

Research Article

Advanced Load Balancing Based on Network Flow Approach in LTE-A Heterogeneous Network

Shucong Jia, Wenyu Li, Xiang Zhang, Yu Liu, and Xinyu Gu

Beijing University of Posts and Telecommunications, Beijing 100876, China

Correspondence should be addressed to Shucong Jia; jiashucongjsc@163.com

Received 20 February 2014; Accepted 21 April 2014; Published 20 May 2014

Academic Editor: Lin Zhang

Copyright © 2014 Shucong Jia et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Long-term evolution advanced (LTE-A) systems will offer better service to users by applying advanced physical layer transmission techniques and utilizing wider bandwidth. To further improve service quality, low power nodes are overlaid within a macro network, creating what is referred to as a heterogeneous network. However, load imbalance among cells often decreases the network resource utilization ratio and consequently reduces the user experience level. Load balancing (LB) is an indispensable function in LTE-A self-organized network (SON) to efficiently accommodate the imbalance in traffic. In this paper, we firstly evaluate the negative impact of unbalanced load among cells through Markovian model. Secondly, we formulate LB as an optimization problem which is solved using network flow approach. Furthermore, a novel algorithm named optimal solution-based LB (OSLB) is proposed. The proposed OSLB algorithm is shown to be effective in providing up to 20% gain in load distribution index (LDI) by a system-level simulation.

1. Introduction

Nowadays, smart phone and tablet users are growing rapidly. The remarkable explosion of mobile internet traffic requires wireless communication systems to support higher data rate. Various kinds of transmission techniques in wireless propagation environment were applied to meet the growing demand, such as the high-order multiple input multiple output (MIMO) [1] and the heterogeneous network, where some lower power nodes are overlaid within a macro network. Long-term evolution advanced (LTE-A), which was standardized by the 3rd generation partnership project (3GPP) [2], is a promising wireless communication system to provide high date rate and spectral efficiency. The bandwidth of LTE-A can be up to 100 MHz by using carrier aggregation technology, which guarantees effective bandwidth allocation to a user through concurrent utilization of radio resources across multiple carriers and efficient carrier scheduling schemes [3]. To improve the service quality of cell edge users, some low power nodes can be deployed at the edge of a cell, creating what is referred to as a heterogeneous network. Besides the transmission techniques mentioned above, some other techniques (e.g., coordinated multipoint transmission and

reception) are applied to improve the performance of LTE-A system. However, there are still some challenges in deploying a real LTE-A system. For example, in LTE-A, the traffic request of some cells may be far higher than an acceptable level, named as “hotspots,” while some of the other cells may have extra resources to serve more users, which would result in load unbalance and user dissatisfaction. As the topology of the LTE-A heterogeneous network is more complex, network planning and optimization bring a heavy burden to LTE-A network operators. Self-optimizing network (SON) is a solution to relieve the burden by selecting and adjusting the key parameters in the LTE-A system automatically [4]. Load balancing (LB), which hands off some users of a heavy payload cell to neighboring comparatively less loaded cells, has been widely discussed to increase the network resource utilization.

2. Related Work

There are a great deal of articles which analyze the load balancing problem of cellular networks. To equalize load among cells, power control algorithms were proposed in [5], which have reduced (or risen) the transmission power to

contract (or expand) the coverage of heavy (or low) payload cells. By controlling beam, coverage patterns of “common signals,” sizes, and shapes of cells can be automatically adjusted to balance cell load [6]. In [7], the cell-specific offset was adjusted automatically based on payloads of the source cell and the neighboring cells. A two-layer mobility LB algorithm was discussed in [8], where the overloaded cell can choose a target cell by considering the situation of its two layer surrounding cells. Authors in [9] selected the appropriate LB method from handover parameter control and cell coverage control according to the situation. To cope with the potential ping-pong load transfer and low convergence issues, authors in [10] proposed a game-theoretic solution to the SON LB. In [11], a multidomain LB framework was proposed, which focuses on reducing the radio resource cost and mitigating the cochannel interference across domains in the heterogeneous network. In [12], authors proposed an ant colony self-optimizing method for LB. In the method, first the load of all cells is estimated; then some users are selected to be handed over to the neighbor cells according to the stimulation intensity of all users in the cell. But none of the above researches analyzes either the optimal target cell for a heavily loaded cell or the optimal number of users that should be transferred between two cells.

There are a lot of articles which analyze wireless network through a Markovian model. In [13], the blocking probability of different types of service in a network is calculated. The authors of [14] performed a stochastic performance analysis of a finite-state Markovian channel shared by multiple users and derived delay and backlog upper bounds based on the analytical principle behind stochastic network calculus.

In this paper, we firstly evaluate the negative impact of load imbalance among cells through a Markovian model. Secondly, we present a mathematical model for LB and introduce the network flow approach to derive the optimal solution. Finally, we present a novel LB algorithm based on the optimal solution. Compared with the previous LB algorithms, our method can not only lighten the load of the busy cells but also avoid handover to much traffic to a low payload target cell and change the target cell into a busy cell.

The rest of this paper proceeds as follows: the scenario of a LTE-A heterogeneous network is described in Section 3. In Section 4, we evaluate the negative impact of load imbalance among cells through a Markovian model. In Section 5, we formulate LB as an optimization problem, analyze load balancing based on the network flow approach, and introduce a novel LB algorithm in an LTE-A heterogeneous network. In Section 6, a system-level simulation model is presented and the simulation results are analyzed. The paper draws a conclusion in Section 7.

