

# Load balancing and control using particle swarm optimisation in 5G heterogeneous networks

Tareq M. Shami, David Grace, Alister Burr

Communication Technologies Research Group, Department of Electronic Engineering  
University of York, York, YO10 5DD, United Kingdom  
{tareq.al-shami, david.grace, alister.burr}@york.ac.uk

**Abstract**—Most of users in heterogeneous networks (HetNets) associate to macro base stations (BSs) due to their high transmission power while small cell BSs are underutilized. Current research has addressed this load balancing problem based on biased user association where all small cell BSs that belong to the same tier, e.g femto or pico BSs are assigned a static biasing value in order to increase their coverage area. This work utilises particle swarm optimisation to assign each small cell BS a certain biasing value with the objective of maximising the achievable throughput and controlling the load per-BS. Simulation results show that the proposed PSO approach achieves better performance in terms of the achievable throughput. In addition, the proposed approach is able to provide fairness among users by controlling the load per-BS.

**Keywords**—Load balancing; User association; heterogeneous networks; particle swarm optimisation;

## I. INTRODUCTION

One of the promising solutions to meet future 5G requirements is to densely deploy heterogeneous networks (HetNets). Traditionally, users associate to the base station (BS) that provides the highest received power signal; however, this conventional user association approach can cause load imbalance where most of the users associate to the macro BSs due to their high transmission power while small cells, e.g., femto and pico BSs, remain lightly loaded.

3GPP has addressed the load imbalance problem in HetNets in Release 10 by introducing the concept of cell range expansion (CRE). In CRE, the coverage of small BSs is increased by adding a bias value to the users' received power from small cells. CRE does not only offload users from macro BSs to small cells BSs, it also improves energy efficiency as shown in [1] since users are forced to associate with small cells that are now connected to BSs that are geographically closer. The biasing concept forms a cell-less architecture where a user is not necessarily connected to the nearest BS. Also, a user can associate with a BS even if it resides out of its nominal cell boundary due to the artificial coverage expansion of BSs.

The work in [2] proved that biased user association can improve capacity. However users that are in the CRE region are not served with the best downlink and they suffer from harmful interference caused by the nearby macrocell [3]. Therefore, it is crucial to carefully choose the bias value that can jointly balance the load and maximise throughput [4].

The aim of this paper is to illustrate how particle swarm optimisation (PSO) can be used to assign a proper bias value to

each small cell BS with the objective of jointly balancing the load, controlling the number of users that can associate with each BS, and maximising the network throughput. The state-of-the-art on biased user association has focused on finding the optimal biasing value per tier, i.e., all BSs that belong to the same tier are assigned the same biasing value [5-8].

The authors in [9] applied Q-learning to allow each UE to decide its own bias value instead of finding the optimal bias value for each tier. In the proposed Q-learning algorithm, each UE can learn its optimal bias value based on its historical experience. The proposed Q-learning approach can perform better in terms of the overall throughput as compared with the approaches in [5, 7] where a common bias value is added. Recently, PSO has been proposed to dynamically generate biasing values in order to balance the load among different tiers [10]. It was shown that the proposed PSO algorithm can perform better than the static biasing approach in terms of maximising the spectral efficiency. However, the authors did not compare the performance when the static biasing value goes above 10 dB. Moreover, the authors did not consider per-BS load control.

Although there exists some research on biased user association, no significant attention has been given to per-BS biasing and per-BS load control. Balancing the load per tier may cause some BSs to be overloaded or lightly loaded. Thus, it is essential to balance the load per-BS. Another important aspect that motivates the per-BS load control is that it aims to provide fairness among users.

This paper considers the downlinks of a HetNet where the objective is to balance and control the load as well as maximise the cell spectral efficiency. To achieve this objective, PSO as an optimisation tool is utilised to assign each BS a bias value dynamically. To validate the effectiveness of the proposed approach, it is compared against the traditional static biasing where all BSs that belong to the same small cell tier, i.e. femto and pico tiers, are assigned a predefined biasing value.

The organisation of this paper is as follows. The system model is described in Section II. Section III explains how PSO can be used to generate the bias value for each small cell BS. Results and discussion are presented in Section IV. Finally, section V provides the conclusion of this work.

