Particle Swarm-Based Cell Range Expansion for Heterogeneous Mobile Networks

HUGO PEREIRA KURIBAYASHI[1,2], **MARCELA ALVES DE SOUZA**[1,2],
DIEGO DE AZEVEDO GOMES[1], **KETYLLEN DA COSTA SILVA**[2],
MARCELINO SILVA DA SILVA[2],
JOÃO C. WEYL ALBUQUERQUE COSTA[2], (Senior Member, IEEE), AND
CARLOS RENATO LISBÔA FRANCÊS[2]

[1]Federal University of Southern and Southeastern Pará (Unifesspa), Marabá 68505400, Brazil
[2]Graduate Program in Electrical Engineering, Federal University of Pará (UFPA), Belém 66075110, Brazil

Corresponding author: Hugo Pereira Kuribayashi (hugo@unifesspa.edu.br)

ABSTRACT Traffic demand in wireless communication systems has emerged as a key issue over recent decades. An ever-increasing trend is projected for the next few years, with explosive data traffic expected to materialize and mobile users imposing new quality of service requirements. This growing traffic demand, combined with increasingly complex heterogeneous network (HetNet) scenarios, has presented ever more challenges for mobile network operators in terms of service, coverage, load balancing, and quality of service. Considering the traditional association mechanism based on maximum power received, HetNets tend to remain unbalanced, making it challenging to satisfy mobile users' traffic requirements. In this paper, instead of trying to maximize the achievable downlink rate per user, we couple the cell range expansion (CRE) technique with a particle swarm optimization (PSO) algorithm to maximize the number of users whose downlink requirements are met. The proposed scheme considers both the loads of base stations and the signal-to-interference-plus-noise ratio (SINR) of user equipment to model an objective function that seeks to compute specific CRE bias values per small base station. The proposed scheme is also compared with some classical PSO implementations. Numerical results validate the performance of the proposed schemes, which effectively fulfill users' data traffic requirements by reducing network imbalance.

INDEX TERMS Cell range expansion, heterogeneous mobile networks, load balancing, particle swarm optimization, user association.

I. INTRODUCTION

The proliferation of multimedia devices and the advent of the Internet of Things have intensified demand for high-speed data connection services. According to Cisco's 2017-2022 Visual Networking Index Report [1], global mobile data traffic is expected to increase sevenfold between 2017 and 2022, reaching 77.5 exabytes per month by 2022.

This growing demand for high volumes of data carried over mobile networks has been driven mainly by augmented and virtual reality applications, high-definition video platforms and tactile Internet services [2]. Thus, this trend is pushing mobile network operators (MNOs) to deploy and

The associate editor coordinating the review of this manuscript and approving it for publication was Qilian Liang.

manage broadband services with more rigorous quality of service (QoS) requirements. It is widely accepted that the current mobile broadband network infrastructure may satisfactorily meet the needs of some of these applications. However, certain services will impose additional requirements for the next generations of mobile networks on MNOs. In this context, there has been a consolidation of the ultra-dense networks (UDN) concept with the adoption of low-powered and short-range small base stations (SBSs) [3]. By deploying additional base stations, operators can extend network coverage close to end-user devices, thereby providing improvements related to system spectral efficiency, performance, and serviceability.

To enhance traffic volume, the 3rd Generation Partnership Project (3GPP) proposes multitier heterogeneous networks

(HetNet), in which SBSs are organized under macro base stations (MBS) to fulfill QoS and traffic volume requirements. The HetNet concept overlays low-power and low-cost base stations with conventional macro cellular networks. By deploying such low-power SBSs to bridge indoor coverage gaps or cell edges, coverage extension can be achieved in a cost-effective way [4]. That said, network densification with different types of BSs tends to cause considerable network imbalance problems, since MBSs typically have much higher transmission power than do SBSs [5]. Given the conventional max-power association mechanism, user equipment (UE) tends to associate with the cell with the highest received downlink power. Consequently, the distribution of users and services over the network becomes imbalanced [6], [7], with user associations mainly concentrated on the MBSs.

There are several definitions of 'load' in mobile cellular networks, which tend to be related to the number of users attached to a BS. If the maximum number of UEs supported by a BS varies and the total number of radio resource blocks used by UEs is constant, then the traffic flow experienced by the BS is directly proportional to the attached number of UEs [8]. Therefore, the load balancing problem becomes one of the central challenges facing heterogeneous mobile networks. The association mechanism between UEs and BSs must achieve the best possible use of network resources to provide a better quality of service to end-users.

The 3GPP standardized technique known as cell range expansion (CRE) is a promising alternative for achieving a better user balance in a mobile network [5], [9]. Through CRE, a bias factor is assigned to each SBS to virtually extend or shrink the coverage area (i.e., the boundary of SBS in Fig. 1), thus making UEs more suitable to connect to an SBS. Hence, it is possible to better exploit the use of radio resources given a higher distribution of users over the mobile network. By considering CRE-related implementation challenges, it is possible to highlight the proper calculation of bias values per HetNet layer [10]–[13] or the computation of a specific bias value for each SBS [12], [14]. Such approaches have been implemented mainly through methods that require the solving of combinatorial optimization problems (e.g., [15], [16] and the references therein), based on maximizing the sum of users' data link rates. In doing so, these approaches do not take into account users' specific requirements in the assessment of CRE adoption. Maximizing system serviceability can be a good strategy. However, trying to meet users' traffic requirements by adopting application-aware approaches could represent a more effective solution.

Moreover, growing mobile traffic demand is leading to even more complexity in network operation processes. Although there are several methods proposed in the literature to improve network performance, how to automatically handle the complexity of mobile networks through evolutionary techniques and algorithms has become an emerging research topic [17]–[21]. Evolutionary approaches and other artificial intelligence- (AI-) related techniques may include
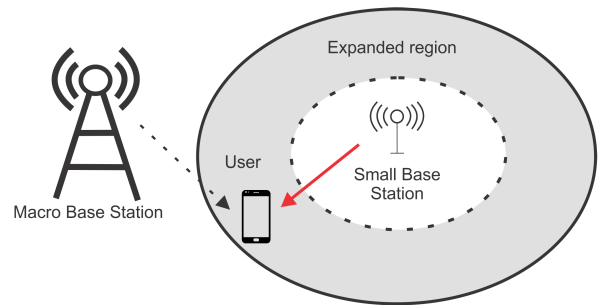

FIGURE 1. Cell Range Expansion of SBSs in downlink HetNets.

multidisciplinary methods from machine learning, bioinspired algorithms, and fuzzy neural networks and have been applied to optimize computer systems in diverse and complex scenarios [20]. The social and collective behavior of species can be used to manage complex systems through bioinspired algorithms. This approach can provide some improvements in designing, maintaining and optimizing self-organized networks (SONs) [22]. These techniques have relatively low complexity, enabled by recursive feedback-based learning and local interactions [20]. Therefore, the integration of bioinspired techniques with network operation techniques has become a promising field to be applied in improving load balance and user association processes in a HetNet.

Furthermore, these techniques may be combined with software-defined networking (SDN) and network function virtualization (NFV) technologies. They are considered two of the most promising enabling technologies to provide better management of network resources [23]–[25]. For instance, by the use of SDN virtual slices with these AI-based schemes, MNOs and mobile virtual network operators (MVNOs) could orchestrate the network components to adjust their user association mechanisms, following the service level agreements (SLAs) defined with end-users.

In addition, no significant attention has been given to the use of user data traffic requirements to define specific bias values per BS. Balancing the load per layer can cause some BSs to be overloaded or lightly loaded. Thus, it is essential to balance the load per BS and aggregate computational-time network resource optimization, which is compatible with real-time mobile network operation without the adoption of additional signaling mechanisms, to better meet traffic and quality of experience requirements for mobile devices.

Hence, we propose the use of a bioinspired algorithm to implement CRE in dense HetNet environments. Through the use of the particle swarm optimization algorithm (PSO), it is possible to perform a search for specific bias values per SBS in a HetNet. This exploration process takes into account the fulfillment of users' traffic requirements to influence the movement of particles in the search space. We present analytical models and parameters utilized to represent the particle's position and velocity in the PSO implementations contemplated. Therefore, the main contributions of this paper are summarized as follows:

- We develop a cell range expansion approach combined with the particle swarm optimization algorithm to consider specific traffic requirements of mobile users and better network balancing control. The adoption of PSO as an optimization tool seeks the unified calculation of all SBSs' bias values.
- We formulate a user association problem that seeks to maximize the number of UEs with fulfilled downlink requirements, unlike related literature that attempts to maximize only the aggregate throughput rate obtained by mobile users.
- We comprehensively analyze the objective function performance under different bias values and investigate the combination of parameters and its influence on the user association problem and network load balancing.
- Through extensive numerical results on a large multi-macro cell HetNet obtained by developing a detailed simulator, we evaluate the performance of our scheme by comparing it to the conventional unified biasing scheme (i.e., a single bias value per tier). The results show that our proposed scheme can be a promising application for ultradense HetNets to achieve satisfactory levels of user service and load balancing in the mobile network.

The remainder of this paper is organized as follows. Related works are discussed in section II. The system model is presented in section III to describe the scenario implementation, decision variables used, and a brief discussion of bioinspired computing techniques to highlight some key features of the particle swarm optimization algorithm. In section IV, the analytical formulation of the problem is described. Section V presents the experiments performed, the parameters used in this work, and a discussion of the numerical results obtained. Finally, section VI presents the final considerations of this paper and discusses future works.

II. RELATED WORK

This section provides a thorough overview of the work related to user association mechanisms and the use of bioinspired approaches in heterogeneous mobile networks.

