SON Conflict Diagnosis in Heterogeneous Networks

Ovidiu Iacoboaiea, Berna Sayrac, Sana Ben Jemaa
Orange Labs
38-40 rue du General Leclerc 92130
Issy les Moulineaux, France
{ovidiu.iacoboaiea,berna.sayrac, sana.benjemaa}@orange.com

Pascal Bianchi
Telecom ParisTech
37 rue Dareau 75014
Paris, France
pascal.bianchi@telecom-paristech.fr

Abstract—In trying to meet the demands of traffic hungry users, mobile network operators are faced with increased CApital EXpenditures (CAPEX) and OPerational EXpenditures (OPEX). The Self Organizing Network (SON) functions have been introduced by 3GPP as a means to cut down these costs. There are mainly 3 categories of such functions: self-configuration, self-optimization and self-healing. In this paper we focus on the second which represents the SON functions performing a run-time optimization of the network. We center our attention on LTE heterogeneous networks. Having several SON functions in a network may lead to conflicts and potentially to bad network Key Performance Indicators (KPIs). Thus a troubleshooting mechanism has to be envisaged. Such a mechanism typically contains 3 steps: fault detection, cause diagnosis and solution deployment. In this paper we tackle the first two and we study the feasibility of using the Naive Bayes Classifier (NBC) in order to build a framework for SON Conflict Diagnosis (SONCD). We provide numeric results proving the feasibility of the framework.

Index Terms—CRE; MRO; eICIC; SON; SON Conflict Diagnosis; LTE; Bayesian networks; SONCD

I. INTRODUCTION

The demand for mobile data is continuously increasing. The International Telecommunications Union (ITU) has released the International Mobile Telecommunications-Advanced (IMT-Advanced) requirements for a 4G network [1]. To help operators meet these demands 3GPP has introduced the heterogeneous networks (HetNets), i.e. besides the Macro cell layer we have a layer of small cells (e.g. Pico cells). However, the use of HetNets leads to increased CApital EXpenditures (CAPEX) and OPerational EXpenditures (OPEX). To bring down the costs, 3GPP has introduced the Self Organizing Network (SON) functions. A SON function is a control loop that tunes the network parameters in order to improve the network Key Performance Indicators (KPIs), see Fig. 1. There are mainly three categories of SON functions: self-coordination, self-optimization and self-healing. In this paper we focus on the second. In the sequel SON refers particularly to this category. Examples of SON functions for HetNets include Mobility Load Balancing via Cell Range Expansion (CRE), Mobility Robustness Optimization (MRO), enhanced Inter-Cell Interference Coordination (eICIC) [1] etc.

Having several SON functions in the network might create conflicts and lead to bad network KPIs [1]. We consider two types of conflicts:

• parameter conflicts: 2 SON functions target the same parameter,
• measurement conflicts: the parameter targeted by one SON function affects the measurements of a different SON function.

Typically a troubleshooting mechanism consists of 3 steps: fault detection, cause diagnosis and solution deployment. In our previous work we have focused mainly on solutions for the third step in the form of a SON conflict resolution [2], [3], [4]. In this paper we focus our attention on the first two steps. We employ a fault detection mechanism which evaluates if we have bad KPIs (making use of some predefined thresholds), and focus mainly on the SON Conflict Diagnosis (SONCD). Diagnosing parameter conflicts is straightforward, i.e. a simple check whether or not there are opposite requests targeting the same parameter could be sufficient. However measurement conflicts are more difficult to diagnose, thus we focus on these conflicts in particular.

The SONCD is responsible for identifying which of the SON functions (and potentially which of their settings) are causing the conflicts. The SONCD concept was introduced fairly recent in 3GPP (Release 10). In a multi-vendor environment the information about the algorithm inside the SON function is typically not shared, i.e. SON functions are seen as black-boxes. In the literature, some tangent work is found in [5], [6], [7] where the focus is on operational troubleshooting based on manual (if-then) rules for how to react on the event of having certain bad KPIs or alarms created by the SON functions. Other works consider the algorithm inside the SON functions to be known, i.e. SON functions are seen as white-boxes, e.g. [8].

In our paper we consider the SON functions as black-boxes. We employ an automated troubleshooting mechanism based on the Naive Bayes Classifier (NBC) [9] as it has proved itself to
be a good diagnosis technique. For example in [10], [11], [12] it has been successfully used to identify bad configurations of the network parameters. In this paper we use NBC to identify the misconfigurations (bad settings) of the SON functions.

