Statistical Learning-based Automated Healing: Application to Mobility in 3G LTE Networks

Moazzam Islam Tiwana, Berna Sayrac, Zwi Altman Orange Labs, RESA/NET 38-40 Rue du General Leclerc 92794 Issy-Les-Moulineaux Cedex 9, France {moazzamislam.tiwana,berna.sayrac,zwi.altman}@orange-ftgroup.com

Tijani Chahed Institut Telecom SudParis; UMR CNRS SAMOVAR 9 rue C. Fourier 91011 Evry CEDEX, France tijani.chahed@it-sudparis.eu

Abstract—Troubleshooting of wireless networks is a challenging network management task. We have developed, in a previous work, a new troubleshooting methodology, which we named Statistical Learning Automated Healing (SLAH). This methodology uses statistical learning, in particular logistic regression, to extract the functional relationships between the noisy Key Performance Indicators (KPIs) and Radio Resource Management (RRM) parameters. These relationships are then processed by an optimization engine so as to calculate the optimized RRM parameters which improve the KPIs of a degraded cell. The process is iterative and converges to the optimum RRM parameter value in few iterations, which makes it suitable for wireless networks. The present work focuses on the adaptation of SLAH for troubleshooting the mobility parameter, namely the handover margin, in 3G Long Term Evolution (LTE) networks. The simulation results, which we obtain for a practical use case, show the advantage of this new, automated troubleshooting methodology.

Keywords: Automated troubleshooting, Statistical learning, Logistic regression, mobility, handover margin, 3G LTE.

I. INTRODUCTION

With the growth of mobile and wireless networks, the troubleshooting costs of the network are increasing, and the growing competition between operators calls for efficient utilization of the network resources and personnel. Troubleshooting may be manual, involving human experts that detect alarms and monitor the (statistical) performance of the network, or automated wherein these tasks are carried out by specialized modules in the Operation and Management Center (OMC).

The three main tasks of troubleshooting are: detecting the problem, identifying the cause (diagnosis) and solving the problem [1][2]. This paper deals with the third step of the troubleshooting process, i.e., problem resolution.

In fact, once the problem has been identified, there can be various ways which can be used to improve the performance of the cells that exhibit a degraded performance. The so-called *steered optimization technique* is one such method introduced in [3]. It uses an interference matrix approach to identify the most coupled eNodeBs (eNBs) with the problematic cell in a cellular (UMTS) environment. Indeed, the knowledge of the interaction between any couple of stations in terms of interference, macrodiversity and load difference allows to, first, accurately identify sectors with poor performance and, second, to suggest corrective measures [4]. The method makes

use of local information from an eNB and its neighbors; the resulting automated healing is thus a local optimization process.

Based on this interference matrix approach and local optimization technique, we have introduced in [5] a new troubleshooting methodology, which we termed Statistical Learning Automated Healing (SLAH), and which uses, as its name indicates, statistical learning to derive the functional relationships between the Radio Resource Management (RRM) parameters, for instance interference mitigation parameter, and the Key Performance Indicators (KPIs), such as Block Call Rate (BCR), file transfer time, etc.

Initially, this relationship between RRM and KPIs is not very accurate because the KPI measurements are inherently noisy. Noise can originate from limited measurement accuracy, but also from traffic fluctuations, varying propagation conditions, etc. The effect of the modification of the RRM parameter on KPIs can be masked partially by unobserved effects thus introducing uncertainty in the relationship between KPIs and RRM parameters. This model is processed by a socalled optimization engine to calculate the optimized value for the RRM parameter. The KPIs corresponding to this new optimized value of the RRM parameter are then calculated. The new and previous values of the RRM parameter, along with the corresponding KPIs values, are next used in the refined model extraction. The optimization process is iterative, and with each step, the accuracy of the model improves. The process is repeated until the calculated value of the RRM parameter converges to an optimum. The advantage of this troubleshooting methodology lies in the fact that it converges in 5 to 10 iterations.

The present work discusses the adaptation of SLAH for the troubleshooting of mobility in 3G LTE, in terms of the handover margin which determines the handover process from one eNB to an adjacent one. We assume that an eNB with degraded KPI performance due to a faulty RRM parameter setting has already been diagnosed. Our aim is to correct this parameter using local information from the degraded eNB and its first tier neighbours.

