

# Fast and Efficient Energy-Oriented Cell Assignment in Heterogeneous Networks

Javier Rubio-Loyola<sup>1</sup> · Christian Aguilar-Fuster<sup>1</sup> · Luis Diez<sup>2</sup> · Ramon Agüero<sup>2</sup> · Juan Luis-Gorricho<sup>3</sup> · Joan Serrat<sup>3</sup>

Received: date / Accepted: date

**Abstract** The cell assignment problem is combinatorial, with increased complexity when it is tackled considering resource allocation. This paper models joint cell assignment and resource allocation for cellular heterogeneous networks, and formalizes cell assignment as an optimization problem. Exact algorithms can find optimal solutions to the cell assignment problem, but their execution time increases drastically with realistic network deployments. In turn, heuristics are able to find solutions in reasonable execution times, but they get usually stuck in local optima, thus failing to find optimal solutions. Metaheuristic approaches have been successful in finding solutions closer to the optimum one to combinatorial problems for large instances. In this paper we propose a fast and efficient heuristic that yields very competitive cell assignment solutions compared to those obtained with three of the most widely-used metaheuristics, which are known to find solutions close to the optimum due to the nature of their search space exploration. Our heuristic approach adds energy expenditure reduction in its algorithmic design. Through simulation and formal statistical analysis, the proposed scheme has been proved to produce efficient assignments in terms of the number of served users, resource

allocation and energy savings, while being an order of magnitude faster than metaheuristic-based approaches.

**Keywords** Cell assignment · resource allocation · metaheuristic · energy efficiency · cellular networks · heterogeneous networks · dense networks

## 1 Introduction

Current forecasts indicate that wireless traffic demand will continuously increase in the near future, reaching more than 4.5 ZB in the next years [8]. In order to provide the required capacity, 5G networks are expected to bring a thousandfold capacity increase, in comparison to current 4G technologies [19]. Among other solutions, network densification stands out as one of the more relevant approaches to ensure such capacity increase [3, 2].

Network densification is achieved by both spectrum and spatial densification. In the former case, potentially non-contiguous frequency chunks are used together by the same or different communication technologies [3, 20], which would lead to multi-connectivity scenarios [20]. On the other hand, spatial densification is obtained by deploying a large number of heterogeneous access elements, such as small-cells or WiFi access points, leading to the so-called heterogeneous networks.

Bearing this scenario in mind, it becomes highly important to adapt network management solutions that will appropriately exploit the additional capacity brought by heterogeneous networks. Among the different techniques, access selection and resource allocation play a fundamental role to ensure an efficient usage of the radio resources. Hence, it is sensible for both techniques to be tightly coordinated in joint solutions [22, 1]. In addition, the appearance of new network techniques and

---

<sup>1</sup>Javier Rubio-Loyola, Christian Aguilar-Fuster  
CINVESTAV-Tamaulipas, Cd. Victoria Tamaulipas  
E-mail: {jrubio,caguilar}@tamaps.cinvestav.mx

<sup>2</sup>Luis Diez, Ramon Agüero  
Communications Engineering Dpt,  
University of Cantabria, Santander, Spain  
E-mail: {ldiez, ramon}@tlmat.unican.es

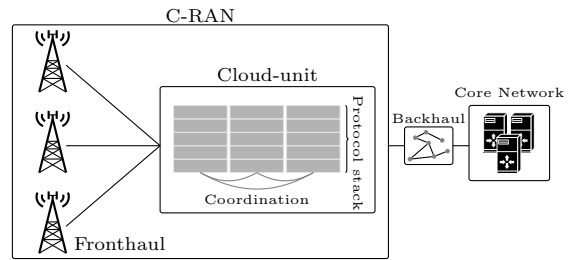
<sup>3</sup>Juan Luis Gorricho, Joan Serrat  
Telematics Engineering Dpt,  
Polytechnical University of Catalunya, Barcelona, Spain  
E-mail: juanluis@entel.upc.edu, serrat@tsc.upc.edu

architectures, such as SDN [28] and C-RAN [18], will allow the use of more centralized algorithms, which would actually facilitate such tight coordination among the elements of the dense network.

Furthermore, the higher communication capacity offered by highly dense networks, along with the corresponding coordination requirements, poses new challenges, with achieving energy efficiency as one of the most relevant. In this sense, compared to traditional networks, where access selection and resource allocation were analyzed by simply considering the wireless access segment, now their impact over backhaul links (access elements towards the network) is also studied, to optimize the efficiency of both radio and energy resources [15, 16]. In particular, the possibility to switch off access elements when they are not needed is seen as one of the most promising techniques [11], since it would reduce energy consumption in both, the access and the backhaul network segments.

This paper models joint cell assignment and resource allocation for cellular heterogeneous networks, and formalizes cell assignment as a combinatorial optimization problem. Generally speaking, combinatorial optimization problems can be solved with exact, heuristic and metaheuristic approaches. Exact algorithms can find optimal solutions however, their execution time increases drastically so that their application is unfeasible with realistic deployments, i.e. larger amounts of base stations and mobile users. It is widely accepted that heuristics are able to find solutions to practical instances of optimization problems with reasonable execution times, but they get usually stuck in local optima hence failing to find appropriate solutions. It is also widely accepted that metaheuristic-based approaches are successful in finding solutions closer to the optimum to combinatorial problems for large instances, at the expense of larger execution times [5].

In this paper we propose a fast and efficient heuristic algorithm for joint cell assignment and resource allocation in dense heterogeneous networks that yields very competitive cell assignment solutions, compared to the ones obtained with three of the most widely-used metaheuristics namely, Harmony Search (HS) [13], Simulated annealing (SA) [26], and Genetic algorithm (GA) [14]. These three metaheuristics have been selected for comparison with our approach since they have been successfully used to solve a wide range of applications, including large-scale combinatorial and constrained optimizations, like the one addressed in this paper. The proposed solution aims to optimize the efficiency of radio resources, while reducing energy consumption by switching off unused cells. Through simulation and formal statistical analysis, the proposed



**Fig. 1** Simplified scheme of the C-RAN architecture.

scheme has been proven to produce efficient assignments in terms of the number of served users, resource allocation and energy savings, while being an order of magnitude faster than metaheuristic-based approaches. As a global optimization solution, the algorithm is conceived for centralized architectures (i.e. C-RAN) where cooperation among access elements is possible. As depicted in Figure 1, in C-RAN the functionalities of the base stations are virtualized and centralized in cloud units. Hence, communication delays between access elements are minimized, enabling tight coordination at different layers of the protocol stack. In this scenario, centralized algorithms may be exploited, since it becomes possible to gather an overall knowledge of the network.

The contributions of this paper are enumerated below:

- General modeling of the joint cell assignment and resource allocation problem for cellular networks.
- Adding energy expenditure reduction within the model to foster cell switching-off.
- Considering actual traffic demand, opposed to full-buffer or saturation condition.
- Proposal of a fast heuristic solution for the cell assignment problem. The algorithm is afterwards extended to include the possibility to switch off cells (energy minimization).
- Performance comparison of the heuristic-based solution with other approaches that are known to produce more efficient solutions (e.g. three metaheuristic-based solutions), in terms of cell assignment efficiency, energy consumption, and execution time.

The most noteworthy contribution is that we use a novel optimization technique, which is able to yield similar results to other metaheuristic-based solutions, but is an order of magnitude faster. We exploit it to solve a joint cellular assignment/resource allocation problem, which has been tackled using different approaches.

The rest of the paper is organized as follows. In Section 2 we revise the current state of the art regarding the joint problem of access selection and resource allocation, and we discuss the need for heuristic solutions.

Then, we define the system model in Section 3, where we pose the aforementioned problem, and we extend it to include energy efficiency. In Section 4 we describe the metaheuristics Harmony Search (HS) [13], Simulated annealing (SA) [26], and Genetic algorithm (GA) [14] addressing the optimization problem modeled in Section 3. Afterwards, our solution is introduced in Section 5, highlighting its different modules for cell assignment and energy saving while addressing the combinatorial optimization problem modelled in Section 3. The evaluation scenario and performance analysis are depicted in Sections 6 and 7, respectively. Finally, Section 8 concludes the paper, summarizing the work and providing an outlook on the aspects that will be tackled in our future research.

## 2 State of the art

Several studies have looked at the joint problem of user association and resource allocation over heterogeneous networks. Worthy of mention among them, are those by Ye *et al.* [31] and Bu *et al.* [7], due to their relevance and similarities with our proposed scheme. As stated in [31], the traditional two-fold problem of cellular networks becomes a complex combinatorial problem: first, the base station that will be associated to each user has to be decided, and afterwards the amount of radio resources to be allocated to each user has to be calculated. To solve this problem, the authors of [31] propose maximizing the sum of transmitted rate logarithms of all active users. This is the most suitable criterion for achieving proportional fairness in resource usage. They also show that with the applied maximization criterion, the optimal resource allocation is the equal allocation, where the base station equally divides the time it uses to transmit to each of its users. They assume full-buffer condition, where all base stations always have frames to be transmitted. This strongly simplifies the modeling of co-channel interference, although it might not be always realistic. This combinatorial problem can be relaxed, allowing any user to be connected to more than one base station, leading to a linear program, which yields solutions close to the optimal one. In this sense, there might be some cases where a user is connected to more than one BS, and this can be avoided, in a second step, by applying different heuristics (for instance, rounding). In this way we can easily find the best assignment for each user, considering those obtained with the aforementioned linear program.

In a similar way, the authors of [7] try to ensure proportional fairness among users, by also using the same utility function (sum of transmission rate logarithms). The main novelty of this second approach is that it

assumes that the base station association is known a priori. Hence, the problem can be reformulated as a maximum weighted bipartite matching problem, to find the matching between all users and all base stations, maximizing the aforementioned criterion. As it is well known, the bipartite matching problem can be solved in polynomial time. Since the authors assumed that the number of connections to each base station was known beforehand, the matching problem needs to be solved with all possible combinations or, at least, considering the feasible combinations according to current users' location. The authors propose using some heuristics to solve it on-line.

As a practical alternative for solving the two-fold problem, while ensuring proportional fairness, some authors, including Ye *et al.* [31], have proposed adding a bias to the pilot signal received by the mobile terminal from small-cells in a heterogeneous network. The resulting pilot strength will be used to decide the base station associated that will be to each user. The authors of [31] conclude that this strategy leads to a behavior that is quite close to that analytically obtained after applying the maximization criterion. In a similar way, the work presented in [23], using the biased signal strength strategy, with an iterative procedure to establish the best bias to be applied to each base station, yields the greatest Jain fairness index, thus ensuring load balancing in heterogeneous LTE networks.

