TRANSACTIONS ON EMERGING TELECOMMUNICATIONS TECHNOLOGIES
Trans. Emerging Tel. Tech. 2016; **27**:1000–1015
Published online 16 May 2016 in Wiley Online Library (wileyonlinelibrary.com). DOI: 10.1002/ett.3043

RESEARCH ARTICLE

A multi-objective performance modelling framework for enabling self-optimisation of cellular network topology and configurations

Hasan Farooq[1]*, Ali Imran[1] and Adnan Abu-Dayya[2]

[1] Department of Electrical & Computer Engineering, University of Oklahoma, Tulsa, OK, USA
[2] Qatar Mobility Innovations Center (QMIC), QSTP, Doha, Qatar

ABSTRACT

Cellular system optimisation, a cornerstone of cellular systems paradigm, requires new focus shift because of the emergence of plethora of new features shaping the cellular landscape. These features include self-organising networks with added flavours of heterogeneity of cell sizes and base station types, adaptive antenna radiation patterns, energy efficiency, spatial homogeneity of service levels and focus shift from coverage to capacity. Moreover, to effectively tackle spatiotemporal dynamics of network conditions, a generic low-complexity framework to quantify the key facets of performance, that is, capacity, quality of service and energy efficiency of the various network topology configurations (NTC), is needed for enabling self-organising networks empowered cellular system optimisation on the fly. In this paper, we address this problem and present a performance characterisation framework that quantifies the multiple performance aspects of a given heterogeneous NTC through a unified set of metrics that are derived as function of key optimisation parameters and also present a cross comparison of a wide range of potential NTCs. Moreover, we propose a low-complexity heuristic approach for holistic optimisation of future heterogeneous cellular systems for joint optimality in the multiple desired performance indicators. The performance characterisation framework also provides quantitative insights into the new tradeoffs involved in optimisation of emerging heterogeneous networks and can pave the way for much needed further research in this area.

***Correspondence**
H. Farooq, Department of Electrical & Computer Engineering, University of Oklahoma, Tulsa, OK, USA.
E-mail: hasan.farooq@ou.edu

Received 2 October 2015; Revised 4 February 2016; Accepted 7 April 2016

1. INTRODUCTION

As the spectral efficiency per link for LTE is approaching the theoretical Shannon limit, it is envisaged that the network densification by small cells (SCs) is among the most promising solutions for realising ambitious goals of infinite capacity and zero latency provision in future 5G networks. To efficiently manage such an ultra-dense, complex, heterogeneous cellular networks, the paradigm of self-organising networks (SON) has recently been investigated heavily to automate cellular system management and maintenance tasks [1–3]. This SON capability allows the cellular network to monitor the key performance indicators (KPIs) and optimise network parameters to adapt itself to spatiotemporal dynamics of networks conditions. This dynamicity of network conditions includes change of traffic patterns over the course of day, relocation of hot spots and cell outages. For example, Figure 1 shows SON-enabled cellular network and variations of traffic pattern observed in real cellular network's data. These variations, which are exhibited through the KPIs and can further be estimated using Minimization of Drive Test reports [4], prompt the SON engine to test each of the possible NTC in a static or dynamic simulator to come up with a new network topology configurations (NTC) that meet specific objectives like spectral efficiency, energy efficiency, quality of service (QoS) or a combination of these. Leveraging on the modern capability of turning base stations (BSs) on and off and smart antennas radiation patterns, SON engine can adapt projected number of sectors, frequency usage and number of SCs on the fly to achieve the desired objectives. In a real network, there are hundreds of possible network parameters configurations (i.e. large search space) characterised by the types of BS, number of sectors per site, number of SCs per site and the frequency reuse, in addition to other configuration parameters including locations, tilts, azimuths and heights. In this ever changing traffic landscape of cellular environment, by the time SON engine comes up with optimum network configuration, the

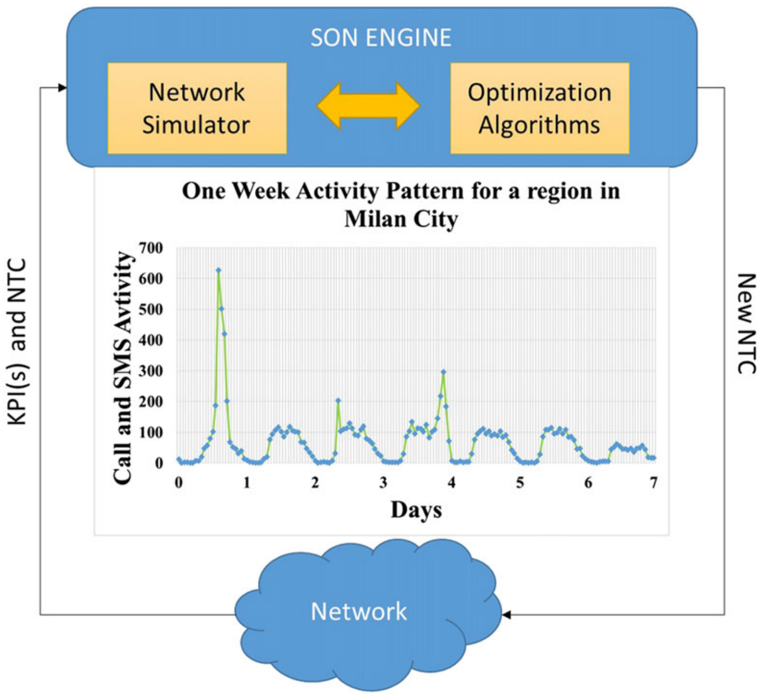

Figure 1. SON engine has to cope with frequent activity variations observed in real cellular network. SON, self-organising network; KPI, key performance indicator; NTC, network topology configuration.

scenario might have already changed, and NTC becomes outdated. This demands for low-complexity performance estimation and then optimisation techniques to cope with the spatio-temporal dynamics of cellular environment in agile fashion. Increasing scarcity of spectrum for LTE is pushing towards more aggressive frequency reuse, leading to new kinds of spectrum reuse, for example, intra-site spectrum reuse [5, 6] that has to be incorporated with SON optimisation objectives. Also, fuelled by performance criteria set forth by 3GPP where spatial fairness of data rate received by the cell edge and cell centre users is being given more and more importance [7], QoS metric cannot be neglected anymore. Similarly in the wake of rising cost of energy and environmental concerns, energy efficiency has also become an important metric [8].

The ambitious goals of zero latency [9] in envisioned future cellular systems require low-complexity cellular system optimisation (CSO) framework to provide agile online multi-objective optimisation of potential NTCs that can judiciously strike the intended balance among the various conflicting goals such as capacity, QoS and energy consumption while taking into account operator's policy. The need for potential search of NTCs, though well-conceived in [10–13], is not fully stated yet, particularly in context of SCs-enhanced cellular system (SC-CS) heterogeneous networks. Another challenge in enabling and evaluating many of the SON use cases in heterogeneous networks is the lack of unified performance quantification framework that can quantify cellular system performance in terms of the aforementioned KPIs. This paper addresses this need by presenting and analysing a holistic framework to quantify the three key KPIs, namely, capacity, QoS and energy efficiency. This framework can act as a key enabler for a number of SON use cases such as capacity and coverage optimisation, inter-cell interference coordination, energy efficiency and load balancing.

1.1. Prior Works

For cellular networks, most of the prior research works on optimisation of network parameters [14–33] use different definitions of a given KPI, for example, coverage and capacity [19–25], QoS [26, 27], cost-efficiency [28] or energy efficiency [16, 29–31], to optimise a single network parameter, for example, BS location [14, 30, 32, 34], or few other parameters such as antenna tilts [22], sectorisation [15, 33] and frequency reuse [17, 18]. Moreover, these KPIs metrics to be used by the SON engine should be able to quantify the long-term average performance of a cellular system by incorporating its dependencies on NTC parameters, while generalizing or averaging out the short-term dynamics of cellular eco-system. An additional requirement is that the metric should be evaluable by the SON engine without resorting to complex dynamic simulators. More precisely, to the best of our knowledge, no previous work has provided a framework to enable a cross comparison among potential NTCs, simultaneously in terms of capacity, QoS and energy efficiency, while taking into account the key deployment factors such as number of sectors per site, number of SCs per site and different variants of intra-site frequency reuse the emerging cellular systems can avail. Furthermore, tradeoff between the capacity and spatial fairness of service level in the coverage area is relatively overlooked. In view of the increasing emphasis by 3GPP on better cell edge throughput rates and better spatial fairness of achievable data rates [27], we also use the proposed performance characterisation framework (PCF) to investigate this under-explored but important tradeoff that various NTCs offer. The presented analysis can be leveraged to design NTCs that can strike operator-intended precise balance in the capacity and spatial fairness while simultaneously taking into account the energy consumption aspect of the given NTC.

Additionally, because optimisation of network parameters is non-deterministic polynomial-time-hard problem, prior works in literature have generally addressed it using metaheuristics such as simulated annealing [34–36], particle swarm [37], genetic algorithms [25, 38], Taguchi's method [39] or ant colony optimisation [15] to obtain near-optimal solutions for a selected set of few parameters. The basic methodology that is generally followed in these works involves a detailed dynamic simulation model that acts as black box between the KPI and the potential parameters of a given NTC. The use of dynamic simulation-based models by the SON engine is not only time-consuming but also provides little insights into system behaviour. On the contrary, our approach builds on a mathematical model to couple the KPIs with the extensive set of NTC parameters and thus helps to obtain better insights into system behaviour. The resultant PCF makes easier the holistic cross comparison of various potential solutions of CSO problem.

