Physical Communication 48 (2021) 101415

Contents lists available at ScienceDirect

Physical Communication

journal homepage: www.elsevier.com/locate/phycom

Review article

Resource allocation trends for ultra dense networks in 5G and beyond networks: A classification and comprehensive survey

Nidhi Sharma *, Krishan Kumar

Department of Electronics and Communication Engineering, National Institute of Technology, Hamirpur 177001, India

A R T I C L E I N F O

Article history:
Received 28 February 2021
Received in revised form 3 June 2021
Accepted 1 July 2021
Available online 10 July 2021

Keywords:
Resource allocation
LTE-U
Cognitive radio
HetNets
Ultra dense networks

A B S T R A C T

With an exaggerating upsurge in mobile data traffic, the wireless networks are confronted with a subtle task of enhancing their network capacity. The shortage of spectrum resources creates a jamming situation in the enhancement of network capacity. To overcome this challenge, cellular networks have been persuaded to seek more fruitful radio spectra. That is why the wireless industry has been experiencing a new evolution through ultra-densification. Ultra dense networks (UDNs) involving LTE-U, cognitive radio networks, heterogeneous networks, cloud-radio access networks, device to device networks and millimeter wave networks appear to be the leading technologies for many more years to come for achieving the distinct capabilities that 5G and beyond networks are expected to provide. Therefore, these technologies will be a crucial enabler for next-generation mobile communications for enhancing capacity. As the resources are scarce which have to be shared by ubiquitous users, therefore it becomes more impelling to follow resource allocation approaches. Hence, the resource allocation in cellular networks inherently makes endeavors for the maximization of resource utilization such as spectrum efficiency, power efficiency etc. In this direction, the article provides a detailed survey of resource allocation approaches for UDNs in 5G and beyond networks. Specifically, in the first phase, this article presents the resource allocation process in different scenarios of UDNs. In the second phase, a taxonomy to classify the resource allocation problem based on approaches, methods, and optimization criteria has been reviewed. The last phase alleviates the main difficulties of the resource allocation process in the wireless network; some prevailing and feasible techniques are presented in detail too. Finally, the emerging technologies, challenges and active research initiatives are outlined which require the attention of the researchers.

Contents

* Corresponding author.
E-mail addresses: nidhi.sharma251291@gmail.com (N. Sharma), krishan_rathod@nith.ac.in (K. Kumar).

https://doi.org/10.1016/j.phycom.2021.101415
1874-4907/

1. Introduction

With the steep advancement of wireless mobile networks in the past few years, the increasing prevalence of smart terminals (e.g., mobiles, laptops) and emerging applications (e.g., internet of things (IoT), artificial intelligence (AI), caching, etc.) has activated an overwhelming growth of mobile data stemming. It has been estimated in the future forecasts that the global mobile data traffic demand is expected to reach 351 Exabyte by 2025 [1]. However, in the 5G and beyond networks, the scanty spectrum levels are imposing a bottleneck situation in further boosting the capacity of wireless communications. To fundamentally breakthrough this predicament, academic and industrial communities paid attention to seek advanced technologies like long term evolution-unlicensed (LTE-U), cognition, etc., with ultra-densification process for improving network capacity [2]. Therefore, the way to boost network capacity and area spectrum efficiency is network densification where access points (APs) and communication links densification occurs to support explosive data traffic. This densification process can be achieved by establishing supplementary low power nodes (such as microcells, femtocells, small-cells, etc.) on every lamp post outside or spaced at a distance of less than 10 m inside, and is displayed in various fashions that coexist congenially in the 5G and beyond networks. As an example, the various such main networks are (a) LTE-U (b) cognitive radio networks (CRNs) (c) heterogeneous networks (HetNets) (d) cloud-radio access networks (C-RANs) (e) device to device (D2D) networks (f) millimeter wave (mmWave) networks, and etc.

As agreed by 3GPP, LTE-Advanced (LTE-A) has emerged as one of the encouraging technologies for addressing the proliferating capacity demand, explosive data traffic and improved user experience with the use of many forefront technologies, such as massive multiple-input multiple-output (MIMO) [3,4], carrier aggregation (CA) [5], link adaptation, and licensed assisted access (LAA) [2], etc. These technologies have provided aid to the traditional licensed spectrum in enhancing their current network performance demands with high spectral efficiency of the current wireless cellular network. As the opposite side of the coin, it closes the door for further capacity improvement as this limited licensed spectrum is discreditably deficient. Cellular networks have been encouraged to tap the abundant unlicensed spectrum for increasing the capacity. So, to overcome the constraint of the limited spectrum, the idea of outstretching LTE to the unlicensed spectrum, also known as LTE-U has been studied on an extensive level in the past few years. For auxiliary usage of the unlicensed spectrum into regularity, another approach called the LAA has been launched by the Third Generation Partnership Project (3GPP).

The second type of promising technology called CRNs can be used for improving network capacity as well as the spectrum efficiency [6]. By definition, a cognitive radio (CR) is a radio which is capable enough to change its transmitter parameters by knowing its environment through information exchange. It has two components, known as the primary and secondary networks. The spectrum is owned initially by the licensed network called as the primary network, whereas the unlicensed secondary network seeks to ingress the licensed spectrum to improve its spectrum efficiency. In order to share the spectrum, the transmission parameters of secondary network ought to be configured in such a way that the primary network remains conserved [7,8].

The third type of promising technology is HetNets in which number of small cells are increased to the legacy macro cells in order to improve the network capacity. These low power small cells can be categorized as femto cell, pico cell, microcell, relay, etc., which helps in traffic offloading from macro base station (MBS), increase the coverage and improve the network capacity [1]. Consecutively, network capacity is improved by using a central wireless cloud network which is known as C-RAN. The main properties of C-RAN comprise sharing of resources, centralized processing and real-time cloud computing [9]. Furthermore, D2D network technology is also being considered as a promising newcomer which could perform a vital job of enhancing the network capacity as well as spectrum efficiency, as peer to peer communication is encouraged between proximate users [10]. Another way to improve network capacity is to move from the traditional radio spectrum to the unused spectrum where wide frequencies from 30 GHz to 300 GHz are available, such as mmWave networks [11].

Despite the manifold transmission manners, nowadays wireless cellular networks trend has transformed from traditional to denser networks, very dense networks, and UDNs [12]. The UDN can be defined as a cellular network where the number of active users is lesser in comparison to the cell density, or quantitatively the cell density is more immense than 10^3 cells/km^2 [13,14]. A lot of essential features come along with various ultra-densified networks in comparison to the traditional networks with following points discuss as: (i) Remarkable augmentation of network performance outputs, e.g., network coverage and spectrum efficiency, etc., can be attained through network densification [15], (ii) It is established that LTE-U UDN technology can utilize the unlicensed spectrum in an efficacious way. Therefore, from the user's viewpoint, to amplify their data rates, the cellular users can approach both licensed and unlicensed spectra by exploiting the existing CA technology, (iii) The coordination between the primary and secondary networks helps to achieve enhanced performance [16] or economic gain [17] and it optimizes the overall performance of CRN [18]. Hence, the problem of spectrum underutilization can be addressed by using CR based UDN technology, (iv) Number of APs, such as base stations (BSs), small cell BSs, relay nodes, remote radio heads (RRHs), antennas, D2D enabled users, machine-type communication, smart devices, future applications, and vehicle to everything (V2X) networks, etc. coexist. In order to handle such a complex and comprehensive system, well-organized resource allocation mechanisms are required. Fig. 1 presents a system model showing resource allocation in UDNs with emerging applications.

Resource allocation is an essential prerequisite, and remarkable research efforts have been undertaken for investigating and designing coherent resource allocation schemes with the aim to serve an ever-increasing number of users and meeting high standards of quality of service (QoS). So, a survey of resource allocation has been provided in this article, in the context of different scenarios of UDNs. To the best knowledge of the authors', this article serves to provide a comprehensive survey on resource allocation in the context of UDNs involving LTE-U, CRNs, HetNets, C-RANs, D2D networks, and mmWave networks. Unlike previous works, this survey acts as a front runner to discuss methodologies and techniques which are efficient in solving the obstacles pertaining to resource allocation in the context of different scenarios of UDNs.

The subject of resource allocation has been widely elaborated but not well studied. For better illustration, this article discusses several review articles on UDNs, and their related topics as summarized in Table 1. Generally, the previous works [19–29] just mention roughly some background of resource allocation in different scenarios of a wireless network, but the difficulties and possible solutions of solving the resource allocation process remain unstated and unsolved.

1.1. Major contribution of the survey

As a part of its important contributions, this article provides the state of the art of resource allocation for different scenarios of UDNs in 5G and beyond wireless networks, discuss approaches and techniques to overcome the obstacles when accomplishing the resource allocation, and debate challenges as well as unfold research spheres. In this survey, numerous hurdles of designing resource allocation schemes generated by the exhaustive properties of future wireless networks, the uncertainty of user deployment, as well as practical deployment, and novel service requests are discussed. The key contributions of this article are summarized as follows:

- The article presents a comprehensive survey of the exiting resource allocation techniques for UDNs in detail. The aim of the article is to fill the research gap exist till date by presenting focused survey on resource allocation techniques in ultra dense LTE-U, ultra dense CRNs, ultra dense HetNets, ultra dense C-RAN, ultra dense D2D networks, and ultra dense mmWave networks scenarios.
- The article presents the comparative analysis of all the resource allocation techniques in detail with their advantages to provide outline of ongoing research in UDNs.
- The article presents emerging technologies, research gaps and challenges that are still open and require the attention of research community.

The rest of the article has been organized in the following sequence: In Section 2, the article briefly introduces literature on the resource allocation process in different scenarios of ultra dense wireless networks. In Section 3, a taxonomy to classify the resource allocation process has been introduced. Efforts have been made to address the main difficulties of resource allocation in ultra dense wireless networks and provide several effective solutions based on different techniques in Section 4. In Section 5, emerging technologies for UDNs have been discussed. In Section 6, challenges have been identified and open research directions of 5G and beyond wireless network design before conclusions are drawn in Section 7. Fig. 2 delineates the organization of the article.

2. Resource allocation in different scenarios of ultra dense networks

Resource allocation has been recognized as an important function for 5G and beyond networks since the disability and alterations of the wireless channels initiate dissimilar proportions of diversity gains in frequency, time, and space, besides multiuser diversity. Hence, the improved utilization of accessible resources could be achieved through the adoption of efficient resource allocation techniques. The resource allocation process

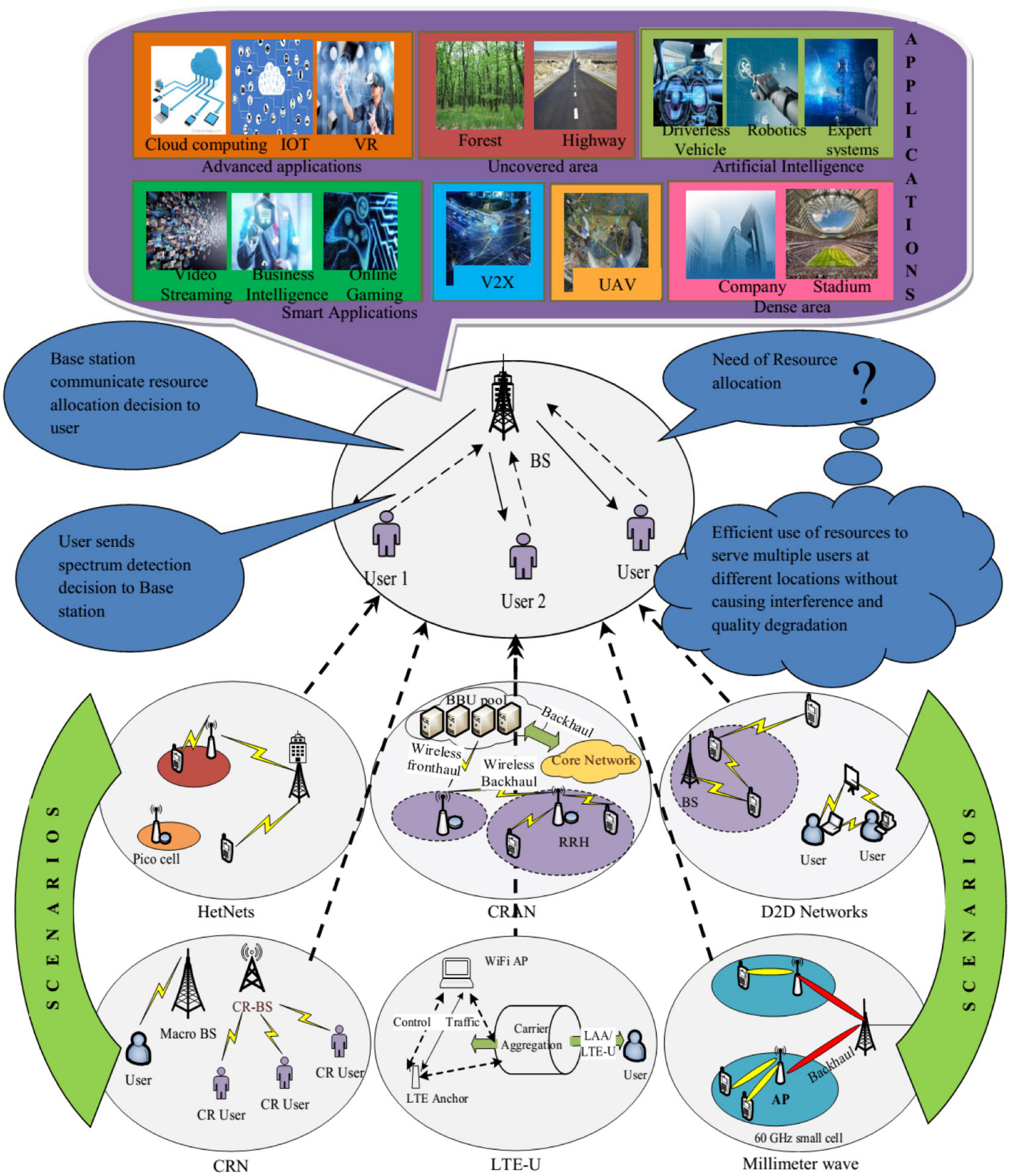

Fig. 1. System model showing resource allocation in UDNs with emerging applications.

can be visualized as shown in Fig. 3. The information (like CSI, interference temperature (IT) threshold) acquired from the neighboring environment linked to both primary/licensed and secondary/unlicensed systems, could be regarded as inputs to the resource allocation process. Though the acquired information could be perfect or imperfect, yet efficient resource allocation techniques should be able to harness both. The running of effective resource allocation algorithms mark the second stage which tends to distribute the resources in hand such as channel, time slots, spectrum bands, transmit and receive antennas, and power, etc. The scheduling parameters such as user assignment, and/or spectrum assignment at each time slot act as a yield of a resource allocation scheme. Even, the power and data rate assigned to each user, and the beamforming matrix become a part of the outputs.

In UDNs, the tally of APs existing in a specified area can be equal or surpass that of users by AP densification, which can be understood and implemented in various designs under different scenarios. The densification process is exhibited variously. Here, resource allocation in six different types of ultra dense scenarios has been considered, namely; ultra dense LTE-U, ultra dense CRNs, ultra dense HetNets, ultra dense C-RANs, ultra dense D2D networks, and ultra dense mmWave networks. These technologies have various emerging applications like cloud computing, virtual reality (VR), V2X, machine to everything (M2X), AI (e.g., driverless car), UAV, etc., as shown in Fig. 1.

2.1. Resource allocation in LTE-U system

To enhance the capacity of cellular networks, there is need to exploit the copious unlicensed spectrum. So, the idea of extending LTE UDN in unlicensed spectrum is known as LTE and WiFi coexistence which is widely investigated in the literature. To have fair concurrence among LTE and WiFi, the user's efficient coexistence mechanisms are required. Furthermore, the regulation of resources in the LTE-U ultra dense system plays a vital role in determining system performance. So, we first discuss the resource allocation in LTE-U ultra dense system. Distributed channel access of WiFi is different from the centralized architecture of LTE; hence

Table 1
Summary of the overviews and surveys of UDNs and their related topics.