More recently, joint user assignment and resource allocation has been tackled in dense networks with mobile edge computing [22]. Although the application scenario differs from the ones used by the works discussed above, the underlying system model boils down to a similar optimization problem. In particular, the authors in [22] propose a sequential solution where user assignment is performed first, to afterwards solve the resource allocation problem, by exploiting quantum-behaved particle swarm optimization (QPSO). Similarly, in [25] the two-fold problem arises in device-to-device scenarios, where user demand needs to be satisfied. Opposed to previous studies, the initial problem is relaxed, leading to a convex version that would then yield a suboptimal solution.

Our work differs from the previously discussed papers in various aspects. First, we propose a more realistic scenario based on how cellular networks operate. In particular, the two-fold problem considers that each individual user has a certain traffic demand, instead of equally sharing the transmission time of each base station among all associated users. Consequently, the co-channel interference cannot be simplified, since not all base stations will be transmitting all the time (i.e. we do not consider full buffer situation). It is worth

highlighting that when user demand is assumed (i.e. no saturation conditions), it has been shown [6] that the coupled resource allocation problem resulting from the co-channel interference cannot be easily solved (i.e. it is not convex).

Other works exploiting different optimization techniques focus on resource allocation, paying special attention to physical layer aspects, thus exploiting low-level techniques such as interference alignment. For instance, in [32] a novel technique is proposed to optimize power consumption in opportunistic communications, while power allocation is addressed in [33] and [34], exploiting cognitive radio systems and artificial noise respectively. Similarly, Wu *et al.* [27] propose a game theoretical solution to optimize the resource allocation in heterogeneous networks considering the required QoS. It is worth noting that these works fundamentally differ from our approach in terms of the level of system abstraction. In this sense, they model low level mechanisms in great detail, while other procedures, happening at a larger pace (for instance, access selection), are assumed to be known.

On the other hand, other studies, such as [29] have focused on optimizing the access selection and handover procedures, while the resource allocation and energy consumption are not considered. Finally, there exist studies that aim to optimize the global energy efficiency in heterogeneous networks, by improving network planning, while assuring some level of quality of service. For instance, the authors of [30] analyze the optimal distribution of small cells in heterogeneous networks, exploiting stochastic geometry techniques, with the goal to reduce energy drain. Although this work shares with our proposal the same overall target, it focuses on network characterization, rather than proposing an algorithm to be exploited to leverage the sought optimization.

In Table 1 we summarize the above comparison, by indicating the particular features of the proposed approach that are covered by the works that are closer to ours. In particular, we analyzed the following parameters: joint access selection and resource allocation; heterogeneous networks scenario; whether energy efficiency is addressed; and whether the proposed solution takes into account the actual traffic demand of users, in contrast to saturation of full-buffer conditions.

In addition, as mentioned earlier, the proposed solution also seeks to minimize the number of required access elements that are needed to ensure user satisfaction. Due to the additional complexity of the proposed working scenario, a novel heuristic approach will be introduced, and its performance will be compared with three of the most widely used *metaheuristic* approaches, as was previously mentioned, when solving

**Table 1** Summary of the features addressed by the related work.

	[31]	[7]	[23]	[22]	[25]	[33]	[34]	[32]	[27]
Joint AS/RA	✓	✓	✓	✓	✓				
HetNet Scen.	✓		✓	✓					✓
Energy Eff.						✓	✓	✓	✓
Traffic Demand						✓	✓	✓	

the corresponding maximization problem. It is worth mentioning that heuristic-based approaches are able to find solutions to practical optimization problems with reasonable execution times, but they usually get stuck in local optima, hence failing to find better solutions, and metaheuristic approaches have been successful in finding closer solutions to the optimum to combinatorial problems [5], such as the one we tackle in this work.

### 3 System Model

We consider a scenario comprising a set of users and base stations,  $\mathcal{U}$  and  $\mathcal{B}$  respectively. We define a set of user categories  $\mathcal{M}$ , so that  $\mathcal{U}_m \subseteq \mathcal{U}$  represents the subset of users belonging to the category  $m \in \mathcal{M}$ , and  $\bigcup_{m \in \mathcal{M}} \mathcal{U}_m = \mathcal{U}$ . Similarly,  $\mathcal{N}$  holds for the set of base station categories, so  $\mathcal{B}_n \subseteq \mathcal{B}$  is the subset of base stations of category  $n \in \mathcal{N}$ , and  $\bigcup_{n \in \mathcal{N}} \mathcal{B}_n = \mathcal{B}$ .

In this work, we assume users run services with a known traffic demand (i.e. non-elastic services), that has to be satisfied by one base station. Then, we define the traffic demand of each user according to the category they belong to, so that the traffic demand of user  $i \in \mathcal{U}$  is defined as  $c_i := \{c_m \mid m \in \mathcal{M}, i \in \mathcal{U}_m\}$ . In a similar way, we define the capacity of each base station in accordance with its category. Therefore, the total amount of radio resources of the base station  $j$ ,  $R_j$ , is defined as follows:  $R_j := \{R_n \mid n \in \mathcal{N}, j \in \mathcal{B}_n\}$ .

To satisfy the traffic demand of each user  $i$ , the corresponding serving base station  $j$  needs to allocate a number of radio resources  $r_{i,j}$ . Then, we define  $\alpha_{i,j}$  as the ratio between the resources allocated to user  $i$ , and the total capacity  $R_j$ . It is worth noting that the mapping between traffic demand and required number of resources is not direct, but it depends on the state of the radio connection.

Then, we model the access selection procedure with a binary variable,  $b_{i,j}$ , that identifies the base station user  $i$  is connected to, as defined in Eq. 1. Besides, for each pair of user and base station,  $(i, j)$ , we denote  $\gamma_{i,j}$  as the power user  $i$  receives from base station  $j$ .

$$b_{ij} = \begin{cases} 1 & j \text{ is the serving base station of user } i \\ 0 & \text{otherwise} \end{cases} \quad (1)$$

The amount of resources  $r_{ij}$  to be allocated to each user depends on the spectral efficiency, that in turn depends on the SINR experienced by such user. In this work, we consider that base stations use a random scheduler, so that the amount of interference power coming from a cell is proportional to the load of such cell. In such case, we can define the SINR of the user  $i$  when connected to base station  $j$  as follows:

$$SINR_{i,j} = \frac{\gamma_{i,j}}{\sigma^2 + \alpha_{i,j} \sum_{k \in \mathcal{B}/j} \left[ \left( \sum_{l \in \mathcal{U}} b_{l,k} \alpha_{l,k} \right) \gamma_{i,k} \right]} \quad (2)$$

where  $\sigma^2$  corresponds to the system white additive noise power. As can be observed in Eq. 2, the interference is modulated by (1)  $\alpha_{i,j}$  which represents the ratio of resources granted by the serving cell, and (2)  $\sum b_{l,k} \alpha_{l,k}$ , which represents the load of interfering base stations. As can be seen, when a cell does not assign any resources, it does not induce interference. Afterwards, the spectral efficiency of the connection between user  $i$  and base station  $j$  can be straightforwardly computed, following the Shannon-Hartley's channel capacity formula [17], as shown in Eq. 3. This parameter, measured in bits per second and hertz, permits calculating the number of resources that should be allocated to each user, according to its traffic demand, the bandwidth of the resources in the system,  $BW$ , and the connection state.

$$\eta_{i,j} = \log_2(1 + SINR_{i,j}) \quad (3)$$

Our aim is to optimize the access selection, considering multiple criteria. First and foremost, we want to maximize the number of users whose traffic demand is satisfied. In addition, we seek to reduce the usage of resources, by selecting efficient connections, which boils down to higher SINR levels. It is worth recalling that the greatest possible SINR values correspond to the situation where a user does not receive any interference, so that the interfering base station does not assign any resource, and would be thus switched off.

To take into account the connection efficiency criterion, we define in Eq. 4 the connection efficiency,  $u_{i,j}$  as the ratio between the experienced SINR and the SNR, corresponding to the situation where interfering base stations do not allocate any resource. As can be observed, the connection efficiency takes the value 1 when

the user does not experience any interference (i.e. minimum amount of allocated resources to satisfy the demand), and tends to 0 as the interfering base station becomes more loaded.

$$u_i = \frac{SINR_{i,j}}{SNR_{i,j}} = \frac{SINR_{i,j}}{\frac{\gamma_{i,j}}{\sigma^2}} = \left\{ 1 + \frac{\alpha_{i,j}}{\sigma^2} \sum_{k \in \mathcal{B}/j} \left[ \left( \sum_{l \in \mathcal{U}} b_{l,k} \alpha_{l,k} \right) \gamma_{i,k} \right] \right\}^{-1} \quad (4)$$

Altogether, we define a coupled assignment and resource allocation problem (Problem 1), where  $b_{i,j}$  are the variables related to the access selection procedure, and  $\alpha_{i,j}$  the ones that define the resource allocation.

### Problem 1 (Coupled user assignment and resource allocation problem)

$$\max \sum_{i,j \in \mathcal{U}, \mathcal{B}} \omega_1 b_{i,j} + \omega_2 (u_{i,j} - \alpha_{i,j}) \quad (5)$$

$$\text{s.t.} \sum_{j \in \mathcal{B}} b_{i,j} \leq 1 \quad \forall i \in \mathcal{U} \quad (6)$$

$$\sum_{i \in \mathcal{U}} \alpha_{i,j} \leq 1 \quad \forall j \in \mathcal{B} \quad (7)$$

$$c_i \leq \sum_{j \in \mathcal{B}} b_{i,j} \alpha_{i,j} R_j BW \eta_{i,j} \quad \forall i \in \mathcal{U} \quad (8)$$

$$0 \leq \omega_t \leq 1 \quad t \in \{1, 2\} \quad (9)$$

The utility function in Eq. 5 is a weighted summation whose first term aims to maximize the number of connected users, while the second one minimizes the number of required base stations and maximizes the spectral efficiency by means of the connection utility. In addition, since both terms are bounded between 0 and 1, it is possible to prioritize one criterion over the other by tuning the weighting parameters. In particular, in this work we will prioritize the number of connections over the other criteria, so that  $\omega_1 \gg \omega_2$ .

