorithm configuring different weights. The last two approaches are classical graph color algorithm and the baseline plan implemented in the live network respectively. These PCI planning schemes are detailed below.

1) 1-1-1 Greedy. It is designed to reduce the collision, confusion, and mod 3 interference in a comprehensive and balanced way. The plan approach is configured with $\omega_1 = 1, \omega_2 = 1, \omega_3 = 1$.
2) 1-1-0 Greedy. The scheme aims to avoid the collision and confusion whereas the PCI mod 3 interference is ignored. This scheme is based on the traditional perspective to assign PCI and the weight configuration is $\omega_1 = 1, \omega_2 = 1, \omega_3 = 0$.
3) 0-0-1 Greedy. It is designed to desperately reduce the PCI mod 3 interference whereas the influence of

TABLE 1. Performance comparison between different PCI assign schemes.

Scheme	Collision	Confusion	Mod 3 interference
Baseline	36	0	12555
Graph Color	0	0	11998 (4.436%)
1-1-0 Greedy	0	0	11953 (4.795%)
0-0-1 Greedy	8942	151008	9220 (26.563%)
1-0-1 Greedy	0	21886	9223 (26.539%)
0-1-1 Greedy	5	0	9245 (26.364%)
1-1-1 Greedy	0	0	9264 (26.213%)

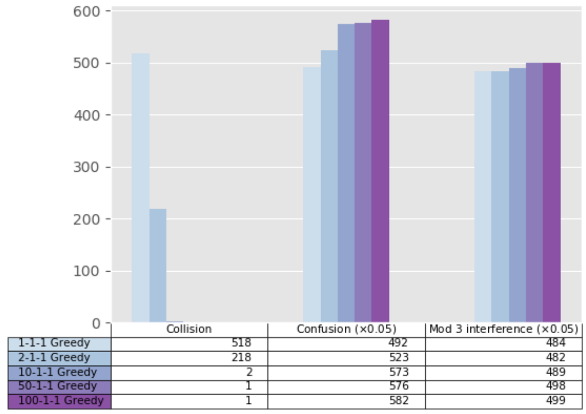


FIGURE 7. The comparison of different ω_1 with $\omega_2 = 1, \omega_3 = 1$.

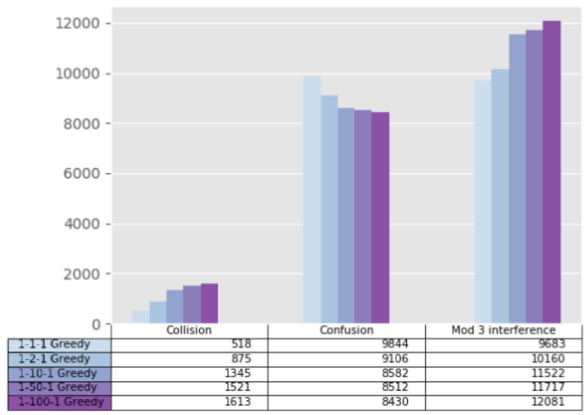


FIGURE 8. The comparison of different ω_2 with $\omega_1 = 1, \omega_3 = 1$.

collision and confusion are ignored. This plan is configured with $\omega_1 = 0, \omega_2 = 0, \omega_3 = 1$.

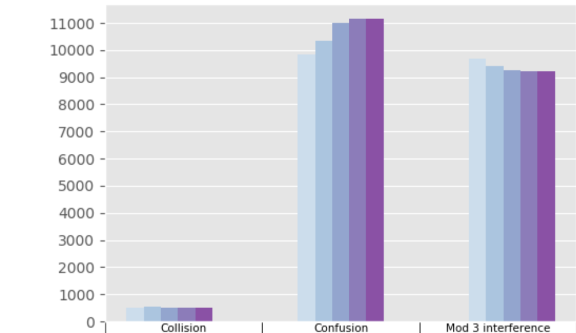
4) 1-0-1 Greedy. It is designed to reduce the collision and PCI mod 3 interference whereas the confusion is ignored. For this purpose, it is configured with $\omega_1 = 1, \omega_2 = 0, \omega_3 = 1$.
5) 0-1-1 Greedy. It is designed to minimize the number of confusion and PCI mod 3 interference whereas the collision is ignored. For this purpose, it is configured with $\omega_1 = 0, \omega_2 = 1, \omega_3 = 1$.
6) Graph Color. The approach is a classical graph theory algorithm for planning PCI and the detail is referred to [48].
7) Baseline. It is the PCI assignment scheme in the live network, which is continuously adjusted and optimized by experienced engineers.

