
©2010 IEEE. Personal use of this material is permitted. However, permission to reprint/republish this material for advertising or promotional purposes or for creating new collective works for resale or redistribution to servers or lists, or to reuse any copyrighted component of this works must be obtained from the IEEE.
A Game-Theoretic Approach to Load Balancing in Cellular Radio Networks

Ahmad Awada and Bernhard Wegmann
Nokia Siemens Networks
Munich, Germany
Emails: {ahmad.awada.ext; bernhard.wegmann}@nsn.com

Ingo Viering
Nomor Research GmbH
Munich, Germany
Email: viering@nomor.de

Anja Klein
Technische Universität Darmstadt
Communications Engineering Lab
Darmstadt, Germany
Email: a.klein@nt-uni-darmstadt.de

Abstract—Game theory provides an adequate methodology for analyzing topics in communication systems that include trade-offs such as the subject of load balancing. As a means of balancing the load in the network, users are handed over from heavily loaded cells to lower loaded neighbors increasing the capacity usage and the Quality of Service (QoS). The algorithm that calculates the amount of the load that each cell should decide either to accept or to offload might differ if the base stations are from distinct vendors, which in-turn may have an impact on the performance of the network. In this paper, we study the load balancing problem using a game-theoretic approach where, in the worst case, each cell decides independently on the amount of load that maximizes its payoff in an uncoordinated way and investigate whether the resulting Nash equilibrium would exhaust the gains achieved. Moreover, we alter the behavior of the players using the linear pricing technique to have a more desirable equilibrium. The simulation results for the Long Term Evolution (LTE) network have shown that the Nash equilibrium point can still provide a remarkable increase in the capacity when compared to a system without load balancing and has a slight degradation in performance with respect to the equilibrium achieved by linear pricing.

Index Terms—Game theory, Load balancing, Self Optimizing Networks, 3GPP LTE Release 9.

I. INTRODUCTION

Game theory provides a set of effective tools to understand the behavior of a system comprising players having conflicting objectives. Recently, it has been applied to study the problem of power control in coded division multiple access based systems [1], [2], [3], the random access to a shared channel [4], [5] and the link adaptation in cellular radio networks [6]. Indeed, it can be used to analyze other interesting topics in wireless communications such as load balancing which is one of the important self optimizing network use cases defined in [7].

In many circumstances, the cellular network may have unequal traffic distribution resulting in having some overloaded cells and other underloaded ones. The users connected to an overloaded cell experience a degradation in the QoS as the cell is unable to satisfy all its users due to its limited number of Physical Resource Blocks (PRBs). To increase the users’ satisfaction level in the network, the User Equipments (UEs) of the overloaded cells that are located at the cell boundaries are handed over to the lower loaded neighbors. Using load balancing, the number of unserved users in the overloaded cells decreases as the other underloaded neighbors are accommodating, if possible, the excess traffic load.

In [8], an algorithm is proposed that balances the traffic load in the network by adjusting the cell-pair specific handover offsets. The load balancing algorithm is applied to LTE network, but other systems may be similarly treated. The 3GPP LTE Release 9 provides the means to exchange the load information among the cells, but the algorithm that computes the amount of load that an overloaded cell should offload and an underloaded cell should accept is not specified as it is implementation specific of the eNodeB and, therefore, vendor-specific [7]. Thus, the game-theoretic approach seems to be an appealing method to understand the impact on the network operation if each cell decides, in the worst-case, autonomously on the load that maximizes its own benefit in a non-cooperative manner.

In this paper, we model the problem of selecting the load that an overloaded cell should offload and an underloaded cell should accept as a game where each player seeks to maximize his payoff. The resulting outcome of the game may not be necessarily the best for the overall network performance. In this case, the pricing technique is used to alter the behavior of the player to act more socially and achieve a better desirable equilibrium state. We also present some simulation results that show the overall performance of the cellular network in LTE downlink for the Nash equilibrium point, in addition to the equilibrium reached by linear pricing.

The paper is organized as follows. The load metrics are defined in section II, followed by a description of the load balancing algorithm in section III. The model of the game is presented in section IV, the strategies of the players leading to the Nash equilibrium are derived in section V and the linear pricing technique is described in section VI. The simulation results for LTE network are discussed in section VII and the work is concluded in section VIII.