The remainder of this paper is organized as follows. In Section II, we describe the proposed troubleshooting scheme block diagram. In Section III, we present the details of the system model for mobility in 3G LTE networks. In Section IV, we explain the adaption of SLAH algorithm to the troubleshooting of the mobility parameter. In Section V, we show the 3G LTE simulation scenario and results. In Section VI, we eventually conclude the paper.

II. TROUBLESHOOTING BLOCK DIAGRAM

As indicated above, we assume that the fault cause has been diagnosed; our focus is only on the problem resolution phase. The SLAH block diagram is presented in Figure 1.



Figure 1. SLAH block diagram

As it can be seen, the system model is composed of four blocks:

- Initialization block: it provides the initial RRM parameter to the faulty eNB in the Network/Simulator block and to the Statistical Learning block
- Network / Simulator block: it represents the real network or the network simulator. It captures (case of real network) or calculates (case of network simulator) a set of KPIs of an eNB and of its neighbors for each new RRM parameter introduced by the Initialization or the Optimization block
- Statistical Learning block: Logistic Regression (LR) [6][7] is the statistical learning approach which we use to extract the functional relationships between the KPIs and the RRM parameter. This is possible as the KPIs are assumed to be well-behaved functions, i.e., not multimodal functions of the RRM parameter.

Specifically, LR fits the data into the functional form of a *logistic function* $f_{log}(z) = \frac{1}{1+e^{-z}}$. $f_{log}(z)$ describes the saturation effects at its extremities, as they are often encountered in KPIs in communication networks.

For $y_{m,i}$ denoting the i^{th} sample value of the m^{th} dependent variable y_m (i.e., the m^{th} KPI) corresponding to the i^{th} sample value x_i of the explanatory variable x (i.e., the RRM parameter), LR models $y_{m,i}$ as follows:

$$y_{m,i} = f_{log}(\eta_{m,i}) + \epsilon_i \tag{1}$$

where $\eta_{m,i} = \beta_{m,0} + \beta_{m,1}x_i$ is the linear predictor representing the contribution of the explanatory variable sample x_i , ϵ_i is the the residual error, and β s are the regression coefficients whose values have to be estimated using maximum likelihood estimation (MLE).

Hence, from Eqn. (1), the functional relationship between \hat{y} , i.e., y estimated by LR, and x can be written as:

$$\hat{y}_m(x) = f_{log}(\beta_{m,0} + \beta_{m,1}x)$$
 (2)

• Optimization block: it calculates the optimal RRM value using the current model. It determines \hat{x} , the value for an RRM parameter x that maximizes a utility function of a set of KPIs denoted as the *optimization set* A_o , subject to constraints on a second set of KPIs denoted as the *constraint set* A_c . Considering that $\hat{y}_m(x)$ has the functional relationship form given in Eqn. (2), the optimization problem can be formulated as:

$$\hat{x} = argmax_x U(x) \tag{3}$$

where $U = \sum_{m \in A_o} w_m \hat{y}_m(x)$ is the utility function and w_m is the weight given to $\hat{y}_m(x)$.

The automated healing process is iterative. At each iteration, a new RRM parameter value is proposed by the Initialization block during the initialization iterations and by the Optimization block during the optimization iterations to update the RRM setting of the faulty eNB in the Network/Simulator block. The performance of the faulty eNB and of its neighbors with this new RRM value is assessed by the Network/Simulator block through a set of KPI values obtained at the end of the measurement period, typically one day. And hence, a data point comprising an RRM parameter value and the corresponding KPIs is obtained. This data point together with the previously obtained data points are used by the Statistical Learning block to refine the model which is then used by the Optimization block to generate the RRM parameter value of the next iteration. And thus, as the iterations progress, on average, the model precision improves and is used by the Optimization block to find a better value for the RRM parameter.

III. SYSTEM MODEL FOR MOBILITY

We now consider user mobility between neighbouring cells. The 3G LTE mobility parameter considered in our study is the Handover Margin (HM). It is the main parameter that governs the Handover (HO) algorithm between two neighboring eNBs and refers to the minimum difference in power, between a neighboring cell and the current one, that is necessary for the mobile to make a HO.

