Theory of Algorithm Suitability on Managing Radio Resources in Next Generation Mobile Networks

H. Fall, O. Zytoune, and M. Yahyai

Abstract—**Beyond 2020, wireless networking model will be radically changed and oriented to business-driven concept as foreseen by the next generation mobile network (NGMN) alliance. As the available spectrum granted to a given operator is physically limited, new radio resource management techniques are required to ensure massive connectivity for wireless devices. Given this situation, in this paper we investigate how the key network functionalities as self-optimizing network (SON) must be thought to meet NGMN requirements. We propose therefore, algorithm suitability theory (AST) combined with the notion of network operator infrastructure convergence. The approach is based on software-defined networking (SDN) principle that allows an adaptability of the load balance algorithm to the dynamic network status. Besides, we use the concept of network function virtualization (NFV) that alleviates the constraint of confining the wireless devices to their home network operator only. Relying on these two technologies, we build AST through a lexicographic optimality criterion based on SPC (Status, Performance, and Complexity) order. Numerical results demonstrate a better network coverage verified by the improvement of metrics such as call blocking rate, spectrum efficiency, energy efficiency and load balance index.**

Index Terms—**Network function virtualization, Software-defined networking, Lexicographic optimality, Self-Organized networking, Radio Resource Distribution.**

I. INTRODUCTION

The evolution of wireless traffic demand and business models will lead to a fully mobile and connected society in the context of 2020 and beyond [1]. In such a situation, the spectral efficiency (SE) becomes one of the key challenges when handling such explosive data traffic on a physically limited bandwidth [2]. Meanwhile, the number of delivered bits per joule (j), which is known as the energy efficiency (EE), represents a relevant parameter in today's mobile network desi-gn, as the connected society must be also a green one. With the random behavior of mobile users, space-time variation in traffic demand causes a non-uniform load distribution among cells, which leverages negatively the SE and EE performances.

The third generation partnership project (3GPP) provided Load Balancing (LB) operation through its self-organizing network (SON) functionality [3]. As soon as the standard LB scheme has been published, it has been demonstrated that its original formulation could be optimized in terms of SE and EE. Therefore, papers [4-6] approach the phenomenon with the same principle by considering the user quality of service (QoS) constraints. Differing from this approach, authors in [7] introduce an EE scaling factor as a criterion for target cell selection in LB procedure. An interference-aware LB solver is studied in [8] where the proposed solution guarantees a low level of Inter-Cell Interference (ICI), which leverages edge user throughputs. While in [8], a network status (ICI) is considered, authors in [9] advocate a Cell-Reselection-based LB scheme where they demonstrate the effectiveness in the environment with a lot of small size data packet services, which is a frequent scenario with the explosion of smart-phones. Without describing the LB scenario between two cells, authors in [25] claim hard reliability guarantees through a distributed and adaptive resources management controller, which allows the optimal exploitation of Cognitive Radio and soft-input/soft-output data fusion in Vehicular Access Network.

By analyzing this non-exhaustive literature review, we realize that LB algorithms suffer mainly from these drawbacks: first, their diversity demonstrates their partial contribution in network performances. The consequence is a non-permanently optimized system. Second, their formulation uses combinatorial optimization approaches, which are often complex. Given the limited capacity of Base Stations (BS), they cause high power consumption and delay degradation. Third, the actual design approach is hardware-oriented and is not coherent with NGMN requirements where scalability will be a performance metric indicator [10]. Fourth, to receive data from the network, users are allowed to establish uplink and downlink communications only with their home access points (AP), and SON operations are performed between two BSs owned by the same operator. Therefore, mobile users are confined to their home available spectrum while today's business model configuration foresee the convergence of network infrastructures [11]. To counter the limitations cited above, algorithm suitability theory (AST) is

proposed as an interesting alternative. The concept tries to optimize permanently the network performance by benefiting from all algorithms thanks to the device programmability aspect of SDN (software-defined network) principle. We define therefore what we call spectro-energy efficiency (SEE), which represents the number of bits received by a mobile per combined energy and frequency unit. A multi-objective function of EE and SE is formulated using scalarization method. By using a Markov model prediction of network status, an SDN controller supervises in real time fashion, the cell conditions. Then, with a lexicographic optimality criterion, it maximizes the objective function by ascribing the resolution of two wireless radio interface operations (Load Balancing and Radio Resource Distribution (RRD)) to predefined optimizers. We associate to AST technique, the concept of full LB operations, i.e. between two APs not necessary belonging to the same operator. In this configuration, the set of candidate users is extended, and not reserved to the edge ones. With the notion of network operator infrastructure convergence allowed by the NFV (Network Function Virtualization) approach, BSs could balance their load with tier parties even if the target AP belongs to another operator but located in the same area.

The remainder of this paper is as follow: in Section 2, we present the system architecture and air interface model. Section 3 develops the proposed AST theory while Section 4 presents the resolution through a lexicographic optimality criterion. We debate in Section 5, on obtained results, draw conclusions, and forecast perspectives in Section 6.

