Distributed Multiple-Service User-Experience Energy Saving Algorithm for Base Station

Yuanyuan Zeng, Yu Zhang, Hao Jiang(✉), and Xian Zhou

School of Electronic Information, Wuhan University, Wuhan, Hubei, China
{zengyy,zyzhangyu,jh,zhouxianzx}@whu.edu.cn

Abstract. The issue about energy consumption of wireless network attracts more and more attention and becomes one of three energy efficiency features in 5G. The user associated base station switching on-off strategy can effectively reduce energy consumptions. The existing strategies mainly focus on Quality of Service for users, without fully considering the influence of user experience for specific applications and the problem of high time complexity of implementing the energy-saving strategy caused by the large scale of the 5G mobile network. This paper proposes Distributed Multiple-service User-experience Energy Saving Algorithm for base station (DMUES). Nonlinear integer programming is utilized to model the above problem, which achieves the tradeoff between user experience and energy consumption. By comparing with other relative energy saving strategy, the results show that DMUES achieves good performance on user experience and energy savings. Furthermore, the time complexity of energy saving strategy is reduced due to community partition.

Keywords: 5G · Base energy saving · Quality of Experience · Energy efficiency · Multi-service user experience

1 Introduction

With the advent of the 5G era, the demand for service applications and user experience has forced network operators to find new technologies to reduce operating costs and increase data transmission rate, thereby improving the quality of user experience. The exponential growth of data services has dramatically increased the demand for network facilities and has also caused large energy consumption. And the 5G communication network has been moving towards the deployment of very large-scale and dense base stations (BSs) [3]. Dense BS deployment results in excessive energy consumptions and carbon emissions, which causes hard burden on electrical energy facilities and affects the ecosystem greatly. Green communication, as one of the important requirements of 5G, requires that the total

V. C. M. Leung et al. (Eds.): 5GWN 2019, LNICST 278, pp. 68–85, 2019.
https://doi.org/10.1007/978-3-030-17513-9_6

energy consumption of the network will not increase when the user data traffic increases by a thousand times. It is expected that by 2020, the end-to-end energy consumption per bit in the future 5G network will need to be reduced to 1/1000. The IMT-2020 proposes that the future 5G will also work to improve the operational energy consumption and cost efficiency of network construction [1]. Therefore, in the future, 5G communication will be necessary to reduce energy consumption while maintaining various types of business growth. The increase in energy consumption costs has also led people to pay more and more attention to ways to reduce energy consumption.

Due to people's living habits and mobility characteristics, the traffic patterns of BSs in cities show tidal phenomena and spatial differences, that is, traffic loads of base stations experience significant peak and trough periods, and there may also be significant differences in the traffic loads of different base stations at the same time. According to the dense deployment of cellular base stations in cities, it is energy efficient that can effectively improve BS utilization by appropriately turning off some BSs and offload users to nearby BSs during non-peak times.

The important difference compared to the 4G research idea is that 5G is considered to be the first important user experience as a research core. In other words, in the upcoming 5G era, the most important thing is not speed, but more application and user experience, and a more adaptive business model. However, experience of users who are offloaded to other base stations may be affected to some extent. Traditional BS energy-saving strategy ensures the quality of service (QoS) of users as a prerequisite, mainly concludes service-related indicators such as delay, throughput, jitter, and packet loss while ignoring the influence of business characteristics on user experience. Quality of Experience (QoE) is a measure of acceptability based on user acceptance, Zhao et al. [18] pointed out that QoE can reflect user's satisfaction degree towards the network services based on quantitative modeling of QoS. In the multi-service oriented application architecture, accurate assessment of user experience toward different services according to service characteristics and demands for transmission resources are helpful to the optimization of system resources. To solve the problem, Gómez et al. [5] provided a detailed introduction to the QoE evaluation methods of services such as voice, video streaming and Web browsing, Khan et al. [9] used the Utility Function from Economics to evaluate the specific business services.

QoE overcomes the deficiencies of QoS metrics that ignore service features, it can better describe users' satisfaction with applications or services. Furthermore, considering that the time complexity of solving the problem is growing exponentially with the number of base stations, in this paper we propose a Distributed Multiple-service User-experience Energy Saving Algorithm for base station (DMUES) adopting QoE utility function to evaluate the multi-service user experience, and nonlinear integer programming is utilized to model the energy-saving strategy. And the community is divided into base stations in the city, and the base station energy-saving strategy is implemented separately in each community. Through experiments, it is proved that the proposed DMUES achieve good user experience and energy saving. Furthermore, it greatly reduces the computational complexity and time complexity.

2 Related Work

In the urgent need to achieve energy saving, the green cellular network has become a hot topic for researchers. The traditional energy-saving measures of cellular networks are mainly divided into four categories, namely, improving the energy efficiency of hardware components, optimizing the energy efficiency of the wireless transmission process, base station hibernation/cell zooming, and deploying heterogeneous units [17]. The way to improve the energy efficiency of hardware components requires network operators to replace cellular network system components on a large scale with high implementation costs. Improving energy efficiency by optimizing the wireless transmission process requires a compromise between energy saving and network performance. The method of deploying heterogeneous units is to reduce energy consumption through the use of plug-and-play small cell base stations in the middle of cellular networks (including microcells, picocells etc.) [7].