3. The Scenario

The scenario we considered is a heterogeneous network composed of macrocells and picocells whose coverage is provided by macrobase stations (Macro eNBs) and picobase stations (pico eNBs), respectively [16], as shown in Figure 1. The combination of one macrocell and some picocells overlaid

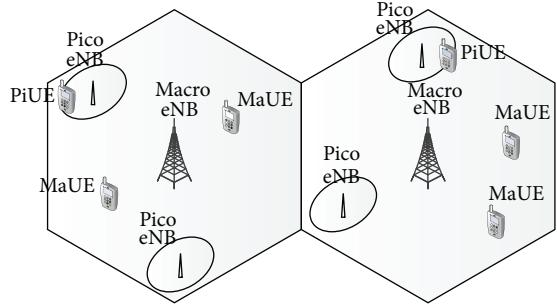


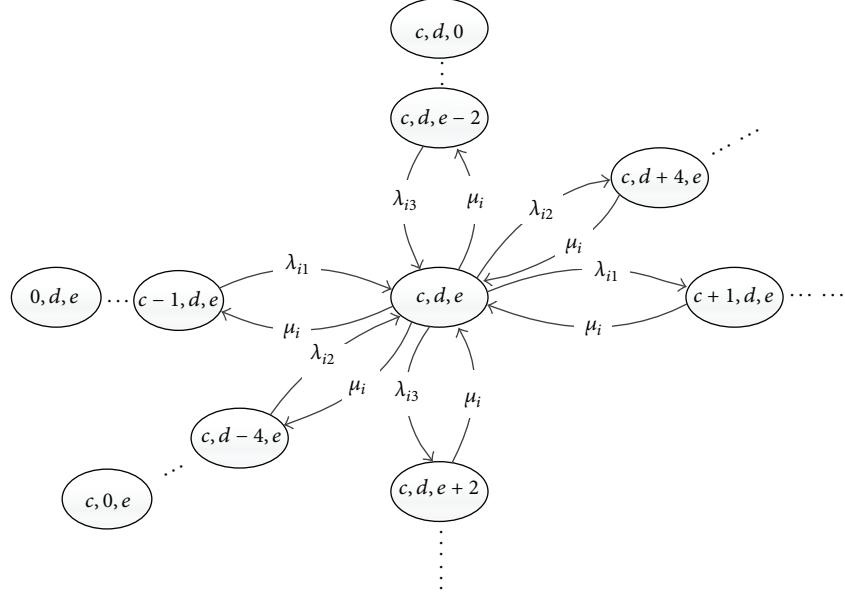
FIGURE 1: The heterogeneous network.

within the macrocell can be named as cell. The users served by a macro eNB are referred to as macrourers (MaUEs) and the users served by a pico eNB are referred as picousers (PiUEs). The system bandwidth of each cell is equal and the frequency spectrums of each cell are divided among macrocell and the picocells to avoid interference between a MaUE and a PiUE.

4. Impact of Load Imbalance

In this section, we evaluate the negative impact of load imbalance among cells through a Markovian model. To simplify the analysis, we consider the load imbalance between two cells (cell 1 and cell 2). The arrival of user is assumed as a Poisson process and the arrival rate of user in cell i is assumed as λ_i . We assume that users are equally distributed across three locations: macrocell center, macrocell edge, and picocell. We assume that the arrival rate of center users of cell 1 is λ_{11} , the arrival rate of edge users of cell 1 is λ_{12} , and the arrival rate of picousers of cell 1 is λ_{13} . Parameters λ_{21} , λ_{22} , and λ_{23} are the same meanings for cell 2. We assume that the service time of users follows negative exponential distribution. The service rate of all users in cell i is assumed as μ_i . The signal to interference and noise ratio of cell center users is larger than that of cell edge users, so the transmission rate of one physical resource block (PRB) for a cell edge user is smaller than that of a cell center user [17]. Therefore, we assume that the number of physical resource blocks (PRBs) needed by a center user is one, the number of PRBs needed by an edge user is four, and the number of PRBs needed by a picouser is two. The total number of PRBs of each cell is assumed to be 100. Then the PRBs occupied by users in cell i can be evaluated by a three-dimensional Markovian model, as shown in Figure 2.

The state (c, d, e) denotes that the number of PRBs occupied by cell center users of cell i is c , the number of PRBs occupied by cell edge users of cell i is d , and the number of PRBs occupied by picousers of cell i is e . The number of PRBs occupied by total users of cell i is $c + d + e$, which can not be larger than the total number of PRBs of each cell, that is, 100. If the free PRB number in a cell is no smaller than the number of PRBs required by a user, the cell will allocate the PRBs to the user. Otherwise, the requirement from the user will be rejected. The probability that n PRBs are used by all

FIGURE 2: The three-dimensional Markovian model for the PRB occupation of cell i .

users can be respected by the stationary distribution $q(n)$ in [13]; $q(n)$ is determined by the recursive formula as follows:

$$q(n) = \sum_{s=1}^S \frac{a_s \cdot b_s}{n} \cdot q(n - b_s) \quad n = 0, 1, \dots, N, \quad (1)$$

where $q(n) = 0$ for $n < 0$ and $\sum_{n=1}^N q(n) = 1$. S is the number of service types, that is, the dimension of the model. In our case, there are three service types: the service of macrocell center user, the service of macrocell edge user, and the service of picocell user. $a_s = \lambda_s / \mu_s$ is the type s offered load. b_s is the number of PRBs required by type s . N is the total number of PRBs of an LTE cell.

The blocking probability P_{b_s} of type s user can be calculated as

$$P_{b_s} = \sum_{n=N-b_s+1}^N q(n) \quad s = 1, 2, \dots, S. \quad (2)$$

Using formulas (1) and (2), we can calculate the blocking probability of users in case of different traffic densities and load distributions.

For example, we consider two load distribution scenarios between two cells. In the first scenario, we assume that the total arrival rate of users in cell 1 is three times larger than the arrival rate of users in cell 2. In the second scenario, we assume that the total arrival rate of users in cell 1 is equal to the total arrival rate of users in cell 2. Besides, we assume that the arrival rate of total users, that is, users in cell 1 with the addition of users in cell 2, is equal in the two load distribution scenarios. Moreover, the resource requirement and the service ratio are assumed to be the same in the two load distribution scenarios and the total number of PRBs of each cell is 100. Some detail parameters are presented in

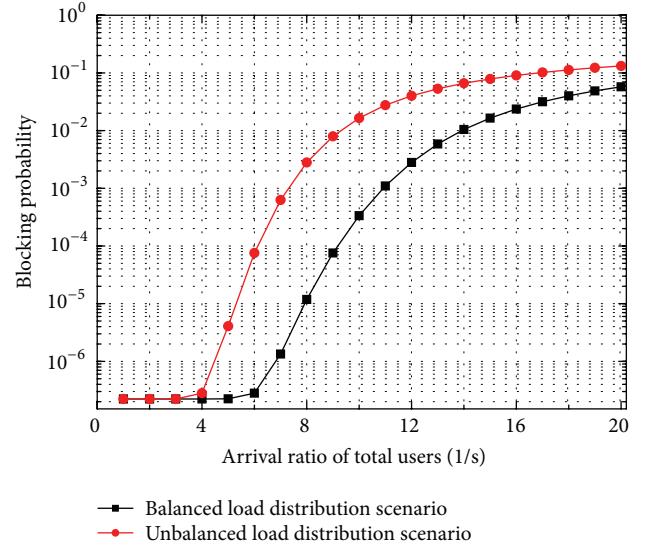


FIGURE 3: The blocking probability of users in two load distribution scenarios versus the arrival rate of total users.