## II. SYSTEM MODEL

The system model of this work considers a downlink three-tier HetNet. Tier 1, tier 2, and tier 3 represent the macrocells, picocells, and femto cells, respectively. BSs that belong to the

same tier have the same transmission power and density. The set  $N = \{1, 2, \dots, n\}$  denotes all the BSs in the system where the first element represents the macrocell and the small cells (pico and femto) are represented by the rest of the elements.

The signal to interference noise ratio (SINR) that is received by user  $i$  from BS  $j$  is calculated as follows:

$$SINR_{ij} = \frac{p_j g_{ij}}{\sum_{l \in A, l \neq j} p_l g_{il} + \sigma^2} \quad (1)$$

Where  $p_j$  indicates the transmission power of the serving BS  $j$ ,  $g_{ij}$  is the channel gain (pathloss and shadowing) between a UE and its serving BS  $j$ ,  $A$  denotes the set of all BSs except the serving BS  $j$ ,  $\sigma^2$  is the noise power.

The transmission rate is modeled according to the truncated Shannon bound model:

$$T = \begin{cases} 0, & SINR < SINR_{min} \\ \gamma \log_2(1 + SINR), & SINR_{min} < SINR < SINR_{max} \\ Th_{max}, & SINR > SINR_{max} \end{cases} \quad (2)$$

where  $T$  is the rate in bps/Hz,  $SINR_{min}$  is the minimum SINR value that is required to guarantee satisfactory QoS,  $\alpha$  is the attenuation factor, and  $SINR_{max}$  is the maximum value of SINR to achieve the highest throughput,  $Th_{max}$ . The TSB parameters [11] are  $\gamma = 0.65$ ,  $SINR_{min} = 1.8\text{dB}$ ,  $SINR_{max} = 21\text{dB}$ ,  $Th_{max} = 4.5\text{bps/Hz}$ .

The throughput achieved by a user  $i$  from BS  $j$  is calculated as follows:

$$Th = B \log_2(1 + SINR_{ij}) \quad (3)$$

where  $B$  is the bandwidth allocated for user  $i$ .

In this work, user association is based on the maximum reference signal received power (RSRP) which is the standardised approach in LTE systems. In biased maximum RSRP, a user associates with the BS that provides the highest biased RSRP. This is expressed mathematically as follows:

$$RSRP' = RSRP' + bias \quad (4)$$

where  $bias$  is the bias value for each small cell BS. To improve the system performance, it is crucial to carefully choose the biasing values.

### III. DYNAMIC BIASING USING PSO

It is essential to develop efficient algorithms that can determine the optimal biasing values because improper biasing values would clearly degrade the system performance in terms of load balance and throughput. The biasing can be generated statically or dynamically, where in the static case the biasing values remain unchanged over time while in the dynamic case they vary over time. Also, the biasing values can be assigned per tier, e.g., all femto BSs are assigned the same biasing value, or per BS. In this work, we utilise PSO [12] to generate biasing values for each BS dynamically.

Generating the biasing values per BS dynamically is a promising approach to balance the load maximise the achieved throughput; however, finding the optimal biasing values for each BS is an NP hard problem. The optimal solution to find the biasing values is to perform an exhaustive search; however this approach is prohibitively complex. Due to its high solution accuracy, fast convergence and few controlling parameters, PSO is applied in this work to dynamically obtain the per-BS biasing values that can balance and control the load as well as maximise the achievable throughput.

PSO consists of a number of particles called a swarm where each particle can fly in the search space in order to find a better position. The initial positions of the particles are randomly created. Then, in each iteration, particles follow a leader known as global best position ( $gbest$ ) which is the best position that has been found so far in the whole swarm. Moreover, particles follow their own historical best position  $Pbest$ . Each particle updates its velocity and position based on the following equations:

$$v_{id} = wv_{id} + c_1 r_1 (Pbest_{id} - x_{id}) + c_2 r_2 (gbest_d - x_{id}) \quad (5)$$

$$x_{id} = x_{id} + v_{id} \quad (6)$$

where  $w, c_1, c_2$  are the three main controlling parameters of PSO known as inertia weight, cognitive acceleration coefficient and social acceleration coefficient, respectively.  $r_1$  and  $r_2$  are two uniform random variables in the range of  $[0, 1]$ .