II. METRIC DEFINITIONS

In this section, we define the metrics that are needed in the rest of the sequel. The load measurements that are processed by the load balancing algorithm are similar to those defined in [9]. In our system, each user $u$ has a constant data rate requirement denoted by $D_u$ and a data rate per PRB given...
by $R(\text{SINR}_u)$ which depends on the Signal to Interference Noise Ratio (SINR) of the user $u$. We could use Shannon’s capacity equation for the throughput function $R(\cdot)$, but we will follow a more realistic approach and use the abstract model presented in [10] which provides results that are close to link level simulations.

The number of required resources by the user $u$ can now be written as

$$Q_u = \frac{D_u}{R(\text{SINR}_u)} \quad (1)$$

We assume that each cell in the network has the same number of resource units that are available to be allocated. We denote the total number of PRBs per frame for all cells by $N_{\text{tot}}$ and define the traffic load of the user $u$ as

$$\kappa_u = \frac{D_u}{R(\text{SINR}_u)} \cdot N_{\text{tot}} \quad (2)$$

According to this definition, the load $\kappa_u$ of the user $u$ is interpreted as the percentage of occupied PRBs per frame needed to make him satisfied, i.e., meeting his data rate requirement.

The load of the cell $c$ is denoted by

$$\rho_c = \sum_{u \mid X(u) = c} \kappa_u \geq 0 \quad (3)$$

where $X(u) = c$ is the connection function that assigns the user $u$ to a single cell $c$. The load of the cell is a “virtual” metric as it can exceed 1, nevertheless, it reflects the level of overload and the QoS. The users connected to the cell $c$ would be all satisfied if the load $\rho_c$ does not exceed 1. Otherwise, there is some fraction of the users that are not satisfied, e.g., $\rho_c = 2$ means half of the users are unsatisfied.

The number of unsatisfied users in a network consisting of $M$ cells is expressed as

$$Z = \sum_{c=1}^{M} \max \left(0, \sum_{u \mid X(u) = c} 1 \cdot \left(1 - \frac{1}{\rho_c}\right)\right) \quad (4)$$

The $\max$ operator is needed as the number of unsatisfied users cannot be negative for the underloaded cell having a load less than 1. In reality, a user would not be admitted to the network if its data rate requirement could not be satisfied and in this work, we assume that the admission control arbitrarily selects the UEs to be served irrespective of their radio conditions.

III. THE LOAD BALANCING ALGORITHM

Having defined the metrics, the load balancing algorithm is briefly described to understand the rules of the game defined in section IV. The selection process of the users to be handed over to the neighbor cells is performed according to the algorithm discussed in [8]. The algorithm described works for any cellular layout, but for illustration purposes, we assume that the overloaded cell has 6 neighbor cells. Let us denote the load of the overloaded cell by $\rho_o$ and the load of each of the $N \leq 6$ neighbor underloaded cells by $\rho_i$ where $i \in \mathcal{S}$, $\mathcal{S} \subseteq \{1, \ldots, 6\}$ and $|\mathcal{S}| = N$.

![TABLE I: EXAMPLE OF A TABLE OF LINK IMBALANCES $\ell_{u,i}$ SET BY AN OVERLOADED CELL TO GENERATE THE HANDOVER CANDIDATE LIST](image)

The overloaded cell having a load $\rho_o > 1$ sets up a table as shown in Table I where the columns correspond to the $N = 4$ neighbor underloaded cells and the rows to its own connected UEs. In this example, the neighbor cells 2 and 5 are overloaded and excluded from the Neighbor Cell List (NCL), i.e., $\mathcal{S} = \{1, 3, 4, 6\}$. The value $y_i$ is the amount of load (expressed as a percentage of $N_{\text{tot}}$) that the underloaded cell $i$ can accommodate. Each entry $(u, i)$ of the table, i.e., $u \mid X(u) = \text{overloaded cell}$, is the link imbalance value defined by the difference in the Reference Signal Received Power (RSRP) levels as

$$\ell_{u,i} = \text{RSRP}_{u,o} - \text{RSRP}_{u,i} \quad (5)$$

where RSRP$_{u,o}$ and RSRP$_{u,i}$ are the average reference signal measurements reported by the user $u$ for the overloaded and underloaded cell $i$, respectively. In many cases, the UE is not able to report the RSRP of some neighbor cells, since the signal of the cell is either too weak (below noise) or drowned in the signal of the overloaded cell. In this case, the link imbalance is not available (N/A).

