Cognitive Policy Based SON Management Demonstrator

Tony Daher, Sana Ben Jemaa, Laurent Decreusefond

Cognitive Policy Based SON Management Demonstrator

Tony Daher, Sana Ben Jemaa
Orange Labs
44 Avenue de la Republique 92320 Chatillon, France
Email:{tony.daher,sana.benjemaa}@orange.com

Laurent Decreusefond
Telecom ParisTech
23 avenue d'Italie, 75013 Paris, France
Email:laurent.decreusefond@mines-telecom.fr

Abstract—**Policy Based SON Management (PBSM) framework has been introduced to manage Self-Organizing Networks (SON) functions in a way that they fulfill all together the operator global goals and provide a unique self-organized network that can be controlled as a whole. This framework mainly translates operator global objectives into policies to be followed by individual SON functions. To cope with the complexity of radio networks due to the impact of radio environment and traffic dynamics, we propose to empower the PBSM with cognition capability. We propose a Cognitive PBSM (CPBSM) that relies on a Reinforcement Learning (RL) algorithm which learns the optimal configuration of SON functions and steers them towards the global operator objectives. The visitor will see how changing the operator objectives leads to a reconfiguration of the SON functions in such a way that the new objectives are fulfilled. The operation of a RL based cognitive management will be illustrated and the exploration/exploitation and scalability dilemmas will be explained.**

I. Introduction

The automation of Radio Access Networks (RAN) management became a market reality with the standardization [1] and the commercial release of Self Organized Networks (SON) solutions. Today, SON functions are running in several operational networks to replace individual tasks such as neighboring relations configuration, load management related parameter settings, etc. Hence several SON functions run simultaneously in the same network, each of them fulfilling a specific objective. These individual objectives may sometimes be conflictual and trade-offs need to be found with respect to the operator strategy. Policy Based SON Management (PBSM) framework has been introduced to orchestrate SON functions in a way that they fulfill all together operator global goals [2]. Finding the optimal mapping between the operator goals and the actual actions or behaviors of the SON functions is a complex problem for several reasons. First, SON functions are provided by RAN vendors as black boxes with limited leverages on their behavior, through the configuration of the SON function parameters such as thresholds, step size, parameters intervals, etc that we note as SON function Configuration Value (SCV) sets. Besides, the behavior of a SON function running on a given cell, for a given SCV set, depends on the radio environment, namely the propagation, and on the traffic distribution and dynamics. To cope with this complexity we propose to empower the PBSM with cognition capability illustrated in figure 1 [3]. The proposed Cognitive PBSM (CPBSM) relies on a Reinforcement Learning (RL) algorithm that learns the optimal configuration of SON functions and steers them towards fulfilling the global operator objectives.
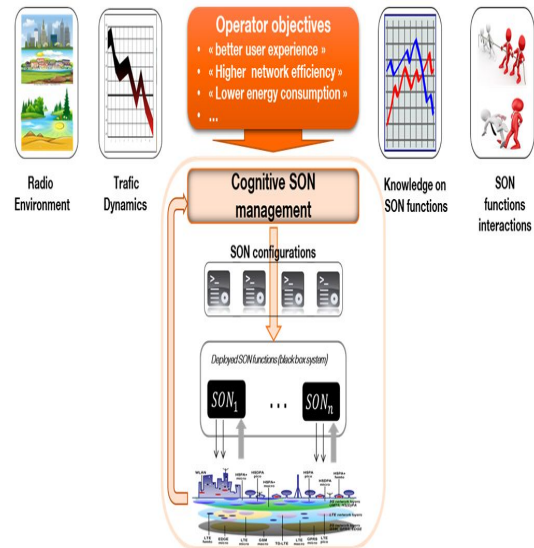

Fig. 1: Cognitive PBSM

The learning technique considered for the demonstrator is a centralized online learning technique, based on Multi-Armed Bandits (MAB) [4].