Cell Zooming is a cell-like breathing concept that changes the base station's coverage area by adjusting the state of the BS. It includes measures such as adjusting base station height, power, sleep, and full shutdown. When a base station is in a low-load state, it can serve users under neighboring base stations by increasing its coverage area and reduce the load on neighboring base stations. Even if all the base stations are under low load, even partial base stations can be completely turned off to achieve energy saving. The BS consumes the largest proportion of energy in the cellular network. By monitoring the traffic of the base station, selectively shutting down some base station transmitting units and cooling equipment (air conditioners, etc.) during off-peak hours, the energy saving effect is very considerable. In 5G heterogeneous networks, the deployment of dense cellular network base stations makes the coverage area of a single base station smaller. The base station's business model is more random, which makes the strategy of shutting down the base station more desirable. Combining 5G heterogeneous network deployment with the base station's sleep mode will yield significant benefits in terms of energy savings [10].

The base station closure policy usually determines the state of the communication unit (base station) of the cellular network by monitoring the traffic load in the network. It is mainly divided into three categories, which are strategies based on traffic load, strategies based on user association and strategies combining traffic load and user association respectively. Han et al. [6] proposed base switching strategy based on traffic load, and solving the problem by centralized greedy decision-making and decentralized autonomous decision-making respectively. Zhu et al. [19] proposed a QoS-aware user association scheme to reconfigure the cell association for energy saving. Oh et al. [13] proposed a practically implementable switching-on/off base energy-saving algorithm considering base load and user association both, and the algorithm can be operated in a distributed manner with low computational complexity. Jiang et al. [8] estimated the aggregate traffic demands of the BS communities and proposed a switch-off strategy while guaranteeing minimal service requirements. Son et al. [16] developed a theoretical (and also practical) framework for BS energy saving that encompasses both dynamic BS operation and user association.

The load based BS on-off switching strategy pays more attention to energy efficiency other than the association status of users. The user association switching strategy needs to switch the BS on-off state frequently according to user association, which brings extra energy consumptions. The hybrid method of joint load and user association based switching strategy avoids frequent status switching with more balanced energy consumptions and better user experience.

For multi-service application architectures, different services have different characteristics and requirements for network transmission quality. The models of users' perceptions of different services are different and need to be evaluated using corresponding models. This will help optimize the resources of the system. The cellular network is a multi-service application architecture. For the characteristics of the business and the demand for network transmission resources, accurate assessment and modeling of different services can help optimize system resources. The European Telecommunications Standards Institute introduced the concept of QoE, and provided a mapping analysis between QoE and network performance indicators for different services. [15] The literature [4] proposes a general formula that links QoE and QoS parameters through an exponential relationship.

3 System Model

In order to describe the energy consumption of the BSs, this section gives the energy consumption model. A BS usually consists of several cells with different radiation azimuths, and cells share the base architecture like cooling system of the BS. Energy consumption of a cell is divided into static consumption (consumption from transmission antenna and power amplifier) and dynamic consumption (consumption related with load/utilization, which varies with cell load), and cells of a BS share the consumption of base architecture.

For the measurement of user experience for specific service types in mobile cellular networks, this paper introduces QoE utility function to model user experience for multi-services.

3.1 Cell Affinity Relationship

To get users' association state to cells, we should first get the coverage area of each cell. Each BS has multiple cells, and the cells belonging to the same BS have the same geographical coordinate. The coverage is different. The coordinate of a point closer to the BS on the center line of the covered cell area with the sector as the position coordinate of the cell is the azimuth, and is a very small positive value. Then we adopt Voronoi tessellation to divide the areas based on the location (latitude and longitude) and antenna radiation azimuth of cells. Cell division example from a typical city of Zhejiang province, China is showed in Fig. 1.

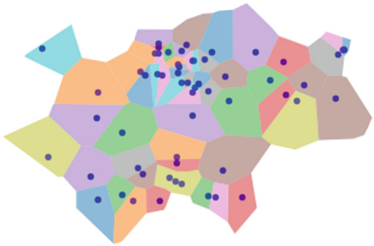

Fig. 1. Coverage areas division of cells

Whether a user can access to a cell depends on user's distance and direction relative to the cell, that is, user-cell affinity relationship. We can construct the affinity graph G_A of users and cells. G_A can be described as a bipartite graph consisting of users and cells. If a user is within the coverage area of a cell, then we can add an edge between them.

In this paper, user association strategy proposed in [2] is adopted. It is assumed that the inter-cell interference and environmental interference are both Gaussian white noises, and the received signal strength of user is only related to the distance. On this condition, users will associate to the nearest active cell it can be associated to. For example, users under cell 1 can associated to cell set, when cell 1 is closed, users under it will be associated to the active cell 7 from which they can receive the best average signal strength.

In a BS switching strategy, users under a closed sector are associated with neighboring cells. In a BS switching strategy, users under a closed cell are associated with neighboring cells. A cooperative relationship between cells is represented by a directed weighted graph. V is the set of sector nodes, and E is the set of edges of nodes. When 80% of users under a cell are under the coverage area of a cell, there are edges in the graph: the size of the edge is the average distance between the user and the sector under the cell. After the cell is closed, users under the sector can associate the set of cells.

3.2 Base Station Energy Comsuption Model

Each BS in the cellular network is consisted of multiple cells with different radiation azimuths, cells share the base station's infrastructure such as the refrigeration equipment. Energy consumption of a cell is consisted of static energy consumption and dynamic energy consumption. Static energy consumption concludes consumption of power amplifiers and antennas and etc. Dynamic energy consumption is needed for processing of the traffic loads, and is positively related to load ρ. Assuming that $\{\rho_1,\rho_2,\rho_3,...,\rho_n\}$ represents the traffic load demands of cells, and $\{\rho_1^*,\rho_2^*,\rho_3^*,...,\rho_n^*\}$ represents cells'current traffic loads with the energy-saving strategy. Power consumption of cells increases linearly with the traffic load:

$$P_{cell}(\rho) = k \cdot \rho + c \tag{1}$$

k is the coefficient, c is the static energy consumption of the cell.