The overloaded cell can estimate, to some extent, the load that the user would produce if handed over to an overloaded cell. If the user $u$ having a local load $\kappa_u$ in the overloaded cell is handed over to a neighbor underloaded cell $i$, its load may decrease/increase if the signal strength to the connected cell is stronger/weaker. As a rough approximation, we assume that a positive link imbalance of $\ell_{u,i}$ dB decreases the average SINR of the user $u$ by $\ell_{u,i}$ dB which in turn reduces the throughput per PRB. The load that the user $u$ would produce if handed over to an underloaded cell is denoted by $\tilde{\kappa}_u$.

The overloaded cell generates the list of handover candidates as follows: It searches for the user having the smallest link imbalance value in the table and checks whether the corresponding target cell can accommodate its estimated load after the handover. If possible, the user is added to the handover list and the overloaded cell searches for the next candidate having the second best link imbalance value and so on. If an underloaded cell is unable to accommodate the load of the user, it is not considered anymore as a target cell for handover.

Having obtained the list of candidates, the overloaded cell executes successively the handover of every user and proposes a new value for the cell-pair specific handover offset to avoid the immediate back handover of the users [8].
IV. GAME MODELING

After stating the necessary metrics and the approach used by the overloaded cells in selecting the candidate users for handover, we model the load balancing game by defining the players, the utility function and the possible strategies.

The players of the game are on one side the overloaded cell having excess load and on the other side all its $N$ underloaded neighbor cells. The game proceeds in time where each underloaded cell $i$ signals first to the overloaded one the amount of traffic load $y_i$ that it is willing to accommodate. Having received all the accepted load values from the neighbor underloaded cells, the overloaded cell should decide on the load amount $\lambda'$ to offload based on the algorithm described in section III.

The underloaded cell $i$ can decide either not to accept any load, i.e., it signals $y_i = 0$, or to accommodate a certain amount of traffic load $y_i > 0$. Similarly, the overloaded cell can decide not to offload or to handover users having a certain load $\lambda' > 0$.

A useful utility function for the game is the number of satisfied users in the cell. One can argue that each player has the incentive to maximize his number of satisfied users as he would have more capacity usage and income resulting from the data rate charging. Let us denote the number of handed over users from the overloaded cell to the underloaded cell $i$ by $x_i \geq 0$. The total amount of load that the $x_i$ users would produce in the underloaded cell $i$ is $\sum_{j=1}^{x_i} \kappa_j$, where $\kappa_j$ is an approximation for the load of the $j^{th}$ handed over user.

The load of the accommodating cell $i$ after load balancing is denoted by $\hat{\rho}_i = \rho_i + \sum_{j=1}^{x_i} \kappa_j$. As long as $\hat{\rho}_i \leq 1$, all the users of the underloaded cell $i$ are satisfied as it has enough PRBs and the utility of the cell is simply the number of connected users. If $\hat{\rho}_i > 1$, not all its connected users are satisfied as the cell is incapable to meet the data rate requirement of every user. In this case, a floor operator is applied to the calculated number of satisfied users as it might not be an integer.

The utility function of the underloaded cell $i$ having a load $\rho_i$ and $U_i \geq 0$ connected users can now be expressed as

$$\text{utility}_i = \begin{cases} U_i + x_i & \text{if } 0 \leq \hat{\rho}_i \leq 1 \\ U_i + x_i & \frac{\rho_i + \sum_{j=1}^{x_i} \kappa_j}{\rho_i + \sum_{j=1}^{x_i} \kappa_j} \end{cases}$$

(6)

Based on the value $y_i$, one rule of the game that the overloaded cell will handover $x_i$ users having a total load $\sum_{j=1}^{x_i} \kappa_j \leq y_i$ to cell $i$, i.e., the overloaded cell is not allowed to handoff a load that exceeds the signaled capacity. A higher value of $y_i$ will most likely lead to a higher $x_i$, since the underloaded cell $i$ can accommodate in this case more users. Unfortunately, the number of handed over users $x_i$ and the estimated user loads $\kappa_j$ are unknown to the underloaded cell $i$ as the information of the link imbalances and the load of the users reside only in the overloaded cell, making the maximization of the utility function challenging.

