Load Balancing in Cellular Networks using Reinforcement Learning: A Real Experience over Operative Networks

I. Alepuz, J. Cabrejas, V. Osa*, J. F. Monserrat, J. López*, V. Soler*

* Instituto de Telecomunicaciones y Aplicaciones Multimedia,
** Universitat Politècnica de València,
6G Building - access D - Camino de Vera s/n - 46022 Valencia (Spain)

* Astellia, Paterna (Spain)

Corresponding author: iralbe@iteam.upv.es

Abstract

Cellular systems are continuously advancing towards the self-organization of their management, operation and maintenance, in what is known as Self-Organizing Networks (SON). In this sense, significant efforts are being made to standardize self-organization for which the 3rd Generation Partnership Project (3GPP) has included some use cases of particular relevance. It is within these use cases where load balancing especially stands, since it is crucial to relieve cell congestion and has an impact on network performance, mainly in interference-limited systems, such as Universal Mobile Telecommunication System (UMTS). This paper provides an overview of practical load balancing in UMTS, describing an actual implementation carried out by Astellia company and showing its impact on an operative network. The results show the enormous potential of self-organizing systems and encourage further research in this direction.

Keywords: Load balancing, self-organizing networks, SON, reinforcement learning.

1. Introduction

In current mobile communication networks, most of network parameters are manually controlled based on a centralized Operation, Administration and Maintenance (OAM) architecture. Traditionally, network optimization is performed based on drive testing. This is about performing a set of controlled trials, which implies moving a benchmarking group on a car while making phone calls and using mobile data connection and then analyse network performance. This approach had the advantage of having a user perspective since the data are collected directly by a user equipment. However, using only drive tests caused that the information available by the operator was completely biased, since the data were not obtained from real users, nor the entire network was covered and, furthermore, the information processing and analysis is slow and usually only manually made.

Due to these reasons, the OAM procedures changed to obtain a more realistic view of network status. New solutions are based on processing information collected by various network equipment, either as counters or even as network traces [1]. These solutions involve less investment in resources and the possibility to cover larger projects.

With these traces collected from the network, the next step forward was the inclusion of self-organization, in which is referred to as SON. SON features aim at automating, as much as possible, the processes of self-configuration, self-optimization and management as a way of reducing installation costs (CAPEX) and operating costs (OPEX). In this sense, SON functionality will involve providing mobile communication networks a level of “intelligence” much higher than current networks. In order to consolidate its usage, the 3GPP started the standardization of SON in Long Term Evolution (LTE) Release 8 and continued in Release 9 and beyond. Although this standardization has been mainly focused on LTE, its functions can be implemented in other systems such as UMTS. The overall process of auto-configuration (basic and radio) and self-optimization of a base station is included in the 3GPP specification TS 36.300 [2]. In addition, the specification 3GPP TS 36.902 [3] lists nine additional use cases.
This paper focuses on one of the above-mentioned algorithms, the MLBO. A significant number of papers in the literature have addressed the correction of congested cells, most of which have focused on improving coverage and capacity in general for Global System for Mobile communications (GSM) [4], Code Division Multiple Access (CDMA) [5] and High Speed Downlink Packet Access (HSDPA) [6]. The authors of [7] propose a scheme for tilt adjustment from a centralized server that has a comprehensive knowledge of all system settings. The problem is that the necessary signalling and data used for this decision is not discussed. On the other hand, in [8] and [9] this change of tilt, for CDMA and HSDPA respectively, is also studied. The tilt mechanism proposed is based, in both cases, on the basic idea of increasing tilt to reduce the coverage and distribute the load to the neighbouring cells. These forms of totally uncoordinated tilt settings are not only suboptimal for system performance, but also have shown ineffective and clearly dangerous since they can adversely affect the load of the cell to which traffic overflows. In [10], the author showed a method for optimizing the tilt based on measurements obtained from terminals. In this paper, an extended local search was carried out and each solution obtained from a cost function is accepted with a given probability. As the search continues, better solutions are accepted and the algorithm runs until no new improvements in the cost function are found. The results presented in the paper can be seen as self-organization, but are far from being implementable in a system operating in real time, and still do not solve the problem of efficient detection and quick reaction to a congested situation.