By considering the increasingly dynamic, heterogeneous, large-scale and complex scenarios associated with the HetNet research field, several works have surveyed the adoption of computational intelligence techniques within the processes of operation and optimization of a HetNet [20]–[22]. In [20], there is a discussion of the application of AI-based techniques for evolving smarter HetNet infrastructure and systems, focusing on the research issues of self-configuration, self-healing, and self-optimization, presenting a discussion of the pros and cons of using each of the related techniques. Additionally, the study cites some PSO-based schemes for the proposition of low-power user association in HetNets and cell outage compensation management. [20] suggests that future research and development of HetNet technologies must address new requirements of M2M (machine-to-machine) and IoT (Internet of Things) services by adopting effective management and optimization techniques inspired by bioecosystems, e.g., GA (genetic algorithms) and ACO (ant colony optimization) of swarm intelligence.

The application of swarm intelligence in communications networks is comprehensively discussed in [21]. Different aspects of bioinspired mechanisms and various algorithms that have been proposed in the literature to improve the performance of artificial systems are surveyed in the paper. It emphasizes the main aspects of swarm intelligence observed in social species to explore intelligent features (e.g., flexibility, robustness, decentralized control, self-evolution, etc.), which can be applied to communication networks. Additionally, some fundamental SON mechanisms and design principles are discussed for wireless communication systems, highlighting the applications of such methods to the maintenance, operation and optimization processes of a network.

In [22], particular attention is given to the application of bioinspired mechanisms to examine the use of various algorithms in artificial SON systems. Some open research issues are shown, including SON design trade-offs, self-X capabilities in LTE-Advanced systems, cognitive machine-to-machine (M2M) self-optimization, cross-layer design, resource scheduling, and power control. It also presents some comparisons of critical SON issues from the perspective of physical-layer and media access control- (MAC-) layer operations. The authors suggest that swarm intelligence can represent a promising solution for managing SONs. However, proposing bioinspired methods that are universally applicable to diverse network environments is still a challenging task.

The authors of [26] propose a PSO-based algorithm (P5G) to explore the optimization of different key performance indicators (KPIs) by allocating virtual elements called reusable functional blocks (RFBs) in the context of software-defined networks (SDNs). PG5 attempts to manage RFBs to deliver high-definition videos to end-users through a typical scenario composed of MBSs, SBSs, and EPC (evolved packet core) nodes. Despite the fact that it does not consider user-specific traffic requirements, the results show that P5G performs close to the optimal solution, and the computation time of P5G is regularly low.

PSO is applied in [27] to dynamically obtain the per-BS biasing values to maximize the achievable throughput. The system model considers a downlink three-tier HetNet, with different path loss models for each type of BS. The PSO implementation uses a decreasing inertia factor to balance the exploration and exploitation phases. Additionally, the proposed approach can provide fairness among users by controlling the load per BS. Nevertheless, user-specific traffic requirements are not considered to influence the movement of particles in the search space.

The authors in [28] analyze the user association problem by applying an algorithm that represents a Bayesian BS selection game to consider the characteristics of the SBSs and the type of user traffic requirements, seeking to increase the chances of an appropriate association that will reduce the end-to-end latency of users. The probability of proper association,

the achieved latency concerning conventional CRE, and the max SINR-based cell selection/user association algorithms used in LTE-Advanced are used to validate the proposed approach. Additionally, considering that this paper works with a low-density scenario of base stations, the technical feasibility of the proposal can be inconclusive due to the use of combinatorial optimization techniques.

In the study presented in [12], stochastic geometry techniques are used to analyze the user association problem to consider optimal bias values that tend to maximize the data rates obtained by mobile users. Additionally, the study proposes bias factors in closed form to explicitly evaluate the impact of different network settings on the virtual coverage areas. Furthermore, the work does not consider specific user traffic characteristics to perform a more accurate analysis of the user association mechanism.

[29] proposes an approach for joint cell selection and resource allocation. This approach aims to reduce the load of MBSs, releasing as many resource blocks (RBs) as possible, pushing the users to SBSs. The results obtained through simulation are promising and show that the proposed scheme may be better than the CRE scheme. One of the main disadvantages observed is that the approach assumes a set of nonstandardized signals, which may represent excessive overhead, making the proposal unsuitable for ultra-dense network scenarios.

The authors of [30] present a load balancing and association approach, based on the knapsack optimization (KO) algorithm, to attempt the distribution of UEs across SBS layers. There is a set of restrictions related to BSs' service capacity and the amount of RBs needed by UEs to satisfy their QoS requirements. However, as the number of users and BSs increases (e.g., in ultradense networks), the solution may not have the scalability and convergence time required for multitier HetNets since KO problems can be considered NP-hard.

In [31], the authors propose the first rule for small cells (SCF). The goal is to force UEs to associate with the nearest SBSs. The receiving SINR should be higher than a predetermined threshold limit. Otherwise, UEs tend to associate with the MBS. By adjusting the threshold parameter, the authors show that the rule can provide much better performance than the best-SINR association mechanism (without the use of CRE).

In addition, most of the aforementioned works are limited to using Shannon's theorem (as opposed to the standard 3GPP-LTE discrete modulation-and-coding scheme (MCS) function), which may unrealistically overestimate network capacity. The use of a discrete MCS function will significantly complicate the selection of the right association for cell-edge users since it does not have the convexity and strictly increasing properties of Shannon's theorem [32].

III. SYSTEM MODEL
In this section, we describe a system model and formulate an optimization problem for the computation of specific CRE

bias per SBS to consider users' traffic requirements. We consider a downlink HetNet, which comprises K independent network tiers of BSs with $K = 1, 2, \ldots, k$ and a typical UE at $\Re^3$. User and BS locations are sampled through independent homogeneous Poisson point process distributions (HPPP). Generally, the kth layer has density λ_k, and its BSs are randomly generated from HPPP $\phi(\lambda_k)$, while the user positioning is generated by HPPP $\phi(\lambda_u)$.

By considering a two-layer model, for instance, we have $K = 2$, and we can consider *tier*-1 as the representation of BSs that have the highest power (MBSs) and low presence density in a given topology, while *tier*-2 may represent randomly deployed base stations, which have low power (SBSs), high density, and a distribution within the proposed scenario. Since the power transmission of BSs in tier-k is represented by P_k, we can assume that $P_1 \gg P_2$. Additionally, to contemplate a UDN scenario, we consider that $\lambda_1 \ll \lambda_2$.

The set of all BSs is denoted by φ, where $\varphi = (\delta \cup \gamma)$. The set of MBSs is represented by $\delta = \{M_1, M_2, M_3, \ldots, M_m\}$, and the set of SBSs is denoted by $\gamma = \{S_1, S_2, S_3, \ldots, S_s\}$, where φ is indexed by $1 \leq j \leq b$ (i.e., $b = m + s$). The set of UEs is denoted by $\pi = \{U_1, U_2, U_3, \ldots, U_u\}$, with $1 \leq i \leq u$, and $\psi = \{\theta_1, \theta_2, \theta_3, \ldots, \theta_s\}$ is the set of bias values for SBSs. Moreover, the ith user requests a class service defined as the tuple $\rho_i = (\eta_i, \tau_i)$, where η_i and τ_i are the average flow throughput (Mbps) and compression factor (ratio of processed-to-raw data), respectively. Hence, the ith UE's required data rate can be expressed by the product $(\eta_i \cdot \tau_i)$.

A. CELL RANGE EXPANSION FOR MAX-SINR
There are several cell association algorithms based on metrics such as RSRQ (Reference Signal Received Quality), RSRP (Reference Signal Received Power) or SINR [33], [34]. RSRP and RSRQ are the ones with the lowest additional communication complexity, since these parameters have already been specified in LTE [35]. In [36], the user association mechanism is evaluated considering these metrics, and it is shown that SINR-based selection can achieve a better downlink rate.

Hence, in this model, we assume that received SINR is a key indicator of the UE rate and outage performance due to its direct relationship with Shannon's theorem [12]. With the Max-SINR association criteria, the ith UE tends to associate with jth BS, such that $j = arg \max(SINR_{ij}), \forall j \in \varphi$. Considering a two-tier HetNet ($K = 2$) and the fact that MBSs typically have much higher transmission power ($P_1 \gg P_2$), the UEs tend to be associated mainly with MBSs. By adding a CRE bias to the SINR of each SBS, the UEs tend to be better distributed between BSs, and each UE can possibly achieve a better long-term rate. When the ith UE tends to associate with the MBS tier (*tier*-1), selecting an MBS $k \in \delta$, the received SINR (ζ) satisfies (1) and (2):

$$\zeta_{ik} > (\zeta_{ij}), \quad \forall j \in \delta, \ k \neq j. \tag{1}$$

$$\zeta_{ik} > (\zeta_{ij} + \theta_j), \quad \forall j \in \gamma, \tag{2}$$

where θ_j represents the CRE bias value for the jth SBS. Additionally, by the adoption of CRE, the ith UE selects the ℓth SBS when the received SINR satisfies (3) and (4):

$$(\zeta_{i\ell} + \theta_\ell) > \zeta_{ij}, \quad \forall j \in \delta. \tag{3}$$

$$(\zeta_{i\ell} + \theta_\ell) > (\zeta_{ij} + \theta_j), \quad \forall j \in \gamma, j \neq \ell \tag{4}$$

By setting proper CRE bias values for the SBSs as described above, the SBSs expand or shrink their downlink coverage, thereby causing more or fewer users to be associated with the SBSs. When the ith UE is associated with the jth BS, the downlink SINR (ζ_{ij}) can be expressed as:

$$\zeta_{ij} = \frac{P_j h_{ij}}{\sum_{k \in \varphi, k \neq j} P_k h_{ik} + P_N}, \tag{5}$$

where P_j is the transmission power of the jth BS, h_{ij} is the effective gain channel between the ith UE and the jth BS and P_N represents the thermal power noise. Hence, the achievable per-channel downlink rate at the ith UE from the jth BS can be expressed as [32]:

$$R_i = e_\ell \cdot \frac{n_{sc} \cdot n_{sym}}{T_{subframe}}, \tag{6}$$

where e_ℓ represents the per-subcarrier efficiency in terms of bits per OFDM symbol for a given threshold SINR. Therefore, e_ℓ is obtained through an MCS function, denoted by $\mu(\zeta_{ij})$. The terms n_{sc}, n_{sym} and $T_{subframe}$ are the number of subcarriers per channel, number of OFDM symbols per subframe and duration of a subframe, respectively. By considering a fair resource allocation scheme, in which the total number of resource blocks (RB) is equally divided between the associated users, the total number of RBs obtained by the ith UE from the jth can be expressed as:

$$n_i^{RB} = \left\lfloor \frac{n_j^{RB}}{L_j} \right\rfloor, \tag{7}$$

where n_j^{RB} represents the total number of resource blocks available in BS j, while L_j represents the load of BS j, i.e., the total number of UEs associated with BS j. Moreover, the number of resource blocks received by a single UE should be higher than a minimum threshold defined by T_B, which seeks to emulate the transport block concept. The notation ($\lfloor \cdot \rfloor$) represents the floor function that gives as its output the greatest integer less than or equal to $\frac{n_j^{RB}}{L_j}$ to guarantee that the number of resource blocks (n_i^{RB}) is an integer value. Finally, Table 1 summarizes the acronyms used.