Table 1. Using formulas (1) and (2), we calculate the blocking probability of user in two load distribution scenarios as the arrival rate of total users λ_{total} increasing from 1 to 20, as shown in Figure 3.

From Figure 3, we can see that although the total traffic is the same in the two load distribution scenarios, the blocking probability of users in case of unbalanced load distribution is larger than the blocking probability of users in case of balanced load distribution. So we need using LB to hands off some users of heavily loaded cell to neighboring comparatively less loaded cells, for the purpose of improving network performance.

TABLE 1: The Markovian model parameters.

Scenario	User type	λ	μ	b_s
Unbalanced load distribution scenario	The center user of cell 1	$0.25 * \lambda_{\text{total}}$	0.2	1
	The edge user of cell 1	$0.25 * \lambda_{\text{total}}$	0.2	4
	The picouser of cell 1	$0.25 * \lambda_{\text{total}}$	0.2	2
	The center user of cell 2	$0.083 * \lambda_{\text{total}}$	0.2	1
	The edge user of cell 2	$0.083 * \lambda_{\text{total}}$	0.2	4
	The picouser of cell 2	$0.083 * \lambda_{\text{total}}$	0.2	2
Balanced load distribution scenario	The center user of cell 1 or 2	$0.167 * \lambda_{\text{total}}$	0.2	1
	The edge user of cell 1 or 2	$0.167 * \lambda_{\text{total}}$	0.2	4
	The picouser of cell 1 or 2	$0.167 * \lambda_{\text{total}}$	0.2	2

5. LB Based on Network Flow Approach

As described in Section 4, load imbalance between two adjacent cells will affect the resource utilization of the two cells. If there are more cells in a system, the problem of how to balance the load among cells is more complex. In this section, we formulate the LB problem as an optimization problem and solve the problem using network flow approach.

5.1. LB Problem Formulation and Analysis. We consider a network consisting of n cells and several users. Among the n cells, the number of physical resource blocks (PRBs) needed by a cell k to support the traffic of users in the cell k is denoted by N_k . Due to traffic distribution imbalance, N_k varies in size. We assume a scenario like Figure 4, where cells A and C have one user, and cells B and D have three users. In Figure 4, we assume each user needs the same number of PRBs to keep the analysis simple. Two LB schemes are shown in Figure 4. In scheme 1, a user in cell B is switched to cell C. After the LB, the numbers of users in four cells are 1, 2, 2, and 3, respectively. In scheme 2, a user in cell B is switched to cell A and a user in cell D is switched to cell C. After the LB, the number of users in each cell is 2. It is obvious that scheme 2 is better than scheme 1.

From the example we think that LB should be analyzed among multiple cells rather than just between two cells. If there are n cells, we need to switch user among cells to make N_k approaching $\bar{N} = (1/n) \sum_{k=1}^n N_k$. \bar{N} is the average number of PRBs needed by a cell. However, switching among cells has some signaling overhead, and the switched user's signal quality may decrease. Therefore, in LB, we should minimize the number of handovers. Then the problem is equivalent to an optimization problem that can be written as

$$(P1) \quad \min \left(\sum_{i=1}^n \sum_{j \neq i} P_{i,j} \right) \quad (1 \leq j \leq n)$$

$$\text{s.t.} \quad N_i - \sum_{j \neq i} P_{i,j} + \sum_{j \neq i} P_{j,i} = \bar{N}, \quad (3)$$

$$(1 \leq i \leq n, 1 \leq j \leq n),$$

where $P_{i,j}$ is the number of PRBs occupied by business which switched from cell i to cell j . It should be noticed

that there may be no solution to the constraint equation of (3) if $P_{i,j}$ are integers, so we assume that $P_{i,j}$ are real numbers in this subsection and round off $P_{i,j}$ in the novel LB algorithm subsection. We analyze the optimization problem in the following cases of Figure 5.

5.1.1. Case 1: Three Cells in a Row. To balance the load of three cells in a row, we firstly need calculate the average load of one cell \bar{N} . Secondly, we start to balance the load from the cell in the left side (denoted as cell 1). If the load of cell 1 is larger than \bar{N} , we transfer services which occupy $N_1 - \bar{N}$ PRBs from cell 1 to the middle cell (denoted as cell 2). If the load of cell 1 is smaller than \bar{N} , we transfer services which occupy $\bar{N} - N_1$ PRBs from cell 2 to cell 1. At last, we balance the load between cell 2 and the cell in the right side (denoted as cell 3). If the load of cell 1 and cell 2 is larger than $2 * \bar{N}$, then we transfer services which occupy $N_1 + N_2 - 2 * \bar{N}$ from cell 2 to cell 3. If the load of cell 1 and cell 2 is smaller than $2 * \bar{N}$, then we transfer services which occupy $2 * \bar{N} - N_1 - N_2$ from cell 3 to cell 2. At last, the solution of (3) is described as follows:

$$P_{i,j} = \begin{cases} \max \left((4-j) \bar{N} - \sum_{x=j}^3 N_x, 0 \right) & i \in \{1, 2\}; j = i+1, \\ \max \left(j \bar{N} - \sum_{x=1}^j N_x, 0 \right) & i \in \{2, 3\}; j = i-1. \end{cases} \quad (4)$$

As a consequence of the objective function in (3), either $P_{i,j}$ or $P_{j,i}$ is zero and those two parameters are nonnegative, so we define other parameters $P'_{i,j}$ which can be negative:

$$P'_{i,j} = \begin{cases} P_{i,j} & (P_{i,j} > 0), \\ -P_{j,i} & (P_{i,j} = 0). \end{cases} \quad (5)$$