Furthermore, we include three constraints within the problem: the first two correspond to the access selection procedure, while the third one related to resource allocation. In particular, Eq. 6 ensures that each user only connects to a single base station, while Eq. 7 ensures that the capacity of base stations is not exceeded. On the other hand, Eq. 8 guarantees that the traffic demand of each user is satisfied. It can be observed, that the access selection strategy depends on the resource allocation, and vice versa, so that it is not possible to separate the problem variables. In addition

the optimization problem has binary integer variables, and it is not convex.

Hence, in order to solve Problem 1 three options are possible [5]: (1) exact algorithms can find optimal solutions to the problem, however their execution time increases drastically with practical instances and they become prohibitive in realistic cell assignment scenarios; (2) heuristics are able to find solutions to practical assignment scenarios of the problem with reasonable execution times, however they usually get stuck in local optima, hence failing to find better solutions; (3) metaheuristic approaches have been successful in finding closer solutions to the optimum to combinatorial problems at the expense of execution time. In this paper we propose an heuristic approach to solve our cell assignment optimization problem with two aims: (1) to be fast so that it can be applied in practical cell assignment scenarios; and (2) to be efficient so that it can find solutions close to those obtained with metaheuristic-based approaches. To evaluate our approach, we consider the following metrics: percentage of served users, profit, active cells, and execution time.

The served users metric indicates the number of effectively connected users. This metric is formally defined as  $\frac{\sum_{i=0}^{|\mathcal{U}|} b_i}{|\mathcal{U}|}$  where  $b_i$  is a binary variable that equals 1 when the user  $i$  has been assigned to one base station that satisfies his demand, and does not exceed the resources of the base station, otherwise it equals 0. We define the profit as the relationship between  $u_{ij}$  and  $\alpha_{ij}$  (utility-cost). Profit is formally defined as  $\sum_{i,j \in \mathcal{U}, \mathcal{B}} (u_{i,j} - \alpha_{i,j})$ .

We define the active cells metric as the number of base stations that have, at least, one assigned user, and we formally define it as  $\sum_{j=0}^{|\mathcal{B}|} b_j$ , where  $b_j$  is a binary variable that equals 1 when the base station  $j$  has at least one user assigned, otherwise it equals 0. Finally, the execution time measures the time that an algorithm takes to solve each test instance, namely to define the most appropriate cell assignment.

## 4 Metaheuristic-based cell assignment

In this Section, we describe basic concepts of metaheuristic-based optimization and the solutions that have been used to compare our proposed cell assignment approach.

### 4.1 Metaheuristic optimization

Optimization refers to the process of searching and comparing feasible solutions until no better solution can be found. It sometimes involves the need to not only reach the best value for a given objective function (or

set of objective functions), but also to satisfy a certain set of predefined requirements, called constraints. Since most engineering and scientific applications involve a complex optimization problem to be solved, a number of optimization methods have been proposed. Metaheuristics encompass a set of optimization techniques that have been found to be effective in locating solutions close to the global optimum in several problems. Optimization metaheuristics can be classified into single-solution-based and population-based [24].

In general, optimization metaheuristics generate new solutions (i.e. cell assignments in the context of this paper) by applying variation operators to previously generated solutions. Then, a selection process between the former and the generated solutions is performed to keep moving the found solutions towards the optimum. Evolutionary computation (EC) comprises a set of metaheuristics that have drawn inspiration from the process of natural evolution [14]. EC techniques have gained popularity in recent years, since they have shown flexibility, adaptability, and good performance [9]

The implementation of the evolution metaphor on EC is described as follows [14]: given a population of individuals, an environment with limited resources and the pressure in the environment causes a natural selection. This mechanism raises the fitness of the population. Given a function to be optimized, a set of candidate solutions is created and evaluated to measure their fitness in the given environment. Based on this ability, a set of solutions is selected to feed the variation operators, to generate new solutions. These newly generated solutions, called offspring, may compete (or not) against their parents to take their place in the next generation. The above process is repeated until a solution with good quality is found (or some criterion is reached).

### 4.2 Specific metaheuristics

This Section describes the metaheuristics that have been compared with our proposed cell assignment approach, namely Harmony Search (HS), Simulated annealing (SA), and Genetic algorithm (GA). These are popular metaheuristics that have been successfully used to solve a wide range of combinatorial problems, exhibiting a good performance [9].

GA is a metaheuristic inspired by the theory of natural evolution [14]. Although there are many variants of GA's, they share basic mechanisms, which are illustrated in Algorithm 1.

In the GA customized for energy-aware cellular assignment, each individual is represented by a binary

**Algorithm 1:** Genetic algorithm.

---

**Data:**  $n, P[n], p_c, p_m$  //  $n$  No. of candidate solutions,  $P[n]$  list of candidate solutions,  $p_c$  crossover probability,  $p_m$  mutation probability

**Result:** the best individual of all populations

```

1  $t \leftarrow 0$ ;
2 Step 1 Initialize population  $P[n]$  randomly with
   $n$  individuals;
3 while  $t \leq iterations_{max}$  do
4    $t \leftarrow t + 1$ ;
5   Step 2 Calculate aptitude  $f(x)$  for each
    individual  $x \in P[n]$ ;
6   Step 3 Select parents of next generation
    based on their aptitude  $C(x)$ ;
7   Step 4 Crossover selected partner based on
     $p_c$  to generate new population;
8   Step 5 Mutate individuals of new generation
    based on  $p_m$ .;
9 end
10 return the best individual of all populations;
```

---

vector  $V2_i \in V2$ , which represents the status of all base stations (see Figure 2). Proportional selection is adopted based on the fitness of each individual, and two genetic operators, crossover and mutation, are used (two-point crossover and uniform mutation). These operators are applied to the selected individuals with fixed probabilities  $p_c$  and  $p_m$ , corresponding to the crossover and mutation operators, respectively.

Harmony Search (HS) algorithm is one of the recent additions to the metaheuristics search techniques for combinatorial optimization problems [10]. This algorithm is inspired by the improvisation process that musicians do to find the optimum harmony [13]. Each candidate solution is represented as an  $n$ -dimensional harmony. The initial harmonies, usually randomly fed, are stored in a harmony memory (HM). A harmony is composed of musical notes which, in turn, represent the variables of the problem to solve (in our case, status of the base stations turned off, and turned on).

The HM has a fixed size  $hms$ , and it stores the harmonies ordered according to the objective function values. A new harmony can be generated using a certain probability from either the elements in the HM or by a tone adjustment of a harmony in HM, which consists of the incorporation of a distinct musical note. The new solution is compared with the worst harmony in HM, if the new one improves the objective function ( $f(x)$ ), then it replaces the worst in  $HM$ . Otherwise, there are no changes in  $HM$ . At the end, the best harmony is returned. Algorithm 2 shows the HS pseudocode.

Simulated Annealing (SA) algorithm is a procedure for solving combinatorial optimization problems [26]. This algorithm is inspired by the physical/chemical process to produce certain alloys of metal, glass, or crys-

**Algorithm 2:** Harmony search.

---

**Data:**  $hms, HM[hms], hmcr, par$  //  $hms$  No. of candidate solutions,  $HM[hms]$  list of candidate solutions,  $hmcr$  probability of selecting a note in  $HM$ ,  $par$  probability of selecting a neighbor tone

**Result:** the best harmony of all harmonies

```

1  $t \leftarrow 0$ ;
2 Step 1 Initialize the  $HM[hms]$  randomly with
   $hms$  harmonies;
3 Step 2 Calculate the value of the objective
  function  $f(x)$  for each harmony;
4 while  $t \leq iterations_{max}$  do
5    $t \leftarrow t + 1$ ;
6   Step 3 Sort  $HM$  according to the value  $f(x)$ ;
7   Step 4 improvise a new harmony  $p^{t+1}$  and
    calculate its value  $f(x)$ ;
8    $p^{t+1} = \emptyset$ ;
9   for  $i = 1$  to  $|B|$  do
10     $r = [0, 1]$  uniformly distributed;
11    if  $r \leq hmcr$  then
12       $p_n^{t+1} = \text{select a value within } HM$ ;
13      if  $r \leq par$  then
14        take the next value, above or
        below the selected value in line
        12;
15      end
16    end
17    else
18       $p_n^{t+1} = \text{allowed range of variables}$ 
        values out of  $HM$ ;
19    end
20  end
21  Step 5 update  $HM$ ;
22  if  $f(p_n^{t+1}) > f(HM^{worst})$  then
23    include  $p_n^{t+1}$  in  $HM$  and exclude
     $HM^{worst}$  of  $HM$ ;
24  end
25 end
```

---

tal. During the annealing process, these materials are heated to a specified temperature and then cooled at a very slow and controlled rate, until they solidify into a perfect crystalline structure. In the searching process, the SA algorithm accepts not only better, but also worse neighboring solutions with a certain probability, in order to escape from local optima. Similarly to the previously reviewed metaheuristics, an SA solution represents the status of all base stations. The SA pseudocode is shown in Algorithm 3.

## 5 Fast and Efficient Cell Assignment

In this section, we describe the proposed heuristic approach to address the cell assignment problem in heterogeneous networks. It consists of an optimization process that finds the minimum number of BSs that meet the service requirements for the greatest number of users.

**Algorithm 3:** Simulated annealing.

---

**Data:**  $Temp$ ,  $\alpha$ ,  $iterations_{max}$  //  $T$  initial temperature,  $\alpha$  temperature-decrement value,  $iterations_{max}$  number of max iterations

**Result:** the best solution found

```

1  $t \leftarrow 0$ ;
2 Step 1 create a initial solution  $x$  ;
3 Step 2 Initialize  $Temp$  with a large value.;
4 while  $t \leq iterations_{max}$  do
5    $t \leftarrow t + 1$ ;
6   Step 3 Apply a random perturbations to the
   state  $x$ ,  $x' = x + \delta x$ ;
7   Step 4 Evaluate  $x'$ ;
8   if  $f(x') > f(x)$  then
9      $x \leftarrow x'$ ;
10  end
11  else
12    accept the new state with probability
     $P = e^{-\frac{f(x) - f(x')}{Temp}}$ ;
13  end
14  Step 5 Update the temperature taking into
    account  $\alpha$ ,  $temp = temp * \alpha$ .;
15 end
16 return the best solution found;
```

---

## 5.1 Cell Assignment Overview

To devise a solution for the cell assignment problem, we propose a heuristic algorithm based on the one proposed in [21]; namely, we opted for changing the BSs contained in  $\mathcal{B}$  to reevaluate the number of satisfied users, instead of altering the users between assignments. For this reason, the algorithm focuses on finding the subset  $\mathcal{B}' \subseteq \mathcal{B}$  achieving the best assignment.

