EUROPEAN COOPERATION
IN THE FIELD OF SCIENTIFIC
AND TECHNICAL RESEARCH

EURO-COST

COST2100 TD(10)10071
Athens, Greece
2010/Febr/03-05

SOURCE: Institut für Nachrichtentechnik,
Technische Universität Braunschweig,
Braunschweig, Germany

Load Balancing in Downlink LTE Self-Optimizing Networks

Andreas Lobinger, Nokia Siemens Networks, München, Germany
Szymon Stefanski, Nokia Siemens Networks, Wrocław, Poland
Thomas Jansen, Technische Universität Braunschweig, Braunschweig, Germany
Irina Balan, Interdisciplinary Institute for Broadband Technology, Ghent, Belgium

Email: andreas.lobinger@nsn.com, szymon.stefanski@nsn.com, jansen@ifn.ing.tu-bs.de, Irina.Balan@intec.ugent.be

Load Balancing in Downlink LTE Self-Optimizing Networks

Andreas Lobinger*, Szymon Stefanski†, Thomas Jansen‡, Irina Balan§

* Nokia Siemens Networks, München, Germany, andreas.lobinger@nsn.com

† Nokia Siemens Networks, Wrocław, Poland, szymon.stefanski@nsn.com

‡ Technische Universität Braunschweig, Braunschweig, Germany, jansen@ifn.ing.tu-bs.de

§ Interdisciplinary Institute for Broadband Technology, Ghent, Belgium, Irina.Balan@intec.ugent.be

***Abstract*—In this paper we present system level simulation results of a self-optimizing load balancing algorithm in a long-term-evolution (LTE) based mobile communication system. Based on previous work [2][1] ,we have evaluated the network performance of this algorithm, that has as input the virtual load of a cell and controls the handover parameters. We then compared the results for different simulation setups: a basic, regular network setup, a non-regular basestation grid with different cell sizes and a realistic scenario based on measurements and realistic traffic setup.**

I. Introduction

In current networks, the running operations and the parameters sets that are based on network planning, already enable a high level of performance. Getting additional performance on top of that, is a challenge. In the area of self-optimizing networks (SON), the use case of load-balancing (LB) tries to go that extra mile in terms of network performance by adjusting the network control parameters in such a way that overloaded cells can offload to low-loaded cells. In a live network, significant load fluctuations may occur and they can only be estimated as overprovisioning by network planning. A SON enabled network, where the proposed SON algorithm monitors the network and reacts to these changes in load, can dynamically achieve better performance [9]. The load balancing algorithm aims at finding the proper handover (HO) offset value between the cell in overload and a possible target where the load could be offloaded. This derived value will allow the maximum number of users to be successfully handed over to a neighbouring cell, thus diminishing the load in the current cell and also, signalling overhead. Simulations were conducted for regular hexagonal and non-regular synthetic and realistic scenarios.

The work has been carried out in the EU FP7 SOCRATES project[1].

II. Definitions and Metric

In [2] a mathematical framework for SON investigations on the downlink is defined.

The main parts for this mathematical framework are

- The definition of the network layout by network nodes -nodeB(s) (eNB)- defining cells c at the position $\vec{p}_c$.
- The users u positions $\vec{q}_u$. In dynamic simulation of course, the user positions can be changing over time.
- The definition of a connection function $X(u)$, following that a user u is served by cell $c = X(u)$ with the constraint (according to the definiton of LTE) that every user is connected exactly to a single cell.
- A pathloss mapping L, $L_{X(u)}(\vec{q}_u)$ defined by the positions of u relative to cell c, $c = X(u)$. The pathloss mapping takes all position dependend channel model effects into account, which are in the closed form distance dependend pathloss, angle dependend antenna patterns and a position dependend additional shadowing term.
- The definition of a load ρ_c, defining the ratio of used ressources -in LTE physical resource blocks (PRBs)- versus the available resources.

With that framework and two addtional terms i.e. N as thermal noise and P_c as transmit power for a cell, we can now define and evaluate for every user, in every time-step a user specific $SINR_u$.

$$SINR_u = \frac{P_{X(u)} \cdot L_{X(u)}(\vec{q}_u)}{N + \sum_{c \neq X(u)} \rho_c \cdot P_c \cdot L_c(\vec{q}_u)} \tag{1}$$

A. Virtual load, and unsatisfied users metric

Based on the long-term $SINR$ conditions of the users before and after load balancing and a given average data rate requirement D_u per user u, the throughput mapping $R(SINR_u)$ as a data rate per physical resource block (PRB) and given $SINR$ is calculated (e.g. based on the concept of a truncated Shannon-Gap mapping curve) and is related to the number of available PRBs M_{PRB}. The evaluation complexity is reduced by the selection of a single service definition for all users, a constant bit rate (CBR) service. The CBR assumption used further on is 512kBit/s.