Consumption of a base station is:

$$P_{base} = \sum_{cell \in base} P_{cell}(\rho_{cell}) + C \tag{2}$$

Among the function, C is the consumption of the base architecture of the BS.

3.3 User Experience Model

QoE was proposed by the ITU-T for measuring user's subjective acceptability to applications or services perceived. For the mapping from objective measurements (in terms of QoS) into subjective metrics (in terms of QoE perceived by the user), we adopt the utility function toward multi-services which is defined as the gain users obtained from a service according to the resource requirements. Mobile cellular network services are mainly classified into four types: VoIP services, streaming services, interactive services, and background services. Each type of services has specific characteristic and different transmission resource requirement. Therefore, different forms of utility function are required for modeling multi-service user experience.

VOIP: such as audio calls and video calls, is a hard real-time service that requires strict end-to-end guarantee and is particularly delay sensitive. Its utility function is a unit step function of the allocated resource r. When the allocated resource r is smaller than the minimum transmission resource requirement r_{min}, the utility value is 0, and when the allocated resource equals to r_{min}, the utility value steps to its max value 1 and will not increase anymore, the utility function is as follows

$$U_{VOIP}(r) = \begin{cases} 0, r < r_{\min} \\ 1, r \geq r_{\min} \end{cases} \tag{3}$$

Streaming (STM): audio and video streaming services that can use data buffering technology, having a certain degree of tolerance for media distortions, and can be accepted when packets exceed the latency limitation slightly or even get some loss. The utility function is in exponential form. When the resource r is smaller than a certain smaller value r_{min}, the utility value increases from zero to r, and its growth rate increases. When r gets greater than r_{min}, the growth rate of utility function decreases. Due to the limited system transmission resource r_{max}, the value of the utility function reaches its maximum value when the resource reaches r_{max} and will no longer increase, the utility function is as follows

$$U_{STM}(r) = 1 - e^{-\frac{k_1 r^2}{k_2 + r}}, r \leq r_{\max} \tag{4}$$

Interactive Services (WEB): Mainly are web browsing, chatting, games and other services that have a certain tolerance for delay. The utility function is an exponential function with the allocated resources r. When r is lower than the minimum resource requirement r_{min}, the utility value is 0; when r_{min} gets greater than r_{min}, the utility value steps to a larger value and begins to increase monotonously with r, and the growth rate keeps decreasing. The utility value reaches its maximum when r reaches the limited system transmission resource r_{max}, the utility function is as follows

$$U_{WEB} = \begin{cases} 0, r < r_{\min} \\ 1 - e^{-\frac{k_3 r}{r_{max}}}, r_{min} \leq r \leq r_{max} \end{cases} \tag{5}$$

Background BK: Services such as file transferring and email, which has high tolerance of delay. The utility function is an exponential function of resource r and increases with the resource r with growth rate decreasing. When the allocated resource r reaches the system's maximum resource limit r_{max}, the utility function will reach its maximum value

$$U_{BK}(r) = 1 - e^{-\frac{k_4 r}{r_{max}}}, r \leq r_{max} \tag{6}$$

Among them, k_1, k_2, k_3, k_4 are positive model parameters used to adjust the growth rate of utility function. The range of value for utility function is [0, 1].

The average transmission rate received by users under cell i is: $r_i = Wlog(1+SINR)$, where W is the channel bandwidth, and $SINR = \frac{p_i+g_i}{\sum\limits_{m \in N_i, m \neq j} g_m + \sigma_0}$ is the Signal to Interference plus Noise Ratio, in which p_i is the transmission power of antenna, σ_0 is white noise from the environment, $\sum\limits_{m \in N_i, m \neq j} g_m$ is the inter-cell interference from the neighbor cells which can also be considered as Gaussian white noise, $g_i = L(d_i) = (1 + (\frac{d_i}{40})^{3.5})^{-1}$ is the channel gain caused by channel fading and Rayleigh fading, d_i is the average distance from users under cell i to the cell.

4 Strategy Formulation

In the last section we have introduced the system model in detail, include the base station energy model and multi-service user experience model. The base station switching strategy is closing some cells and offloading users under them to their neighbor cells according to cells cooperation relationship, and ensuring that the user experience be guaranteed. Our goal is to save as much energy as possible at the expense of a smaller loss of user experience, that is, base station energy saving is a multi-objective optimization problem of base station energy and multi-service user experience.

4.1 Optimization Model of DMUES

Assume that the status vector of cells is $\{x_1, x_2, x_3, ..., x_n\}$, $x_i = \{0, 1\}$, let $\{\rho_1, \rho_2, \rho_3 ..., \rho_n\}$ be the cell load demand vector, $\{\rho_1{}^*, \rho_2{}^*, \rho_3{}^* ..., \rho_n{}^*\}$ be the current cell load vector. $\rho_i{}^*$ is related with the status of cell i and status of its neighbors.