MAB is a RL problem, formulated as a sequential decision problem, where at each iteration an agent is confronted with a set of actions called arms, each when pulled, generates a reward. The agent is only aware of the reward of the arm after pulling it. The objective is to find a strategy that permits to identify the optimal action, while maximizing the cumulated rewards obtained during the process. In our case, a learning agent on top of the SON functions takes actions by configuring all the instances of the SON functions in the network, as shown in figure 2.
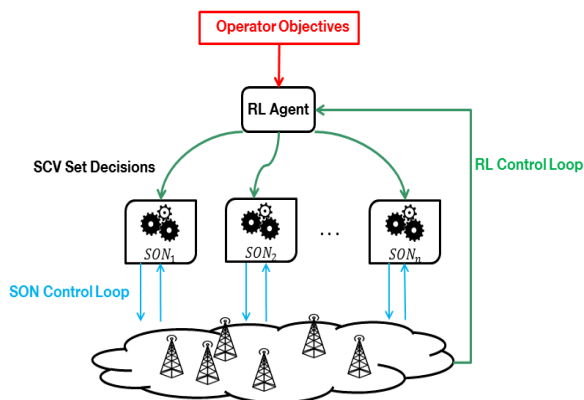

Fig. 2: Cognitive PBSM Functional Scheme

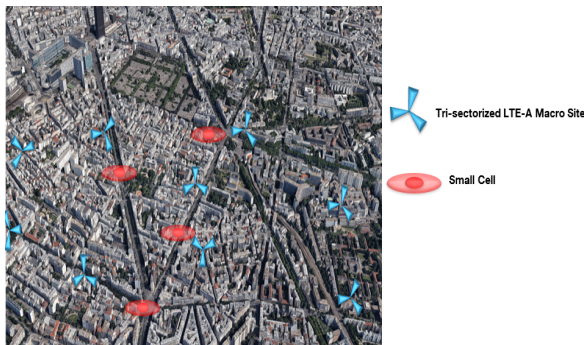## II. DEMONSTRATOR DESCRIPTION

A. Scenario

Fig. 3: Network Topology

The demonstration shows a cognitive management system for radio access mobile networks. Based on operator-defined objectives, such as maximizing end-user average bit-rates and minimizing average cell loads, the cognitive management system automatically configures and controls the operation of SON Functions. The demonstrator is based on an LTE-A radio access network simulator with several SON functions deployed on a heterogeneous network, with realistic topology and propagation modeling. The considered network portion is presented in figure 3. The macro-cellular layer corresponds to a portion of a real network topology with real-like network parameters and accurate ray tracing based propagation. Small cells are added to the network using standard propagation. We consider a stationary and unbalanced traffic distribution. The demonstration scenario considers 3 SON functions:

- Mobility Load Balacing (MLB) deployed on each macro-cell, in charge of setting the mobility parameters to balance the load between neighboring macro cells.
- Cell Range Expansion (CRE) deployed on each small cell, in charge of balancing the load between the small cell and the neighbor macro cell.
- Enhanced Inter Cell Interference Coordination (eICIC) deployed on each macro cell having at least 1 small cell deployed in its coverage region. The eICIC's task is to protect small cell edge users from the high levels of interference generated by the macro cell, by tuning the number of Almost Blank Subframes (ABS) transmitted by the macro cell (ABS contains only control signals transmitted at reduced power).

For this scenario, we will define the following KPIs:

- $L_{i,c,t}$ is the load of cell i
- $\overline{L}_{c,t}$ is the average load in the considered section
- $\overline{T}_{c,t}$ is the average user throughput in the considered section
- $\overline{T'}_{c,t}$ is the average pico cell edge user throughput in the considered section

$c \in C$, where C is the set of actions, which is in this case the set of SCV set combinations for the deployed SON instances in the considered section. t represents the iteration. The variables were normalized between 0 and 1. The reward function reflecting the operator's objective, and that the learning algorithm seeks to maximize, is therefore:

$$r_{c,t} = \omega_1(1 - \sigma_{c,t}) + \omega_2\overline{T}_{c,t} + \omega_3\overline{T'}_{c,t} \qquad (1)$$

Where the load variance is:

$$\sigma_{c,t} = \frac{\sum_{i=0}^{B}(L_{i,c,t} - \overline{L}_{c,t})^2}{B}$$

B is the number of cells in the considered section and ω_1, ω_2 and ω_3 are priority weights set by the operator.

B. Demonstrator

The CPBSM demonstrator consists of 3 main components that are: the *Operator Objectives Panel*, the *Learning Panel* and the *Performance Evaluation Panel*.