This results in the virtual cell load that can be expressed as the sum of the required resources of all users u connected to cell c by connection function $X(u)$ which gives the serving cell c for user u.

$$\widehat{\rho}_c = \frac{1}{M_{PRB}} \cdot \sum_{x|X(u)=c} \frac{R(SINR_u)}{D_u} \tag{2}$$

The evaluation of virtual cell load takes an assumption about the actual scheduling policy into account, in our case the scheduling policy is a throughput-fair policy. With the scheduling policy and the virtual load for a cell we can derive a

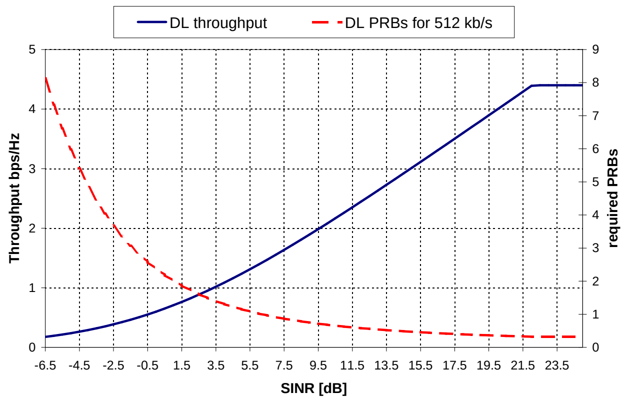

Fig. 1. Required PRBs for transmission of 512 kbps as a function of SINR in 10 MHz transmission bandwidth

metric for the overloaded cell defining a number of unsatisfied users i.e. users that can not achieve the target service bitrate (CBR).

Total number of unsatisfied users in the whole network (which is the sum of unsatisfied users per cell, where number of users in cell c is represented by M_c) can be written as:

$$z = \sum_{\forall c} max\left(0, M_c \cdot \left(1 - \frac{1}{\widehat{\rho}_c}\right)\right) \tag{3}$$

B. Load and throughput mapping

We assume that the best modulation coding scheme (MCS) is used for given SINR and the maximum of the theoretical throughput for given SINR $Thr(SINR)$ is represented by Shannon formula.

$$Thr(SINR) = \log_2 \cdot (1 + SINR) \tag{4}$$

For better approximation to realistic MCS, the mapping function is scaled and bounded by maximum available bitrate (4.4 bps/Hz) and minimum required SINR (-6.5 dB), detail description of which can be found in [4]. Base on the achievable throughput at given SINR, the necessary number of PRBs for the required throughput Thr_{req} and the transmission bandwidth BW can be obtained from the following equation:

$$N_{PRB} = \frac{Thr_{req}}{Thr(SINR) \cdot BW} \tag{5}$$

Figure 1 presents the relationship between SINR, throughput and required load in 10 MHz bandwidth DL transmission.

C. Load estimation

Load balancing is achieved by handing over users from the overloaded cells to cells able to accommodate additional load. After HO, these users may generate a different load in the new cel (TeNB) than in previous cell (SeNB), but this load should no exceed load reported as available by TeNB. The problem of limited resources in TeNB can be solved by admission and congestion control mechanisms. This solution, however, may increase the number of rejected LB HO requests and the required time to achieve best load distribution through the LB functionality and also unnecessary increase