This paper presents the current version of the Astellia MLBO algorithm based on the use of standardized mobility parameters such as reselection/handover parameters, Common Pilot Channel (CPICH) or a hybrid solution of these parameters. This function distributes the load from congested cells towards neighbour cells with a lower congestion index. This allows that load of cells is balanced while increasing capacity of the system and under minimum human intervention in network management and optimization tasks. The solution presented in this paper is based on a Markov Decision Process (MDP), namely reinforcement learning-based algorithm. This algorithm, inspired by behaviourist psychology, models a decision-making process that chooses different alternatives with the objective of maximizing some cumulative reward. Note that the solution presented in this document only focuses on an intra-frequency and intra-Radio Access Technology (RAT) approach.

2. Load Balancing Algorithm

This section details the algorithm implementation. The MLBO algorithm is structured in five different parts, as can be seen in Figure 1: data loading, pre-processing, Call-trace UMTS Prediction (CUP), transition matrix construction and load balancing.

As shown afterwards, the load balancing algorithm is a MDP-based algorithm in which different actions are taken to maximize a cost function. In this case, the actions are taken based on a transition matrix known hereafter as $T$ matrix. This matrix is used to select the best action from the load balancing point of view.

First, in the data loading part, data are taken from topology databases and from the call trace information. Then, in the pre-processing part, data are prepared for being efficiently ingested by next phases. Third, the CUP solution is used to obtain the total Uplink/Downlink (U/DL) cell load as well as the load that each individual call added to the cell. Afterwards, in the transition matrix construction part, a $T$ matrix is filled for each cell with an estimation of the effect on cell load of a certain change in the radio conditions. Finally, the load balancing algorithm is executed. The details of the CUP algorithm and the procedures of the MDP are presented in Sections 2.2 and 2.3. Section 2.4 is dedicated to the load balancing execution. Before that, the detection procedure is detailed. It will identify which cells need to be adjusted by the MBLO algorithm.

![Figure 1. Overall description of the MLBO algorithm.](image-url)
2.1 Congestion detection

The main input of the MLBO algorithm is the cell congestion; to this aim, each cell is categorized for each temporal slot to different congestion states. The different congestion states sorted from lower to higher congestion are: Normal Operation (NO), Light Congestion (LiC), No Problematic Congestion (NPC) and Hazard Congestion (HaC). The main objective of the MLBO algorithm is to reduce the HaC cell states trying to balance the network load.

A Fuzzy algorithm has been developed with the aim of obtaining the different cell states. This algorithm uses Key Performance Indicators (KPI) obtained from traces during the studied temporal period. Each KPI is obtained per cell and per temporal slot. KPI such as Receive Total Wideband Power (RTWP), Total Carrier Power (TCP), Radio Resource Control (RRC) connection requests, RRC rejections, drop calls and input Handovers (HO) are used in the Fuzzy algorithm [11].

2.2 CUP

In order to prove whether the recommendations for each cell are optimal and solve the problem without being harmful to the overall network state, there is a necessity to know how the network evolves when a certain action is applied. This verification task is complex because it depends on many factors and these changes may have greater or lower impact depending on the network evolution. Therefore, changes should be gradual to see whether the influence is positive or negative, but also to be able to come back to the initial state when the event ends.

In order to estimate the level of aggressiveness of the changes, and also to isolate in an online phase the impact of changes, a network simulation tool is needed. With this tool, recommendations will be applied and the hypothetical network behavior will be evaluated. Up to date, the solutions found in the literature are based on radio predictions and the subsequent system simulation. However, with the advent of SON, an alternative is needed to take advantage of the available call-trace information.

This paper proposes a new approach to the problem that combines the available call-trace information with the mathematical UMTS system characterization using load concepts. This new solution, called CUP, will allow speeding up the estimation of a change effect while maintaining a realistic view thanks to the radio parameters characterization adjustment.

The system will start from a network snapshot, extracted from the call-trace received from the manufacturer. From this information, all the calls will be processed, obtaining for each one this information

- Serving cell.
- RACH-propagation delay of the serving cell.
- Call type (voice or data).