The cell assignment approach presented in this paper differs from the state of the art (including [21]) in that it appropriately deals with the critical nature of heterogeneous environments, where the base stations have different capacities (e.g. power, resource blocks, etc) and the users have different rate requirements. In addition, our work considers the joint user assignment and resource allocation problem, exploiting a realistic model, with varying interference, which in turn is proportional to the amount of resources that are allocated in the assignments.

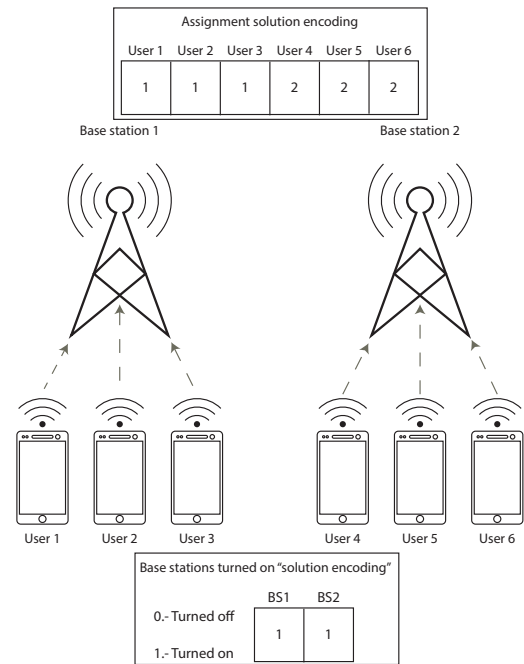
To handle the cell assignment problem, we adopted two specific representations: one for the user assignment, and another for the set of turned-on base stations. Each assignment solution is depicted by a vector  $V$  of size  $|V| = |\mathcal{U}|$  (number of users in the scenario). The element  $V_i \in V$  holds the index  $j | j \in \{1, 2, \dots, |\mathcal{B}|\}$  of the base station assigned to user  $i$ .

In the same way, the second representation is a vector  $V2$  of size  $|V2| = |\mathcal{B}|$  (number of BSs in the scenario), where each position  $V2_i \in V2$  represents the

status of the base station (0.- turned off, and 1.- turned on). An example of both representations is shown in Figure 2.

With the objective to determine the quality of each solution, we considered the following ordered list of criteria for the evaluation function:

- Maximize the number of satisfied users.
- Minimize the number of turned-on base stations.
- Maximize the profit (utility - cost).



**Fig. 2** Solution encodings

## 5.2 Detailed Description

---

**Algorithm 4:** Outline of the proposed algorithm - ExCAP-Het.

---

**Data:** Users, requirements and location. Base stations, capacity and location.

**Result:** The best found assignment  $s^*$ .

```

1 /* THE BEST ASSIGNMENT USING  $\mathcal{N}$  */
2  $s^* = \text{Assignment}(Users, BS)$ ;
3 /* TURN OFF OF BASE STATIONS */
4  $s^* = \text{TurnOff}(s^*)$ ;
5 return  $s^*$ ;
```

---

In this section we describe the ExCAP-Het algorithm proposed in this paper. For readability purposes, the outline of the ExCAP-Het algorithm is depicted



(high level perspective) in Algorithm 4. The details of the two main functions, namely the *Assignment* (see line 2 in Algorithm 4) function and the *Turn-off* function (see line 4 in Algorithm 4) are described hereafter.

### 5.2.1 Assignment Function

This function determines which base stations should be turned on ( $\mathcal{B}' \subseteq \mathcal{B}$ ). It also finds the best assignment, i.e. the one that meets the requirement of the maximum number of users at the lowest cost.

The following steps are needed to obtain and evaluate a particular assignment.

1. **Determine the Distance Matrix.** In this step, we calculate the distance of each user to every base station. Line 2 of Algorithm 5.
2. **Determine the LoS Matrix.** To know whether or not user  $i$  has a line of sight (LoS, NLoS respectively) to base station  $j$ , the heuristic computes the probabilities for every user-base station pair, considering the probability functions of Table 3. Line 4 of Algorithm 5.
3. **Determine the Power Matrix.** In this step, the heuristic computes for each user the power received from each base station,  $\gamma_{i,j}$ , taking into account the *LoS Matrix* and the propagation loss. We consider a propagation model for each type of BS (see Table 2). Line 5 of Algorithm 5.
4. **Pre-assign users to base stations.** The heuristic pre-assigns each user to the base station with the highest received power. The base stations that have, at least, one assigned user is considered to be turned on. These base stations create the subset ( $\mathcal{B}' \subseteq \mathcal{B}$ ). Lines 6-7 of Algorithm 5.
5. **Calculate the SNR and Resource Blocks (RB).** In this step the heuristic performs a resource allocation to satisfy demand under ideal conditions, i.e. without interference. Lines 9-11 of Algorithm 5.
  - 5.1. **Determine the SNR Matrix.** First, the heuristic calculates the value of *SNR* received by user  $i$  from base station  $j$  without considering the interference from the rest of the base stations in the system:  $SNR = \frac{\gamma_{ij}}{\sigma^2}$ .
  - 5.2. **Determine the RB needed.** Then, we compute the resource blocks needed by the base station  $j$  to meet the demand of assigned user  $i$ , taking into account its value of *SNR*.
6. **Calculate the real values of the pre-assignment with interference.** We calculate the RB needed to meet user demand using the *SINR*, i.e. mutual interference is considered. Matrix Lines 13-16 of Algorithm 5
  - 6.1. **Determine the SINR Matrix.** We calculate the *SINR* received by each user (see Eq. 2), taking into account the interference generated by the other turned-on base stations and the resource blocks calculated in step 5.2.
  - 6.2. **Determine the RB needed.** We recalculate the resource blocks (RBs) needed by the base station  $j$  to meet the demand of assigned user  $i$ , taking into account the *SINR* value (i.e. considering mutual interference).
  - 6.3. **Determine the values of thermal noise.** We calculate the thermal noise figure considering the percentage of resources blocks needed by the base station  $j$  to meet the demand of user  $i$ .
  - 6.4. **Determine the SINR Matrix.** We calculate the *SINR* received by each user (see equation 2), taking into account the interference generated by the other turned-on base stations, the thermal noise and the resource blocks calculated in step 6.2.
7. **Assign users iteration.** The capacity of the BSs may be exceeded due to many factors: large number of users, high interference causing the provision of more RBs, far away users, etc. The algorithm performs iterations to reallocate users who cannot be served due to the lack of resource blocks. Lines 17-19 of algorithm 5

**Table 2** Propagation models

Scene	Range (m)	$L$ (dB)	$\sigma$ (dB)
$Ma_{NLoS}$	$10 \leq d \leq 5000$	$139.1033 + 39.0864 \cdot (\log_{10}(d) - 3)$	6
$Ma_{LoS}$	$10 \leq d \leq 328.4211$ $328.4211 \leq d \leq 5000$	$36.2995 + 22 \cdot \log_{10}(d)$ $40 \cdot \log_{10}(d) - 10.7953$	6 4
$Pi_{NLoS}$	$10 \leq d \leq 5000$	$145.48 + 37.5 \cdot (\log_{10}(d) - 3)$	3
$Pi_{LoS}$	$0 \leq d \leq 600$	$103.8 + 20.9 \cdot (\log_{10}(d) - 3)$	6

**Table 3** Probabilities of line of sight

BS	Probability function based on distance
Macro	$P_{LoS} = \min\left(\frac{18}{d}, 1\right) \cdot \left(1 - e^{-\frac{d}{36}}\right) + e^{-\frac{d}{36}}$
Pico	$P_{LoS} = 0.5 - \min\left(0.5, 5 \cdot e^{-\frac{156}{d}}\right) + \min\left(0.5, 5 \cdot e^{-\frac{d}{30}}\right)$

The iterative procedure designed to re-assign users when the resources of a base station are exceeded is described hereafter (see Algorithm 6).

1. Make up the subset  $\mathcal{B}$  from those base stations that have available resources. Line 2 Algorithm 6.

**Algorithm 5:** *Assignment* function.

---

**Data:** The location of each user and base station, and the base stations considered for the assignment.  
**Result:** An optimal assignment  $s$  using all nearest base stations to each user.

```

1 /* IDEAL VALUES (WITHOUT INTERFERENCE) */;
2 ObtainDistanceMatrix( $Users, BS$ );
3 /* DETERMINE THE LINE OF SIGHT (LoS) PROBABILITY */;
4 ObtainLosProbabilities( $Users, BS$ );
5 ObtainPowerMatrix( $Users, BS$ );
6  $s = \text{AssignmentToHighestPowerBS}(\mathit{Users}, \mathit{BS})$ ;
7  $\mathit{TurnedOnBS} = \text{TurnOnBS}(s)$ ;
8 /* CONSIDERING ONLY THE TURNED ON BS WITHOUT INTERFERENCE */;
9 ObtainSNRmaxMatrix( $Users, \mathit{TurnedOnBS}$ );
10 /* CONSIDERING THE PRE-ASSIGNMENT MATRIX */;
11 ObtainNeededRB( $s$ );
12 /* REAL VALUES (WITH INTERFERENCE) */;
13 ObtainSINRMatrix( $Users, \mathit{TurnedOnBS}$ );
14 /* CONSIDERING THE PRE-ASSIGNMENT MATRIX */;
15 ObtainNeededRB( $s$ );
16  $\mathit{ExceededBS} = \text{CalculateCosts}(s)$ ;
17 /* REASSIGNMENT USERS */;
18 if  $|\mathit{ExceededBS}| > 1$  then
19   |  $s = \text{ReassignmentUsers}(s)$ ;
20 end
21 return  $s$ ;
```

---

2. For all base stations that exceeded resources repeat Steps 1.1 and 1.2. Line 3 Algorithm 6.
  - 2.1. Sort assigned users in ascending order with respect to costs.
  - 2.2. ‘‘Select best users’’ (i.e. those with highest  $SINR$ ) and put aside those who cannot be served due to lack of resources (RBs).
3. If  $|\mathcal{B}| \geq 1$  execute Steps 3.1 and 3.2. Lines 7-16 Algorithm 6.
  - 3.1. Reallocate users without service to the base station (from  $\mathcal{B}$ ) with the highest SINR.
  - 3.2. Reset all system values and recalculate them, considering the new user allocation.
4. If any base station of  $\mathcal{B}$  exceeded its resources go to Step 1.