Description	Focus	Refs.
1. A brief overview of deployment challenges in small cell and UDNs • Small cell and UDN architecture and enabling technologies discussed; • Different types of small cell and UDN classified; • Challenges in resource management and interference management addressed.	Small cell and UDNs	[19]
2. An overview of energy efficient techniques for network management in HetNet UDNs • Energy efficiency techniques and enabling strategies are categorized; • Key design issues and future research direction.	HetNet UDNs	[20]
3. A detailed overview of design agreements for LTE-U system • LAA channel selection, downlink LAA framework; • Coexistence enhancements for LAA listen before talk (LBT); • Radio resource management and channel state information (CSI) Measurements.	Network coexistence (LTE-U)	[21]
4. An overview of dense small cell networks • Discussed densification model types and techniques; • Enabling technologies and research issues addressed.	UDNs	[22]
5. A detailed survey on UDN and emerging technologies • Security issues in Massive MIMO, D2D, IoT and, visible light communication systems; • Security threats and their possible solutions in UDNs are discussed.	UDNs	[23]
6. Resource management in LTE-U system • Single small base station (SBS), multiple SBSs, D2D networks, vehicular networks, and unmanned aerial vehicle (UAV) systems; • Research issues addressed for resource management in LTE-U.	Network coexistence (LTE-U)	[24]
7. An overview of resource allocation techniques in CRNs • Various design approaches and CR optimization methods are presented for the resource allocation problem; • Discussed QoS criteria for the physical and the medium access control layers; • Dynamic spectrum allocation and aggregation, and frequency mobility.	CRNs	[25]
8. Recent advances in resource allocation methods in underlay CRNs • Design of resource allocation process and its components; • Resource allocation algorithm based on the methodologies, parameters and constraints, common techniques, and network architecture.	Underlay CRNs	[26]
9. An overview of interference control, resource allocation, and self-organization techniques in underlay HetNets. • Spatial interference coordination at transmitter and the interference cancelation at the receiver; • Comprehensively discussed the multi-dimensional optimization, cross-layer optimization, and cooperative radio resource management methods; • The self-configuration, self-optimization, and self-healing techniques.	HetNets	[27]
10. State of art of resource allocation techniques in UDNs • Resource allocation approaches in UDNs scenario; • Classification of resource allocation methods/techniques with a feasible solution; • Emerging technologies and future research directions.	UDNs	[28]
11. Survey of state of the art of UDNs • Challenges in UDNs including resource management, mobility management, interference management, and security; • UDNs on emergent applications like IoT, security and privacy, modeling and realistic simulations, and relevant techniques.	UDNs	[29]

the fair coexistence of the two networks is a great challenge. To overcome this challenge the two approaches LTE-U and LAA are widely exploited in literature which combine unlicensed band with the aggregator in the licensed band as shown in Fig. 4.

LTE-U resource allocation cross-layer framework has been proposed in [30] where proportional fairness among all the users of LTE-U and WiFi networks has been achieved by using centralized approach. Due to the inherent extreme density of 5G\6G cellular networks, their centralized architectures face lack of scalability and an enhanced information exchange overhead, which in turn questions their suitability. In this context, the game-theoretic and decentralized approaches which have ease of scalability and less overhead are beneficial. A self-organized decentralized approach has been followed to reduce overhead and increase system performance [31]. In this approach, the system autonomously learns and decides the allocation of the licensed and unlicensed band to each mobile station (MS) in small cell network. To proactively allocate the LTE-LAA resources over the WiFi spectrum, a reinforcement learning based long short term memory (RL-LSTM) cells algorithm has been used [32], and these resources are formulated as a non-cooperative game to maximize throughput. The scheme ensures fairness among users, and hence boosts the throughput and spectrum efficiency. The average packet stopover time has been utilized as the performance metric for solving the problem of allocating licensed and unlicensed spectrum in HetNets [33]. For an unlicensed band, LBT queueing model has been proposed to capture its reliability, spectral efficiency, and additional delay. As an extension of this work, the technique of spectrum allocation in downlink HetNets with numerous radio access technologies over different bands has been adopted to obtain optimal network utility [34]. Further, the authors in the above discussion do not consider the hidden terminal problem and the effect of user action prediction. The WiFi hidden node problem in the LAA system arises because of LBT, which results in interference in the LAA system. To overcome the problem, the authors have proposed the hidden node aware resource allocation algorithm, which maximizes the throughput and guarantees the QoS of the LAA system [35].

All of the above approaches focus on spectrum efficiency, throughput maximization, and maintaining the QoS, whereas the aspect of energy efficiency is not considered. So, an energy-efficiency criterion is developed in [36] to reduce the downlink transmission power and increase the user throughput by allocating modulation and coding scheme, resource blocks, and transmit power to users. This leads to the allocation of different transmission powers to different users according to their QoS requirements and channel conditions. Moreover, energy-efficient spatial reuse and intercell interference coordination are achieved in the LTE system. A collision in unlicensed channel and

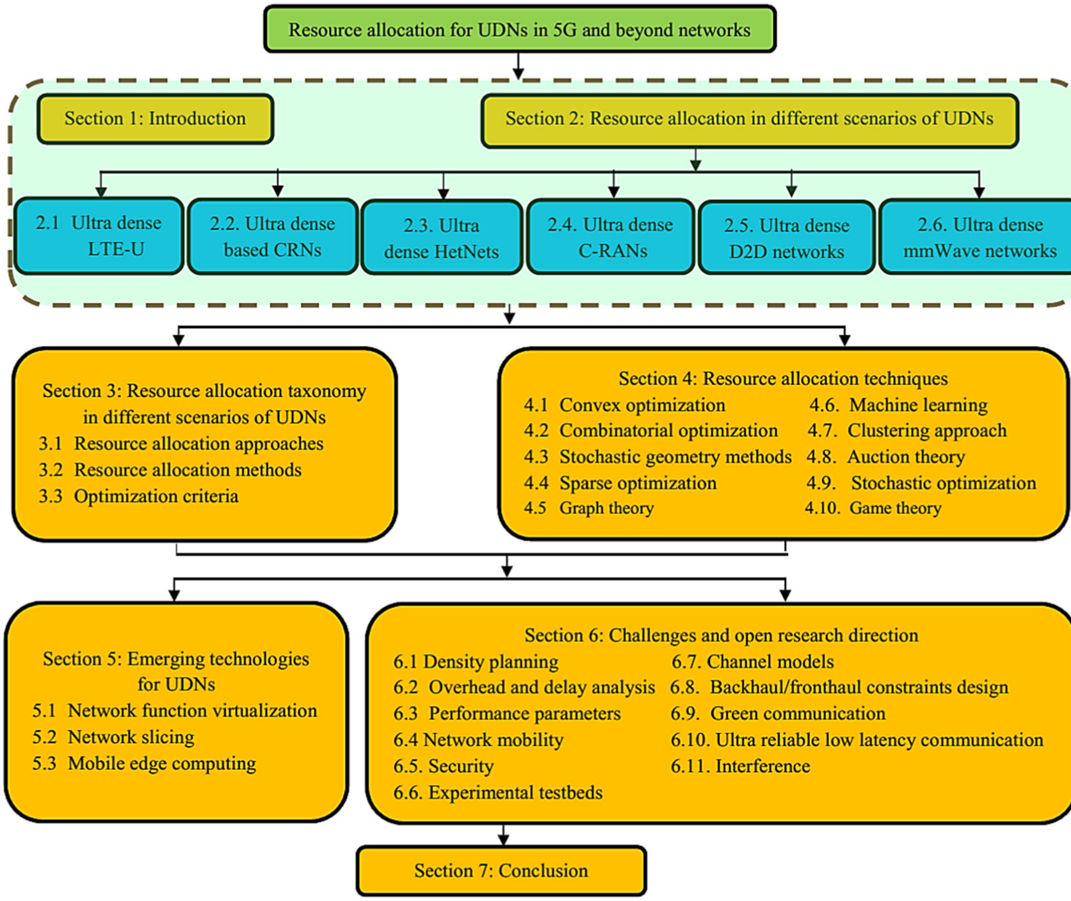

Fig. 2. Article organization.

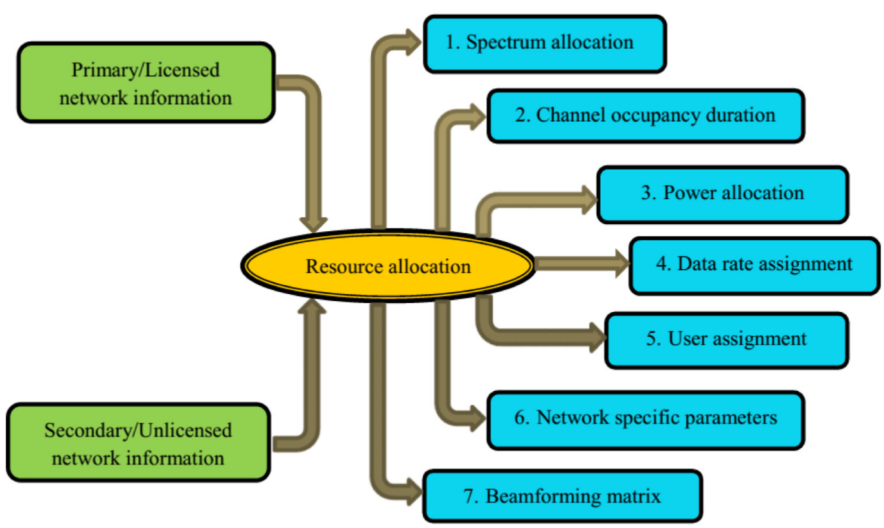

Fig. 3. Resource allocation process.

interference in a licensed channel can arise due to network densification and overlapping of dynamic WiFi nodes, which results in degradation of QoS, and thus an increase in energy consumption [37]. So, the authors have proposed the spectrum access and power allocation scheme based on the online energy-aware algorithm, which reduces power consumption in the LAA system by considering dynamic traffic load and time-varying channel conditions.

A joint power and spectrum allocation of both LTE and WiFi bands could be initialized to maximize spectrum efficiency and maintain the QoS of small cell users while guarantying fair co-existence between LTE and WiFi system [46]. For a multi-mode scenario, a joint channel and resource allocation problem for the LTE-U system has been presented to maximize network throughput, which is solved by delay column generation method and greedy algorithm [47]. Moreover, the hardware limitation of user equipment (UE) in the LTE-U system on an unlicensed spectrum

Table 2
Survey of resource allocation process in the LTE-U scenario.

Scenario characteristics	Performance parameters	Proposed scheme	Refs.
Blank subframe allocation for effective coexistence	• Spectrum efficiency • LTE delay	• Proposed a Q-learning algorithm to dynamically allocate blank subframes	[38]
LTE-U Resources allocation	• Rate improvement • Reduce collision probability	• Proposed a stochastic programming model for the allocation of resources	[39]
Channel allocation	• Fairness • Throughput	• Proposed an optimal channel selection algorithm model for resource allocation	[40]
Spectrum and transmission power allocation	• Spectrum efficiency	• A hybrid adaptive channel access scheme is proposed by taking the advantage of duty-cycle muting and LBT methods	[41]
User association, spectrum allocation, and load balancing	• Rate improvement	• Proposed a decentralized expected Q-learning algorithm for spectrum allocation	[42]
Spectrum allocation	• Spectrum efficiency • Fairness	• Proposed adaptive channel access mechanisms based on LBT methods	[43]
Radio resources allocation for both licensed and unlicensed bands	• Maximize user's utility • Throughput	• Proposed one-to-one and many-to-one matching algorithms for radio resource allocation	[44]
Power control and licensed and unlicensed spectrum allocation	• Spectrum efficiency • QoS	• The convex optimization method is proposed to solve the allocation problem	[45]

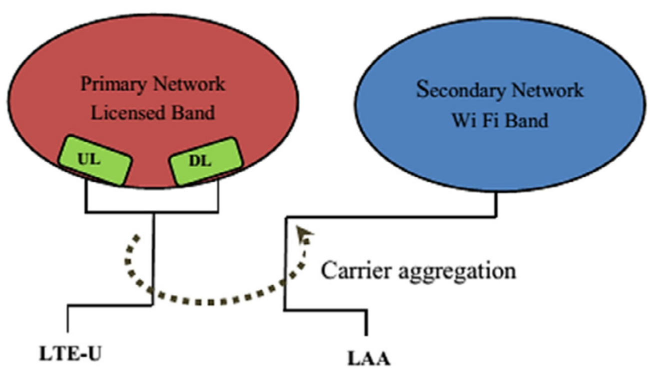

Fig. 4. The relationship between LTE-U and LAA.

is taken into consideration, which is ignored generally in current research on the LTE-U network. However, the schemes in [46] and [47] do not consider the constraint of the capacity-limited backhaul link. So, the authors in [48] have proposed the channel and power allocation for LTE in the unlicensed and licensed band by maintaining the constraint of a minimum data rate of small cell users, interference limit, and congestion-free backhaul links. In this, the Lagrangian scheme of relaxation combined with the matched game results in low complexity and fast convergence and increases overall network utility. Another extended work in [49] reuses the unlicensed band resources to improve the transmission quality, and hence maximizes the throughput of both LTE and WiFi users based on matching theory. Besides, to handle the external effect, an inter-channel cooperation subroutine is given. This task has been expanded to dynamic resource allocation [50], wherein the tradeoff between licensed and unlicensed users has been modeled. Like an interactive matching game, the dynamic resource allocation in LTE-U is handled by the use of Gale–Shapley algorithm and random path to stability algorithm. Moreover, the inter-channel cooperation algorithm is introduced to address the external effect which re-stabilize the system as well as increases the network throughput. Hence, the matching theory method for wireless resource allocation can address a few limitations of game theory and optimization. Furthermore, to have a better understanding of the resource allocation process in the LTE-U scenario, some more papers have been discussed in Table 2.

2.2. Resource allocation in CRNs

An essential factor to be considered in ultra dense CRN is the obstacle of augmenting the lowest transmission rate in between multiple source–destination pairs. It has been proposed that cooperative communication is a viable solution in a CRN [51]. Here, a joint relay assignment and channel allocation are presented to increase spectrum efficiency under the interference constraint. In another study, an efficient resource allocation scheme has been utilized to maximize the throughput for the secondary user (SU) under the primary interference power constraint, the secondary rate outage constraint, and the peak power constraint [52]. Another approach that has been proposed emphasized on using a resource allocation algorithm, which gives more proclivity to the users with large participation in the spectrum sensing process and allocating resources, in comparison to those with best channel conditions only [53]. The approach results in increased throughput and fairness among SUs.

However, the aspect of energy efficiency is missing in the above discussion. So, a cooperative energy harvesting resource allocation scheme, where SU is located close to SBS harvest energy in CRN has been proposed to increase the sum rate of the system under the interference constraint of the primary user (PU) and limited power budget constraint at the SBS [54]. In a similar study, the authors in [55] proposed resource allocation for SUs with a particular overlapping region in which heterogeneous PUs operate simultaneously in multi-radio access technology-based CRNs. It led to an increase in network capacity, spectrum efficiency, and energy efficiency. The energy-efficient aspect of spectrum sharing and power allocation is formulated by the Stackelberg game technique [56]. Here, interference has not been considered. The problem of spectrum assignment and allocation has been modeled as a worldwide optimization problem by accounting the interference between PUs and SUs, and the interference between SUs [57]. A modified binary artificial bee colony algorithm has been proposed to work out this optimization problem. The result shows the significant spectrum utilization of the allocation.