CUP calculates the evolution of the network behavior taking into account that users remain static. This approach allows working without the need for radio predictions that will be replaced with the user real measurements, ensuring the validity of the prediction. To that end, the channel effect will be calculated between the cell $i$ and the user $m$, $\gamma_{i,m}$ as

$$\gamma_{i,m} = S_{i,m} \cdot \frac{P_{C,i}}{P_{T,C}}. \quad (1)$$

This $\gamma_{i,m}$ value will be constant during all the prediction.

It is also important to know the consumed power of the cell $i$, this value will be obtained directly from the data base as

$$P_i = TCP[\%] \times P_{max,i}, \quad (2)$$

where $P_{max,i}$ is the maximum transmitted power of the cell $i$. If the $P_{max,i}$ value is not included in the topology information, it will be taken as 20 W.

It is necessary to calculate the DL load factor for each cell. The load system calculation is

$$\eta_{DL,i} = \frac{TCP[\%]}{100}. \quad (3)$$

The procedure starts with theoretical values of the load factor per service in the DL [12] [13]. Through this information, the consumed power of the cell $i$ to support the user $m$ is

$$P_{i,m} = \frac{\gamma_{i,m} \theta_{i,m} P_i + \sum_{j \neq i} \gamma_{j,m} P_j + P_N}{\gamma_{i,m} \theta_{i,m} \alpha_{m} + \frac{\gamma_{i,m}}{\rho_{m}}}, \quad (4)$$

where the $\theta_{i,m}$ is the orthogonality between the cell $i$ and the user $m$, $\alpha_{m}$ is the activity factor of the user $m$, and $\rho_{m}$ is the received Signal to Noise Ratio (SNR) of the user $m$. The noise power value, $P_N$, varies depending on the particular base station and its temperature. Indeed, the thermal noise level measured by each base station can be processed to take a real approximation. However, if this information is not available, a noise power of -102 dBm can be supposed.

This paper provides an overview of practical load balancing in UMTS, describing an actual implementation carried out by Astellia company and showing its impact on an operative network.

- In case it is a data call, mean throughput in UL and DL
- Call duration, to measure this call impact.
- Relation of the monitored cell in a Measurement Report (MR), with the mean Received Signal Code Power (RSCP) level. In case there are several MR with the same cell, the RSCP values will be averaged in a single value. For one cell $i$, the mean RSCP value received by a user $m$ will be noted as $S_{i,m}$.

Moreover, for all the topology cells the RTWP and TCP mean values will be extracted as well as the CPICH power of the cell $i$, $P_{C,F}$.

In order to estimate the level of aggressiveness of the changes, a network simulation tool is needed. This verification task is complex because it depends on many factors and these changes may have greater or lower impact depending on the network evolution. Therefore, changes should be gradual to see whether the influence is positive or negative, but also to be able to come back to the initial state when the event ends.

In order to estimate the level of aggressiveness of the changes, a network simulation tool is needed. With this tool, recommendations will be applied and the hypothetical network behavior will be evaluated. Up to date, the solutions found in the literature are based on radio predictions and the subsequent system simulation. However, with the advent of SON, an alternative is needed to take advantage of the available call-trace information.

This paper proposes a new approach to the problem that combines the available call-trace information with the mathematical UMTS system characterization using load concepts. This new solution, called CUP, will allow speeding up the estimation of a change effect while maintaining a realistic view thanks to the radio parameters characterization adjustment.

The system will start from a network snapshot, extracted from the call-trace received from the manufacturer. From this information, all the calls will be processed, obtaining for each one this information

- Serving cell.
- RACH-propagation delay of the serving cell.
- Call type (voice or data).
The power of the estimated cell is
\[
P_i = \sum_m P_{i,m} .
\] (5)

It is necessary that \( P_i = P_t \) and, thus, an adjustment of the factor \( p_m \) is used. In particular, the following constant is used
\[
\tau = \frac{P_i}{\sum_m P_{i,m}}.
\] (6)

Therefore, the target signal to noise ratio of each user is
\[
\rho_m = \frac{\tau^i_{i,m} \gamma_{i,m}}{\sum_j \gamma_{j,m} (p_{i,j} \tau P_{i,j})} + \gamma_{i,m} P_j + P_N.
\] (7)

### 2.3 Calculation of the transition matrix \( T \) for the Markov Decision Process

The \( T(s,a,s') \) is the probability of being in the \( s \) state and applying the action \( a \), passing to the \( s' \) state. At first, this matrix is filled using the CUP solution. For each \( s \) state, all actions \( a \) are applied and the \( T \) matrix is updated. These initial data are based on predictions and only when the load balancing is operating in closed loop \( T \) matrix will have real information.