The following explains how PSO is used to generate the biasing values per BS in order to balance and control the load. At the initialisation stage of PSO, a swarm of particles is generated where each particle represents a candidate solution. The dimension of each particle is the number of small cell BSs and the value of each dimension is limited by the search space (0dBm-30dBm). Then, in each iteration of the PSO process, each particle attempts to find a better solution by updating its velocity and position according to equations (5) and (6). After that, in order to control the load, the particles are divided into two groups: valid and invalid particles. A valid particle is a particle that satisfies the constraint in equation (9) and its velocity and position are updated in each iteration according to equation (5) and (6). To elaborate further, a particle is valid only if it can control the load per-BS. After this filtering process, the fitness of each valid article is calculated based on equation (8). The best historical valid particles are recorded as  $Pbest$  whereas the best valid particle that has obtained the best performance so far is recorded as  $gbest$  ( $gbest$  is the best particle among all  $Pbest$ ). The PSO algorithm continues until a stopping condition such as the maximum number of iterations is met. After the maximum number of iterations is reached, the  $gbest$  particle that represents the best achieved biasing values per BS is returned. The PSO process to dynamically generate the biasing values is illustrated in Fig. 1.

In this work, PSO searches for the best particle that can control the number of users that can associate to each BS while the CSE is still maximised. The CSE can be mathematically calculated as follows:

$$System\ throughput = \sum_{k=1}^N \sum_{i=1}^M D_{ki} Th_{ki} \quad (7)$$

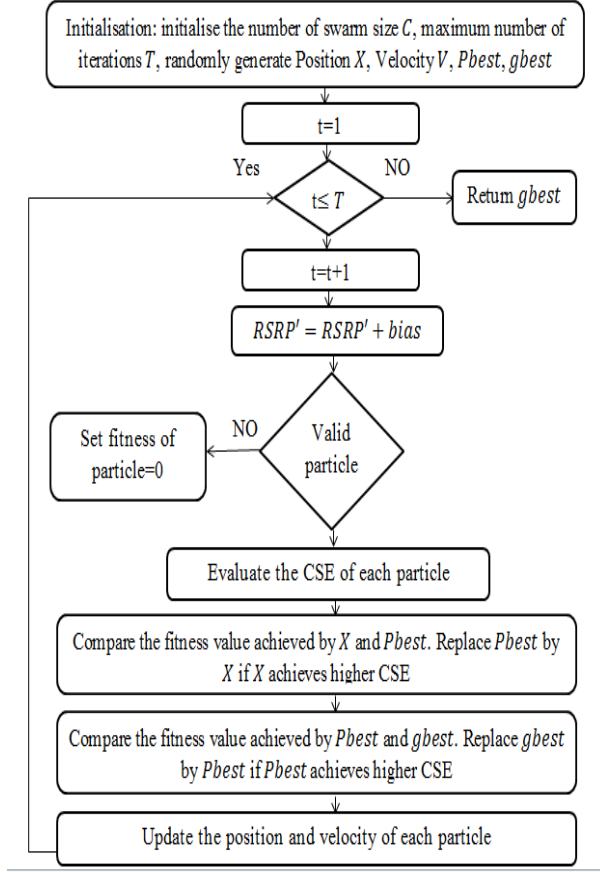


Fig. 1. The PSO process to generate dynamic biasing per-BS

$$CSE = \frac{\text{System throughput}/BW}{N} \quad (8)$$

where

$$D_{ki} = \begin{cases} 1 & , \text{ if a user } i \text{ is connected to BS } k \\ 0 & , \text{ if a user } i \text{ is not connected to BS } k \end{cases}$$

$N$  is the number of total BSs,  $M$  is the number of users,  $BW$  is the bandwidth. The value of  $D_{ki}$  is directly influenced by the biasing values of each particle where each particle generates different  $D_{ki}$ .