The LTE standard has adopted hard HO wherein a mobile terminal will not be simultaneously connected to the current cell and the new cell [8]. Hard HO is implemented in LTE using an algorithm similar to the one used in GSM and is based on the comparison between the received signal strength from the serving cell and from the neighboring cells. It actually operates as follows.

While in communication, the mobile terminal periodically measures the received power from its serving eNB and from the neighboring eNBs. The mobile terminal, initially connected to eNB_i , triggers a HO to eNB_j if the following conditions are satisfied:

1) The Power Budget Quantity (PBQ) is higher than the HM: a mobile connected to eNB_i triggers a HO to an adjacent eNB_i if:

$$PBQ = P_j - P_i \ge HM_{ij} + Hysteresis \qquad (4)$$

where P_j is the received power from eNB_j expressed in dB and HM_{ij} is the outgoing HM of eNB_i towards eNB_j . Hysteresis is a constant independent of the eNBs and of the mobile stations, and is fixed here to 0

- The power being received from the target cell must be higher than a given threshold
- There are sufficient number of resources/chunks in the target cell. Otherwise, the mobile terminal is bounced back to the original cell

IV. ADAPTATION OF SLAH TO MOBILITY PARAMETER

We now consider the adaptation of the troubleshooting algorithm to the mobility parameter of the 3G LTE network by optimizing the HM parameter of the degraded eNBs. We assume that the cause of the degraded performance has been diagnosed and is related to a bad mobility parameter setting of the mobile terminals attached to the problematic eNB and its first tier neighboring eNBs.

Let us denote by eNB_c (c standing for *central*) an eNB with degraded performance. Let NS1 and NS2 denote the set of first and second tier neighbouring eNBs of eNB_c , respectively. Hence, eNB_j where $j \in NS1$ and eNB_t where $t \in NS2$, denote the first and second tier neighbours of eNB_c respectively.

To troubleshoot eNB_c , HM_{jc} , $j \in NS1$, are updated and optimized. In order to reinforce the effect of optimization of HM_{jc} on the KPIs, the HM_{jt} and HM_{cj} for the immediate geographical neighbours are calculated as a function of HM_{jc} , given as: $HM_{cj} = HM_{max} - HM_{jc}$ and $HM_{jt} = HM_{jc}$, where HM_{max} is the maximum HM value between any two eNBs.

We use the notion of coupling between eNB j and c which is expressed in terms of the interference that eNB_j produces on the mobile terminals connected to eNB_c and can be written in terms of the interference matrix element I_{cj} [3][4]. Hence the larger I_{cj} , the stronger the coupling between the two eNBs. In this work, the matrix element I_{cj} is equal to the sum of the interferences perceived by the mobile terminals attached to eNB_c and generated by the downlink transmissions to the mobile terminals of eNB_j .

The use of SLAH to jointly optimize all the elements of the vector (HM_{jc}) , $j \in NS1$, is not a simple task. Denote by s, $s \in NS1$, the index of the eNB which is the most coupled with eNB_c , namely $s = argmax_j(I_{cj})$, s and $j \in NS1$. To reduce the complexity of the SLAH process and to enhance its scalability, we propose to adjust the HM_{jc} parameter according to the degree of coupling between eNB_j and eNB_c .

Hence, we define a functional relationship between HM_{sc} and HM_{jc} , $HM_{jc} = g_j(HM_{sc})$, that accounts for the coupling terms I_{cs} and I_{cj} mentioned above, as follows:

$$HM_{jc} = g_j(HM_{sc}) = HM_{sc} + (1 - HM_{sc})(1 - \frac{I_{cj}}{I_{cs}})$$
(5)

By using Eqn. (5), just one parameter, HM_{sc} , needs to be optimized.

The process is scalable in the sense that the automated healing can be performed simultaneously on any number of eNBs provided they are not direct neighbours.

Two KPIs are used in the SLAH process: the Average Bit Rate (ABR) and the Block Call Rate (BCR) of the mobile terminals attached to an eNB. SLAH aims at maximizing the ABR for eNB_c and of its direct neighbours while verifying the constraints on BCR_j , $j \in c \cup NS1$.