Table 1 lists the performance compared results of these seven PCI plans, including the number of collision, confusion, and mod 3 interference. To show the comparison results more clearly, the improvement ratio is calculated by the comparison of baseline plan and it is presented in parentheses as a supplementary explanation. Obviously, Graph Color algorithm completely eliminates collision and confusion despite the slight improvement of mod 3 interference. The performance of 1-1-0 Greedy is very similar to that of Graph Color and it reduces 0.359% mod 3 interference more than Graph Color. However, if ω_3 is not zero, the mod 3 interference phenomena will be significantly improved. The last four PCI assign scheme all reduce more than 26% mod 3 interference compared with baseline plan. 0-0-1 Greedy simply consider the effects of mod 3 interference leads to serious deterioration of collision and confusion. Moreover, 1-0-1 Greedy ignores the effects of PCI confusion and finally leads to the rapid increase of confusion. In the meantime, collision under 0-1-1 seems to be not too bad as the number of assignable PCI $m = 504$ is relatively enough to the number of cells $n = 1131$. Finally, 1-1-1 Greedy completely eliminates conflicts and confusion meanwhile the amount of interference reduction is also acceptable. Ignoring any one of collision, confusion, and mod 3 interference will all lead to network performance deterioration thereby PCI planning should comprehensively consider the impact.

Furthermore, we focus on the effects of single weight changes on network performance under the premise of comprehensive consideration. Since the number of 504 PCIs is sufficient relative to the amount of 1131 cells, the number of PCI is reduced to 30 in experiments of below groups for better observation. In this case, the effects of weight on the change of conflict and confusion are more significant..

Fig. 7 shows the impact of increasing conflict weight ω_1 on network performance under comprehensive considerations of these three conditions. Without loss of generality, the value of ω_1 is set to 1, 2, 10, 50 and even 100. Since the conflict value is relatively too small, the values of confusion and mod 3 interference are both shrunk by 20 times for the convenience of display. Proportional scaling does not affect the order of values, thereby the trend of each group of histograms will violate the facts. Additionally, to distinguish the slight difference between the height of histograms in Fig. 7, the detail value is also added to the figure in the form of a table. Obviously, the value of conflict will decrease with the increase in ω_1 until it no longer changes. In the meantime, the value of confusion and mod 3 interference are increasing because they become more and more insignificant with the increase in ω_1.

Fig. 8 and 9 show the impact of confusion weight ω_2 and mod 3 interference weight ω_3 on network performance, respectively. They are basically the same as shown in Fig. 7. However, the small difference is that they both show the original value of confusion and mod 3 interference without



	Collision	Confusion	Mod 3 interference
1-1-1 Greedy	518	9844	9683
1-1-2 Greedy	529	10336	9429
1-1-10 Greedy	510	10992	9247
1-1-50 Greedy	508	11155	9225
1-1-100 Greedy	508	11155	9225

FIGURE 9. The comparison of different ω_3 with $\omega_1 = 1, \omega_2 = 1$.

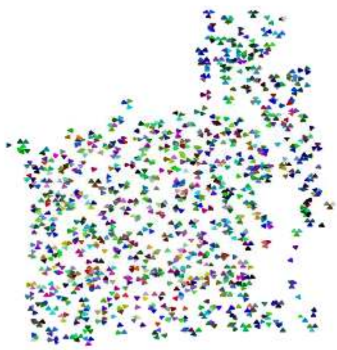


FIGURE 10. The schematic diagram of PCI assignment scheme.

shrinking as the emphasis is no longer placed on conflict. Likewise, the table under the histogram provides detailed numerical information. Fig. 8 reflects a phenomenon similar to Fig. 7, Namely, as the confusion's weight ω_2 increases, the phenomenon of PCI confusion becomes more and more improved whereas the conflict and mod 3 interference are deteriorated continuously. However, Fig. 9 shows that reducing the mod 3 interference also slightly improves conflict. This interesting phenomenon is correlated with their definitions and logical inference. Looking back at their definition above, if the PCI between two cells conflicts, then the mod 3 interference between them will be inevitable. For instance, the original PCIs of neighboring cells A and B are both 3, then the conflict and mod 3 interference between them both exist. Supposed the PCI of A is unchanged meanwhile improve the PCI of B, then its value is completely impossible to equal 3 if the mod 3 interference between them is eliminated.