2. BACKGROUND AND SYSTEM MODEL

2.1. CSO objectives and proposed solution approach

For emerging cellular systems, the optimisation problem has multiple target objectives like maximisation of capacity, coverage and fairness of service in the coverage area, spectral efficiency, spectrum reuse efficiency, throughput, minimisation of cost, energy consumption and/or outage and so on. However, all these objectives can be boiled down to three main categories of performance measures:

(1) Capacity-oriented performance measures: these include cellular capacity, spectral efficiency, spectrum reuse efficiency, throughput or goodput.
(2) QoS-oriented performance measures: rate fairness and outage are typical QoS measures.
(3) Cost-oriented performance measures: total cost of ownership of a cellular system over its life has three further major factors:
   (a) capital cost: cost of hardware, software and deployment labour cost;
   (b) maintenance cost: cost of labour required for operation, optimisation and maintenance of sites and the switching network; and
   (c) energy consumption: energy consumed to keep the cellular system running is increasingly becoming a very significant factor of operational cost.

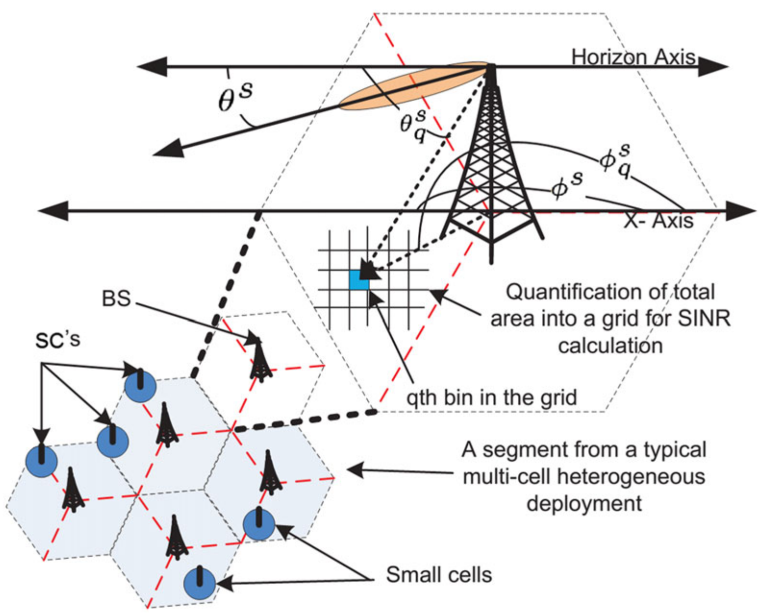

Figure 2. Generic system model used for SINR calculation. SINR, signal-to-interference-plus-noise ratio; BS, base station; SC, small cell.

Trans. Emerging Tel. Tech. **27**:1000–1015 (2016) © 2016 John Wiley & Sons, Ltd.
DOI: 10.1002/ett

In this paper, we derive the PCF to quantify each of the three listed aspects of performance of a cellular systems as function of NTC parameters. Under the cost-oriented performance, we only focus on the energy efficiency, as energy consumption has recently become highly important, particularly because of rising costs of energy and concern for CO_2 emissions. For the treatment of other two cost factors, interested reader is referred to previous works in [10], which deal with capital cost reduction by introduction of low-cost BSs (relays or femto or pico BSs) and [1] which provide a comprehensive review of self-organising networks as a major maintenance cost reduction approach.

The main idea of the proposed solution in this paper is that in order to cope with spatiotemporal changes in traffic or cellular environment, SON engine will dynamically switch to a suitable NTC, based on an adaptive utility function that incorporates major system objectives, for example, spectral efficiency, fairness and power consumption, and can prioritise among these objectives. To overcome size and complexity of holistic CSO problem, we propose to exploit a hybrid approach, that is, a detailed mathematical system model is first constructed and extensive system level simulations are performed to generate the whole solution space for all feasible NTCs consisting of number of sectors per site 'S', spectrum reuse factor 'F' and number of SCs per site 'R'. Because in a practical cellular system, possible combinations of 'S', 'F' and 'R' are not very large, and in fact, only configurations listed in Figure 3 are technically the most feasible ones, so SON engine can effectively search over this confined solution space easily and adapt the utility to set an optimisation target and switch to most suitable NTC in time-efficient manner.

2.2. System model and holistic CSO problem formulation

We consider a generic cellular system model as illustrated in Figure 2.

We divide the whole area to be optimised by SON engine into set of Q bins denoted by $\mathcal{Q}$, where q denotes q^{th} bin, such that $\sum_{q=1}^{Q} a_q = A$, and $\frac{A}{Q} = a_q, \forall q \in \mathcal{Q}$ where A is the total area and area a of the bin is so small that shadowing and path loss can be considered constant within it. Now, using the notation defined in Table I, the problem of holistic joint optimisation of the three performance objectives identified perviously can be formulated as a multi-objective optimisation problem:

$$\max_{\mathcal{Q}_b,\mathcal{Q}_r,H_r,H_s,S,R,P^s,P^r,\Upsilon_f,\theta} f(\Upsilon,\Lambda,\Omega) \tag{1}$$

subject to feasibility and range constraints on the optimisation parameters. The definition of the parameters in (1) is presented in Table I.

The expression in (1) is a holistic CSO problem in which the location of BS and SC, the number of sectors per BS, the number of SC per BS, the antenna heights, the transmission powers, the antenna azimuth, the antenna tilts and the frequency reuse have to be optimised to achieve the best possible performance in terms of all three KPIs. Sub-problems of such CSO problem have been shown to be non-deterministic polynomial-time-hard in a number of studies [37, 40–42]; therefore, generally metaheuristic techniques are utilised to partially explore the solution space of the CSO problem, in order to find an acceptable solution. From (1), we can obtain some useful insights into the solution space of the problem. Let us take a simple example of only $19 \times 3 = 57$ sectors cellular system and focus on solving for only one NTC parameter, for example, the optimal sector azimuth angle. With an over simplifying assumption that the azimuth can only take 10 possible values centred around the nominal azimuth of the sector, a brute force-based solution will have to search among 10^{57} possible azimuth angle combinations. If system level evaluation of the KPIs of interest as a function of azimuth angles that is generally carried through a simulation tool takes time τ_e (that can be in order of minutes),

Table I. Notation for system model.

Symbol	Description
b,s,r,q	As subscript or superscript denote association to base station, sector, small cell or q^{th} bin, respectively
B	Set of all base stations in systems where $\lvert\mathcal{B}\rvert = B$
A	Total area of interest
Q	Set of Q bins that constitute A
q	q^{th}bin, $\sum_{q=1}^{Q} a_q = A$, & $\frac{A}{Q} = a_q, \forall q \in \mathcal{Q}$
Q_b	Set of bins in which BS are located, $\mathcal{Q}_b \subseteq Q$
S	Set of all sectors in the systems,where $\lvert\mathcal{S}\rvert = S$
S_b	Total number of sectors b^{th} BS has
$\mathcal{S}$	$\mathcal{S} = \{S_1, S_2, S_3 \ldots S_B\}$, $S = \lvert\mathcal{S}\rvert = \sum_{b=1}^{B} S_b$
h_s	(Antenna) height of s^{th} sector antenna on BS
Υ_f	Number of times spectrum is reused within site
R_b	Number of SC in b^{th} BS
R	$\mathcal{R} = \{R_1, R_2, R_3 \ldots R_B\}$, $R = \lvert\mathcal{R}\rvert = \sum_{b=1}^{B} R_b$
Q_r	Set of bins in which SC are located, $\mathcal{Q}_r \subseteq Q$
H_r	Set of all SC antenna heights
h_r	Height of r^{th} SC antenna
$\boldsymbol{\phi}$	Vector of azimuth angles of all sectors
$\boldsymbol{\phi}^s$	Azimuth angle of s^{th} sector
θ	Set of tilt angles of all sectors
θ^s	Tilt angle of s^{th} sector
P^s	Set of transmission powers of all sectors
p^s	Transmission power from s^{th} sector
P^r	Set of transmission powers of all SC
p^r	Transmission power from r^{th} SC
G_q^s	Gain from the s^{th} sector antenna to q^{th} bin
$\boldsymbol{\alpha}$	Path loss co-efficient including long term shadowing
$\boldsymbol{\beta}$	Pathloss exponent
δ_q^s	Shadowing from s^{th} to q^{th} bin
φ_v	Vertical beamwidth of the antenna
φ_h^s	Horizontal beamwidth of s^{th} sector antenna
Υ	Capacity-wise KPI
Ω	Energy consumption-wise KPI
Λ	Service area fairness-wise KPI
$\mathcal{X}\backslash y$	Means all elements of $\mathcal{X}$ except y

finding an optimal solution may take as long as $\frac{10^{57}}{1/\tau_e}$! minutes. Obviously, the actual size of the solution space of a typical holistic CSO problem represented by (1) is far more gigantic.

If we apply one of the aforementioned evolutionary metaheuristics used in literature [14, 15, 17, 19, 21, 23, 24, 26–30, 32, 36–39], the search space C_p of the holistic CSO problem in (1) can be reduced by a factor ϵ. The solution still would require time $\tau = \frac{\left(\frac{C_p}{\epsilon}\right)}{1/\tau_e}$ minutes, and yet may not be guaranteed to be optimal. Contrary to most of the works in open literature on CSO, which propose variations and combinations of different metaheuristics to only increase ϵ to reduce the solution time τ, the framework that we present in this paper exploits a bi-pronged approach for increasing the efficiency of CSO process by reducing both τ_e and C_p instead. Firstly, through PCF, it eliminates the need for dynamic simulator needed for KPI evaluation at each iteration of search. This is expected to substantially reduce τ_e that will ultimately reduce the τ irrespective of the metaheuristic used to factorise C_p by ϵ. Secondly, by preponderance of (1), it is clear that different parameters have different significance in CSO. Building on further insights into this observation provided by the PCF, we propose a simple algorithm for holistic CSO problem that can substantially reduce the C_p itself. This bi-pronged approach can improve the quality of solutions obtained, by allowing conventional metaheuristic to be more thorough, while significantly reducing the complexity of the holistic CSO problem.