B. DECISION VARIABLES
X represents a binary matrix that expresses the association between the ith UE and the jth BS. Hence, the x_{ij} element can be expressed by Eq. (8):

$$x_{ij} = \begin{cases} 1 & \text{if the } i\text{th UE is associated with the } j\text{th BS;} \\ 0 & \text{otherwise.} \end{cases} \tag{8}$$

In addition, Y denotes an array of binary values that represents the fulfillment of the UE's downlink requirements,

TABLE 1. Brief description of acronyms.

Notation	Description
φ	Set of all BSs
δ	Set of MBSs
γ	Set of SBSs
π	Set of user equipment
ψ	Set of SBS CRE bias
λ_k	Density of BSs
λ_u	Density of UEs
$\phi(\varphi_k)$	HPPP to constitute the k-tier
$\phi(\varphi_u)$	HPPP to constitute the UEs' positions
ρ_i	Required class service of the ith UE
η_i	ith user's average flow throughput
τ_i	Ratio of processed-to-raw-data for the ith UE
ζ_{ij}	Downlink SINR at ith UE from jth BS
P_j	Transmission power of a BS j
h_{ij}	Effective gain channel between ith UE and jth BS
P_N	Thermal power noise
R_i	Per-channel downlink rate at ith UE from jth BS
e_ℓ	Per-subcarrier efficiency
n_{sc}	Number of subcarriers per channel
n_{sym}	Number of OFDM symbols per subframe
$T_{subframe}$	Subframe duration
$\mu(.)$	Discrete MCS function
n_j^{RB}	Total of RBs available at the jth BS
T_B	RB threshold parameter
L_j	Load of the jth BS

i.e., the element $y_i = 1$, if the ith UE has its download requirements met, according to Eq. (9):

$$y_i = \begin{cases} 1 & \text{if the } i\text{th UE's DL requirement is met;} \\ 0 & \text{otherwise.} \end{cases} \tag{9}$$

Moreover, Z represents an array of binary values that the z_j element denotes when the jth BS is serving at least one UE, as described by Eq. (10):

$$z_j = \begin{cases} 1 & \text{if the BS } j \text{ is serving at least one UE;} \\ 0 & \text{otherwise.} \end{cases} \tag{10}$$

Finally, S_π represents the sum of the UEs' downlink rate, as an indication of the aggregate total flow obtained by users, as described by Eq. (11):

$$S_\pi = \sum R_i, \quad \forall i \in \pi. \tag{11}$$

C. PARTICLE SWARM OPTIMIZATION ALGORITHM
The field of bioinspired computing attempts to replicate how biological organisms and suborganisms operate using ideas from abstract computing of biological systems [37]. In general, bioinspired computing seeks to optimize a problem to iteratively examine the improvement of a candidate solution over a given measure of quality. The particle swarm optimization (PSO) algorithm, which is inspired by animal group behavior, is one of these techniques. In PSO, the population is called a swarm, and each individual in the swarm is known as a particle [38].

The purpose of PSO is to perform a biased stochastic exploration of the global optimum solution through the search space of a problem. A group of particles is used to exploit the problem represented in a multidimensional space. At the

beginning of execution, the particles are randomly scattered in the search space, so each particle represents a candidate solution with an aptitude value that serves to rank its quality and select the best solutions. The particles must move according to equations (12) and (13) until the end of the processing conditions [39].

$$V_i(t+1) = V_i(t) + c_1r_1(P_i - X_i) + c_2r_2(G - X_i) \quad (12)$$
$$X_i(t+1) = X_i(t) + V_{i+1}(t+1), \quad (13)$$

where V_i and X_i represent the velocity vector and location of the ith particle, respectively. The term c_1 adjusts the cognitive particle behavior, and c_2 controls the social particle behavior. In addition, r_1 and r_2 are random factors belonging to interval [0, 1]. Finally, P_i represents the previous best position of the ith particle, while G denotes the current global best position of the swarm.

Since the first release of the PSO algorithm, some significant improvements have been made by other research efforts, and the algorithm has been successfully applied to many problems, such as function optimization and others [40], [41]. Due to the ease of implementation and the fast convergence to acceptable solutions, PSO has received much more attention in recent years [40] and has been implicated in numerous PSO variants based on different velocity/position updating rules [42], different parameter values [43], the use of dynamic and adaptive parameters [44], [45] and population sizing [46].

For instance, [47] introduces into the original PSO a parameter referred to as inertia weight (IW) (ω), which moderates the current particle position, changing the definition of Eq. (12), to that presented in Eq. (14). [48] proposes a decreasing-IW PSO, where the value of the inertia parameter decreases along with the algorithm's interaction number. The authors of [43] propose the adoption of a constriction factor (χ), which redefines the velocity update process according to (15) and (16) to create better cohesion and balance between exploration and exploitation within the process search space.

$$V_i(t+1) = \omega V_i(t) + c_1r_1(P_i - X_i) + c_2r_2(G - X_i) \quad (14)$$
$$V_i(t+1) = \chi(V_i(t) + c_1r_1(P_i - X_i) + c_2r_2(G - X_i)) \quad (15)$$
$$\chi = \frac{2}{|2 - c - \sqrt{c^2 - 4c}|} \quad (16)$$

where c is defined by the following expression $c = c_1 + c_2$:

IV. PROBLEM FORMULATION
In this section, considering the performance criteria and decision variables adopted in this paper, we formulate an optimization problem modeled by the objective function as follows:

$$\text{Maximize} \quad \alpha \cdot \sum_{i \in \pi} y_i + \beta \cdot \sum_{j \in \varphi} z_i \quad (17)$$

The goal of (17) is to maximize the fulfillment of UEs by maximizing the number of users whose downlink requirements are met and the number of BSs that have users connected. When trying to maximize the parameter (z_i), it is

expected that the number of BSs associated with users and, hence, the number of RBs used by these UEs increase. The parameters α and β balance the contributions of the objective function components. Additionally, the goal of maximization is based on the choice of set values $\psi = \{\theta_1, \theta_2, \theta_3, \ldots, \theta_s\}$, which directly influences the values obtained from y_i and z_i. In addition, the objective function is subject to the following restrictions:

$$\sum_{j \in \varphi} x_{ij} = 1, \quad \forall i \in \pi, \quad (18)$$
$$\sum_{i \in \pi} n_i^{RB} \leq n_j^{RB}, \quad \forall j \in \varphi \quad (19)$$
$$n_i^{RB} \geq T_B, \quad \forall i \in \pi \quad (20)$$

Constraint (18) guarantees that each UE is associated with at most one BS, i.e., it is not considered a coordinated multipoint transmission scenario (CoMP). Constraint (19) ensures that the number of resource blocks used by the ith UE is less than or equal to the total number of RBs available at the jth BS. Finally, the constraint (20) guarantees the feasibility of the solution by assuring that the number of RBs received by a single UE should be higher than a minimum threshold T_B.

V. SIMULATION MODELS AND EXPERIMENTS
In this section, we present the simulation models and experiments used for performance evaluations. Then, we describe the numerical results obtained.

A. SIMULATION MODELS
We use the distance-based path loss model and simulation parameters recommended by 3GPP [49]. Observations on how performance changes are based on Monte Carlo simulation in an attempt to emulate the long-term behavior of the proposed scenario. In these simulations, we consider a two-tier HetNet ($K = 2$) with $\lambda_2 = 10\lambda_1$ and $\lambda_u = 150\lambda_1$ to reflect a UDN scenario.

We assume that MBSs (deployed outdoors, at rooftop levels) are capable of covering a vast area and supporting a very high number of UEs. Each MBS site uses fiber or microwave for backhaul, and the sites are deployed by the MNO using engineered deployment strategies. On the other hand, we consider SBSs as randomly deployed by the end-users (similarly to Wi-Fi systems). Hence, the SBS-based deployment does not follow any engineered strategy, which favors the adoption of HPPP distribution. These SBS systems are usually backhauled via the users' existing broadband infrastructure, i.e., digital subscriber line (DSL), cable modem, ethernet, or fiber [50].

Moreover, the transmission power is 46.0 dBm for an MBS and 23.0 dBm for an SBS (i.e., we consider SBSs to be femto systems). We run 1, 000 simulations for each scenario. We generate the locations of BSs and UEs using their respective HPPP densities over an area of 1 km^2. We then calculate the SINR of the possible links and connect UEs to their associated BSs using the Max-SINR association policy. Through

TABLE 2. Physical-layer parameters [49].