Secure communication is an essential parameter for the efficiency of ultra dense CRNs as the flexibility and openness of CRNs instigate the latest security threats. These threats either result in excessive interference at PU network, or under-utilized spectrum. Hence, they degrade the system performance. In one of the studies, the resource allocation scheme has been investigated in the presence of a PU emulation attack [58]. As false alarm probability decreases the PU detection, so the double threshold soft detection fusion scheme for CRNs has been formulated. A joint spectrum sensing and resource allocation problem has been introduced to deal with the spectrum sensing data falsification attack and SUs QoS in CRNs [59]. Here, for cooperative spectrum sensing of SUs a reinforcement learning method has been presented. Hence, the proposed scheme improves system robustness and system utility gain. To secure the transmission for both PU and SU a cooperative paradigm is maintained in CRNs [60]. For this, transmission power, relay, and time duration are selected to escalate the secondary secrecy rate, simultaneously securing the primary secrecy rate in an imperfect CSI environment.

Under the condition of a time-varying channel in a CRN, stochastic uncertainty, different estimation errors, time delay, random changing of users, and different QoS requirements exist. Thus, the allocation of power becomes a vigorous process, wherein each active SU is required to dynamically regulate the transmit power on grounds of an instantaneous objective function, all available information, and the changing environment. Dynamic control theory acts as a powerful technique to deal with such kind of situations with high precision. While considering the effect of the exogenous disturbance, parameter uncertainties, and time-delay, etc., the linear quadratic gaussian (LQG) control and H^{∞} control method can be used as a working tool for power allocation. In one of the studies, a distributed closed-loop power allocation algorithm has been used, which is based on LQG regulator with safety switching and weight adjustment mechanisms [61]. In this case, the power allocation hurdle has been constructed in the form of a state–space model with inputs as time changing CSI and few parameter measurement errors and is solved by tracking power control algorithm. This algorithm guarantees the signal to interference plus noise ratio (SINR) necessity of the SUs and controls the IT constraint of all the PUs below a threshold and perform better in term of QoS of SUs, lower computational complexity, and signal overhead. The issue of time varying channel gain and online channel estimation for a given CRN has been solved by considering a scattered power allocation algorithm based on H^{∞} state feedback control [62]. This algorithm results in better communication performance in terms of QoS of SUs, interference control, and lower computational complexity. However, it is easier to apply the H^{∞} control-based method as compared to the LQG control regulator method. Whereas the LQG control method deals with external interference in the control process, the H^{∞} control does not require any statistical information about external interference, besides the bounded energy constraint [63]. Another method based on the dynamic controller design has been studied which delineates the intrinsic transient behavior of the system. This method formulates a robust power allocation problem for distributed CRNs with channel perturbations and dynamic environment to amplify the data rate of SUs under the limitations of maximum allowable interference, and total power budget of SUs [64]. Afterward, this allocation problem is solved by a robust controller, based on a distributed projected dynamic system, which provide stable and higher transmission rate.

During the allocation of resources, it is necessary to consider the QoS provision for some delay-sensitive users or applications. So, PU's minimum constant average transmission rate is not an effective approach for delay-sensitive services. Hence, robust power allocation and relay selection methods have been formulated with the arrangements for QoS to each SU, considering probabilistic and worst-case scenarios [65]. These allocation schemes amplify the output of the system while gratifying the constraint of interference for PU in an imperfect CSI environment. To provide a statistical delay, QoS provisioning for PUs alongside optimizing the SUs performance is a great challenge in highly stochastic wireless channels. So, as the extension of the work to optimize both PU and SU performances, it has been proposed that the resource allocation problem be optimized to amplify SU's average throughput while fulfilling PU's statistical delay QoS demand, and SU's average and peak transmit power limitations [66].

Along with time delay and QoS provision, some of the work also consider energy efficiency. The concept of energy efficiency has been studied in [67] wherein the joint power and time allocation for multi-user secure CRNs with the guarantee of time delay and QoS requirements is proposed by considering perfect and imperfect CSI. Another similar work where a multi-objective resource allocation problem has been proposed is solved by two algorithms [68]. Enhanced fuzzy C-means algorithm based cluster formation for spectrum sensing and multi-objective random walk grey wolf optimization algorithm is used to choose the optimal routing path between PU and SUs. The scheme results in better throughput, delivery ratio, network lifetime, delay awareness, energy efficiency, and fairness index.

Most of the previous work considers only one type of service so, dynamic resource allocation for SUs supporting heterogeneous service has been formulated as a mixed-integer problem by considering the minimum data rate and proportional fairness constraints [69]. As it is cumbersome to have the ideal knowledge of dynamic radio environment, so dynamic channel and power allocation has been done by considering imperfect spectrum sensing which has been designed as a mixed-integer programming problem and solved as Lagrangian dual method and discrete stochastic optimization method [70]. Both of these studies do not consider the economic efficiency and the spatial reusability at the same time, so to increase spectrum efficiency additionally, these constraints are considered [71]. Here, the coalitional double auction mechanism is incorporated for full-economic spectrum having reclaimable allocation among SUs for CRNs. Furthermore, to have a better understanding of the resource allocation process in the CRNs scenario, more papers have been discussed in Table 3.

2.3. Resource allocation in HetNets

In order to enhance the capacity of cellular networks, dense deployment of various SBSs aggravates the situation towards HetNets [80]. With the evolution of wireless internet and swift universalization of smart devices, heterogeneous cellular networks turn out to be an efficient method for handling the escalating demand for seamless network coverage, enhanced capacity, minimized latencies, and inflated data rate (uplink and downlink) to end-users [81]. For supporting the demand of a thousandfold capacity increase in the next-generation wireless communication network, nowadays HetNets are emerging as more firmly packed, thereby becoming ultra dense, thus being called as ultra dense HetNets. Hence, in ultra dense HetNets, a wide range of high-power macrocells, low power small cells (i.e., pico, femto, and microcells), and supplementary non-cellular heterogeneous communication systems (i.e., WiFi, UAV, Low-power wide-area network, D2D, etc.) are transformed into a dense structure to encounter the high capacity requirements in different environments. So, the ultra dense HetNets remarkably improve the capacity and coverage even in blind wireless areas and hotspots by candidly benefitting from the enhanced spatial reuse of the scarce frequency resources and availing more ease in its installation process [82].

In ultra dense HetNets, efficient interference management methods are required. Resource sharing method based on reinforcement learning theory has been proposed in [83] to solve multi-armed bandit problem. The objective is to guide the decision of each cell in choosing the most suitable resources autonomously with minimum interference level. This has to be done while ensuring reactivity to the possible changes that can occur in resource usage. The scheme results in throughput maximization and interference control. Further, due to the dense deployment in 5G and beyond networks, two types of bottlenecks arise in HetNets: the interference issue and the cell-edge effect. Literature reveals that some of the studies have discussed these issues jointly. In order to mitigate these issues, interference cancellation techniques like the power control, cell-free BS-user association (BUA), and dynamic user-side have been exploited [84]. The cell-free BUA technique yields a new degree of freedom to utilize

Table 3
Survey of resource allocation process in the CRNs scenario.

Scenario\characteristics	Performance parameters	Proposed scheme	Refs.
Spectrum allocation	• Throughput	• Proposed online flow control, scheduling, and spectrum allocation algorithm for resource allocation	[72]
Sub-channel and power control in cooperative CRNs	• Throughput • Fairness	• Proposed a suboptimal centralized heuristic and optimal distributed algorithm for resource allocation	[73]
Power allocation	• Minimize symbol error rate	• Proposed Langrangian multiplier method-based resource allocation algorithm	[74]
Channel and slot allocation	• Energy efficiency • Throughput • Fairness	• Proposed a particle swarm optimization algorithm for resource allocation	[75]
Channel allocation	• Spectrum efficiency • Interference mitigation	• Proposed interference aware resource allocation algorithm for channel allocation	[76]
Power allocation	• Data rate	• Proposed a recursive algorithm for power allocation	[77]
Channel allocation	• Lower computational complexity	• Proposed a heuristic algorithm for channel allocation	[78]
Channel allocation	• Fairness • Throughput	• A heterogeneous optimal algorithm for channel allocation	[79]

the multi-BS diversity to reduce the interference and cell-edge effect by selecting the best BS satisfying the channel conditions and system loadings for each user in HetNets. Joint clustering and inter-cell resource allocation optimization problem has been solved in [85] by the game theory and the graph coloring algorithms respectively. The scheme results in cell average and cell edge throughput maximization. Besides, the macro diversity coordinated multipoint technique is employed to deal with the inter-cell interference issue. In another extensive study, the game theory approach has been used to mitigate the user association and resource allocation optimization problem, to coordinate transmission power and eliminate inter-cell interference among cell edge [86].

The work can further be enhanced by using the concept of energy efficiency and by adding some CR capability. This will result in enhanced coverage capacity and reliability. In this regard in a study, a power-availability-aware cell association for BS bearing small cell, with a hybrid energy supply network has been introduced [87]. This proposed framework mutually inherits the effect of the battery fluctuations and the users' power requirement based on stochastic geometry tools. Along with energy efficiency consideration, some of the work focuses on spectrum efficiency also. Recently a study has been conducted on frequency and power allocation to maximize spectrum efficiency and energy efficiency subject to the constraint of the QoS requirement of user [88].

Although, in the aforementioned studies the backhaul capacity constraint between BSs and the core network is not reviewed, yet with the network densification of the future wireless networks, huge data transmissions over backhaul links will occur, which will affect capacity. A study has been conducted under the constraints of backhaul capacity limit and power consumption to jointly optimize user association, spectrum allocation, power control, and the number of activated antennas to maximize energy efficiency and spectrum efficiency while ensuring proportional rate fairness [89]. One of the studies considered jointly optimizing the cell association and bandwidth allocation problem to maximize user rate under the wireless backhaul link constraint [90]. Global and local backhaul bandwidth allocation has been considered.

2.4. Resource allocation in C-RAN

C-RAN with ultra-densification is a promising network architecture for the upcoming next-generation wireless communication systems [91]. In the context of C-RAN, the traditional BSs are substituted by modest complexity and shallow-power RRHs that are coordinated by a centralized base band unit (BBU) pool, which serves as the latest cheap way to realize network densification. This process improves spectrum efficiency, energy efficiency, and link reliability. Moreover, the use of cloud computing technologies as the infrastructure of the central processor (i.e., BBU pool) enormously enhances hardware utilization [92].

Several studies in the literature focus on energy efficiency and increased network capacity in the C-RAN by network densification process. A cross-layer resource allocation scheme has been studied to reduce the overall system power utilization in the BBU pool, fiber links, and the RRHs in the C-RAN system [93]. Here, the resource allocation problem considers RRH selection, elastic service scaling, and joint beamforming altogether and is solved by the Shaping-and-Pruning (SP) algorithm, and results in weighted sum-rate maximization. Under the aspect of throughput increase and energy efficiency, a study has been conducted to evaluate the benefits of small cell on/off and adaptive baseband unit sharing under UDNs with C-RAN architecture [94]. A framework has been designed for an energy-efficient C-RAN [95] having characteristics of a joint RRH selection and power minimization through coordinated beamforming. Here bi-section group sparse beamforming framework (GSBF) algorithm is used for large-scale C-RAN, whereas iterative GSBF algorithm is used for medium-sized networks to achieve optimal points. In C-RAN, as the RRHs grow denser, the fronthaul capacity becomes critical, which restricts the performance of the network [96]. Here, a combination of hard and soft transfer modes methodology has been exploited to maximize the delivery rate, simultaneously fulfilling front haul capacity and per-enhanced RRH power limitations.

End-to-end resource allocation methods, from data centers to users, are required to ensure the QoS for users. In C-RAN, optimizing the end-to-end performance during the resource allocation process is imperative. To this end, the authors have proposed a multi-pair two-way transmission, using lattice based compression method to optimize end-to-end user data rate in C-RANs [97]. In reference [98] hypervisor has been used to allocate resources dynamically among various virtual operators. Hence, the scheme improves throughput and minimizes end-to-end delay for delay sensitive applications. For minimization of both network power and spectrum consumption of hybrid C-RAN, an optimization framework has been proposed to reduce end-to-end latency constraint from central cloud to the end-user [99]. In yet another study to obtain an end-to-end QoS for end users, a queuing method prioritizing the improvement in time and space has been followed by using the radio access network buffer management concept [100]. Further, resource reservation is an important concept for network management as it involves traffic predictions for resource allocation. Therefore, it helps to maintain better network performance from data centers to users. In this, the transmission resources in virtualized RAN should be sliced and reserved to aid unseen traffic requests. In this context, a block coordinate descent algorithm has been proposed to solve the resource reservation problem by using multi-path routing [101]. Further, reservation of link capacities in backhaul and transmission resources in RAN is obtained to minimize outage of wireless links and to maximize the total expected traffic load.

Additionally, this article provides a summary of related works on the resource allocation process in HetNets and C-RANs in Table 4.

2.5. Resource allocation in D2D networks

D2D technology is used to deal with the turbulently increased data traffic problem. In this technique, radio resource blocks are reutilized to enhance the spectrum efficiency in an ultra dense scenario. In other words, D2D communications offload the traffic load of a BS and a pair of D2D devices can reutilize a frequency, used by another pair of D2D devices, to enhance the spectrum efficiency and system capacity when the two pairs of D2D devices have no intervention amongst themselves.

A coloring algorithm has been proposed to allocate radio resource blocks to D2D users initially and cellular users afterward in such a way that each radio resource block can be effectively reutilized [110]. The system capacity, as well as the spectrum efficiency, is improved remarkably. A joint resource block and power allocation in ultra dense D2D network scenario could also ensure QoS of the cellular UEs while maximizing the sum rate of the D2D tier [111]. Along with network improvement in D2D communication, some studies consider the interference issue, which arises as a result of the concurrence of multiple tiers of BSs. For instance, a sequential max search algorithm could be used for D2D resource channel allocation to reduce the interference among D2D, cellular, and small cells while augmenting the overall throughput of the network [112]. The heuristic resource allocation algorithm is another technique to maximize system throughput and satisfy SINR for all D2D users and small cellular users [113]. To avoid service degradation and network utilization during the time-domain muting a heuristic resource allocation technique has been developed which is established on the traffic load at the SBSs. It amalgamates improved inter-cell interference coordination (eICIC) with D2D communication [114].

2.6. Resource allocation in mmWave networks

Today's traditional wireless communication system that uses the frequencies of 300 MHz to 3 GHz has the problem of spectrum scarcity. Therefore, the mmWave communication network that uses frequencies ranges from 30 to 300 GHz is a better alternative. With the increase in frequency, the mm-wave channel becomes more affected by shadowing and path loss effects during mmWave propagation. However, due to extremely short wavelength, high gain beamforming could be obtained by using a large number of antennas. It will overcome the problem of path loss and shadowing by tuning signal power in the desired direction, by using multiple directional antennas [115].

To design beamforming, beamwidth selection, user association, and resource allocation in an ultra dense mmWave network is challenging since multiple nodes are present. Numerous discussions in the literature study about the directional transmission, beamforming, etc. in ultra dense mmWave networks. For indoor mmWave networks, the problem of user association, beamwidth selection, and power allocation has been studied [116]. Here, mmWave communications characteristics namely beam alignment policy and directional transmission has been used to achieve the objective of maximization in minimum throughput with fairness factor. Another study focusing over interference management in beam domain channels obtained user association and beamforming design for access links via the weighted sum-rate maximization [117]. Afterwards, wireless backhaul link design and time resource allocation are optimized to maximize the end-to-end weighted sum rate performance. A clustering method has been proposed [118] to increase line of sight (LoS) connectivity and to reduce co-channel interference in ultra dense mmWave femto networks. Clustering is performed on both femto users and femto BSs by considering the maximum probability of LoS link connectivity. Finally, the user association, subchannel allocation and power allocation are done to maximize the sum rate of the network. Another study [119] proposes the best option first algorithm and many-to-one matching with externalities algorithm for beam assignment and sub-band allocation respectively to manage inter-cell interference and improve sum rate of the network.

3. Resource allocation taxonomy in different scenarios of ultra dense networks

In this section, a taxonomy has been provided for the resource allocation process for the UDNs. Particularly, as the first step, the resource allocation problem has been categorized by following an approach that could be centralized, distributed, and partially distributed. Secondly, resource allocation problems can be bifurcated into different methods according to the information type\availability, demand requirements, and modeling patterns. Third, based on the problem objective, these schemes are also categorized based on numerous performance optimization parameters.