II. Reference Architecture and System Model

A. NFV principle and network operator infrastructure convergence

Performances such as much greater throughput, much lower latency, ultra-high reliability, much higher connectivity density, and higher mobility range are the main NGMN requirements, which are unexpected with today's network design and operations [1]. To address these challenges, NGMN design should include embedded flexibility to optimize the network usage. A higher connectivity density requires an availability of spectrum when needed. Physically speaking, a given network operator cannot ensure absolutely this availability every time and everywhere due to the limited bandwidth budget and space-time variation in traffic demand. Therefore, one needs sometimes to balance the network load between neighboring BSs. The classical SON operation links two BSs owned by the same network operator [3] and transfers edge user status from one AP to a neighboring one by adjusting handover parameters. The problem with this approach is twofold: first, the target candidate BS accepts the transfer if and only if the requested resource is available enough to ensure user (QoS). Second, as the users are able to decode their home network operator signals, the ICI phenomenon degrades the throughput performances. A question of fundamental importance is how to enhance quantitatively and qualitatively, the set of candidate target APs. The former is related to the number of candidate target BSs and the last concerns the signal quality and/or availability of resources. The network

architecture presented in Fig. 1 allows to the user, the possibility to be connected to different radio access network
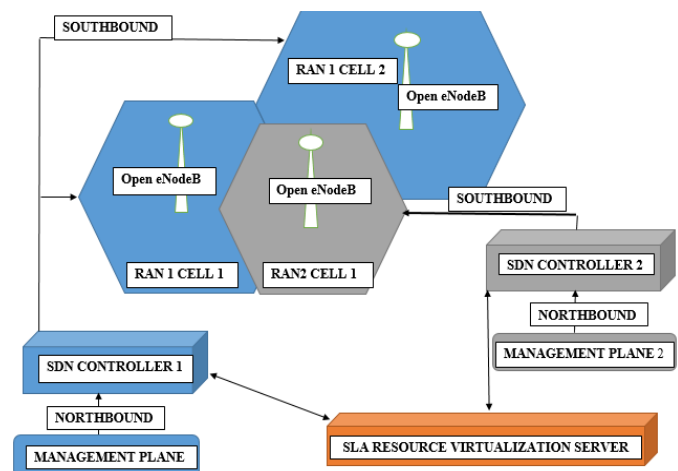

Fig. 1. Reference architecture

(RAN) owned by different network service provider as proposed in [11]. By virtualizing the RAN, the balance can be made even if the cells control the same geographical region. As two mobile network service providers operate in different sub-bands, the constraint of ICI is quasi-alleviated in this LB procedure approach. In Fig. 1, the service level agreement (SLA) server is the common network device shared by two mobile service providers. Therein, the resource sharing policies are defined according to certain conditions based merely on resource availability in target candidate APs. With a Markov prediction, at every predefined time transmission interval (TTI) between stakeholders, Open BSs forward their bandwidth usage ratio (BUR) to their corresponding SDN controller that informs the SLA server. Instead of balancing load always to neighboring cells (from cell 1 RAN 1 to cell 2 RAN 1), a tradeoff could be studied between cost, or resource (spectrum or energy) in balancing load to another RAN owned by a different operator (RAN 2 cell 1). When the LB procedure is generated by the user mobility, the SON operations take place between two cells belonging to the same RAN [3]. Herein, the SDN controller constitutes the network brain and adapts the LB algorithm to the open eNodeB relying on network status (ICI, load, Energy mode, etc…).

B. SDN concept and LB algorithm adaptability

SDN concept (Fig. 2) is the most serious candidate technology for future networking in terms of management and exploitation [12]. It differs from the actual networking design by decoupling the forwarding plane from the decision plane. Build through tree layers; it makes the network scalable by providing a programmability feature of devices (Open eNodeB, Fig. 1). For wireless cellular networks, radio protocols could be defined on Management plane where their possible improvement is possible. Once the management plane defines a suite of decisions, the control plane has to monitor the forwarding devices by adapting the decisions to the matching traffic requirements. The APs and switches, which form the data plane, have no embedded intelligence, and through a
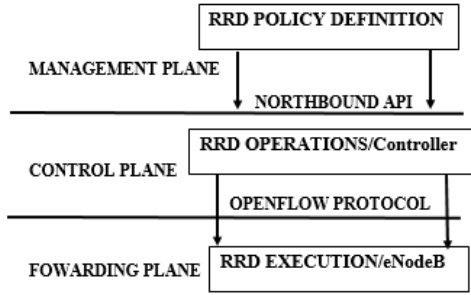
Fig. 2. SDN Network architecture

telemetry system, they upload the network status permanently to the controller. Therefore, the system is flexible, scalable and open. The programmability aspect of SDN is used in this paper to enhance resource (energy and spectrum) efficiency. Prior to this optimization, we present the system model in next sub-section.

C. Spectrum, time and energy model

A common understanding of NGMN is that it should integrate the long-term evolution-advanced (LTE-A) RAN in the definition of its global air interface (AI) as developed in [13]. The overall AI is composed of multiple AI variant (AIV). In this paper, we consider the LTE-A AIV in load balance procedure between two RAN belonging to the same network operator or not.

Consider a wireless cellular deployment and a set B of neighboring BSs. Due to frequency agility of LTE-A radio interface, each BS can operate randomly with a given bandwidth among the set of available spectra [14]. Let W_b be the available bandwidth at a given BS b. The access mode to LTE-A radio interface is based on Orthogonal Frequency Division Multiple Access (OFDMA). However, in NGMN vision, the notion of SoDeMA (Software-Defined Multiple Access) allows an adaptation of radio access method based on network status requirements (for example channel bandwidth or propagation conditions).

In OFDMA environment, every user k in the set K of mobiles turns a random number of services (VoIP, Streaming Video, Online gaming, etc…). The bandwidth W_b is shared in a set N of physical resource block (PRB) made with 12 subcarriers. Then, the resource allocation is submitted to the following constraint:

$$\sum x_{k,n,s} w_n \leq W_b \qquad (1)$$

$x_{k,n,s}$ is an assignment parameter equal to 1 when the PRB w_n is allocated to the mobile k on its service s and 0 otherwise. The bandwidth usage ratio (BUR) μ is defined in (2) as:

$$\mu_b = \frac{\sum x_{k,n,s} w_n}{W_b} \qquad (2)$$

According to [15], when $70\% \leq \mu_b \leq 100$, the cell is heavily loaded, while $\mu_b \geq 100$ characterizes an overloaded cell. Load Balance is recommended to preserve network performances and a mobile user k is attached to only one BS b in the context of equation (3):

$$\sum_b x_{k,b} = 1 \qquad \forall \ k \qquad (3)$$

Let K^e denotes a subset in K and represents the cell edge mobile users. At cell edge regions, the throughput of mobile users is leveraged by the SINR (signal to interference plus noise ratio) α_{b,k^e} given in equation (4) by:

$$\alpha_{b,k^e} = \frac{P_{b,k^e} H_{b,k^e}}{\sum P_{b',k^e} H_{b',k^e} + \delta} \qquad (4)$$