Parameter	Value
MBS power transmission	46.0 dBm
SBS power transmission	23.0 dBm
Carrier frequency	2.0 GHz
Environment	Urban
$T_{subframe}$	1 ms
n_{sc}	12
n_{sym}	14
MBS antenna gain	15.0 dBi
Noise power	-174.0 dBm/Hz
Subchannel BW	180 kHz
UE antenna gain	0.0 dBi
SBS antenna gain	5.0 dBi
MBS-UE path loss	$128.0 + 37.6 \log_{10}(\max(d, 35)/1000)$
SBS-UE path loss	$140.7 + 36.7 \log_{10}(\max(d, 10)/1000)$

TABLE 3. Simulation parameters.

Parameter	Value	Ref.
Area	$1.0 \ km^2$	-
Number of simulations	1,000	-
Number of tiers (K)	2	-
MBS density (λ_1)	$2.0/km^2$	-
SBS density (λ_2)	$20.0/km^2$	-
UE density (λ_u)	$300.0/km^2$	-
Minimum threshold (T_B)	5	-
Height of antenna for MBS/SBS/UE	$30.0/10.0/1.5 \ m$	[34]
SCF threshold	15.6 dB	[32]

the SINR, it is possible to compute the UEs' downlink rate using Eq. (6), based on the number of RBs allocated per UE. Finally, the major simulation parameters are summarized in Tables 2 and 3.

Furthermore, we use the 15-rate MCS available in LTE, as shown in Table 4, to parametrize the rate function $\mu(\zeta_{ij})$, in accordance with Eqs. (5) and (6). In addition, the class service requested by each UE, defined as the tuple $\rho_i = (\eta_i, \tau_i)$, is sampled from a random uniform integer distribution parametrized by the integer interval [1, 8], following values shown by Table 5. For all simulations, an alternative scenario with CRE bias absence is considered to better evaluate the application of CRE and the degree of network imbalance based on the metrics and decision variables considered in the objective function of Eq. (17). Therefore, the impact of the user association policy based on Max-SINR can be assessed to perform a quantitative analysis based on the number of UEs meeting DL requirements (y_i) and the quantity of BSs that serve users (z_i). Moreover, a sensitivity analysis approach is used to determine the values for α and β.

B. EXPERIMENTS
We consider multiple experiments to assess the adoption of PSO as an efficient approach to implementing the CRE technique and evaluate improvements in the user association mechanism.

1) EXPERIMENT I: UNIFIED CRE BIAS
In this experiment, all SBSs adopt a unified CRE bias value [52]. Thus, the main objective of this experiment is to evaluate the adoption of ordinary CRE bias values, without any type of optimization technique, to observe its impact on the network balance and consequently, the UEs' achievable rates, in compliance with the objective function of Eq. (17).

The CRE bias values used are set in the discrete range [−10.0, 85.0] dB, with intervals/steps of 5.0 dB. Through the experiment, we expect to evaluate the network imbalance level, which should be lower than that observed in the scenario without the adoption of CRE bias, implying higher downlink rates for the UEs and a better distribution of load among the BSs. Additionally, the proposed experiment can assess the most suitable CRE bias values for a unified approach in the proposed HetNet model. In the remainder of this work, the unified cell range expansion bias approach is denoted by UCB for convenience reasons.

2) EXPERIMENT II: PSO-BASED CRE BIAS
In this experiment, the PSO algorithm is proposed to compute specific CRE bias values for each SBS. The proposed implementation is denoted as PCB. Thus, unlike the UCB approach, by the use of PCB, we seek to adopt specific CRE bias values per SBS, which represents the definition of the elements of the set $\psi = \{\theta_1, \theta_2, \theta_3, \ldots, \theta_s\}$. Additionally, the PCB implementation uses a decreasing inertia weight (Decrease-IW PSO) with a variable population behavior.

Moreover, each particle is a candidate solution to the optimization problem in a d-dimensional space, where $d = | \gamma |$. Thus, the position of the ith particle is denoted by a vector $x_i = [x_{i1}, x_{i2}, \ldots, x_{id}]$, where $i = 1, 2, \ldots n$, and n represents the number of particles in the swarm. The PCB fitness function evaluates the swarm particles based on the objective function defined by Eq. (17). Hence, at the end of the processing conditions, the best overall position reached (G) represents the proposed solution for the CRE bias set ψ.

In addition, most metaheuristics consider the trade-off between exploration and exploitation features. Both features are critical and should be explored appropriately during the algorithm execution time. In this context, [53]–[55] discuss the contradiction of maintaining high diversity and obtaining fast convergence as simultaneous goals of any metaheuristic. Additionally, the problem with premature convergence often persists in multimodal optimization, which results in significant performance loss and suboptimal solutions. This issue seems to be formed by the fast information flow among the swarm, resulting in critical outcomes, such as difficulties of escaping local optima, low diversity of the population, and fitness stagnation [44].

Hence, in this PSO variation, the average swarm particle evaluation at the ith step (E_i) represents evidence of convergence of the solution. Hence, when $| E_i - E_{i-1} |$ is below a certain threshold ϵ ($\epsilon \rightarrow 0$) for at least k steps, there is an indication of swarm stabilization. Thus, the PCB approach resets any particle ranked below the mean swarm evaluation to a new random position. This operation provokes a decreasing behavior in the overall swarm evaluation. However, by generating new random candidate solutions,

TABLE 4. Threshold SINR (dB) to efficiency e_ℓ (bits/symbol) [32].

SINR	-6.5	-4	-2.6	-1	1	3	6.6	10	11.4	11.8	13	13.8	15.6	16.8	17.6
e_ℓ	0.15	0.23	0.38	0.6	0.88	1.18	1.48	1.91	2.41	2.73	3.32	3.9	4.52	5.12	5.55

Algorithm 1 PCB Algorithm

---

1 Input: X, V, P, E;
   Result: ψ
2 **begin**
3   initialize $X, V, P \leftarrow X, c$;
4       $G \leftarrow getBest(P)$;
5   **repeat**
6   **for** $i \leftarrow 1$ **to** N_p **do**
7               $V_i \leftarrow getUpdate(X_i, V_i, P_i, G)$;
8               $X_i \leftarrow X_i + \omega_i V_i$;
9               $E_i \leftarrow getEvaluation(X_i)$;
10   **if** $f(X_i) > f(P_i)$ **then**
11                  $P_i \leftarrow X_i$;
12   **if** $f(X_i) > f(G)$ **then**
13                      $G \leftarrow X_i$
14   **end**
15   **end**
16              $\omega_i \leftarrow getUpdate(\omega_i)$;
17   **if** $| E_i - E_{i-1} | < \epsilon$ **then**
18                  $c \leftarrow c + 1$;
19   **end**
20   **if** $c > k$ **then**
21                  $c \leftarrow 0$;
22                  $X_i \leftarrow getUpdate(X_i, E_i)$;
23   **end**
24   **end**
25   **until** *Stop criteria not matched*;
26      $\psi \leftarrow setBias(G)$;
27 **end**

TABLE 5. Application profile [51].

#	Application profile	η_i (Mbps)	τ_i
1	AR/VR	100	0.6
2	Factory automation	1	0.8
3	Data backup	1	1
4	Smart grid	0.4	0.3
5	Smart home	0.001	1
6	Medical	2	0.2
7	Environmental monitoring	1	0.1
8	Tactile internet	200	0.8

the increase of swarm evaluation on the following iterations is presumed.

The process of PCB is presented in Algorithm 1. Lines 2-4 initialize all the variables and calculate the best position among all particles (G), respectively. From lines 5-25, the algorithm executes the main loop. When a stop criterion matches the predefined criteria, e.g., the maximum number of steps, then the main loop is finished. The inside loop (lines 6-15) updates the velocity, position, swarm evaluation, and fitness of each particle with the influence of the decreasing inertia weight (ω_i). Additionally, it checks if the new position is better than the previous one (line 10). If true, the new position turns into the previous best position of the particle found so far (line 11). The global best position is checked for possible updates (lines 12–14). Then, the inertia weight is updated (line 16). After that, the inside loop (lines 17-23) checks for indications of convergence. In the case of stabilization, the algorithm resets the particles' rank position above the threshold ϵ (line 22). The main loop finishes when the maximum number of steps N_p is reached (line 25). The result of the algorithm is the global best position G (line 26), which defines the set ψ.

According to the objectives of this paper, the PCB approach can be divided into three main stages (summarized in Fig. 2), which are as follows: first, initialization of particles and velocity; second, updating of particle position, updating of particle velocity and evaluation of solution fitness; third, assessment of the convergence of the swarm and random reboot of some particles. To obtain the time complexity of the proposed mechanism, we analyze the time complexity of the compound stages.

In the first stage, the most complex calculations are the initialization of particle position and particle velocity. The first phase generates a random population of N_p particles. Then, the algorithm performs the computation of the fitness function for each particle and the initialization of the inertia weight factor (ω). Later, this stage performs the computation of the best position of each particle (P). Finally, this phase computes G as the best fitness of all particles. Considering the number of particles (N_p), the time complexity for this computation is $\mathcal{O}(N_p \cdot log(N_p))$.