C. ALLOCATION SCHEME

Finally, to present the actual allocation of PCI and PSS, all tested cells on the region map are dyed with different colors to distinguish. In accordance with the previous experimental results, the 1-1-1 Greedy algorithm is a comprehensive and balanced strategy. Accordingly, the allocated PCI and PSS are selected by the 1-1-1 Greedy algorithm and the PCI number

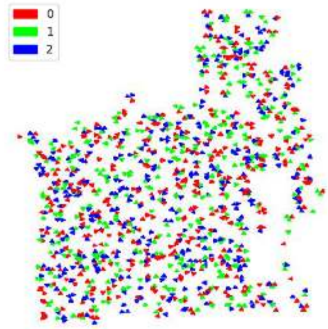


FIGURE 11. The schematic diagram of PSS assignment scheme.

in the simulation environment is 504. Thus, Fig. 10 provides 504 different colors, each of which corresponds to one PCI value. However, the number of PSS are always 3 then Fig. 11 only provides three different colors to label 0, 1, 2.

VI. CONCLUSION

In this study, a new approach to establish neighborhood relations was proposed based on the actual mobile networks, which is more reliable comparing with the conventional approach based on simulation results or idealistic assumptions. In addition, we established combinatorial optimization model to describe collision, confusion, and mod q interference comprehensively and quantitatively. Through rigorous formula derivation and mathematical proof, PCI planning problem was mapped as a BQP model, which is known to be NP-hard. Unfortunately, the difficulty is further increased by the non-convexity of objective functions. Based on the BQP optimization model, a Greedy algorithm was developed to automatically configure PCIs to each cell in the whole network. The time and space complexity of this algorithm were analyzed in detail. To evaluate the optimization performance of the proposed algorithm and verify the impact of weights, numerical simulations based on MR data were performed compared with the scheme implemented in the current network and the classical graph coloring algorithm. The experimental results demonstrated that the Greedy algorithm has a great advantage in reducing the collision, confusion, and mod 3 interference according to different weights in various scenarios. Moreover, the performance improvements of the Greedy algorithm configured different weights allow planners to make comprehensive and reasonable trade-offs in various scenarios. Finally, the allocation scheme of PCI and PSS provided 1-1-1 Greedy algorithm are both illustrated intuitively as examples. In fact, some results of our algorithms have been widely applied in CMBJ. In the meantime, the proposed algorithm can be applied in the 5G network when the mod 6 interference is considered for the MIMO system.

APPENDIX A
PROOF OF THE THEOREM 1

According to the constraint (1) in model (I), there is a k' satisfying $x_{i,k'} = 1$ whereas $x_{i,k''} = 0$ for all $k'' \neq k'$ and $k'' \in \{1, 2, 3, \cdots, m\}$.

Then for $j \neq i$, if $x_{j,k} = x_{i,k}$ are always satisfied for all k then $x_{j,k'} = x_{i,k'} = 1$ and $x_{j,k''} = x_{i,k''} = 0$. Thereby $u_{i,j} = 1$ for all node i and j. In addition,

$$\sum_{k=1}^{m} x_{i,k}x_{j,k} = \underbrace{x_{i,k'}x_{j,k'}}_{=1} + \underbrace{\sum_{k''\neq k'} x_{i,k''}x_{j,k''}}_{=0} = 1$$

Therefore, $u_{i,j} = \sum_{k=1}^{m} x_{i,k}x_{j,k}$ in this case.

Otherwise, if some $\tilde{k} \in$ satisfies $x_{j,\tilde{k}} \neq x_{i,\tilde{k}}$, then there will be two possible scenarios for $\tilde{k}$ will occur: 1) $\tilde{k} = k'$; 2) $\tilde{k} \neq k'$. Certainly, $u_{i,j}$ are both equal 0 in these two scenarios.