### 5.2.2 The Turn Off Function

After the *Assignment* function finds the best assignment  $s^*$  from the set  $\mathcal{B}$ , the algorithm tries to outperform the solution ( $\mathcal{B}' \subseteq \mathcal{B}$ ) turning off BSs from  $\mathcal{B}'$ . Therefore, a new set ( $\mathcal{B}'' \subseteq \mathcal{B}'$ ) is found as a result of turning off the base station  $i$ ,  $\forall i \in \{1, 2, \dots, |\mathcal{B}'|\}$  base station.

On each iteration, this function turns off one BS from the set  $\mathcal{B}'$ , such that  $|\mathcal{B}''| = |\mathcal{B}'| - 1$ , then the assignment  $s''$  (from  $\mathcal{B}''$ ) is compared with the best

**Algorithm 6:** Outline of the *ReassignmentUsers* function.

---

**Data:** The current assignment  $s$ .  
**Result:** A new assignment for users that were not served by lack of resources.

```

1 while  $|\mathit{ExceededBS}| > 1$  do
2    $\mathcal{B} = \text{IdentifyAvailableBS}(\mathit{TurnedOnBS})$ ;
3    $\mathit{NonServedRB} = \text{SelectBestUsers}(\mathit{ExceededBS})$ ;
4   if  $|\mathit{ExceededBS}| == 0$  OR  $|\mathcal{B}| == 0$  then
5     | break;
6   end
7    $s = \text{ReassignmentToSINR}^{\max}(\mathit{NonServedRB}, \mathcal{B})$ ;
8   ResetSystemValues();
9   /* CALCULATE THE IDEAL VALUES */;
10  ObtainSNRmaxMatrix( $Users, \mathit{TurnedOnBS}$ );
11  ObtainNeededRB( $s$ );
12  /* CALCULATE THE REAL VALUES */;
13  ObtainSINRmaxMatrix( $Users, \mathit{RB}, \mathit{TurnedOnBS}$ );
14  ObtainNeededRB( $s$ );
15  ObtainTN( $s, \mathit{RB}$ );
16  ObtainSINRMatrix( $Users, \mathit{RB}, \mathit{TN}, \mathit{TurnedOnBS}$ );
17   $\mathit{ExceededBS} = \text{CalculateCosts}(s)$ ;
18 end
19 return  $s$ ;
```

---

solution  $s^*$ . If the solution  $s''$  improves the best solution  $s^*$ , the best solution is updated and  $\mathcal{B}'$  is also updated with the set  $\mathcal{B}''$  of the new solution  $s^*$ . The algorithm, including the turn off iterations with the updated set  $\mathcal{B}'$ , is shown in Algorithm 7.

The above procedure is based on the fact that turning-off a base station may result in lower interference among the remainder set  $\mathcal{B}'$ , improving the  $SINR$  of users and reducing the number resource blocks needed for BSs to serve users, and consequently increasing the number of served users.

## 6 Evaluation Setup

This section describes the evaluation setup. We initially describe the testing instances and their characteristics. Following on, we describe the execution platform used to carry out the analysis and finally, the metaheuristics parameters used to generate the data for comparing the performance of our metaheuristic approach.

### 6.1 Scenario Parameters

The simulations have been run in 16 different testing instances (see Table 4). Each instance is composed of the set  $\mathcal{B}$ , with two subsets of base stations  $\mathcal{B}_1 \cup \mathcal{B}_2 = \mathcal{B}$ , consisting of Macro and Pico Base Stations respectively. On the other hand, users  $\mathcal{U}$  were increased from 50 to

**Algorithm 7:** Outline of the *TurnOff* function.

---

**Data:** The turned-on base stations from  $s^*$ .  
**Result:** The best-found assignment  $s^*$ .

```

1 repeat
2   Improved = 0;
3   TurnedOnBS = TurnOnBS( $s^*$ );
4   forall TurnedOnBS do
5     /* TURN OFF EACH BS FROM  $s^*$  */;
6     TurnedOnBSnew =
7       TurnOffBS(TurnedOnBS);
8      $s$  = Assignment(Users, TurnedOnBSnew);
9     Profit = ObtainProfit( $s$ );
10    if NumOfServedC > NumOfServedC* then
11       $s^*$  =  $s$ ;
12      Improved = 1;
13    else
14      if NumbOfServedC == NumbOfServedC*
15        AND Profit > Profit* then
16         $s^*$  =  $s$ ;
17        Improved = 1;
18    end
19  end
20 until Improved = 0;
21 return  $s^*$ ;

```

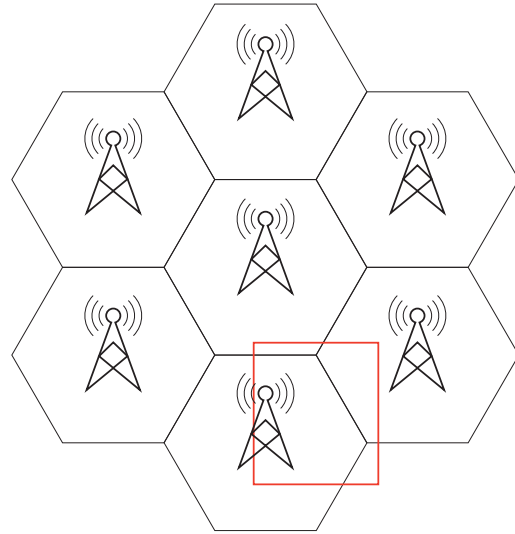
---

200 in each subset of  $\mathcal{B}$  for the different instances. All users had fixed capacity demand of 1.75 Mbps.

**Table 4** Testing instances characteristics.

Instance	# Users		# BSs	
	Macro	Pico	Macro	Pico
1	50	50	7	20
2	50	100	7	20
3	50	150	7	20
4	50	200	7	20
5	100	50	7	20
6	100	100	7	20
7	100	150	7	20
8	100	200	7	20
9	150	50	7	20
10	150	100	7	20
11	150	150	7	20
12	150	200	7	20
13	200	50	7	20
14	200	100	7	20
15	200	150	7	20
16	200	200	7	20

The number of BSs used for the experiment was set to  $|\mathcal{B}| = 27$ , 7 were Macro BSs, while the rest were 20 Pico BSs. Figure 3 illustrates the locations of the Macro base stations, while the Pico BSs were randomly placed in the red square between the Macro BSs in the lower-right quadrant of the Figure. Each Macro BS had coverage of 4.5 Km, and each Pico one had a limited coverage of 600 meters. The users were randomly placed on the covered area by the Macro and Pico BSs. Table 5 summarizes the parameters used in each testing instance.

**Fig. 3** Locations of users and base stations.**Table 5** Scenario parameters

Parameter	Value
<b>Scenario parameters</b>	
Central frequency ( $f_c$ )	2.6 GHz
Avg. streets width ( $W$ )	20 m
Avg. buildings height ( $h$ )	20 m
Avg. base stations height ( $h_{BS}$ )	25 m
Avg. user terminals height ( $h_{UT}$ )	1.5 m
<b>Device parameters</b>	
BS $T_x$ gain ( $G_{TX}$ )	15.0 dB
UT $R_x$ gain ( $G_{RX}$ )	0.0 dB
( $MCL$ )	-70.0 dB
Macro $T_x$ power ( $P_{TX_{MA}}$ )	46.0 dBm
Pico $T_x$ power ( $P_{TX_{PI}}$ )	36.0 dBm
<b>Noise parameters</b>	
Thermal noise density ( $f$ )	-173.9772 dBm
Noise figure ( $NF$ )	6.0 dB
Thermal noise figure ( $TN$ )	$NF + 10 * \log_{10} *$ $(\alpha_{i,j} * 180kHz) + f$
Control Overhead ( $CO$ )	1.0 dB

## 6.2 Execution Platform

ExCAP-Het and the metaheuristics algorithms were implemented in C programming language following *c99* standard (compilation has been done with *GCC-4.6.3*). For statistical validation purposes, we executed thirty-one runs of each algorithm for each instance of Table 4. This analysis demands a considerable amount of computing resources. For this reason, the experiments were carried out in a platform with 64 cores Intel Xeon CPU E5-4620 2.20GHz, 512 GB RAM, and Ubuntu 14.04 LTS 64-bit Operating System.

### 6.3 Metaheuristics parameters

The parameters used for the metaheuristic algorithms are: crossover probability  $p_c = 0.9$ , and a mutation probability  $p_m = 0.05$  for the genetic algorithm. For harmony search, we set the probabilities of selecting a note in HM ( $hmcr$ ) and selecting a neighbor tone ( $par$ ) to 0.9 and 0.1, respectively. The population size for both metaheuristics was set to 10. The temperature used for simulated annealing was 100, and the temperature-decrement value was 0.88. We set the stop criterion in 1000 evaluations of the objective function to make a fair comparison.

## 7 Evaluation Results and Discussion

This section presents the cell assignment results, a statistical analysis of the assignment results and a complexity discussion to justify the outperforming performance of the proposed approach.

### 7.1 Cell assignment results

All instances were executed using our algorithm with two versions: (1) ExCAP-Het without using the TurnOff function, which is similar to a standard best-effort solution (No switching off cells); (2) ExCAP-Het using the TurnOff function (Switching-off cells). Then, we compared our algorithm with the three metaheuristics described above, all solving the optimization problem.

To analyze the results, we grouped the test instances according to the amount of users located in each macro cell, yielding four groups: small density - 50 users in instances 1 to 4; medium density - 100 users in instances 5 to 8; high density - 150 users in instances 9 to 12; and very high density - 200 users in instances 13 to 16. For each group, inside the area covered by the pico cells we placed 50, 100, 150 and 200 users yielding sixteen test instances (see Table 4). The metrics used in this paper are the percentage of served users, profit, energy savings (active cells), and execution time.

The average number of served users after 31 executions of the five algorithms is graphically depicted in Figures 4 to 7. The results show that the performance of all algorithms worsens when the number of users increases. We can see that the algorithm with the worst performance in most tests is the one labelled as *No switching-off*. The algorithm with the best performance according this metric is based on the metaheuristic harmony search “HS”, with a slightly better performance in instances 5 and 13 (tests with the lowest number of users in pico cells). Otherwise, the

“HS” approach and the genetic algorithm, “GA”, exhibit the same performance, and compared with our turn-off function *Switching-off* approach, their performance is rather alike. Even in tests instances 5 and 13, our approach is very competitive in the average number of served users metric. More importantly, our approach shows a slightly better performance in instances 2, 6, 7, 11 and 16. This demonstrates that it is efficient in terms of the quality of assignments produced, when compared with metaheuristic-based approaches, which are known to yield efficient assignments.