In the second stage, the main computation is to update all particle positions and velocities and to evaluate the fitness solutions at each iteration (step). The time complexity of this phase depends on the number of steps (t) and particles (N_p). Thus, the time complexity can be equal to $\mathcal{O}(t \cdot N_p)$. In the last stage, the main computation is to occasionally update the swarm population by replacing particles ranked above the global evaluation. If the global evaluation becomes stagnant ($| E_i - E_{i-1} | < \epsilon$) for at least k steps, the replacement process is triggered. In this case, we define the worst and the best running time. Let m be the number of times the particle population is partially reset. In the best case, $m = 0$, which means there was no stagnation of the global evaluation; thus, the time complexity is negligible (i.e., the operations declared in lines 21 to 22 are not executed). In the worst case, when $arg \max m$ ($m > 0$), the time complexity can be equal to $\mathcal{O}(N_p)$.
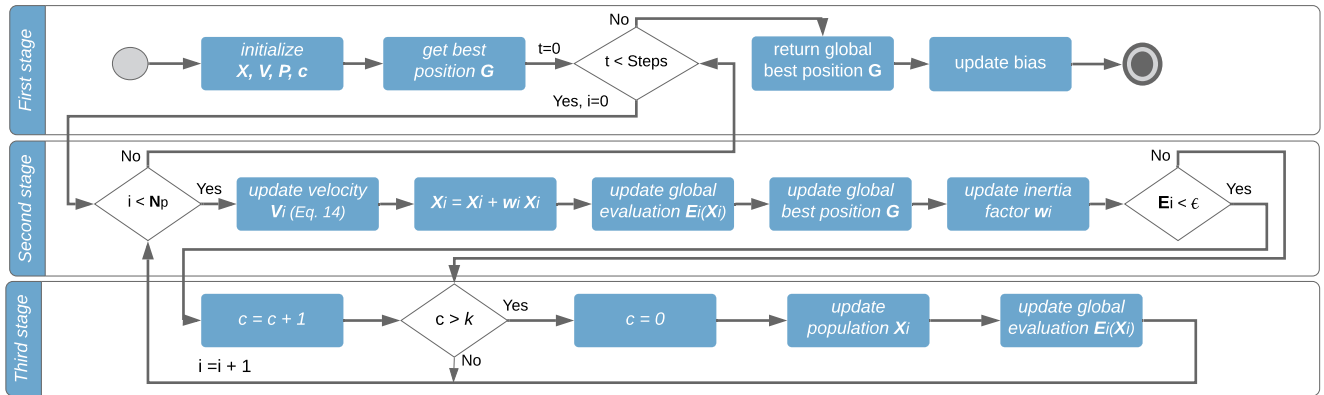
FIGURE 2. Flowchart representation of the PCB algorithm.

TABLE 6. PCB parameters.

PCB parameter	Value
Swarm size (n)	60
Maximum number of interactions (i)	100
Inertia weight interval (ω)	$\omega = 0.9 \rightarrow 0.6$
Random factor interval for r_1 and r_2	[0,1]
Cognitive particle behavior factor (c_1)	2.05
Social particle behavior factor (c_2)	2.05
Threshold factor (ϵ) and k	0.01, 5% of N_p

FIGURE 3. UCB over α and β values.

FIGURE 4. UCB over several bias values.

FIGURE 5. Ratio of UEs per SBS.

The overall time complexity of the PCB algorithm is $\mathcal{O}(N_p \cdot log(N_p) + t \cdot N_p)$. Moreover, we note that time complexity is mainly governed by the input parameters, which may be chosen following the considered scenario. Hence, Table 6 presents the parameters used in the PCB approach.

Additionally, the experiment compares the PCB approach with some additional classical PSO variants. Thus, it is possible to assess whether these implementations can browse the search space for better solutions, given the inherent characteristics of the problem and decision variables analyzed. Therefore, this experiment considers the use of CoPSO [43], Increase-IW PSO [56] and Stochastic-IW PSO [57]. However, the use of these PSO variants involves several
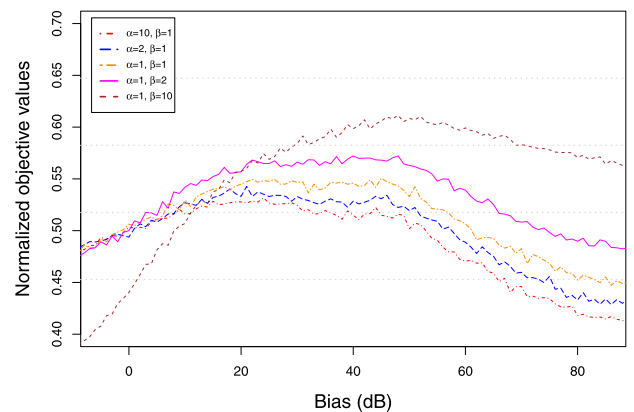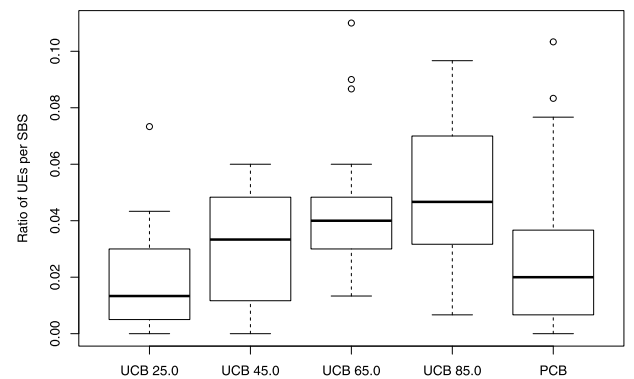
specific parameters, which may vary depending on the implementation considered. Hence, the determination of such parameters may represent a secondary optimization problem. Therefore, by considering these PSO variants in this experiment, the parameter values used in this work are presented in Table 7, following the evaluated values recommended by related literature [40].

Moreover, all PSO variations have similar mechanisms of operation. However, they differ in updating the inertia weight
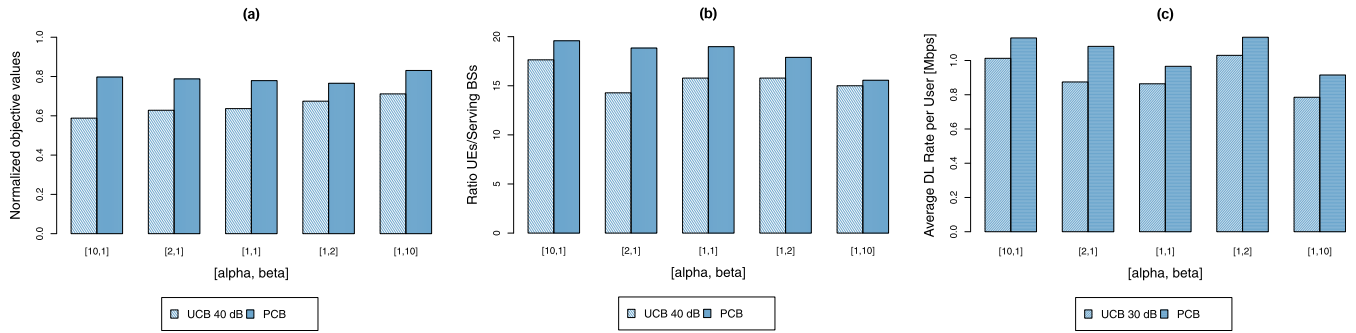
FIGURE 6. (a) Normalized numerical values over α and β values (b) Sum of the UE downlink rate (c) Average downlink rate per UE.

TABLE 7. PSO variant parameters.

PSO variant	Velocity updating parameters	Reference
CoPSO	$c_1 = c_2 = 2.05$	[43]
Increase-IW PSO	$\omega = 0.4 \rightarrow 0.9, c_1 = c_2 = 2.0$	[56]
Stochastic-IW PSO	$\omega = [0.5, 1], c_1 = c_2 = 2.05$	[57]

and constriction factor routine. By considering the algorithm complexity of classical PSO, we assume that all considered implementations have equivalent complexity, $\mathcal{O}(N_p \cdot t \cdot log(N_p))$.

C. NUMERICAL RESULTS

In this section, the simulation results are presented to provide a comparison between the experiments examined to assess the adoption of PSO as a viable approach for CRE implementation. We first evaluate the efficiency of UCB as a suitable and straightforward solution to the problem formulated in Eq. (17) by comparing it to the solutions obtained through scenarios with CRE bias absence. We then evaluate the PCB approach, comparing it with UCB and with the SCF [31] results. Next, we explore some PSO variants to try to obtain better results and observations than those obtained through the PCB approach.

Fig. 3 shows the average performance of the UCB approach over different values of CRE bias (i.e., 20.0, 40.0, and 60.0 dB). The values obtained through objective function (12) are normalized according to the influence of the α and β parameters. Thus, considering a unified approach (UCB), we expect to evaluate the most suitable CRE bias values in the proposed HetNet model. The use of CRE provides better results than those observed in scenarios with CRE bias absence. The CRE technique tends to provide better user balancing across the HetNet, performing an offloading process by shifting (*tier*-1) MBS users to (*tier*-2) SBSs. Additionally, for all combinations of α and β, the 40.0 dB range is among the best CRE bias values observed, except for [$\alpha = 1, \beta = 10$], where the 60.0 dB range generates equivalent results. Hence, it is possible to confirm that the application of CRE is a promising technique to improve load balancing in the network.

Furthermore, Fig. 4 presents the performance of the UCB approach for different values defined in the interval
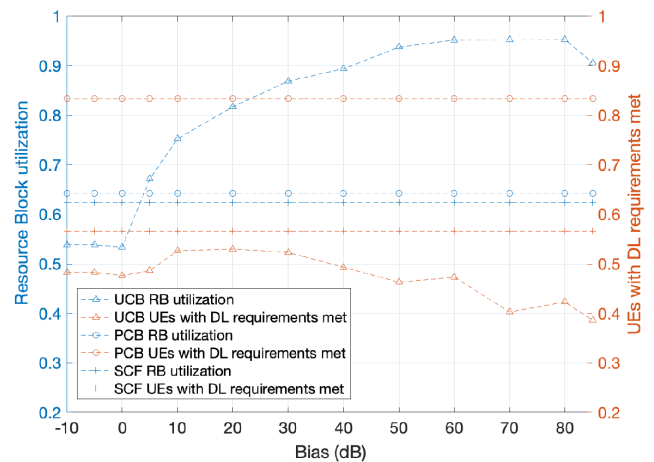

FIGURE 7. PCB and UCB over several bias values.