If $\tilde{k} = k'$, binary variable $x_{j,\tilde{k}}$ should equal 0 due to $x_{j,\tilde{k}} \neq x_{i,\tilde{k}}$ and $x_{i,\tilde{k}} = x_{j,k'} = 1$. Then,

$$\begin{aligned}\sum_{k=1}^{m} x_{i,k}x_{j,k} &= x_{i,\tilde{k}}x_{j,\tilde{k}} + \sum_{k\neq\tilde{k}} x_{i,k}x_{j,k} \\ &= \underbrace{x_{i,\tilde{k}}\underbrace{x_{j,\tilde{k}}}_{=0}}_{=0} + \underbrace{\sum_{k''}\underbrace{x_{i,k''}}_{=0}x_{j,k''}}_{=0} = 0 = u_{i,j}\end{aligned}$$

If $\tilde{k} \neq k'$, binary variable $x_{i,\tilde{k}} = 1$ and $x_{i,\tilde{k}} = 0$ will be satisfied for some k'' due to $x_{j,\tilde{k}} \neq x_{i,\tilde{k}}$ and $x_{i,\tilde{k}} \neq x_{i,k'} = 1$. In addition, $x_{j,\check{k}} = 0$ for all $\check{k} \neq \tilde{k}$ due to $\sum_{k} x_{j,k} = x_{j,\tilde{k}} + \sum_{\check{k}\neq\tilde{k}} x_{j,\check{k}} = 1$. Then,

$$\sum_{k} x_{i,k}x_{j,k} = \underbrace{\underbrace{x_{i,\tilde{k}}}_{=0}x_{j,\tilde{k}}}_{=0} + \underbrace{\sum_{\check{k}\neq\tilde{k}} x_{i,\check{k}}\underbrace{x_{j,\check{k}}}_{=0}}_{=0} = 0 = u_{i,j}$$

In the end, $u_{ij} = \sum_{k=1}^{m} x_{i,k}x_{j,k}$ are always satisfied if constraints (1) and (3) in model (I) are satisfied.

APPENDIX B
PROOF OF THE THEOREM 2

On the one hand, if there is a g' satisfying $x_{i,q\times g'+h} = 1$ for all h, then $y_{i,h} = 1$ because node i is allocated the PSS h in this scenario. According to the constraint (1) in model (I), $x_{i,q\times g''+h} = 0$ should be satisfied for all $g'' \neq g'$. Therefore

$$\begin{aligned}\sum_{g=0}^{m/q-1} x_{i,q\times g+h} &= \underbrace{x_{i,q\times g'+h}}_{=1} + \underbrace{\sum_{g''\neq g'} x_{i,q\times g''+h}}_{=0} \\ &= 1 = y_{i,h}\end{aligned}$$

On the other hand, if there exists a h satisfying $x_{i,q\times g+h} = 0$ for all $g \in \{0, 1, 2, \cdots, m/q-1\}$, then $\sum_{g=0}^{m/q-1} x_{i,q\times g+h} = 0 = y_{i,h}$.

APPENDIX C
PROOF OF THE THEOREM 3

Let $g(t) = F(X) = \omega_1 Tr((Z+tV)^T A(Z+tV)) + \omega_2 Tr((Z+tV)^T B(Z+tV)) + \omega_3 Tr(L^T(Z+tV)^T A(Z+tV)L)$, where $Z, V \in R^{nm}$. Then the convexity of $F(X)$ is consistent with $g(t)$ according to the conclusion in [49]. Therefore, the gradient $\nabla g(t)$ and *Hessian* matrix $\nabla^2 g(t)$ are presented as follows.

$$\begin{aligned}\nabla g(t) = {} & \omega_1 Tr(Z^T AV + V^T AZ + 2tV^T AV) \\ & + \omega_2 Tr(Z^T BV + V^T BZ + 2tV^T BV) \\ & + \omega_3 Tr(L^T Z^T AVL + L^T V^T AZL + 2tL^T V^T AVL)\end{aligned}$$

and

$$\begin{aligned}\nabla^2 g(t) = {} & 2\omega_1 Tr(V^T AV) + 2\omega_2 Tr(V^T BV) \\ & + 2\omega_3 Tr(L^T V^T AVL)\end{aligned}$$

where the positive definiteness depend on matrix A and B. Due to $\omega_1 \geq 0$, $\omega_2 \geq 0$, $\omega_3 \geq 0$, then

$$\nabla^2 g(t) \begin{cases} \geq 0, & A \geq 0 \text{ and } B \geq 0 \\ \leq 0, & A \leq 0 \text{ and } B \leq 0 \\ \textit{indefinite}, & \text{Otherwise} \end{cases}$$

Therefore, the convexity of $g(t)$ and $F(X)$ depend on the positive definiteness of A and B.