From the user-cell affinity graph we can get the cell set N_i that users under cell i can associated to, and the cell set C_i whose serving users can associated to cell i. When cell j in C_i is closed, and neighbor cells closer to cell j than cell i are all closed, cell i will be the nearest active cell of cell j, then the load of cell j will be offloaded to cell i, that is, ${\rho_i}^*$ is related to state of N_i and state of neighbor cells of cells in N_i, ${\rho_i}^*$ can be expressed as follows

$$\rho_{\mathrm{i}}^* = x_i \cdot \rho_i + (1 - x_j) \cdot \sum_{j \in C_i} \rho_j \cdot \prod_{m \in N_j, d_{jm} < d_{ji}} (1 - x_m) \tag{7}$$

As cell consumption is related with load, cell consumption will be

$$E_i = P_{cell}(\rho_i^*) \tag{8}$$

And BS infrastructure consumption is related with status of all cells under it

$$E_{\mathrm{b}} = \mathrm{C} \cdot (1 - \prod_{i \in \mathrm{b}} (1 - x_i)) \tag{9}$$

We adopt the average distance from users' random location in the cell's coverage area to the cell as the distance from user to its accessing cell. Distance from user u under cell i to its accessing cell d_u is related to the state of cell i and N_i

$$d_u = x_i \cdot d_{ii} + (1 - x_i) \cdot \sum_{j \in N_i} x_j \cdot \prod_{m \in N_i, d_{im} < d_{ij}} (1 - x_m) \tag{10}$$

In formula (10), d_{ii} is the distance from user u to cell i, d_{ij} is the distance from user u to its neighbor cell j.

And QoE of user u under cell i whose accessing content type is *cont* is

$$q_u = U_{cont}(r(d_u)) \tag{11}$$

As value of QoE utility function (3–6) is normalized with a maximum of 1, we define the QoE loss as

$$c_u = 1 - U_{cont}(r(d_u)) \tag{12}$$

User set with different accessing content type under cell i is U_i, sum of QoE loss under cell i is

$$Q_i = \sum_{u \in U_i} c_u \tag{13}$$

The purpose of this paper is to achieve energy savings of BSs with less loss of user experience, so the goal can be transferred into getting the minimum value of the optimization function $min(f)$ which combines the energy consumptions and QoE losses.

$$\min(f) = (\sum_{i \in B} E_i + \sum_{b \in B} E_b)/P_i + \eta \cdot \sum_{i \in B} Q_i \tag{14}$$

$$s.t.\rho_{\mathrm{i}}^{*} < \rho_{\max}, \forall i \in \{1, 2, ..., n\} \tag{15}$$

$$s.t.x_i + \sum_{m \in N_i} x_m >= 1, \forall i \in \{1, 2, ..., n\} \tag{16}$$

Among them, η is a parameter for adjustment, by adjusting the value of η we can change the proportion of energy consumption and user experience in the optimization. Increasing the value of η means increasing the importance of user experience. Decreasing the value of a corresponds to increase in importance of energy consumption. For example, if η equals to value 1, it indicates that user's optimal experience is equivalent to a cell's energy consumption. Formula (15) gives the burden limit of cells, and means that traffic load of each cell should not overpass its service capability ρ_{max}. Formula (16) gives the user association protection to ensure that each user should be able to connect to at least one cell.

The multi-objective programming problems of energy consumption and user experience are converted into a joint optimization goal min(f) in this paper. The energy-saving scheme is modeled as a 0-1 integer nonlinear programming problem of the cell status vector $\{x_1, x_2, x_3, ..., x_n\}$, $x_i = \{0, 1\}$. And the goal is to get a set $\{x_1, x_2, x_3, ..., x_n\}$ to achieve the joint optimization of energy consumption and user experience. And the tradeoff between the energy consumption and the user experience will be achieved.

This paper uses the classical branch-and-bound algorithm to solve the 0-1 integer nonlinear programming problem. The thought of branch-and-bound algorithm [14] is to relax discrete 0-1 variables into continuous variables range from 0 to 1, and decompose the original problem into disjoint relaxation sub-problems. Target value corresponding to the relaxation solution is taken as the upper bound of the original problem, after iterations, we can get the optimal solution of the original problem. In the end, we use the optimization software Lingo to solve the above 0-1 integer nonlinear programming problem.

4.2 Problem Solving

The optimization problem in this paper is a 0-1 polynomial programming problem. Branch-and-bound method solves the deterministic solution of nonlinear integer programming and is more efficient in solving hybrid nonlinear integer programming.The basic idea is to generate a branch and delimitation tree and a sequence of upper and lower bounds. By decomposing the original problem into a series of disjoint relaxed sub-problems, the target value corresponding to the solution of the relaxed sub-problem is taken as the upper bound of the original problem, and the iterative solution to the original problem is obtained [14].

The integer programming problem in this paper is a convex nonlinear 0-1 integer programming problem, which is a pure integer programming problem. Pure Integer Nonlinear Programing (PINLP). We use an improved linear/nonlinear branch and bound method for 0-1 integer programming(LP/NLP based branch-and-bound, LP/NLP-BB).

Relaxation strategies include relaxation of the original problem to continuous space (NLP sub-problems) and linearization of nonlinear constraints (ILP sub-problems).

