A Resource Allocation Mechanism for Cloud Radio Access Network Based on Cell Differentiation and Integration Concept

Zainab H. Fakhri, M. Khan, Firas Sabir, and H. S. Al-Raweshidy, *Senior Member, IEEE*

Abstract—A Self-Organising Cloud Radio Access Network (C-RAN) is proposed, which dynamically adapt to varying capacity demands. The Base Band Units and Remote Radio Heads are scaled semi-statically based on the concept of cell differentiation and integration (CDI) while a dynamic load balancing is formulated as an integer-based optimisation problem with constraints. A Discrete Particle Swarm Optimisation (DPSO) is developed as an Evolutionary Algorithm to solve load balancing optimisation problem. The performance of DPSO is tested based on two problem scenarios and compared to an Exhaustive Search (ES) algorithm. The DPSO deliver optimum performance for small-scale networks and near optimum performance for large-scale networks. The DPSO has less complexity and is much faster than the ES algorithm. Computational results demonstrate significant throughput improvement in a CDI-enabled C-RAN compared to a fixed C-RAN, i.e., an average throughput increase of 45.53 and 42.102 percent, and a decrease of 23.149 and 20.903 percent in the average blocked users is experienced for Proportional Fair (PF) and Round Robin (RR) schedulers, respectively. A power model is proposed to estimate the overall power consumption of C-RAN. A decrease of $\approx 16\%$ is estimated in a CDI-enabled C-RAN when compared to a fixed C-RAN, both serving the same geographical area.

Index Terms—Base band unit (BBU), cloud radio access network (C-RAN), particle swarm optimisation (PSO), remote radio head (RRH), self-optimising network (SON)

1 Introduction

In the past few years, the proliferation of personal handheld mobile computing devices such as tablets and smartphones, along with the growing volume of data-demanding services and applications, has produced and a great need for wireless access and high-speed data transmission. Internet access anywhere and everywhere has triggered the formation of radio *hot-spot* networks. The major challenge in cellular networks is managing the available resources in a way to achieve 1) Optimum returns on investment, 2) User's service demands satisfaction, and 3) High levels of network QoS. Unaware of the cell load, a user equipment (UE) associates itself to the cell providing the strongest signal. The spatial distribution of users and their capacity demands vary with respect to time, causing unbalanced traffic loads and wasteful utilisation of network's resources. Therefore, it is important to self-optimise the network resources dynamically.

To overcome the aforementioned challenges, C-RAN [1], [2], [3] has been proposed as a novel architecture that can address some significant challenges the Mobile Network Operators (MNOs) are facing with today. C-RAN architecture is composed of three parts: 1) The Base Band Units (BBUs) collected into a virtualised BBU cloud/pool for centralised processing, 2) The Remotely distributed Remote Radio Heads (RRHs) in the radio access network, and 3) An optical transport network (OTN) that connects the BBUs to the RRH. C-RAN can achieve significant cost and energy savings by dynamically scaling the BBUs with respect to changing traffic conditions [4] and adjusting the logical BBU-RRH links using suitable resource allocations schemes.

C-RAN with Self-optimising ability can provide MNOs with a flexible network regarding network dimensioning, adaptation to non-uniform traffic, and efficient utilisation of network resources. However, before a full commercial C-RAN deployment, several challenges need to be addressed. First, the front-haul technology used must support enough bandwidth for delivering delay sensitive signals. Second, the proper BBU-RRH assignment in C-RAN to not only support collaboration technology like Cooperative Multipoint Processing (CoMP) [5] but also enabling dynamic load balancing and power saving in the network.

The main motivation of this paper is to exploit the capacity routing ability of C-RAN by employing self-optimisation for efficient resource utilisation with high levels of QoS and a balanced network load. Inspired by the concept of cell splitting in biological sciences, a two-stage design is proposed for real-time BBU-RRH mapping and power saving in C-RAN. The main contributions of the proposed scheme is as follows:

1) The proposed mechanism monitors the load on each cell in a given geographical area and divides it into multiple small cells and vice versa if the load in a cell exceeds or falls a certain threshold.

• *The authors are with the Brunel University, Uxbridge, Middlesex UB8 3PH, United Kingdom. E-mail: {zainab.fakhri, muhammad.khan, Firas.sabir, hamed.al-raweshidy}@brunel.ac.uk.*

Manuscript received 6 July 2017; revised 13 Sept. 2017; accepted 15 Sept. 2017. Date of publication 16 Oct. 2017; date of current version 11 Dec. 2018. (Corresponding author: Zainab Hassan.)
Recommended for acceptance by X. Cao.
For information on obtaining reprints of this article, please send e-mail to: reprints@ieee.org, and reference the Digital Object Identifier below.
Digital Object Identifier no. 10.1109/TNSE.2017.2754101

2) The fitting number of BBUs required to serve all RRHs in the given geographical area is assigned based on the actual load on the network. A key challenge of initial BBU-RRH mapping before identifying an optimum BBU-RRH mapping is also addressed.
3) An Evolutionary Algorithm (EA) is proposed to find the optimum BBU-RRH configuration to balance the network load for enhanced QoS. Therefore, the resources can be utilised efficiently.

This paper is organised as follows: Section 2 presents a survey of related work. Section 3 presents the system model. Section 4 illustrates the formulation for dynamic BBU-RRH allocation problem; Section 5 presents the RRH clustering constraint; Section 6 presents the proposed C-RAN power model; Section 7 defines the CDI algorithm. Computational results are discussed in Section 8. Finally, the paper is concluded in Section 9.

2 Related Work

Artificial Intelligence (AI) techniques facilitate a network to automatically re-configure system parameters for optimum network performance and adaptively learn necessary system parameters to perform upgrades and maintenance routines along with recovering from failures. Since AI, is the basis of self-organising and machine learning network technologies, it can lead to a significant paradigm shift by driving the ongoing efforts in next-generation wireless network (5G) standardisation [6].

Most recent studies on resource management in C-RAN mainly focus on schemes related to RRH-UE mapping and only limited work addresses the BBU-RRH configuration schemes. Some related works on the former schemes are briefly discussed in [7], [8], [9]. In [7], the authors propose a QoS-aware radio resource optimisation solution for maximising downlink system utility in C-RAN. User grouping, virtual base stations clustering, and beamforming for multiuser, multicell distributed MIMO networks were investigated. In line with this work, the authors of [8] propose an efficient resource allocation scheme in heterogeneous C-RAN. A weighted minimum mean square error (WMMSE) approach is used to solve network-wide beamforming vectors optimisation and identify proper RRH-UE clusters. Moreover, minimising the number of active BBUs is formulated as a bin packing problem for energy saving. The work of [9] expresses a mixed integer nonlinear programming (MINLP) problem aiming joint RRH selection to minimise power consumption via beamforming, where the transport network power is determined by the set of active RRHs. Regarding the BBU pool in C-RAN, some studies are described in [10], [11], [12], [13]. A joint-scheduling strategy for resource allocation in C-RAN is proposed in [14] where the time/frequency resources of multiple base stations are jointly optimised to schedule network users concurrently for network throughput improvement. However, the authors did not consider BBU-RRH mapping and focused mainly on joint scheduling in C-RAN .The authors of [10] initially investigated semi-static and adaptive BBU-RRH switching schemes for C-RAN. The authors of [11] then proposed a lightweight, scalable framework that utilises optimal transmission strategies via BBU-RRH reconfiguration to cater dynamic user traffic profiles. A dynamic BBU-RRH mapping scheme is introduced in [12] using a borrow-and-lend approach in C-RAN. Overloaded BBUs switch their supported RRHs to underutilised BBUs for a balanced network load and enhanced throughput. The authors of [13] propose a load balancing technique which considers load fairness as an optimisation problem. When the load fairness exceeds an alarming threshold, the given geographical area is divided into small compact zones based on an infinite optimisation formulation. The author's previous work address a blocking probability based load balancing problem in C-RAN via evolutionary algorithms [15]. However, power saving in C-RAN was not addressed.

Regarding other related work, there have been attempts to develop Network Function Virtualisation (NFV) and Software Defined Network (SDN) solutions for C-RAN [16], [17], [18]. Although SDN and NFV are not the primary focus of this paper, they are presented in this section for a complete introduction of C-RAN. Moreover, an in-depth review of the principles, technologies and applications of C-RAN describing innovative concepts regarding physical layer, resource allocation, and network challenges together with their potential solutions are highlighted in [1], [19]

To sum up, the existing resource allocation mechanisms does not take full advantage of the centralised BBU pool concept in C-RAN. This paper extends the scope of C-RAN by introducing the concept of Cell Differentiation and Integration (CDI) with dynamic BBU-RRH mapping for load balancing and efficient resource utilisation. The system model in this article allows combining self-optimising feature of SON and capacity routing ability of C-RAN for a more centrally managed network operations.

3 System Model

3.1 Proposed C-RAN Architecture

A self-optimised C-RAN architecture is presented in Fig. 1. The BBUs are decoupled from the RRH and migrated to a centralised BBU-pool, whereas the RRHs are left on the cell sites. A SON controller is introduced inside the BBU cloud which monitors the BBU-pool resource utilisation as well as controlling the switch. Since an optical switch can only support one-to-one switching, soft switching (one-to-one and one-to-many) is enabled indirectly by using optical splitters and multiplexers [11]. The SON controller dynamically assigns BBU pool resources to the independent RRHs based on traffic demands. However, each BBU allocates its radio resources (PRBs) only to the RRHs assigned to it at a particular time.

At extremely low traffic load conditions, only a high power macro-BS serves the given geographical area. As the traffic load increases and the macro cell reaches its load limit, the geographic area is differentiated into $\mathbf{C}$ equally sized small cells serving the same coverage area. Each $\mathbf{C}$ cell can further differentiate into $\mathbf{c}$ more small cells by activating the CDI supporting RRHs deployed to accommodate capacity demands. The actual number of RRHs are determined by the coverage area, users density, and other environment-related factors. However, both $\mathbf{C}$ and $\mathbf{c}$ are considered to be seven as a reasonable example.

Furthermore, the CDI concept is realised by considering three tiers of RRHs deployment as shown in Fig. 2, i.e., tier-3 RRH deployment imitates a high-power base station serving a Macro cell as in traditional cellular systems. Tier-2 and tier-1

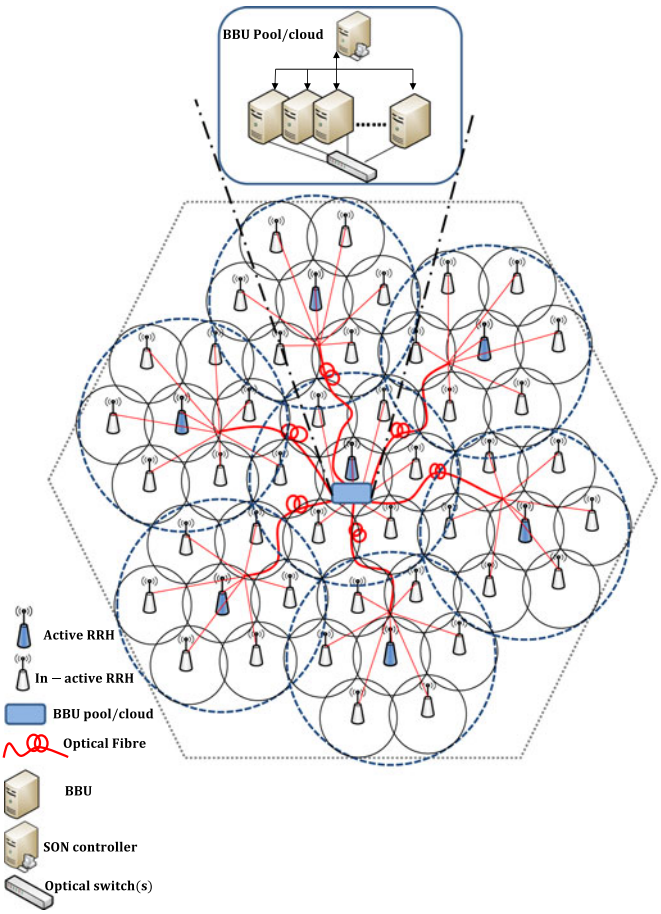


Fig. 1. Structure of a cloud radio access network represented as SON.