We define the utility function used for optimization as:

$$U = ABR_c + \sum_{j \in NS1} \omega_j ABR_j \tag{6}$$

It is noted that ABR_j is a function of HM_{jc} and hence, via equation (5), of HM_{sc} . ABR_j also depends on the interference from its neighbouring eNB_j . The weighting coefficients ω_j depend on the relative contribution of I_{cj} with respect to the sum on all eNBs in NS1 and are given by:

$$\omega_j = \frac{I_{cj}}{\sum_{k \in NS1} I_{ck}} \tag{7}$$

satisfying the condition $\sum_{j \in NS1} \omega_j = 1$.

The optimization problem can now be formulated as follows:

$$HM_{sc} = argmax_{HM'_{sc}}U(HM'_{sc}) \tag{8}$$

subject to $BCR_j < BCR_{th}$; $j \in c \cup NS1$.

 BCR_{th} is the threshold for BCR_j . The ABR and BCR indicators in equations (6) and (8), are given in the form of a Logistic Regression (LR) function (Eqn. (2)) obtained using the Statistical learning block defined above.

Denote a data point p_k^j as the vector:

$$p_k^j = (HM_{jc}, ABR_j, BCR_j)_k \ ; \ j \in c \cup NS1$$
(9)

The set of k data points for an eNB_j , $j \in c \cup NS1$, is denoted by P_k^j . The SLAH algorithm is written below:

Initialization: 1. Identify the most coupled eNB s with eNB c among the neighbours in NS1 2. For each eNB_j , $j \in c \cup NS1$, compute an initial set of k data points P_k^j Repeat until convergence: 3. For each eNB_i , compute the statistical model using LR for ABR and BCR using the corresponding data points in P_k^j 4. Compute a new vector (HM_{jc}) , $j \in NS1$ (using equations (5) and (8)) 5. Apply (HM_{ic}) in the network/simulator and observe (ABR_i) and (BCR_i) , $j \in c \cup NS1$. Compute new data point p_{k+1}^{j} (equation (9)) 6. Update $P_{k+1}^j: P_{k+1}^j = P_k^j \cup p_{k+1}^j$ 7. k = k + 1End Repeat

The initial set P_k^j of data points in Step 2 is obtained by applying k (HM_{jc}) vectors to the network/simulator one by one and obtaining the corresponding KPIs.

V. SIMULATIONS AND RESULTS

A. Simulation Scenario

We consider a 3G LTE network composed of 45 eNBs in a dense urban environment, as depicted in Figure 2.



Figure 2. The network diagram of the simulated system

We consider downlink transmissions. The simulation parameters are as listed in Table I. A MATLAB-based LTE simulator described in [9] has been used.

The simulator performs Monte Carlo correlated snapshots of the simulated network in order to account for the time evolution of the traffic. At the end of each time step of one second, new mobile terminal positions are updated, HO events are processed, new users are admitted to the system according to the conditions of access and others leave the network, either after the completion of their service or because they were dropped.

Parameters	Settings			
System bandwidth	5MHz			
Cell layout	45 eNBs, single sector			
Maximum eNB transmit power	32 dBm			
Inter-site distance	1.5 to 2 KM			
Subcarrier spacing	15 kHz			
PRBs per eNB	21			
Path loss	L=128.1 + 37.6 $\log_{10}(R)$, R in kilo-			
	meters			
Thermal noise density	-173 dBm/Hz			
Shadowing standard deviation	6 dB			
Traffic model	FTP			
File size	5700 Kbits			
PRBs assigned per mobile	1 to 3 (First-come, first-serve basis)			
Mobility of mobiles	90%			
Mobile speed	15 m/s			
HM_{max}	12dB			
Table I				

SYSTEM LEVEL SIMULATION PARAMETERS

The simulations are run for 3300 time steps, with a fixed HM value, and the KPIs are averaged using the interval between 500 and 3300 seconds to account for transient effects. The interference matrix elements used in equations (5) and (7) are calculated once for the reference solution, described next, during a longer time interval varying from 500 to 7000 seconds to achieve higher accuracy.