The contribution of this work is summarized as follows: i) our SONCD design is operator-centric, considering the SON functions as black-boxes, ii) we base our SONCD design on NBC and iii) we provide a study case in a HetNet scenario with 3 SON functions: CRE, MRO and eICIC.

In section II we present a short description of NBC, section III provides the system description and section IV contains the diagnosis methodology. Results are contained in section V and section VI concludes the paper.

II. NAIVE BAYES CLASSIFIER

Consider a network running several SON functions and which is experiencing some faults, i.e. bad network KPIs. Let \( C \) be the set of potential causes, e.g. SON function misconfigurations. Let \( S = (S_i) \) be the set of symptoms (e.g. network KPIs) that are relevant for helping identify the most probable cause for the faults, i.e. diagnosis (see fig. 2). Thus we have to calculate the conditional probability of having a cause \( C \in C \) given some symptoms \( S: P(C|S) \). According to Bayes’ rule we can write:

\[
P(C|S) = \frac{P(S|C)P(C)}{P(S)}
\]

(1)

Note that \( P(S) \) is the same for all causes so, in order to identify the most probable cause, we do not necessarily need to calculate it. Thus we focus only on the numerator \( P(C) \) is simply the probability that a given cause may occur, so it can be easily obtained. Calculating/estimating \( P(S|C) \) can be quite difficult and impractical especially if we have a lot of symptoms. For this reason we assume i) the symptoms are independent and ii) there can be only one cause at a time. Even if in reality this is rarely the case, still, significant results have been obtained under these assumptions (e.g. [10]). The simplicity of the model has its clear benefits. Consequently eq. (1) becomes:

\[
P(C|S) = \prod_i P(S_i|C) \frac{P(C)}{P(S)}
\]

(2)

The above equation represents a so-called Naive Bayes Classifier (NBC), and it serves as our diagnosis model. The cause inference method simply outputs the most probable cause according to eq. (2).

III. SYSTEM DESCRIPTION

Consider a cell cluster formed by a Macro cell and \( N \) slave Pico cells under its coverage meant to boost the capacity (see Fig. 3). We index the macro cell with \( n = N + 1 \) and the pico cells with \( \mathcal{N} = \{1,...,N\} \).

Let the Cell Individual Offset (CIO), the HO Hysteresis (HYS) and the number of Almost Blank Sub-Frames (nABS) be the \( P = 3 \) network parameters of interest, indexed (in this order) by \( \mathcal{P} = \{1,...,P\} \). The CIO and the HYS are two parameters that are used in mobility management as follows: a UE that wants to transmit data will attach to cell \( n_o = \arg \max_{n \in \mathcal{N}} (RSRP_n + CIO_n) \) where \( RSRP_n \) is the Reference Signal Received Power [13] from cell \( n \) and \( CIO_n \) is the CIO of cell \( n \). When attached to a serving cell \( n_S \) (\( \forall n_S \in \mathcal{N} \)) a UE performs a HO to a target cell \( n_T \neq n_S \) if \( n_T = \arg \max_{n \in \mathcal{N}} (RSRP_n + CIO_n + HYS_{n_S}) \) (1) where \( HYS_{n_S} \) is the CIO of cell \( n_S \). The nABS establishes the number of sub-frames where the cell will not transmit data in order to reduce interference to the neighbors [14]. It is used to protect the cell edge UEs attached to neighboring cells.

We employ \( Z = 3 \) SON functions (\( Z = \{1,2,3\} \)): the CRE function (\( z = 1 \)), the MRO function (\( z = 2 \)) and the eICIC function (\( z = 3 \)). We consider them to be synchronized, i.e. they perform the desired network parameter changes periodically at the end of a time interval of size \( T = 5 \) min. Briefly (we omit for now the time index):

a) CRE : takes as input measurements the vector of loads (LD) of all cells \( m_{LD,n}^1, \ldots, m_{LD,n}^{N+1} \). Its task is to reduce the maximum load (among all cells), i.e. avoid having overloaded cells. Let \( m_{LD} = \max_{n} m_{LD,n} \). For this purpose CRE tunes the CIO of the Pico cells (increase/ decrease / maintain). The settings of the CRE are: the threshold used to trigger parameter changes (\( T_{LD} \)) and the CIO step-size (\( \Delta_{CIO} \)).