The optimization goal of Eq. (14) can be expressed as follows:

$$\begin{aligned} & Z_{INLP} = \min f(x), \\ & s.t. g(x) \leq 0, \\ & m(x) \leq 0, \\ & x \in \{0, 1\}^n \end{aligned} \tag{17}$$

$g(x) = \{g_1(x), g_2(x), ..., g_n(x)\}$ is the nonlinear constraint of Formula (15), $m(x) = \{m_1(x), m_2(x), ...m_n(x)\}$ is the linear constraint of Formula (16), $x = \{x_1, x_2, ..., x_n\}$

ILP sub-problems: If the original target (14) is non-linear, its optimal solution may be the interior point in the convex hull. Such a problem should not be solved directly. Therefore, we introduce an auxiliary variable to convert a nonlinear target into a linear target, and the original objective function is treated as a constraint to obtain the equivalent INLP problem.

$$\begin{aligned} & Z_{INLP} = \min \eta, \\ & s.t. f(x) \leq \eta, \\ & g(x) \leq 0, \\ & m(x) \leq 0, \\ & x \in \{0, 1\}^n, \eta \in R \end{aligned} \tag{18}$$

Due to the convexity of the objective function and nonlinear constraints, so at the current point $\hat{x}$,the inequalities $f(x) + \nabla f(\hat{x})^T(x - \hat{x}) \leq f(x)$ and $g(x) + \nabla g(\hat{x})^T(x - \hat{x}) \leq g(x)$ are established. So we can convert constraints to linear constraints, then the original goal is relaxed as follows:

$$\begin{aligned} & Z_{INLP} = \min \eta, \\ & s.t. f(x) + \nabla f(\hat{x})^T(x - \hat{x}) \leq 0 \\ & g(x) + \nabla g(\hat{x})^T(x - \hat{x}) \leq 0 \\ & m(x) \leq 0 \end{aligned} \tag{19}$$

NLP sub-problems: Another relaxation relaxes the integer variable of the original problem to continuous space. Assume that each component meets the limit $0 \leq l_I \leq x \leq u_I \leq 1$, The original problem will be transformed into a general nonlinear programming sub-problem.

$$\begin{aligned} & Z_{NLPR}(l_I, u_I) = \min f(x), \\ & s.t. g(x) \leq 0 \\ & m(x) \leq 0 \\ & 0 \leq l_I \leq x \leq u_I \leq 1 \end{aligned} \tag{20}$$

If (l_I, u_I) is the upper bound of the feasible domain (17), then the corresponding target value of (20) is to provide an effective lower bound for the original problem. The upper bound of (17) is provided by a feasible solution to the relaxed sub-problem. In combination with the linear relaxation of constraints and the continuation of integer variables, when the branch-and-bound method is solved, if an integer point is considered as a feasible solution to the original problem, the original problem becomes a series of LP sub-problems.

$$\begin{aligned}
&Z_{LP/NLP} = \min \eta, \\
&s.t. f(x) + \nabla f^T(\hat{x})(x - \hat{x}) \leq 0, \\
&g(x) + \nabla g^T(\hat{x})(x - \hat{x}) \leq 0, \\
&m(x) \leq 0, \\
&x \in \{0,1\}^n, \eta \in R
\end{aligned} \tag{21}$$

Relaxation algorithm provides a basis for the delimitation of the branch and bound method. The specific LP/NLP-branch-demarcation algorithm steps are as follows:

First, we use a heuristic algorithm to find the initial feasible solution to the optimization objective function. namely the root node, and then relax the integer variable to $l_I \leq \hat{x} \leq u_I$.

Step 1: Select node. Starting from the root node, we select a node in the branch-and-bound tree to solve the loose sub-problem. If this problem is not feasible, we delete this node and search for the child nodes of the branch tree. Otherwise we assume the solution $\hat{x}$.

Step 2: Pruning. If the value of the objective function corresponding to the solution at this node is greater than the current upper bound, the feasible region of this portion does not contain the optimal solution, and the branch is cut off.

Step 3: Check integer constraints. If the point $\hat{x}$ does not satisfy the integer constraint, we select a variable that does not satisfy the integer constraint and add the left and right branch constraints $\hat{x}_i = 0$ and $\hat{x}_i = 1$. Otherwise, if $\hat{x}$ satisfies the integer constraint condition and the objective function value is less than the current optimal value, the upper bound is updated, and the branch whose objective function value is greater than the current upper bound is trimmed.

Step 4: Check if the branch delimiter tree is empty. If the branch and delimitation tree is not empty, then it returns to step 1. Otherwise, the algorithm will be terminated and the current optimal solution will be output.

4.3 Implementation of DMUES

The branch-and-bound method finds the integer optimal solution in the feasible domain of linear relaxation model of integer programming according to certain search rules. The time complexity of its solution will still increase exponentially with the number of base stations. To implement base station energy-saving strategy in large-scale cellular network clusters, the time complexity is very high. We consider the community division of BS clusters based on the cooperative relationship between cells. BS energy saving strategy is implemented separately in each community BS to achieve distributed computing to reduce time complexity.

The BSs in the city are deployed according to the functional areas of the city and the users' demands, which are distributed unevenly. We use space cooperation network to describe the cooperation between BS cells in urban space.

We use space cooperation network G to describe the cooperation between cells in urban space. The graph is usually expressed as $G = (V, E)$. It consists of nodes and edges. V represents a set of nodes and E represents a set of edges. The network consists of the cell node $<V_i, V_j>$ and the associated edges W_{ij} between the sectors. When there is a cooperative relationship between sectors, increase the edge in .The edge weight is a measure of the tightness between cells and is inversely proportional to the distance $W_{ij} = e^{-{d_{ij}}^2/(2 \cdot delta^2)}$.

When there is a cooperative relationship between sectors, the smaller the distance between the sectors, the greater the connection tightness between them; the greater the distance between base stations, the smaller the connection tightness.

In the connection relationship between base station sectors in the main urban area of Jinhua City, the distribution of base stations in cities is uneven. And the closer we get to the city center, the more dense the base station is. The distance between the BS is close. The connection is more and the network structure, has certain community structure. In the suburban areas, the deployment of BSs is sparse and the connection between base stations is sparse, presenting a distinctly small community structure.