The action \( a \) consists in decreasing the CPICH power, or reducing the range of the HaC cell through the Cell Individual Offset (CIO) parameter in active state or the Quality offset (Qoffset) for cell reselection in idle state. The modification of these parameters creates a load decanting between the HaC cell and its neighbour cells. The value of the corresponding radio signal offset is defined per cell relation. After applying such offset another cell could increase its load. Of course, this load cannot exceed a certain value, prefixed as input parameter.

The \( T \) matrix is firstly filled after the CUP stage, where for each slot and for all cells, all actions \( a \) are applied and the \( T \) matrix is updated. More specifically, the reselection/HO parameter consists in adding to all neighbour cells a virtual offset to the measured RSCP values of the User Equipment (UE) to decide the best serving cell. The offset values go from -10 to -0.5. The CPICH offset is subtracted to the CPICH of a congested cell to expel traffic to the neighbour cells. The CPICH values go from -10 to 50 dBm. Based on the specific parameter to be modified, the following actions can be executed:

- **CIO/Qoffset parameter:** Like with the CPICH parameter, a \( \Delta \)CPICH is subtracted to the received signal level but those congested cells will be blocked with the CIO/Qoffset. The CIO/Qoffset between the serving cell and the congested cell shall have an offset= \( \Delta \)CPICH.

As a last means, a reduction of resources to the data users can be applied. This action is named throughout the paper as data block action.

The above actions can be applied under these two conditions:

- When the number of congested cells after applying the action is lower than before applying the action but the load of each neighbour cell is lower than 100%.
- When the number of congested cells before and after applying the action is equal but the total cluster congestion has been reduced and the load of each neighbour cell is lower than 100%.

### 2.4 Load balancing decision

The goal of the load balancing algorithm is to decongest the HaC cells of the network without causing any other problem. The state of each cell is classified in each temporal slot (5 minutes, in this study). Therefore, the load balancing algorithm has to be executed each slot. For one slot, the block diagram is depicted in Figure 2.

![Figure 2. Block diagram load balancing function slot N.](image)

As it can be seen, an action is tried first. It is important to note that the \( T \) matrix will be updated continuously with an updating function. Then, clusters will be created. Finally, an action for each HaC cell will be found.

#### 2.4.1 Creation of clusters

In order to simplify the MLBO problem, it is important to limit the modifications to a certain cluster of cells that confine the congested cells allowing the absorption of the transferred load. Therefore, a cell clustering is necessary.

The involved parameters for the clustering are: the state of the cell, the congestion state in the neighbourhood of a cell, the distance to the problematic cell, the neighbouring rings of the problematic cell and the minimum congestion in a cluster.
Different clusters are obtained considering the above parameters, first, taking into account the cells with congestion problems (HaC cells) and, in turn, with the highest neighbourhood congestion state. Therefore, the clustering algorithm starts with the HaC cell whose neighbourhood is the most congested. After that, several cells are included with a maximum order, where the root order is defined as 1, its neighbourhood has order 2, the neighbours of neighbour cells have order 3, and so on. The maximum order is set to limit the size of the cluster and then allow for a more efficient execution of load balancing. Moreover, at least, all the root neighbour cells will be included, where the root is the first cell added to the cluster.

Next, those neighbour cells of nodes with order two and with a state HaC or NPC will be also included. For those neighbour cells of the order-2 LiC or NO state cells, the neighbour congestion will be studied and compared with a threshold. If the congestion is higher than this threshold, these cells will be also included in the cluster. It is important to emphasize that it is not possible to have a leaf node with a HaC state, where a leaf node is the cell with the highest order in the cluster. If this happens, the three lowest congested neighbour cells will be added to the cluster. Finally, it must be guaranteed that a cell does not belong to two different clusters. An example of clustering is shown in Figure 3.