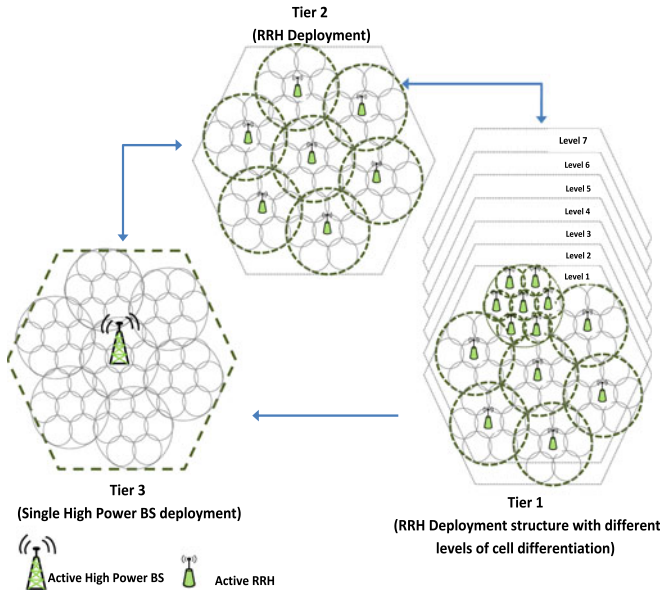


Fig. 2. Cell differentiation and integration with multiple tiers of RRH deployment.

represents a structure with universal frequency reuse, where each cell is surrounded by a continuous tier of $6+E$ and $6\times[1+j]+E$ cells, respectively. Where E represents the number of other external macro cells and j accounts for the level of differentiation. A set $\mathbf{S}_i=\{\mathrm{RRH}_{i1},\mathrm{RRH}_{i2},\ldots,\mathrm{RRH}_{ic}\}$ is maintained for each cell C_i in tier-2 RRH deployment, which contains a group of RRHs responsible for differentiating cell C_i into c small cells provided that the sum of transmit powers of all RRHs covers C_i coverage area. The central RRH of each cell C_i is represented as RRH_{i1}. Where i represents the cell number in tier-2 RRH structure. The SON server is responsible for cell differentiation and integration with proper BBU-RRH configurations, whereas the optical switch is in charge of realising the settings via server commands. Note that, with small cell deployment in C-RAN, a high inter-cell interference is inevitable. Therefore, a clustering based interference mitigation technique is adopted to avoid network performance degradation. RRHs served by the same BBU are grouped together based on a proximity property [15].

3.2 Channel Model

In this paper, Guaranteed Bit Rate (GBR) users with QoS requirements are considered. The frequency reuse factor is 1, and the time-frequency resources are equal for all BBUs. The basic unit of time-frequency resources that can be allocated to users per time slot (0.5 ms) of an LTE subframe is known as the Physical Resource Block (PRB). Each PRB consists of 12 consecutive sub-carriers with a sub-carrier spacing of 15 kHz, corresponding to 0.5 milli-seconds in time domain and 180 kHz in frequency domain. Let $\mathbf{M}$ and $\mathbf{N}$ represent the number of active BBUs and RRHs in the network, respectively, such that K_{in} represents the total number of users in cell i served by RRH n. Each user reports Channel Quality Information (CQI) to its serving BBU every two subframes (i.e., 2 milli-seconds) for proper PRB assignment. The channel model considered in this paper is a composite fading channel which involves path-loss and both small and large scale fading, given as

$$\mathrm{H}_{k_{in}}=\mathrm{h}^{*}_{k_{in}}\,l_{k_{in}}\left[\mathrm{AD}^{-\delta}_{k_{in}}\right], \tag{1}$$

where $\mathrm{h}^{*}_{k_{in}}$ and $l_{k_{in}}$ represent the small and large scale fading channel between the RRH n and user k in cell i, respectively. The small scale fading is assumed to be a Rayleigh random variables with a distribution envelop of zero-mean and unity-variance Gaussian process. $\mathrm{AD}^{-\delta}_{k_{in}}$ reflects the path-loss between RRH n and user k in cell i, where A is a constant which depends on the carrier frequency f_c and $\mathrm{D}_{k_{in}}$ is the distance between user k and RRH n in cell i and a path-loss exponent of δ. In this paper, a path-loss of $(\mathrm{A},\delta)=(1.35\times10^{7},3)$ is considered [20] . The large scale fading is assumed to be lognormal random variable with a standard deviation of 10 dB and is typically modelled with a probability density function of [21]

$$\rho(l)=\frac{\zeta}{\sqrt{2\pi}\sigma_l l}\exp\left[-\frac{(10\log_{10}l-\mu_l)^2}{2\sigma_l^2}\right], \tag{2}$$

where $\zeta=10/\ln10$, and μ_l and σ_l are the mean and the standard deviation of l, both expressed in decibels.

The instantaneous Signal-to-Interference-and-Noise-Ratio γ based on CQI received from user k in cell i served by RRH n at time-slot t is expressed as

$$\gamma_{k_{in}}(t)=\frac{\mathrm{H}_{k_{in}}(t)\mathrm{P}_{in}(t)}{\mathrm{N}_0+\sum_{j\in C}\sum_{a\in c,a\neq n}\mathrm{H}_{k_{ja}}(t)\mathrm{P}_{ja}(t)}, \tag{3}$$

where $\mathrm{P}_{in}(t)$ and $\mathrm{H}_{k_{in}}(t)$ are the transmit power and channel gain between the serving RRH n of user k at time-slot t in cell i . N_0 is the power of Additive White Gaussian Noise per PRB and $\sum_{j\in C}\sum_{a\in c,a\neq n}\mathrm{H}_{k_{ja}}(t)\mathrm{P}_{ja}(t)$ represents the inter-cell interference power received from all other active RRHs a at time-slot t in cells j except the serving RRH n of user k in cell i.

Assuming the best modulation coding scheme, the highest data rate achieved by user k served by RRH_n in cell i for a given SINR at time-slot t can be expressed by Shannon formula

$$\vartheta_{k_{in}}(t) = \log_2(1 + a\gamma_{k_{in}}(t)), \tag{4}$$

where a is the constant bit error rate (BER) defined as $a = -1.5/ln(5\times 10^{-6})$ [22]. The total PRB required by the user can be now be determined by the achievable throughput of the user k at a given SINR, the demanded data rate ϕ_k of user k, and the bandwidth P_{BW} of a single PRB (i.e., 180 KHz) from the following:

$$\mathrm{N}^k_{\mathrm{RB}}(t) = \left\lceil \frac{\phi_k(t)}{\mathrm{p}_{\mathrm{BW}}.\vartheta_{k_{in}}(t)} \right\rceil, \tag{5}$$

where p_{BW} represents the bandwidth of a PRB and the notion $\lceil . \rceil$ is the ceil function.

4 Dynamic BBU-RRH Configuration and Formulation

For a Self-optimising C-RAN architecture shown in Fig. 1, it is essential to balance the network load amongst the active BBUs by proper BBU-RRH configuration. After each CDI cycle, the network may reconfigure itself by scaling the BBUs and RRHs with respect to traffic load. However, during the process, the BBU-RRH mapping might not satisfy the QoS requirement. Therefore, If the BBU-RRH configuration at time t is known then it is necessary to adjust the BBU-RRH configuration at time $t+1$ to adaptively balance the variance in traffic demands. Note that, the time between t and $t+1$ is longer than that of a subframe (i.e., one millisecond) and is called the load balancing cycle. A user location indicator vector $\mathbf{u} = \{\mathrm{u}_1, \mathrm{u}_2, \ldots, \mathrm{u}_\mathrm{K}\}$ is defined which shows users association with RRHs such that $\mathrm{u}_\mathrm{k} = \{\mathrm{r}_{in} | \mathrm{r}_{in} \in \mathbb{Z}^+ : i, n = 1, 2, 3, \ldots C\}$, where $\mathrm{u}_\mathrm{k} = r_{in}$ if user k is associated with RRH_n of cell C_i. To indicate RRHs association with BBUs, a vector $\mathbf{r} = \{\mathrm{r}_{11}, \mathrm{r}_{12}, \ldots, \mathrm{r}_{in}\}$ is defined, where $\mathrm{r}_{in} \in \{1, 2, \ldots, \mathrm{M}\}$ and $\mathrm{r}_{in} = m$ indicates RRH_n of cell C_i is being served by BBU_m. Whereas, $\mathrm{r}_{in} = 0$ indicates that RRH_n of cell C_i is not active. If the user location indicator vector $\mathbf{u}$ is given, then the problem is to identify the new RRH allocation vector $\mathbf{r}$.

4.1 Number of BBUs Required in the Network

The required number of BBUs to serve the offered traffic load at a particular time t can be calculated using actual load $\eta(t)$ on the network. Let $\eta_m(t)$ be the load on BBU_m at time period t, which is represented as

$$\eta_m(t) = \frac{\sum_{k=1}^{\mathrm{K}} \mathrm{I}_{m,k}(t)\mathrm{N}^k_{\mathrm{RB}}(t)}{\mathrm{P}_{\mathrm{RB}}}. \tag{6}$$

Where $\mathrm{I}_{m,k}$ is a binary indicator such that $\mathrm{I}_{m,k} = 1$ if user k is served by BBU_m. However, an important constraint $\sum_{m=1}^{M} \mathrm{I}_{m,k} = 1, \forall k$ defines that each user k is served by only one BBU at time period t. Note that, all BBUs are assigned the same number of PRBs (P_{RB}). Another important constraint is that $\sum_{k=1}^{\mathrm{K}} \mathrm{I}_{m,k}(t)\mathrm{N}^k_{\mathrm{RB}}(t) \leq \mathrm{P}_{\mathrm{RB}}, \forall m$, which states that the number of PRBs assigned to users served by the same BBU should not exceed the BBU PRB limitation. The total load on the network at time t is represented as the aggregated load on each active BBU at time t, which is given by

$$\eta(t) = \sum_{m=1}^{\mathrm{M}} \eta_m(t). \tag{7}$$

Now the number of required BBUs (M) in the network at a particular time t can be given as

$$\text{No. of BBUs} = \mathrm{M} = \begin{cases} \lceil \eta(t) \rceil & if\ \eta(t) < \mathrm{M}_{\mathrm{total}} \\ |\mathrm{M}_{\mathrm{total}}| & if\ \eta(t) \geq \mathrm{M}_{\mathrm{total}}, \end{cases} \tag{8}$$

where $\mathrm{M}_{\mathrm{total}}$ is the total number of BBUs available in the BBU pool and the notation $\lceil . \rceil$ is the ceil function. Moreover, the load contributed by an active RRH_n of cell i in the network is given by

$$\eta_{\mathrm{RRH}_{in}}(t) = \sum_{k=1}^{\mathrm{K}} \mathrm{I}_{k,in}(t)\mathrm{N}^k_{\mathrm{RB}}(t). \tag{9}$$

Network performance determined by Key Performance Indicators (KPIs). Based on these KPIs, the SON server identifies optimum BBU-RRH conguration by utilising the existing number of active BBUs and RRHs, to achieve a highly stable network with highest achievable QoS with respect to load demand. Following are the important KPIs considered for BBU-RRH mapping problem;

4.2 Key Performance Indicator for Load Fairness Index

In this paper, a Jains fairness index ψ is monitored, which determines the level of load balancing in the network at a particular time t and is defined as

$$\psi(t) = \frac{\left(\sum_{m=1}^{\mathrm{M}} \eta_m(t)\right)^2}{\mathrm{M}\left(\sum_{m=1}^{\mathrm{M}} \eta_m^2(t)\right)}, \tag{10}$$

where M is the number of active BBUs. The range of ψ is in the interval $[\frac{1}{M}, 1]$, with a higher value representing a highly balanced load distribution amongst all active BBUs. Therefore, maximising ψ is one of the objectives to achieve a highly balanced load in the C-RAN.