The Louvain algorithm is used to divide the network into communities, and the base station energy-saving strategy is implemented separately in each community, which can effectively reduce the time complexity and can be used to save energy in large-scale cellular networks.

5 Experiment Results

In this section, we numerically evaluate the proposed MUES by comparing its performance to several reference algorithms based on the real China Mobile UDR dataset in Jinhua District of Zhejiang province. We use the method of community division to divide the BS clusters in the city. And then we implement energy-saving strategies in each community to reduce the time complexity.

5.1 Description of Dataset

We analyze the cellular network of the urban areas in a city of Zhejiang province, China. 59 dense cells of a square area is selected for experiment, the coverage areas of cells is shown in Fig. 1. The UDRs used in this paper is based on users' Internet access behavior which conclude fields of session time, cell location information, URL and traffic load, etc. Table 1 shows the main fields and record example of UDRs.

5.2 Experiment Results and Analysis

Time Complexity. The optimization problem that needs to be solved in the BS energy-saving strategy in this paper is the 0-1 combination planning problem. The time complexity of the solution is exponentially increasing with the number

Table 1. Field of the usage detail records

userID	Session time	Location	URL	Traffic load/B
69201446765	2014-11-21 20:18:24	6893_379A	news.baidu.com	107031

of targets. Firstly, The community energy-saving strategy optimization model with different sectors is modeled. Then branch and bound method is used to solve the optimization problem. The logarithmic curve of the solution time and the number of cells is shown in Fig. 2.

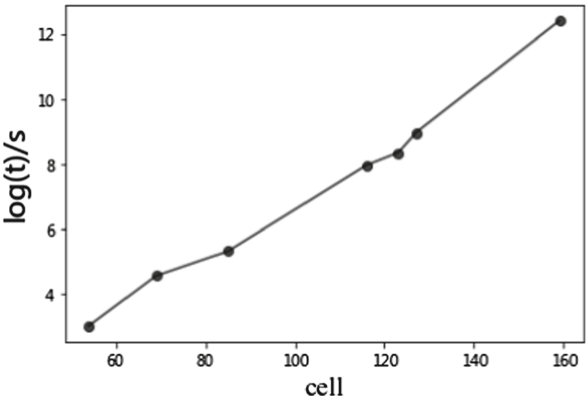

Fig. 2. Time logarithmic curve

The time complexity of the energy saving strategy optimization problem is an exponential function of the number of cells: $T(n) = k \cdot 2^n$. The branch-and-bound solver used in this paper runs on ordinary computer windows systems. When the number of cells is 127, the average optimization solution time reaches 20 min, which is more than 1/3 of the implementation period (1 h) of the energy saving strategy, and the implementation efficiency is low. When the number of cells is 160, the complete iterative search process time for the entire optimization problem reaches 1.5 h, which is much larger than the implementation period of the energy-saving strategy, and there is no possibility of implementation. Therefore, We consider the community division of BS clusters based on the cooperative relationship between cells. Community division can reduce the complexity of the network, so the efficiency of energy-saving strategies after community division is higher than that of energy-saving strategies based on global networks. For example, if we divide the BS into m communities and implement the energy saving strategy in each BS, then the time complexity of the energy saving strategy is $D(n) = k \cdot m \cdot 2^{\frac{n}{m}}$. And before the community is divided, the time complexity of the energy saving strategy is $T(n) = k \cdot 2^n$.Obviously, T(n) is greater than D(n), so community division can reduce the time complexity of energy saving strategy.

Community Division Results. In order to solve the problem of large-scale cellular networks and energy efficiency of energy-saving strategies, this paper adopts the method of community division of BS clusters to realize the distributed implementation of energy-saving strategies. The construction of the base station network and the edges status are described in Sect. 3.1. The Louvain algorithm was used to divide the BS network into communities. The results of the community division are shown as follows: The number of nodes is 1444. The number of edges is 3821. The number of communities is 39. The maximum number of BSs included in the community is 107. The minimal number of BS included in the community is 2. The average number of BSs included in the community is 37. The modularity 0.8745.

The total number of nodes of the BS in the network is 1,444, and the total number of edges of the network is 3,821, which is divided into 39 independent communities. The modularity range is $[-0.5, 1)$, The greater the degree of modularity is, the stronger the community characteristics of the network and the greater the independence of the divided communities are. That is, the better the result of community division is. Studies have shown that when the value of modularity is greater than 0.5, the results of community division are better. The modularity of the BS network is as high as 0.8745. It shows that the BS clusters in cities have strong community characteristics and strong independence between communities. At the same time, in the divided communities, there is no situation where the number of BSs in the community is particularly large, and the community is divided more evenly, which facilitates the implementation of distributed computing.

With arcgis you can project points in space into real geographic space. Figure 3 projects the nodes of 1444 BSs in Lucheng District of Jinhua City onto the streets of Jinhua City. Different colors are used to mark different communities so that the distribution of sector nodes in the network and the community structure of the network can be seen. It can be seen that the nodes of different communities are separated in geographical space. The BS in the same community are very close. The characteristics of the community structure are obvious, and the number of nodes in a single community is relatively uniform, which is conducive to the distributed implementation of the energy-saving strategy.