b) MRO : takes as input measurements the percentage of too late HOs (TL) \( m_{TL,n}^1, \ldots, m_{TL,n}^{N+1} \) and of ping pong HOs (PP) \( m_{PP,n}^1, \ldots, m_{PP,n}^{N+1} \) (over all HOs) originating only from the Pico cells. We do not look at the ones originating from the Macro cells as typically they are considerably less frequent. Its task is to reduce the averages (over all Pico cells) of the two measurements. Let \( m_{TL} = \sum_n m_{TL,n} / N \) and \( m_{PP} = \sum_n m_{PP,n} / N \). For this purpose MRO tunes the HYs of the Pico cells (increase/ decrease / maintain). The settings of the MRO are the thresholds used to trigger parameter changes (\( T_{TL} \) and \( T_{PP} \)) and the HYs step-size (\( \Delta_{HYS} \)).

c) eICIC : takes as input measurements the throughput ratio (TR), specifically the ratio between the average throughput of the macro users and the average throughput of all the protected users \( m_{TR} \). By protected users we refer to all Pico
users that would be attached to the Macro cell if the CIO and HYS were null. Its target is to keep this ratio between certain bounds in order to guaranty some degree of fairness among the network users (i.e. comparable throughput). For this it tunes the nABS on the Macro cell (increase/ decrease / maintain). The settings of the eICIC are: the threshold used to trigger parameter changes ($\Gamma_{TR}$) and the nABS step-size ($\Delta_nABS$).

For more details see Appendix A. Thus there are $M = 4$ measurements of interest: $LD$, $TL$, $PP$ and $TR$ which we index (in this order) by $M = \{1, ..., M\}$.

IV. SON conflict diagnosis

In this section we first present the potential conflicts, then we establish the set of probable causes and define the symptoms considered for the diagnosis process. Finally we outline the diagnosis methodology.

A. Potential conflicts

Let $\Gamma_{SP}$ be the $Z \times P$ matrix where $\Gamma_{SP}(z, p) = 1$ if SON $z$ tunes the parameter $p$ and $\Gamma_{SP}(z, p) = 0$ otherwise. Let $\Gamma_{MS}$ be the $M \times Z$ matrix where $\Gamma_{MS}(m, z) = 1$ if SON function $z$ takes as input measurement $m$ and $\Gamma_{MS}(m, z) = 0$ otherwise. In our example (see Fig. 4) we have:

$$\Gamma_{SP} = \begin{bmatrix}
0 & 1 & 0 \\
1 & 0 & 0 \\
0 & 1 & 0 \\
0 & 0 & 1
\end{bmatrix}, \quad \Gamma_{MS} = \begin{bmatrix}
1 & 0 & 0 \\
0 & 1 & 0 \\
0 & 1 & 0 \\
0 & 0 & 1
\end{bmatrix}$$

Next, let $\Gamma_{PM}$ be the matrix where $\Gamma_{PM}(p, m) = 1$ if parameter $p$ is expected to have a significant impact on measurement $m$ and $\Gamma_{PM}(p, m) = 0$ otherwise. This matrix is obtained through expert knowledge. In our example we propose the following matrix:

$$\Gamma_{PM} = \begin{bmatrix}
1 & 1 & 0 & 1 \\
1 & 0 & 0 & 1 \\
0 & 1 & 1 & 0 \\
0 & 0 & 1 & 0
\end{bmatrix}$$

where we consider that the CIO does not significantly impact the PP, the HYS does not significantly impact the LD and TR, and nABS does not significantly impact the TL and the PP.

In order to find out which SON functions may end up conflicting we simply calculate, using Boolean algebra ($1 \equiv \text{true}$, $0 \equiv \text{false}$):

$$\Gamma = \Gamma_{SP}\Gamma_{PM}\Gamma_{MS} = \begin{bmatrix}
1 & 1 & 1 \\
0 & 1 & 0 \\
0 & 1 & 0 \\
0 & 1 & 0
\end{bmatrix}$$

If $\Gamma(z_1, z_2) = 1$ it means that SON function $z_1$ may have a negative impact (conflict) on SON function $z_2$. Note that CRE may impact MRO and eICIC, and eICIC may impact CRE.