3. A PERFORMANCE CHARACTERISATION FRAMEWORK

In this section, we derive quantitative measures for the three KPIs of interest, that is, Υ, Λ and Ω in terms of the key NTC parameters, which can be evaluated with low complexity, that is, without resorting to black box type complex dynamic simulators. The signal-to-interference-plus-noise ratio (SINR) perceived in the q^{th} bin from s^{th} sector (Figure 2) can be given as follows:

$$\gamma_q^s = \frac{p^s G_q^s \alpha (d_q^s)^{-(\beta)} \delta_q^s}{\sigma^2 + \sum_{\forall s' \epsilon S} (p^{s'} G_q^{s'} \alpha (d_q^{s'})^{-(\beta)} \delta_r^{s'} . u(\Upsilon_f))} \quad (2)$$

where $\{s, s'\} \epsilon S$, $q \epsilon Q$ and $u\left(\Upsilon_f\right)$ are unit functions that determine whether or not the q^{th} bin will receive interference from a particular sector depending on the frequency reuse. Note that we assume full load scenario, that is, all sub-carriers allocated to a cell are simultaneously under use. With this assumption, in calculating SINR, the impact of dynamic scheduling can be omitted, and only static frequency reuse that is part of NTC can be used to determinethe inter carrier collision and hence interference at a given location. Here, d_q^s is distance the between the q^{th} bin andthe s^{th} sector antenna located in $q_b^{th} \in \mathcal{Q}_b$ bin, given by

$$d_q^s = \sqrt{\left(x_{q_b} - x_q\right)^2 + \left(y_{q_b} - y_q\right)^2 + (h_s - z_q)^2} \quad (3)$$

Three-dimensional antenna gain can be modelled as in [43]

$$G_q^s = G(\zeta, D) \times 10^{-1.2\left(\lambda_v\left(\frac{\theta_q^s - \theta^s}{\varphi_v}\right)^2 + \lambda_h\left(\frac{\varphi_q^s - \varphi^s}{\varphi_h^s}\right)^2\right)} \quad (4)$$

where θ_q^s is the vertical angle in degrees from s^{th} sector to q^{th} bin and can be given as $\theta_q^s = \arctan\left(\frac{h_s}{d_q^s}\right)$(Figure 2). ϕ_q^s is the horizontal angle in degrees on s^{th} sector to q^{th} bin with respect to positive x-axis. λ_h and λ_v represent the weighting factors for the horizontal and vertical beam pattern of the antenna in three-dimensional antenna model [44], respectively. As indicated in (4), the maximum antenna gain G is the function of antenna efficiency ζ and directivity D and can be written as $G = \zeta D$, where D can be further approximated as $D = \frac{4\pi}{\varphi_h^s \varphi_v}$.

For the practical cellular antennas, the relationship between the horizontal beamwidth of sector antenna and the number of sectors S_b per b^{th} BS site can be modelled as $\varphi_h^s = \frac{360}{\mu * S_b}$, where μ is a factor representing the overlap between the sectors. Thus, using (4) in (2), the SINR can be determined as in (5).

$$\gamma_q^s = \frac{P^s \alpha (d_k^s)^{-(\beta)} \delta_q^s . \left(\frac{4\pi\rho}{\left(\frac{360}{\mu * S_b}\varphi_v\right)}\right) .10^{-1.2\left(\lambda_v\left(\frac{\theta_q^s - \theta^s}{\varphi_v}\right)^2 + \lambda_h\left(\frac{\varphi_q^s - \varphi^s}{\left(\frac{360}{\mu * S_b}\right)}\right)^2\right)}}{\sigma^2 + \sum_{\forall s' \epsilon S}\left(p^{s'} \alpha (d_q^{s'})^{-(\beta)}\right) \delta_q^{s'} . \left(\frac{4\pi\rho}{\left(\frac{360}{\mu * S_b^{s'}}\varphi_v\right)}\right) .10^{-1.2\left(\lambda_v\left(\frac{\theta_q^{s'} - \theta^{s'}}{\varphi_v}\right)^2 + \lambda_h\left(\frac{\varphi_q^{s'} - \varphi^{s'}}{\left(\frac{360}{\mu * S_b^{s'}}\right)}\right)^2\right)} .u(\Upsilon_f)} \quad (5)$$

As desired, the SINR in (5) is the function of key parameters of a given NTC. Similarly, the SINR from r^{th} SC in q^{th} bin can be given as

$$\gamma_q^r = \frac{p^r \alpha \left(d_q^r\right)^{-(\beta)} \delta_q^r}{\sigma^2 + \sum_{\forall r' \in \mathcal{R}} \left(p^{r'} \alpha \left(d_q^{r'}\right)^{-(\beta)} \delta_q^{r'}\right)} \quad (6)$$

where $\{r, r'\} \in \mathcal{R}$ and $q \in \mathcal{Q}$. Note that for SC, the antenna gain can be assumed as unity; therefore, it is omit-

Trans. Emerging Tel. Tech. **27**:1000–1015 (2016) © 2016 John Wiley & Sons, Ltd.
DOI: 10.1002/ett

ted in the SINR expression. Also, because BS have much higher Tx powers than SC, SC have to duplex with BSs in time or frequency to avoid excessive interference from BS. With this assumption, only interference from other SC is considered in (6). Because frequency reuse of one is assumed among SC, therefore, no exclusive term to capture the frequency reuse as in (2) is needed in (6).

3.1. Quantifying Υ: reflecting capacity-wise performance from CSO perspective

We propose a metric, namely, effective spectral efficiency (ESE) to quantify capacity-wise performance denoted by Υ. This metric has semantics similar to the area spectral efficiency, but it does not require throughput estimation for its calculation; rather, it can be determined through simple semi analytical approach. A key advantage of ESE is that it can also serve as the basis for calculation of other two KPIs, that is, Λ and Ω. This is useful in modelling the coupling between these contradicting CSO objectives. In the succeeding discussion, we explain the calculation of ESE.

Because the sub-carrier bandwidth in emerging cellular systems (e.g. LTE) is fixed, so the throughput on single sub-carrier in a given BS-user link and hence the total throughput of the system depends on average achievable modulation coding efficiency (MCE) on each link in the system. Over the long term, the MCE in turn depends on SINR available on that link, whose long-term average value (in full load scenario as assumed previously) in turn depends mainly on NTC as derived in (5) and (6).

Let $\mathcal{L} = \{0, 1, 2, 3, \ldots L\}$ be the set of modulation coding schemes (MCS) available in the standard under consideration. MCE_l denotes the MCE of l^{th} MCS, $l = 0$ means an MCS with zero spectral efficiency, that is, no link, representing outage and L is the MCS with the highest spectral efficiency. Invoking the bin grid concept, an easily evaluable metric can be given as

$$\Upsilon_{MCE_e} = \sum_{l=0}^{L} \left(MCE_l \times \frac{Q_l}{Q} \right) \tag{7}$$

where

$$Q_l = \sum_{\forall q \in \mathcal{Q}} U_l \left(\gamma_q \right) \tag{8}$$

Here, γ_q denotes the SINR perceived in q^{th} bin from the best serving BS sector or SC (whichever is greater), and the unit function $U_l(\gamma_q)$ is defined as follows:

For $l \epsilon \mathcal{L} \setminus \{0, L\} L$:
$$U_l(\gamma_q) = \begin{cases} 1 & T_{l-1} < \gamma_q < T_{l+1} \\ 0 & otherwise \end{cases}$$
For $l = L$: $U_l(\gamma_q) = \begin{cases} 1 & T_{l-1} < \gamma_q \\ 0 & otherwise \end{cases}$

And for $l = 0$: $U_l(\gamma_q) = \begin{cases} 1 & \gamma_q < T_0 \\ 0 & otherwise \end{cases}$

where T_l is the threshold SINR required to use l^{th} modulation and coding scheme from set $\mathcal{L}$ and T_0 is the threshold of minimum γ in the succeeding discussion, which link cannot be maintained with the lowest modulation and coding pair implemented in the standard and all such points in coverage area constitute the outage area. Note that

$$\sum_{l=0}^{L} Q_l = Q \tag{9}$$

A key advantage of quantifying spatial spectral efficiency in this manner is that it has the potential to reflect geographical areas of high importance with weighting factors to pronounce their importance in capacity optimisation and reflect them in the ESE measure proportionally. This provides freedom to tailor this KPI for CSO process in order to reflect operator's policy. For setting different coverage priorities for different regions, Q in (7) can be replaced with the sum of weights associated with each bin, that is,

$$\Upsilon_{MCE_w} = \sum_{l=0}^{L} \left(MCE_l \times \frac{\sum_{i=l'}^{l} w_{q_{l'}}}{\sum_{q=1}^{Q} w_q} \right) \tag{10}$$

where Υ_{MCE_w} denotes weighted average MCE and w_q denotes weight assigned to the q^{th} bin in proportion to its relative importance in the area of interest. Thus, these weights can be used to model QoS requirements of different demographic groups or differentiate areas with different user densities. $w_{q_{l'}}$ denotes the weight of q^{th} bin using l'^{th} MCS, where $l' \in \mathcal{L}$. If not enough data is available so that precise weight to individual bins can be assigned and operator in general wants to make sure that the spatially fair data rates are available throughout the coverage area, instead of using arithmetic mean in (7), harmonic mean can be used. Unlike the arithmetic mean, the harmonic mean will aggravate the impact of bins with low spectral efficiency and will damp down the impact of bins having very large spectral efficiency, while representing the overall spectral efficiency of system. In this case,

$$\Upsilon_{MCE_h} = \frac{Q}{\sum_{q=1}^{Q} \left(\frac{1}{MCE_q} \right)}, MCE_q > 0 \tag{11}$$

where Υ_{MCE_h} denotes harmonic mean spectral efficiency in the area of interest and MCE_q denotes the spectral efficiency achievable in q^{th} bin based on the SINR γ_q perceived in that bin. Note that unlike Υ_{MCE_e}, Υ_{MCE_h} cannot take into account the outage in the coverage area. While Υ_{MCE} reflect link spectral efficiencies achievable with a particular NTC and can be used as an aspect of capacity, for holistic quantification of capacity, an important means of cellular capacity, that is, spectrum reuse also has to be taken into account.