Performance Analysis. In order to verify the performance difference between implementing the BS energy-saving strategy for the BS cluster of the entire cellular network and implementing the energy-saving strategy for each community after dividing the community, this paper chooses to implement DMUES as a distributed implementation solution for two geographically adjacent communities respectively. We combine two communities as a community to implement DMUES as a centralized implementation plan and compare the results of the two scenarios. The number of cells selected for the community 1 is 54, and the number of cells selected for the community 2 is 59. The total number of cells after the merger is 113.

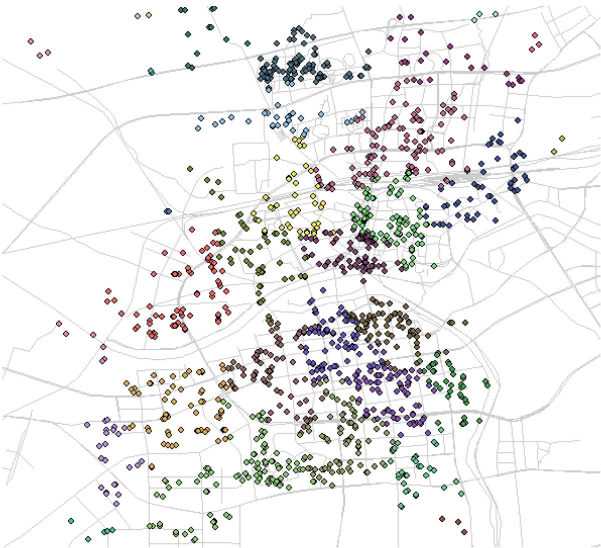

Fig. 3. Network community structure of base station

From Fig. 2, we can see that the complexity of solving the optimization problem has an exponential growth relationship with the number of variables. The experimental results show that the number of sectors in community 1 is 54. And the average calculation time is 8.14 s. The number of sectors in community 2 is 59. And the average calculation time is 23.7 s. But when merging the adjacent communities in these two spaces into a community with 113 sectors, the optimization solution time increases to 327 s, which is more than ten times.

We partition a day into 24 hourly intervals, the experiment is conducted on the off-peak hours from 0 o'clock to 6 o'clock. We first statistic the user numbers, access service types and traffic loads of cells. Then we construct the energy consumption model and user experience model according to the user-cell affinity graph. The object optimization problem is formulated and we solve this problem using branch-and-bound algorithm. Our strategy is compared with the energy saving strategy proposed in [11] that only considered the flow-level performance (ES-FLD) and the classical energy-saving strategy that used the blocking rate as the user experience (ES-BR) in literature [12]. The energy saving strategy is compared with the aspects of the number of active cells, energy saving ratio and user QoE loss. And the result is shown in Fig. 4. From the experimental results, it can be seen that the energy-saving strategy ES-FLD achieves a maximum of 86% energy saving. It only considers the flow-level performance. The QoE loss increases several times, which seriously affects the user experience. The energy-saving

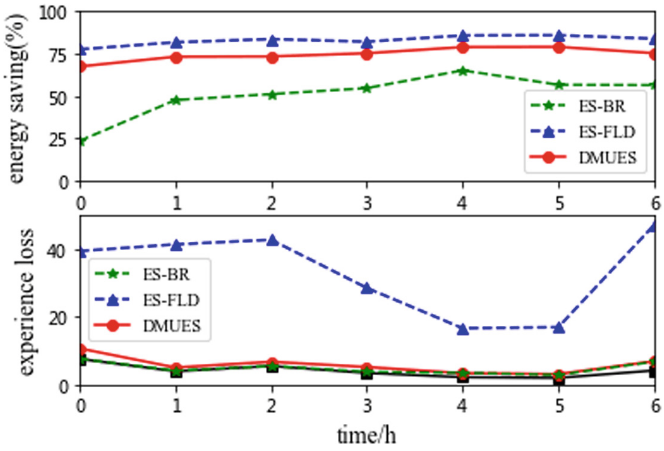

Fig. 4. Method performance at non-peak times

strategy in [12] achieves 50% energy saving. The growth rate of the QoE loss in [12] is approximately the same as DMUES proposed in this paper. DMUES proposed in this paper achieves a maximum of 79% energy saving, with very little increase in user experience loss with better performance.

Therefore, compared with ES-FLD, which considers only the flow-level performance cost, the DMUES achieves a higher energy saving rate with only a little increase in the user experience cost. Compared with the energy-saving strategy in [12], under the premise of guaranteeing the cost of user experience, more than 20% of energy is saved. In general, compared with traditional base station energy-saving strategies, DMUES achieves a good energy-saving effect and guarantees the quality of user experience, and achieves the best combination of the two.

At the same time, this paper discusses the load utilization of active cells before and after implementing the DMUES at 0 o'clock, which is shown in Fig. 5. The results show that most of the cells in the city are under low utilization at 0 o'clock. And after the MUES is implemented, most of the cells with low utilization are turned off while cells whose utilizations are high are basically been retained. The strategy also ensures that load of a single cell is in the affordable range. In conclusion, the strategy not only improves the energy efficiency of cellular network base station, but also ensures the load balance of the active cell. Furthermore, the computational complexity and time complexity is greatly reduced. This proves that the DMUES proposed in this paper is very effective and feasible.