4.3 Key Performance Indicator for Average Network Load

Minimising the average network load can avoid handovers of users with poor channel conditions in the system. A user (s) associated to an RRHs may have imperfect channel conditions with more PRBs requirement to meet desired data rate. Failure to meet the user's PRB demand, the BBU has to perform a handover operation. Therefore, to avoid unnecessary handovers, minimising the average network load is considered as a second objective and is given as

$$\eta_{ave}(t) = \frac{\sum_{m=1}^{\mathrm{M}} \eta_m(t)}{\mathrm{M}}, \tag{11}$$

where $\eta_m(t)$ is the load on a BBU_m defined in (6) and M is number of active BBUs calculated from (8).

4.4 Key Performance Indicator for Handovers

Network transition to a new BBU-RRH configuration may require significant forced handovers. An increased number of forced handovers in the system is undesirable and leads to performance degradation. Allocating an RRH to a new BBUs at a particular time results in forced handovers of all users associated with the RRH. Since inter-BBU handovers not only involves BBUs but a signalling overhead between the Serving Gateway (S-GW) and Mobility Management Entity (MME). Therefore, it is desirable to achieve a new optimum BBU-RRH configuration with a minimum required handovers. A handover index $\mathrm{h}(t)$ is monitored as a third objective for load balancing problem and is given as

$$\mathrm{h}(t) = \frac{1}{2}\left(\frac{\sum_{m=1}^{\mathrm{M}}\sum_{k=1}^{\mathrm{K}} |\mathrm{I}_{m,k}(t) - \mathrm{I}^{\circ}_{m,k}(t)|}{\mathrm{K}}\right), \tag{12}$$

where $\mathrm{I}^{\circ}_{m,k}(t)$ is a binary variable that indicates a user's association to BBU in previous BBU-RRH configuration, i.e., $\mathrm{I}^{\circ}_{m,k}(t) = 1$, if user k is served by BBU_m in previous BBU-RRH configuration.

5 RRH Clustering

Proper BBU-RRH allocation can provide enhanced flexibility in C-RAN network management. However, an important limitation to consider is the reliable operation of C-RAN regarding BBU-RRH mapping for high system performances. Neighbouring RRHs must be assigned to the same BBU [12] to support advance LTE-A features like Coordinated MultiPoint transmission (CoMP) and to avoid unnecessary handovers among network cells. RRHs clustering is an approach to support CoMP for interference mitigation in LTE-A and C-RAN [23]. Therefore, this paper considers that the RRHs served by the same BBU forms a compact cluster. This compactness and consistently connected RRHs in a cluster, not only minimises frequent handovers among cells but also reduces the inter-cell interference among them. This is because compact an RRH group shares fewer common boundaries with other RRH groups. Therefore, the RRHs proximity is defined by introducing a binary variable A_{ij}, where $\mathrm{A}_{ij} = 1$, if RRH_i and RRH_j are adjacent else A_{ij}= 0. If a cluster has multiple RRHs, then the RRHs in that cluster must be adjacent and connected. To formulate the connectedness of a cluster and proximity of the RRHs, let S1 be any proper subset of the set of RRHs served by BBU_m (Z_m), such that $\mathrm{S1} \subset \mathrm{Z}_m, \mathrm{S1} \neq \emptyset$, and $\mathrm{S1} \neq \mathrm{Z}_m$. Let S2 be another subset of Z_m such that, $\mathrm{S2} = \mathrm{Z}_m - \mathrm{S1}$, i.e., S2 is the complementary set of S1. To confirm that the RRHs in Z_m are connected, the following property must be satisfied.

$$\sum_{i\in \mathrm{S1}}\sum_{j\in \mathrm{S2}} \mathrm{A}_{ij} \geq 1. \tag{13}$$

For proper BBU-RRH configuration, a QoS function is needed which is the weighted combination of KPIs defined in (10), (11), and (12). The multiple objectives are combined into a single QoS objective function. This paper represents QoS as the following maximisation problem with constraints

$$\begin{aligned} \mathrm{Max} \quad & \mathrm{QoS}(t) = \alpha\psi(t) - \beta\eta_{ave}(t) - (1-\alpha-\beta)\mathrm{h}(t) \\ s.t. \quad & \mathrm{C}_1 : \sum_{i\in \mathrm{S1}}\sum_{j\in \mathrm{S2}} \mathrm{A}_{ij} \geq 1, \forall \mathrm{S1}, \mathrm{S2} \in \mathrm{Z}_m, \forall m \in \{1, 2, \ldots \mathrm{M}\} \\ & \mathrm{C}_2 : \sum_{k=1}^{\mathrm{K}} \mathrm{I}_{m,k}(t)\mathrm{N}^{k}_{\mathrm{RB}} \leq \mathrm{P}_{\mathrm{RB}}, \forall m \in \{1, 2, \ldots \mathrm{M}\} \\ & \mathrm{C}_3 : \sum_{m=1}^{\mathrm{M}} \mathrm{I}_{m,k}(t) = 1, \forall k \in \{1, 2, \ldots, \mathrm{K}\}. \end{aligned} \tag{14}$$

Both α and β are control parameters of the QoS function. The main objective is to maximise the QoS function.

TABLE 1
BBU Operations and Their Scaling Values with Transmit Antennas and Bandwidth

Processing type, i	GOPS	$\mathrm{P}^{\mathrm{ref}}_{i,\mathrm{BBU}}$ [W]	x^A_i	x^w_i
Time Domain Processing	360	9.0	1	0
Frequency Domain Processing	60	1.5	2	1
Forward Error Correction	60	1.5	1	0
Central Processing Unit	400	10.0	1	0
Common Public Radio Interface	300	7.5	1	0
Leakage	118	3.0	1	0

6 Power Model for C-RAN

This section explains the necessary aspects needed to assess the power consumption of C-RAN. However, a more detailed description of the components involved in a C-RAN power model is given in [24]. The three most important parts considered for the power model are described as follows.

6.1 BBU Power Estimation Model

The BBU performs a different set of functions (I_{BB}) which includes scheduling of PRBs, Forward error correction, FFT and OFDM specific processing, filtering, modulation/demodulation, and transport link related functions, etc. These features can be measured in Giga Operations per Second (GOPS) and then translated into power figures. About 40 GOPS per Watt is estimated as the power cost of a large BBU [25]. The power model for the BBU can be given as

$$\mathrm{P}_{\mathrm{BBU}} = \sum_{i\in I_{\mathrm{BB}}} \mathrm{P}^{\mathrm{ref}}_{i,\mathrm{BBU}} \mathrm{A}^{x^A_i} \mathrm{W}^{x^w_i}, \tag{15}$$

where $\mathrm{P}^{\mathrm{ref}}_{i,BBU}$ in Watts represents the power consumption of BBU with respect to BBU functions. A is the number of antenna chains/RF transceivers with x^A_i scaling exponent. W is the bandwidth share used in transmission with a scaling exponent x^w_i. In [26], the authors model BBU operations with exact scaling components and reference values to calculate BBU power consumption, shown in Table 1.

6.2 RRH Power Estimation Model

An RRH consist antenna chains/ RF transceivers, each with its own power amplifier (PA). The PA is main element of

consideration as it consumes most of the power within an RRH. The power consumption of a PA is affected by its power efficiency (η_{PA}). The power consumed by the PA can be given as $\mathrm{P_{PA}} = \frac{\mathrm{P_{TX}}}{\eta_{\mathrm{PA}}(\sigma_{\mathrm{feed}})}$, where $\mathrm{P_{TX}}$ is the output power of the PA, which depends on the bandwidth share (χ), i.e., the actual number of physical resource blocks ($\mathrm{N_{RB}}$) used for transmission and the output power of the antenna $\mathrm{P_{out}}$ ($\mathrm{P_{TX}} = \mathrm{P_{out}}\chi$). σ_{feed} represents the feeder loss. Moreover, the RF transceiver units of an RRH are responsible for functions like signal modulation/demodulation, voltage controlled oscillation and mixing, AC-DC and DC-AC conversions, and low noise, gain amplification. The power consumed by an RRH can be modelled as

$$\mathrm{P_{RRH}} = \sum_{a=1}^{A} (\mathrm{P_{PA}} + \mathrm{P_{RF}}), \tag{16}$$

where $a \in \{1, \ldots, A\}$ denotes the number of antenna/RF chains.

6.3 Optical Transceiver Power Estimation Model

In C-RAN architecture, the *front-haul* connectivity with high bandwidth, low cost, and low latency requirements for transport networks is challenging. Several factors influence the operation of optical transceivers such as the technology used, the operating conditions, and the output power required, which in turn affect the power consumption. From a power consumption perspective, the optical transceivers can be divided into two modules. The optical transmitter module, in which the OFDM electrical signals are modulated over optical carriers using an external or direct modulated lasers. And a receiver module which detects the optical OFDM signals either by direct detection or coherent detection. The power consumption of the optical transceiver as described in [27] can be given as

$$\mathrm{P_{TRANS}} = \left(\mathrm{P_{laser}} + \mathrm{P_{driver}} + \mathrm{P_{I/O}}\right)_{\mathrm{TX}} + \left(\mathrm{P_{PD}} + \mathrm{P_{amp}} + \mathrm{P_{I/O}}\right)_{\mathrm{RX}}, \tag{17}$$

where $\mathrm{P_{laser}}, \mathrm{P_{driver}}, \mathrm{P_{I/O}}, \mathrm{P_{PD}}$, and $\mathrm{P_{amp}}$ are the powers consumed by direct-modulated laser, electronics driving the laser, the electrical input/output interface, photodetector, and the trans-impedance and limiting amplifiers. This paper consider a point-to-point (PtP) transceivers rather than point-to-multipoint, because the PtP link loss is driven by distance and used operating wavelength only, i.e., the link loss of PtP is as low as 6 dB with a 20 km network reach [28].