$P_{b,k}$ and $H_{b,k}$ denote respectively the power seen by the mobile k from BS b, the channel gain and, δ is the Additive White Gaussian Noise (AWGN). For an interfered channel, the maximum available rate on a given PRB n for a mobile user k is given in equation (5) and, for a minimum rate $r_{k,s}$, on its service s, the required QoS follows the constraint in (6):

$$R_{k^e,n} = W_n log_2(1 + \alpha_{b,k^e}) \qquad (5)$$

$$\sum x_{k,n,s} R_{k^e,n} \geq r_{k^e,s} \qquad (6)$$

Time is a granular resource in LTE-A air interface procedure. Indeed, the resource allocation is performed at every TTI or scheduling period, which measures *1 ms*. Mobile radio conditions, queue length, service priority are analyzed every TTI before PRB allocation. LB also has a cycle duration when it is required [16].

For energy characterization, the power seen by a mobile k from BS b is the sum of total powers received in every PRB n. Then, the downlink transmission for all users is submitted to relation (7):

$$\sum \sum P_{k,n} \leq P_{max} \qquad (7)$$

P_{max} is the overall power budget available at the BS (Downlink transmission power and reciprocal of drain efficiency of power amplifier) for a BS b. The SE is defined as the number of bits received by a mobile per unit bandwidth as seen in equation (8):

$$SE = \frac{R}{W_b} \qquad (8)$$

Where $\quad R = \sum R_{k,n}$

The EE is defined as the number of bits received by a mobile per unit energy as seen in equation (9):

$$EE = \frac{R}{P_{max}} \qquad (9)$$

III. ALGORITHM SUITABILITY THEORY

SE and EE are increasing functions of power and bandwidth respectively (Fig. 3) and, their optimization may present two conflicting objectives [17]. LB Algorithms, which are based on QoS constraints [4-6], optimize the SE as the throughput requirement (Equation 6) relies on an efficient use of bandwidth. As far as that goes, the solvers taking into account the SINR [8], walk in the same way because a low level of ICI means a good rate (Equation 5). The EE scaling factor [7] relies on power mode of target BS and, balance load efficiently.
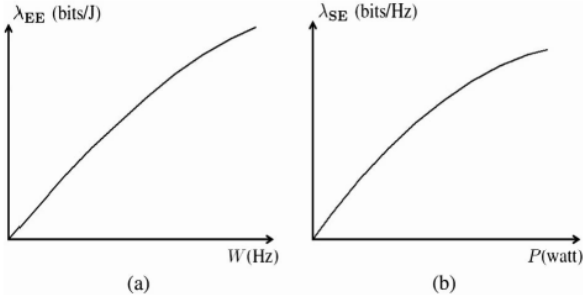
Fig. 3. Evolution of EE and SE in function of frequency and power [17]

It reduces therefore the power Consumption of the system. Without being an LB scheme, resource efficiency presented in [17] makes a combination of conflicting objectives as shown in Equation (10). However, this scheme considers a perfect channel state information, i.e. without taking into account SINR phenomenon.

$$\text{Max } (F = \gamma_1 SE + \gamma_2 EE) \qquad (10)$$

$$\text{s.t.} \quad (1), (6), (7)$$

Equation (10) is a summation of two parameters with different dimensions [(bit/Hz) and (bit/joule)]. However, this weighted sum with $\gamma_1 + \gamma_2 = 1$ could be interesting if we introduce the following parameters:

$$\beta_{EE} = \frac{EE}{W_b} \qquad (11)$$

$$\beta_{SE} = \frac{SE}{P_{max}} \qquad (12)$$

Interestingly, (11) and (12) measure the number of bit per combined unit energy and bandwidth [bit/(Hz*joule)]. Therefore, two contributors of the same performance parameter are formulated. This metric represents both SE and EE. Then, let us rewrite the objective function in (10) as follow:

$$\text{Max } F_{SEE} = \gamma_1 \beta_{SE} + \gamma_2 \beta_{EE} \qquad (13)$$

$$\text{s.t.} \quad (1), (6), (7)$$

As seen in (11) and (12), an EE optimizer only use as bandwidth as possible (Fig. 3). Then, the denominator of the first term in (11) increases and, therefore decreases the first term of (13). The same reasoning can be done for the second term in (13). At the best of our knowledge, there is no LB algorithm, which fulfills all network status and QoS requirements. Consider that this algorithm exists; it will be very complex for the limited capacity of BSs, knowing that LB procedure operates between BSs. Furthermore, this complexity causes delay degradation. With the actual design approach, improving this algorithm means manual programming which sounds quasi-impossible as the next generation cellular networks will be characterized by high node density in some area like urban zones. For these reasons, we tackle the problem at two levels (LB and RRD) with a lexicographic criterion optimality developed in the next Section.

IV. SPC-BASED LEXICOGRAPHIC OPTIMALITY APPROACH OF AST

A. Markov Model Prediction of network conditions

Let S, be a set of network status (Interference level, bandwidth usage ratio, user diversity, energy mode, etc…). The mobile users return at every TTI, the channel conditions through a parameter called Channel Quality Indicator (CQI), which informs about SINR and RSSI (Received Signal Strength Indicator). As a time scheduling period is very small (*1 ms*), we assume that two consecutive TTIs do not differ well in terms of SINR and RSSI. Thus, at every TTI, the BSs have to forward the following parameters to the SDN controller:

- the bandwidth usage ratio μ_b
- the average SINR of cell edge user defined as:

$$\alpha_b = \frac{1}{K^e}\sum_{k^e} \alpha_{b,k^e} \qquad (14)$$

Those parameters received at frame (*i*), are the matching information in the processes of predicting the network state at frame (*i+1*). As the system next state is predicted based only on the precedent one, the process can be qualified as a Markov process. The following table (TABLE I) gives the proposed bandwidth usage ratio (BUR) transition Probability. The different network states are the ones defined in subsection *C*.