In the backdrop of need for aggressive frequency reuse, we propose to reuse spectrum within a site. By exploiting the fact that aggressive sectorisation can provide significant isolation among cells projected from same BS, spectrum can be reused within a site among sectors pointing in opposite directions as well as among alternative sectors pointing in different directions as illustrated in various NTCs sketched in Figure 3. To quantify the spectrum reuse gain in capacity obtained from such spectrum reuse, we define Υ_f as 'number of times spectrum is reused within a site'. Thus, Υ_f can be calculated as

$$\Upsilon_f = \begin{cases} \rho^b \times \frac{S}{F} + \rho^r \times R & \text{if R} > 0 \\ \frac{S}{F} & \textit{otherwise} \end{cases} \tag{12}$$

where ρ^b and ρ^r are the factors with which spectrum is shared between BS and SC such that $\rho^b + \rho^r = 1$. F is the number of parts into which spectrum allocated to the BS (excluding the spectrum allocated to SC) is divided. Although intra-site spectrum reuse is expected to increase interference thus decrease Υ_{MCE}, it would be interesting to investigate how gain in capacity through higher Υ_f trades against the loss in capacity because of lower Υ_{MCE}. To incorporate the impact of both of these factors in the cellular capacity, we define the desired capacity-wise KPI named ESE as

$$\Upsilon = \Upsilon_{MCE} \times \Upsilon_f \tag{13}$$

Υ_{MCE} can be modelled using (7), (10) or (11) depending on the CSO objectives and service priorities of operator. Υ_{MCE} effectively reflects the capacity gain via spectral efficiency. Υ_f on the other hand essentially reflects capacity gain via spectrum reuse efficiency that might come from intra-site frequency reuse (or inter site frequency reuse or even fractional frequency reuse, not covered in this paper). Thus, the Υ quantifies the intended capacity-wise KPI from CSO perspective by incorporating the effect of key NTC factors.

3.2. Quantifying Λ: reflecting service area fairness from CSO perspective

From CSO perspective, the QoS has two aspects: (i) achievable data rates and (ii) spatial fairness of achievable data rates. An explicit metric to quantify only the second aspect is needed from CSO perspective, as first aspect is already covered in our definition of Υ. However, for an appropriate measure of fairness that has to be used in CSO process as an optimisation objective, we have to significantly depart from the conventional notion of fairness that is considered when designing very short time scale adaptive mechanisms, for example, scheduling or power allocation. Generally for long-term traffic variations, such short-term dynamics can be generally neglected as they are

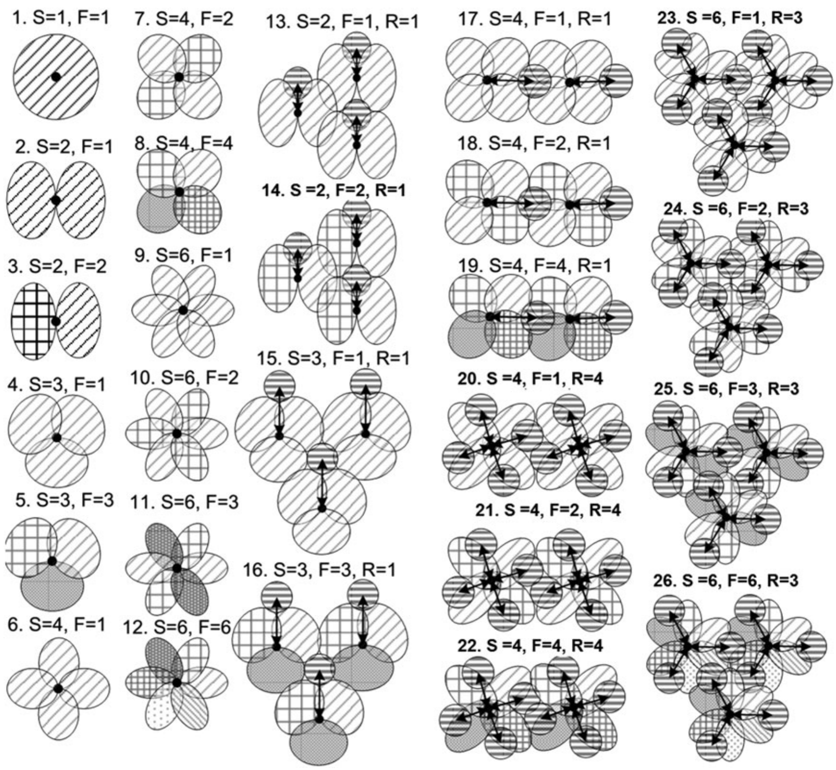

Figure 3. Twenty six different NTCs with varying S, F and R, which are investigated in this paper. Dots in the centre of each site represent base station locations. Oval shapes represent sectors and small circular shapes represent small cells attached to a site. Filling patterns represent frequency reuse pattern, whereas arrows represent backhaul links between base stations and small cells.

Trans. Emerging Tel. Tech. **27**:1000–1015 (2016) © 2016 John Wiley & Sons, Ltd.
DOI: 10.1002/ett

averaged out. Thus, it is fairness in space rather than classic fairness in time that is dependent more heavily on NTC and therefore has to be considered and evaluated during the CSO. More precisely, this spatial fairness of data means the homogeneity of the level of service that can be provided in the coverage area. We build on derivations in last section and define a metric to reflect the service area fairness (SAF) effectively as the inverse of the standard deviation of the spatial distribution of MCE as

$$\Lambda = 1/\sqrt{\frac{1}{Q}\sum_{q=1}^{Q}\left(MCE_q - \sum_{l=0}^{L}\left(MCE_l \times \frac{Q_l}{Q}\right)\right)^2} \quad (14)$$

Note that, similar to ESE, SAF can also be evaluated using the SINR expressions derived earlier. Having explicit spatial connotation instead of temporal, SAF gives the cell edge users judiciously higher importance because more bins lie farther from the cell centre. Thus, the advantage of SAF is that it is capable to explicitly capture the cell-centre and cell-edge rate disparity. In case, a finite bound-based estimation of SAF is required; Jain's fairness index can also be adapted to estimate the fairness of the service area as follows [45] :

$$JSAF = \frac{\left(\sum_{q=1}^{Q} MCE_q\right)^2}{N\sum_{q=1}^{Q}\left(MCE_q\right)^2} \quad (15)$$

3.3. Quantifying Ω: reflecting energy consumption-wise performance from CSO perspective

Energy consumption in cellular system has many complicated and interrelated components. Considering the scope of this paper, we focus on five selected elements of NTC that mainly determine the energy consumption of given NTC, that is, types of access points (BS or SC), number of sectors per site, and number of SC per site, transmission powers and sector overlap. These are the main parameters that make energy consumption in various cellular system NTCs different from each other. To this end, we model power consumption on a site while incorporating both fixed and variable power consumption per site that in turn depends on the type of BSs. Fixed power consumption is the power that is consumed in keeping the circuitry of BS sectors alive no matter if there is traffic or not. Fixed power remains non-zero until all sectors and SCs associated to a BS are completely switched off. Variable power consumption is the power required for transmission on air interface and varies with the traffic load. Thus, total power consumption in b^{th} BS site (including that of all sectors and associated SCs) can be written as (16), where subscripts f, v and t denote fixed, variable and transmission powers, respectively. For sake of simplicity, we do not consider any stray losses, for example, feeder loss, connectors loss as they are negligible for the purpose of this analysis. Variable power consumption within each sector or SC further depends on the transmission power p_t^s and p_t^r, traffic loading factors for sectors and SCs (between 0 to 1) η^s and η^r, respectively, and antenna gain G of sector and SCs, respectively. Antenna gain is further a function of antenna efficiency ζ and directivity D. The directivity of the antenna determines its gain and hence the transmission power required to provide a certain coverage and service level.

$$p = \sum_{s=1}^{S_b}\left\{p_f^s + p_v^s\left(G\left(\zeta^s, D^s\right), p_t^s, \eta^s\right)\right\} + \sum_{r=1}^{R_b}\left\{p_f^r + p_v^r\left(G\left(\zeta^r, D^r\right), p_t^r, \eta^r\right)\right\} \quad (16)$$