Two instances of the utility function for the underloaded cell are depicted in Fig. 1. As long as the accepted load $y_i$ is less than $1 - \rho_i$, the utility function increases by $x_i$. When the accepted load $y_i$ exceeds $1 - \rho_i$, only a fraction of the users is satisfied as $\hat{\rho}_i$ would be greater than 1. For example, the underloaded cell $i$ having originally 6 users and $\rho_i = 0.45$ can increase its utility to 9 satisfied users by signaling $y_i = 1 - \rho_i$ and accommodating $x_i = 3$ users. If it signals a higher load than $1 - \rho_i$, e.g., $y_i = 0.74$, the underloaded cell $i$ would accommodate in this case $x_i = 4$ users and its utility would decrease from 9 to 8 satisfied users.

In this example, it happens that the utility function of the cell is maximum for a $y_i$ value which is less or equal to $1 - \rho_i$, however, it may also occur that the utility function peaks for a $y_i$ which is greater than $1 - \rho_i$, i.e., $\hat{\rho}_i > 1$. This is illustrated by the utility function of the other cell having $U_i = 5$ users and $\rho_i = 0.52$, where the number of satisfied users is maximal for $\hat{\rho}_i = 1.11$.

Two instances of the utility function of the underloaded cell are shown in Fig. 2. The cell having $U_o = 18$

![Fig. 1. Examples of utility functions of underloaded cells.](image)
users and $\rho_o = 1.17$ will seek to increase its utility from 15 to 16 satisfied users by offloading $X = 0.21$, i.e., decreasing its load from $\rho_o = 1.17$ to 0.96. Interestingly, it is not always beneficial for the overloaded cell to offload. As an example, the overloaded cell having $U_o = 9$ users would be better paid off if it does not handover any user. The utility of the overloaded cell would decrease from $[9/1.12] = 8$ satisfied users to $[8/1.05] = 7$ if it executes the handover of 1 user having a load $\kappa_u = 0.07$.

The value of the threshold strongly depends on the load of the underloaded cell $i$. If it is highly occupied, it is most likely that no user is handed over to the cell if it signals $y_i = 1 - \rho_i$ as the load of the user would probably not fit. In this case, the underloaded cell $i$ is indifferent and signals $y_{\text{max}}$ as it has a chance to increase its utility. Hence, the load signaled by the underloaded cell $i$ is summarized by

$$y_i = \begin{cases} 1 - \rho_i & \text{if } \rho_i < \rho_t \\ \frac{U_i + 1}{U_i} - \rho_i, & \text{otherwise} \end{cases}$$  \hspace{1cm} (9)$$

On the other hand, the overloaded cell should decide on the optimal load value, denoted by $X_{\text{opt}}$, to offload to the neighbor underloaded cells. Having received the available capacities from each neighbor, the overloaded cell sorts the users according to their link imbalances and keeps only those who fit in the target cell after calculating their estimated loads. As a result, the overloaded cell generates a list containing the users that are candidates for handover as explained in section III. For clarity, Fig. 3 shows an example of a set of handover candidate users with their corresponding target cells generated by the overloaded cell. The first two users shown on the left side can be handed over to the first neighbor cell, the third user to neighbor cell 6 and so on.

The overloaded cell can compute now all the utility values corresponding to the handover of the first $x_o$ users, e.g., utility$(1, 1, 6)$ is the payoff of the overloaded cell after executing the handover of the first 3 users. The overloaded cell selects the number of users $x_o = x_{o, \text{max}}$ that maximizes its payoff by comparing the utility values, without excluding the payoff if does not offload at all, and sets $X_{\text{opt}}$ to the sum of their respective local loads, i.e., $X_{\text{opt}} = \sum x_{o, \text{max}} n_j$. As a result, the overloaded cell will handover only the first $x_{o, \text{max}}$ users having a total load $X_{\text{opt}}$.

VI. LINEAR PRICING

In the Nash equilibrium, the player has no incentive to deviate from selecting the best strategy that maximizes its utility, which might not be necessarily advantageous for the overall network performance. Linear pricing is a powerful technique that can be used to adapt the action of each player, in favor of a better community-based achievement.