The total power consumption of a C-RAN ($\mathrm{P_{C\text{-}RAN}}$) can be estimated by summing the power consumed by three main parts of the network along with power consumed by other components $\mathrm{P_{OTHER}}$ such as power conversions (AC-DC, DC-DC) and cooling, i.e.,

$$\mathrm{P_{C\text{-}RAN}} = \sum(\mathrm{P_{BBU}} + \mathrm{P_{TRANS_B}}) + \sum(\mathrm{P_{RRH}} + \mathrm{P_{TRANS_R}}) + \mathrm{P_{other}}. \tag{18}$$

Where $\mathrm{P_{TRANS_B}}$ and $\mathrm{P_{TRANS_R}}$ indicates the power consumption of PtP transceivers located at each BBU and RRH, respectively. According to [29], base stations with a total power consumption ≤ 500 Watts do not require a cooling system. This can be applied to RRH in C-RAN if its components

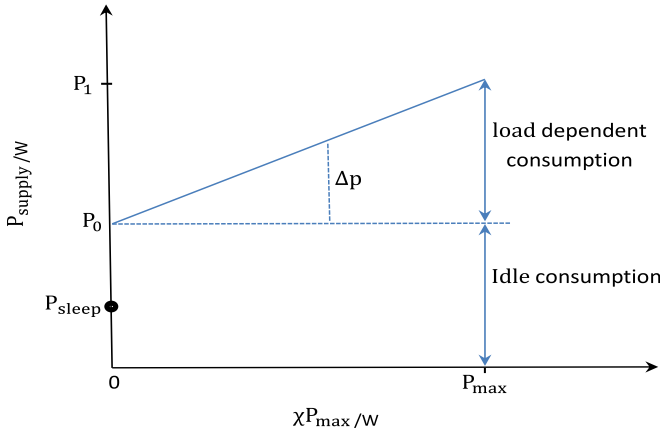


Fig. 3. Structure of a cloud radio access network represented as SON.

(i.e., PA, RF, and optical transceiver) require an overall power less than 500 Watts. In this paper, the cooling power for RRH is ignored considering supply power as the only overhead.

From [24], the supply power required for a base station can be estimated as an affine function of transmitting power. The power consumption can be expressed by a load-dependant part that linearly increases with a power gradient (slope) $\Delta\mathrm{p}$ and a static load independent part $\mathrm{P_{static}}$ as shown in Fig. 3. Moreover, the supply power reaches a maximum $\mathrm{P_1}$ when the transmitting power reaches the maximum limit $\mathrm{P_{max}}$. A base station may enter an idle mode (sleep mode), with minimum power consumption ($\mathrm{P_{sleep}}$) when it is not transmitting. The total supply power for a base station can be formulated as

$$\mathrm{P_{supply}}(\chi) = \begin{cases} \mathrm{P_1} + \Delta\mathrm{p}\ \mathrm{P_{max}}(\chi - 1) & \text{if } 0 < \chi \leq 1 \\ \mathrm{P_{sleep}} & \text{if } \chi = 0, \end{cases} \tag{19}$$

where $\mathrm{P_1} = \mathrm{P_{static}} + \Delta\mathrm{p}\mathrm{P_{max}}$. χ is a scaling parameter which indicates the bandwidth share, i.e., $\chi = 1$ indicates that the system is transmitting with full power and bandwidth whereas $\chi = 0$ represents an idle system. The basic power model presented in (19) is parameterised to understand the contribution of different parameters. Parameters which are assumed to be constant or having negligible effects are also highlighted. The following approximations are made:

- Both the BBU and Radio Frequency (RF) power consumption, linearly scales with the number of Antennas (A) and bandwidth (W), i.e., $\mathrm{P_{BBU}} = \mathrm{A}(\frac{\mathrm{W}}{\mathrm{BW_{Total}}})\mathrm{P_{BBU}^{pm}}$ and $\mathrm{P_{RF}} = \mathrm{A}(\frac{\mathrm{W}}{\mathrm{BW_{TOTAL}}})\mathrm{P_{RF}^{pm}}$. Where $\mathrm{P_{BBU}^{pm}}$ and $\mathrm{P_{RF}^{pm}}$ are parameterised power consumption of BBU and RF, respectively.
- Each antenna unit of an RRH has a power amplifier. The power consumed by a PA depends on the maximum power transmission per antenna unit ($\frac{\mathrm{P_{max}}}{A}$) and its efficiency ($\eta_{\mathrm{PA}}$). Losses between the antenna and PA are known as feeder losses (σ_{feed}) which may be ignored since PAs are placed close to the antennas [30].
- The loss factors of DC-DC, AC-DC conversions, main supply units (MS), and cooling power consumption for the BBU pool are approximated by $\sigma_{\mathrm{DC,POOL}}$, $\sigma_{\mathrm{MS,POOL}}$, and $\sigma_{\mathrm{COOL,POOL}}$. Whereas for the RRHs, the loss factors are approximated by $\sigma_{\mathrm{DC,RRH}}$ and $\sigma_{\mathrm{MS,RRH}}$. Moreover, the optical fibre losses

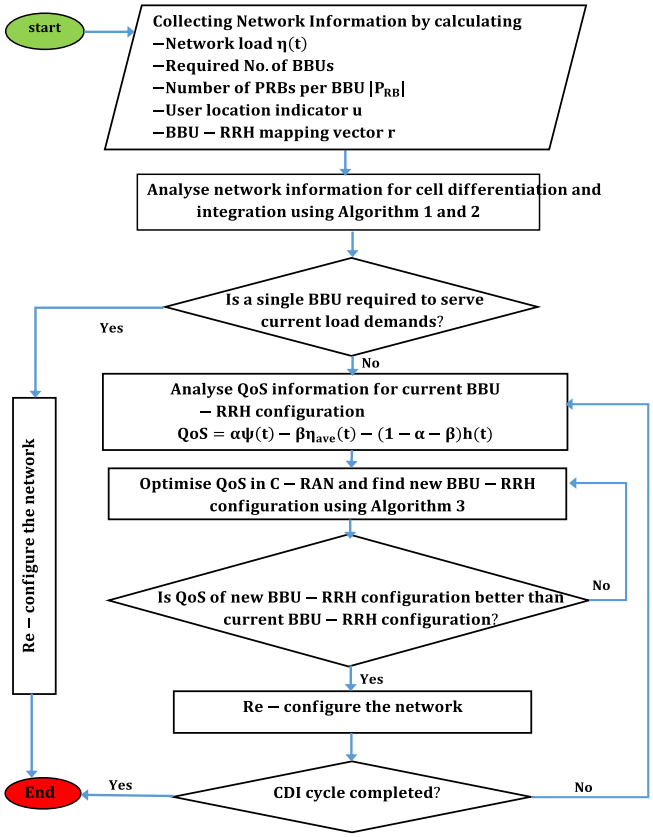


Fig. 4. Block diagram of CDI algorithm for one CDI cycle.

between BBUs and RRHs are approximated by a loss factor σ_{optical}.

- Power consumption of the optical transceivers linearly scales with the number of BBUs and RRHs.

If the power consumed by a single BBU serving a single RRH is

$$\mathrm{P}_1 = \mathrm{P}_{\mathrm{BBU}} + \mathrm{P}_{\mathrm{RRH}} \tag{20}$$

$$\mathrm{P}_1 = \frac{\mathrm{A}(\frac{\mathrm{W}}{\mathrm{BW}_{\mathrm{TOTAL}}})\mathrm{P}^{\mathrm{pm}}_{\mathrm{BBU}} + \mathrm{P}_{\mathrm{TRANS_B}}}{(1-\sigma_{\mathrm{DC,POOL}})(1-\sigma_{\mathrm{MS,POOL}})(1-\sigma_{\mathrm{COOL,POOL}})} + \frac{\mathrm{A}(\frac{\mathrm{W}}{\mathrm{BW}_{\mathrm{TOTAL}}})\mathrm{P}^{\mathrm{pm}}_{\mathrm{RF}} + (\mathrm{P}_{\max}/\mathrm{A}.\eta_{\mathrm{PA}}) + \mathrm{P}_{\mathrm{TRANS_R}}}{(1-\sigma_{\mathrm{DC,R}})(1-\sigma_{\mathrm{MS,R}})(1-\sigma_{\mathrm{optical}})}. \tag{21}$$

Then the total power consumed by all active BBUs and RRHs in a C-RAN network can be modelled as

$$\mathrm{P}_{\mathrm{supply}} = \sum_{m=1}^{\mathrm{M}} \left(\mathrm{P}_{\mathrm{BBU}} + \sum_{n \in Z_m} \mathrm{P}_{\mathrm{RRH}} \right), \tag{22}$$

where M represents the number of active BBUs in the network and Z_m represents the list of RRHs handled by BBU_m

7 Cell Differentiation and Integration (CDI) Algorithm

According to the intuitive analysis above, a CDI algorithm is proposed in this section and Fig. 4. Network information is collected in the first step and analysed for proper cell differentiation and integration. The algorithm seeks to utilise the network resources efficiently by calculating the necessary number of BBUs and RRHs to serve capacity demands at the end of each CDI cycle. Apart from a single BBU required to serve load requirements, proper BBU-RRH configuration is adjusted at the end of optimisation step by comparing the analysed and optimised QoS values. Note that the QoS metrics can be different depending on load intensity and the number of active BBUs and RRHs in the network. For the optimisation part of the algorithm, a Discrete Particle Swarm Optimisation (DPSO) algorithm is developed as an Evolutionary Algorithm to solve the BBU-RRH configuration problem and is explained in the next section. The optimisation process continues until the CDI cycle is completed. Note that, the CDI algorithm shown in Fig. 4 is triggered at the beginning of each CDI cycle.

The pseudo-codes for semi-static cell differentiation and integration are given in Algorithms 1 and 2, respectively. An important consideration is the first association of RRHs to the required number of BBUs during cell differentiation and integration, before identifying a proper BBU-RRH mapping in the optimisation phase. Algorithms 4 and 5 are supporting algorithms for Algorithms 1 and 2 which covers all possible cases of initial BBU-RRH assignment during differentiation or integration of cells along with cases where the number of BBUs are increased, decreased or remain unchanged. The initial BBU-RRH mapping is necessary for utilising the available BBU resources in an efficient manner so as to prevent high blocking rate. The blocking rate of the network at time t can be measured as

$$\text{Blocking rate} = \left[1 - \frac{\sum_{m=1}^{\mathrm{M}} \sum_{k=1}^{\mathrm{K}} \mathrm{I}_{m,k}(t)}{\mathrm{K}} \right] \times 100, \tag{23}$$

where $\mathrm{I}_{m,k}(t)$ as discussed earlier, is a binary indicator such that $\mathrm{I}_{m,k} = 1$, if user k is served by BBU_m at time t. Note that, users are served based on the choice of scheduler used by a BBU. Moreover, the amount of resource shortage (or PRB shortage) in the network based on users PRB demand can be estimated as follows

$$\text{Resource Shortage} = \sum_{m=1}^{\mathrm{M}} \max\Big[(\eta_m(t) - 1), 0\Big] \times 100. \tag{24}$$

Note that, the CDI algorithm triggers Algorithms 1 and 2 sequentially, i.e., Algorithm 2 is triggered immediately after the Algorithm 1 is executed. In the interest of simplicity and understanding, the CDI algorithm is divided into separate pseudo-codes.