The objective function that PSO aims to optimise is formulated as follows:

$$\begin{aligned} & \max \quad CSE \\ \text{Subject to} \quad & \frac{M}{N} - \alpha M \leq K \leq \frac{M}{N} + \alpha M \end{aligned} \quad (9)$$

where  $\alpha$  is the spread load control parameter that is used to guarantee that the number of users ( $K$ ) that are associated with each BS does not go below or above a certain limit. The constraint in equation (9) is to ensure that a BS is not

overloaded or lightly loaded. If this constraint is satisfied, better load balance would be achieved.

#### IV. RESULTS AND DISCUSSION

The results in this work are obtained based on a MATLAB snapshot simulation. A total of 20 BSs (one macro BS, four pico BSs, and 15 femto BSs) are deployed in the area. The number of users represents the network load and it is considered in this work to be 500/km<sup>2</sup>. Table 1 presents the simulation parameters that are considered in this work. The PSO parameters that include the swarm size, maximum number of iterations, and controlling parameters are shown in Table 2.

TABLE 1: SIMULATION PARAMETERS

Parameter	Value
Bandwidth	20MHz
Tx Power (macro, pico, femto)	( 46dBm, 30dBm, 20dBm)
Macro pathloss [13]	128.1 + 37.6log <sub>10</sub> (R), R in km
Pico pathloss [13]	140.7 + 36.7log <sub>10</sub> (R), R in km
Femto pathloss [13]	127 + 30log <sub>10</sub> (R), R in km
Shadowing std. dev.	8dB (macro), 10dB (pico), 10dB (femto)
Noise power level	-174 dBm/Hz
Scheduler	Round robin
Traffic model	Full buffer

TABLE 2: PSO PARAMETER SETTINGS

Parameter	Setting
Swarm size	40
Maximum number of iterations	100
$c_1$	2
$c_2$	2
$w$	0.9-0.4

The traditional static biasing approach for values ranging from 0 dBm to 30 dBm is compared against the proposed dynamic approach that takes per-BS load control into account. The reason for choosing high biasing values up to 30 dBm is to ensure an optimal biasing value is included. The work in [8] have applied the same biasing range for the same reason. The comparison includes the user association per BS and the achievable throughput per user. The obtained results are averaged over 100 snapshots and they are shown in box plots form. Fig. 2. shows the number of users that are associated with each BS for the static biasing approach and for the proposed approach. From this figure, it is clear that when no bias is added (0dBm) there are some small cell BSs that are

lightly loaded while the macro cell is heavily loaded (an outlier that shows that the macro cell is overloaded was not shown in Fig.2 in order to improve the graph readability). The load starts to be balanced when a bias value of 10dBm is implemented; however, the macro cell is still heavily loaded. As the bias value goes above 10 dBm, users start to offload from the macro cell to the small cell BSs; nevertheless, a high bias value (>25 dBm) causes the macro cell BS to be lightly loaded. For the proposed approach, when the load control parameter  $\alpha$  is 3%, the maximum, minimum, and mean number of users is 38, 14, 25 respectively. This indicates that PSO is able to control the number of users per BS and this number follows the constraint in equation 9.

Fig. 3. compares the number of users per BS when the load control parameter  $\alpha$  is set to 3%, 3.5%, 4%, 4.5%, and 5%. The maximum value of the control parameter is set to be 5% because in the case of having 500 UEs/km<sup>2</sup> (the case considered in this work) the value of  $\frac{M}{N} - \alpha M$  goes below zero if  $\alpha$  is set higher than 5%. The maximum value of the load control parameter  $\alpha$  can be adjusted to have higher values if the number of users increases. As illustrated by Fig. 3., increasing the load control parameter causes load imbalance where some BSs become under-loaded and some BSs become heavily loaded. Therefore, it is important to keep the load control spread parameter as small as possible. Fig. 4. shows the user throughput for the static approach and the PSO with control. Comparing the static biasing for values from 0dBm to 30 dBm, 20 dBm outperforms all other static biasing values in terms of the achievable throughput where 75% of the total users can obtain throughput up to 1.64 Mbps. The reason that 20 dBm achieves better throughput is because it can achieve better load balance, as can be seen from Fig. 2.