B. Causes

From eq. (5) we identify that 2 SONs may cause conflicts, namely CRE and eICIC. Thus we define a $1^{\text{st}}$ order cause-set $C_1 = \{\text{CRE, eICIC}\}$ (with cardinality $|C_1| = 2$). We note that the problem may come from one of the different settings of the SON functions like: the triggering thresholds ($T$) and the step size for the tuned parameters ($\Delta$), and so we define the $2^{\text{nd}}$ order cause-set $C_2 = \{\text{CRE, eICIC}\} \times \{T, \Delta\}$ ($|C_2| = 4$). Moreover, if we go into even more details we can also target to identify the alteration degree of the problematic SON setting: low ($1$), medium ($m$), high ($h$) and so we define the $3^{\text{rd}}$ order cause-set: $C_3 = \{\text{CRE, eICIC}\} \times \{T, \Delta\} \times \{1, m, h\}$ ($|C_3| = 12$). In this paper we compare the diagnosis for the 3 orders of cause-sets.

C. Symptoms

As mentioned the SON functions look at the measurements and change the network parameters accordingly at the end of every time step (of size $T$). We use $I = 4$ symptoms calculated at every time step $t \in \mathbb{N}$ each of which is based on one of the 4 measurements. First we calculate the average of $m_i$ within a sliding time window of size $b_i$ (see Fig. 5), for any $i \in \{LD, TL, PP, TR\}$, as:

$$m_i(t) = \frac{1}{b_i} \sum_{j=0}^{b_i-1} m_i(t-j)$$

Next we define the corresponding symptoms as the percentage of time the average $m_i(t)$ is larger than a threshold $c_i$, considering a sliding time-window of size $a_i$:

$$S_i(t) = \frac{100}{a_i} \sum_{j=0}^{a_i-1} I(m_i(t-j) > c_i)$$

where $I$ is the indicator function ($I(\text{true}) = 1, I(\text{false}) = 0$). The values of $a_i$, $b_i$ and $c_i$ ($\forall i \in \{LD, TL, PP, TR\}$) are defined later on in section V.

We define $S_i$, $\forall i \in \{LT, TL, PP, TR\}$, to be the targeted performance w.r.t. the symptoms. At time $t$ we say that there is a network fault if $\exists i \in \{LT, TL, PP, TR\}$ s.t. $S_i(t) > S_i$. 

![Figure 4. SON interaction](image-url)
D. Diagnosis methodology

Considering a network which is experiencing bad KPIs, the diagnosis methodology is composed of 2 stages:

1) We perform a manual diagnosis on a first set of samples, and we learn the diagnosis (qualitative) model, i.e. the probabilities $P(C|S)$ and $P(C)$ in eq. (2).
2) We apply the inference method on a second set of samples.

More details are included in the results Section V.

V. Simulation results

In this section we target to analyze the feasibility of using NBC for the design of the SONCD considering the 3 different orders of (cause) details. For this we employ a network with 9 macro cells. In one of the Macro cells (call it ‘M’) we place randomly 2 slave Pico cells (call them ‘P1’ and ‘P2’). We make use of a wraparound technique. The SON functions will be applied to the cell cluster formed by M, P1 and P2.

We make use of a wraparound technique. The SON functions will be applied to the cell cluster formed by M, P1 and P2.

We consider an FTP-like traffic. User Equipments (UEs) arrive randomly in the network according to a Space Poisson Point Process (SPPP). We consider a UE arrival rate $\lambda$ for the design of the SONCD considering the 3 different settings and for each we run simulations of 2400 hours. One and then leave the network. For details see Table I.

More details are included in the results Section V.

We consider 13 scenarios with different SON function settings and for each we run simulations of 2400 hours. One of them is the reference scenario and the other 12 correspond to the causes in $C_2$ (and implicitly in $C_3$ and $C_4$). For details see Table II. We make use of the symptom parameters in Table III. Fig. 7 shows, for all the scenarios, the percentile of time each fault occurs (LD, TL, PP, TR) and the percentile of time at least one fault occurs (LD/TL/PP/TR). Consequently these faulty samples represent the input to our diagnosis as described in Section IV-D.

We select 1000 faulty time samples from each of the scenarios except REF. Next say that we focus on the $i^{th}$ order cause-set: $C_i$. We get $P(S|C_i)$ and $P(C_i)$ simply by making use of histograms.