TABLE I
BUR SYSTEM TRANSITION PROBABILITY

STATUS	S1	S2	S3	S4
S1	0.5	0.25	0.25	0
S2	0.125	0.5	0.25	0.125
S3	0.125	0.25	0.5	0.125
S4	0	0.25	0.25	0.5

B. Standard Formulation of lexicographic optimality

Lexicographic optimality is an optimization approach where several objectives, in competition, are classified according to a specified order of importance [18]. It can be formulated as follows:

Consider a combinatorial optimization problem. The objective functions f_i with $i = 1, … … . p$, where p is the number of objective functions, are classified in such a way that, when $i < j$, f_i has higher priority than f_j in lexicographic philosophy [19].Then, we solve the problem by optimizing the first objective function f_1; the obtained optimum is imposed as a constraint in the process of optimizing f_2. An iteration is performed until the general optimum is found. The process can be formulated as shown in (15), with $(MOP_{lex,i})$, a Multi-objective Optimization Problem with lexicographic criterion. The index i represents a given objective function in the lexicographic order.

With $f_j(x_j^*) = f_j^*$ and f_j^* the better solution found by optimizing the f_j objective function. Ω is the set of feasible solutions.

$$(MOP_{lex,i}) = \begin{cases} \min f_i \\ s.t \\ x \in \Omega \\ f_1(x) = f_1^* \\ f_2(x) = f_2^* \\ . \\ . \\ . \\ . \\ f_{i-1}(x) = f_{i-1}^* \end{cases} \quad (15)$$

C. Algorithm Suitability Theory

1. General considerations

Using these above mathematical theories, we propose the following LB scheme by considering these hypotheses: let A be a set of algorithms resolving the two LTE-A air interface procedures (TABLE II): Load balancing witch is intrinsically combined to radio resource distribution (RRD). Subsection $A)$ describes the network state and its model of prediction.
For load balancing process, a cell can experience four states:

- State 1 : $\mu_b < 70\%$, normal network operation.
- State 2 : $70\% < \mu_b < 100\%$ and $\alpha_b \geq \varphi$, the cell is heavily loaded but ICI level is acceptable for good transmission.
- State 3: $70\% < \mu_b < 100\%$ and $\alpha_b < \varphi$, the cell is heavily loaded and ICI level is high.
- State 4: $\mu_b \geq 100\%$, the cell is overloaded.

For ICI level and for a user k^e, the SINR α_{b,k^e} must verify:

$$\alpha_{b,k^e} \geq \varphi, \quad (16)$$

Where φ is the minimum required signal level for guaranteeing 1% BLER (bloc error rate) [20].
We assume that when $\alpha_b \leq \varphi$, the ICI starts to destroy transmissions, mainly for edge user k^e governed by a BS b (TABLE III).

TABLE II
RADIO RESOURCE MANAGEMENT ALGORITHMS USED IN THIS PAPER

Algorithm	Operation	Characteristics		
		state	performance	Complexity
$CR - LB$	LB	SSPE-O	acceptable	Low
$IA - LBA$	LB	ICI-O	acceptable	high
$EE - LBA$	LB	ESM-O	acceptable	Low
PSO	RRD	Load-O	acceptable	high
WF^2Q	RRD	ICI-O	acceptable	average
$QA - EERS$	RRD	EE-O	acceptable	high

2. Ordering objectives through SPC (Status-Performance-Complexity).

Load balancing means transferring some users from a heavily or over-loaded cell to a slightly loaded neighboring one by adjusting handover parameter. Therefore, technically speaking, all LB algorithms have the same principle [3]. However, they differ from mathematical formulations, triggering parameters, cell performance metric considerations, etc…. Consider the LB algorithm treated in [8], it worries about interference level and provides good performance by reducing the ICI at cell edge regions (TABLE II). Likewise, for balancing load, the energy mode of potential target cells can be considered as in [7]. Therefore, there is a network status dimension in the formulation of algorithms. Given the random variability of network conditions, we assume that optimization process of load balancing must follow the network state. Then, the first objective function (f_1) in our lexicographic order represents the network state. This first criterion is submitted as a constraint in the second where the performances of algorithms are assessed. Then, the second objective function (f_2) represents the performances of a given algorithm in a given state. Algorithms differ also by the mathematical approach. Given a state, algorithms offering the maximum of performances in SE and/or EE with less complicity are more efficient. Then, the complexity of the algorithm represents the third objective function (f_3) in the lexicographic criterion. We can resume algorithm suitability as follows: ***In a given cell status, which algorithm offers more performances with less complexity (SPC).*** Fig. 4 describes the proposed process.