[$-10.0, 85.0$] dB. We observe that all curves in the subinterval [25.0, 50.0] dB have the most suitable values of the normalized objective function. Additionally, the subinterval [$-10.0, 20.0$] dB indicates low values of y_i, which corresponds to a scenario without the application of CRE. We observe that SBSs' virtual coverage area in this interval has been excessively shrunk or insignificantly extended, provoking no adjustments to the user association mechanism.

By considering the subinterval [50.0, 85.0] dB, we observe a decreasing inclination, possibly indicating an excessive offloading of users from the MBSs, i.e., the UEs are being pushed towards the SBS layer rather than a specific SBS. Hence, given the excessive increase of SBSs' load (L_j), the expected number of RBs per UE is reduced, leading to decreasing downlink rates. Moreover, when $\alpha \geq \beta$, we observe similar behavior between the curves. Otherwise, when $\alpha \ll \beta$, the variable z_i has a higher influence on the final value of Eq. (17). This peculiarity associated with the dimensionality of the variable z_i, in which the maximum equals the size of set φ, tends to explain this behavior.

Fig. 5 describes the percentage of UEs per SBS in the [$\alpha = 1, \beta = 1$] case. We observe growing behavior as the bias values increase. The ratio of UEs per SBS reaches up to 10% for UCB 85.0 dB, which may represent an overload scenario. This excessive offload helps to explain the reducing
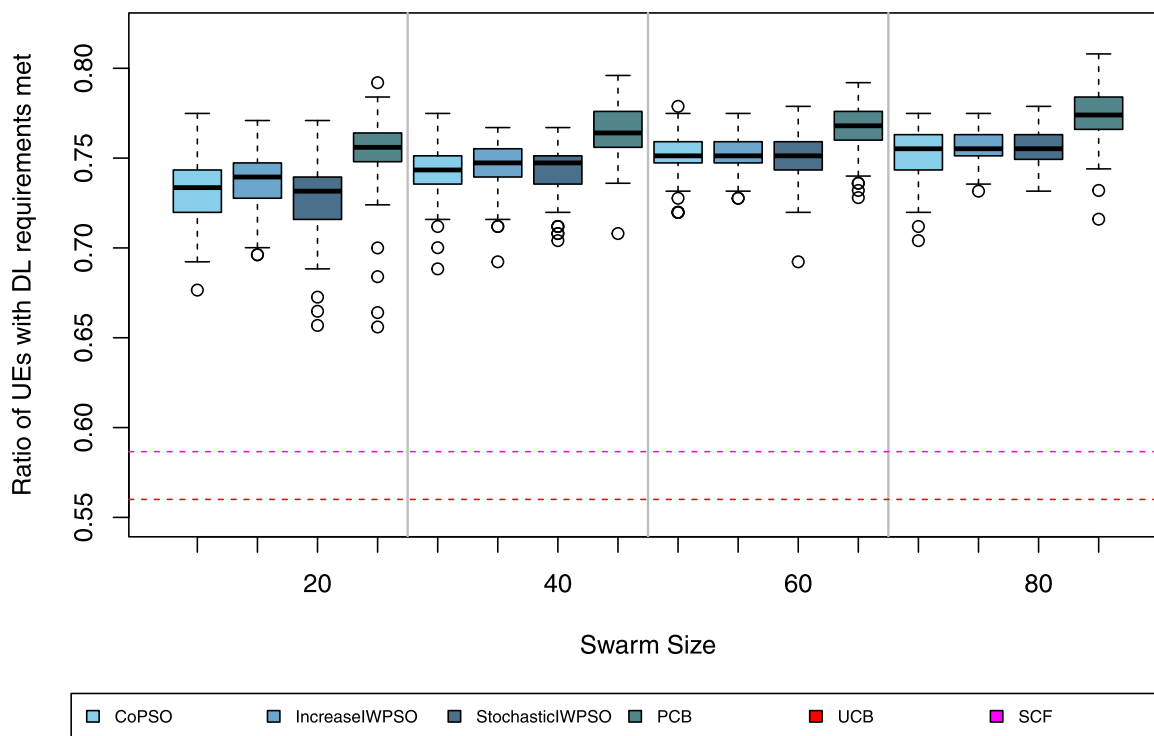
FIGURE 8. Comparison among the PSO variants.

pattern observed in the objective values presented by Fig. 4. Hence, more moderate proportions of UEs per SBS may be suitable for producing a better load balance among BSs in the proposed model.

Additionally, Fig. 5 introduces results from the PCB approach by showing a substantial fraction of SBSs serving between 2.0 and 7.5% of the UEs (the interval within the median and superior limit). Moreover, the bottom part of the PCB boxplot reports that 50% of SBSs serve approximately 2% of the UEs. Nevertheless, the UCB boxplots and the PCB boxplot have different shapes, which may suggest different effects on the user association mechanism through these approaches. Furthermore, Fig. 6(a) presents the normalized numerical values obtained by applying the UCB 40.0 dB (i.e., the best UCB configuration found) and PCB approaches. For all scenarios modeled, the PCB approach presents higher results concerning the definition of Eq. (17).

Based on Fig. 6(b), the PCB approach shows promising results through the sum of the UEs' downlink rate (i.e., Eq. 11). Additionally, the average downlink rate per UE presented by Fig. 6(c) supports this observation. Therefore, we may assume that the solutions generated by the PCB approach tend to produce even better results than those observed in the UCB approach. Every configuration of α and β tends to provide similar results, with a slight advantage when [$\alpha = 10$, $\beta = 1$], especially when considering the ratio of UEs per serving BS and the average downlink rate per UE. However, for the remaining analysis, we discuss the results by considering the case [$\alpha = 1$, $\beta = 1$] to consider an equal influence on the objective function components.

Fig. 7 presents an analysis of the UCB, SCF, and PCB approaches based on the ratio of RB utilization and the proportion of UEs with downlinks meeting the requirements. For the UCB approach, even with the increase in RB utilization, the proportion of UEs with downlink requirements has a decreasing tendency. Since MBSs have much higher transmission power compared with SBSs, we assume poor channel conditions for UEs associated with SBSs. Hence, the number of bits per OFDM symbol (e_ℓ) decreases, which implies a low level of UEs with their downlink requirements met. Additionally, the SBS layer (tier-2) may be overloaded, which indicates idle MBSs or MBSs with few associated UEs.

Considering the PCB approach, we observed the best results obtained, with promising levels of UEs with their requirements met, albeit with lower levels of RB utilization compared to the UCB approach. This result suggests that the PCB approach could exploit bias values that provide better channel conditions for user association between UEs and BSs. Furthermore, the SCF approach provided better results than UCB, but with a disadvantage from the PCB approach, when considering the ratio of UEs meeting the downlink requirements.

Fig. 8 shows a comparative analysis of the PSO variants, while the swarm size ranges between 20, 40, 60 and 80 particles. This analysis also includes the maximum values obtained by the UCB and SCF approaches and considers the ratio of UEs with downlink requirements met to compare the results. An increasing evaluation behavior is perceived as the swarm population size becomes more extended. We can presume, considering the HetNet model used in these
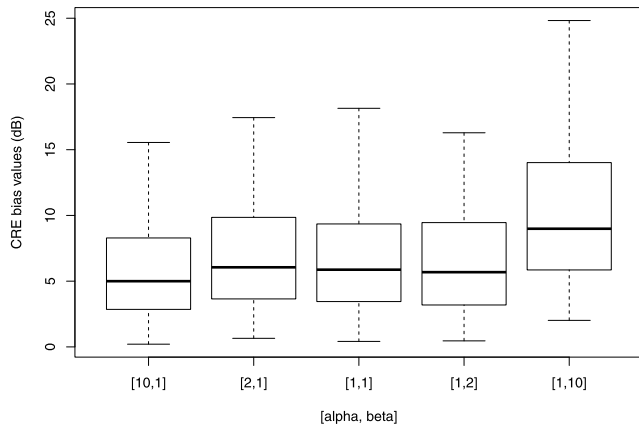
FIGURE 9. CRE bias values for PCB approach over α and β values.

simulations, that swarms with 60 or 80 particles could more properly exploit and explore the d-dimensional space, where $d = |\gamma|$.

Additionally, when the swarm has 80 particles, we observe the highest values achieved. Moreover, the Stochastic-IW PSO variant presents the worst results observed, mainly when the swarm has 20 particles. Nevertheless, as the population size is set to higher values, we can recognize that the observed results are statistically representative to demonstrate the effectiveness of the technique employed.

Finally, Fig. 9 shows the dispersion of CRE bias values obtained by the PCB approach. For all combinations of α and β, the CRE bias calculated values are mainly in the range [0, 20.0] dB. Thus, it represents a substantial distinct range of values from the UCB unified approach, whose best results are approximately 30.0 dB. For all cases, the median bias is placed in the [5.0, 10.0] dB interval, thus confirming that the PCB approach tends to select solutions with low CRE bias values compared to those observed in previous experiments. Such behavior suggests that few dBs may be sufficient to achieve a promising network balance. The burden on the MBSs is thereby reduced, and the SBS resources tend to be better utilized.

VI. CONCLUSION AND FUTURE WORK
In this paper, we considered a HetNet environment, proposing user association schemes and load balancing through a bioinspired approach using a particle swarm optimization algorithm. In the K-tier HetNet, UEs traditionally connect to the highest received signal power, resulting in lower association with SBSs, thus leading to a network imbalance process. By imposing a PSO-based user association mechanism that considers the volume of UEs meeting downlink requirements, the proposed scheme offloads the UEs toward SBSs. The application of this technique may lead to a better balance in the network without the resolution of combinatorial optimization problems or the use of excessive and nonstandard signaling. This PSO-based approach shows promising results by reducing the network imbalance and providing satisfying throughput levels per UE.