It can be written as

$$D = 4\pi/\left(\frac{\int_0^{2\pi}\int_0^{\phi} f(\theta,\phi)\sin\theta d\theta d\phi}{f(\theta,\phi)\,|_{\max}}\right) \quad (17)$$

where $f(\theta,\phi)$ is function representing radiation pattern of antenna as function of spherical co-ordinate angles θ and ϕ. For almost all commercial antennas used in cellular systems, the denominator of (17) can be approximated by-product of half power beam widths φ_h and φ_v in horizontal and vertical plane [46]. Thus, (17) can be approximated as

$$D \approx \frac{4\pi}{\varphi_h\varphi_v} \quad (18)$$

In commercial cellular systems, the typical vertical beam width of antenna is around $\varphi_v \approx \pi/18$ radians, and horizontal beam width depends on the number of sectors per access point. For BS with three sectors and six sectors, beam width of around 70° and 35° are generally used respectively. Using μ defined earlier as the factor determining the overlap between the adjacent sectors, we can write horizontal beam width as a function of S_b as $\varphi_h = \mu\pi/S_b$. Using these values of φ_h and φ_v, (18) can be written as

$$D \approx \frac{72S_b}{\mu\pi} \quad (19)$$

Typical value of μ can be assumed to be $\mu = 1.1$. To achieve a desired effective isotropic radiated power (EIRP) in the coverage area, less transmission power p_t will be required for antennas with higher gains as

$$EIRP = \zeta \times D \times p_t \quad (20)$$

If p_d is the power required to achieve desired $EIRP_d$ with an omnidirectional antenna

$$p_d = \frac{EIRP_d}{\zeta D} \tag{21}$$

Therefore, for given coverage level, if more sectors per site are used, less transmission power per sector would berequired because of high directivity and hence higher gains of the antennas. Thus, the variable circuit power per sector for desired $EIRP_d$ can be written in dB as

$$p_v^s = 10\log_{10} p_d^s - 10\log_{10}\left(\frac{4\zeta^s S_b}{\mu\varphi_v}\right) + 10\log_{10}\eta^s \tag{22}$$

Similarly, the variable circuit power on a SC can be written as

$$p_v^r = 10\log_{10} p_d^r - 10\log_{10}\left(\frac{4\zeta^r}{\varphi_v}\right) + 10\log_{10}\eta^r \tag{23}$$

Substituting (22) and (23) to (16) and re-arranging, we get (24).

$$\Omega = \left(\sum_{s=1}^{S_b}\left\{p_f^s + \mu\left(\frac{\eta^s\varphi_v^s p_d^s}{4\zeta^s S_b}\right)\right\} + \sum_{r=1}^{R}\left\{p_f^r + \frac{\eta^r\varphi_v^r p_d^r}{4\zeta^r}\right\}\right) \tag{24}$$

$$\frac{\Omega}{\Upsilon} = \frac{\left(\sum_{s=1}^{S_b}\left\{p_f^s + \mu\left(\frac{\eta^s\varphi_v^s p_d^s}{4\zeta^s S_b}\right)\right\} + \sum_{r=1}^{R}\left\{p_f^r + \frac{\eta^r\varphi_v^r p_d^r}{4\zeta^r}\right\}\right)}{\sum_{l=0}^{L}\left(MCE_l \times \frac{Q_l}{Q}\right) \times \Upsilon_f} \tag{25}$$

Equation (24) provides a simple metric to quantify the power consumption in a NTC as a function of number of sectors per site, number of SC per site and transmission powers and sector overlap and antenna beamwidths. This metric also takes into account an additional factor, that is, the traffic load factor which is not direct part of NTC but can affect the power consumption heavily. The split ratio between the fixed power consumption and transmission power can be used to model various BS types as well. On the other hand, Equation (25) provides metric to quantify the long-term average energy efficiency in $\frac{Joules}{bits}$ for a given NTC.

4. PERFORMANCE EVALUATION OF DIFFERENT NTCS

In this section, we evaluate the performance of a range of potential NTCs using the PCF.

4.1. System model for performance evaluation

A total of 26 NTCs with generally feasible combinations of key NTC parameters F, S and R (Figure 3) are evaluated, while other parameters are kept fixed at values listed in table II of [11]. Two tiers of cells are modelled for each NTC to consider realistic amount of interference in multi cellular scenario. Shadowing and appropriate path loss models for BS and SC similar to [47] are used in order to model a realistic cellular system environment. In SC-CS, SCs are located at half of inter site distance where the SINR is minimum, that is, where the far end corners of adjacent sectors join. To map the SINR γ_q to the long-term average link spectral efficiency, we refer to SINR thresholds for MCSs used in LTE. A given NTC is denoted by number of sectors per site S, frequency reuse F and number of SC per site R. Thus, for example, NTC denoted by '25. S=6,F=3, R=3' means NTC No. 25 has six sectors per site and spectrum allocated to BS (after splitting with SC) is divided in three equal parts, each part is allocated to three adjacent sectors, and the pattern is repeated for other three sectors on the site such that sectors using the same spectrum are pointing apposite to each other. And the site has three SCs. Thus, in NTC 25, the spectrum is reused $\Upsilon_f = \rho^b \times \frac{S}{F} + \rho^r \times R = 0.5 \times \frac{6}{3} + 0.5 \times 3 = 2.5$ times within a site area. For brevity, onward analysis will use Υ_{MCE_e} as measure of capacity and Λ as a measure to reflect SAF.

4.2. Analysing capacity-wise performance

Figure 4 plots Υ, Λ and Ω evaluated for all 26 NTCs under consideration, normalised by their maximum values. It can be seen that different NTCs offer different tradeoff among different KPIs. For ease of discussion while probing into these tradeoffs, we first focus on NTC 9–12, all with $S = 6$. It can be seen that from NTC = 9 to NTC = 12, as frequency reuse is made less tight with other parameters being fixed, the overall capacity of system, that is, Υ still increases (Figure 4). This is because the increase in Υ_{MCE_e} due to decreased interference overweighs the loss in Υ_f. Hence as a net result Υ is larger in NTC = 10, 11,12 compared to NTC = 9. However, there is payoff of this gain. It can be seen that Λ (i.e. SAF) continuously decrease from NTC = 9 to NTC = 12. The reason for this will be discussed in the next subsection. By comparing the Υ for SC-CS with those for macro-cells only cellular system, it can be easily seen that SCs bring a significant improvement in overall capacity. This improvement is due to two reasons: because of much smaller height and lower transmission power of SC cause and suffer from much lesser interference resulting in better Υ_{MCE_e}. Secondly, in addition to higher Υ_{MCEe}, there is another positive contribution of SC towards higherΥ that is explained as follows: let us assume that three SC are working in a cell. In this case, spectrum is divided into two parts for sharing between BS and SC. This reduces Υ_f by half only compared with scenario with three sectors where Υ_f will be reduced by a factor of three. These two

Trans. Emerging Tel. Tech. **27**:1000–1015 (2016) © 2016 John Wiley & Sons, Ltd.
DOI: 10.1002/ett

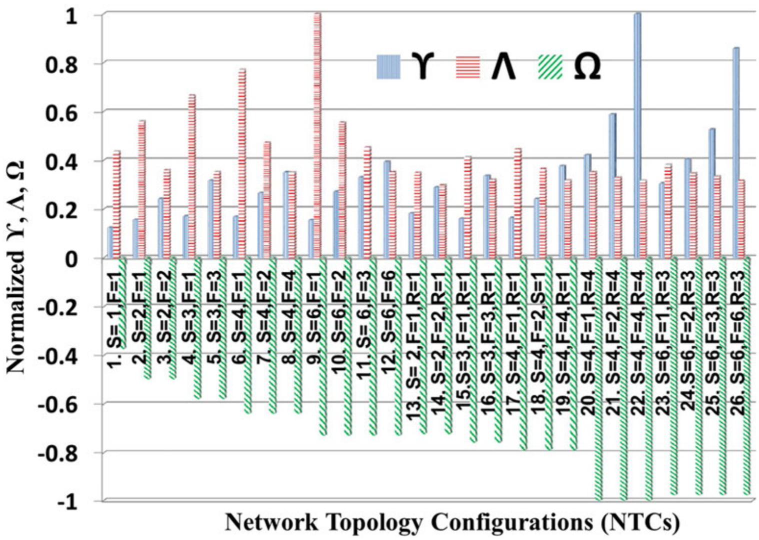

Figure 4. Comparison of different NTCs in terms of their capacity Υ, service area fairness Λ and power consumption Ω. NTC, network topology configuration.

reasons together make SC more advantageous method to boost capacity compared with adding more sectors. However, there is payoff for this gain in capacity achieved by SC in terms of both SAF as well power consumption. It can be seen from Figure 4 that SC-CS in general has lower SAF and higher power consumption compared with macro-cells only cellular system.

4.3. Analysing service area fairness

From the results in Figure 4, it can be noted that SAF increases with increase in number of sectors, but it decreases with increase in F (or in other words decrease in Υ_f). This is because increasing the number of sectors in general decrease the cell edge interference thus makes geographical distribution of data rates more uniform in a cell. In other words low Υ_f means less intera-site reuse, and therefore, interference mostly comes from adjacent sites rather than adjacent sectors, leading towards classic scenarios where cell edge suffers much more interference then cell centre users and hence low SAF. On the other hand, SAF in SC-CS is noticeably lower than that in cellular system because of the drastic change in distribution of data rates brought by SC.

4.4. Analysing power consumption

It is clear from results in Figure 4 that as expected, total power consumption increase as both mean of increasing capacity, that is, sectors or SC are added. Thus, although SC offer good means to increase capacity as seen previously, slightly higher power consumption is another payoff for them in addition to poorer SAF. Figures 5 plot the total power consumptions, respectively, for a range of R and S using (24) and preceding analysis. $\eta^s = \eta^r = 1$ is assumed because we are considering full load scenario. Antenna efficiency of commercial antennas is used. That is, $\zeta^r = \zeta^s = 60$ per cent. $P_f^s = 15W$ with $P_f^r = 0.5P_f^s$ is used because of reasons explained in [10]. It can be seen that in addition to the spectral efficiency and spatial fairness of data rates, power consumption also varies with S and R and thus adds a third dimension to the capacity-QoS tradeoff in dimensioning NTC. Figure 5 shows that power consumption per site increases more rapidly with the increase in number of SC (i.e. R) than increase in number of sectors per site (i.e. S). This is mainly because each SC has an omnidirectional antenna, so there is no compensating factor as in case of sectors.