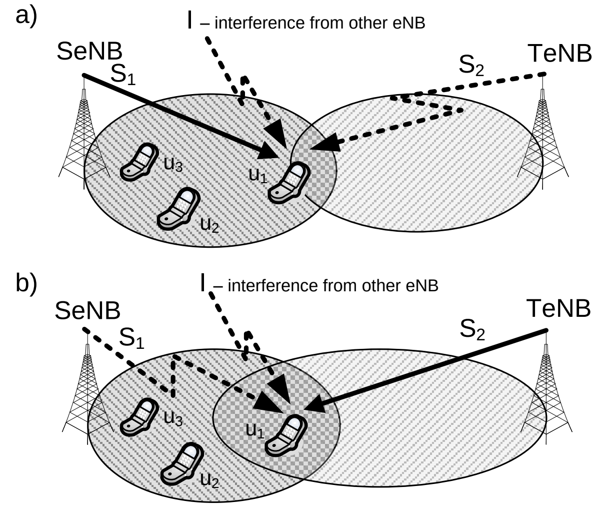

Fig. 2. Signals received by UE a) before HO, b) after HO

signalling overhead. We propose a prediction method for the load required at TeNB side, based on SINR estimation after LB HO, taking into account only UE measurements like RSRP and RSSI. For simplification we do not consider additional factors related with the current load at SeNB and TeNB which may have an impact on SINR. Before LB HO the user u_1 is connected to the SeNB, $SeNB = X(u_1)$ with the strongest received RSRP signal (signal S_1 in Figure 2 a). The RSRP signal from TeNB (S_2 in Figure 2a) is a component of total interference as well as signals originated from other eNBs (represented by I in Figure 2a). After HO user u_1 to TeNB, received signal S_1 from previously serving SeNB now belongs to the interference signal at SeNB and signal S_2 from TeNB start to be a serving one (see Figure 2b).

We assume that during the time of HO execution, the user's u_1 position $\vec{q}_{u_1}$ does not change and we can also assume no changes of received signal power by the user u_1. We can extract signals S_1 and S_2 from the interference part of $SINR_{SeNB}$ and $SINR_{TeNB}$ equations and after combaining them, this relation can be written as follows:

$$SINR_{u_1,TeNB} = \frac{S_2}{\frac{S_1}{SINR_{u_1,SeNB}} + S_1 - S_2} \tag{6}$$

III. Algorithm

We are assuming in our investigations modifying the cell specific Handover (HO) offset which may force group of users to hanover from serving eNB (SeNB) to a target eNB (TeNB). It is very important to derive proper HO offset value, as not exceeds the acceptable load at TeNB after LB HO procedure. The main goal of the presented algorithm 1 is to find optimum HO offset that allows the maximum number of users to change cell without any rejections by admission control mechanism at TeNB side. Before appluing the LB procedure, the SeNB needs to create list of potential targets for HO, collect measurements reports from the served UEs and the available resources reports from neighbouring cells. These preparation actions are included in steps 1 - 5 of algorithm 1. For each adjusted values of HO offset T, SeNB sorts list of the potential TeNB regarding to the number of possible LB HOs. Subsequently for given HO offset T and cell C from the list L load after HO $\widehat{\rho}_c$ is estimated. If predicted load does not exceed acceptable level ρ_{Thld}, HO offset to this cell is adjusted to the T value and virtual load at SeNB ρ_{SeNB} is reduced by the amount generated by the users handoverd with this offset. Algorithm works until load ρ_{SeNB} at SeNB is higher than accepted level $\rho_{Thld,SeNB}$ and HO offset is below the maximum alowed value.

Algorithm 1 HO offset based LB algorithm

Require: List L of potential Target eNB (TeNB) for LB HO

1. collect measurements from users; RSRP to potential TeNB is reported,
2. group users corresponding to the best TeNB for LB HO (criterion is the difference between SeNB and TeNB measured signal quality),
3. get information from TeNB on available resources,
4. estimate number of required PRBs after LB HO for each user in the LB HO group,
5. $T \Leftarrow 0$
6. **while** $\rho_{SeNB} > \rho_{Thld,SeNB} \wedge T < T_{max}$ **do**
7. $i \Leftarrow 1$
8. $T \Leftarrow T + step$
9. $L \Leftarrow sort$ (TeNB according to number of users allowed to LB HO with given T, descending order)
10.   **while** $\rho_{SeNB} > \rho_{Thld,SeNB} \wedge i \leq size(L)$ **do**
11. $C \Leftarrow L(i)$ {take next cell from list}
12.   estimate $\widehat{\rho}_c$ after HO for given T
13.   **if** $\widehat{\rho}_c < \rho_{Thld,C}$ **then**
14. $\rho_{SeNB} \Leftarrow \rho_{SeNB} - \rho_{SeNB,T}$ {update load in overloaded cell by substract handed over load}
15. $T_{C,u} \Leftarrow T$
16.   **end if**
17. $i \Leftarrow i + 1$
18.   **end while**
19. **end while**
20. adjust HO offsets T_C

IV. Simulation Scenarios

As already mentioned a standard LTE DL system of 10MHz bandwidth is simulated, following the simulation assumptions in the LTE 3GPP definitions [7]. For both the synthetic and real scenario a simulation time-step of 500ms have been used. So all internal signals, evaluations and updates are carried out with averaging over the 500ms step. Also the LB algorithm is called in every time-step.