3.1. Resource allocation approaches

The literature suggests that the problem of resource allocation can be addressed by using the following techniques: centralized, distributed, and partially distributed. Each one has its advantages and disadvantages. Based on the resource allocation problem type and requirement, one of these could be utilized.

3.1.1. Centralized approach

This approach centrally survives on the existence of a principal body such as BS, eNodeB, separate node, etc. Their function is the collection of information and operation control. Information from the network is gathered by the principal body and subsequently sent to different users to harmonize their access, and the terminal decision is taken based on the measurements. As the entire network shares this global view, the ideal or near-ideal response of a performance metric (e.g., QoS, energy efficiency, spectrum efficiency, throughput improvement, interference minimization, etc.) can be achieved. Thus, the overall network performance increases.

Although the centralized scheme is more efficient due to its approach to global information, yet it is more subtle also, due to the enormous volume of information exchange, which enhances the signaling overhead. Moreover, the resource allocation process of the network in this approach entirely depends on a central node. Therefore, the failure of this central node can harshly deteriorate the network execution. Moreover, the centralized algorithms used to address the resource allocation problem require all required data related to the current network (e.g., CSI) with the presumption that calculative errors might not exist in all of known data, so that they can execute in a better way to attain the QoS of the network. Nevertheless, this assumption is unrealistic when contemplating over an actual wireless communication system as it needs substantial overhead to send the information. Hence, bandwidth requirement increases. Due to excessive information, the calculations increase manifold times which results in more time delays and errors. This scheme has been broadly discussed in literature [86,95,120].

Table 4
Survey of resource allocation process in HetNets and C-RANs.

Theme	Scenario\characteristics	Performance parameters	Proposed scheme	Refs.
Ultra dense HetNet	• User association and resource allocation	• Maximize network utility • Proportional fairness	• Lagrange dual decomposition method is used to address the resource allocation problem	[102]
	• Resource allocation in UAV-HetNets	• Energy minimization • QoS	• A heuristic approach is used to compute proportional weights and after that weighted power allocation scheme is implemented	[103]
	• Frequency and time slot selection in HetNets	• Throughput • Interference mitigation	• Game theory is used to solve resource allocation	[104]
	• Resource allocation of two-tier HetNets	• Capacity maximization	• Proposed distributed and formulated power allocation Q-learning algorithms for resource allocation	[105]
Ultra Dense CRAN	• Subchannel allocation in a cache-enabled C-RAN	• Maximize weighted network sum rate	• Proposed matching algorithm for subchannel allocation	[106]
	• Joint power allocation, user association and energy management in H-CRAN	• Maximizes the cost efficiency	• Lagrange dual decomposition method is used to solve power allocation, user association and energy management problem	[107]
	• Resource assignment and power allocation in H-CRANs	• Energy efficiency • Spectrum efficiency • Interference mitigation	• Lagrange dual decomposition method is used for resource block and power allocation	[108]
	• Coordinated cross-cell radio resource allocation in H-CRAN	• Throughput • Fairness	• Proposed resource optimization with cooperative radio resource manager	[109]

3.1.2. Distributed approach

In the distributed approach, one does not need any central entity for controlling the scheduling process. Instead, each user decides individually. In the distributed schemes, the resource allocation problems can be transformed into numerous easier subproblems which usually just improve a small set of the decision variables on the basis of local information. Hence, the distributed approach flaunts larger flexibility and resilience. Moreover, the volume of overhead and latency with simultaneous computation is reduced. However, since the distributed algorithms used to solve the resource allocation problem count on limited knowledge or information, the perfect solution can hardly be obtained. Hence, based on the above discussion, it could be stated that the distributed resource allocation methods which use local information are more dependable and realistic than those of the centralized methods. A lot of distributed schemes are exploited for resource allocation in a wireless cellular network, such as distributed algorithm based on Lagrangian relaxation method [48], distributed iterative algorithm based on the fixed-point theorem [121], alternating direction method of multipliers (ADMM) [122,123]., cooperative online learning [124], etc.

3.1.3. Partially distributed

The partially distributed approach combines the benefits of centralized and distributed design approaches altogether and overcomes the drawbacks. Hence, it has high network capacity, low control, and computational overhead. This approach has been widely discussed in the literature [73,125,126].

3.2. Resource allocation methods

In literature, different resource allocation methods are presented, which can be categorized into deterministic, stochastic, data-driven adaptive methods, and two time-scale methods, and are discussed as following:

3.2.1. Deterministic and stochastic methods

Based on the availability of information, the resource allocation process can be deterministic or stochastic. The deterministic process depends upon precise knowledge of the network, such as CSI, queue state information, cache state information, available resources, etc. Even though being ordinary and simple, the deterministic information-based schemes become too unrealistic in the real practical system and difficult to achieve. So, user position calculation can be used despite CSI [127]. Furthermore, most resource allocation methods can allow a certain degree of failure arises due to imperfect information. Hence, probabilistic information can be used with respect to deterministic. It will decrease the computational intricacy while guaranteeing optimal resource allocation results. Probabilistic resource allocation has gained attention recently, and hence several probabilistic information-based methods have been studied in the literature (e.g., [52,65]).

3.2.2. Data-driven adaptive methods

Resource allocation based on the reservation policy leads to the problem of underutilization of resources. So, there is a need to apply dynamic resource allocation methods where the demand requirements of users are varying over time. For this type of scenario, the resource demand for each timeslot is specified by start time, end time, capacity, and cost function. Further, the data-driven adaptive methods are used for dynamic resource allocation so that high reliability, resource utilization, and low latency in the communication process can be achieved.

For fully resource utilization the data-driven approach for online forecasting of resource demand which considers the amount of uncertainty in the prediction has been presented in [128]. The scheduling method namely, the Prophet has been proposed in [129] to overcome the problem of resource fragmentation and over allocation. Here, time-varying resource demand within each executor has been taken into the scheduling decision. An intelligent mobile traffic prediction and control model by using LSTM based deep learning algorithm has been proposed in [130]. In this, the peak traffic predicted at each time instant is transferred to a remote server, and resources are allocated dynamically based on the traffic adaptation using a cognitive engine. This results in achieving high reliable low latency communication.

3.2.3. Two time-scale methods

In wireless communication, the network resources are allocated in two different time-scales based on different requirements. In certain scenarios, when there are different dimensions of randomness these methods are effective. For instance, besides channel state fluctuation, the amount of harvested energy and power price also results in high temporal–spatial variations and are hard to be predicted accurately. Moreover, the time scales of channel, energy, and power price variations are different. The channel state changes in seconds whereas the state of harvested energy and power price mostly changes in minutes. This results in a difference in time-scale, and thus motivates to apply two or multi-scale resource allocation methods.

For the time-scale difference between energy arrival and data rate, the authors in [131] have proposed two time-scale cross-layer resource allocation online algorithm based on matching theory and Lyapunov optimization. The scheme results in improved data rate and energy utilization. In another study [132], an online two time-scale resource allocation algorithm has been proposed by using Lyapunov optimization. It determines the channel allocation and data collection on a small time scale and the harvested energy and electricity price on a large time scale.

A multiple time-scale coordination management scheme has been proposed for densely deployed self-organizing networks (SONs) [133]. Different SONs have different time scales, so to guarantee efficient network operation M time-scale Markov decision model has been developed. Here, the SON decisions made in each time scale make an influence on SON decisions in other M-1 time scales on the network. In continuation of [133], the authors in [134] have presented a Q-learning based algorithm for SON functions in the multiple time-scale coordination management scheme to obtain a stable control policy by learning from past experience. The approach remarkably enhances the network utility with diverse quality of experience requirements.

3.3. Optimization criteria

In the literature, several performance criteria could be earmarked by the resource allocation problem in UDNs. These criteria may include energy efficiency, spectrum efficiency, fairness, interference, throughput, computational complexity, etc. The criteria may be taken as a fundamental purpose of optimization or as a limitation that should be satisfied. In Fig. 5 taxonomy of optimization criteria or constraints related to resource allocation process has shown.

3.3.1. Energy efficiency

Energy-efficient communication has received enormous research inclination in the past few years from both academia and industry due to the continuously increasing wireless devices operating in the small cells. Energy efficiency is defined as information bits per unit of transmitted energy. In literature, various energy-efficient resource allocation schemes are discussed which focused on energy consumption minimization [56,68,135,136], energy harvesting [54,87] and energy-saving [94]. Energy consumption includes both transmitted power consumption as well as hardware or equipment energy consumption. For instance, in [137] minimization in the weighted sum of backhaul link cost, and signal power under the constraint of QoS in cache-enabled C-RAN.

Past studies suggest that the energy efficiency maximization problem can be subjected to numerous constraints such as QoS [81,136,138] interference [58], etc. Further, harvesting energy resources is yet another alternative for obtaining energy efficiency. A proposed scheme in [139], exploits harvested energy to reduce the on-grid power. A two-stage energy-aware traffic offloading scheme has been introduced for the manifold-secondary tier case, considering different operating characteristics of secondary tiers with different power sources. The harvesting method has been used in a study [140] in which SBSs harvest energy from non-conventional sources, in addition to the commonly used power grid. By efficient allocation of the available energy over time across the network, energy cooperation between cells, and thus energy efficiency is increased.

3.3.2. Interference

In future wireless communication, interference is recognized as one central problem that affects reliable communication. Interference can occur between primary\licensed and secondary\unlicensed users, or between secondary\unlicensed users, and hence affects the usage of the entire network. To resolve this problem, interference coordination with resource allocation among interfering users\nodes uproot in various domains like frequency, time, space, power, etc. Interference coordination mechanism as a frequency domain for interference mitigation can be achieved by using orthogonal frequency channels in either for dynamic or static resource allocation. Multiple access techniques (e.g., TDMA, FDMA, SDMA, CDMA) are utilized to harmonize access among users and sustain orthogonal transmissions, and hence mitigate interference. In the time-domain multiplexing technique, unlicensed spectrum eICIC mechanism is used for interference reduction for the sharing of an unlicensed channel by multi-operator LTE-U small cells [141]. Furthermore, other interference coordination methods such as power control [54], almost blank subframe [113], adaptive beamforming, and expected interference ratio [76], etc., are supposed to be a productive method of reducing the interference. Spreading based spectrum and power allocation techniques could be used to maximize spectrum usage and minimizing co-channel interference [135]. The multi-BS diversity gain and user cell interference cancellation techniques have been used to eliminate co-tier and cross-tier interference respectively [84]. In HetNets, the soft frequency reuse and co-channel reverse time division duplex framework have been introduced for interference management [90].

3.3.3. Throughput

Maximizing network throughput is an important constraint that is considered in the resource allocation problem. The resource allocation problem may broadly deal with maximizing the individual throughput of the user [36], total throughput of the network [47,94,136,142,143], network capacity [53,144], sum rate of the users [54,67,68,140], weighted sum rate [84], throughput maximization of unlicensed band [135,145] throughput maximization of all users [49], cell average and cell edge throughput maximization [85,86], and cell capacity [146]. It is known that the throughput maximization problem is subjected to numerous constraints, such as statistical PUs delay QoS requirement, SU's average and peak transmit power constraints [66], minimum SINR [147], minimum interference [65], the secondary rate outage constraints and the peak power constraints [52], limited power budget at the SBS and the interference constraint of the primary network [54], QoS in terms of success probability and per-tier minimum rate [148].

3.3.4. Quality of service

Resource allocation algorithms should assure QoS provisioning which is another important parameter to consider. The reliability of the QoS of a cellular network is significant for effective wireless communication owing to its time-varying nature. Hence, a lot of work has been done to enhance system performance while maintaining specific QoS constraints for the secondary network [65]. Since the QoS criteria demand that the system is improved application-wise, so the QoS may require improving single or multiple specific performance metrics, e.g., delay [62,66,67] (in case of online video streaming and gaming), reliability (for transportation network and electronic health monitoring needs), battery life (subjecting to smart meters), false alarm probability (when a disaster or calamity monitoring is required), rate outage probability [139,149], SINR [62,150], etc., on the basis of the type of application. The delay and throughput have been quantified into QoS metrics for LAA-LTE/WiFi coexistence System [151]. Similarly in another study, the throughput of the system is quantified into QoS metrics for both LTE and WiFi band [35].

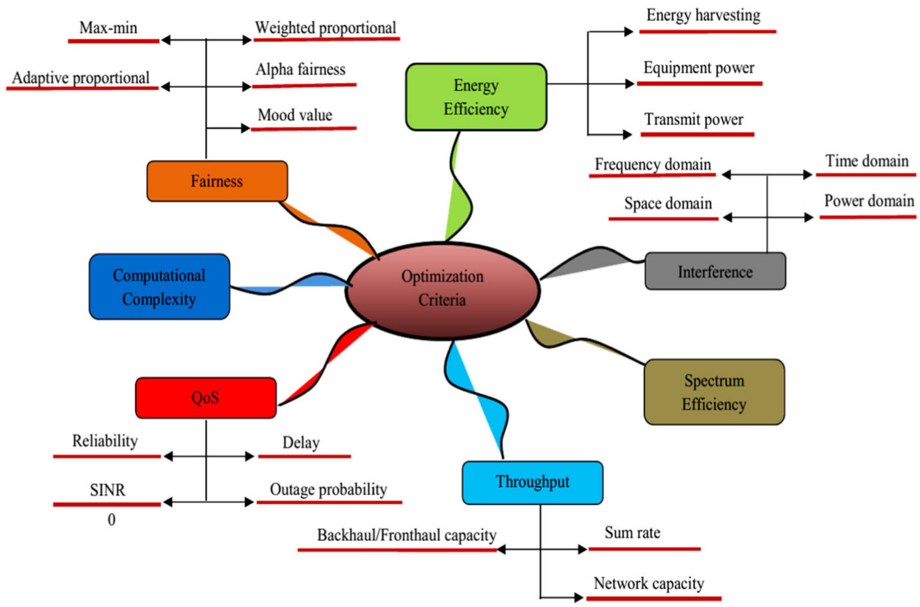

Fig. 5. Categorization of optimization criteria/constraints.

3.3.5. Spectrum efficiency

Spectrum efficiency is an important parameter to be considered in a wireless communication system. It has an important relationship with the resource allocation problem. There are numerous methods to overcome the problem of wireless scarcity, e.g., LTE communications in an unlicensed spectrum using LAA-LTE [32], cognition enabled communication [71], etc. Proactive mechanisms can be used to serve the delay tolerant data, and hence results in the efficient spectrum utilization. Additionally, several studies also consider energy efficiency and spectrum efficiency together to investigate the bargain between them [120, 152]. It can be investigated that sacrifices in terms of energy efficiency can be modified into the gain in terms of spectrum efficiency or vice versa. Another study proves that the network performance can be improved by balancing the tradeoff between spectrum efficiency and energy efficiency during different load conditions in the heterogeneous radio access technology environment [153]. A recent study shows that considerable improvement in network performance could be achieved by normalizing energy efficiency-spectrum efficiency tradeoff based on the overlay approach in heterogeneous cellular networks [154].

3.3.6. Fairness

Possible unfairness problems may arise due to the discrepancy in channel quality, that is why generally numerous fairness rules are used in resource allocation problems, including: max–min fairness [79,125,145], proportional fairness [70] weighted proportional fairness [32,52,59], adaptive proportional fairness [109, 155], alpha-fairness [123,156], power fairness among users [67], mood value [157], and time delay [53], etc. The coalition game based cooperative resource allocation algorithm has been used in [146] to serve all users fairly. In another study, fairness among the users is obtained by optimizing the least ratio between the wanted and achieved rates [125].