ACKNOWLEDGMENT
The MR data in actual network is supported by CMBJ.

REFERENCES

[1] J. Byun, B. W. Kim, Y. K. Chang, and J. W. Byun, "4G LTE network access system and pricing model for IoT MVNOs: spreading smart tourism," *Multimedia Tools Appl.*, vol. 76, no. 19, pp. 19665–19688, 2016.
[2] H. Du, Q. Zheng, W. Zhang, and Y. Huang, "LTE-EMU: A high fidelity LTE cellar network testbed for mobile video streaming," *Mobile Netw. Appl.*, vol. 22, no. 3, pp. 454–463, 2017.
[3] A. Damnjanovic *et al.*, "A survey on 3GPP heterogeneous networks," *IEEE Wireless Commun.*, vol. 18, no. 3, pp. 10–21, Jun. 2011.
[4] A. Toskala, H. Holma, K. Pajukoski, and E. Tiirola, "Utran long term evolution in 3GPP," in *Proc. IEEE Int. Symp. Pers., Indoor Mobile Radio Commun.*, Sep. 2006, pp. 1–5.
[5] E. Park and A. P. del Pobil, "Modeling the user acceptance of long-term evolution (LTE) services," *Ann. Telecommun.*, vol. 68, nos. 5–6, pp. 307–315, 2013.
[6] Z. Shen, A. Papasakellariou, J. Montojo, D. Gerstenberger, and F. Xu, "Overview of 3GPP LTE-advanced carrier aggregation for 4G wireless communications," *IEEE Commun. Mag.*, vol. 50, no. 2, pp. 122–130, Feb. 2012.
[7] S. Jimaa, K. K. Chai, Y. Chen, and Y. Alfadhl, "LTE-A an overview and future research areas," in *Proc. IEEE Int. Conf. Wireless Mobile Comput., Netw. Commun.*, Oct. 2011, pp. 395–399.
[8] T. Jiang and Y. Wu, "An overview: Peak-to-average power ratio reduction techniques for OFDM signals," *IEEE Trans. Broadcast.*, vol. 54, no. 2, pp. 257–268, Jun. 2008.
[9] F. B. Frederiksen and R. Prasad, "An overview of OFDM and related techniques towards development of future wireless multimedia communications," in *Proc. Radio Wireless Conf. (Rawcon)*, 2002, pp. 19–22.
[10] E. Pateromichelakis, M. Shariat, A. U. Quddus, and R. Tafazolli, "On the evolution of multi-cell scheduling in 3GPP LTE/LTE-A," *IEEE Commun. Surveys Tuts.*, vol. 15, no. 2, pp. 701–717, 2nd Quart., 2013.
[11] D. Lopez-Perez, I. Guvenc, G. de la Roche, M. Kountouris, T. Q. S. Quek, and J. Zhang, "Enhanced intercell interference coordination challenges in heterogeneous networks," *IEEE Wireless Commun.*, vol. 18, no. 3, pp. 22–30, Jun. 2011.