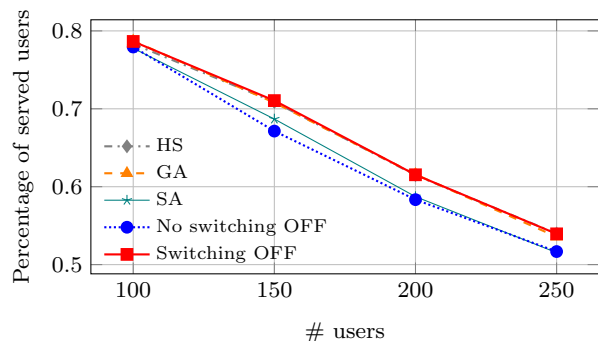


Fig. 4 Comparison of served users in instances 1 to 4.

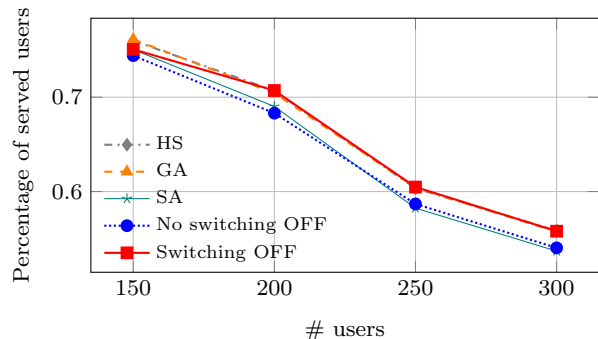


Fig. 5 Comparison of served users in instances 5 to 8.

Figures 8 to 11 show the average of 31 executions of the algorithms in terms of the profit metric. At a first glance, the worst performing algorithm in this metric is the *No-switching-off* in all instances, which means that it is unable to properly manage the resources. On the other hand, our *Switching-off* approach presents very competitive results in this metric, outperforming in all instances the HS-based approach, and performing very close for most instances to the SA- and GA-based schemes. It is worth mentioning that both, the metaheuristic and our *Switching-off* approach present the same behavior in this metric, which means that our switching-off approach succeeds in managing

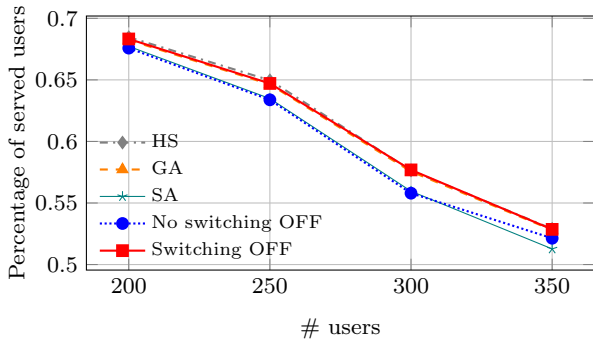


Fig. 6 Comparison of served users in instances 9 to 12.

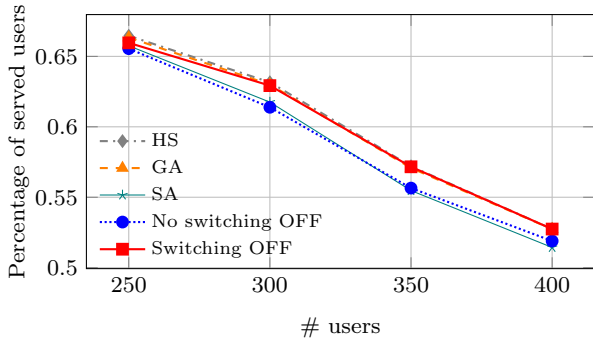


Fig. 7 Comparison of served users in instances 13 to 16.

the resources similarly to the metaheuristic-based approaches, albeit exploring a small fraction of the search space in the combinatorial problem, as we will see in the execution time results. The small differences in the performance of the metaheuristic-based and our *Switching-off* approach in this metric are due to the different assignments found by each algorithm, i.e. assignments with the same number of users can have different costs to meet users' demand. The algorithm with the best performance in more instances is based on a genetic algorithm (GA).

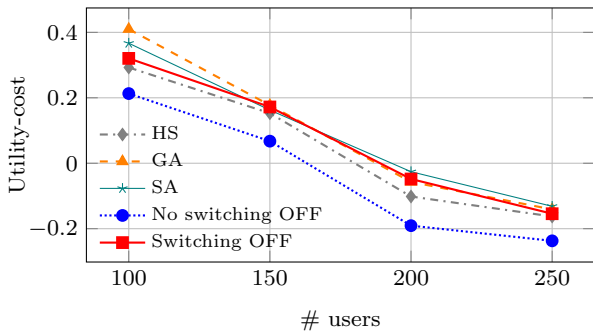


Fig. 8 Comparison of profit in instances 1 to 4.

The performance results in terms of the metric active-cells are depicted in Figures 12 to 15. The results show

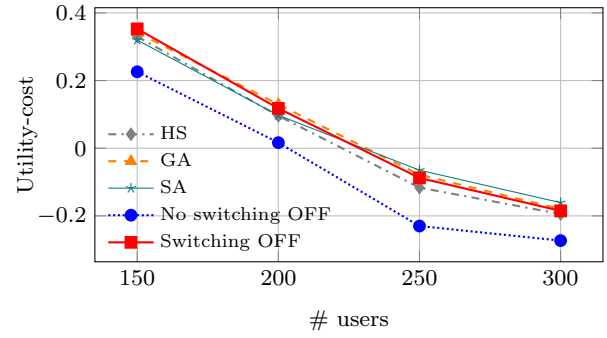


Fig. 9 Comparison of profit in instances 5 to 8.

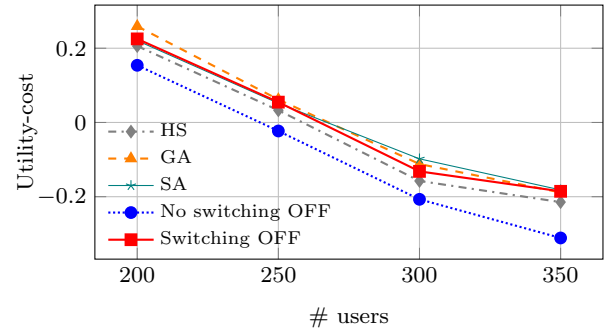


Fig. 10 Comparison of profit in instances 9 to 12.

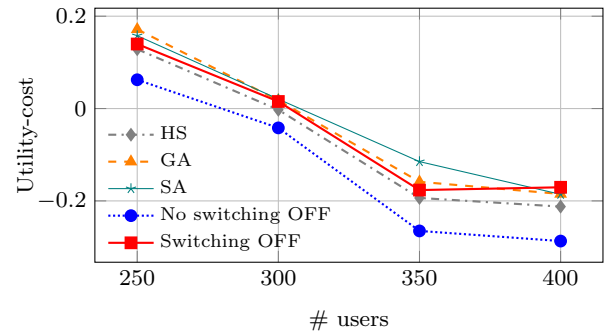


Fig. 11 Comparison of profit in instances 13 to 16.

that the *No-switching-off* algorithm needs more base stations than the other algorithms to satisfy a lower number of users in all instances, which involves a high energy consumption. For this reason the *No-switching-off* is the algorithm with the worst performance for this metric. The algorithm with the best performance in almost all instances is the GA. However, our approach has practically the same performance than the GA algorithm in instances 2, 3, 7, 10, 12, 15 and 16. In instances like 3, 4, 7, 8, and 15, the best algorithm was the simulated annealing "SA" approach, which yields poor results, like the *No-switching-off* approach in the served users metric. Finally, it is important to mention that our algorithm outperforms the "HS" algorithm in all cases, which in turn produced slightly better results in the served users metric.

Finally, we present results of how much time is spent (execution time) on average for each algorithm, to solve each instance (see Figures 16 to 19). In terms of this metric, we can see that all metaheuristics have a high time consumption in comparison with the approach proposed in this paper, with a difference of an order of magnitude. This demonstrates that our *Switching-off* approach is efficient, yielding similar results to traditional metaheuristic-based solutions in key metrics like served users, profit and active cells, while requiring much shorter time. Considering the metric execution time, the best algorithm is *No-switching-off*, and the worst is the metaheuristic harmony search (HS). Nevertheless, lower execution times are signs of better performance in such metric only. Low times do not necessarily lead to better cell assignments, better profit or, better energy-savings.

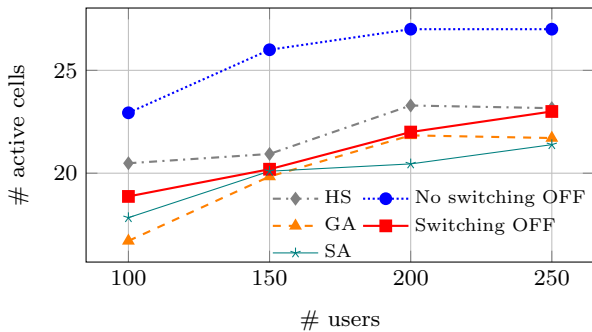


Fig. 12 Comparison of active cells in instances 1 to 4.

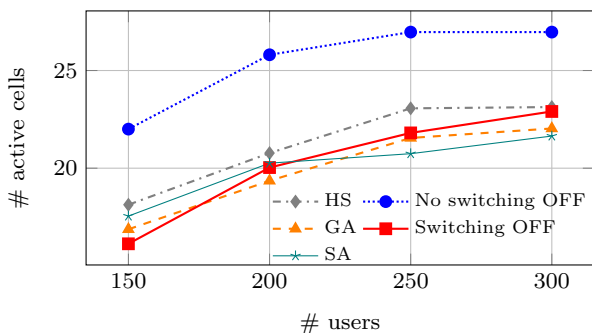


Fig. 13 Comparison of active cells in instances 5 to 8.

## 7.2 Statistical Analysis

This section presents a formal analysis of the results provided to determine if there is a significant difference between the performance of the studied algorithms. For this purpose, we have used statistical significance tests.