3.3.7. Computational complexity

For the 5G and beyond networks, there is a need to increase capacity and coverage as the number of users is increasing stupendously. So, a lot of research is going on in the field of ultra network densification. As this field has an unavoidable propensity, the setup of such a huge number of APs/users is demanding in terms of computational difficulty. In literature, various methods like convex optimization [52], game theory [48], group sparse beamforming [95], graph theory [125], stochastic geometry methods [145], stochastic optimization [70], reinforcement learning [120,150], grouping/clustering method [158], etc., are exploited to reduce computation complexity during resource allocation. Several algorithms like primal decomposition-based algorithm for the multi-objective problem [89], SP algorithm to obtain a sparse solution [93], fixed β and optimum power allocation algorithm [159] have been investigated in the literature to reduce computational complexity during resource allocation in different cellular environments. The resource allocation problem can be divided into many sub-optimal problems or addressed bit by bit, thereby diminishing the computational complexity.

4. Resource allocation techniques

The densification of cellular network enhances the performance gains, like, coverage probability, throughput, energy efficiency, spectrum efficiency, QoS, etc., but a high computational intricacy and significant signaling overhead are also reflected when resource allocation algorithms are performed. To overcome these limitations several acceptable alternatives like convex optimization, combinatorial optimization, graph theory, game theory, stochastic optimization, sparse optimization, and grouping/clustering approach are discussed in this section. Fig. 6 shows techniques discussed in literature to solve resource allocation problem to satisfy wide range of objectives or constraints.

4.1. Convex optimization

Convex optimization is an immensely effective method because it guarantees the optimality of the achieved result. In this technique, the convex objective function is minimized or maximized over a convex feasible set. One of the applications of a convex optimization method in wireless cellular networks is to solve the resource allocation problems for the effective utilization of accessible network resources. The optimization problems introduced with different purposes could exist as convex/non-convex problem, continuous/discrete or linear/non-linear problem. Examples of such problems are energy minimization, fairness, coverage maximization, interference minimization, minimum data

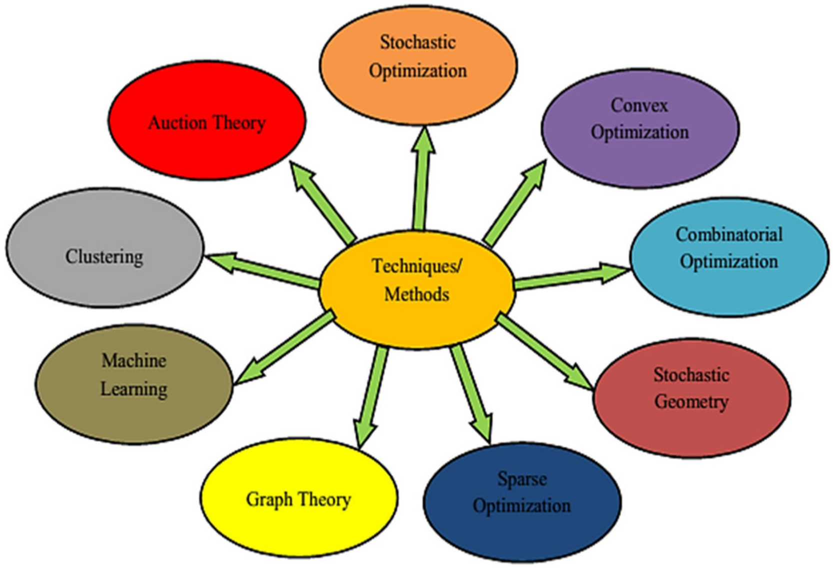

Fig. 6. Techniques for optimizing resource allocation problem.

rate, maximum system capacity, weighted sum-rate maximization, etc. In general, mathematically, the optimization problem could be represented as [160]:

minimize $f_0(x)$

subject to: $f_i(x) \leq b_i, i = 1, \ldots, m$

whereas the problem components are as follows:

- $x = (x_1, \ldots, x_n)$: optimization variables
- $f_0 : \mathrm{R}^n \rightarrow \mathrm{R}$: objective function
- $f_i : \mathrm{R}^n \rightarrow \mathrm{R}, i = 1, \ldots, m$: constraint functions
- and the optimal solution x^* has the smallest value of f_0 among all vectors that satisfy the constraints.

For convex problems including the standard linear programming, non-linear programming, quadratic programming, mixed integer programming, semi-definite programming problems, etc., the optimal solutions can be achieved by applying numerous normal optimization methods, such as Lagrange duality method, duality theory, descent methods, etc. For instance, resource allocation at the relay station (RS) has been designed as a convex optimization problem [149] that minimizes the consumed power of RS.

The non-convex optimization problems should be transformed into convex ones or solved by the heuristic method, successive convex approximation (SCA), greedy algorithm, etc., to obtain an acceptable ideal solution. The non-convex hurdle of the CRN scenario has been transformed into a convex problem by using convex hull and probabilistic transmission theories [66]. As the convex optimization function achieves the perfect power allocation strategy, the scenario adjusts to both PU's delay QoS requirements and channel conditions. Additionally, it has been proposed that the probabilistic power allocation problem could be first transformed into a deterministic non-convex problem and then it could be solved by the SCA algorithm [52]. The proposed algorithm achieves optimal performance and has fast convergence. Joint power allocation and backhaul bandwidth allocation problem could be designed as a non-convex nonlinear programming problem that gets additionally dissociated into two convex sub-problems [159].

Although, obtaining the optimal resource allocation scheme by optimization theory is quite popular, but most optimization problems consider a large network with several numbers of APs, and users are also on a large-scale. In such cases obtaining the optimal solution with lower complexity is a tedious task. Thus, sub-optimal approaches with reduced complexity prove to be noteworthy in addressing the optimization problems in wireless cellular networks. For instance, a non-convex and non-linear optimization problem has been solved by distributed joint cell association wireless backhauling bandwidth allocation algorithm which results in enhanced spectrum efficiency and network utility. Further, low complexity heuristic algorithm for cell association is used to provide a sub-optimal solution [90]. A non-linear non-convex optimization problem has been solved by SCA and novel heuristic method, and by using these methods a practical sub-optimal distributed algorithm has been developed [161]. Additionally, a convex optimization problem is divided into two sub-optimal problems and a staggered two-step joint clustering and scheduling scheme which is suitable for large networks and overlapping clusters has been introduced [85], to maximize cell average and cell edge throughput. To provide a better illustration of the optimization technique, the brief information of some similar studies has been presented in Table 5 which showcases the formulated optimization type problem, objective function, and the solution approach of this technique.

4.2. Combinatorial optimization

In a combinatorial optimization problem, mathematical techniques are used to find optimal solutions within a finite set of possible solutions. It is used to solve the discrete optimization problem and is related to algorithm theory and computational complexity theory. A combinatorial optimization problem that hikes up the spectrum usage of the allocation solution for a CRN has been solved by a modified binary artificial bee colony algorithm [57]. In this approach, the binary variables of spectrum assignment are encoded as bit strings. Further, the solution pool is selected on the grounds of various selection pressure schemes. Afterward, the latest solutions are obtained by the application of crossover and mutation operations. This algorithm improves the spectrum usage efficiency and reduces interference among PUs and SUs. Some of the work discussed mixed optimization techniques. For instance, a mixed non-convex and combinatorial optimization problem has been presented [84]. Optimization problem modification based on a virtual network and weighted minimum mean square error (WMMSE) algorithm has been used to address

Table 5
A summary of convex optimization-based resource allocation algorithms.

Problem type	Convex objective function	Solution approach	Refs.
Non-convex	• Maximizing good put	• The probabilistic and non-convex problem is transformed into deterministic and convex optimization • Adopting Hungarian algorithm to solve resource allocation problem	[65]
Non-convex	• Maximize sum-rate	• The problem is transformed into a convex optimization problem • Proposing duality theory and the simplex method-based solutions	[54]
Non-convex	• Maximize energy efficiency • Spectrum efficiency	• A low complex algorithm is developed based on the primal decomposition method	[89]
Non-convex	• Maximize energy efficiency	• Proposed iterative resource allocation algorithm to obtain the solution for resource allocation problem • A suboptimal low-complexity algorithm is also developed	[159]
Non-convex	• Maximize energy efficiency	• SCA and gradient search algorithm is adopted	[154]
Non-convex	• Maximize energy efficiency	• Hierarchical resource allocation algorithm is adopted • Further to reduce complexity uniform pricing scheme is used	[142]
Non-convex	• Maximize energy efficiency • Minimize backhaul capacity	• The transformed optimization problem in subtractive form. Adopting efficient iterative resource allocation algorithm for allocation	[152]
Non-convex	• Maximizing throughput benefit per unit energy cost	• Adopting a clustering-based method and iterative algorithm	[158]
Convex	• Maximize energy efficiency	• Optimal resource allocation of transmit power and bandwidth power allocation is done based on Lagrange duality principles	[162]

the power optimization problem. Further, the greedy algorithm is applied to find the optimal solution. Additionally, [154] proposes a dual-layer resource allocation approach where a complex mixed-combinatorial and non-convex optimization problem has been contrived as a function of the quasi-concavity of the energy efficiency function and has been solved using the difference of two concave functions approximation.

4.3. Stochastic geometry methods

Stochastic geometry model is a powerful mathematical technique which is developed to study wireless networks, where their use considerably increases in the fields of mobiles ad hoc networks, vehicular ad hoc networks, sensor networks and several types of ultra dense cellular networks such as heterogeneous cellular networks, LTE-U, CRNs, etc. The main goal of the stochastic geometry method is to presume that the locations of users or the network structure are irregular in nature due to the size and unreliability of users in wireless cellular networks. So, the use of stochastic geometry processes like homogeneous Poisson point processes (HPPPs) [148], hard core point process (HCPP), Poisson cluster process (PCP), etc., help to get the closed or semi closed-form expressions of network performance without utilizing simulation technique or deterministic models. The network performance and QoS of the network are mainly based on the SINR which forms the mathematical basis for defining coverage and connectivity of the network. A brief review of stochastic geometry methods has been presented in the following paragraph.

Firstly, the effective resource allocation algorithms can be designed by partial statistical property-based performance parameters (interference, throughput, etc.), which is derived by stochastic geometry methods. For instance, the spectrum and channel allocation problem has been formulated to seek throughput maximization under the constraints of transmission success probability for both open and closed femtocell access policies. In each tier, the calculation of the transmission success probability for shared or unshared subchannel is done by using stochastic geometric methods [148].

Successively, based on the performance parameters obtained by stochastic geometry techniques, different issues can be managed, including modeling node\user position, user association, state control, power measurements, etc. For instance, in [87] small-cell infrastructures are randomly deployed, resulting in irregularly-shaped or unpredictable networks. So, modeling the node location as random variables stochastic geometry and Markov chain modeling have been exploited to study the network performance in terms of the power outage and the coverage probability. Here, the BSs are distributed according to an independent HPPP, and at each time period, a random number of UEs are taken from a Poisson distribution which are uniformly distributed over the specified network area. Additionally, a stochastic geometry method that contemplates the variations in transmission powers and data transmission demands between cognitive secondary base stations (CSBSs) and primary BSs is used to model a HetNet scenario [145]. The primary BS and CSBS are randomly deployed and independent of each other and can be modeled as two independent Voronoi tessellations, where the primary BS, the CSBS and the associated UE follow the independent HPPP. So, by using the stochastic process the tier connection probability is calculated and, then the average outage probability of PU and cognitive SU is obtained.

4.4. Sparse optimization

As the small cell UDNs bring unforeseen complexity as well as handling difficulties, so for the optimization problem simple and approximate solutions are more useful than complex exact solutions. Hence, the data can be categorized as a set of sparse or nearly sparse coefficients to obtain a simple optimal solution. Sparse optimization in the cellular wireless network is useful for compressing the signal, minimizing the power consumption, and computational complexity. As energy consumption is the main problem in ultra dense C-RANs. So, the GSBF framework is designed [95]. Firstly, the greedy selection algorithm is used to yield an optimal solution. To further minimize the computational complexity the sparse beamforming algorithm has been introduced which is based on minimizing the weighted $l1/l2$ norm. A bi-section GSBF algorithm and an iterative GSBF algorithm have been proposed with different complexity. Similarly, the sparse multicast beamforming (SBF) problem has been formulated in [137] to reduce backhaul cost and transmitted power consumption by adopting smoothed-norm approximation. The SBF problem is obtained into the difference of convex programs and then solved using the convex–concave procedure algorithms. In this reference, a recent study presented a robust sparse beamforming vector to minimize the total power consumption of the network in C-RAN and introduced an iterative reweighted sparse beamforming algorithm to resolve the problem [163]. So, by using Bernstein-type inequality, the chance SINR constraints are obtained to be convex, and the $l1$ norm approximation technique has been utilized to get a sparse solution for the problem. The drawback of the robust sparse beamforming problem is that the chance SINR constraints do not have closed-form expression;

hence, the optimization problem cannot generate sparse values directly. Therefore, to deal with the situation Bernstein-type inequality is utilized to approximate the chance SINR constraints to static constraints. However, this proposed algorithm performs better in terms of total network power consumption compared to the scheme presented in [95].

Further, the aspect of network performance has also been considered along with energy efficiency in terms of throughput maximization [164]. Along with BS clustering, the user scheduling, and beamforming design problem has also been considered from the perspective of maximization of network utility with per-BS backhaul capacity constraint in green C-RANs. This constraint is approximated using the iterative weighted $l1$ norm method. WMMSE method is used to solve the weighted sum-rate maximization problem. In addition, elastic service scaling, RRH selection, and joint beamforming problem has been considered altogether from the perspective of sum-utility maximization function and is solved by a less complex SP algorithm to obtain a sparse solution for the active remote radio headset [93]. In this study, it can be analyzed that the average sparsity of the solution obtained by the SP algorithm is more energy-efficient as well as less complex compare to the greedy selection algorithm used in [95].

4.5. Graph theory

Graph theory is a mathematical tool that is broadly used to model pairwise interactions among objects. In this, a highly computational complex problem could be solved by designing and analyzing the graph model. Therefore, the graph theory is widely used in different scenarios of UDNs. Generally, the graph comprises a set of vertices and edges and is represented as $G = (V, E)$, where V is known as vertices and E is known as edges. These components vary according to problem type. Mostly, the vertices indicate the entities in the network. Whereas, edge existing between every two vertices shows some kind of "relationship". Graph algorithms are an efficient tool for resource assignment in UDNs. A vertex-coloring graph, conflict graph, and bipartite graph are the three types of graph which are used to solve the resource allocation problem in a wireless cellular network. The graph can be weighted or unweighted. The complexity in a weighted graph is more but it is better than an unweighted graph until the edge weights rightly consider the magnitude of the interference between any two APs. For example, in the case of a vertex coloring graph, if the interference graph is sparse and each node is to at most N nodes, where N represents the total number of possible colors then optimal coloring is possible with the low complexity algorithms. To this end, breadth first search algorithm can be used for coloring with the complexity of $O(|V|+|E|)$ with $|E| = O(|V|)$, where $|E|$ and $|V|$ are the cardinality of edges and vertices, respectively.

Vertex list-coloring problem is a popular method in graph theory that intends to allocate resources like channel assignment among APs in which two interfering nodes (nodes connected with an edge) should not be allocated the same color. As the graph coloring problem is NP-hard an optimal solution should be designed, to allocate the resource properly by a global scheduler or distributed optimization techniques. Distributed solutions are naturally preferred for a large-scale network to decrease computational complexity. For instance, a low complexity based greedy algorithm also known as a heuristic algorithm has been proposed [125]. In this algorithm at every iteration, the vertex adjacent to the largest number of variably colored neighbors is colored, possibly with a new one if required. The merit of the proposed scheme is that by carrying out the graph coloring step in a distributed manner a fully distributed scheme is achieved. Another distributed fair coloring algorithm (DFC) has been proposed [165]. This is an iterative algorithm that enhances the utilization of the resources while sharing as equally as possible between the users. For every iteration, it gives one resource (color) to each user as long as there are sufficient remaining colors, and that this color is not used by any of its neighboring users. So, the resources are shared fairly.