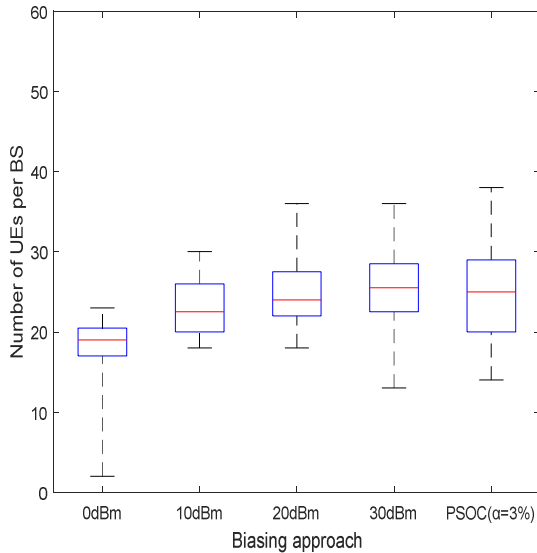


Fig. 2. A comparison between the static biasing and the PSO biasing with control in terms of number of users associated to each BS

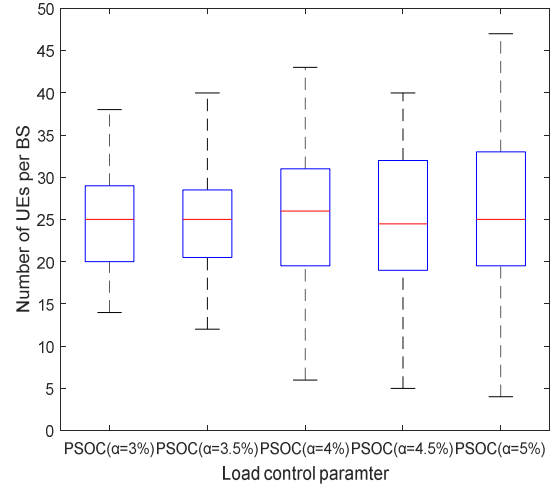


Fig. 3. Number of users per BS for different load control spread parameter.

However, the proposed PSO with 3% control outperforms the 20 dBm static biasing since it can provide throughput up to 2.04 Mbps to 75% of users.

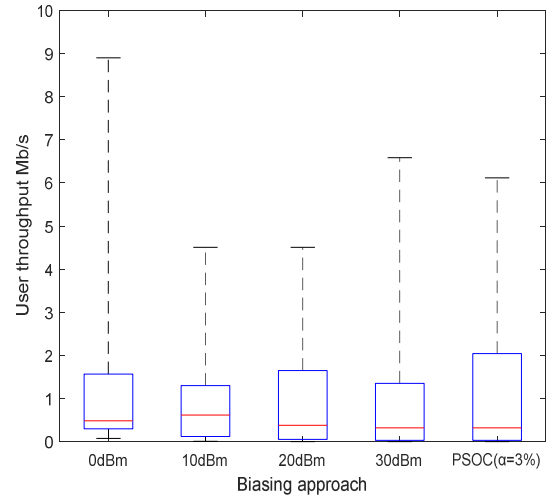


Fig. 4. A comparison of the static biasing and dynamic biasing with control in terms of user throughput

Fig. 5. shows the user throughput PSO with control for different values of the load control spread parameter. The figure shows that as the load control parameter increases, a small number of users would enjoy very high throughput; however, the achievable throughput of the majority of users decreases. This is because for a high spread value, some BSs will have few users, meaning that these users are assigned a large proportion of the available radio resources (physical resource blocks). On the other hand, some BSs will be highly loaded and users associated with these BSs will receive less radio resources, degrading their achievable throughput.