Property ($P'_{i,j} = -P'_{j,i}$). Combining formulas (4) and (5), the simultaneous equation of LB model is as follows:

$$N_1 - P'_{1,2} = \bar{N}, \quad (6)$$

$$N_2 - P'_{2,3} + P'_{1,2} = \bar{N}, \quad (7)$$

$$N_3 + P'_{2,3} = \bar{N}. \quad (8)$$

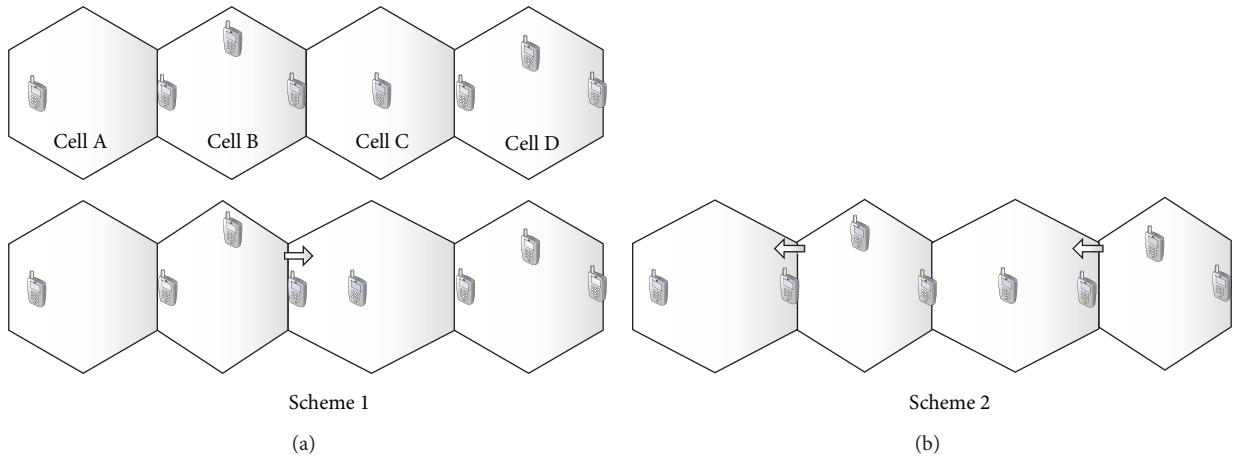


FIGURE 4: A LB scenario and two corresponding schemes.

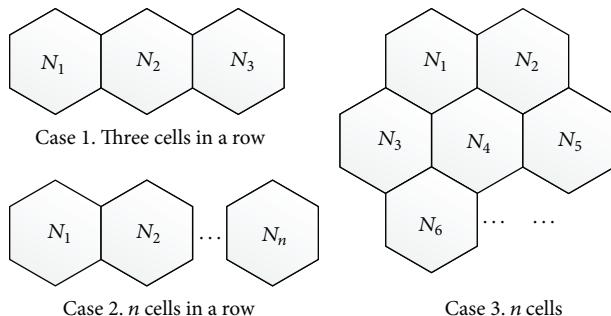


FIGURE 5: Network layout cases.

Formula (6) means that after transferring some service from cell 1 to cell 2 if $P'_{1,2} \geq 0$ (or from cell 2 to cell 1 if $P'_{1,2} < 0$), cell 1 has an average load level. Formulas (7) and (8) are the same meanings to cell 2 and cell 3. We sum up (6), (7), and (8) and obtain $N_1 + N_2 + N_3 = 3\bar{N}$ which is an identity. There are two unknown numbers in the above simultaneous equation, and the number of linearly independent equations is two; that is, (6) and (8) are linearly independent. So the solution (4) is the optimal solution because of the uniqueness of the solution.

5.1.2. Case 2: n Cells in a Row. In this case, using the same method in case 1, we can calculate the solution of (3), which is described as

$$P_{i,j} = \begin{cases} \max\left((n-j+1)\bar{N} - \sum_{x=j}^n N_x, 0\right) \\ \quad i \in \{1, 2, \dots, n-1\}; \quad j = i+1, \\ \max\left(j\bar{N} - \sum_{x=1}^j N_x, 0\right) \\ \quad i \in \{2, 3, \dots, n\}; \quad j = i-1. \end{cases} \quad (9)$$

Similarly to case 1, we can demonstrate the solution (9) is the optimal solution.

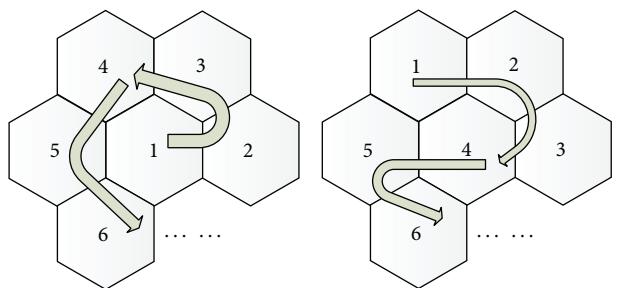


FIGURE 6: Two methods to mark n cells with numbers.

5.1.3. Case 3: n Cells. When there are n cells, we can mark them in accordance with the sequence in Figure 6. If only cells with adjacent sequence numbers can switch users, then case 3 is equal to case 2. So (9) is a solution to case 3. But it is not the optimal solution. In the remainder of this paper, we use network flow approach to obtain the optimal solution of the optimization problem.

5.2. LB Base on Network Flow Approach. In this section, firstly, we describe the optimization problem by graph theory. In graph theory, network flow has been rapidly expanding since the work of Ford and Fulkerson [18] on flow in 1962. The broad applicability in different systems of network flow optimization has brought great interest in it. A network flow is a directed graph composed of nodes and edges. Each edge receives a flow which cannot exceed the edge's capacity. Nodes are classified as three types: source, middle, and sink.

In imbalance traffic distribution scenario, some heavily loaded cells have more than \bar{N} PRBs occupied. Therefore, some users need to switch to adjacent low payload cells. We term those heavy payload cells as the source nodes while those low payload cells are termed as the sink nodes, in graph theory. The target is to transfer occupied PRBs in sink nodes exceeding \bar{N} to the sink nodes. Handoff from a cell to another ($i \rightarrow j$) is viewed as an arc. The arc is bidirectional

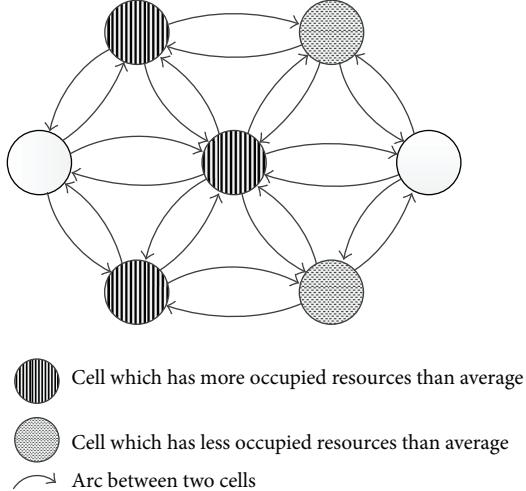


FIGURE 7: Seven nodes representing 7 cells.

because handoff between two cells is bidirectional, as shown in Figure 7.