An optimal default value for HM is chosen as 6dB for all eNBs in the network and serves as the reference (default) solution. This reference solution is used as a starting point for the automated healing process. The default HM value is determined in [10] by varying it uniformly from 0.05 to 12 in steps of 0.15 for all the eNBs. For each HM value, the network performance is assessed in terms of the mean Ping Pong Rate (PPR), mean BCR and mean Drop Call Rate (DCR). If these three KPIs are aggregated, the global optimum HM value occurs around HM = 6dB. Hence, the value of HM = 6dB is selected as the reference (default) HM value for eNBs in the network.

B. Troubleshooting Scenario

We now select, for the sake of illustration, $eNB_{c=13}$ in the simulated network of Figure 2 as the problematic eNB. The eNB_j , where $j \in NS1 = \{14, 15, 22, 23, 43, 45\}$, are the six first tier geographical neighbours of $eNB_{c=13}$. The set NS2 of the second tier neighbours of the problematic eNB consists of eNB_1 , eNB_{10} , eNB_{11} , eNB_{16} , eNB_{18} , eNB_{24} , eNB_{37} and eNB_{44} .

Denote by the optimization zone the subnetwork comprising $eNB_{c=13}$ and its first tier NS1, and by the evaluation zone the subnetwork comprising the $eNB_{c=13}$ and its first two tiers NS1 and NS2. $eNB_{s=45}$ is the eNB most coupled with $eNB_{c=13}$.

C. Results

The first five values of $HM_{sc,s=45}$ in Table II are chosen in the initialization phase of SLAH. The next seven values (Phase-II in Table II) are calculated iteratively by the SLAH algorithm. The values of $HM_{jc,j=14}$, $HM_{jc,j=15}$, $HM_{jc,j=22}$,

	HM_{jc}						
Phase I	j = 14	j = 15	j = 22	j = 23	j = 43	j = s = 45	
	11.7	11.5	11.8	11.9	11.6	11.4	
	10.5	9	10.9	11.6	9.9	8.7	
	9.2	6.8	9.9	11.3	8.2	6	
	8	4.4	9	11	6.4	3.3	
	6.7	2	8	10.7	4.8	0.6	
п	9.2	6.8	9.8	11.3	8.0	5.7	
	9	6.4	9.8	11.3	7.9	5.6	
	8.9	6.2	9.7	11.2	7.8	5.4	
ase	8.9	6.2	9.7	11.2	7.7	5.3	
Ph	9	6.4	9.8	11.2	7.9	5.6	
	8.9	6.2	9.7	11.2	7.8	5.4	
	8.9	6.2	9.7	11.2	7.8	5.4	

Table II Phase-I shows the initially chosen HM_{jc} values. Phase-II shows the HM_{jc} values calculated during optimization.

 $HM_{jc,j=23}$ and $HM_{j=43}$ are calculated using equation (5). In spite of the noisy data, one can see from Phase-II that $HM_{sc,s=45}$ converges in a few iterations. $HM_{sc,s=45} = 5.4$ is chosen as the optimized solution for the next Figures.



Figure 3. Mean KPI values and LR regression curves as a function of HM_{sc} for $eNB_{j=13}$, $eNB_{j=22}$ and $eNB_{j=43}$ (a) mean BCR (b) mean ABR

Figure 3(a) shows the mean BCR (points as well as extracted LR curves) for $eNB_{c=13}$, $eNB_{j=22}$, $eNB_{j=23}$ and

 $eNB_{j=43}$ as a function of $HM_{sc,s=45}$ after convergence. For the sake of clarity, the mean BCR curves for $eNB_{j=14}$, $eNB_{j=15}$ and $eNB_{j=45}$ are not shown; they are below the maximum allowable mean BCR threshold of 5%.

The concentration of mean BCR data points around $HM_{sc,s=45} = 5.4$ indicates the convergence of the SLAH algorithm.

The mean ABR along with its LR curves for $eNB_{j=13}$, $eNB_{j=15}$, $eNB_{j=43}$ and $eNB_{s=45}$, after convergence, are shown as a function of $HM_{sc,s=45}$ in Figure 3(b).