7.1 Discrete Particle Swarm Optimisation (DPSO)

PSO utilises a population (or swarm) of particles, where each individual particle represents a solution [31], namely BBU-RRH association vector $\mathbf{r}$ defining the BBU-RRH configuration. As the QoS represented in (14) is considered as the main objective function, PSO seeks to maximise the QoS function by finding the best solution vector $\{\mathrm{r}_{11}, \mathrm{r}_{12}, \ldots, \mathrm{r}_{in}\}$. PSO utilises a group of particles (or solutions) to probe the solution space in a random way with different velocities. To direct the particles to their best fitness values, the velocity of each particle is changed stochastically at each iteration. The velocity update of each particle j depends on the historical

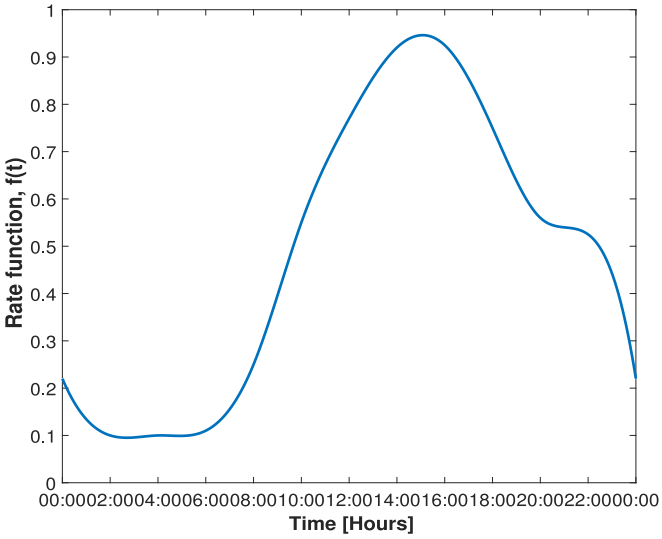


Fig. 5. Rate function for time in-homogeneous user arrivals.

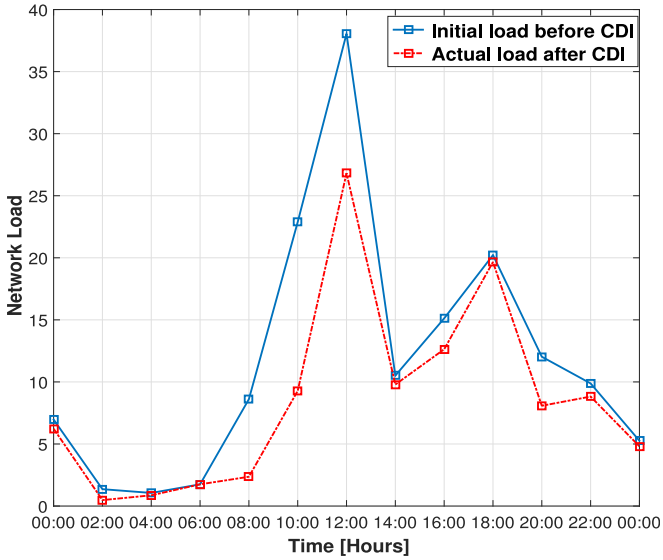


Fig. 6. Actual network load with respect to time.

best position experience (pbest) of the particle itself and the best location experience of neighbouring particles, i.e., the global best position (gbest) [32] and is given as

$$\mathrm{v}_j^{\mathrm{I}} = w\mathrm{v}_j^{\mathrm{I}-1} + c_1\varepsilon_1\left(\mathrm{pbest}_j^{\mathrm{I}} - x_j^{\mathrm{I}}\right) + c_2\varepsilon_2(\mathrm{gbest}_j^{\mathrm{I}} - x_j^{\mathrm{I}}) \quad 1 \leq j \leq |\Delta|, \tag{25}$$

where $|\Delta|$ represents the population (or swarm) of particles and I represents the Iteration number. x_j^{I} is the current position of particle j in iteration I and $\varepsilon_1, \varepsilon_2$ are random numbers between 0 and 1. Both c_1 and c_2 are acceleration constants that pulls the particle towards best position. Values in the range 0-5 are chosen for c_1 and c_2. The inertial weight w represents the effect of preceding velocity on the updated velocity. Choosing an optimum value for w can assist a balanced proportion between global and local exploration of the search space. Usually values between 0-1 are selected for w [33]. A value of 0.9 for w is selected in this paper. The new position of particle j for the next iteration $\mathrm{I}+1$ will be

$$x_j^{\mathrm{I}+1} = x_j^{\mathrm{I}} + \mathrm{v}_j^{\mathrm{I}}. \tag{26}$$

PSO terminates if a stopping criterion is satisfied, e.g., after reaching a predefined number of iterations I_{max}. Since the solution vector, $\mathbf{r}$ (or particle) should be real-valued, the standard PSO algorithm can not be applied to solving this discrete optimisation problem. In this paper, a Discrete Particle Swarm Optimisation is developed to solve the QoS maximisation problem defined in (14). The pseudo code of DPSO is given in Algorithm 3

8 Computational Results and Analysis

To make the simulation more realistic, the user arrivals in Fig. 6 follows a Poisson process with rate λ. However, due to the dynamic spatial and temporal nature of user traffic, the user arrival is modelled as a time-inhomogeneous process. This is achieved by multiplying the time-homogeneous Poisson process with traffic intensity parameter λ and the rate function $f(t)$ shown in Fig. 5. The rate function is unitless and reshapes the traffic from constant intensity to an analogous time varying profile that reflects typical traffic patterns in a real cellular network. If users arrive in the system following a Poisson process with intensity λ users/min, with a constant service time of h (60 sec), then the number of users at time t is calculated as $\mathrm{K}(t) = \chi h f(t)$. Where $\chi \sim \mathrm{Poiss}(\lambda)$ is a random variable with mean λ (i.e., $\lambda = 200$). Moreover, different data rate requirements are assumed for end users based on 3GPP standard simulation parameters [34], i.e., 4-25 kbps for audio, 32-384 kbps for video, 28.8 kbps for data, and 60 kbps for real-time gaming services. Based on uniform user distribution and network load shown in Fig. 6, an actual number of active BBUs and RRHs with respect to time is shown in Fig. 7.

The BBU-RRH association vector $\mathbf{r} = \{\mathrm{r}_{11}, \mathrm{r}_{12}, \ldots, \mathrm{r}_{in}\}$ is maintained and updated after each CDI cycle. Newly activated RRHs and BBUs in the network are mapped according to Algorithms 4 and 5. In this paper, a maximum of 49 RRHs and 5 BBUs are deployed in the network to support semi-static cell differentiation and integration. The initial BBU-RRH mapping at the beginning of a CDI cycle might degrade the network QoS. Therefore, dynamic BBU-RRH mapping is proposed to identify proper BBU-RRH mapping.

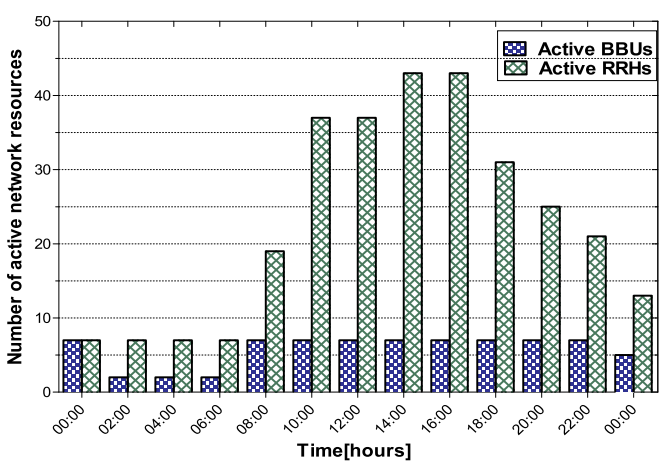


Fig. 7. Number of active BBUs and RRHs with respect to network load/time.

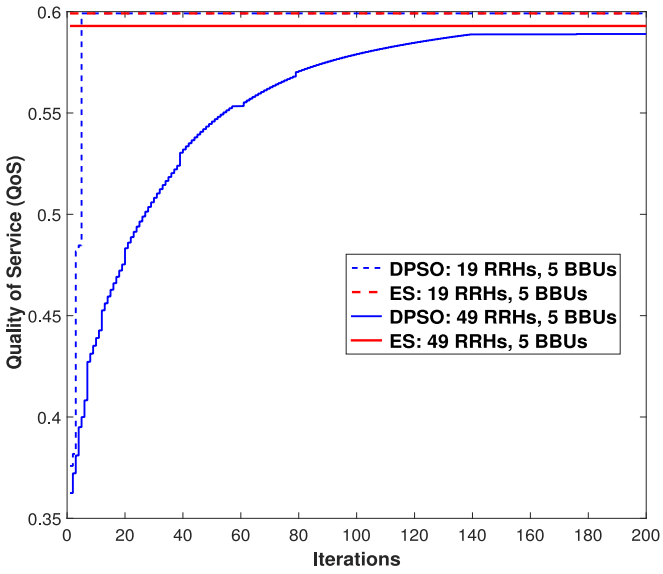


Fig. 8. QoS values for DPSO and ES.

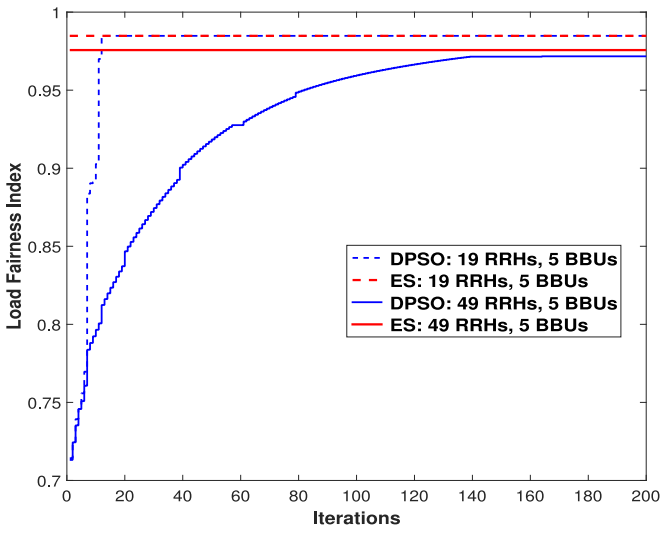


Fig. 9. Load fairness index values for DPSO and ES.

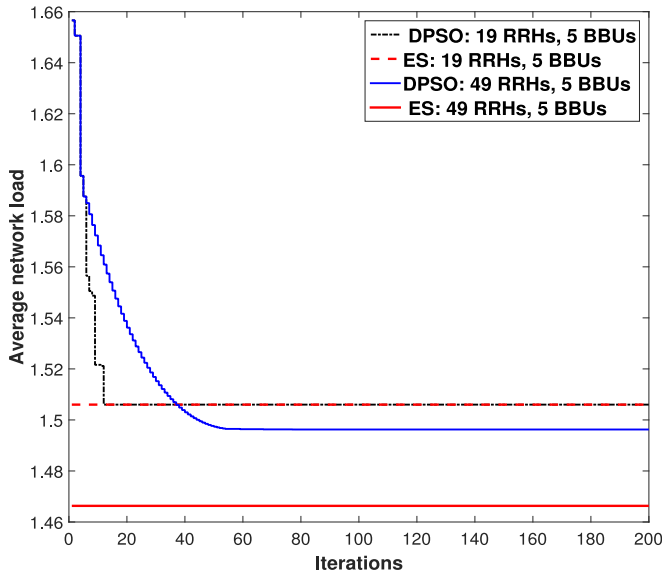


Fig. 10. Average network load for DPSO and ES.