The number of PRBs occupied by the users switched between two cells $P_{i,j}$ is compared as the amount of flow on the arc. The number of PRBs occupied by the users which can be switched between two cells is the capacity of the arc. The capacity of an arc ($i \rightarrow j$) is $\omega \cdot N_i$, where $0 < \omega < 1$. ω is used to indicate that only a part of users at the edge of a cell can be switched to adjacent cell. Then the LB is equivalent to the following optimization problem:

$$(P2) \quad \min \quad \left(\sum_{i=1}^n \sum_{j \neq i} P_{i,j} \right) \quad (1 \leq j \leq n)$$

$$\text{s.t.} \quad N_i - \sum_{j \neq i} P_{i,j} + \sum_{j \neq i} P_{j,i} = \bar{N}, \quad (10)$$

$$(1 \leq i \leq n, 1 \leq j \leq n)$$

$$P_{i,j} \leq \omega \cdot N_i$$

Now the optimization problem is a transportation network flow problem including multiple source multiple sink nodes. Each source node needs to transfer out services which occupy $N_k - \bar{N}$ PRBs and consequently each sink node needs to take over those services. We need to assign flow distribution in each arc of the graph. In the flow distribution, the difference of the amount of flow from a source node (i.e., cell i) to other nodes and the amount of flow from other nodes to the source node is equal to $N_i - \bar{N}$ so that the source node will have \bar{N} PRBs occupied by users after the handover process. For a sink node (i.e., cell j), the difference of the amount of flow from other nodes to the sink node and the amount of flow from the sink node to other nodes is equal to $\bar{N} - N_j$ so that the sink node will have \bar{N} PRBs occupied by users after the handover process.

A multiple source nodes and sink nodes problem is harder than a single source node and sink node problem. By using a virtual source node and a virtual sink node, the problem

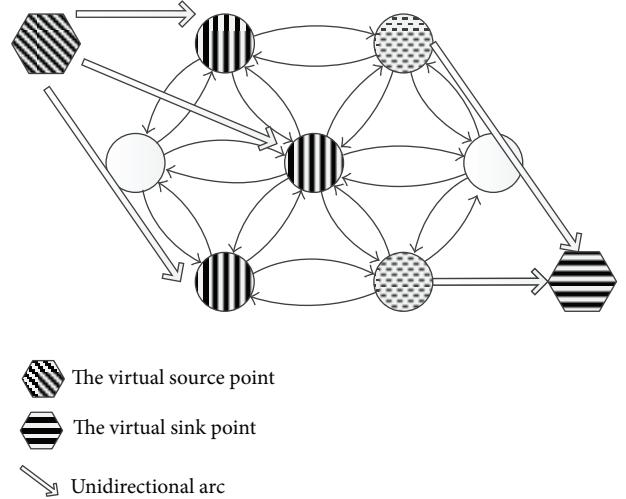


FIGURE 8: Arcs between the virtual source node and source nodes (sink nodes and the virtual sink node).

is transformed into a single source node and single sink node transportation network flow problem. In order to make the above two problems equivalent, a unidirectional arc is assumed from the virtual source node to each source node. The capacity of each unidirectional arc is the number of PRBs occupied in each heavily loaded cell minus \bar{N} . Similarly, a unidirectional arc is assumed from each sink node to the virtual sink node. The capacity of each unidirectional arc is \bar{N} minus the number of PRBs occupied in each low payload cell, as shown in Figure 8. If an arc between the virtual source node and a source node is saturated, then the source node will have \bar{N} PRBs occupied by users after the handover process, because the flow in the arc between the virtual source node and a source node is equal to the amount of flow from the source node to other nodes minus the amount of flow from other nodes (except the virtual source node) to the source node. Similarly, If an arc between a sink node and the virtual sink node is saturated, then the sink node will have \bar{N} PRBs occupied by users after the handover process. Now the multiple source nodes and sink nodes problem is equal to the single source node and single sink node problem.

If we give a flow distribution, of which the amount of flow out of the virtual source node is the amount of PRBs occupied by users in all source node subtract $l \cdot \bar{N}$ (i.e., arcs between the virtual source node and all source nodes are saturated), where l is the number of source nodes, then all nodes will have \bar{N} PRBs occupied after the handover process and the flow distribution is a solution to the optimization problem (10). Judging the existence of the solution to (10) is equal to judging if there is a flow distribution where arcs between the virtual source node and all source nodes are saturated. So we need to calculate the maximum flow between the virtual source node and the virtual sink node. If the maximum flow is equal to the sum of capacities of arcs between the virtual source node and all source nodes (i.e., the amount of PRBs occupied by users in all source node subtract $l \cdot \bar{N}$), solution to (10) exists. We can calculate network maximum

flow from the virtual source node to the virtual sink node by using algorithm of Ford and Fulkerson [18], which is to find augmented paths by using the idea of finding independent paths, and at least one arc will reach saturated state. The computation complexity of calculating the maximum flow by using independent paths method is $O(|E| * f)$, where E is the number of edges in the graph and f is the maximum flow in the graph. If aforementioned ω is close to 1 (i.e., a large part of users are located at cell edge and can be switched to adjacent cell), the maximum flow will be calculated as equal to the amount of PRBs occupied by users in all source node subtract $l \cdot \bar{N}$ and the solution to (10) will exist. So we assume that ω is close to 1 to ensure the solution to (10) will exist and to keep analysis simple. In future work, we will analyse some scenario where solution to (10) does not exist. After making sure the solution to (10) exists, we need to find a flow distribution which needs the least sum of flow between each two adjacent nodes (i.e., the least sum of handover between each two adjacent cells). To keep analysis simple, we assume that the cost for switching per PRB unit traffic is equal for all cells and users. Now, finding a flow distribution which needs the least sum of flow between each two adjacent nodes is a minimum cost flow problem. The optimal solution to the minimum cost flow problem can be figured out by using Orlin algorithm [19], and the computation complexity of Orlin algorithm is $O(m \log n(m + n \log n))$, where n is the number of nodes and m is the number of arcs. It is not complicated to calculate the maximum flow and the minimum cost flow because they all can be solved by polynomial algorithm. So we can use the optimal solution solved by Orlin algorithm to guide load balance.