[12] H. Yang, A. Huang, R. Gao, T. Chang, and L. Xie, "Interference self-coordination: A proposal to enhance reliability of system-level information in OFDM-based mobile networks via PCI planning," *IEEE Trans. Wireless Commun.*, vol. 13, no. 4, pp. 1874–1887, Apr. 2014.
[13] P. Szilágyi, T. Bandh, and H. Sanneck, *Physical Cell ID Allocation in Multi-Layer, Multi-Vendor LTE Networks*. Berlin, Germany: Springer, 2012.
[14] F. Ahmed, O. Tirkkonen, M. Peltomäki, J.-M. Koljonen, C.-H. Yu, and M. Alava, "Distributed graph coloring for self-organization in LTE networks," *J. Elect. Comput. Eng.*, vol. 2010, p. 5, 2010.
[15] M. Amirijoo, P. Frenger, F. Gunnarsson, H. Kallin, J. Moe, and K. Zetterberg, "Neighbor cell relation list and physical cell identity self-organization in LTE," in *Proc. IEEE Int. Conf. Commun. Workshops (ICC)*, May 2008, pp. 37–41.
[16] M. M. Abdulkareem, S. A. Yaseen, and L. M. Abdullah, "Matrix based graph coloring algorithm for LTE-PCI assignment and reassignment reduction," in *Proc. IEEE 8th Control Syst. Graduate Res. Colloq. (ICSGRC)*, Malaysia, Aug. 2017, pp. 41–46.
[17] L. Chrost and K. Grochla, *Conservative Graph Coloring: A Robust Method for Automatic PCI Assignment in LTE*. Berlin, Germany: Springer, 2013.
[18] Y. Liu, W. Li, H. Zhang, and W. Lu, "Graph based automatic centralized PCI assignment in LTE," in *Proc. Comput. Commun.*, 2010, pp. 919–921.
[19] Y. Wei, W.-B. Wang, M.-G. Peng, and S.-J. Min, "Graph theory based physical cell identity self-configuration for LTE-A network," *J. China Univ. Posts Telecommun.*, vol. 20, no. 1, pp. 101–107, 2013.
[20] L. M. Abdullah, M. D. Baba, S. G. A. Ali, A. O. Lim, and Y. Tan, "New graph colouring algorithm for resource allocation in large-scale wireless networks," in *Proc. Control Syst. Graduate Res. Colloq.*, 2014, pp. 233–238.
[21] L. I. Panxing and J. Wang, "PCI planning method based on genetic algorithm in LTE network," *Telecommun. Sci.*, vol. 32, no. 3, p. 2016082, 2016.
[22] T. Wu, L. Rui, A. Xiong, and S. Y. Guo, "An automation PCI allocation method for eNodeB and home eNodeB cell," in *Proc. Int. Conf. Wireless Commun. Netw. Mobile Comput.*, 2010, pp. 1–4.
[23] A. H. Zahran, "Extended synchronization signals for eliminating PCI confusion in heterogeneous LTE," in *Proc. Wireless Commun. Netw. Conf.*, 2012, pp. 2588–2592.
[24] J. Lim and D. Hong, "Management of neighbor cell lists and physical cell identifiers in self-organizing heterogeneous networks," *J. Commun. Netw.*, vol. 13, no. 4, pp. 367–376, Aug. 2011.
[25] N. Li, C. Huang, and M. Zhuang, "Performance optimization and simulation verification of LTE network planning based on micro coverage," in *Proc. IEEE Int. Conf. Anti-Counterfeiting, Secur., Identificat.*, Sep. 2017, pp. 126–130.
[26] J. Zyren and W. McCoy, "Overview of the 3GPP long term evolution physical layer," Freescale Semicond., White Paper, 2007, vol. 22.
[27] R. Acedo-Hernández, M. Toril, S. Luna-Ramírez, I. de la Bandera, and N. Faour, "Analysis of the impact of PCI planning on downlink throughput performance in LTE," *Comput. Netw.*, vol. 76, pp. 42–54, Jan. 2015.
[28] R. Acedo-Hernández, M. Toril, S. Luna-Ramírez, and C. Úbeda, "A PCI planning algorithm for jointly reducing reference signal collisions in LTE uplink and downlink," *Comput. Netw.*, vol. 119, pp. 112–123, Jun. 2017.
[29] M. R. Garey and D. S. Johnson, *Computers and Intractability: A Guide to the Theory of NP-Completeness*. San Francisco, CA, USA: Freeman, 1983.
[30] E. Boros and P. L. Hammer, "The max-cut problem and quadratic 0–1 optimization; polyhedral aspects, relaxations and bounds," *Ann. Oper. Res.*, vol. 33, no. 3, pp. 151–180, 1991.
[31] G. A. Kochenberger, F. Glover, B. Alidaee, and C. Rego, "An unconstrained quadratic binary programming approach to the vertex coloring problem," *Ann. Oper. Res.*, vol. 139, no. 1, pp. 229–241, 2005.