We then select 400 other faulty time instances and we test the SONCD. The percentiles of correct identification of the fault are presented in Fig. 8 for all the 3 orders. We have performances of over 91.7% for detecting the SON function causing the problem ($1^{st}$ order cause), of over 82.5% if we also want to detect the problematic setting ($2^{nd}$ order cause) and of over 52% in the case where we want to detect the alteration degree of the setting ($3^{rd}$ order cause). As expected, detecting only the problematic SON is less subject to errors than identifying altogether the problematic SON, the problematic setting and the degree of alteration.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number cells</td>
<td>9 Macro cells, 2 Pico cells</td>
</tr>
<tr>
<td>Propagation Model</td>
<td>OIFF Case 3 [15],</td>
</tr>
<tr>
<td>Bandwidth</td>
<td>10 MHz</td>
</tr>
<tr>
<td>TxPow</td>
<td>-49 dBm / 30 dBm</td>
</tr>
<tr>
<td>Traffic type</td>
<td>FTP-like, constant file size,</td>
</tr>
<tr>
<td>$F S$</td>
<td>$10[Mb/s]/UE$</td>
</tr>
<tr>
<td>$\lambda_{C2}\cdot F S$</td>
<td>$12.5[Mb/s]/Macro\ Hexagon\ Area$</td>
</tr>
<tr>
<td>$\lambda_{HS}\cdot F S$</td>
<td>$12.5[Mb/s]/Pico\ HotSpot\ Area$</td>
</tr>
<tr>
<td>HotSpot radius</td>
<td>100m</td>
</tr>
</tbody>
</table>

Table I

<table>
<thead>
<tr>
<th>Symptom summary</th>
</tr>
</thead>
<tbody>
<tr>
<td>$T_{FL} = 0.1$, [ratio]</td>
</tr>
<tr>
<td>$T_{PP} = 0.2$, [ratio]</td>
</tr>
<tr>
<td>$\Delta_{HY S} = 1$, [dB]</td>
</tr>
</tbody>
</table>

Table II

<table>
<thead>
<tr>
<th>Setting</th>
<th>$t_{LD}$ [ratio]</th>
<th>$t_{PP}$ [ratio]</th>
<th>$t_{FL}$ [ratio]</th>
<th>$\Delta_{C2}$ [dB]</th>
<th>$\Delta_{PR}$ [dB]</th>
<th>$\Delta_{ABS}$ [dB]</th>
</tr>
</thead>
<tbody>
<tr>
<td>REF</td>
<td>0.85</td>
<td>1</td>
<td>2.1</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CRE(\cdot P1)</td>
<td>0.80</td>
<td>1</td>
<td>2.1</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CRE(\cdot P2)</td>
<td>0.75</td>
<td>1</td>
<td>2.1</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CRE(\cdot P1)</td>
<td>0.25</td>
<td>1</td>
<td>2.1</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CRE(\cdot P2)</td>
<td>0.25</td>
<td>1</td>
<td>2.1</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CRE(\cdot P1)</td>
<td>0.85</td>
<td>6</td>
<td>2.1</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CRE(\cdot P2)</td>
<td>0.85</td>
<td>1</td>
<td>2.1</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CRE(\cdot P1)</td>
<td>0.85</td>
<td>1</td>
<td>2.1</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CRE(\cdot P2)</td>
<td>0.85</td>
<td>1</td>
<td>2.1</td>
<td>1</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table III

<table>
<thead>
<tr>
<th>Symptom parameters</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\theta$</td>
</tr>
<tr>
<td>$\alpha$</td>
</tr>
<tr>
<td>$\delta$</td>
</tr>
<tr>
<td>$\gamma$</td>
</tr>
<tr>
<td>$\sigma$</td>
</tr>
<tr>
<td>$\beta$</td>
</tr>
<tr>
<td>$c_i$</td>
</tr>
<tr>
<td>$m_i$</td>
</tr>
<tr>
<td>$n_i$</td>
</tr>
<tr>
<td>$\Delta$</td>
</tr>
</tbody>
</table>
In this paper we tackle the problem of SON Conflict Diagnosis. First we provide the methodology to identify all potential conflict sources. Second we present a framework based on NBC that helps identify the most probable cause. As expected, for an increased level of details describing the cause it is harder to correctly identify the source of a conflict. The performance of detecting only the problematic SON function is the highest followed by the performance of detecting also the amplitude of its (the setting’s) alteration. Thus we find that it is harder to correctly identify the source of a conflict. The research leading to these results has been carried out within the FP7 SEMAFOUR project [16] and has received funding from the European Union Seventh Framework Programme (FP7/2007-2013) under grant agreement no 316384.

VI. CONCLUSIONS

ACKNOWLEDGMENT

The research leading to these results has been carried out within the FP7 SEMAFOUR project [16] and has received funding from the European Union Seventh Framework Programme (FP7/2007-2013) under grant agreement no 316384.

REFERENCES