5.3. A Novel LB Algorithm. In this section, a novel LB algorithm is presented based on the optimal solution. The novel LB algorithm is called optimal solution-based LB (OSLB). Firstly, the optimal solution to the numbers of PRBs occupied by communication service transfer between cells for load balancing will be figured out by the following procedure.

- (1) Each cell calculates the PRBs needed by each user in the cell through channel measurements and counts the total PRBs needed, which is denoted as N_i .
- (2) Each cell transmits the N_i to its ambient two layer cells, so that each cell knows its own N_i and its ambient two layer cells' N_j . Then, by the Orlin method [19], $P_{i,j}'$ can be calculated.
- (3) Using (5), $P_{i,j}'$ can be calculated. Then each cell transmits the $P_{i,j}'$ to its ambient one layer cells.
- (4) All cells average their own $P_{i,j}'$ and the $-P_{j,i}'$ received from their ambient cells, and the averaged values are defined as $P_{i,j}''$, which are the final amount of PRBs occupied by the users that should be transferred. At last, $P_{i,j}''$ are rounded down if they are not integers.

Secondly, we pick out cell i which has $P_{i,j}''$ greater than zero. Then we sequence the users in cell i according to the size of $D_n = \text{RSRP}_{n,j} - \text{RSRP}_{n,i}$. $\text{RSRP}_{n,i}$ is reference

signal received power (RSRP) between cell i and user n . The sequence number of user is indicated by M , and the PRBs occupied by user n are indicated by PRB_n . PRB_n vary among users because the modulation and coding mode is different among users with the base station.

Lastly, handover for load balancing in the above mentioned cell i will be implemented. According to [20], a handover event is initiated when user detects that a neighbor cell offers a better signal quality than its current serving cell. This condition is referred to as measurement event A3, which has been formulated as

$$M_s + \text{Oc}_{s,t} + \text{Hyst} < M_t, \quad (11)$$

where M_s and M_t are the signal strength or quality values for serving cell s and target cell t , and Hyst is cell-specific hysteresis value. $\text{Oc}_{s,t}$ is the specific offset for RSRP between cell s and cell t . As can be seen from (11), when $\text{Oc}_{s,t}$ is small, it is easy for users to migrate to the cell t rather than camp on the cell s . So the coverage of different cells is adjustable by changing the specific offset Oc among cells. In our LB algorithm, LB is performed by automatically adjusting $\text{Oc}_{s,t}$ based on $P_{i,j}''$. $\text{Oc}_{s,t}$ can not be very large because an unsuited value will cause some users to switch to an unsuited cell. Oc_{\max} is the upper limit of $|\text{Oc}_{s,t}|$. We define that $D'_n = D_n - \text{Hyst}$. The handover for load balancing of cell i is implemented by the procedures in Figure 9, where $\sum_{m=1}^M \text{PRB}_m$ is the number of PRBs required by M users who are switched from the heavily loaded cell to the low payload cell.

If the process comes to an end with $\sum_{m=1}^M \text{PRB}_m < P_{i,j}''$ by the reason of $-D_n < \text{Oc}_{\max}$, that is to say, there are not enough users at the edge of the heavily loaded cell and the target low payload cell, then the load imbalance problem is not completely solved. In this case, we search picocells at the edge of the heavy payload cell and switch those picocells and the users served by those picocells to the adjacent low payload cell to balance load.

6. Simulation and Performance Analysis

In this section, system-level simulation for the LTE-A cellular network is carried out to evaluate the performance of the proposed algorithm. There are 2 reference scenarios: no LB and the load-based MLB method as presented in [7]. The simulation platform contains 37 regular hexagonal cells, and the cell radius is 577 m. Some detail simulation parameters are presented in Table 2. In order to avoid boundary effects, wrap-around technique is applied. For simplicity, only one eNB is located in the cell center, and no sectors are divided. In each cell, there are two to four picocells located randomly at the cell edge. User numbers of 22 normal cells are 10. User numbers of other 15 busy cells are varying from 10 to 40. It is assumed that the traffic type of users is constant bit rate (CBR). The constant target date rate for each user is 1 Mbps. Detailed simulation assumptions and parameters are given in Table 1.

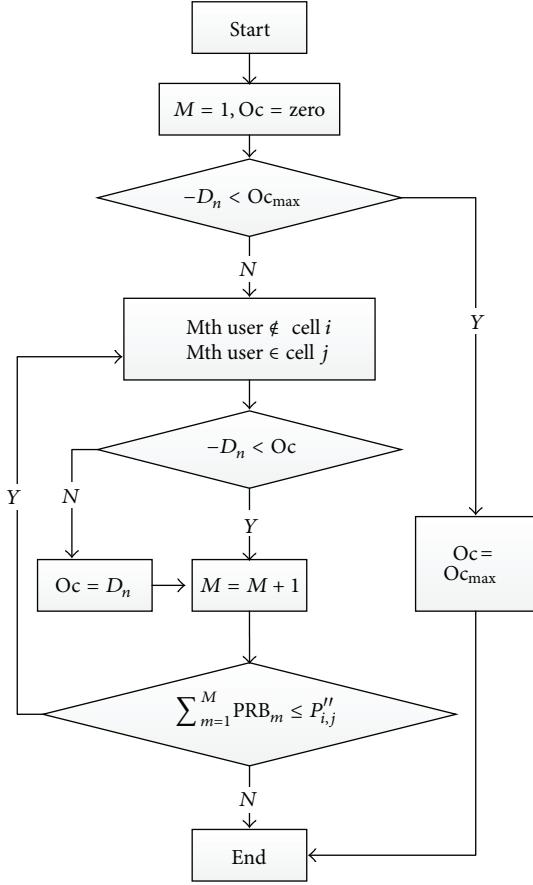


FIGURE 9: The load balancing flow chart.

TABLE 2: Simulation parameters.