Another graph-based method to solve the resource allocation problem is the conflict graph, which represents logical relations between binary variables. A matrix graph (a network-like conflict graph) approach has been used to solve the frequency allocation problem with interference constraints [166]. Here, an asymptotically optimal algorithm has been proposed based on the theoretical properties of matrix graphs, which is the key to reducing the multi-coloring problem. A subchannel assignment and interference alignment optimization problem to maximize the satisfactory user ratio in MIMO femtocell networks is proposed in reference [167]. The problem is solved by transform rule to form the conflict graph and results in interference mitigation and increases the network throughput.

To model matching problem other graph-based technique called bipartite graphs, in which vertex set can be divided into two disjoint subsets in such a manner that each edge connects two vertices in different subsets, is mostly used in the cellular communication system. For instance, a polynomial-time proportionally fair resource allocation scheme that can optimally assign resource blocks to the D2D pairs and can work in time-varying channel conditions [155]. Here, the resource block allocation is solved by using the bipartite matching graph theory approach. To provide a better illustration, for the graph-based resource allocation method, a brief description of related works has been presented in Table 6. Here, we discuss the graph model type, graph mapping to resource allocation problem, algorithm, and the performance parameters.

4.6. Machine learning

Machine learning is an efficient tool that is used in the wireless communication system and networking. It is anticipated as the potential solution for an effective resource allocation and traffic offloading. In the reinforcement learning-based approach of machine learning, a specific model or any causal information is not required, but they learn their environment model by socializing with the environment and then making the strategies accordingly. The following paragraph discusses the different types of reinforcement learning methods used in the resource allocation process.

Online learning utilizes environment awareness to control co-tier and cross-tier interference and also allocate frequency resource blocks [124]. In this scheme, secondary tier (femtocell D2D, picocell) takes actions while they are learning, and hence it does not require a central controller, which results in reduced overhead, and thus less complexity. Additionally, in [120] online learning model considers compact state representation to reduce the complexity by decreasing the size of the state space, improve the algorithm convergence, and handle dimensionality. Similarly, in [150] the intuition-based online scheme utilizes a brief representation feature that considerably reduces the computation complexity by approximating the Q-value of online learning. In reference [126], using multi-agent online learning, macro BSs learn cooperatively about the environmental conditions. Further, the macro BSs estimate the SINR, based on the allocated power and the CSI reported by the network nodes and execute the sophisticated online learning algorithm for resource allocation. As the wireless channel conditions have stochastic properties and the environment's dynamics are not known, the

Table 6
A summary of graph-based resource allocation algorithms.

Graph model	Purpose	Resource allocation problem to graph mapping	Algorithm	Performance parameter
Vertex list-coloring graph	Inter cell resource allocation in UDN	• Vertex: Set of MSs in the systems • Edge: Between two nodes i and j if j is a neighbor of i • Color: Resources.	• Scheduling graphic coloring algorithm	Fairness, Goodput [85]
	Channel allocation	• Vertex: Set of MSs in the systems • Edge: Between two nodes i and j if j is a neighbor of i • Color: Channel resources	• DFC algorithm	Fairness [165]
	UE Grouping	• Vertex: Corresponds to the UEs • Edge: Represents the downlink interference conditions between vertex in each femto BS-cluster or pico BS-cluster	• UE grouping algorithm	Mitigate intra-cluster interference [168]
Conflict graph	Frequency allocation	• Vertex: Communication links as conflicting agents • Edge: Connect each two vertices if they are within the interference range • Color: Frequency band	• Approximation algorithm and floor dividing method	Interference mitigation, Complexity, Frequency reuse [166]
Bipartite graph	Interference cancellation	• Vertex: Variable nodes and factor nodes • Edge: Exists if there is a connection between variable nodes and factor nodes	• Successive interference cancellation algorithm	Uplink throughput improvement, Interference mitigation, Channel estimation[169]

authors in [80] have proposed a model-free reinforcement learning method that learns the optimal policy upon their interaction with the environment. A policy-gradient based actor-critic algorithm has been used to solve the resource allocation problem. For the LTE-U scenario, LSTM cells based deep reinforcement learning algorithm has been presented for proactively allocating LTE-LAA resources over the unlicensed spectrum [32]. This algorithm converges and attains a mixed-strategy Nash equilibrium, which results in significant improvement in rate.

4.7. Clustering

The clustering method is categorized as a powerful tool to tackle high computational complexity and unsupervised learning based problems. In the highly dense scenario as the number of SBSs increases, the overhead of information exchange among SBSs will increase and results in high signaling overhead. So, the computational complexity can be considerably reduced by organizing objects into clusters where the network members share some kind of similarities with some techniques.

Clustering-based methods are efficient in small cell networks. For that, the resource allocation optimization problem is divided into two phases. In the first phase, the SBSs are clustered based on certain criteria. Thereafter, the original optimization problem is decomposed into multiple complete sub-problems. In the second phase, the proposed resource assignment scheme is carried out for each cluster separately. Fig. 7 shows the clustering process in term of its types, features, and techniques.

Disjoint clustering and user-centric clustering are the two types of BS clustering methods. The disjoint clustering method is exploited in [170] where the entire network is decomposed into non-overlapping clusters and the BSs in each cluster cooperatively serve all the users within the coverage range. Moreover, the disjoint clustering method is effective in reducing the interference, yet the users at the cluster edge do not perform well [171]. Whereas, in the case of user-centric clustering, there is no explicit cluster edge. Here, different clusters for the different users may overlap, and hence each user is served by an independently chose subset of neighboring BSs. In reference [164] dynamic and static user-centric clustering scheme has been implemented according to different user scheduling time slots. When the BS cluster for each user can change over time, a dynamic clustering scheme is adopted which gives more freedom to fully avail the backhaul resources, but at the cost of increased overhead, as new BS-user associations required to be established continuously. When the BS-user association does not change over time, a static clustering scheme is used. The scheme results in efficient utilization of backhaul resources, and hence network gain can be improved.

A cluster strategy can be made, based on different features or criteria. These criteria can be QoS requirement, velocity, direction, geographical location, data rate requirement, interference, and distance, etc. For instance, in a study to solve the clustering problem for dense-small cell networks, a graph-based method has been adopted [158]. The minimum data rate requirement is the criteria selected for making the clustering strategy. Resource allocation based on the iterative algorithm is performed to each cluster separately. For pattern recognition problems, the enhanced fuzzy C-means algorithm could be used for the clustering process. Hence, in [68] an enhanced fuzzy C-means algorithm has been employed for spectrum sensing by scrutinizing the tradeoff between the high spectrum sensing decision reliability and network overhead.

4.7.1. K-means-class clustering algorithm

The K-means clustering technique has been applied by using the group-based graph coloring resource allocation algorithm in [172] that exploits unused spectrum areas to improve cellular services. In this technique, the feedback parameter has been considered which act as weights of the links that need excessive power and is updated in each iteration to obtain different clusters every time. Wireless resource management has been investigated in UDNs by using a modified K-means clustering algorithm scheme to seek maximization in the sum throughput parameter [147]. In [168] the cluster-based energy-efficient resource allocation scheme has been presented to mitigate interference and improve energy efficiency for UDN with low complexity. In the clustering phase, a modified K-means clustering algorithm is exploited which combines subtractive clustering and K-means clustering algorithm, and in the resource allocation phase a two-step resource blocks allocation algorithm has been proposed.

4.7.2. Game theory based clustering algorithm

When the game theory approach is used in the clustering process, the users are always acting as players, and the cooperating APs are regarded as strategies. Nevertheless, the utility function differs according to different network scenarios. In [173] distributed game-theoretic clustering algorithm has been proposed, wherein the network lifetime is extended by maintaining the equilibrium of the energy utilization for the wireless network, thereby attenuating the problem of a hotspot. Here, the cluster size is decided adaptively by using the game theory and the cooperation between cluster heads. Inside a cluster, the cell load of the serving BSs may vary so, based on load information the user-centric clustering has been obtained by using game theory techniques, which aim to maximize the system utility function [85].

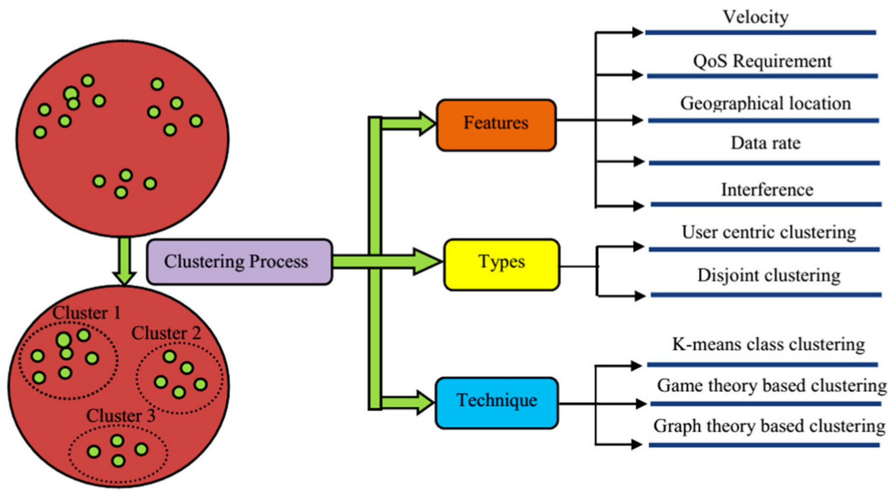

Fig. 7. Framework of clustering method.

4.7.3. Graph theory based clustering algorithm

The clustering algorithm has been proposed based on an interference graph where each vertex shows a femtocell and edge shows the interference relationship between two adjacent femtocells [174]. Here, a dynamic cell clustering strategy has been presented by establishing the various features amongst femtocells in such a manner that varying types of femtocells utilize different sub-channels to reduce mutual interference. In another study, an interference graph-based multi-cell scheduling framework has been presented to effectively reduce downlink inter-cell interference in small cell orthogonal frequency division multiple access networks [175]. Here, dynamic clustering incorporated with channel-aware resource allocation which maintains QoS and significantly improves the user's spectral efficiency of the cellular system. Improved spectrum efficiency, better throughput, and mitigated interference could be achieved through coloring based cluster resource allocation algorithm based on the graph theory [176].

4.8. Auction theory

An auction can be defined as the method of resource assignment and price learning based on bids from auctioneers'. It is broadly used as schemes for resource allocation (e.g., channel allocation, power level, and time allocation), spectrum reusable allocation, job scheduling, resource assignment for cloud computing, etc., in different UDNs scenarios. In the wireless cellular networks different auction algorithms e.g., Vickrey–Clarke–Groves (VCG) auction [177] (a type of the sealed-bid auction in which auctioning of multiple items has been done to maximize social welfare), Vickrey's auction [178] (a type of the sealed-bid auction in which winning bidder pays the second-highest price), combinatorial auction [179] (multiple items are auctioned simultaneously) are mostly used. In order to solve the power allocation problem in C-RANs an auction-based distributed scheme has been presented [180]. In CRNs the coalitional double auction mechanism has been proposed in [71] to maximize the spectrum utilization.

4.9. Stochastic optimization

Stochastic optimization is a powerful mathematical tool used to solve resource allocation problems in different scenarios of UDNs. This technique generates and use random variables and minimize or maximize objective function in the presence of randomness in the optimization process. In the resource allocation problem there may be uncertainty or randomness, which occurs due to random noise or interference and random fading channels. A problem based on uncertainty exploits stochastic optimization-based methods, and hence the optimization problem is formulated using random variables, and thus pertains random objective functions or random constraints. For instance, to track the non-stationary channels by relearning the channel statistics over time a stochastic optimization algorithm has been introduced [67], which results in efficient and fair resource allocation in CRNs. Another study is presented in [37], where a queue aware stochastic optimization method has been introduced to reduce the average power consumption of SBSs by optimizing the licensed and unlicensed subcarriers and power altogether. Here, stochastic optimization problem is solved by using the SCA method by exploiting the Lyapunov technique, thus developing an online energy-aware optimal algorithm called as a drift-plus-penalty algorithm.

4.10. Game theory

Game theory is a powerful mathematical framework used to solve the resource allocation problem in UDNs. Typically, the game model incorporates a group of players (decision-makers) who provide recommendations (actions) to enhance the performance measure function (e.g., throughput, spectrum efficiency, energy efficiency, network capacity, etc.). So, in this section, a survey of certain resource allocation schemes has been entailed which were executed based on game-theoretic approaches.

To compensate abrupt cell outage and enable self-healing in small cell networks, a mixed-integer non-convex optimization problem is solved by the coalition game, which can increase the total payoff without a reduction in payoffs for the individual players'. So, a coalition game based resource allocation algorithm has been proposed in [146], which manages the network failure and improves network performance in terms of network capacity as well as user fairness. A non-cooperative game model has been introduced in which SBSs act as homoegualis agents that could predict a sequence of future actions, and hence can improve weighted fairness and throughput [32].

The interference is the main problem in future wireless networks. In ultra dense LTE-U system scenario, a one to one matching game theory has been utilized in which small cell user compete to get matched with the chunk channels. Hence, it significantly reduces the complexity of the system and increases total network utility by controlling interference and traffic congestion at the Wi-Fi access point [48]. Along with interference control, some work focused on the energy efficiency parameter since it is essential to utilize energy-efficient techniques in HetNets for prolonging battery life and reducing power consumption in the network. A two-level dynamic energy-efficient resource allocation scheme based on a non-cooperative game-theoretic approach has been proposed which reduces the intercell-interference and maximizes the network efficiency. So, a distributed iterative algorithm based on the fixed-point theorem is formulated to attain the equilibrium of the game in the HetNets [121]. Additionally, the power allocation problem to maximize energy efficiency is transformed into a two-stage Stackelberg game, where MBS is a follower and all the SBSs are leaders [142].

Table 7 summarizes comparison of all the techniques exploited in resource allocation process. The comparison is done in terms of classification, examples, advantages, and complexity analysis.

5. Emerging technologies for ultra dense networks

In this section, we highlight a number of emerging technologies including network function virtualization (NFV), network slicing and mobile edge computing (MEC) for 5G and beyond networks, which act as turning wheel for the development cycle of UDNs.

5.1. Network function virtualization

With the increase in the number of wireless services and applications, network virtualization (NV) has emerged as a promising technology for ultra dense wireless networks. NV provides decoupling the control and data planes, and hence network management is easy. By using NV technology, diverse network functions can be performed on the same hardware platform by a software defined network (SDN). This will help many virtual networks to perform on a unified infrastructure with no interference due to the presence of enough isolation among virtual units. Following this method, traditional networks can possibly be fragmented into many separated virtual network units that coexist on the same environment with an advantage of minimal interference because of isolated networks. It makes the resource allocation process more flexible and greatly reduces energy consumption. Moreover, due to the programmability, customization, and spectrum virtualization and sharing of the virtual networks, various network technologies and services can be integrated without changing physical infrastructure characteristics and their interfaces.

The current researches in NV mainly focus on virtualization in RAN networks. Various issues of existing traditional networks including uncertainty in traffic and information exchange in the UDN can be effectively overcome by using NV with C-RANs. To this end, authors in [184] have proposed a scheme that allows spectrum sharing among various mobile network operators and RRH by mitigating inter-tier interference. Then resource allocation is performed by using the binary integer programming method to maximize network throughput and minimize the delay. In another study based on the use of virtualization, the authors have used coordinated multipoint transmission/reception and SDN technologies to form virtualized BS on per cell or user basis by assigning virtualized resources on demand [185]. Therefore, the cross-layer framework for virtualized C-RANs improves throughput, reduces handover, and energy consumption.

5.2. Network slicing

Due to the rise of cloud dependency, a surge in data traffic of 5G and beyond networks is needed with very vast and mega requirements. These networks are envisaged to provide multiple services with various characteristics. The services can be categorized based on their specifications such as massive machine type communications, enhanced mobile broadband, and ultra-reliable low latency communication (URLLC). To support these multi-services of 5G and beyond networks, network slicing technology is needed. In network slicing, the multiple dedicated virtual networks, known as network slices could operate on the same physical network infrastructure [186]. For the implementation of network slicing the SDN and NFV are the key technologies that are needed. A network slice has several virtual network functions (VNFs) that give flexibility, scalability, and programmable network service with reduced capital and operational expenditures by proficiently controlling and managing VNFs. On the other side, there are a number of challenges in terms of isolation, elasticity, end-to-end coordination, and customization. For instance, in the process of resource isolation where sharing and isolation of radio resources is not that easy due to the varying communication environment, and hence results in low multiplexing gains [187].