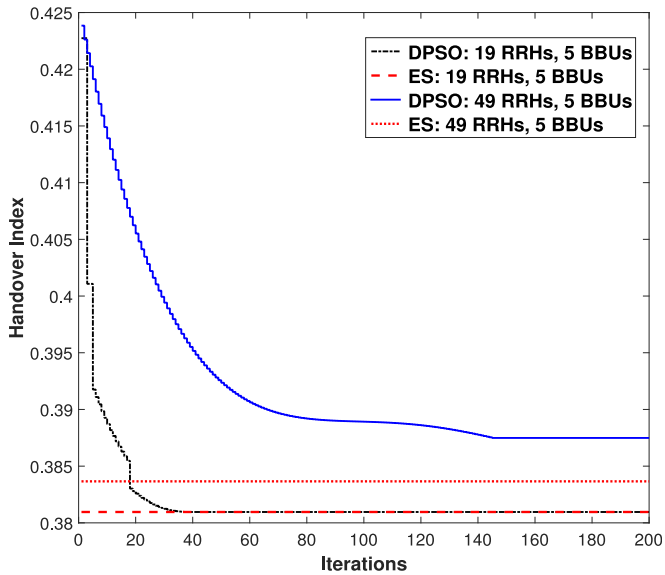


Fig. 11. Average handovers for DPSO and ES.

Before going to a more thorough analysis of the proposed CDI concept, the efficiency of DPSO over two different problem scenarios, $\mathrm{P}_1, \mathrm{P}_2$, and compared with Exhaustive Search (ES) algorithm. Both scenarios consists of 5 active BBUs with 19 active RRHs including two differentiated cells (Tier 1, level 2, RRH structure) for P_1, and 49 active RRHs (Tier 1, level 7, RRH structure) for P_2, respectively. The aim is to analyse DPSO performance for small and large networks. User distribution within each cell is uniform where 6 and 25 users are considered for non-dense and high dense cells, respectively.

For Monte Carlo analysis, the DPSO and ES algorithms are repeated 50 times with different initial BBU-RRH settings for each problem and results obtained are averaged. The load fairness index, averaged network load, and handover index are represented in Figs. 9, 10, and 11, respectively, over 200 iterations for both P_1 and P_2. The optimum values shown in the figures and Table 2, are achieved by exhaustively searching for all possible solutions N^{M}) using ES algorithm, which helps in demonstrating the improvement in each iteration of the DPSO algorithm. Note that, ES algorithm is independent of iterations.

TABLE 2
Computational Results for DPSO and ES

		P_1 (19 RRH)	P_2 (49 RRH)
Quality of Service	DPSO	0.599142	0.588970793
	ES	0.599142	0.592940793
Load Fairness Index	DPSO	0.984797	0.97168
	ES	0.984797	0.97556
Average Network Load	DPSO	1.506	1.4962
	ES	1.506	1.4663
Handover Index	DPSO	0.38095	0.38748
	ES	0.38095	0.383659

Fig. 8 shows that the DPSO algorithm converges to the optimum solution in P_1 with a Convergence Rate (CR) of 0.825. Where CR is defined as the number of times, the

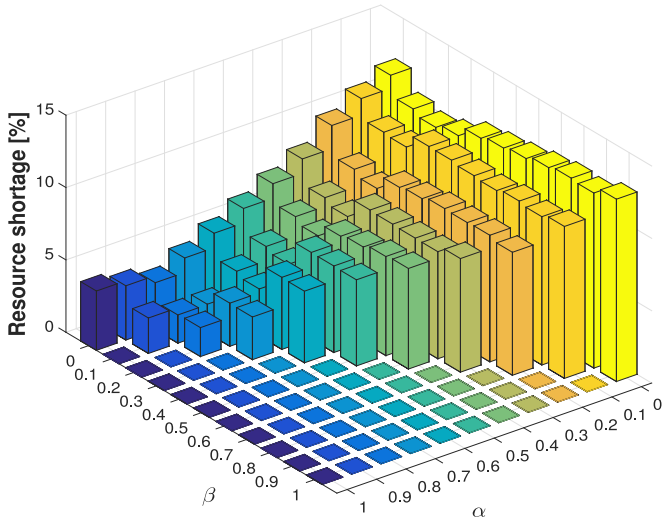


Fig. 12. Resource shortage for different α and β.

DPSO finds a best or optimum solution during the entire number of iterations. This implies that over 200 iterations, the optimum solution is achieved 165 times for P_1. For P_2, the CR of DPSO algorithm is 0.12. However, the optimum solution is not reached over 200 generations. DPSO algorithm achieves the best value 24 times, i.e., after 176 iterations and 176×|Δ| fitness evaluations, which is still 99.53 percent of the optimum value achieved by ES algorithm after an enormous 5^{49} (M^N) fitness evaluations.

Fig. 9 shows that the DPSO algorithm converges to the optimum load fairness index value after 13^{th} iteration in P_1. However, in P_2, the optimum value can not be found over 200 iterations, and the best load fairness index value is achieved after only 176 iterations and 176×|Δ| fitness evaluations, which is 99.57 percent of the optimum value found by ES algorithm. ES algorithm performs 5^{49} fitness evaluations to find the optimum value which is a considerable amount of fitness evaluations.

Figs. 10 and 11 displays the convergence of DPSO algorithm to the optimum value for average load value and handover index in both P_1 and P_2. In P_1, optimum are achieved after 12 and 38 iterations for average network load and handovers, respectively. For P_2, the DPSO algorithm could not find the optimum value over 200 iterations. However, the best possible value achieved for average network load and handover index are 98 and 99.01 percent of the optimum value found by ES algorithm, respectively. ES algorithm determines the optimum value after performing 5^{49} enormous fitness evaluations whereas the DPSO algorithm performs 67×|Δ| and 145×|Δ| to find the best value for average network load and handover index, respectively. Note that, the α and β control parameters in (14) are selected by performing an exhaustive search algorithm to identify the optimal BBU-RRH setting for P_1. Both α and β values are orderly set to 0, 0.1, ... , 1 with a constraint $\alpha + \beta \leq 1$ as shown in Fig. 12. An optimal BBU-RRH setting is found using ES algorithm for each pair of α and β. It is observed that setting a higher value for load fairness index (until $\alpha = 0.8$) not only reduces the resource shortage but also improves network balance. Setting values for $\alpha > 0.8$ results into improper BBU-RRH mapping which implies that maximising network load balance is overly considered compared to minimising average network load and handovers, resulting into an increased resource shortage. This paper considers $\alpha = 0.8$ and $\beta = 0.1$ which means assigning a 10 percent weight to handover minimisation.

Algorithm 1. Pseudo-Code for Semi-Static Cell Differentiation

Input: Current network load $\eta(t)$ from (7)
BBU-RRH mapping vector **r**
Required number of BBUs from (8)

if *No. of active BBUs* = 1 **then**
  if $\eta(t) \geq |P_{RB}|$ **then**
    -Activate required No. of BBUs
    -Differentiate cell into tier-2 RRH structure by BBU-RRH mapping using Algorithm 4
    -Update BBU-RRH mapping vector **r**
    for *i=1 to C* **do**
      -Select set S_i
      -Compute $\eta_{\text{RRH}_{i1}}(t)$ from (9)
      if $\eta_{\text{RRH}_{i1}}(t) > |P_{RB}|$ **then**
        -R← S_i {Add S_i to R}
        -Diffentiate cell C_i by activating all RRHs in S_i an map them to BBUs according to Algorithm 4.
        -Update BBU-RRH mapping vector **r**.
      end
    end
  else
    -No cell differentiation required.
    -Tier-3 RRH structure remains.
  end
else
  if *No. of active BBUs* ≤ *No. of required BBUs* **then**
    if *All possible RRHs deployed in the network are active* **then**
      -Activate the required No. of BBUs.
      -Cells can not be differentiated further.
      -Update BBU-RRH mapping vector **r**
    else
      -Activate required number of BBUs
      for *i= 1 to C* **do**
        -Select set S_i
        -Compute $\eta_{\text{RRH}_{i1}}(t)$ from (9)
        if $\eta_{\text{RRH}_{i1}}(t) > |P_{RB}|$ **then**
          - R← S_i {Add S_i to List R}
          -Differentiate cell C_i further to tier-1 RRH structure by mapping newly activated RRHs to active BBUs using Algorithm 4
          -Update BBU-RRH mapping vector **r**
        end
      end
    end
  end
end

For a more thorough analysis, the proposed CDI concept is compared to a fixed C-RAN scenario (F-CRAN). The BBU cloud holds five BBUs in both cases. However, the fixed C-RAN scenario does not support cell differentiation or integration, and only 7 RRHs serves the entire macrocell coverage area. The dynamic BBU-RRH mapping is enabled in the

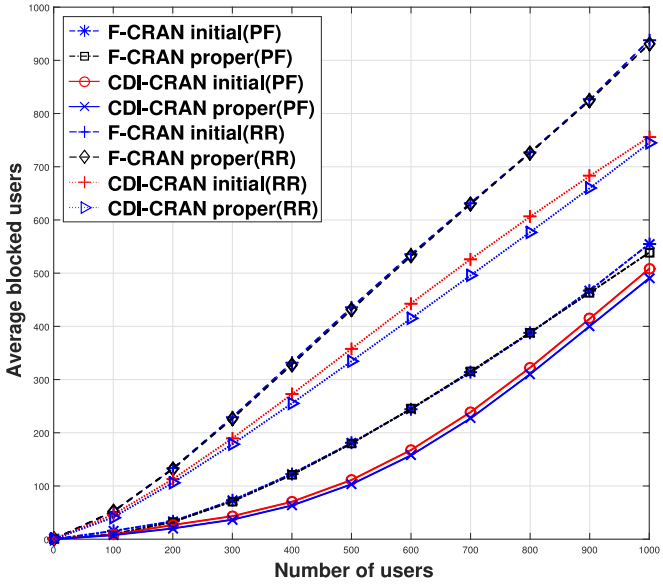


Fig. 13. Average blocked users for fixed and CDI-enabled C-RAN.

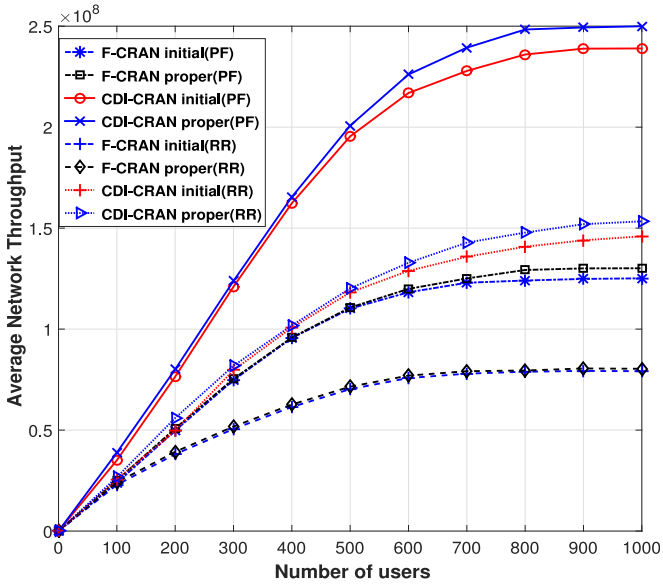


Fig. 14. Average network throughput for fixed and CDI-enabled C-RAN.

fixed C-RAN scenario which shows 5^7 possible BBU-RRH mapping solutions to choose from at the beginning of each CDI cycle. The number of possible BBU-RRH mapping solutions for CDI scenario at the start of each CDI cycle is M^{N}, where M and N represents the number of active BBUs and RRHs, respectively. Moreover, an increasing user arrival is considered in the network with random data rates requirement as explained earlier. However, a Monte-Carlo analysis is performed, where 100 uniformly distributed users are envisaged for each instance, and the average of all distributions are taken into account regarding network load, throughput, blocked users, and resource shortage analysis. Figs. 13 and 14 shows the relative performances regarding average blocked users and average network throughput with Proportional Fair (PF) and Round Robin (RR) scheduling techniques. Since the CDI algorithm includes 2 phases of BBU-RRH mapping, i.e., the initial BBU-RRH assignment during cell integration/differentiation and the optimum BBU-RRH setting achieved by DPSO in the second step, the results of both phases are analysed.