In our context, the utility function of the players is based on the number of satisfied users which does not correspondingly
consider the number of unsatisfied users in the network. From a system’s perspective, the capacity should be fully exploited, which in-turn, implies that the number of unsatisfied users should be minimized as much as possible. To emphasize that cells with unsatisfied users must be more “punished” than unloaded ones, we modify the utility function by introducing an additional pricing term. Thus, the variant utility function maximized by the players is defined as

$$\text{utility}^c = \text{utility} - \beta \max\{N_{\text{tot}} \cdot (\bar{\rho} - 1), 0\}$$ (10)

where $\beta$ is a tuning positive scalar and $\bar{\rho}$ is the load of the cell after load balancing, i.e., $\bar{\rho} = \hat{\rho}_i$ for the underloaded cell $i$ and $\bar{\rho} = \hat{\rho}_o$ for the overloaded cell. If $\beta$ is set to a high value, then each player is extremely penalized by a reduction factor proportional to his surplus load. In principle, $\beta$ should be adjusted to achieve the “natural” actions, where the overloaded cell seeks to offload all its excess load and the underloaded cell $i$ to accept users as long as its new load $\hat{\rho}_i$ does not exceed 1, i.e., it signals $y_i = 1 - \rho_i$. These load values, achieved by linear pricing, are the ones that would be recommended by a fully cooperative approach defining a sort of an upper bound of overall network performance.

VII. SIMULATIONS

In this section, we will evaluate the performance of the overall network for the Nash equilibrium in LTE downlink system. The parameters values are set according to the reference settings for LTE simulations defined in [12].

A. Layout and parameters

The cellular network is composed of $M = 57$ hexagonal cells separated from each other by 500 m and a wrap around is assumed. Every cell is served by one of the 3 sectors of a single eNodeB.

The maximum eNodeB transmission power is 40 W or equivalently 29 dBm per PRB, i.e., 10 MHz system with $N_{\text{tot}} = 50$ PRBs. The path loss offset and exponent are set to 128.1 dB and 3.76 dB respectively. The penetration loss is assumed to be 20 dB and the thermal noise power is $-114$ dBm. The standard deviation of shadowing is set to 8 dB and the decorrelation distance to 50 m. The transmit antenna has a beam width of $70^\circ$ and a backward attenuation of 20 dB.

The effect of the height of the base station and the antenna downtilt are not considered in the simulation, i.e., are set to 0. The handover hysteresis is 3 dB for all the cells and only the users having a link imbalance smaller than 5 dB are considered as candidates for handover. We also assume that every user has a constant bit rate of $D_u = 512$ kbps.

B. User positions

To demonstrate the effect of load balancing, we will generate heavy concentration of users in cells $\{4, 5, 6, 34, 35, 36, 43, 44, 45\}$ as shown in Fig. 4. The number of UEs dropped in the 9 hotspots is varied from 20 to 40 users. In the rest of the network, 192 UEs are randomly dropped. By following this distribution of UEs, we are creating many overloaded and underloaded cells and the load balancing game is played multiple times in the whole network. For every scenario, we average the number of unsatisfied users in the network over 20 different user drops.

C. Evaluation

The percentage of unsatisfied users in the network is shown in Fig. 5 as a function of the number of users in the hotspots for no load balancing, Nash equilibrium and pricing cases. The load threshold $\rho_t$ is set to 0.9 and $\beta$ to 4 which is high enough to achieve the “natural” actions.

According to the graph, a higher level of unsatisfaction is experienced in the network as the number of users increases...
in the hotspots. This is due to the fact that the underloaded cells are not able to accommodate high numbers of users when the overloaded cells are heavily occupied. The percentage of unsatisfied users in the network roughly increases by 25% for all the three cases when the number of the users in the hotspots doubles from 20 to 40.

Interestingly, there is a remarkable gain even if the cells behave selfishly in the load balancing game (bold dashed line). For 20 users dropped in the hotspots, the Nash equilibrium point achieves to accommodate most of the unsatisfied users in the network whereas 3.9% of the users are not satisfied if no load balancing is performed. Furthermore, the gain is visible for all the scenarios and the capacity usage can increase up to 6.7%.

If the players maximize the modified utility (solid line), the percentage of unsatisfied users slightly declines when compared to the Nash equilibrium curve without pricing. This certainly reflects that the “natural” behavior of the cells in load balancing relies, to some extent, on an inherent selfish actions.

VIII. Conclusion

We have presented a game-theoretic analysis for load balancing where we have defined the rules of the game, the players and their strategies. Also, we have modeled the utility function maximized by each player and defined the actions leading to the Nash equilibrium point.

The simulation results have shown that load balancing can remarkably increase the capacity usage in the network even when the cells act in a non-cooperative way. If the amount of load to accept or to offload is decided independently by each cell, we would expect that the attained Nash equilibrium point achieves most of the gain intended from load balancing. This indeed paves the way for the possibility of considering the deployment of different load balancing algorithms by various manufacturers as the loss in performance would be negligible.

REFERENCES