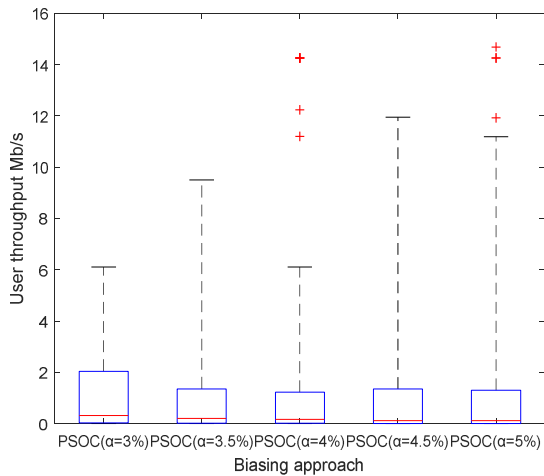


Fig. 5. User throughput for different load control values

## V. CONCLUSION

This work has utilised particle swarm optimization in order to balance and control the load per BS in 5G heterogeneous networks. The proposed approach was compared against the conventional static biasing approach in terms of load balance and achievable throughput. The simulation results showed that PSO can balance and control the load while the cell spectral efficiency is still maintained, with its performance being better than the static biasing method. The results also shows that the best performance of PSO is achieved when the load control parameter value is 3% (tight control) and the worst case is when the load control parameter value is 5% (loose control). PSO with load control parameter value of 3% can improve the user throughput where 75% of users can achieve up to 2.04 Mbps.

## ACKNOWLEDGMENT

This work is supported by the European Union's Horizon 2020 research programme through the 5G-AURA project under grant 675806.

## REFERENCES

- [1] R. Thakur, A. Sengupta, and C. S. R. Murthy, "Improving capacity and energy efficiency of femtocell based cellular network through cell biasing," in *Modeling & Optimization in Mobile, Ad Hoc & Wireless Networks (WiOpt)*, 2013 11th International Symposium on, 2013, pp. 436-443: IEEE.
- [2] I. Guvenc, "Capacity and fairness analysis of heterogeneous networks with range expansion and interference coordination," *IEEE Communications Letters*, vol. 15, no. 10, pp. 1084-1087, 2011.
- [3] H.-S. Jo, Y. J. Sang, P. Xia, and J. G. Andrews, "Outage probability for heterogeneous cellular networks with biased cell association," in *Global*

- Telecommunications Conference (GLOBECOM 2011)*, 2011 IEEE, 2011, pp. 1-5: IEEE.
- [4] E. Hossain, M. Rasti, H. Tabassum, and A. Abdelnasser, "Evolution toward 5G multi-tier cellular wireless networks: An interference management perspective," *IEEE Wireless Communications*, vol. 21, no. 3, pp. 118-127, 2014.
- [5] 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 *Wireless Conference 2011-Sustainable Wireless Technologies (European Wireless)*, 11th European, 2011, pp. 1-6: VDE.
- [6] I. Guvenc, M.-R. Jeong, I. Demirdogen, B. Kecioglu, and F. Watanabe, "Range expansion and inter-cell interference coordination (ICIC) for picocell networks," in *Vehicular Technology Conference (VTC Fall)*, 2011 IEEE, 2011, pp. 1-6: IEEE.
- [7] D. López-Pérez and X. Chu, "Inter-cell interference coordination for expanded region picocells in heterogeneous networks," in *Computer Communications and Networks (ICCCN)*, 2011 Proceedings of 20th International Conference on, 2011, pp. 1-6: IEEE.
- [8] 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 Communications*, vol. 21, no. 2, pp. 18-25, 2014.
- [9] T. Kudo and T. Ohtsuki, "Cell range expansion using distributed Q-learning in heterogeneous networks," *EURASIP Journal on Wireless Communications and Networking*, vol. 2013, no. 1, p. 61, 2013.
- [10] Y. Wang, S. Chen, H. Ji, and H. Zhang, "Load-aware dynamic biasing cell association in small cell networks," in *Communications (ICC)*, 2014 IEEE International Conference on, 2014, pp. 2684-2689: IEEE.
- [11] T. Jiang, A. Papadogiannis, D. Grace, and A. Burr, "Bungee deliverable: D4. 1.1 interim simulation," *Deliverable*, 2011.
- [12] R. Kennedy, "J. and Eberhart, Particle swarm optimization," in *Proceedings of IEEE International Conference on Neural Networks IV*, pages, 1995, vol. 1000.
- [13] 3GPP Organizational Partners, 3rd Generation Partnership Project; technical specification group radio access network; evolved universal terrestrial radio access (E-UTRA); further advancements for E-UTRA physical layer aspects (release 9). 3GPP TR 36.814 Release 9 V9.0.0 (2010-03) (2010). <http://www.quintillion.co.jp/3GPP/Specs/36814-900.pdf>