Parameters	Assumptions
Carrier frequency	1.9 GHz
Bandwidth	40 MHz
Cell radius	577 m
Antenna type	Omnidirectional
Macrocell transmitter power	46 dBm
Picocell transmitter power	3 dBm
Channel model	SCM model [15]
Shadow fading (SF)	Log-normal
SF correlation distance	10 m
Noise	Thermal noise
Noise power spectral density	-174 dBm
Number of Tx antenna	1
Number of Rx antenna	2

Figure 10 shows the load distribution index (LDI), which is similar to Jain et al.'s fairness index [21] and is defined as follows:

$$\text{LDI} = \frac{(\sum_{i=1}^n N_i)^2}{|n| \sum_{i=1}^n (N_i)^2}. \quad (12)$$

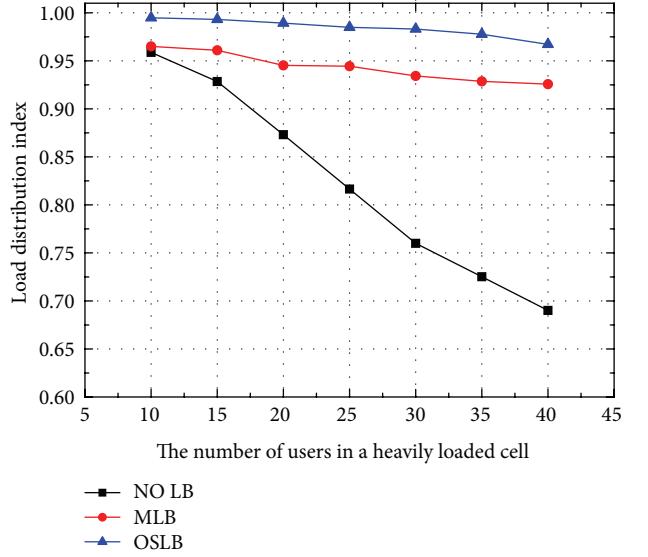


FIGURE 10: The load distribution index versus the number of users in a heavily loaded cell.

This index measures the degree of similarity of load among cells. The more closer the payload values are, the closer the index is to 1. When the call arrival rate of busy cells is increased, the load among cells is imbalanced. So this value is decreasing in the three lines of Figure 11. But the OSLB scheme achieves the highest load distribution index among three scenarios, so OSLB outperforms references in terms of LB result. When user numbers of heavy-payload cells are 10, PRBs needed of cells may vary because PRBs needed by users are different (some users in cell edge need more PRBs than users in cell centre). So the LDIs of MLB and NO LB are not 1 when all cells have 10 users. The LDI of OSLB is very close to 1 when all cells have 10 users.

The average resource occupied ratio of heavily loaded cells versus the user number of a heavily loaded cell is given in Figure 11. The less resource occupied ratio of heavily loaded cells is directly proportional to the lower service failure ratio of heavy payload cells. So this parameter should be as close as possible to the line with inverted triangle which means the average resource occupied ratio of all cells. Compared with the NO LB scenario, busy cells in OSLB have smaller resource occupied ratio. So in OSLB, the service failure ratio of busy cells is less than the NO LB one, since proper boundary users are switched to neighboring idle cells and more resources are reserved to remaining and new coming users. The resource occupied ratio of heavily loaded cell in the MLB scenario is similar to the OSLB one, which means that the two methods have similar effect on lightening the load of the busy cell.

As depicted in Figure 12, it is obvious that the proposed scheme outperforms the compared schemes in terms of the resource occupied ratio of a particular low payload cell, which accepts most users switched from heavily loaded cells and has the largest PRBs occupied after handover process among all old low payload cells. If the resource occupied ratios of some low payload cells after LB are high, new busy cells are

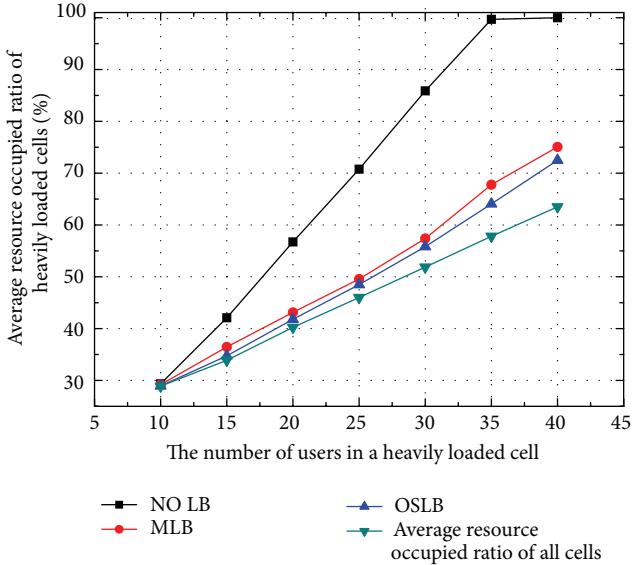


FIGURE 11: The average resource occupied ratio of heavily loaded cells versus the number of users in a heavily loaded cell.

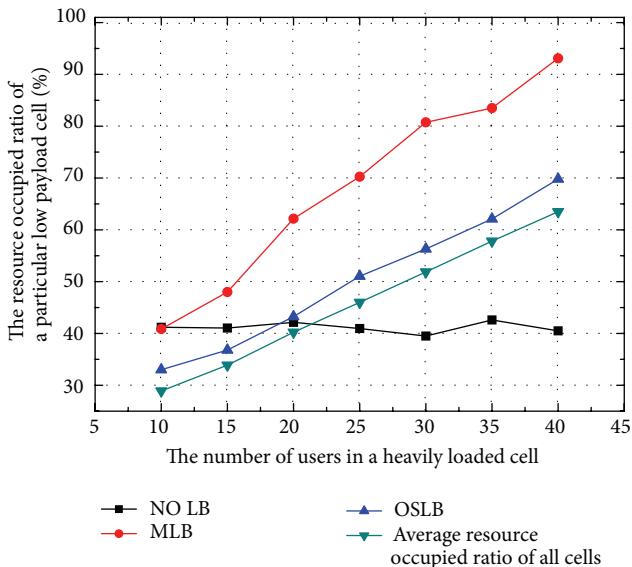


FIGURE 12: The resource occupied ratio of a particular low payload cell versus the number of users in a heavily loaded cell.

brought in. So the ratio in Figure 12 should be low and should be as close as possible to the line with inverted triangle which means the average resource occupied ratio of all cells. When LB is not used, this value does not vary and remains near 40%. When MLB method in [7] is used, this value increases and is larger than average resource occupied ratio of heavily loaded cells in the same scenario, which signifies that new heavily loaded cells are brought in. When OSLB is used, this value increases but is smaller than average resource occupied ratio of heavily loaded cells in the same scenario and is a little bigger than the average resource occupied ratio of all cells,

which signifies that new heavily loaded cells are not brought in.