Future works include the adoption of dynamic clustering methods to reduce search space dimensionality and improve the effectiveness of the results obtained. CoMP scenarios may also be analyzed to consider multiple BS-UE associations to assess energy and spectral efficiency with high user mobility. We have foreseen a wider spatial distribution of additional BSs and UEs, and future works may also evaluate the adoption of the proposed scheme in device-to-device (D2D) communications to improve the association process and to stimulate cooperation behavior among the UEs.

REFERENCES

[1] S. Jose, "Cisco visual networking index (VNI) global mobile data traffic forecast update, 2017–2022, white paper," Cisco, San Jose, CA, USA, White Paper c11-738429, 2019.

[2] R. Vannithamby and S. Talwar, *Introduction.* Hoboken, NJ, USA: Wiley, 2017.

[3] A. Osseiran, F. Boccardi, V. Braun, K. Kusume, P. Marsch, M. Maternia, O. Queseth, M. Schellmann, H. Schotten, H. Taoka, H. Tullberg, M. A. Uusitalo, B. Timus, and M. Fallgren, "Scenarios for 5G mobile and wireless communications: The vision of the METIS project," *IEEE Commun. Mag.*, vol. 52, no. 5, pp. 26–35, May 2014.

[4] D. Lopez-Perez, I. Guvenc, and X. Chu, "Mobility management challenges in 3GPP heterogeneous networks," *IEEE Commun. Mag.*, vol. 50, no. 12, pp. 70–78, Dec. 2012.

[5] D. Lopez-Perez, I. Guvenc, G. de la Roche, M. Kountouris, T. Quek, and J. Zhang, "Enhanced intercell interference coordination challenges in heterogeneous networks," *IEEE Wireless Commun.*, vol. 18, no. 3, pp. 22–30, Jun. 2011.

[6] D. Liu, L. Wang, Y. Chen, M. Elkashlan, K.-K. Wong, R. Schober, and L. Hanzo, "User association in 5G networks: A survey and an outlook," *IEEE Commun. Surveys Tuts.*, vol. 18, no. 2, pp. 1018–1044, 2nd Quart., 2016.

[7] Q. Kuang, J. Belschner, Z. Bleicher, H. Droste, and J. Speidel, "A measurement-based study of handover improvement through range expansion and interference coordination," *Wireless Commun. Mobile Comput.*, vol. 15, no. 14, pp. 1784–1798, Jan. 2014.

[8] B. B. Kumar, L. Sharma, and S.-L. Wu, "Online distributed user association for heterogeneous radio access network," *Sensors*, vol. 19, no. 6, p. 1412, Mar. 2019.

[9] A. Damnjanovic, J. Montojo, Y. Wei, T. Ji, T. Luo, M. Vajapeyam, T. Yoo, O. Song, and D. Malladi, "A survey on 3GPP heterogeneous networks," *IEEE Wireless Commun.*, vol. 18, no. 3, pp. 10–21, Jun. 2011.

[10] J. G. Andrews, S. Singh, Q. Ye, X. Lin, and H. S. Dhillon, "An overview of load balancing in hetnets: Old myths and open problems," *IEEE Wireless Commun.*, vol. 21, no. 2, pp. 18–25, Apr. 2014.

[11] X. Nie, Y. Wang, J. Zhang, and L. Ding, "Coverage and association bias analysis for backhaul constrained HetNets with eICIC and CRE," *Wireless Pers. Commun.*, vol. 97, no. 4, pp. 4981–5002, Aug. 2017.

[12] G. Hattab and D. Cabric, "Rate-based cell range expansion for downlink massive MIMO heterogeneous networks," *IEEE Wireless Commun. Lett.*, vol. 7, no. 3, pp. 296–299, Jun. 2018.

[13] O. G. Aliu, A. Imran, M. A. Imran, and B. Evans, "A survey of self organisation in future cellular networks," *IEEE Commun. Surveys Tuts.*, vol. 15, no. 1, pp. 336–361, 1st Quart., 2013.

[14] X. Dong, F.-C. Zheng, X. Zhu, and J. Luo, "HetNets with range expansion: Local delay and energy efficiency optimization," *IEEE Trans. Veh. Technol.*, vol. 68, no. 6, pp. 6147–6150, Jun. 2019.

[15] Q. Ye, B. Rong, Y. Chen, M. Al-Shalash, C. Caramanis, and J. G. Andrews, "User association for load balancing in heterogeneous cellular networks," *IEEE Trans. Wireless Commun.*, vol. 12, no. 6, pp. 2706–2716, Jun. 2013.

[16] S. Singh and J. G. Andrews, "Joint resource partitioning and offloading in heterogeneous cellular networks," *IEEE Trans. Wireless Commun.*, vol. 13, no. 2, pp. 888–901, Feb. 2014.

[17] C. Zhang, P. Patras, and H. Haddadi, "Deep learning in mobile and wireless networking: A survey," *IEEE Commun. Surveys Tuts.*, vol. 21, no. 3, pp. 2224–2287, 3rd Quart., 2019.

[18] S. Harizan and P. Kuila, *Evolutionary Algorithms for Coverage and Connectivity Problems in Wireless Sensor Networks: A Study*. Singapore: Springer, 2020, pp. 257–280.

[19] S. Singh and S. Prakash, "A survey on software defined network based on architecture, issues and challenges," in *Proc. 3rd Int. Conf. Comput. Methodol. Commun. (ICCMC)*, Mar. 2019, pp. 568–573.

[20] X. Wang, X. Li, and V. C. M. Leung, "Artificial intelligence-based techniques for emerging heterogeneous network: State of the arts, opportunities, and challenges," *IEEE Access*, vol. 3, pp. 1379–1391, 2015.

[21] Z. Zhang, W. Huangfu, K. Long, X. Zhang, X. Liu, and B. Zhong, "On the designing principles and optimization approaches of bio-inspired self-organized network: A survey," *Sci. China Inf. Sci.*, vol. 56, no. 7, pp. 1–28, Jul. 2013.

[22] Z. Zhang, K. Long, J. Wang, and F. Dressler, "On swarm intelligence inspired self-organized networking: Its bionic mechanisms, designing principles and optimization approaches," *IEEE Commun. Surveys Tuts.*, vol. 16, no. 1, pp. 513–537, 1st Quart., 2014.

[23] S. Fichera, M. Gharbaoui, P. Castoldi, B. Martini, and A. Manzalini, "On experimenting 5G: Testbed set-up for SDN orchestration across network cloud and IoT domains," in *Proc. IEEE Conf. Netw. Softwarization (NetSoft)*, Jul. 2017, pp. 1–6.

[24] R. Martínez, A. Mayoral, R. Vilalta, R. Casellas, R. Muñoz, S. Pachnicke, T. Szyrkowiec, and A. Autenrieth, "Integrated SDN/NFV orchestration for the dynamic deployment of mobile virtual backhaul networks over a multilayer (Packet/Optical) aggregation infrastructure," *J. Opt. Commun. Netw.*, vol. 9, no. 2, pp. A135–A142, Jan. 2017.

[25] A. Mayoral, R. Vilalta, R. Muñoz, R. Casellas, and R. Martínez, "SDN orchestration architectures and their integration with cloud computing applications," *Opt. Switching Netw.*, vol. 26, pp. 2–13, Nov. 2017.

[26] M. Shojafar, L. Chiaraviglio, N. Blefari-Melazzi, and S. Salsano, "P5G: A bio-inspired algorithm for the superfluid management of 5G networks," in *Proc. IEEE Global Commun. Conf. (GLOBECOM)*, Dec. 2017, pp. 1–7.

[27] T. M. Shami, D. Grace, and A. Burr, "Load balancing and control using particle swarm optimisation in 5G heterogeneous networks," in *Proc. Eur. Conf. Netw. Commun.(EuCNC)*, Jun. 2018, pp. 1–9.

[28] M. Elkourdi, A. Mazin, and R. D. Gitlin, "Towards low latency in 5G HetNets: A Bayesian cell selection / user association approach," in *Proc. IEEE 5G World Forum (5GWF)*, Jul. 2018, pp. 268–272.

[29] W. K. Lai and J.-K. Liu, "Cell selection and resource allocation in LTE-advanced heterogeneous networks," *IEEE Access*, vol. 6, pp. 72978–72991, 2018.

[30] K. M. Nasr and K. Moessner, "Knapsack optimisation versus cell range expansion for mobility load balancing in dense small cells," in *Proc. Eur. Conf. Netw. Commun. (EuCNC)*, Jun. 2018, pp. 1–9.

[31] D. Fooladivanda and C. Rosenberg, "Joint resource allocation and user association for heterogeneous wireless cellular networks," *IEEE Trans. Wireless Commun.*, vol. 12, no. 1, pp. 248–257, Jan. 2013.

[32] A. Shaverdian, J. Ghimire, and C. Rosenberg, "Simple and efficient network-aware user association rules for heterogeneous networks," *Comput. Netw.*, vol. 156, pp. 20–32, Jun. 2019.

[33] H.-S. Jo, Y. J. Sang, P. Xia, and J. G. Andrews, "Heterogeneous cellular networks with flexible cell association: A comprehensive downlink SINR analysis," *IEEE Trans. Wireless Commun.*, vol. 11, no. 10, pp. 3484–3495, Oct. 2012.

[34] K. Da Costa Silva, Z. Becvar, and C. R. L. Frances, "Adaptive hysteresis margin based on fuzzy logic for handover in mobile networks with dense small cells," *IEEE Access*, vol. 6, pp. 17178–17189, 2018.