2.4.2 Selection of the best actions

For each HaC cell of a cluster, the algorithm tries to find the best action that minimizes:

\[ F = \arg \min_a \sum_{i=1}^{N} T(s, a, s_i') \cdot s_i' \],

(10)

where \( s \in [0,9] \) is the current cell state and \( s_i' \in [0,9] \) is the final cell state after applying the action \( a \). For each cluster and for each HaC cell of the cluster, all actions are tested getting a sorted list of actions that minimize \( F \).

3. Results

In this section, simulation statistics and Keyhole Markup Language (KML) results of the MLBO algorithm are presented.

3.1 Statistics

The scenario under study is located in a European city where traces have been collected during 287 temporal slots, lasting each of them five minutes. There are 1149 HaC cells distributed among 96 slots, with an average of 12 HaC per slot. As aforementioned, there are three types of proposed actions that imply the modification of CIO/Qoffset, CPICH or CPICH+CIO/Qoffset. Table 1 shows the number of proposed actions, data block actions and cells solved without proposing any action. Note that a data block is proposed when it is not possible to choose a reselection/HO or CPICH action due to:

- The actions proposed are bad actions; an action is applied when some conditions are fulfilled as explained in Section 2.3.
- The offset increment, for each pair of cells, has already reached the maximum value in previous slots.
- \( T(s, a, s_i') \) is not filled.

Likewise, a cell is solved without an action when, due to the execution of the actions proposed in previous slots, the cell has reduced its congestion state. Therefore, this cell is not going to propose any more actions because no action is needed.

Table 2 shows the percentage of cells that reduce the congestion from HaC state and the percentage of cells whose state is changed to a lower congestion level. In this study, only the HaC cells with an action proposed related with reselection /HO and CPICH are taken into account.
account. That is, neither cells solved without action nor cells solved with data block are not used in the results. It is observed that the CPICH action is the one for which a higher number of cells see a congestion state reduction.

Table 3 shows, for each one of the proposed actions, the percentage of slots where congestion has been reduced as well as the mean congestion decrease of those slots. In this case, the CIO/Qoffset action outperforms the other actions at network level per slot. In particular, the averaged congestion is reduced about 12%.

3.2 KML Results

In Figure 4 and Figure 5, an example of the decongestion is presented. In Figure 4, the starting point of the load balancing test is shown, where no actions are applied, so a HaC cell is shown in red colour. The same temporal slot but with optimization actions applied, is represented in Figure 5, where the HaC cell has turned into NPC and the NO operation cell placed at the right side of the figures has turned into a LiC cell. Therefore, the main goal of the MLBO algorithm is accomplished since the network congestion is balanced.
5. Conclusions

This paper has presented the current version of the Astellia MLBO algorithm, which is based on the modification of reselection/handover and CPICH parameters. These actions distribute the load from congested cells towards neighbour cells with a lower congestion index. In particular, the goal is to balance the network load by decongesting HaC cells while cells in NO or LiC increase their loads. The scenario has been clustered in order to isolate the congestion problem to a group of cells in which the problem is solved. The impact of the proposed actions over the load of the different cells in a real network have been evaluated and compared being the state of the cells estimated through call traces collected from network elements.

It has been observed that the CPICH action is the most effective one, showing the highest reductions in the number of cells with congestion. However, from the global network congestion point of view, the CIO/Qoffset outperforms the other actions. The results highlight the enormous potential of self-organizing systems and encourage further research in this direction.

References


Biographies

Irene Alepuz obtained her MSc. in Telecommunication Engineering by Universitat Politècnica de València (UPV), Spain in 2010 by defending her Master Thesis through the Erasmus Grant at the Norwegian University of Science and Technology in Trondheim, Norway (NTNU). In 2011, she received a second MSc. in Telecommunication Technologies, Systems and Networks and she did her second Master Thesis at the Nanophotonic Technology Centre (NCT) at UPV. In 2012 she joined the Mobile Communication Group (MCG) in the Institute of Telecommunications and Multimedia Applications (iTEAM) from the UPV. At the iTEAM, her work is focused on R+D for optimization of mobile networks and location-based algorithms in collaboration with Astellia company.