3. AST Description

The LB scheme proposed with AST vision is based on the considerations mentioned in Subsection C.2. Given the diversity of LB algorithms (see TABLE II), we assume that an implementation with only one scheme couldn't provide a permanent system optimization. The idea is "right LB algorithm at right status". We rely therefore on the schemes presented in TABLE II which are used for the development of AST technique and present two sets of algorithm. The first category concerns LB operation and the second one is related to the radio resources distribution (RRD) functionality. RRD algorithms are also characterized by their diversity (load-oriented, ICI-oriented or EE-oriented). AST objective is the optimization of ratios in Equation (11) and (12) by the selection of schemes (see Fig 4), which provide better performance in the experienced status. For each ratio, when the required LB scheme is a spectral

efficiency one, we equilibrate the related metric with an energy efficiency RRD scheme such that the SEE parameter is maximized and vice versa.
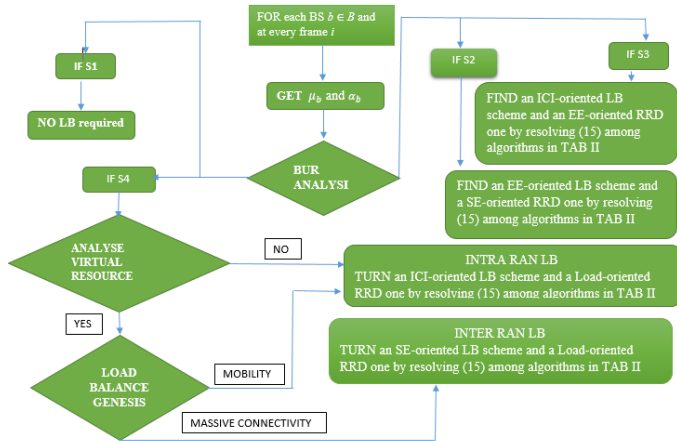

Fig. 4. AST procedure

As seen in Fig. 4, at every frame, each BS b uploads the instantaneous level of μ_b and α_b, which represent the bandwidth usage ratio (BUR) and the experienced ICI level. TABLE I gives the system transition probabilities between the states that the cell could experience. As proposed in subsection 4.A, LB and RRD algorithms performed by the AP at frame $i+1$ are based on data received by the SDN controller at frame i. When the system is in S1, the BUR is under 70%, radio resources are sufficient and LB is not required. When it passes from S1 to S2, the cell is heavily loaded, but ICI level is acceptable for good transmissions. In such a status, it is more interesting to prioritize an EE-oriented LB algorithm (EE-LBA) and a SE-based RRD (PSO) one. Indeed, while the former optimizes the system energy, the last provides a spectrum efficiency. In the context of equations 11 and 12, the two ratios are both enhanced at the same time, and finally the general SEE in Equation (13). If the system passes from S1 to S3, the ICI level degrades the transmission, we equilibrate the load between cells (IA-LBA) by taking into account the interference and distribute the radio resource efficiently in EE side (QA-EERS). When the system passes from S2 or S3 to S4, the network is overloaded. In this configuration, we question the tier party's resource availability. If the LB is generated by the user mobility, it is cost efficient to perform an intra- RAN LB by adjusting handover parameters. If we face massive connectivity with slight mobility, i.e. the mobile users remain in the cell region; it is interesting to use the available resource in neighboring cell partner, expecting a good signal strength from it. The choice of algorithms is done with the same principle.

V. RESULTS AND DISCUSSIONS

The performances of a given LB algorithm is evaluated through call blocking rate (CBR), load balance index (LBI) and fifth percentile throughput. In addition to these parameters, we assess EE and SE behaviors of Algorithm suitability in comparison with some reference algorithms. Prior to discuss on obtained result, let us present the simulation platform in the following sub-section.
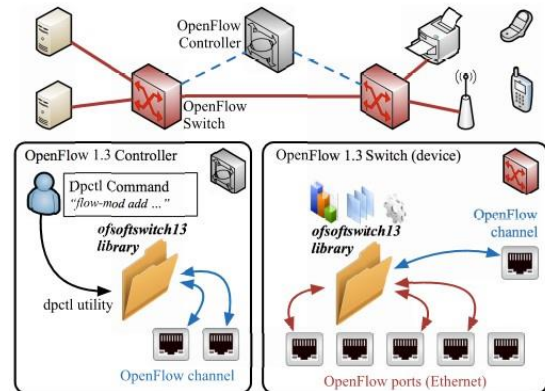
1. Simulation Platform

Fig. 5. SDN-enabled LTE-A network [27]

Actually, the SDN-enabled LTE environment is not yet deployed in real scenario but it exists some test beds allowing the simulation of SDN-enabled LTE networks. A reasonable choice would be using the ns-3 network simulator [28], which is an open-source discrete event simulator, designed to work not only with simulations, but also integrated with virtualized test-bed environment. It is structured in modules, and among them, there is a complete LTE-EPC implementation. We use the OpenFlow 1.3 module [27] for ns-3 (Fig. 5) to simulate an SDN-enabled LTE network. However, it is worth to note that this module is designed for the evolved packet core (EPC) and do not integrate the RAN functionality. Although, according to [27], the controller can be extended to implement any desired feature, such as those necessary to control an SDN-based RAN network. This is because, since ns-3 is free software, it is possible to modify the protocols in any desired way to proper integrate both technology and evaluate new architectures.

2. Call blocking rate (CBR)

A call is blocked when the cells are not able to admit new calls due to reduced available resources. As seen in TABLE II, the solvers do not take into account the variability of the radio conditions. The CBR decreases when SINR increase. In Fig. 6, the mean SINR is fixed at $\varphi = 3.8\ dB$ (TABLE III). Therefore, when we are below φ, AS-LB handles ICI-LB that has the capability of optimizing networks resource in those situations. Moreover, the scheme optimizes also the RRD distribution by proposing an EE-aware resource scheduling as QA-EERS, which uses efficiently the spectrum with reduced energy. (Fig. 6) represents the CBR in function of different levels of interference and, the reduced performance of EE-LB is due to its formulation, which the load distribution principle presents few sensibility to ICI.
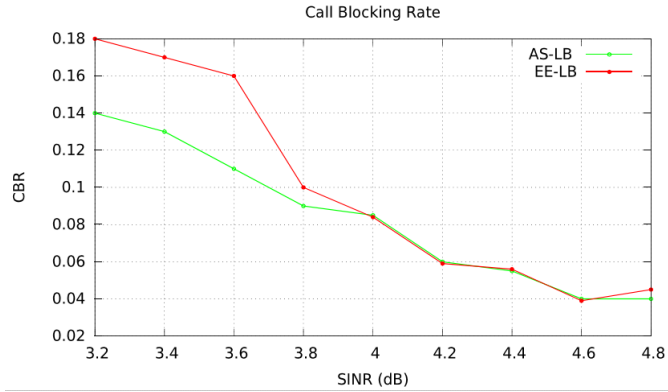
Fig. 6. Call Blocking Rate vs. SINR. $\varphi = 3.8\ dB$

3. Load balance index (LBI)

Load balance index measures how well the distribution of users among BSs is. A high LBI means a balanced allotment among APs while a low LBI describes some heavily loaded cells alongside with slightly loaded ones. (Fig. 7) presents the evolution of LBI for three algorithms.