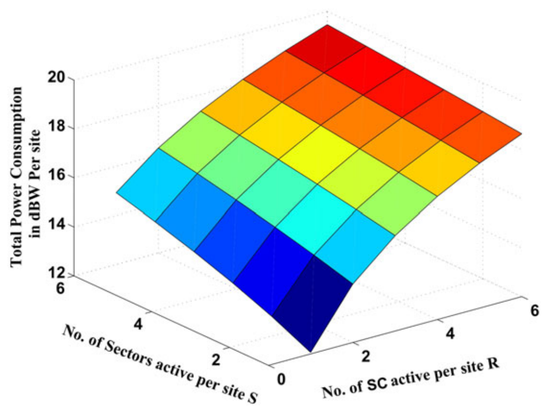

Figure 5. Total power consumption per site.

4.5. Trade off among the Three Performance Aspects

Finally, from results in Figure 4, it can be seen that no single NTC is simultaneously optimal in all three performance aspects. Here, the key observation to be made is there exists certain pareto optimality in which one objective generally improves only with loss in other. Therefore, the PCF's capability to precisely quantify this tradeoff with computation efficiency can actually help to design a NTC that is optimal to simultaneously meet the multiple CSO objectives in a priority order intended by the operator. Although regular topology has been assumed in the analysis for sake of simplicity, in case of realistic irregular topologies, PCF can build on the real user Minimization of Drive Test reports, and as a result, the KPIs produced by PCF and hence the optimal NTCs determined will still be optimal for real networks.

5. APPLICATION OF PCF IN HOLISTIC OPTIMISATION

Despite the fact that PCF can reduce τ_e and thus can reduce the overall solution time, holistic optimisation of all parameters together remains a daunting task. In this section, we propose a simple three step heuristic algorithm for SON engine to simplify the holistic optimisation by using the PCF.

5.1. CPS: a pragmatic heuristic for holistic optimisation

Classify parameters, prioritise objectives and solve subproblems (CPS) algorithm has the following three steps:

(1) Classify parameters of interests into hierarchical groups based on their impact on the KPIs. For example, parameters that substantially determine network performance can be classified in a group named gross parameters (GP), and the parameters that fine tune network performance in another group named fine tuning parameters (FTP). This grouping can be performed by examining the role of a particular parameter through PCF.

(2) Prioritise objectives: model the optimisation objective of the holistic CSO problem using PCF. This modelling should reflect operator's priorities for each KPI. This step will be explained in detail through a case study in the succeeding discussion.

(3) Solve the subproblem

   (a) Start from the highest group in the parameter hierarchy (resulted from Step 1). Optimise the objective function defined in Step 2, for the parameters in this group, considering it as subproblem independent of groups below it.

      (i) For solving this optimisation subproblem, normalise the KPIs to make them unitless to bring them to same scale.

      (ii) Use these normalised values of the KPIs in the objective function defined in Step 2 and solve the subproblem using exhaustive search or metaheuristic depending on the parameter group size.

   (b) Once a group of parameter is optimised, lock all parameters in that group at their optimal values and repeat Step 3 for the lower group until all groups are optimised.

CPS algorithm is further explained in the following through a case study.

5.2. A case study for CPS

As a case study, we consider joint optimisation of four key NTC parameters of F, S, R and θ that has been largely overlooked in the literature. Thus, the CSO problem under consideration can be written as

$$\max_{F,S,R,\theta} \left\{ \Upsilon\left(F,S,R,\theta\right), \Lambda\left(F,S,R,\theta\right), \Omega\left(F,S,R\right) \right\} \quad (26)$$

From previous section, we know that no single NTC is optimal simultaneously for Υ, Λ and Ω. This also implies that (26) is non-convex hence difficult to solve with analytical approaches. In the succeeding discussion, we apply the CPS to find a solution with low complexity. We place F, S and R in GP group and θ in FTP group. This grouping is quite intuitive and also can be inferred from the expressions in (5) and (6) that show that F, S and R have more profound impact on the SINR and hence the KPIs associated with it, than θ.

5.2.1. Optimizing GP.

Gross parameter optimisation problem can be written as

$$\max_{F,S,R} \left\{ \Upsilon\left(F,S,R\right), \Lambda\left(F,S,R\right), \Omega\left(F,S,R\right) \right\} \quad (27)$$

Because the mutual priority of these objectives and their target values are strongly dependent on the operator's policy [1], we propose to use the multi-objective optimisation as used in [48] by representing the three objectives simultaneously as a single utility function. That is,

$$v = \begin{cases} v_g(\Upsilon,\Lambda,\Omega) & , \textit{General Optimisation} \\ v_t(\Upsilon,\Lambda,\Omega) & , \textit{Targeted Optimisation} \end{cases} \quad (28)$$

where the subscripts g and t denote the general and targeted cases, respectively, as further explained in the following:

(1) Case 1 (general optimisation): this case represents a scenarios where the operator has no specific target values for the KPIs but has certain priority for each KPI. In this case, the optimisation problem can be modelled as

Trans. Emerging Tel. Tech. 27:1000–1015 (2016) © 2016 John Wiley & Sons, Ltd.
DOI: 10.1002/ett

$$\max_{F,S,R} \upsilon_g (\Upsilon, \Phi, \Omega) = \max_{F,S,R} (\lambda_1 \Upsilon + \lambda_2 \Lambda - \lambda_3 \Omega) \tag{29}$$

This utility function can reflect the mutual priority among these objectives. In the succeeding discussion, we present some exemplary rules to manifest these priorities:

(a) If the operator has equal priority for all the KPIs, in (29), set:

$$\lambda_1 = \lambda_2 = \lambda_3 = 1/3 \tag{30}$$

(b) If operator wants to maximise some objective (d^{th} objective), while neglecting others, in (29), set

$$\lambda_i = \begin{cases} 1 & if\ i = d \\ 0 & otherwise \end{cases}, \quad i = 1, 2, 3 \tag{31}$$

(c) If operator has specific priority for each objective, it can be represented by weights such that

$$\lambda_1 + \lambda_2 + \lambda_3 = 1 \tag{32}$$

(2) Case 2 (targeted optimisation): this case represents the scenario where the operator has specific target values to be achieved in each performance aspect: In this case the optimisation problem can be written as (33).

The rules for utility adaptation are as follows:

(a) If operator wants to achieve desired targets in each metric with same priority, substitute (30) in (33).
(b) If operator has desired target value in one objective but has no priority in others, substitute (31) in (33).
(c) If operator has specific values of each metric as target but has different priority of each target to be met, substitute (32) in (33).

Figure 4 provides the solution space for the problem in (27) obtained by the normalisation of the KPIs with their respective maximum values.

Figure 6 plots utility υ_g for four sets of different objective priorities. With equal priority of all three objectives, we can see that GP values in NTC = 9 are optimal. When capacity has highest priority, that is, 80 per cent and fairness and energy efficiency have lower and equal priorities of 10 per cent each, GP values in NTC = 22 are optimal. On the other hand, when fairness has highest importance, that is, 80per cent, and the capacity and energy efficiency have lower and equal priorities of 10 per cent, GP values in NTC = 9 become optimal. When energy efficiency is the most important target with 80 per cent importance factor, and fairness and spectral efficiency are lower priorities with importance of just 10 per cent, the optimal GP choice is given by NTC = 1.

Figure 7 plots υ_t for three different set of target values of the three objectives, each having the same priority, that is, $\lambda_1 = \lambda_2 = \lambda_3 = 1/3$. The first case (blue) represents the CSO scenario when operator wants capacity and fairness

$$\min_{F,S,R} \upsilon_t (\Upsilon, \Phi, \Omega) = \min_{F,S,R} \left| \sqrt{\lambda_1 (\Upsilon - \Upsilon_t)^2 + \lambda_2 (\Phi - \Phi_t)^2 + \lambda_3 (\Omega - \Omega_t)^2} \right| \tag{33}$$

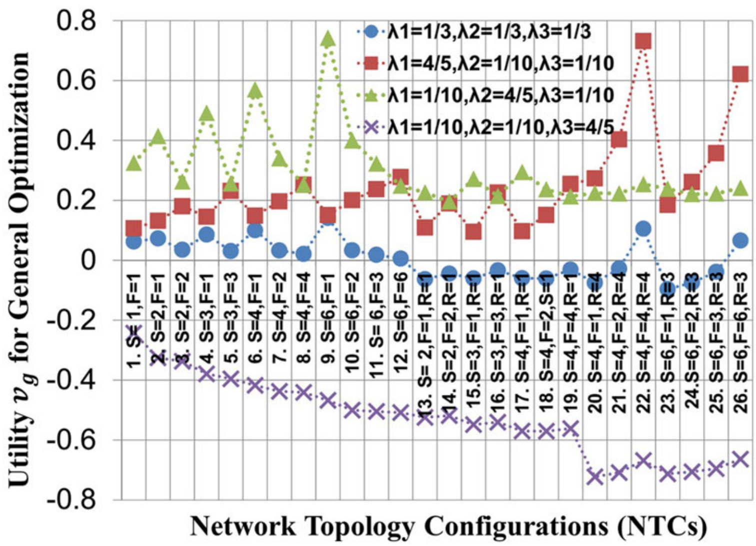

Figure 6. Solution space for general optimisation. NTC, network topology configuration.