The simulation results demonstrate the advantage of using CDI-enabled C-RAN (CDI-CRAN) instead of a F-CRAN setting. When a fixed C-RAN is considered, the average blocked users in the network are much higher with significantly lower average throughput, using any scheduling technique, as shown in Figs. 13 and 14, provided that the dynamic BBU-RRH mapping is also enabled. However, an interesting observation is the significant drop in the averaged blocked users and the necessary increase in average network throughput in CDI-CRAN compared to F-CRAN. This indicates that during cell differentiation, an overloaded cell divides into multiple smalls cells, and not only reduces the user to RRH distances but also the PRB demands resulting from high SINR and low path loss values. A further decrease in average network load is observed after proper BBU-RRH mapping, providing a balanced network load across the active BBUs. Note that, cell differentiation increases the number of RRH interferers in the network. However, RRHs served by the same BBU does not contribute to the overall interference experienced by users served by the same BBU.

Algorithm 2. Pseudo-Code for Semi-Static Cell Integration

Input: Current network load $\eta(t)$ from (7)
BBU-RRH mapping vector **r**
Required number of BBUs from (8)
if *No. of active BBUs =1* **then**
  -No cell integration required.
  -A high-power BS serves the geographical area.
else
  if *No.of required BBUs=1* **then**
    -Integrate all cells into tier-3 RRH structure, i.e., a high power BS should serve the geographical area.
    -Switch-off remaining BBUs.
    -Update BBU-RRH mapping vector **r**.
  else
    for *i=1 to C* **do**
      -Select set S_i
      for *j=1 to end of* S_i **do**
        -Compute load $\eta_{\mathrm{RRH}_{ij}}(t)$ from (9)
        -Sum=Sum+$\eta_{\mathrm{RRH}_{ij}}(t)$
      end
      if $Sum \leq \mathrm{P_{RB}}$ **then**
        -Integrate all cells by switching-off all RRHs in set S_i except RRH_{i1}.
        -Offload RRHs to required number of BBUs according to Algorithm 5.
        -Update BBU-RRH mapping vector **r**.
      end
    end
    -Run Algorithm 5
    {Case of BBU reduction and no integration}
  end
end

From the results shown in Fig. 14, it is observed that the average network throughput increases by 45.53 percent in the CDI-CRAN compared to F-CRAN, both

TABLE 3
Comparison Results for Fixed and CDI-Enabled C-RAN

	Blocking Rate[%]				Resource Shortage	
	Initial		Proper		Initial	Proper
	PF	RR	PF	RR		
F-CRAN	35.99	81.66	35.34	81.10	16.44×10^3	16.42×10^3
CDI-CRAN	26.87	67.33	25.809	62.13	38.79×10^2	38.47×10^2

enabled with PF schedulers. Whereas with RR schedulers, an increase of 42.102 percent is observed. Moreover, the average throughput difference between initial and optimum BBU-RRH mapping in a CDI-enabled C-RAN, with PF and RR scheduling is 4.0219 and 4.126 percent, respectively. This indicates efficient resource utilisation during cell differentiation and integration, ensuring minimum blocked users until a proper BBU-RRH setting is identified. Note that, the initial BBU-RRH mapping supported by Algorithms 4 and 5, is an important consideration in the overall CDI concept (as explained earlier). About 23.149 percent reduction in the average number of blocked users with PF scheduler and 20.903 percent with RR is observed in Fig. 13. Moreover, the average resource shortage drastically decreases in the CDI-CRAN, compared to fixed C-RAN as shown in Table 3 provided that both scenarios have an equal amount of resources available (i.e., 5 BBUs, 5×100 PRBs). A 76.57 percent decrease in average PRB shortage is estimated with CDI-CRAN compared to F-CRAN.

Fig. 15 shows the average power consumed by the C-RAN network for both CDI and fixed setting for different schedulers. Despite the fact that the geographical area is served with more RRHs in CDI-CRAN, the total power consumed by the network is still lower. An average decrease of $\approx 15.28\%$ and $\approx 16.02\%$ in the average power consumption is estimated in the CDI-CRAN with PF and RR schedulers, respectively, compared to a fixed C-RAN setting. Fig. 16 shows the SINR thresholds versus the probability of coverage results. The CDI activated C-RAN performs well compared to fixed C-RAN regarding coverage performances for 1000 users, provided that no interference mitigation techniques are applied in this work. Note that, the path-loss models for different tiers of cell differentiation are not identical,

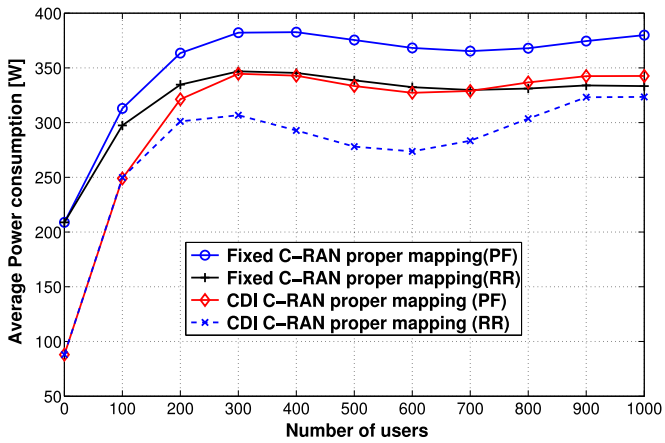


Fig. 15. Average power consumed by fixed and CDI-enabled C-RAN.

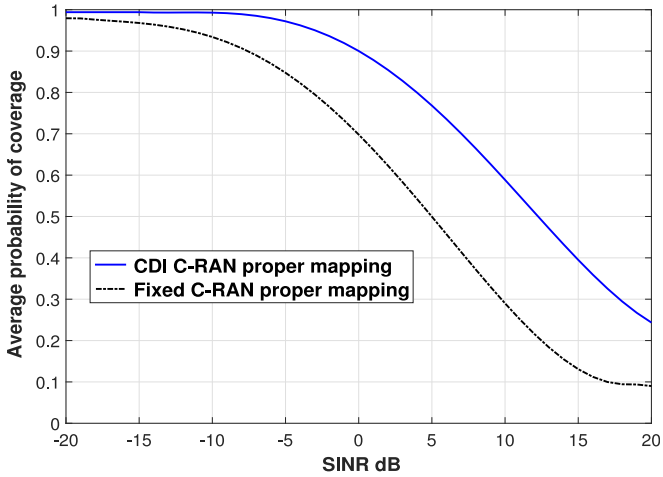


Fig. 16. Probability of coverage versus SINR threshold.

the probability of coverage still has a linear behaviour for SINR thresholds within the range of 0 dB to 5 dB for both cases.

Algorithm 3. Discrete Particle Swarm Optimisation (DPSO) Algorithm

-I=0;
-Generate initial Swarm with random position and velocity $(r_1^I, r_2^I, \ldots, r_{|\Delta|}^I)$.
for $i = 1$ *to* $|\Delta|$ **do**
  {Initialise best positions of each particle}
  $\text{pbest}_i = \text{r}_i^I$
end
-QoS=0;
-{Find QoS of each particle using (14), i.e., }
for *i=1 to* $|\Delta|$ **do**
  -Select r_i^I from pbest_i
  if $f(r_i^I) >$ QoS **then**
    QoS=$f(r_i^I)$
  end
end
gbest=QoS;
-Update positions and velocities of all particles in the swarm using equations (25) and (26)
while $\text{I} < I_{max}$ **do**
  for *i=1 to* $|\Delta|$ **do**
    -Select r_i^I from the swarm
    F_1=$f(r_i^I)$
    -Select r_i^I from pbest
    $F_2 = f(r_i^I)$
    if $F_1 \geq F_2$ **then**
      $\text{pbest}_i = r_i^I$
      QoS=F_1
      if *QoS > gbest* **then**
        gbest=QoS;
        $x = i$;
      end
    end
    -Update positions and velocities of all particles in the swarm using equations (25) and (26)
  end
  I=I+1;
end

Algorithm 4. Initial RRH Association to Active BBUs During Cell Differentiation

Input: List A of newly activated BBUs
List R containing sets of RRHs supporting cell differentiation

if *A is not empty* **then**
  for *m=1 to No. of active BBUs* **do**
    -Compute $\eta_m(t)$ from (6)
    if $\eta_m(t) \leq$ lower limit **then**
      A $\leftarrow$ BBU$_m$ {Add BBU$_m$ to List A}
    end
  end
  I=1;
  while *not the end of List R* **do**
    -Select Ith set from list R
    $m = 1$;
    for *j=1 to end of set* S_i **do**
      if m $>$ $|A|$ **then**
        $m = 1$
      end
      $BBU_m \leftarrow RRH_{ij}$ {Map RRH_{ij} to BBU_m except R_{1j}}
      $m = m + 1$;
    end
    I=I+1;
  end
else
  for *m=1 to No. of active BBUs* **do**
    -Compute $\eta_m(t)$ from (6)
    if *lower limit* $\leq \eta_m(t) \leq$ *Upper limit* **then**
      A$\leftarrow$ BBU$_m$ {Add BBU$_m$ to A}
    end
  end
  if *A is still empty* **then**
    A $\leftarrow$ All active BBUs
  end
  -Sort A in increasing order of BBU loads
  I=1;
  while *not the end of List R* **do**
    -Select Ith set from List R
    $m = 1$;
    for *j=1 to end of set* S_i **do**
      if m $>$ $|A|$ **then**
        m=1;
      end
      BBU$_m \leftarrow$ RRH$_{ij}$ {Map RRH$_{ij}$ to BBU$_m$ except RRH$_{1j}$}
    end
    I=I+1;
  end
end

9 Conclusion

The concept of cell differentiation and integration in C-RAN is examined with an objective to utilise network resources efficiently without degrading the overall network QoS. An energy efficient C-RAN network is considered to accommodate traffic by scaling the BBUs and RRHs as well as maintaining a balanced network via proper BBU-RRH mapping, formulated as a constrained integer programming problem. A CDI algorithm is developed for C-RAN and tested for comparison with a fixed C-RAN setting. Computational results based on Monte Carlo analysis shows an average throughput increase of 45.53 with 23.149 percent decrease in average blocked users and 76.57 percent reduction in average resource (PRBs) shortage. The CDI algorithm hosts a DPSO algorithm which is developed to find optimum BBU-RRH configuration dynamically. The performance of DPSO is tested and compared to ES. Using two benchmark problems, the DPSO delivered noticeably faster convergence compared to ES, which makes the CDI algorithm more reliable for a self-organised C-RAN. Moreover, the power model for C-RAN is proposed to estimate the overall network power consumption. It is noticed that despite deploying a higher number of RRHs (49) in a given geographical area for CDI enabled C-RAN, the power consumption of a fixed C-RAN for the same geographic area is still higher by $\approx 15.28\%$ and 16.02 % for PF and RR schedulers, respectively.