TABLE III
MODULATION AND CODING SCHEME (MCS) WITH REQUIRED SINR [23 SCHOENEN]

Index	MCS	SE (bit/s/Hz)	Min SINR (dB)
0	Outage	0	< 0.9
1	QPSK 1/3	0.75	0.9
2	QPSK 1/2	1	2.1
3	QPSK 2/3	1.25	3.8
4	16QAM 1/2	2	7.7
5	16QAM 2/3	2.75	9.8
6	16QAM 5/6	3.25	12.6
7	64QAM 2/3	4	15.0
8	64QAM 5/6	5	18.2

Given that effectiveness in SE and EE is the main goals of NGMN, we argue that this parameter should not be maximized at any moment. Algorithm suitability technique presents an average LBI as it handles a set of schemes and keeps the mean value of their LBIs. Thus, we notice that its LBI tends at stabilization while other algorithms present continuously increased LBI. Meanwhile, solvers as $EE - LB$, which relies on the target BS energy saving mode [7] presents often reduced LBI. Indeed, when the target BS is in energy saving mode or begins to start an energy saving mode, a tradeoff between load balance and power saving is performed and the candidate users are not automatically transferred.
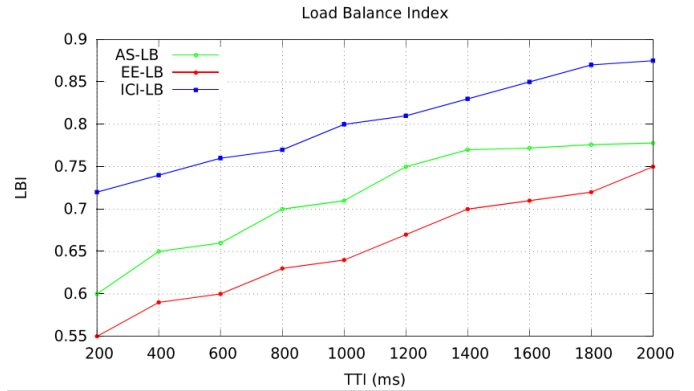
4. Spectral efficiency

Fig. 7. Load Balance Index in function of time

SE is the principal challenge of NGMN because the spectrum is scarce while we must ensure massive connectivity. Meaning the number of bit per unit bandwidth, SE must be as greater as possible. Algorithm suitability outperforms all the other algorithms because it benefits from all advantages (Fig. 8).

In Fig. 7, it is worth to notice that AS-LB performance is located in a high SE gap [1.5:2.6] in comparison with the fixed MCS SE [1.25, (TABLE III)], and regardless of the transmission power. Obviously, when the power increase, other algorithms present increasing SE, i.e. they degrade energy at the same time. An ICI-oriented LB scheme may not be poor in term of SE in low level interfered environments but there are alternative schemes, which present greater performances in those conditions. The algorithm suitability exploits this opportunity enabled by the time granularity of LTE-A air interface operations. When a solver do not integrate a given network condition in its formulation (**EE-LB)**, it is often inefficient in that state. The high contribution of AST to spectrum efficiency could be justified by the possible availability of tier party's bandwidth resource. Indeed, when this partnership is handled, the users remain in the same region like their home APs, benefiting therefore from the partner cell, a good signal strength. A good RSSI mean more bits with reduced frequency as the system operates with high-level MCS order (TABLE III).

5. Energy Efficiency

The same reasoning as in SE can be done in EE side. Knowing that the throughput depends on radio conditions, and bad radio conditions increase transmission power [24], when we optimize the rate, we participate on EE findings. Algorithm suitability takes into account all these considerations. Moreover, a second level of optimization is introduced in the LB process. Indeed, the RRD, which follows also network status, is optimized on demand. Therefore, the presented scheme outperforms evenly the other solutions in terms of EE (Fig. 9). More interestingly, it formulates, by combination of EE and SE, a single optimization parameter referred to as SEE,

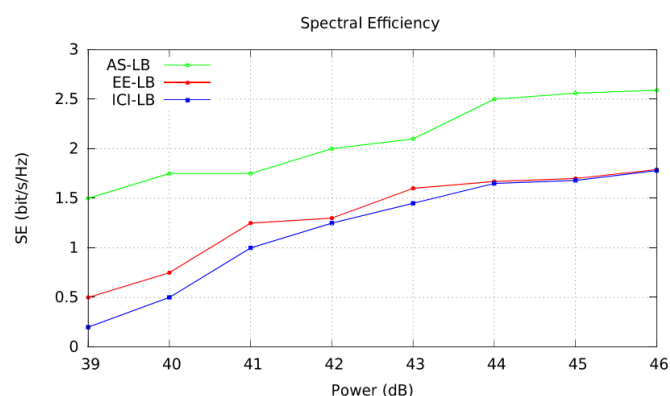which is more representative as network performance metric indicators.