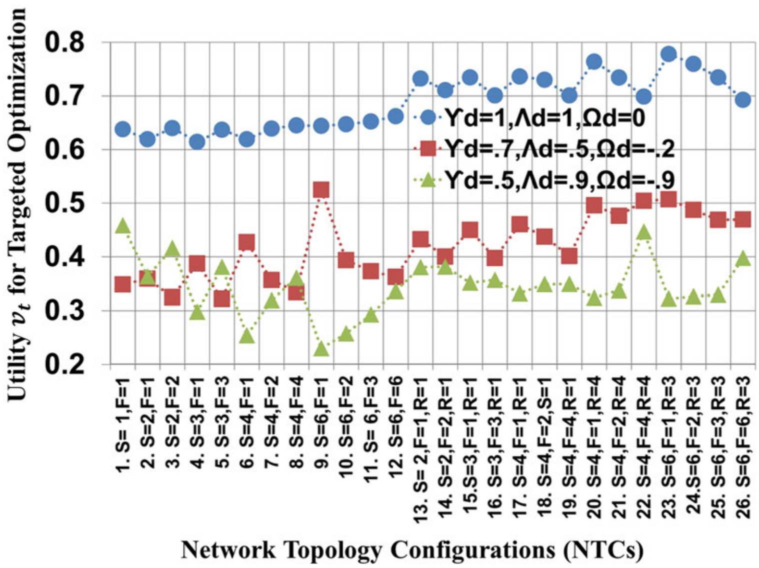

Figure 7. Solution space for targeted optimisation. NTC, network topology configuration.

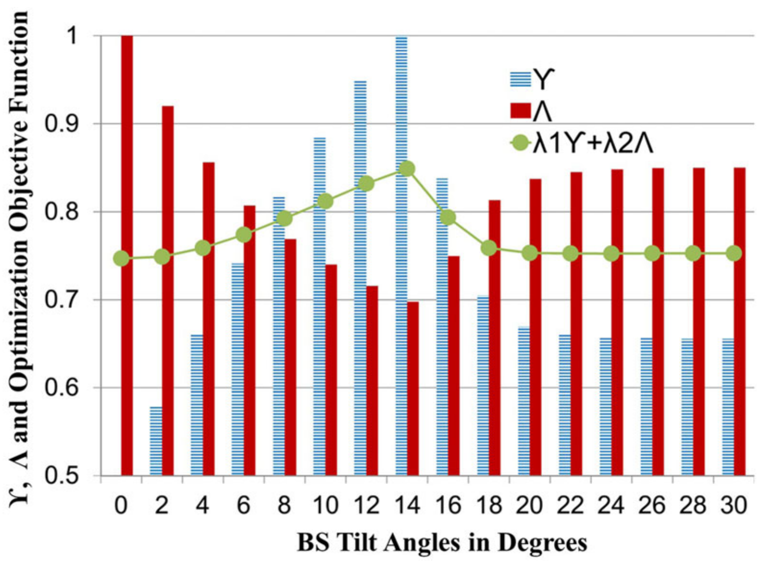

Figure 8. Υ, Λ and optimisation objective function of tilt angle, for NTC = 9. NTC, network topology configuration.

wise performance both to be closest to their absolute optimal values but has some flexibility in energy efficiency. In this case out of the 26 GPs combinations explored, the optimal solution is NTC = 23. The second case (red) represents a scenario where operator desires energy consumption to be closest to optimal, followed by spectral efficiency, followed by fairness. Now, the NTC = 9 can be seen to be the optimal solution. The last case (green) represents scenarios where operator wants SAF to be closest to its absolute optimal and can tolerate middle level performance in capacity but can compromise fully on the energy efficiency. In this case, NTC = 1 gives the optimal GP values to meet these priorities.

5.2.2. Optimizing FTP.

Assuming that the operator's business model requires all three KPIs to be equally important, this policy will be modelled with utility 1, with $\lambda_1 = \lambda_2 = \lambda_3 = 1/3$. In this case, the GP optimisation, that is, solution of subproblem in (27) will return a solution $(F = 1, S = 6, R = 0)$.The next step of holistic CSO problem according to CPS algorithm now can be written as

$$\max_{\boldsymbol{\theta}} \{\lambda_1 \Upsilon(\gamma(\boldsymbol{\theta})) + \lambda_2 \Lambda(\gamma(\boldsymbol{\theta}))\} \tag{34}$$

Note that because Ω is not a function of θ, therefore, it does not have to be included in the optimisation problem. The KPIs Υ and Λ are function of SINR γ, which is

Trans. Emerging Tel. Tech. **27**:1000–1015 (2016) © 2016 John Wiley & Sons, Ltd.
DOI: 10.1002/ett

further function of the $\boldsymbol{\theta}$, which is vector of tilt angles of all sector in the system as is modelled in (5). Note that in our particular case, each site has the same F and S. Therefore, from the insights obtained from (5), it is clear that optimal tilt angles being dependent on height and S and F (GPs) will be the same across the network. With this additional simplification similar to previous subsection, PCF can be used to quickly draw the solution space of (34), which is shown in Figure 8. It can be seen that again, there is strong tradeoff among the two KPIs and no single tilt is optimal for both Υ and Λ. Using the same utility-based approach as proposed previously, optimal value of the FTP that meets the operators defined objective the solution can be easily found. For example, in case where $\lambda_1 = \lambda_2$ reflects operator priorities, solution is $\theta = 14^0$. Thus, the solution of our CSO problem in (26), for given KPI priorities set by the operator, is $(F = 1, S = 6, R = 0, \theta = 14)$.

5.3. Complexity of PCF-based and CPS-based holistic CSO approach

Because the grouping of parameters reduces the search space size substantially and PCF reduces τ_e compared with traditional dynamic simulation-based SON approaches, the CPS algorithm can reduce the solution complexity of holistic CSO problem substantially. More specifically, if conventional approach takes time $\tau = \frac{\left(\frac{V^M}{\epsilon}\right)}{1/\tau_e}$ for solving CPS problem with M optimisation parameters each of which can take V different values, the CPS will take time:

$$\tau' = \frac{\tau}{\left(\frac{V^M}{\sum_{i=0}^{G}\left(V^{\frac{M}{gi}}\right)} \times \frac{\tau_e}{\tau_e'}\right)} \tag{35}$$

where τ_e' is time required for single evaluation of KPIs using PCF. G is number of groups in which CPS divides the parameters. This implies generally $\tau >> \tau'$. For our particular case study, the feasible combinations of F, S and R were as low as 26, and τ_e' on regular desktop computer was less than 1 s. Therefore, it took less than a minute to explore search space for GP and almost the same time for FTP.

6. CONCLUSION

In this paper, we presented a framework to quantify, analyse and optimise three major KPIs used for holistic optimisation of SON-enabled heterogeneous cellular systems, that is, capacity, SAF and energy efficiency. The PCF that we proposed in this paper can model the KPIs of interest as functions of a comprehensive set of optimisation parameters like spectrum reuse factor, number of sectors per site, number of SCs per site, adaptive coding and modulation. The metrics derived in the PCF can be quickly evaluated semi-analytically and thus can facilitate the solution of multi-objective holistic optimisation problem that otherwise is tackled using black box-type complex dynamic simulation models. Using PCF, we also evaluated and compared 26 different network topologies and quantified their relative gains. We analysed respective tradeoffs offered by each NTC in terms of capacity, SAF and energy efficiency. Our results also showed that contrary to common notion, NTCs with highest spectrum efficiency are not necessarily those that resort on full frequency reuse. The insights given by proposed framework can help to address new requirements from future heterogeneous cellular networks. Building on these insights, we proposed a heuristic CPS algorithm for holistic optimisation. We demonstrated through a case study, how PCF and CPS together can be used for a wide range of cellular optimisation scenarios with low complexity.

ACKNOWLEDGEMENTS

This work was made possible by NPRP grant no. 5-1047-2-437 from the Qatar National Research Fund (a member of The Qatar Foundation). The statements made herein are solely the responsibility of the authors.

REFERENCES

1. Aliu O, Imran A, Imran M, Evans B. A survey of self organisation in future cellular networks. *IEEE Communications Surveys Tutorials* 2012; **PP**(99): 1–26.
2. Ramiro J, Hamied K. *Self-organizing Networks (SON): Self-planning, Self-optimization and Self-healing for GSM, UMTS and LTE* (1st edn). Wiley Publishing: Chichester, UK, 2012.
3. Seppo Hämäläinen CSE, Cinzia S, Henning S (eds.) *LTE Self-organising Networks (SON): Network Management Automation for Operational Efficiency*. Wiley: Chichester, UK, 2012. ISBN 978-1-1199-7067-5.
4. 3GPP. Universal mobile telecommunications system (UMTS); LTE; universal terrestrial radio access (UTRA) and evolved universal terrestrial radio access (E-UTRA); Radio measurement collection for minimization of drive tests (MDT); overall description; stage 2. *Technical Report,* April 2011.
5. Boustani A, Khorsandi S, Danesfahani R, Mirmotahhary N. An efficient frequency reuse scheme by cell sectorization in OFDMA based wireless networks. In *Fourth International Conference on Computer Sciences and Convergence Information Technology (ICCIT)*, Seoul, Korea, 2009; 800–805.
6. Imran A, Yaacoub E, Dawy Z, Abu-Dayya A. On the capacity and spatial fairness trade-off in planning sectorization and frequency reuse. In *International Conference on Communications and Information Technology (ICCIT)*, Beirut, Lebanon, 2013; 406–410.