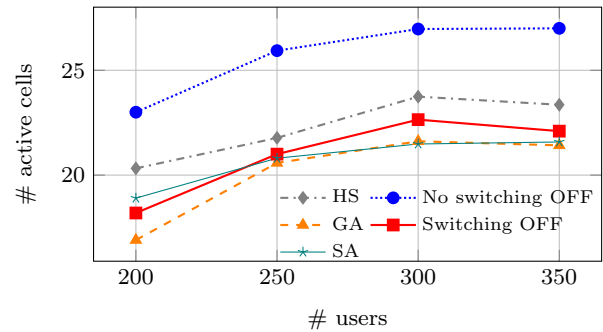


Fig. 14 Comparison of active cells in instances 9 to 12.

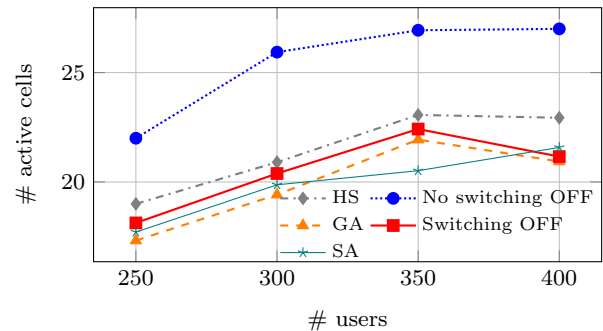


Fig. 15 Comparison of active cells in instances 13 to 16.

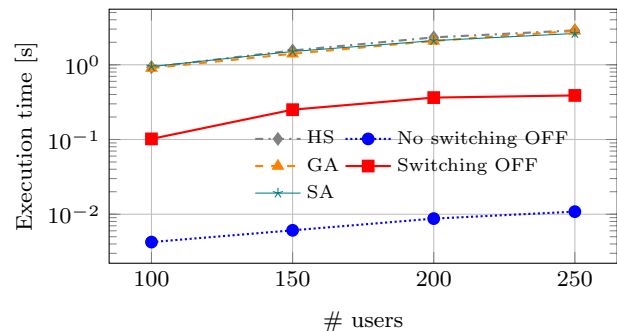


Fig. 16 Comparison of execution time in instances 1 to 4.

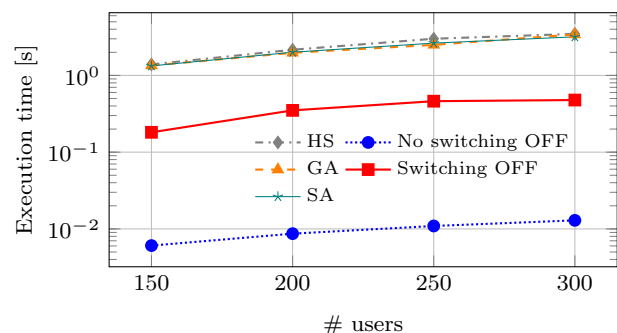


Fig. 17 Comparison of execution time in instances 5 to 8.

Our intention here is to analyse and detect statistical significance on the performance of the algorithms for each metric. To accomplish this, we applied the Friedman test [12], which is a non-parametric statistical test

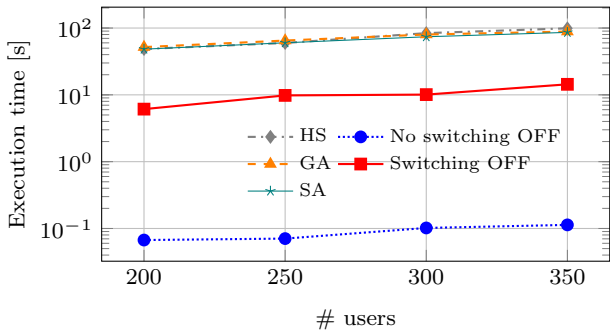


Fig. 18 Comparison of execution time in instances 9 to 12.

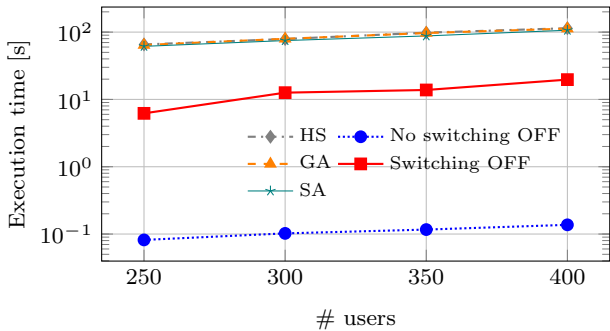


Fig. 19 Comparison of execution time in instances 13 to 16.

that detects differences in treatments across multiple attempts.

The procedure of Friedman test begins by rank-ordering the algorithms (1 to  $k$  algorithms) for each instance (row). The smallest score corresponds to the algorithm with the best performance, and the biggest score to the algorithm with the worst performance. To detect an overall statistically significant difference between the performance of the algorithms, a hypothesis and its corresponding null hypothesis are formulated: the null hypothesis ( $H_0$ ) suggests that all sets of ranks come from algorithms with identical performance, i.e. the algorithms do not have a statistical significance difference. The alternate hypothesis ( $H_a$ ) holds that, at least, one algorithm has a different performance. Finally, a significance level of  $\alpha = 0.05$  is considered.

For each metric, we applied the Friedman test to find out whether there is statistical difference between the algorithms and to rank their performance. The average rank of the Friedman test obtained by the algorithms in each metric is shown in Table 6. The algorithms with the best performance for each metric are underlined and in bold. The Friedman test proved that the algorithms in all metrics have statistically different results. However, we only know that there are differences somewhere between the related algorithms, but we do not know exactly where those differences lie. To examine where the differences actually occur, we ap-

plied the Wilcoxon test to our data, using the Bonferroni adjustment [4], to detect the algorithms that perform significantly different compared with the performance achieved by the best algorithm in each metric, namely “HS” in served users, “GA” in profit and active cells, and *No – switching – off* in execution time.

Table 6 Average ranks of Friedman test for each metric

Metric	Algorithms				
	GA	SA	HS	Off	No off
Served users	2.44	4.44	<b><u>1.66</u></b>	1.91	4.56
Profit	<b><u>1.75</u></b>	1.94	3.94	2.38	5.00
Active cells	<b><u>1.44</u></b>	1.88	4.00	2.69	5.00
Execution time	3.88	3.44	4.69	2.00	<b><u>1.00</u></b>

The results of this test are presented in Table 7. Each cell in this table compares the algorithm for such column with the best algorithm of the row (metric). For example, consider that we want to compare algorithms A and B, the latter being the best in a given metric. If a significant performance difference exists between A and B for a particular metric, the corresponding cell of algorithm A is marked with  $\checkmark$  otherwise it is marked with  $\times$ .

Table 7 Statistically significant difference and performance rank of algorithms for each metric

Metric	Algorithms				
	GA	SA	HS	Off	No off
Served users	2.44 $\times$	4.44 $\checkmark$	<b><u>1.66</u></b>	1.91 $\times$	4.56 $\checkmark$
Profit	<b><u>1.75</u></b>	1.94 $\times$	3.94 $\checkmark$	2.38 $\times$	5.00 $\checkmark$
Active cells	<b><u>1.44</u></b>	1.88 $\times$	4.00 $\checkmark$	2.69 $\checkmark$	5.00 $\checkmark$
Execution time	3.88 $\checkmark$	3.44 $\checkmark$	4.69 $\checkmark$	2.00 $\checkmark$	<b><u>1.00</u></b>

The Bonferroni statistically significant test shows that our *Switching – off* approach does not perform significantly different compared to the algorithms with the best performance in key metrics like served users and profit. This is of particular relevance because it demonstrates that our heuristic approach yields a performance comparable to the best metaheuristic-based approaches in these key metrics. Regarding the active cells metric, our *Switching – off* approach performs significantly different to the best metaheuristic (GA), which in turn performs similar to “SA” metaheuristic. Moreover, our *Switching – off* approach outperforms metaheuristic “HS” and, naturally, the *No–switching–off* solution with statistical significance in the active cells metric. Finally, our approach is significantly faster than all traditional metaheuristics.



### 7.3 Computational complexity

During the executions of the algorithms, the computations of all the elements in the model described in Section 3 are carried out in the same manner for all the algorithms evaluated (i.e. ExCAP-Het and the three metaheuristics). For the optimization problem addressed in this paper, the pivotal element that determines the complexity of the algorithms is the procedure for determining the base stations that are *on* and *off*.

ExCAP-Het defines an initial pre-assignment (based on the SINR of each user), which will determine a set of base stations that will be initially *on*. From this set, ExCAP-Het proceeds to turn off base stations on a one-by-one basis. For example, if an assignment under evaluation has 5 active base stations, BS 1 would be switched off, and the assignment would be evaluated with the remaining 4. If the evaluation with the 4 BSs is better than the previous best, that assignment is considered the new best solution. In the next iteration the base station 1 would be turned *on* again, and BS 2 would be turned *off*, to evaluate the assignment under such conditions, and so on. In the worst case, the initial assignment would be to turn on all the base stations, so this process would consist of traversing the entire set of base stations  $|\mathcal{B}|$ . On the other hand, if an improvement occurs when switching off a given base station, the previous process will be repeated with the reduced subset  $B' = |\mathcal{B}| - 1$ . In the worst case, the second process will be repeated until the subset is equal to 1, that is, when all base stations are turned off except one, which in turn would cause the evaluation to be executed  $|\mathcal{B}|$  times. Based on the above, the computational complexity of ExCAP-Het would be in the worst case  $\mathcal{O}(|\mathcal{B}|^2 A) = \mathcal{O}(|\mathcal{B}|^2 A)$ , where  $\mathcal{B}$  is the subset of base stations and  $A$  is the cost of the initial allocation.

Due to their search space exploration nature, metaheuristics perform a number of evaluations of the objective function given by the product of the population size and the number of generations. The metaheuristics use a binary representation for the base stations in *on* mode. Thus, in the worst case, the metaheuristic can generate all possible binary chains of size  $|\mathcal{B}|$ , that is  $2^{|\mathcal{B}|}$  evaluations. Therefore the computational complexity of the metaheuristics are in the worst-case  $\mathcal{O}(2^{|\mathcal{B}|} A)$ , where  $|\mathcal{B}|$  is the number of base stations and  $A$  is the cost (constant) of evaluating an assignment.