Fig. 8. Spectral Efficiency vs. Transmission Power with $\varphi = 3.8\ dB$ MSC=QPSK 2/3

In Fig. 8, there are two zones: before 70% BUR and after. In the first region, which corresponds to normal network operations, LB is not necessary. Therefore, the QoS-oriented LB schemes, which offer high throughput, are handled. Transferring data with high rate reduces the energy consumption. Besides, after 70% of BUR, the EE decreases because the cell starts to be heavily loaded. EE-LB and AS-LB present still high effectiveness, as they are equivalent. The greater performance of AS-LB is explained by the second level of optimization: When the BUR increase, the ICI-LB combined to the QA-EERS or the EE-LB combined to ICI-oriented RRD, keep the performance of AS-LB at higher values. On the other hand, a single LB algorithm implementation as ICI-LB degrades the energy efficiency because it uses often power adaptation to counter resource unavailability. It is worth to note that, for Vehicular clients (high mobility), the configuration developed in [26], where authors advocate an Energy-Efficient adaptive resources management for real-time cloud services, can be investigated for implementation in the set of schedulers (TABLE II) used for deployment with AST approach.

VI. CONCLUSION

This paper has dealt with the load balancing issue to improve the spectrum efficiency required by the NGMN alliance in the context of 2020 and beyond. By realizing that the solver performances (EE and SE) vary according to network conditions, AST has been proposed as an alternative solution in order to optimize simultaneously the spectrum efficiency and the energy efficiency. In particular, we proposed the full load balance concept operated between two cells belonging to different operators. We have seen through simulations that differently from intra-RAN SON cases, the proposed scheme optimizes permanently the system. Based on SDN theory, AS-LB makes the system scalable and energy efficient, which is actually an important network performance metric indicator in view of the green society.
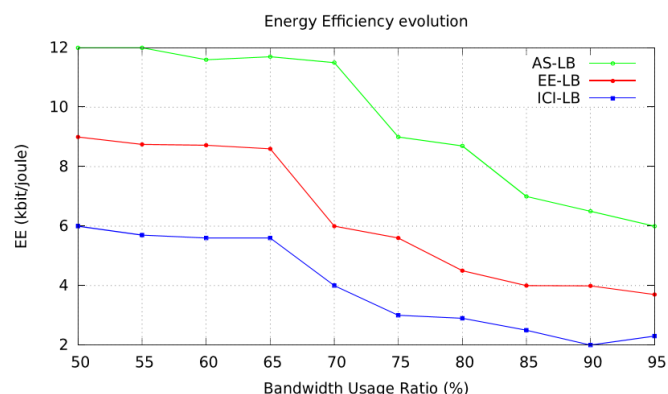

Fig. 9. Energy Performance vs. Bandwidth

REFERENCES

[1] NGMN Alliance, "NGMN 5G White paper," Feb. 2015.
[2] L. Dai et al., "Non-Orthogonal Multiple Access for 5G: Solutions, Challeng- es, Opportunities, and Future Research Trends," IEEE Commun. Mag., vol. 53, no. 9, doi: 10.1109/MCOM.2015.7263349 Sept. 2015, pp. 74–81
[3] 3GPP. (2011) TR 36.902 self-configuring and self-optimizing network (SON) use cases and solutions, Rel-9. [Online]. Available: http://www.3gpp.org
[4] Li Z., Wang H., Pan Z., Liu N., You X., "QoS and channel state aware load balancing in 3GPP LTE multi-cell networks", Science China, Vol. 56, 042309:1–042309:12 doi: 10.1007/s11432-013-4820-y, April 2013.
[5] Rangisetti A. K., Pasca T. V. S., Tamma B. R, QoS Aware load balance in software defined LTE networks, Computer Communications, Volume 97, 1 January 2017, doi: 10.1016/j.comcom.2016.09.005, Pages 52-71, ISSN 0140-3664
[6] Gang C., Fanfan M., Li S., "Qos-Priority Based Load Balancing Agorithm for LTE system with mixed user", Journal of China University of Post and Telecommunications, 22(3):9-17, doi: 10.1016/S1005-8885(15)60647-1, June 2015.
[7] Xinyu G., Shucong J., Wenyu Li., Lin Z., "Energy Efficient Load Balancing in LTE Self- Organization Network", 2013 IEEE 24th International Symposium on Personal, Indoor and Mobile Radio Communications: Workshop on Self-Organizing Networking (SON) in HetNets, doi: 10.1109/PIMRCW.2013.6707844, pp 96-100.
[8] Giovanidis A., Liao Q., Stanczak S., "A Distributed Interference-aware Load balancing Algorithm for LTE Multi-Cell Networks" , 2012 International ITG Workshop on Smart Antennas (WSA), doi: 10.1109/WSA.2012.6181222, pp 28-35
[9] Yamamoto T., Komine T., Konishi S., "Mobility Load Balancing Scheme based on Cell Reselection", ICWMC 2012 : The Eighth International Conference on Wireless and Mobile Communications, pp 381-387, INRIA.
[10] A. Reaz R. Boutaba, "Design Consideration for Managing Wide Area Software Defined Networks," IEEE Communication. Mag., vol. 52, issue 7, 2014, doi: 10.1109/MCOM.2014.6852092, pp. 116–123.
[11] J. Menglan, X. Dionysis, C. Salvatore, P. Nikos, M. Toktam, Radio Resource Sharing as a service in 5G: A software-defined networking approach, Journal of Computer Communications, Volume 107, 2017, doi : 10.1016/j.comcom.2017.03.006, Pages 13-29, ISSN 0140-3664
[12] A. Ian F, S. Chun Lin, P. Wang, Wireless sofware-defined networks (W-SDN) and network function virtualization (NFV) for 5G cellular systems: An overview and qualitative evaluation, Journal of Computer Networks, vol 93, 2015, doi: 10.1016/j.comnet.2015.10.013 pp 66-79
[13] P. Marsch et al., "5G Radio Access Network Architecture: Design Guidelines and Key Considerations," in IEEE Communications Magazine, vol. 54, no. 11, pp. 24-32, November 2016. doi: 10.1109/MCOM.2016.1600147CM
[14] B. R. Sonia, N. Nidal, T. Sami, A novel resource allocation scheme for LTE network in the presence of mobility, Journal of Network and Computer Applications, vol 46, No 11 (2014), doi, 10.1016/j.jnca.2014.07.017, pp 352-361.
[15] N. Tabia, A. Gondran, O. Baala, A. Caminada. Interference model and evaluation in LTE networks. WMNC 2011, 4th Joint IFIP Wire- less and