Concerning network slicing, current researches broadly focus on many aspects like slice association, reconfiguration, creation, and activation. An association relationship among user-BS-network slicing has been established in reference [188], so that the slice users whose QoS need to be satisfied are identified. After that slice association and bandwidth allocation has been done for the identified users to minimize bandwidth consumption. Matching theory has been used in [189] to find the user association, and the isolation character of slicing is used to mitigate co-tier and cross-tier interferences. This scheme improves the QoS and energy efficiency of users and eliminates interferences. In network slicing, to configure the challenge of isolation and to improve network performance during varying traffic conditions, there is a need to reconfigure network slice adaptively. In this regard, the deep reinforcement learning-based method has been used to reconfigure the core network to get maximization in bandwidth utilization [190]. Further, during network slicing, one of the main concerns is to minimize slice setup time. In this regard, a metaheuristic genetic algorithm for scheduling the VNFs to optimize service creation time has been presented in reference [191]. In brief, network slicing is a key enabler for ultra dense next-generation networks and could solve the problem of resource allocation and optimization concerning diverse application domains.

5.3. Mobile edge computing

The evolution of UDN has exhibited an inexorable trend. Many important challenges also appear, like increased energy consumption, user traffic, resource requirement demand, and resource allocation complexity. To solve these challenges MEC is a promising technique that allows computation intensive task to be executed at the edge of the network. Hence, MEC provides ultra-low latency, high bandwidth, and real time network access to resource limited mobile devices. Moreover, it reduces energy consumption, network load, and cost.

Concerning MEC, current researches broadly focus on the combination of mobile computing and wireless communication, resulting in a broad range of techniques for task offloading and network resource allocation. Authors have proposed channel resource allocation by using a differential evolution algorithm to optimize energy consumption under the constraint of delay [192]. The scheme takes efficient offloading decisions for task execution

Table 7
Comparative analysis of different resource allocation techniques.

Technique	Definition	Classification	Examples of the applicable problem	Complexity analysis and other advantages
Convex optimization	• Convex optimization is a branch of mathematical optimization in which the resource allocation problem is solved as a function of some maxima or minima by considering the number of constraints	Distributed/ centralized/ deterministic/ probabilistic/ stochastic	• A convex optimization technique is used when the resource allocation problem is convex or non-convex (e.g., [52,54,65,66,85,89,90,140,142, 149,152,154,158–162])	• A convex and non-convex problem can be solved by using Lagrange duality method, SCA, greedy algorithm, distributed methods such as ADMM, etc., which reduces the computational complexity
Combinatorial optimization	• The resource allocation problem is solved to find optimal solutions within a finite set of possible solutions	Distributed/ centralized/ deterministic/ probabilistic/ stochastic	• It can be used for resource allocation of a discrete optimization type problems (e.g., [57,84,154])	• A mixed combinatorial and convex resource allocation problem can be solved, and hence this dual perspective can lead to structure insights and better algorithms to reduce computational complexity
Stochastic geometry	• Allows study of random network structure by modeling and analyzing the average behavior over many spatial realizations of the system	Distributed/ centralized/ stochastic	• Stochastic geometry processes like HPPPs, HCPP, PCP, etc., are used to analyze network performance due to the randomly distributed users (e.g., [87,145,148])	• Resource allocation problem in random distributed network is solved by stochastic geometry methods helps to reduce the complexity of the network
Sparse optimization	• Based on compressive sensing theory and composite minimization framework the sparse feature is exploited to solve the resource allocation problem	Distributed/ centralized/ deterministic/ stochastic	• In the resource allocation problem, the nodes can be grouped as a set of sparse or nearly sparse coefficients to obtain a simple optimal solution (e.g., [93,95,137,163,164])	• The use of sparse optimization in the wireless cellular network helps in compressing the signal, minimizing the power consumption, and computational complexity, etc.
Graph theory	• Graph theory consists of nodes connected by edges and used to model the interactions in the network.	Distributed/ centralized/ deterministic	• The resource allocation problem is solved by a vertex-coloring graph, conflict graph, and bipartite graph models (e.g., [85,125,155,165–169])	• By using global scheduler and distributed optimization methods a low complex optimal solution is achieved
Machine learning	• Learn network patterns, user demands, user behaviors, resource usage, etc., by interacting with the environment and then make predictions or decisions	Distributed/ centralized/ deterministic/ stochastic	• Online learning, Q-learning, actor-critic algorithm, RL-LSTM cells based deep reinforcement learning algorithm are adopted to solve the resource allocation (e.g., [32,59,80,120,124,126,150, 181,182])	• Machine learning techniques results in efficient solutions to deal with complex problems with significant computational complexity • Moreover, it is model-free learning which facilitates its usage in a dynamic network
Clustering method	• The resource allocation process is performed separately by grouping the users into different clusters based on certain criteria	Distributed/ centralized/ deterministic/ stochastic	• K-means class clustering, game theory and graph theory-based clustering methods are efficiently used (e.g., [68,85,147,158,164,168, 170–176,183])	• The clustering technique is useful to tackle highly computational complexity and unsupervised learning based problems by dividing the larger number of users into a group of clusters • Hence, the complexity and exchange of overhead reduce
Auction theory	• Auctions theory deal with the concept of buying and selling service to maximize revenue in terms of cost, spectrum efficiency, etc.	Distributed/ centralized/ deterministic/ stochastic	• VCG auction, Vickrey's auction, combinatorial auction are widely used in the resource allocation process (e.g., [71,177–180])	• Auction theory works efficiently in the following environment: • When the resource allocation process is needed to be done in a decentralized manner • To match dynamic patterns of demand and supply • Operating when limited or no network state and utility information is available
Stochastic optimization	• The resource allocation problem is formulated as a minimize or maximize objective function in the presence of randomness in the optimization process	Distributed/ centralized/ stochastic	• The resource allocation problems which deals with time dependent queue state model are solved by using stochastic optimization. (e.g., [37,67])	• Used in the uncertain environment scenario where deterministic optimization does not work
Game theory	• The game theory incorporates a set of decision-makers who make the actions to maximize the performance measure function	Distributed/ centralized/ deterministic/ stochastic	• In the resource allocation problem, the game-theoretic technique is used by players to arrive at an optimal strategy (e.g., [32,48,111,121,142,146])	• Game theory explains the concept of bargaining and coalition-formation, and hence suitable in the situation of conflicting interests

either locally or on the edge server. However, a single edge server environment has been considered, which is the main drawback of this study. In a multi-edge server environment, SDN based framework for MEC has been proposed which considers the load at the edge server while offloading the task [193]. In the UDN for ultra-low latency applications, it is important to minimize delay for both uplink and downlink transmission. In this regard, a joint uplink and downlink task offloading and resource allocation method serves as an aid that considers transmission overhead and computation load at the edge server [194]. In an effort to optimize the beamforming vectors at the users and beamforming matrices at the BS, a mmWave MEC resource allocation and beamforming algorithm has been proposed that involves computation offloading and resource allocation at the edge server to minimize overall network latency [195].

6. Challenges and open research direction

There are many research gaps and challenges in the area of resource allocation for UDNs that need to be addressed. The main challenges and open research direction are summarized in the following:

6.1. Density planning

In today's scenario, the wireless cellular network considers different types of cell structures (e.g., femto, macro, micro, etc.) and users that operate on different frequencies. There is a different level of data traffic requirement for a divergent number of applications and different data rate requirements for each user. So, the main question arises that how to design an efficient density plan i.e., the number of different types of users and cells in the network system is the critical challenge for future cellular network planning. Mainly, the efficient density design of the network greatly depends on the data rate requirement of users, which is highly dynamic in nature, and hence requires prediction. For instance, in the ultra dense LTE-U SBSs system, two main problems arise i.e., co-channel interference between licensed and unlicensed spectrum and the increase in collision probability of Wi-Fi systems during LTE-U transmission. Furthermore, in ultra dense CRNs with the increasing mobile data traffic, the issue of spectral congestion has also become unavoidable. These problems can be solved by using stochastic geometry tools, game-theoretic approaches, etc.

6.2. Overhead and delay analysis

In UDNs, various critical challenges should be addressed for the resource decision making process. Mostly, in future cellular networks, resource allocation decision is taken when accurate CSI is present. Although, the overhead to obtain complete CSI while considering dense implementation is always high. So, the solution to the problem is to investigate appropriate probabilistic resource allocation methods with partial CSI or only user position or distance information. Another solution is to use a distributed approach over a centralized approach. As this approach-based algorithms reduce the delay of exchanging information and the amount of overhead in the system. Moreover, other emerging methods such as large network decompositions (clustering approach), AI, deep learning, etc., may promise marvelous solutions for the future wireless cellular system.

6.3. Performance parameter

In future wireless cellular networks, QoS of users can be measured on distinct performance indicators, such as latency, cost, and reliability. Although, most prevalent research mainly concentrates on relevant indicators such as network utility, interference management, fairness, spectrum efficiency, and minimum data rate requirement with the perspective of resource allocation. Moreover, there is an inherent tradeoff between the data rate, cost, and reliability. For instance, as nominal cost usually results in more unlicensed spectrum resources, the data rate and reliability of the network will sharply reduce because the unlicensed spectrum is less reliable than a licensed spectrum. Hence, by achieving the right balance among these performance parameters to obtain QoS in UDNs is an important research area to explore. Moreover, the new emerging technologies, such as VR, caching, SDN, energy harvesting, MEC, non-orthogonal multiple access, AI, telehealth, etc., need higher QoS in the network that demands extremely low latency, ultra-high reliability, security, and privacy, etc. These performance parameters put forward critical requirements for both hardware structure and resource management policies.

6.4. Network mobility

Earlier most of the work considers static environment whereas network mobility is greatly ignored in literature so the effect of users' mobility on the performance of the resource allocation methods can be investigated more deeply in the UDNs. For instance, in the mobility-based wireless network model, the SUs could be mobile during the entire process, and hence can change their space, frequency, and time coordinates. During the process, the SUs may be nearer to the PU, which could result in interference, and hence limit the performance of PU transmission. Therefore, the issue of network mobility and its influence on the optimality of the resource allocation method needs further research. In the case of the future wireless cellular network environment, the new mobility management aware protocols are needed to assure high flexibility, high availability, high reliability, and resource efficiency. Therefore, realistic user and efficient traffic distribution schemes are required for maintaining the reliability of the communication process.

6.5. Security

The security issues play a vital role in UDNs, so obtaining a secure communication is an important design objective for reliable wireless communications. The dynamic and time varying nature of wireless channel makes it more agreeable for the untrusted intruders' eavesdropping as long as they are in the transmitters range. Thus, the security performance is a pivotal challenge that require efficient techniques to ensure the data safety in wireless communication. Mainly, there exist two approaches to secure the information from the malicious intruders: upperlayer encryption and physical layer security. In upperlayer encryption approach the computation cost is high due to the use of encryption/decryption methods. Whereas in physical layer security network secrecy rate performance is analyzed. For instance, in ultra dense HetNets scenario number of tier (pico, femto, macro etc.) are available for different users. So, to provide secure multi-tier communication the main challenge is to investigate efficient strategy to properly provide user access to multiple tiers. In this direction, the precoder design are discussed in literature to enhance the network secrecy rate.

For instance, if we consider the ultra dense CRNs scenario, the uniqueness of the CR technology makes these networks more vulnerable to security threats in comparison to other traditional cellular wireless communication system. On account of the property of a CR device which enables it to sense a broad range of frequency systems, eavesdropping in CRN has mainly become candid for a user with malign intent. Hence, it is critical to develop the resource allocation techniques that tackle this issue effectively by maximizing the secrecy rate. Most of the works in literature considered the secure resource allocation problems which either provide PUs' security or SUs' security, not both. Hence, the design of resource allocation techniques that focus on providing secure communications for both PUs and SUs simultaneously along with encrypted approach is the need of the hour.

6.6. Experimental testbeds

The real-time testing of the wireless system to check its performance is an essential part of any work invention. This would spotlight on practical realization issues of wireless cellular communication that need further exploration and require to be solved to have real-time practical usage. So, designing experimental testbeds can provide an efficient tool for investigators to experimentally approve their logical results in a real-time scenario.

These testbeds are essential, and hence encouraged to perform rigorous, transparent, and replicable research process. Thus, the simulation of the resource allocation algorithms on the experimental testbeds (system level) is the foremost important step. It will enhance the accuracy of obtained analytical results by addressing factors such as real channel, mobility models, inter-cell and intra-cell interference, etc. The various platforms can be used to design experimental testbeds. For instance, in [196] a testbed using LabVIEW communications system design suite software to universal software radio peripheral (USRP-2954) is exploited to analyze the real implantation issues of the resource allocation process in CR networks. Similarly, in reference [197] performance of decentralized cloud assisted cooperative multi-agent reinforcement learning based spectrum allocation scheme is analyzed on USRP-2954R by using LabVIEW design suite software.

6.7. Channel models

Mostly the proposed resource allocation methods in a wireless cellular network are based on a Rayleigh fading channel model. However, some works considered the Rician channel model, Jakes' channel model, and Clarkes' model. So, it becomes essential to evaluate the prevalent methods and introduce fresh methods for various channel models according to the channel characteristics (i.e. LoS, non-line of sight (NLoS)). Particularly, NLoS transmission occurs when there is no visual LoS between transmitting and receiving end. It is common in office environments and central business districts. Whereas, in the case of a LoS transmission quite little distance exist between a receiver and a transmitter. So, Rician fading channels would be more appropriate in such environment. Since, for the present scenarios of the wireless system, the UDNs architecture is more prevailing. In this network, the distance between transmitter and receiver decreases due to the densification of small cells, and hence chances of a LoS transmission increases. Therefore, more advanced propagation models including both LoS and NLoS transmissions would be more relevant, which will affect the analytical performance and the optimality of the resource allocation method.

6.8. Backhaul/fronthaul constraints design

The network elements are connected to core networks through a link known as backhaul. This link should satisfy the extreme requirements in terms of capacity, latency, availability, energy, and cost efficiency in order to capture the diverse performance aspects of the network. Due to the requirement of additional computing resources as well as efficient backhaul design, the innovatory resource allocation methods that includes the newly emerging techniques such as SDN, C-RAN intelligence, software optimizing network, and caching capabilities are needed in the wireless cellular network design. These technologies are building block to backhaul and fronthaul management for the infrastructure sharing in the ultra dense heterogeneous network structure, hence leading to increased energy and cost efficiency. However, an SDN framework may results in the issue of network security; a challenge that should be resolved without affecting flexibility, reliability, and adaptability of the network. Another challenge exist is the design of efficient compression methods for fronthaul and backhaul links in a network. To address this, it is required to find the effects of the latency of the links on the performance of the network. Furthermore most of the work in literature consider unified backhaul allocation [159]. However, it is important to investigate optimal non unified backhaul bandwidth allocation since in heterogeneous environment link latency and capacities are different. Moreover, based on the traffic load, network topology, and user location, the link should be re-configurable and adaptable.

6.9. Green communication

In future wireless cellular communications, except energy efficiency all other requirements like throughput, capacity, connectivity, data rate, etc., are increasing with speed ranging from 100 to 1000 times. In the present-day scenario, the target to achieve energy efficiency gain in the densified network is more critical. Hence, there is need to modify and analyze the present energy efficient algorithm so that the energy efficient gain could increase dramatically. The modification should be done in such a way that it should support heterogeneous applications along with throughput, delay, and capacity requirements. The key design methods in this regard is deployment of virtualized network framework, cloud and cooperative radio networks. Furthermore, combining energy harvesting techniques with cellular communication system improve energy efficiency. This can be achieved by using renewable energy resources or wireless power transfer techniques in which nodes can charge their batteries by utilizing electromagnetic energy. Furthermore, massive-MIMO, mmWave, and beamforming technologies also provide promising solution to increase energy efficiency, which need further investigation.