7. Physical layer aspects for evolved universal terrestrial radio access (UTRA). *Technical Report TR 25.814*, 3GPP Std.
8. Alsedairy T, Qi Y, Imran A, Imran MA, Evans B. Self organising cloud cells: a resource efficient network densification strategy. *Transactions on Emerging Telecommunications Technologies* 2015; **26**(8): 1096–1107.
9. Imran A, Zoha A. Challenges in 5G: how to empower SON with big data for enabling 5G. *Network, IEEE* 2014; **28**(6): 27–33.
10. Imran A, Tafazolli R. Evaluation and comparison of capacities and costs of multihop cellular networks. In *16th International Conference on Telecommunications (ICT)*, Marrakesh, Morocco, 2009; 160–165, Available from: http://portal.acm.org/citation.cfm?id=1700234.1700264 [30 April 2016].
11. Imran A, Yaacoub E, Dawy Z, Abu-Dayya A. Planning future cellular networks: a generic framework for performance quantification. In *Proceedings of the 19th European Wireless Conference (EW)*, Guildford, UK, 2013; 1–7.
12. Koutitas G, Karousos A, Tassiulas L. Deployment strategies and energy efficiency of cellular networks. *IEEE Transactions on Wireless Communications* 2012; **PP**(99): 1–12.
13. Velez FJ, Nazir MK, Aghvami AH, Holland O, Robalo D. Cost/revenue tradeoff in the optimization of fixed WiMAX deployment with relays. *IEEE Transactions on Vehicular Technology,* 2011; **60**(1): 298–312.
14. Abdel Khalek A, Al-Kanj L, Dawy Z, Turkiyyah G. Optimization models and algorithms for joint uplink/downlink UMTS radio network planning with SIR-based power control. *IEEE Transactions on Vehicular Technology* 2011; **60**(4): 1612–1625.
15. Berrocal-Plaza V, Vega-Rodriguez MA, Gomez-Pulido JA, Sanchez-Perez JM. Artificial bee colony algorithm applied to WiMAX network planning problem. In *11th International Conference on Intelligent Systems Design and Applications (ISDA)*, Cordoba, Spain, 2011; 504–509.
16. Zhang X, Zhang Y, Yu R, Wang W, Guizani M. Enhancing spectral-energy efficiency for LTE-advanced heterogeneous networks: a users social pattern perspective. *Wireless Communications, IEEE* 2014; **21**(2): 10–17.
17. Elayoubi SE, Ben Haddada O, Fourestie B. Performance evaluation of frequency planning schemes in OFDMA-based networks. *IEEE Transactions on Wireless Communications* 2008; **7**(5): 1623–1633.
18. Imran A, Imran MA, Tafazolli R. A novel self organizing framework for adaptive frequency reuse and deployment in future cellular networks. In *Proceedings IEEE 21st Int Personal Indoor and Mobile Radio Communications Symp., (PIMRC'10)*, Istanbul, Turkey, 2010; 2354–2359.
19. Bobkov DD, Zgurskiy AS, Kozlov OS, Kolomicev MA, Sakhnevich AL, Sklyar MA. 3GPP LTE access network planning. In *20th International Crimean Conference on Microwave and Telecommunication Technology (CriMicO)*, Sevastopol, 2010; 433–434.
20. Engels A, Reyer M, Xu X, Mathar R, Zhang J, Zhuang H. Autonomous self-optimization of coverage and capacity in LTE cellular networks. *Vehicular Technology, IEEE Transactions on* 2013; **62**(5): 1989–2004.
21. Guo W, O'Farrell T. Relay deployment in cellular networks: planning and optimization. *IEEE Journal on Selected Areas in Communications* 2012; **PP**(99): 1–10.
22. Imran A, Imran MA, Abu-Dayya A, Tafazolli R. Self organization of tilts in relay enhanced networks: a distributed solution. *Wireless Communications, IEEE Transactions on* 2014; **13**(2): 764–779.
23. Beyer J, Isensee U, Droste H. A measurement based approach to predict the MIMO throughput of the LTE downlink in RF planning tools. In *IEEE Vehicular Technology Conference (VTC Fall)*, San Francisco, California, 2011; 1–5.
24. Gu J, Ruan Y, Chen X, Wang C. A novel traffic capacity planning methodology for LTE radio network dimensioning. In *IET International Conference on Communication Technology and Application (ICCTA)*, Beijing, China, 2011; 462–466.
25. Zhang J, Sun C, Yi Y, Zhuang H. A hybrid framework for capacity and coverage optimization in self-organizing LTE networks. In *2013 IEEE 24th International Symposium on Personal Indoor and Mobile Radio Communications (PIMRC)*, London, UK, 2013; 2919–2923.
26. Gordejuela-Sanchez F, Zhang J. LTE access network planning and optimization: a service-oriented and technology-specific perspective. In *IEEE Global Telecommunications Conference, (GLOBECOM)*, Hawaii, USA, 2009; 1–5.
27. Louvros S, Aggelis K, Baltagiannis A. LTE cell coverage planning algorithm optimising uplink user cell throughput. In *Proceedings of the 11th International Conference on Telecommunications (ConTEL)*, Graz, Austria, 2011; 51–58.
28. Hurley S, Allen S, Ryan D, Taplin R. Modelling and planning fixed wireless networks. *Wireless Networks: The Journal of Mobile Communication, Computation and Information* 2010; **16**(3): 577–592.
29. El-Beaino W, El-Hajj AM, Dawy Z. A proactive approach for LTE radio network planning with green considerations. In *19th International Conference on Telecommunications (ICT)*, Jounieh, Lebanon, 2012; 1–5.

Trans. Emerging Tel. Tech. **27**:1000–1015 (2016) © 2016 John Wiley & Sons, Ltd.
DOI: 10.1002/ett

30. Niu Z, Zhou S, Hua Y, Zhang Q, Cao D. Energy-aware network planning for wireless cellular system with inter-cell cooperation. *IEEE Transactions on Wireless Communications* 2012; **PP**(99): 1–12.
31. Xiao X, Tao X, Lu J. Energy-efficient resource allocation in LTE-based MIMO-OFDMA systems with user rate constraints. *Vehicular Technology, IEEE Transactions on* 2015; **64**(1): 185–197.
32. Li Z, Li S. LTE network planning based on game theory. In *International Conference on Computer Science and Service System (CSSS)*, Nanjing, China, 2011; 3963–3966.
33. Trichias K, Litjens R, Tall A, Altman Z, Ramachandra P. Self-optimisation of vertical sectorisation in a realistic LTE network. In *2015 European Conference on Networks and Communications (EuCNC)*, Paris, France, 2015; 149–153.
34. Yaacoub E, Dawy Z. LTE BS placement optimization using simulated annealing in the presence of femtocells. In *Proceedings of 20th European Wireless Conference, European Wireless 2014*, Barcelona, Spain, 2014; 1–5.
35. Peyvandi H, Imran A, Imran MA, Tafazolli R. A target-following regime using similarity measure for coverage and capacity optimization in self-organizing cellular networks with hot-spot. In *Proceedings of 20th European Wireless Conference, European Wireless 2014*, Barcelona, Spain, 2014; 1–6.
36. Hurley S. Planning effective cellular mobile radio networks. *IEEE Transactions on Vehicular Technology* 2002; **51**(2): 243–253.
37. Tsilimantos D, Kaklamani D, Tsoulos G. Particle swarm optimization for UMTS WCDMA network planning. In *3rd International Symposium on Wireless Pervasive Computing (ISWPC)*, Santorini, 2008; 283–287.
38. Yang H, Wang J, Song X, Yang Y, Wang M. Wireless base stations planning based on GIS and genetic algorithms. In *9th International Conference on Geoinformatics*, Shanghai, China, 2011; 1–5.
39. Awada A, Wegmann B, Viering I, Klein A. Optimizing the radio network parameters of the long term evolution system using Taguchi's method. *IEEE Transactions on Vehicular Technology* 2011; **60**(8): 3825–3839.
40. Amaldi E, Capone A, Malucelli F, Signori F. UMTS radio planning: optimizing base station configuration. In *IEEE 56th Vehicular Technology Conference, (VTC Fall. 2002)*, Vancouver, Canada, 2002; 768–772.
41. Galota M, Glasser C, Reith S, Vollmer H. A polynomial-time approximation scheme for base station positioning in UMTS networks. In *Proceedings of the 5th International Workshop on Discrete Algorithms and Methods for Mobile Computing and Communications*. ACM, New York, NY, USA, 2001; 52–59.
42. Weicker N, Szabo G, Weicker K, Widmayer P. Evolutionary multiobjective optimization for base station transmitter placement with frequency assignment. *IEEE Transactions on Evolutionary Computation* 2003; **7**(2): 189–203.
43. Imran A, Imran MA, Tafazolli R. Relay station access link spectral efficiency optimization through SO of macro BS tilts. *IEEE Communications Letters* 2011; **15**: 1326–1328.
44. Viering I, Dottling M, Lobinger A. A mathematical perspective of self-optimizing wireless networks. In *IEEE International Conference on Communications (ICC '09)*, Dresden, Germany, 2009; 1–6.
45. Jain R. *The Art of Computer Systems Performance Analysis*, Techniques for experimental design, measurement, simulation and modeling. Wiley: Chichester, UK, 1991.
46. Pozar D. Directivity of omnidirectional antennas. *IEEE Antennas and Propagation Magazine* 1993; **35**(5): 50–51.
47. Jacobson KR, Krzymien WA. System design and throughput analysis for multihop relaying in cellular systems. *IEEE Transactions on Vehicular Technology* 2009; **58**(8): 4514–4528.
48. Bao Y, Jiang H, Huang Y, Hu R. Multi-objective optimization of power control and resource allocation for cognitive wireless networks. In *Proceedings IEEE 8th/ACIS International Conference on Computer and Information Science, (ICIS'09)*, Shanghai, China, 2009; 70–74.