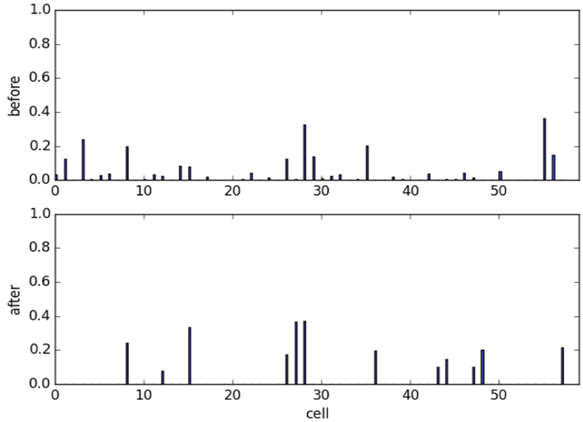

Fig. 5. Traffic loads of cells before and after implementing the DMUES

6 Conclusion

Current base station energy-saving strategy has not fully considered the impact of service characteristics on user experience and the problem of high time complexity of implementing the energy-saving strategy caused by the large scale of the 5G mobile network. In this paper, we introduce user experience quality as the measure of the subjective experience of users on the received service and propose a Distributed Multiple-service User-experience Energy Saving Algorithm for base station. And we use a complex network graph model to describe the cooperation relationships between BSs based on the proximity of BSs and the closeness of interaction in the city. Using the method of community division, the urban BS is divided into multiple communities, and the energy-saving strategy is implemented separately for each community. Part of the cellular network in a city of China is selected, and the experiment was conducted on the non-peak hours in this scenario based on real UDRs. Compared with the existing energy-saving strategies, the results show that the DMUES proposed in this paper can achieve good energy savings under the premise of less user experience loss. Furthermore, the time complexity of energy saving strategy is reduced due to community partition.

References

1. Budzisz, L., Ganji, F., Rizzo, G., Marsan, M.A.: Dynamic resource provisioning for energy efficiency in wireless access networks: a survey and an outlook. IEEE Commun. Surv. Tutor. **16**(4), 2259–2285 (2014)
2. Ding, M., Wang, P., López-Pérez, D., Mao, G., Lin, Z.: Performance impact of LoS and NLoS transmissions in dense cellular networks. IEEE Trans. Wirel. Commun. **15**(3), 2365–2380 (2016)

3. Feng, M., Mao, S., Jiang, T.: Base station ON-OFF switching in 5G wireless networks: approaches and challenges. IEEE Wirel. Commun. **24**(4), 46–54 (2017)
4. Fiedler, M., Hossfeld, T., Tran-Gia, P.: A generic quantitative relationship between quality of experience and quality of service. IEEE Press (2010)
5. Gómez, G., de Torres, E., Lorca, J., García, R., Pérez, Q., Arias, E.: Assessment of multimedia services QoS/QoE over LTE networks. In: Obaidat, M.S., Filipe, J. (eds.) ICETE 2012. CCIS, vol. 455, pp. 257–272. Springer, Heidelberg (2014). https://doi.org/10.1007/978-3-662-44791-8_16
6. Han, T., Ansari, N.: Enabling mobile traffic offloading via energy spectrum trading. IEEE Trans. Wirel. Commun. **13**(6), 3317–3328 (2014)
7. Hu, R.Q., Qian, Y.: Macro-Femto Heterogeneous Network Deployment and Management. Wiley, Hoboken (2013)
8. Jiang, H., et al.: Data-driven cell zooming for large-scale mobile networks. IEEE Trans. Netw. Serv. Manag. **15**(1), 156–168 (2018)
9. Khan, M.A., Toseef, U.: User utility function as quality of experience (QoE). In: Proceedings of the ICN, vol. 11, pp. 99–104. Citeseer (2011)
10. Li, W., Zheng, W., Xie, Y., Wen, X.: Clustering based power saving algorithm for self-organized sleep mode in femtocell networks. In: International Symposium on Wireless Personal Multimedia Communications, pp. 379–383 (2012)
11. Liu, B., Zhao, M., Zhou, W., Zhu, J.: Flow-level-delay constraint small cell sleeping with macro base station cooperation for energy saving in hetnet. In: Vehicular Technology Conference, pp. 1–5 (2015)
12. Niu, Z., Wu, Y., Gong, J., Yang, Z.: Cell zooming for cost-efficient green cellular networks. IEEE Commun. Mag. **48**(11), 74–79 (2010)
13. Oh, E., Son, K., Krishnamachari, B.: Dynamic base station switching-on/off strategies for green cellular networks. IEEE Trans. Wirel. Commun. **12**(5), 2126–2136 (2013)
14. Quesada, I., Grossmann, I.E.: An LP/NLP based branch and bound algorithm for convex MINLP optimization problems. Comput. Chem. Eng. **16**(10–11), 937–947 (1991)
15. Shaikh, J., Fiedler, M., Collange, D.: Quality of experience from user and network perspectives. Ann. Telecommun. - annales des télécommunications **65**(1–2), 47–57 (2010)
16. Son, K., Kim, H., Yi, Y., Krishnamachari, B.: Base station operation and user association mechanisms for energy-delay tradeoffs in green cellular networks. IEEE J. Sel. Areas Commun. **29**(8), 1525–1536 (2011)
17. Wu, J., Zhang, Y., Zukerman, M., Yung, K.N.: Energy-efficient base-stations sleep-mode techniques in green cellular networks: a survey. IEEE Commun. Surv. Tutor. **17**(2), 803–826 (2015)
18. Zhao, T., Liu, Q., Chen, C.W.: Qoe in video transmission: a user experience-driven strategy. IEEE Commun. Surv. Tutor. **19**(1), 285–302 (2017)
19. Zhu, Y., Kang, T., Zhang, T., Zeng, Z.: QoS-aware user association based on cell zooming for energy efficiency in cellular networks. In: IEEE International Symposium on Personal, Indoor and Mobile Radio Communications, pp. 6–10 (2014)