Jorge Cabrejas received his MSc. in Telecommunications Engineering in 2008, a second MSc. in Telecommunication Technologies, Systems and Networks in 2009, and a Ph.D. in Telecommunications in 2016, all of them from the Universitat Politècnica de València (UPV). In 2007, he joined the Mobile Communication Group (MCG) in the Institute of Telecommunications and Multimedia Applications (iTEAM) from the UPV. In 2009, he took part in the best research project awarded to the MCG by the Valencia city council. In 2012, he was a guest researcher at the Department of Systems and Computer Engineering, Carleton University, Ottawa, Canada. He also participated in the METIS/METIS-II project on non-coherent techniques for the 5G technology. He is currently working on a collaboration project with the company Astellia about location-based algorithms. His research focuses on the study of new mobile communications techniques related to non-coherent communications.
Vicente Osa received his MSc. in Telecommunications Engineering in 2005 from Universitat Politècnica de València (UPV), Spain. He began his career as a software engineer in different companies. In 2008 he joined the Mobile Communication Group (MCG) in the Institute of Telecommunications and Multimedia Applications (iTEAM) from the UPV to manage the collaboration framework between UPV and Ingenia Telecom S.L., participating in different research projects at European level (ICARUS, PROSIMOS) and national/regional level (AUPA, ERGOLTE, OPTIFREL). In 2013, he received the Ph.D. in Telecommunications on Automatic Planning and Optimization of LTE Networks and joined Ingenia Telecom. Currently he is with Astellia S.A., where he is involved in funded research projects like SATIX and the design and prototyping of call-trace-based automatic optimization algorithms and Nova RAN Optimizer® features.

Jose F. Monserrat received his MSc. degree with High Honors and Ph.D. degree in Telecommunications Engineering from the Universitat Politècnica de València (UPV) in 2003 and 2007, respectively. He was the recipient of the First Regional Prize of Engineering Studies in 2003 for his outstanding student record receiving also the Best Thesis Prize from the UPV in 2008. In 2009 he was awarded with the best young researcher prize of Valencia. He is currently an associate professor in the Communications Department of the UPV. His current research focuses on the design of future 5G wireless systems and their performance assessment. He has been involved in several European Projects, being especially significant his participation in NEWCOM, PROSIMOS, WINNER+ and METIS/METIS-II where he led the simulation activities. He also participated in 2010 in one external evaluation group within ITU-R on the performance assessment of the candidates for the future family of standards IMT-Advanced. He co-edited two special issues in IEEE Communications Magazine on IMT-Advanced and 5G technologies and is co-editor of the Wiley book “Mobile and wireless communications for IMT-Advanced and beyond” and the Cambridge book “5G Mobile and Wireless Communications Technology”. Jose Monserrat is senior member of the IEEE, holds 6 patents and has published more than 40 journal papers. Currently the group headed by Prof. Jose F. Monserrat consists of 5 Postdoctoral fellows, 8 Ph.D. students and 2 Master students.

Javier López received his MSc. in Telecommunications Engineering from the Universitat Politècnica de València (UPV) in 2005. He defended his Master Thesis with Highest Honours through the Socrates-Erasmus Grant at Strathclyde University, Glasgow (UK). In 2005, he joined Accenture España S.L. as analyst programmer. In 2006, he joined the Mobile Communication Group (MCG) in the Institute of Telecommunications and Multimedia Applications (iTEAM) from the UPV, where he participated in R&D projects for Teltronic and Vossloh companies. In 2008, he joined Ingenia Telecom S.L., a leading company of Mobile Communications, as a partner. Here he carried out multitude functions of telecom analyst, developer, project manager among others. In 2014, he joined Astellia Spain, where he is currently developing his professional career as a Senior Innovation Manager.

Vicent Soler received his MSc. in Telecommunication Engineering from the Universitat Politècnica de València (UPV) in 2001. He obtained The National Prize for the Final Degree Project from the Universidad Politécnica de Madrid (UPM) in charge of Fondo de Cooperación Amena-Auna with the project titled ‘Sistema Móvil de Transmisión de Imágenes en Directo a Internet’. In 2001, he joined the Mobile Communication Group (MCG) in the Institute of Telecommunications and Multimedia Applications (iTEAM) from the UPV to manage the collaboration framework between UPV and Motorola Spain, where he developed multitude projects related to mobile communications such as UMTS, IMS and GSM networks. Moreover, in the following years he managed other collaborative agreements with companies such as Teltronic or Vossloh. Currently, he is the Technical Director of Ingenia Telecom-Astellia.