[32] D. J. Laughhunn, "Quadratic binary programming with application to capital-budgeting problems," *Oper. Res.*, vol. 18, no. 3, pp. 454–461, 1970.
[33] J. C. T. Mao and B. A. Wallingford, "An extension of Lawler and bell's method of discrete optimization with examples from capital budgeting," *Manage. Sci.*, vol. 15, no. 2, pp. B51–B60, 2011.
[34] R. D. Mcbride and J. S. Yormark, "An implicit enumeration algorithm for quadratic integer programming," *Manage. Sci.*, vol. 26, no. 3, pp. 282–296, 1980.
[35] J. Shi and J. Malik, "Normalized cuts and image segmentation," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 22, no. 8, pp. 888–905, Aug. 2000.
[36] S. X. Yu and J. Shi, "Segmentation given partial grouping constraints," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 26, no. 2, pp. 173–183, Feb. 2004.
[37] M. Heiler, J. Keuchel, and C. Schnörr, "Semidefinite clustering for image segmentation with *a-priori* knowledge," in *Proc. DAGM Conf. Pattern Recognit.*, 2005, pp. 309–317.
[38] S. Guattery and G. L. Miller, *On the Quality of Spectral Separators*. Philadelphia, PA, USA: SIAM, 1998.
[39] R. Kannan, S. Vempala, and A. Veta, "On clusterings: Good, bad and spectral," in *Proc. Symp. Found. Comput. Sci.*, 2000, p. 367.
[40] K. J. Lang, "Fixing two weaknesses of the spectral method," in *Proc. 18th Int. Conf. Neural Inf. Process. Syst. (NIPS)*. Cambridge, MA, USA: MIT Press, 2005, pp. 715–722. [Online]. Available: http://dl.acm.org/citation.cfm?id=2976248.2976338
[41] J. Park and S. Boyd, "A semidefinite programming method for integer convex quadratic minimization," *Optim. Lett.*, vol. 63, no. 2, pp. 1–20, 2015.
[42] M. M. Amini, "A scatter search approach to unconstrained quadratic binary programs," *New Ideas Optim.*, pp. 317–329, 1999.
[43] F. Glover, G. A. Kochenberger, and B. Alidaee, "Adaptive memory tabu search for binary quadratic programs," *Manage. Sci.*, vol. 44, no. 3, pp. 336–345, 1998.
[44] P. Merz and B. Freisleben, "Greedy and local search heuristics for unconstrained binary quadratic programming," *J. Heuristics*, vol. 8, no. 2, pp. 197–213, 2002.
[45] M. W. Carter, "The indefinite zero-one quadratic problem," *Discrete Appl. Math.*, vol. 7, no. 1, pp. 23–44, 1984.
[46] N. Van Thoai, "Solution methods for general quadratic programming problem with continuous binary variables: Overview," in *Advanced Computational Methods for Knowledge Engineering*. Springer, 2013, pp. 3–17.
[47] S. Nyberg, "Physical cell id allocation in cellular networks," Tech. Rep., 2016.
[48] T. Bandh, G. Carle, and H. Sanneck, "Graph coloring based physical-cell-id assignment for LTE networks," in *Proc. IWCMC*, 2009, pp. 116–120.
[49] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge, U.K.: Cambridge Univ. Press, 2004.

JIHONG GUI is currently pursuing the Ph.D. degree in operational research and control theory with the University of Chinese Academy of Sciences, Beijing, China. His research interests are communication network optimization and algorithm design.

ZHIPENG JIANG received the Ph.D. degree in operational research and control theory from the University of Chinese Academy of Sciences in 2011. He is currently a Lecturer with the University of Chinese Academy of Sciences. His research interests are mobile communication network optimization, smart grid, and software defined networking.

SUIXIANG GAO received the Ph.D. degree in mathematics from the Institute of Applied Mathematics, Chinese Academy of Sciences, in 1998. He is currently a Professor with the School of Mathematics, University of Chinese Academy of Sciences, Beijing, China. His research interests are optimization theory and algorithms, communication networks optimization, graph theory, and network flows.