In the *Operator Objectives Panel*, the operator defines its objectives by giving priorities to several proposed criteria such as load balance, end user throughput or interference level at cell edge. An example of operator objective panel is illustrated in figure 4.
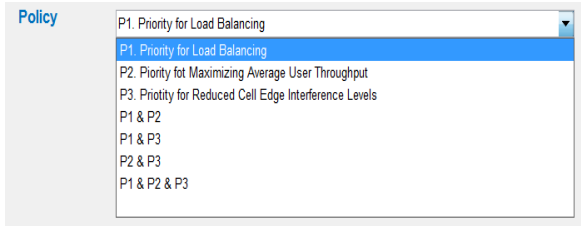

Fig. 4: Operator Objectives Panel

The *Learning Panel* shows the evolution of the CPBSM's action selection statistics, as well as the generated rewards. This panel, illustrated in figure 5, shows how the learning process progresses. The action selection statistics graph shows the evolution of the online learning agent decisions. In the first stages of the learning, the demonstrator shows an almost random action selection, due to the lack of knowledge of the learning agent. After several iterations, the agent's decisions start converging towards the optimal action. This is also shown in the reward function evolution graph, where we can see that the generated rewards keep increasing with time, to finally converge to a stationary state. Furthermore, the demonstration shows that the CPBSM, through its learning agent, is able to adapt to objective changes, i.e. when the operator changes its objectives, the CPBSM will perform a brief learning phase, before converging back to a new optimal action.
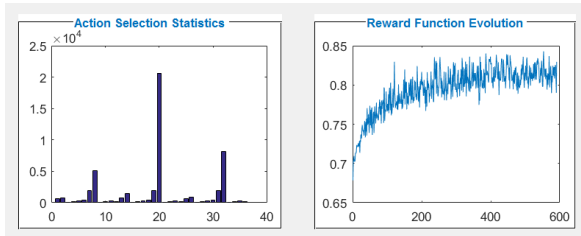

Fig. 5: Learning Process Evolution

The *Performance Evaluation Panel* compares the KPIs performances of the CPBSM with a none-objective driven, static configuration of SON functions as shown in figure 6. It is important to note that the static configuration of SON functions is a typical configuration deployed on real field networks. The objective is to illustrate the gain that we can obtain by adopting the CPBSM with respect to typical performances that can be obtained today with SON.
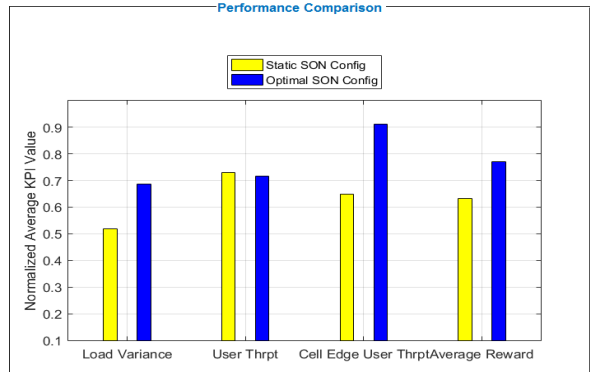

Fig. 6: Performance Comparison

III. CONCLUSION

The demonstrator shows how cognition capability allows the RAN management system to always make optimal decisions. The visitor will see how changing the operator objectives leads to a reconfiguration of the SON functions in such a way that the new objectives are fulfilled. The operation of a RL based cognitive management will be illustrated and the exploration/exploitation and scalability dilemmas will be explained.

REFERENCES

[1] 3GPP TR 36.814, "Technical Specification Group Radio Access Network; Evolved Universal Terrestrial Radio Access (E-UTRA); Further advancements for E-UTRA physical layer aspects (Release 9)", March 2010.

[2] SEMAFOUR project web page http://fp7-semafour.eu/.

[3] T. Daher, S. Ben Jemaa and L. Decreusefond, "Cognitive Management of Self-Organized Radio Networks Based on Multi Armed Bandit," *IEEE Personal Indoor and Mobile Radio Communications (PIMRC)*, 2017.

[4] P. Auer, N. Cesa-Bianchi and P. Fischer, "Finite-time analysis of the multiarmed bandit problem," *Machine learning, 47*(2-3), pp. 235-256, 2002