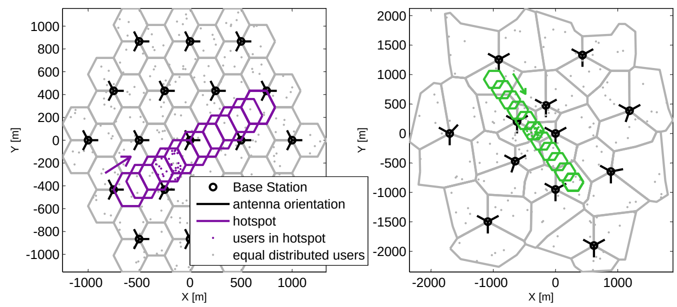

Fig. 3. Regular and non-regular simulation scenarios with hotspot path

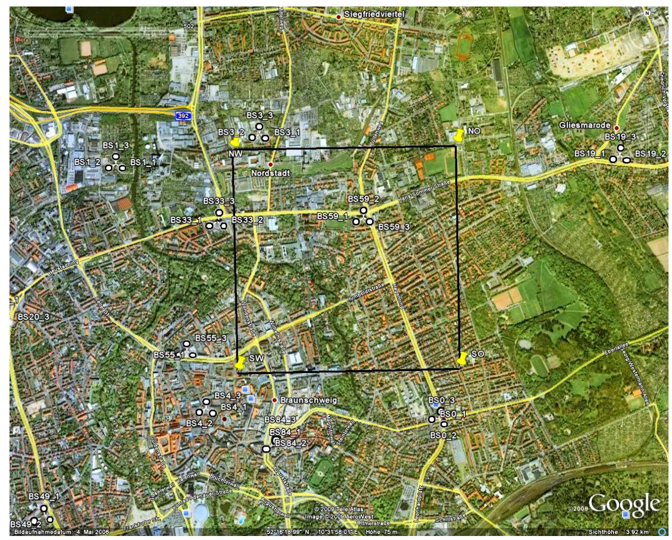

Fig. 4. Realistic Scenario, area with basestation positions

A. Synthetic scenario

Following the standard simulation assumptions a simulation setup with 19 sites in a regular (hexagonal) grid, 3 sectors per site and 57 cells have been defined. Additional to that a non-regular grid with 12 sites, 3 sectors per site and 36 cells is used as comparison taking real network effects like different cell sizes, number of neighbor cells and interference situations into account.

For employing localized higher load in the system a simulation scenario is used here with a setup of background load in all cells with a low number of users -so they are satisfied in any position of the network- and an additional hotspot in which new, additional users are dropped over time. The hotspot area is moved over time on a path through the network as depicted in figure 3.

All the channel model is defined in closed form (see standard [7]), so the pathloss mapping L is continuos. The movement of the users (whether background or dropped in the hotspot) is a constant velocity, random waypoint model.

B. Real scenario

The realistic reference scenarios in the SOCRATES project is an LTE network based on the real layout of the existing

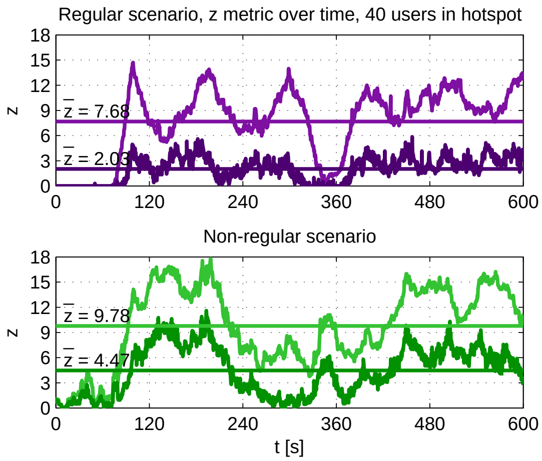

Fig. 5. Load Balancing performance over time, comparison

2G and 3G macro networks provided by a network operator. An area of 72 km x 37 km has been chosen. The resulting network comprising 103 sites and 309 cells is shown partly in 4. On this a area a detailed mobility model generated by SUMO simulator [10] is used to simulate user movement. In this scenario the moving hotspot is defined by a bus with a set of users moving together along a road.

Based on operator measurements and path prediction tools a pathloss mapping L is available for background users (static + moving) and the users in the hotspot (bus). For this pathloss mapping the 30 significant cells -for both connected cell and interfering cells- are taken into account.