The mean ABR curves for $eNB_{j=14}$, $eNB_{j=22}$ and $eNB_{j=23}$ are, again, not shown for the sake of clarity; they show a similar trend.



Figure 4. KPI of the eNBs in the optimization zone for the reference solution (white) and optimized (black) network conditions; mean BCR (a) and mean ABR (b)

Figure 4(a) shows the gain brought about by the SLAH algorithm for the optimization zone (set NS1 of eNBs). The mean BCR of the problematic $eNB_{c=13}$ is reduced below the threshold of 5% to 4.73% with respect to the mean BCR value of the reference solution of 9.52%. Similarly, after optimization, the mean BCR of each eNBs in NS1 is below the 5% threshold.

As of the mean ABR, the improvement obtained by the SLAH algorithm in the optimized zone with respect to the



Figure 5. KPIs in descending order for the eNBs in the evaluation zone, mean BCR (a) and mean ABR (b)

reference solution is shown in Figure 4(b). The mean ABR of $eNB_{c=13}$ is increased by 2.79% whereas the average mean ABR of NS1 is increased by 2.51%.

Figures 5(a) and 5(b) show, in a descending order, the mean BCR and the mean ABR, respectively, for the reference (square) and the optimized (circle) eNBs in the evaluation zone. It is noted that the order of the stations in the two curves of each figure may not be preserved. One can see that, on average, the mean BCR and the mean ABR in the evaluation zone are improved. The average improvement of the mean ABR in the evaluation zone is equal to 3.51%.

VI. CONCLUSION

We presented, in this paper, the adaptation of our SLAH troubleshooting algorithm to troubleshoot the HM mobility parameter in 3G LTE networks. Statistical learning is used to model the functional relationship between KPIs and RRM parameter. This model information is processed by the statistical learning block in order to calculate the optimum RRM parameter to improve the performance of the degraded eNB. A case study of eUTRAN HO algorithm has been considered; it shows that our model improves after each iteration and that the RRM parameter converges to the optimal value in a limited number of iterations.

REFERENCES

- T. Halonen, J. Romero, J (eds). Melero: GSM, GPRS and EDGE Performance. Wiley: Chichester, 2003
- [2] R. Barco, V. Wille, L. Diez: System for automatic diagnosis in cellular networks based on performance indicators, European Trans. Telecommunications, vol. 16, no. 5, Oct. 2005, pp. 399-409.
- [3] S. Ben Jamaa, Z. Altman, J.-M. Picard, A. Ortega: Steered Optimization Strategy for Automatic Cell Planning of UMTS Networks, IEEE 61st Vehicular Technology Conference, 2005 Spring.
- [4] S. Ben Jamaa, Z. Altman, A. Ortega, B. Fourestie: UMTS design strategies based on indicator matrix approach, IEEE Int. Conf. on Communications, ICC 2004, Paris, France, June 20-24, 2004.
- [5] M. I. Tiwana, B. Sayrac, and Z. Altman: "Statistical approach for automated troubleshooting: application to LTE interference mitigation", submitted to IEEE Trans. on Vehicular Technology.
- [6] D.G. Kleinbaum, M. Klein,: Logistic Regression: A Self-Learning Text, 2nd Edition, Springer-Verlag New York, 2002.
- [7] W.D. Hosmer, S. Lemeshow,: Applied Logistic Regression, 2nd Edition, Chichester, Wiley, New York, 2000, pp. 1-30.
- [8] R. Nasri, Z. Altman and H. Dubreil, WCDMA downlink load sharing with dynamic control of soft handover parameters, IEEE International Symposium VTC, Melbourne, Australia, 7-10 May, 2006.
- [9] R. Nasri, Z. Altman: Handover Adaptation for Dynamic Load Balancing in 3GPP Long Term Evolution Systems., MoMM'2007-The Fifth International Conference on Advances in Mobile Computing and Multimedia, 3-5 December 2007, pp. 145-154.
- [10] M. I. Tiwana, B. Sayrac, Z. Altman, T. Chahed: Troubleshooting of 3G LTE mobility parameters using iterative statistical model refinement, IFIP Wireless Days 2009, Paris, Dec. 2009.