7. Conclusion

In this paper, LB is optimized by carefully designing a novel algorithm named OSLB. Different from the previous literatures, the network flow approach is used to derive the optimal solution, which plays a large part in OSLB. The OSLB algorithm is evaluated and compared by system-level simulation. Results show that the load distribution index and the resource occupation ratio of cells are significantly improved. In this paper, we assume that each PRB of a cell can only be allocated to macrocell users or picousers; that is, those two kinds of users can not work on the same PRB simultaneously. In the future, we will investigate the scenario where a macrocell user and a picouser can share the same PRB, which is more realistic and more efficient.

Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

Acknowledgments

This work has been partially sponsored by EU FP7 IRSES MobileCloud Project (Grant no. 612212), the 111 Project (no. B08004), the major project of Ministry of Industry and Informationization of China (2013ZX03001026), and the Fundamental Research Funds for the Central Universities.

References

- [1] A. P. Vasudevan and R. Sudhakar, "A low complexity near-optimal MIMO antenna subset selection algorithm for capacity maximisation," *International Journal of Antennas and Propagation*, vol. 2013, Article ID 956756, 11 pages, 2013.
- [2] 3GPP TR 36.912 v10.0.0 Feasibility study for Further Advancements for E-UTRA (LTEAdvanced), 2011.
- [3] K. Zheng, F. Liu, W. Xiang, and X. Xin, "dynamic downlink aggregation carrier scheduling scheme for wireless networks communications," *IET*, vol. 8, no. 1, pp. 114–123, 2014.
- [4] M. Peng, D. Liang, Y. Wei et al., "Self-configuration and self-optimization in LTE-advanced heterogeneous networks," *IEEE Communications Magazine*, vol. 51, no. 5, pp. 36–45, 2013.
- [5] S. Das, H. Viswanathan, and G. Rittenhouse, "Dynamic load balancing through coordinated scheduling in packet data systems," in *Proceedings of the 22nd Annual Joint Conference on the IEEE Computer and Communications Societies (INFOCOM '03)*, pp. 786–796, San Francisco, Calif, USA, April 2003.
- [6] H. Zhang, X. S. Qiu, L. M. Meng, and X. D. Zhang, "Achieving distributed load balancing in self-organizing LTE radio access network with autonomic network management," in *Proceedings of the 25th IEEE Globecom Workshops*, pp. 454–459, Miami, Fla, USA, December 2010.
- [7] R. Kwan, R. Arnott, R. Paterson, R. Trivisonno, and M. Kubota, "On mobility load balancing for LTE systems," in *Proceedings of the IEEE 72nd Vehicular Technology Conference Fall (VTC-Fall '10)*, pp. 1–5, Ottawa, Canada, September 2010.

- [8] L. Zhang, Y. Liu, M. R. Zhang, S. C. Jia, and X. Y. Duan, "A Two-layer mobility load balancing in LTE self-organization networks," in *Proceedings of the International Conference on Communication Technology (ICCT '11)*, pp. 925–929, Jinan, China, September 2011.
- [9] T. Warabino, S. Kaneko, S. Nanba, and Y. Kishi, "Advanced load balancing in LTE/LTE-A cellular network," in *Proceedings of the 23rd IEEE International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC '12)*, pp. 530–535, Sydney, Australia, September 2012.
- [10] M. Sheng, Y. Chungang, Y. Zhang, and J. Li, "Zone-based load balancing in LTE self-optimizing networks: a game theoretic approach," *IEEE Transactions on Vehicular Technology*, 2013.
- [11] B. Li, C. Zhang, and X. Wang, "Multi-domain Load resource optimization for heterogeneous network in LTE-A," in *Proceedings of the 23rd IEEE International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC '12)*, pp. 215–219, Sydney, Australia, September 2012.
- [12] W. Bo, S. Yu, Z. Lv, and J. Wang, "A novel self-optimizing load balancing method based on ant colony in LTE network," in *Proceedings of the 8th International Conference on Wireless Communications, Networking and Mobile Computing (WiCOM '12)*, pp. 1–4, September 2012.
- [13] J. S. Kaufman, "Blocking in a shared resource environment," *IEEE Transactions on Communications*, vol. 29, no. 10, pp. 1474–1481, 1981.
- [14] K. Zheng, F. Liu, L. Lei, C. Lin, and Y. Jiang, "Stochastic performance analysis of a wireless finite-state Markov channel," *IEEE Transactions on Wireless Communications*, vol. 12, no. 2, pp. 782–793, 2013.
- [15] 3GPP TR 25. 996 v6. 1. 0, Spatial Channel Model for Multiple Input Mutiple Output (MIMO) Simulations, 2003.
- [16] A. Damnjanovic, J. Montojo, Y. Wei et al., "A survey on 3GPP heterogeneous networks," *IEEE Wireless Communications*, vol. 18, no. 3, pp. 10–21, 2011.
- [17] X. Zhang, X. Gu, W. Li, L. Zhang, J. Shen, and Y. Wan, "The study of indoor and field trials on MIMO architecture in TD-LTE network," *International Journal of Antennas and Propagation*, vol. 2013, Article ID 181579, 9 pages, 2013.
- [18] L. R. Ford Jr. and D. R. Fulkerson, *Flows in Networks*, Princeton University Press, Princeton, NJ, USA, 1962.
- [19] J. B. Orlin, "A faster strongly polynomial minimum cost flow Algorithm," in *Proceedings of the 20th ACM Symposium on the Theory of Computing*, pp. 377–387, 1988.
- [20] 3GPP TS 36. 331 v8. 4. 0, Evolved Universal Terrestrial Radio Access (E-UTRA), Radio Resource Control (RRC), Protocol Specification (Release 8), 2008.
- [21] R. Jain, D. Chiu, and W. Hawe, "A quantitative measure of fairness and discrimination for resource allocation in shared systems," Tech. Rep. DEC-TR-301, Digital Equipment Corporation, 1984.