[35] H. Ramazanali, A. Mesodiakaki, A. Vinel, and C. Verikoukis, "Survey of user association in 5G HetNets," in *Proc. 8th IEEE Latin-Amer. Conf. Commun.(LATINCOM)*, Nov. 2016, pp. 1–6.

[36] J. Sangiamwong, Y. Saito, N. Miki, T. Abe, S. Nagata, and Y. Okumura, "Investigation on cell selection methods associated with inter-cell interference coordination in heterogeneous networks for LTE-advanced downlink," in *Proc. 11th Eur. Wireless Conf. Sustain. Wireless Technol. (Eur. Wireless)*, 2011, pp. 1–6.

[37] A. K. Kar, "Bio inspired computing—A review of algorithms and scope of applications," *Expert Syst. Appl.*, vol. 59, pp. 20–32, Oct. 2016.

[38] J. Kennedy and R. Eberhart, "Particle swarm optimization," in *Proc. IEEE Int. Conf. Neural Netw.*, vol. 4, Nov. 1995, pp. 1942–1948.

[39] S. Kefi, N. Rokbani, P. Kromer, and A. M. Alimi, "Ant supervised by PSO and 2-Opt algorithm, AS-PSO-2Opt, applied to traveling salesman problem," in *Proc. IEEE Int. Conf. Syst., Man, Cybern. (SMC)*, Oct. 2016, pp. 4866–4871.

[40] M. R. Bonyadi, Z. Michalewicz, and X. Li, "An analysis of the velocity updating rule of the particle swarm optimization algorithm," *J. Heuristics*, vol. 20, no. 4, pp. 417–452, Apr. 2014.

[41] R. Poli, J. Kennedy, and T. Blackwell, "Particle swarm optimization: An overview," *Swarm Intell.*, vol. 1, no. 1, pp. 33–57, 2007.

[42] S. Ghosh, S. Das, D. Kundu, K. Suresh, B. K. Panigrahi, and Z. Cui, "An inertia-adaptive particle swarm system with particle mobility factor for improved global optimization," *Neural Comput. Appl.*, vol. 21, no. 2, pp. 237–250, Apr. 2010.

[43] M. Clerc and J. Kennedy, "The particle swarm-explosion, stability, and convergence in a multidimensional complex space," *IEEE Trans. Evol. Comput.*, vol. 6, no. 1, pp. 58–73, Aug. 2002.

[44] R. Santos, G. Borges, A. Santos, M. Silva, C. Sales, and J. C. W. A. Costa, "A semi-autonomous particle swarm optimizer based on gradient information and diversity control for global optimization," *Appl. Soft Comput.*, vol. 69, pp. 330–343, Aug. 2018.

[45] Z.-H. Zhan, J. Zhang, Y. Li, and H. S.-H. Chung, "Adaptive particle swarm optimization," *IEEE Trans. Syst., Man, Cybern. B, Cybern.*, vol. 39, no. 6, pp. 1362–1381, Dec. 2009.

[46] D. Chen and C. Zhao, "Particle swarm optimization with adaptive population size and its application," *Appl. Soft Comput.*, vol. 9, no. 1, pp. 39–48, Jan. 2009.

[47] Y. Shi and R. Eberhart, "A modified particle swarm optimizer," in *Proc. IEEE Int. Conf. Evol. Comput., IEEE World Congr. Comput. Intell.*, May 1998, pp. 69–73.

[48] Y. Shi and R. C. Eberhart, "Parameter selection in particle swarm optimization," in *Evolutionary Programming VII* (Lecture Notes in Computer Science), vol. 1447, V. W. Porto, N. Saravanan, D. Waagen, and A. E. Eiben, Eds. Berlin, Germany: Springer, 1998, doi: 10.1007/BFb0040810.

[49] *Evolved Universal Terrestrial Radio Access (E-UTRA); Further Advancements for E-UTRA Physical Layer Aspects (Release 9)*, document 3GPP TR 36v9, 2010.

[50] M. Fiorani, S. Tombaz, F. S. Farias, L. Wosinska, and P. Monti, "Joint design of radio and transport for green residential access networks," *IEEE J. Sel. Areas Commun.*, vol. 34, no. 4, pp. 812–822, Apr. 2016.

[51] E. Sturzinger, M. Tornatore, and B. Mukherjee, "Application-aware resource provisioning in a heterogeneous Internet of Things," in *Proc. Int. Conf. Opt. Netw. Des. Modeling (ONDM)*, May 2017, pp. 1–6.

[52] Z. Gao, Z. Li, N. Wang, D. Wang, and X. Mu, "On SINR biasing for optimal cell range expansion in two-tier wireless hetnets," in *Proc. 7th IET Int. Conf. Wireless, Mobile Multimedia Netw. (ICWMMN)*, 2017, p. 13.

[53] M. J. Islam, X. Li, and Y. Mei, "A time-varying transfer function for balancing the exploration and exploitation ability of a binary PSO," *Appl. Soft Comput.*, vol. 59, pp. 182–196, Oct. 2017.

[54] O. Olorunda and A. P. Engelbrecht, "Measuring exploration/exploitation in particle swarms using swarm diversity," in *Proc. IEEE Congr. Evol. Comput., IEEE World Congr. Comput. Intell.*, Jun. 2008, pp. 1128–1134.

[55] F. Chen, X. Sun, D. Wei, and Y. Tang, "Tradeoff strategy between exploration and exploitation for PSO," in *Proc. 7th Int. Conf. Natural Comput.*, vol. 3, Jul. 2011, pp. 1216–1222.

[56] Y. L. Zheng, L. H. Ma, L. Y. Zhang, and J. X. Qian, "Empirical study of particle swarm optimizer with an increasing inertia weight," in *Proc. Congr. Evol. Comput. (CEC)*, 2003, pp. 221–226.

[57] R. C. Eberhart and Y. Shi, "Tracking and optimizing dynamic systems with particle swarms," in *Proc. IEEE Congr. Evol. Comput.*, vol. 1, May 2001, pp. 94–100.

HUGO PEREIRA KURIBAYASHI received the B.S. and M.S. degrees in computer science from the Federal University of Pará (UFPA), Belém, Brazil, in 2007 and 2011, respectively, where he is currently pursuing the Ph.D. degree in electrical engineering with an emphasis on applied computing. From 2013 to 2017, he was the Director of the Information and Communication Technology Center, Federal University of South and Southeast Pará (Unifesspa). From 2017 to 2019, he was the Pro-Rector of administration. He is currently an Assistant Professor with the Faculty of Computing and Electrical Engineering, Unifesspa. His research interests are next-generation networks, energy efficiency, smart radio resource management for wireless mobile networks, and the Internet of Things.

MARCELA ALVES DE SOUZA received the B.S. degree in information systems from the Federal University of Pará (UFPA), Marabá, Brazil, in 2012, and the M.S. degree in electrical engineering with emphasis on applied computing from UFPA, Belém, Brazil, in 2016, where she is currently pursuing the Ph.D. degree. Since 2014, she has been an Information Technology Analyst with the Information and Communication Technology Center, Federal University of South and Southeast Pará (Unifesspa). She has experience in information systems, software analysis, and development, process modeling, database administration, and programming languages. Her research interests include energy efficiency, wireless mobile networks, energy harvesting, and technoeconomic analysis.

DIEGO DE AZEVEDO GOMES received the Ph.D. degree in telecommunications from the Federal University of Pará, Brazil, in 2017. He is currently a Professor with the Faculty of Computer and Electrical Engineering, Federal University of Southern and Southeastern Pará. His research interest includes MIMO communication systems and digital signal processing applied to machine learning.

KETYLLEN DA COSTA SILVA received the B.Sc. degree in computer science from Amazon University, in 2010, and the M.Sc. and Ph.D. degrees in electrical engineering, with an emphasis in applied computing, from the Federal University of Pará, in 2014 and 2018, respectively. She completed internships with the Technological Institute of Aeronautics, São Paulo, in 2012, and Czech Technical University, Prague, in 2018. She is currently a member of the Laboratory of High Performance Networks (LPRAD). Her research interests include radio access technology, performance evaluation, mobility management, and heterogeneous networks.

MARCELINO SILVA DA SILVA received the bachelor's degree in computer engineering from the Federal University of Pará (UFPA), Belém, Brazil, in 2006, and the master's and Ph.D. degrees in electrical engineering from UFPA, in 2008 and 2014, respectively. He is currently a Professor with the Telecommunication and Computer Engineering Department, UFPA. His research interests include optimization, performance evaluation, Markov processes, and numerical models.

JOÃO C. WEYL ALBUQUERQUE COSTA (Senior Member, IEEE) received the B.Sc. degree in electrical engineering from the Federal University of Pará (UFPA), Belém, Brazil, in 1981, the M.Sc. degree in electrical engineering from the Pontifical Catholic University of Rio de Janeiro, Rio de Janeiro, Brazil, in 1989, and the Ph.D. degree in electrical engineering from the State University of Campinas, Campinas, Brazil, in 1994. He is currently a Professor with the Institute of Technology, UFPA, and also a Researcher with the Brazilian Research Funding Agency National Council for Scientific and Technological Development, Brasília, Brazil. His current research interests include broadband systems and optical sensors.

CARLOS RENATO LISBÔA FRANCÊS received the B.S. degree in computer science from the Federal University of Pará (UFPA), in 1995, and the M.S. and Ph.D. degrees from the University of São Paulo, Brazil, in 1998 and 2001, respectively. He held a postdoctoral position with the Institute for Systems and Computer Engineering, Technology and Science, Porto, Portugal, in 2012. He is currently an Associate Professor with the Telecommunication and Computer Engineering Department, UFPA. His research interests include performance evaluation, optimization, Markov processes, and artificial intelligence models.

• • •