V. Results

A. Synthetic Scenarios

In figure 5 a timeline of the z-metric for each simulation scenario with a reference system (light curve) versus load-balancing (dark curve) performance is depicted. The operating point used here is 5 users in background(equal) load and 40 users dropped in hotspot. The performance of the load-balancing is, as expected, dependend on the hotspot position in the network, and the users position within the hotspot and therefore changing over time. The averaged unsatisfied users metric $\bar{z}$ is depicted as horizontal line. In the following table, some more operating points and the average results as overview.

scenario	regular	non-reg.	regular	non-reg.
users in hotspot	$\bar{z}$ reference system		$\bar{z}$ load-balancing	
30	2.77	3.20	0.26	0.65
40	7.67	9.77	2.02	4.46
50	13.60	16.72	6.39	10.55
60	22.02	23.15	15.50	17.39

B. Real Scenario

In figure 6 a timeline of the z-metric for a real simulation scenario with a reference system versus load-balancing performance is depicted. As described in section IV-B the users in the hotspot move together on a bus along a street. Different than the operating point(s) choosen for the synthetic scenarios here a significant background load with 100 static and 100 slow moving users.

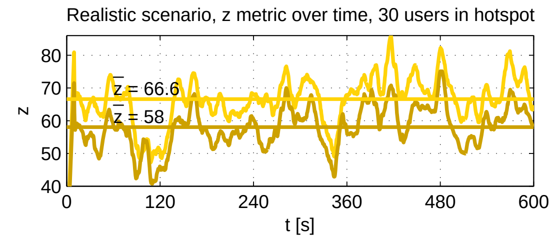

Fig. 6. Load Balancing performance over time

users in hotspot	$\bar{z}$ reference system	$\bar{z}$ load-balancing
20	57.7	49.3
30	66.5	58.0
40	75.7	67.3

The gain in the z metric (unsatisfied users) could be seen as low - in absolut numbers around 8 - but in relation of the CBR service of 512kBit/s this means a gain in network throughput of around 4Mbit/s.

VI. Conclusion and Outlook

The work described in [2][5] enables us to simulate detailed and realistic LTE network scenarios, in which the load situation changes dynamically. The proposed algorithm deals with the overload in a suitable way and reduces the overload significantly. The algorithm works on the measurements, information elements and control parameters defined in 3GPP for LTE Release9 [8] and would in this form enable a decentralized load-balancing, as well as a centralised solution. The work was carried out in both the SOCRATES project for the realistic scenario and as well as in 3GPP standardisation for LTE. Further on the work will be extended to take also uplink simulation into account and the integration of SON algorithms beyond load-balancing.

References

[1] SOCRATES, *Self-optimisation and self-configuration in wireless networks*, European Research Project, http://www.fp7-socrates.eu.

[2] I. Viering, M. Döttling, A. Lobinger, A mathematical perspective of self-optimizing wireless networks, *IEEE International Conference on Communications 2009 (ICC)*, Dresden, Germany, May 2009.

[3] 3GPP, *Self-configuring and self-optimizing network use cases and solutions*, Technical Report TR 36.902, available at http://www.3gpp.org.

[4] 3GPP, *3rd Generation Partnership Project;Technical Specification Group Radio Access Network;Evolved Universal Terrestrial Radio Access (E-UTRA);Radio Frequency (RF) system scenarios (Release 8)*, Technical Report TR 36.942, available at http://www.3gpp.org.

[5] M. Amirijoo, R. Litjens, K. Spaey, M. Döttling, T. Jansen, N. Scully, and U. Türke, "Use Cases, Requirements and Assessment Criteria for Future Self-Organising Radio Access Networks," *Proc. 3rd Intl. Workshop on Self-Organizing Systems*, IWSOS 08, Vienna, Austria, December 10-12, 2008.

[6] Next Generation Mobile Networks, *Use Cases related to Self Organising Network, Overall Description*, available at http://www.ngmn.org.

[7] 3GPP, *Physical Layer Aspects for evolved Universal Terrestrial Radio Access (E-UTRAN)*, Technical Report TR 25.814, available at http://www.3gpp.org.
[8] 3GPP, *Evolved Universal Terrestrial Radio Access Network (E-UTRAN); X2 Application Protocol (X2AP)*, Technical Specification TS 36.423, available at http://www.3gpp.org.
[9] U. Türke, and M. Koonert, "Advanced site configuration techniques for automatic UMTS radio network design," *Proc. Vehicular Technology conference VTC 2005 Spring*, vol. 3, pp. 1960-1964, Stockholm, Sweden, May 2005.
[10] H. Schumacher, M. Schack and T. Kürner, "Coupling of Simulators for the Investigation of Car-to-X Communication Aspects" 7th COST2100 Management Committee Meeting, TD(09)773, Braunschweig, Germany, February 2009.