Mobile Networking Conference, Oct 2011, Toulouse, France. pp 1-6, 2011, doi: 10.1109/WMNC.2011.6097237. <hal-00938517>

[16] Oh S., Kim H., Na J., Kim Y., Kwon S. (2016) Mobility Load Balancing Enhancement for Self-Organizing Network over LTE System. In: Galinina O., Balandin S., Koucheryavy Y. (eds) Internet of Things, Smart Spaces, and Next Generation Networks and Systems. NEW2AN 2016. Lecture Notes in Computer Science, vol 9870. Springer, doi: 10.1007/978-3-319-46301-8_17 Cham

[17] Tang J., So D. K. C, Alsua E., Hamdi K. A., "Ressource Efficiency: A New Paradigm on Energy Efficiency and Spectral Efficiency Tradeoff," IEEE Transactions on Wireless Communications, Vol 13., NO. 8. AUG 2014, doi: 10.1109/TWC.2014.2316791.

[18] Markana A., Nitin P., Kannan M., "Lexicographic optimization based MPC: Simulation and experimental study, Computers & Chemical Engineering", Volume 88, 8 May 2016, doi 10.1016/j.compchemeng.2016.02.002, Pages 135-144, ISSN 0098-1354

[19] Abraham P. Punnen, Y.P. Aneja, "Lexicographic balanced optimization problems", Operations Research Letters, Volume 32, Issue 1, January 2004, doi: 10.1016/S0167-6377(03)00065-8, Pages 27-30, ISSN 0167-6377

[20] B. A. Salihu et al, Scheduling in Interference-Limited Environment for LTE-A Systems, International Journal of Future Generation Communication and Networking, Vol.7, No.5 (2014), doi: 10.14257/ijfgcn.2014.7.5.15, pp. 179-190.

[21] L. Su, P. Wang., F Liu. , Particle swarm optimization based resource block allocation algorithm for downlink LTE systems. Presented at Asia Pacific Conference on Communication (APCC), (2012, October), doi: 10.1109/APCC.2012.6388227.

[22] X. Xiao, X. Tao and J. Lu, "QoS-Aware Energy-Efficient Radio Resource Scheduling in Multi-User OFDMA Systems," in *IEEE Communications Letters*, vol. 17, no. 1, doi: 10.1109/LCOMM.2012.112012.121910 pp. 75-78, January 2013.

[23] R. Schoenen, W. Zirwas and B.H Walke., "Capacity and coverage analysis of a 3 GPP-LTE multihop deployment scenario," Proc. IEEE International Conference On Communications Workshops, May 2008, doi: 10.1109/ICCW.2008.11.

[24] G. Takoua, J. Badii, "Energy consumption evaluation for LTE scheduling algorithms". Presented at International symposium on Network, computer and communications (ISNCC), (2015, May), doi: 10.1109/ISNCC.2015.7238594.

[25] N. Cordeschi, A Danilo, M Shojafar, E. Baccarelli, "Distributed and adaptive resource management in Cloud-assisted Cognitive Radio Vehicular Networks with hard reliability guarantees", Vehicular Communication, Vol. 2, Issue 1, doi: 10.1016/jvehcom.2018.08.004, pp 1-12, 2015.

[26] M Shojafar, N. Cordeschi, E. Baccarelli, " Energy-efficient Adaptive Resource Management For Real-time Vehicular Cloud Based Services", IEEE Transactions on Cloud Computing, Volume PP, Issue 99, 2016, doi: 10.1109/TCC.2016.2551747

[27] Chavez L. J., Eichemberger V. M., Garcia I. C., Madeira E. R. M., "Integrating Openflow to LTE: Somme Issues Towards Software-Defined Mobile Networks" 7th IEEE International Conference on New Technologies, Mobility and Security (NTMS), 2015, September, Paris.

[28] Nsnam.org

Ouadoudi Zytoune received his Ph.D. in 2010 from the University of Mohammed V - Agdal, Rabat, Morocco. He is with the ENSA (National School of Applied Sciences) of Kenitra, Ibn Tofail University, Morocco. He became an HDR Professor on 2014; he is a member of LGS (Laboratoire Genie des Systèmes) lab. His current research interests include Wireless Communications, QoS in Wireless communication and Wireless Sensor Networks.

Mohamed Yahyai received his PhD from the University of Metz in France in 1995 and has joined the Faculty of Sciences at Ibn Tofail University since this year. His research interests include geometry and mathematical analysis. He managed several teams (Geometry; functional analysis, complex analysis, harmonics and potential; mathematical modelling) within the Laboratory of analysis, geometry and applications in the 2010-2014 session of accreditation. From 2000 to 2012, he taught algebraic and mathematical analysis lectures and became the chief of the mathematical and computer science department in 2012. He is actually the vice Dean of the faculty of sciences at Ibn Tofail University, in charge of pedagogic affairs. Phone : +212 6 13 72 07 77.

Fall Hachim is at the Mathematics and Computer Sciences Department at Ibn Tofail University. He obtained his master degree in 2010 in Micro- electronics from Ibn Tofail University and his PhD in March 2018 from Ibn Tofail University. He is actually with the Algebraic and Geometrical Analyze Laboratory at Ibn Tofail University. He is working on the wireless cellular fifth generation Radio Access Network (RAN). His fields of interest are Wireless Cellular Network, Radio Resource Management in Wireless Cellular Software-Defined Networking. Phone : +212 675038976.