Algorithm 5. Initial RRH Association to Active BBUs During Cell Integration

Input: List A of No. of active BBUs
No. of required BBUs
BBU-RRH mapping vector r

if *No. of required BBUs* $<$ |A| **then**
  for *m=1 to* |A| **do**
    for *i=1 to C* **do**
      for *j=1 to c* **do**
        -Select RRH$_{ij}$ from BBU-RRH vector **r**
        if $RRH_{ij} = m$ **then**
          $Z_m \leftarrow$ RRH$_{ij}$
          {Z_m is a List of RRHs handled by BBU$_m$}
        end
      end
    end
  end
  -Sort List A in decreasing order of BBU loads
  for *m=1 to end of List A* **do**
    if m $\neq$ |No. of required BBUs| **then**
      $\overline{A} \leftarrow$ BBU$_m$ {$\overline{A}$ is a List of required BBUs}
    else
      $\overline{B} \leftarrow$ BBU$_m$
      {$\overline{B}$ is a List of BBUs to be switched off}
    end
  end
  -Sort List $\overline{A}$ in increasing order of BBU loads
  for *i=1 to end of List* $\overline{B}$ **do**
    -Select ith BBU from List $\overline{B}$
    -Select List Z_m of the ith BBU
    m=1;
    for *j=1 to end of* Z_m **do**
      if m $>$ $|\overline{A}|$ **then**
        m=1;
      end
      -Select RRH at jth index in List Z_m
      -Select BBU at mth index of List $\overline{A}$
      -BBU$_m \leftarrow$ RRH^j {Assign RRH j to BBU m}
      m++
    end
  end
  -Switch off all BBUs in List $\overline{B}$
end

REFERENCES

[1] T. Q. Quek, M. Peng, O. Simeone, and W. Yu, *Cloud Radio Access Networks: Principles, Technologies, and Applications*. Cambridge, U.K.: Cambridge University Press, 2017.

[2] M. Peng, Y. Sun, X. Li, Z. Mao, and C. Wang, "Recent advances in cloud radio access networks: System architectures, key techniques, and open issues," *IEEE Commun. Surv. Tutorials*, vol. 18, no. 3, pp. 2282–2308, Apr.–Jun. 2016.

[3] A. Checko, et al., "Cloud RAN for mobile networks; A technology overview," *IEEE Commun. Surv. Tutorials*, vol. 17, no. 1, pp. 405–426, Jan.–Mar. 2015.

[4] M. Khan, R. Alhumaima, and H. Al-Raweshidy, "Reducing energy consumption by dynamic resource allocation in C-RAN," in *Proc. Eur. Conf. Netw. Commun.*, 2015, pp. 169–174.

[5] Y. Cai, F. R. Yu, and S. Bu, "A decision theoretic approach for clustering and rate allocation in coordinated multi-point (CoMP) networks with delayed channel state information," in *Transactions on Emerging Telecommunications Technologies*. Hoboken, NJ, USA: Wiley, 2017, vol. 28, no. 1, Art. no. e2831.

[6] C. Jiang, H. Zhang, Y. Ren, Z. Han, K. C. Chen, and L. Hanzo, "Machine learning paradigms for next-generation wireless networks," *IEEE Wireless Commun.*, vol. 24, no. 2, pp. 98–105, Apr. 2017.

[7] X. Huang, G. Xue, R. Yu, and S. Leng, "Joint scheduling and beamforming coordination in cloud radio access networks with QoS guarantees," *IEEE Trans. Veh. Technol.*, vol. 65, no. 7, pp. 5449–5460, Jul. 2016.

[8] K. Wang, W. Zhou, and S. Mao, "On joint BBU/RRH resource allocation in heterogeneous cloud-RANs," *IEEE Internet Things J.*, vol. 4, no. 3, pp. 749–759, Jun. 2017.

[9] Y. Shi, J. Zhang, and K. B. Letaief, "Group sparse beamforming for green cloud-RAN," *IEEE Trans. Wireless Commun.*, vol. 13, no. 5, pp. 2809–2823, May 2014.

[10] S. Namba, T. Warabino, and S. Kaneko, "BBU-RRH switching schemes for centralized RAN," in *Proc. 7th Int. ICST Conf. Commun. Netw. China*, 2012, pp. 762–766.

[11] K. Sundaresan, M. Y. Arslan, S. Singh, S. Rangarajan, and S. V. Krishnamurthy, "FluidNet: A flexible cloud-based radio access network for small cells," *IEEE/ACM Trans. Netw.*, vol. 24, no. 2, pp. 915–928, Feb. 2016.

[12] Y. S. Chen, W. L. Chiang, and M. C. Shih, "A dynamic BBU-RRH mapping scheme using borrow-and-lend approach in cloud radio access networks," *IEEE Syst. J.*, vol. PP, no. 99, pp. 1–12, 2017.

[13] C. Ran, S. Wang, and C. Wang, "Optimal load balancing in cloud radio access networks," in *Proc. IEEE Wireless Commun. Netw. Conf.*, Mar. 2015, pp. 1006–1011.

[14] M. Awais, et al., "Efficient joint user association and resource allocation for cloud radio access networks," *IEEE Access*, vol. 5, pp. 1439–1448, 2017.

[15] M. Khan, R. S. Alhumaima, and H. S. Al-Raweshidy, "QoS-aware dynamic RRH allocation in a self-optimized cloud radio access network with RRH proximity constraint," *IEEE Trans. Netw. Serv. Manage.*, vol. 14, no. 3, pp. 730–744, Sep. 2017.

[16] W. He, J. Gong, X. Su, J. Zeng, X. Xu, and L. Xiao, "SDN-Enabled C-RAN? an intelligent radio access network architecture," in *New Advances in Information Systems and Technologies*. Berlin, Germany: Springer, 2016, pp. 311–316.

[17] M. Condoluci, T. Mahmoodi, and G. Araniti, "Software-defined networking and network function virtualization for C-RAN systems," in *5G Radio Access Networks: Centralized RAN, Cloud-RAN and Virtualization of Small Cells*. Boca Raton, FL, USA: CRC Press, 2017, Art. no. 117.

[18] S. Gu, Z. Li, C. Wu, and H. Zhang, "Virtualized resource sharing in cloud radio access networks through truthful mechanisms," *IEEE Trans. Commun.*, vol. 65, no. 3, pp. 1105–1118, Mar. 2017.

[19] H. Dahrouj, A. Douik, O. Dhifallah, T. Y. Al-Naffouri, and M. S. Alouini, "Resource allocation in heterogeneous cloud radio access networks: Advances and challenges," *IEEE Wireless Commun.*, vol. 22, no. 3, pp. 66–73, Jun. 2015.

[20] J. Zhang, R. Zhang, G. Li, and L. Hanzo, "Distributed antenna systems in fractional-frequency-reuse-aided cellular networks," *IEEE Trans. Veh. Technol.*, vol. 62, no. 3, pp. 1340–1349, Mar. 2013.

[21] M. K. Simon and M.-S. Alouini, *Digital Communication Over Fading Channels*. Hoboken, NJ, USA: John Wiley & Sons, 2005, vol. 95.

[22] C. Bouras, G. Diles, V. Kokkinos, K. Kontodimas, and A. Papazois, "A simulation framework for evaluating interference mitigation techniques in heterogeneous cellular environments," *Wireless Personal Commun.*, vol. 77, no. 2, pp. 1213–1237, 2014.

[23] H. Zhang, C. Jiang, J. Cheng, and V. C. M. Leung, "Cooperative interference mitigation and handover management for heterogeneous cloud small cell networks," *IEEE Wireless Commun.*, vol. 22, no. 3, pp. 92–99, Jun. 2015.

[24] R. S. Alhumaima, M. Khan, and H. S. Al-Raweshidy, "Component and parameterised power model for cloud radio access network," *IET Commun.*, vol. 10, no. 7, pp. 745–752, 2016.

[25] C. Desset, et al., "Flexible power modeling of LTE base stations," in *Proc. IEEE Wireless Commun. Netw. Conf.*, 2012, pp. 2858–2862.

[26] H. A. Holtkamp, "Enhancing the energy efficiency of radio base stations," Ph.D. dissertation, University of Edinburgh, South Bridge, Edinburgh, U.K., 2014.

[27] R. S. Tucker and K. Hinton, "Energy consumption and energy density in optical and electronic signal processing," *IEEE Photonics J.*, vol. 3, no. 5, pp. 821–833, Oct. 2011.

[28] K.-L. Lee, B. Sedighi, R. S. Tucker, H. K. Chow, and P. Vetter, "Energy efficiency of optical transceivers in fiber access networks [invited]," *J. Optical Commun. Netw.*, vol. 4, no. 9, pp. A59–A68, 2012.

[29] A. Chatzipapas, S. Alouf, and V. Mancuso, "On the minimization of power consumption in base stations using on/off power amplifiers," in *Proc. IEEE Online Conf. Green Commun.*, 2011, pp. 18–23.

[30] G. Auer, et al., "How much energy is needed to run a wireless network?" *IEEE Wireless Commun.*, vol. 18, no. 5, pp. 40–49, Oct. 2011.

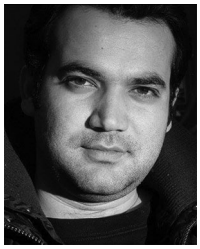
[31] A. Kaveh, "Particle swarm optimization," in *Advances in Metaheuristic Algorithms for Optimal Design of Structures*. Berlin, Germany: Springer, 2017, pp. 11–43.

[32] A. Passaro and A. Starita, "Particle swarm optimization for multimodal functions: A clustering approach," *J. Artif. Evol. Appl.*, vol. 2008, 2008, Art. no. 8.

[33] J. Vesterstrom and R. Thomsen, "A comparative study of differential evolution, particle swarm optimization, and evolutionary algorithms on numerical benchmark problems," in *Proc. Congress Evol. Comput*, 2004, vol. 2, pp. 1980–1987.

[34] 3GPP, "Technical Specification Group Services and System Aspects; Services and service capabilities (Release 13)," in *3rd Generation Partnership Project (3GPP), TS V13.0.0*. Memphis, TN, USA: Books LLC Dec. 2015.

Zainab Fakhri received the BSc degree in electrical engineering from the Al-Mustansiriya University Baghdad and the MSc degree from the University of Technology Baghdad, in 2004 and 2006. She is currently working toward the PhD degree with Brunel University London. Her research interests include cloud radio access network, self-organized networks, and millimeter wave communications.



Muhammad Khan received the BEng degree in electronics engineering from the COMSATS Institute of Information Technology (CIIT), Abbottabad, Pakistan, in 2008. He is currently working toward the PhD degree in wireless communications with Brunel University London, United Kingdom, under the supervision of Prof. Hamed Al-Raweshidy. He joined the Department of Computer Science, Abdul Wali Khan University Mardan, Pakistan, as a lecturer in 2009. His main research interest is next generation wireless communications and cloud radio access networks, particularly in self-optimisation, and load balancing.

Firas Sabir received the PhD Degree in electrical and electronic engineering from the University of Technology, Baghdad, in 2008. He is an associate professor in the computer engineering department university of Baghdad. He is currently a visitor researcher with the Wireless Networks and Communications Centre (WNCC), Brunel University London, United Kingdom. His areas of specialty and interest include the next generation of mobile communication, cloud radio access networks, and self-organised networks.

Hamed Al-Raweshidy is professor of Communications Engineering and has been awarded his PhD, in 1991 from Strathclyde University in Glasgow/United Kingdom. He was with Space and Astronomy Research Centre/Iraq, PerkinElmer, Carl Zeiss/Germany, British Telecom/United Kingdom, Oxford University, Manchester Met. University and Kent University. He is currently the director of the Wireless Networks and Communications Centre (WNCC) and director of PG studies (ECE), Brunel University, London, United Kingdom. His current research area is 5G and beyond such as C-RAN, SDN, IoT, M2M and Radio over Fibre. He is a senior member of the IEEE.