## 8 Concluding Remarks and Future Work

This paper has presented a fast and efficient cell assignment approach for heterogeneous networks with en-

ergy expenditure reduction to foster cell switching off. A joint cell assignment and resource allocation model was defined as a pivotal element for the formulation of a complex combinatorial optimization problem. The proposed solution is intended to be applied in centralized systems such as those brought by C-RAN and SDN. To solve the problem, a heuristic was defined to address the optimization problem, which has also been solved with three of the most widespread metaheuristics, namely Genetic Algorithm (GA), Harmony Search (HS) and Simulated Annealing (SA). Through extensive simulations and formal statistical analysis we have demonstrated that it is possible to achieve cell assignments close to the optimum, compared to those achieved with metaheuristic-based approaches at much lower execution times, thus favoring practical applicability. Concretely, our heuristic approach allows the production of cell assignments of the same quality as those provided by the metaheuristic of the best performance in key metrics like served users and profit. Regarding energy expenditure reduction, our heuristic provides competitive results, yielding energy-savings comparable to the three metaheuristics evaluated. Overall, we have shown that the performance of the proposed heuristic is similar to the best metaheuristic approach, with statistical significance at much lower execution times.

Our future work will be directed towards exploring other approaches to maximize the performance of the cell assignment problem as well as evolving the problem into a multi-objective one, to explore other computational methods to maximize the use of radio resources. In addition, part of our future work will be to analyze the performance of the cell assignment in more complex scenarios. In particular, we will include elastic services and users with multiple traffic demands, which undoubtedly would increase the complexity of the problem. We will also consider other aspects such as mobility and handovers, which may have a strong impact on the users perceived QoS, and which would impose additional constraints to our problem instances. The proposed scheme will be used to solve individual snapshots, where both user positions and their willingness to establish a connection will be updated based on the corresponding mobility and traffic patterns. Other use cases, also leveraged by SDN, such as the one discussed in [28], will be also considered.

## 9 Acknowledgment

This paper has been supported by the National Council of Research and Technology (CONACYT) through grant FONCICYT/272278 and the ERANetLAC (Network of the European Union, Latin America, and the



Caribbean Countries) project ELAC2015/T100761. This paper is partially supported also by the ADVICE project, TEC2015-71329 (MINECO/FEDER) and the European Union's Horizon 2020 research and innovation programme under grant agreement No 777067 (NECOS project).

## References

1. Abdelnasser A, Hossain E (2016) Resource allocation for an ofdma cloud-ran of small cells underlying a macrocell. *IEEE Transactions on Mobile Computing* 15(11):2837–2850, DOI 10.1109/TMC.2015.2513052
2. Andrews JG, Buzzi S, Choi W, Hanly SV, Lozano A, Soong ACK, Zhang JC (2014) What Will 5G Be? *IEEE Journal on Selected Areas in Communications* 32(6):1065–1082, DOI 10.1109/JSAC.2014.2328098
3. Bhushan N, Li J, Malladi D, Gilmore R, Brenner D, Damnjanovic A, Sukhavasi R, Patel C, Geirhofer S (2014) Network densification: the dominant theme for wireless evolution into 5g. *IEEE Communications Magazine* 52(2):82–89, DOI 10.1109/MCOM.2014.6736747
4. Bland JM, Altman DG (1995) Multiple significance tests: the bonferroni method. *Bmj* 310(6973):170
5. Blum C, Roli A (2003) Metaheuristics in combinatorial optimization: Overview and conceptual comparison. *ACM Comput Surv* 35(3):268–308, DOI 10.1145/937503.937505, URL <http://doi.acm.org/10.1145/937503.937505>
6. Boche H, Naik S, Alpcan T (2011) Characterization of convex and concave resource allocation problems in interference coupled wireless systems. *Signal Processing*, *IEEE Transactions on* 59(5):2382–2394, DOI 10.1109/TSP.2011.2112652
7. Bu T, Li L, Ramjee R (2006) Generalized proportional fair scheduling in third generation wireless data networks. In: *Proceedings IEEE INFOCOM 2006. 25TH IEEE International Conference on Computer Communications*, pp 1–12, DOI 10.1109/INFOCOM.2006.145
8. CISCO (2017) Cisco visual networking index: Global mobile data traffic forecast update. URL [www.cisco.com/c/en/us/solutions/collateral/service-provider/visual-networking-index-vni/white-paper-c11-741490.pdf](http://www.cisco.com/c/en/us/solutions/collateral/service-provider/visual-networking-index-vni/white-paper-c11-741490.pdf)
9. Coello CAC, Lamont GB, Van Veldhuizen DA, et al (2007) *Evolutionary algorithms for solving multi-objective problems*, vol 5. Springer
10. Dréo J, Pétrowski A, Siarry P, Taillard E (2006) *Metaheuristics for hard optimization: methods and case studies*. Springer Science & Business Media
11. Feng M, Mao S, Jiang T (2017) Base station on-off switching in 5g wireless networks: Approaches and challenges. *IEEE Wireless Communications* 24(4):46–54, DOI 10.1109/MWC.2017.1600353
12. Friedman M (1937) The use of ranks to avoid the assumption of normality implicit in the analysis of variance. *Journal of the american statistical association* 32(200):675–701
13. Geem ZW, Kim JH, Loganathan GV (2001) A new heuristic optimization algorithm: harmony search. *simulation* 76(2):60–68
14. Goldberg DE, Holland JH (1988) Genetic algorithms and machine learning. *Machine learning* 3(2):95–99
15. Mesodiakaki A, Adelantado F, Antonopoulos A, Alonso L, Verikoukis C (2016) Energy and spectrum efficient user association in 5g heterogeneous networks. In: *2016 IEEE 27th Annual International Symposium on Personal, Indoor, and Mobile Radio Communications (PIMRC)*, pp 1–6, DOI 10.1109/PIMRC.2016.7794877
16. Mesodiakaki A, Zola E, Kassler A (2017) User association in 5g heterogeneous networks with mesh millimeter wave backhaul links. In: *2017 IEEE 18th International Symposium on A World of Wireless, Mobile and Multimedia Networks (WoWMoM)*, pp 1–6, DOI 10.1109/WoWMoM.2017.7974342
17. Mogensen P, Na W, Kovacs IZ, Frederiksen F, Pokhariyal A, Pedersen KI, Kolding T, Hugl K, Kususela M (2007) Lte capacity compared to the shannon bound. In: *Vehicular Technology Conference, 2007. VTC2007-Spring*. *IEEE 65th*, pp 1234–1238, DOI 10.1109/VETECS.2007.260
18. Peng M, Wang C, Lau V, Poor HV (2015) Fronthaul-constrained cloud radio access networks: insights and challenges. *IEEE Wireless Communications* 22(2):152–160, DOI 10.1109/MWC.2015.7096298
19. Qualcomm (2013) The 1000x data challenge. URL <https://www.qualcomm.com/invention/1000x>
20. Ravanshid A, Rost P, Michalopoulos DS, Phan VV, Bakker H, Aziz D, Tayade S, Schotten HD, Wong S, Holland O (2016) Multi-connectivity Functional Architectures in 5G. In: *2016 IEEE International Conference on Communications Workshops (ICC)*, pp 187–192
21. Rubio-Loyola J, Gonzalez-Hernandez L, Diez L, Agüero R, Serrat J (2014) An energy-oriented optimization algorithm for solving the cell assignment problem in 4g-lte communication networks. In: *2014 IFIP Wireless Days (WD)*, pp 1–4, DOI 10.1109/WD.2014.7020851

22. Seng S, Li X, Ji H, Zhang H (2018) Joint access selection and heterogeneous resources allocation in udns with mec based on non-orthogonal multiple access. In: 2018 IEEE International Conference on Communications Workshops (ICC Workshops), pp 1–6, DOI 10.1109/ICCW.2018.8403502
23. Siomina I, Yuan D (2012) Load balancing in heterogeneous lte: Range optimization via cell offset and load-coupling characterization. In: 2012 IEEE International Conference on Communications (ICC), pp 1357–1361, DOI 10.1109/ICC.2012.6364075
24. Talbi EG (2009) Metaheuristics: from design to implementation, vol 74. John Wiley & Sons
25. Tan Z, Li X, Yu FR, Chen L, Ji H, Leung VCM (2017) Joint access selection and resource allocation in cache-enabled hcns with d2d communications. In: 2017 IEEE Wireless Communications and Networking Conference (WCNC), pp 1–6, DOI 10.1109/WCNC.2017.7925732
26. Van Laarhoven PJ, Aarts EH (1987) Simulated annealing. In: Simulated annealing: Theory and applications, Springer, pp 7–15
27. Wu D, Wu Q, Xu Y, Liang Y (2017) Qoe and energy aware resource allocation in small cell networks with power selection, load management, and channel allocation. IEEE Transactions on Vehicular Technology 66(8):7461–7473, DOI 10.1109/TVT.2017.2650949
28. Wu D, Nie X, Asmare E, Arkhipov D, Qin Z, Li R, McCann J, Li K (2018) Towards distributed sdn: Mobility management and flow scheduling in software defined urban iot. IEEE Transactions on Parallel and Distributed Systems pp 1–1, DOI 10.1109/TPDS.2018.2883438
29. Xiao Z, Li T, Ding W, Wang D, Zhang J (2016) Dynamic pci allocation on avoiding handover confusion via cell status prediction in lte heterogeneous small cell networks. Wirel Commun Mob Comput 16(14):1972–1986, DOI 10.1002/wcm.2662, URL <https://doi.org/10.1002/wcm.2662>
30. Xiao Z, Liu H, Havyarimana V, Li T, Wang D (2016) Analytical study on multi-tier 5g heterogeneous small cell networks: Coverage performance and energy efficiency. Sensors 16(11), DOI 10.3390/s16111854, URL <http://www.mdpi.com/1424-8220/16/11/1854>
31. Ye Q, Rong B, Chen Y, Al-Shalash M, Caramanis C, Andrews JG (2013) User association for load balancing in heterogeneous cellular networks. IEEE Transactions on Wireless Communications 12(6):2706–2716, DOI 10.1109/TWC.2013.040413.120676
32. Zhao N, Yu FR, Leung VCM (2015) Opportunistic communications in interference alignment networks with wireless power transfer. IEEE Wireless Communications 22(1):88–95, DOI 10.1109/MWC.2015.7054723
33. Zhao N, Yu FR, Sun H, Li M (2016) Adaptive power allocation schemes for spectrum sharing in interference-alignment-based cognitive radio networks. IEEE Transactions on Vehicular Technology 65(5):3700–3714, DOI 10.1109/TVT.2015.2440428
34. Zhao N, Cao Y, Yu FR, Chen Y, Jin M, Leung VCM (2018) Artificial noise assisted secure interference networks with wireless power transfer. IEEE Transactions on Vehicular Technology 67(2):1087–1098, DOI 10.1109/TVT.2017.2700475