6.10. Ultra-reliable low latency communication

The prime requirement of URLLC is to achieve reliability and tactile internet delay. URLLC points at delivering fast data transferring services with guaranteed latency less than 1 ms. To achieve such a stringent delay is a great challenge because the busty URLLC requests can result in heavy network traffic or congestion in the random-access phase. Furthermore, in the densification process as the number of BSs increases, the transmission and scheduling delays increases. The transmission and scheduling delays can directly affect real time processing capabilities in the network; hence it is essential to look over these parameters. The edge computing, cooperative cloud edge processing, and proactive caching are the recent emerging technologies used by researchers in order to get ultra-low latency and very high reliability in the network. Additionally, cooperative cloud edge processing technique combine the benefits of both edge and cloud technologies, and hence enable real time processing of tactile applications along with handling large amount of data and user traffic at the cloud server.

6.11. Interference

The future wireless networks are moving towards the technology known as densification, and hence the distance between the users and interfering nodes/APs has reduced. So, when considering resource allocation problem, this deployment has produced stronger interference sources, and hence results in co-tier and cross-tier interference which affects the network efficiency and QoS requirements. Therefore, in order to control interference future research should aim at designing efficient algorithms for power control and cell association in multi-tier structure, support user association to different BSs simultaneously, and analyze trade-off between reduced interference level and enhanced resource usage, etc.

7. Conclusion

It is anticipated that 5G and beyond wireless networks will depend on ultra dense network technology, to counter the evergrowing demand for network capacity and data rate requirement, and hence provide the answer to the issue of spectrum scarcity and resource management. In this article, a detailed review of resource allocation issues in various layouts of ultra dense network has been provided. Particularly, it provides insight details about the resource allocation process explored in the latest technologies of future wireless networks. Besides this, the terminologies for prevalent resource allocation schemes have been detailed from the outlook of approaches, methods, and optimization criteria. Continuing further, a categorization of current resource allocation techniques has been provided, which incorporates convex optimization, combinatorial optimization, stochastic geometry-based methods, stochastic optimization, sparse optimization, graph theory, machine learning-based methods, clustering approach, game theory, and auction theory. These techniques underline their operational variation with their respective design principle. Based on the different performance metrics yardsticks and implemented design principles, the resource allocation techniques have been studied. Finally, on the grounds of current research endeavours and other leading technologies, the article ends with a discussion of many budding challenges and opens new research horizons.

Unluckily, existing resource allocation techniques for ultra dense networks still need the adequate reflection of different network parameters with comprehensive research for implementation in real environment. So, future research in this direction should focus on designing the self-organizing intelligent resource allocation techniques that should truly be beneficial in time varying dynamic environment network conditions. Therefore, the resource allocation in ultra dense networks is still a challenging area which requires the attention of the researchers. Hence, our far-reaching review on resource allocation in the 5G and beyond ultra dense wireless network presented in this article will conceivably shed light on important references and specifications for further in-deep investigation in this area.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

[1] B.U. Kazi, G.A. Wainer, Next generation wireless cellular networks: ultra-dense multi-tier and multi-cell cooperation perspective, Wirel. Netw. 25 (4) (2019) 2041–2064.
[2] Y. Gu, Y. Wang, Q. Cui, A stochastic optimization framework for adaptive spectrum access and power allocation in licensed-assisted access networks, IEEE Access 5 (2017) 16484–16494.
[3] X. Xiao, X. Tao, J. Lu, Energy-efficient resource allocation in LTE-based MIMO-OFDMA systems with user rate constraints, IEEE Trans. Veh. Technol. 64 (1) (2015) 185–197.
[4] A.R. Utami, Iskandar, Resource allocation analysis with genetic algorithm in LTE MIMO-OFDMA cellular system, in: Proc. 13th International Conference on Telecommunication Systems, Services, and Applications, 2019, pp. 182–185.
[5] W. Fu, Q. Kong, Y. Zhang, X. Yan, A resource scheduling algorithm based on carrier weight in LTE-Advanced system with Carrier Aggregation, in: Proc. Wireless and Optical Communications Conference, 2013, pp. 1–5.
[6] K. Kumar, A. Prakash, R. Tripathi, A spectrum handoff scheme for optimal network selection in cognitive radio vehicular networks: A game theoretic auction theory approach, Phys. Commun. 24 (2017) 19–33.
[7] K. Kumar, A. Prakash, R. Tripathi, Spectrum handoff in cognitive radio networks: A classification and comprehensive survey, J. Netw. Comput. Appl. 61 (2016) 161–188.
[8] A. Kumar, K. Kumar, Multiple access schemes for cognitive radio netwoks: A survey, Phys. Commun. 38 (2020) 100953.
[9] W. Ejaz, S.K. Sharma, S. Saadat, M. Naeem, A. Anpalagan, N.A. Chughtai, A comprehensive survey on resource allocation for CRAN in 5g and beyond networks, J. Netw. Comput. Appl. 160 (2020) 102638.
[10] M.K. Pedhadiya, R.K. Jha, H.G. Bhatt, Device to device communication: A survey, J. Netw. Comput. Appl. 129 (2019) 71–89.
[11] X. Wang others, Millimeter wave communication: A comprehensive survey, IEEE Commun. Surv. Tutor. 20 (3) (2018) 1616–1653.
[12] M.A. Adedoyin, O.E. Falowo, Combination of ultra-dense networks and other 5g enabling technologies: A survey, IEEE Access 8 (2020) 22893–22932.
[13] H. Zhang, Y. Dong, J. Cheng, J. Hossain, V.C.M. Leung, Fronthauling for 5g LTE-u ultra dense cloud small cell networks, IEEE Wirel. Commun. 23 (6) (2016) 48–53.
[14] M. Kamel, W. Hamouda, A. Youssef, Ultra-dense networks: A survey, IEEE Commun. Surv. Tutor. 18 (4) (2019) 2522–2545.
[15] Yunas S. Fahad, M. Valkama, J. Niemelä, Spectral and energy efficiency of ultra-dense networks under different deployment strategies, IEEE Commun. Mag. 53 (1) (2015) 90–100, IEEE.
[16] W. Su, J.D. Matyjas, S. Batalama, Active cooperation between primary users and cognitive radio users in heterogeneous ad-hoc networks, IEEE Trans. Signal Process. 60 (4) (2012) 1796–1805.
[17] I. Ahmad, Z. Wei, Z. Feng, Y. Bai, Q. Zhang, P. Zhang, Joint price and power allocation under interference constraint for dynamic spectrum access networks, in: Proc. IEEE International Symposium on Dynamic Spectrum Access Networks, 2014, pp. 141–144.
[18] W. Zong, S. Shao, Q. Meng, W. Zhu, Joint user scheduling and beamforming for underlay cognitive radio systems, in: Proc. 15th Asia-Pacific Conference on Communications, 2009, pp. 99–103.
[19] Nidhi, A. Mihovska, Small cell deployment challenges in ultradense networks: Architecture and resource management, in: proc. 12th International Symposium on Communication Systems, Networks and Digital Signal Processing, 2020, pp. 1–6.
[20] O. Alamu, A. Gbenga-Ilori, M. Adelabu, A. Imoize, O. Ladipo, Energy efficiency techniques in ultra-dense wireless heterogeneous networks: An overview and outlook, Eng. Sci. Technol. Int. J. 23 (6) (2020) 1308–1326.
[21] A. Mukherjee others, Licensed-assisted access LTE: coexistence with IEEE 802.11 and the evolution toward 5g, IEEE Commun. Mag. 54 (6) (2016) 50–57, IEEE.
[22] M. Kamel, W. Hamouda, Ultra-dense networks: A survey, IEEE Commun. Surv. Tutor. 18 (4) (2019) 2522–2545.
[23] G. Chopra, R. Kumar Jha, S. Jain, A survey on ultra-dense network and emerging technologies: Security challenges and possible solutions, J. Netw. Comput. Appl. 95 (2017) 54–78.
[24] R. Liu, Q. Chen, G. Yu, G.Y. Li, Z. Ding, Resource management in LTE-u systems: Past, present, and future, IEEE Open J. Veh. Technol. 1 (2019) (2019) 1–17.
[25] G.I. Tsiropoulos, O.A. Dobre, M.H. Ahmed, K.E. Baddour, Radio resource allocation techniques for efficient spectrum access in cognitive radio networks, IEEE Commun. Surv. Tutor. 18 (1) (2016) 824–847.
[26] M. El Tanab, W. Hamouda, Resource allocation for underlay cognitive radio networks: A survey, IEEE Commun. Surv. Tutor. 19 (2) (2017) 1249–1276.
[27] M. Peng, C. Wang, J. Li, H. Xiang, V. Lau, Recent advances in underlay heterogeneous networks: interference control, resource allocation, and self-organization, IEEE Commun. Surv. Tutor. 17 (2) (2015) 700–729.
[28] Y. Teng, M. Liu, F.R. Yu, V.C.M. Leung, M. Song, Y. Zhang, Resource allocation for ultra-dense networks: A survey, some research issues and challenges, IEEE Commun. Surv. Tutor. 21 (3) (2019) 2134–2168.
[29] W. Yu, H. Xu, H. Zhang, D. Griffith, N. Golmie, Ultra-dense networks: Survey of state of the art and future directions, in: Proc. 25th International Conference on Computer Communications and Networks, 2016, pp. 1–10.
[30] H. He, H. Shan, A. Huang, L.X. Cai, T.Q.S. Quek, Proportional fairness-based resource allocation for LTE-u coexisting with wi-fi, IEEE Access 5 (2017) 4720–4731.
[31] M. Chen, W. Saad, C. Yin, Optimized uplink-downlink decoupling in LTE-U networks: An echo state approach, in: Proc. IEEE ICC Mobile and Wireless Networking Symposium, 2016, pp. 1–6.
[32] U. Challita, L. Dong, W. Saad, Proactive resource management for LTE in unlicensed spectrum: A deep learning perspective, IEEE Trans. Wirel. Commun. 17 (7) (2018) 4674–4689.
[33] Z. Zhou, D. Guo, M.L. Honig, Allocation of licensed and unlicensed spectrum in heterogeneous networks, in: Proc. Australian Communications Theory Workshop, 2016, pp. 59–64.
[34] Z. Zhou, D. Guo, M.L. Honig, Licensed and unlicensed spectrum allocation in heterogeneous networks, IEEE Trans. Commun. 65 (4) (2017) 1815–1827.
[35] T. Zhang, J. Zhao, Y. Chen, Hidden node aware resource allocation in licensed-assisted access systems, in: Proc. IEEE Global Communications Conference, 2017, pp. 1–6.
[36] D. Lopez-perez, X. Chu, A.V. Vasilakos, H. Claussen, On distributed and coordinated resource allocation for interference mitigation in self-organizing LTE networks, IEEE/ACM Trans. Netw. 21 (4) (2013) 1145–1158.
[37] Y. Gu, Q. Cui, Y. Wang, S. Soleimani, Energy-Aware adaptive spectrum access and power allocation in LAA networks via Lyapunov optimization, in: Proc. IEEE Vehicular Technology Conference, 2017, pp. 1–6.
[38] Y.Y. Liu, S.-J. Yoo, Dynamic resource allocation using reinforcement learning for LTE-U and WiFi in the unlicensed spectrum, in: Proc. International Conference on Ubiquitous and Future Networks, 2017, pp. 471–475.

[39] H.V. Lima, A.S. Abdallah, E.F. Bueno, K.V. Cardoso, A stochastic programming model for resource allocation in LTE-U networks, in: Proc. IEEE Canadian Conference on Electrical and Computer Engineering, 2018, pp. 4–8.
[40] J. Dai, C. Shen, Adaptive resource allocation for LTE/WiFi coexistence in the unlicensed spectrum, in: Proc. International Conference on Computing, Networking and Communications, 2018, pp. 457–462.
[41] S. Liu, R. Yin, G. Yu, Hybrid adaptive channel access for LTE-u systems, IEEE Trans. Veh. Technol. 68 (10) (2019) 9820–9832.
[42] Y. Hu, R. Mackenzie, M. Hao, Expected Q-learning for Self-organizing resource allocation in LTE-U with downlink-uplink decoupling, in: Proc. 23rd European Wireless Conference., 2017, pp. 185–190.
[43] R. Yin, G. Yu, A. Maaref, G.Y. Li, LBT-Based adaptive channel access for LTE-u systems, IEEE Trans. Wirel. Commun. 15 (10) (2016) 6585–6597.
[44] Y. Gao, Y. Wu, H. Hu, X. Chu, J. Zhang, Licensed and unlicensed bands allocation for cellular users: A matching-based approach, IEEE Wirel. Commun. Lett. 8 (3) (2019) 969–972.
[45] Y. Xu, R. Yin, Q. Chen, G. Yu, Joint licensed and unlicensed spectrum allocation for unliencensed LTE, in: Proc. IEEE International Symposium on Personal, Indoor and Mobile Radio Communications, 2015, pp. 1912–1917.
[46] Y. Xu, R. Yin, Q. Chen, G. Yu, Joint licensed and unlicensed spectrum allocation for unlicensed LTE, in: Proc. IEEE 26th International Symposium on Personal, Indoor and Mobile Radio Communications, 2015, pp. 1912–1917.
[47] X. Zhen, H. Shan, G. Yu, Y. Cheng, L.X. Cai, A. Huang, Joint resource allocation for LTE over licensed and unlicensed spectrum, in: Proc. IEEE Vehicular Technology Conference, 2017, pp. 1–6.
[48] T. Leanh, N.H. Tran, D.T. Ngo, Z. Han, C.S. Hong, Orchestrating resource management in LTE-unlicensed systems with backhaul link constraints, IEEE Trans. Wirel. Commun. 18 (2) (2019) 1360–1375.
[49] Y. Gu, Y. Zhang, L.X. Cai, M. Pan, L. Song, Z. Han, Exploiting student-project allocation matching for spectrum sharing in LTE-unlicensed, in: Proc. IEEE Global Communications Conference, 2015, pp. 1–6.
[50] Y. Gu, C. Jiang, L.X. Cai, M. Pan, L. Song, Z. Han, Dynamic path to stability in LTE-unlicensed with user mobility: A matching framework, IEEE Trans. Wirel. Commun. 16 (7) (2017) 4547–4561.
[51] P. Li, S. Guo, W. Zhuang, B. Ye, On efficient resource allocation for cognitive and cooperative communications, IEEE J. Sel. Areas Commun. 32 (2) (2014) 264–273.
[52] W. Chen, W. Lin, H. Tsao, C. Lin, Probabilistic power allocation for cognitive radio networks with outage constraints and one-bit side information, IEEE Trans. Signal Process. 64 (4) (2016) 867–881.
[53] S. Dwivedi, O.P. Meena, An improved resource allocation algorithm for OFDM overlay cooperative cognitive radio network, in: Proc. International Conference on Recent Innovations in Signal processing and Embedded Systems, 2017, pp. 75–77.
[54] Z. Ali, G.S.A. Sidhu, M. Waqas, L. Xing, F. Gao, A joint optimization framework for energy harvesting based cooperative CR networks, IEEE Trans. Cogn. Commun. Netw. 5 (2) (2019) 452–462.
[55] Y. Jiao, I. Joe, Energy-efficient resource allocation for heterogeneous cognitive radio network based on two-tier crossover genetic algorithm, J. Commun. Netw. 18 (1) (2016) 112–122.
[56] R. Xie, F.R. Yu, H. Ji, Y. Li, Energy-efficient resource allocation for heterogeneous cognitive radio networks with femtocells, IEEE Trans. Wirel. Commun. 11 (11) (2012) 3910–3920.
[57] X. Zhang, X. Zhang, L. Han, Utilization-oriented spectrum allocation in an underlay cognitive radio network, IEEE Access 6 (2018) 12905–12912.
[58] D. Das, S. Das, Intelligent resource allocation scheme for the cognitive radio network in the presence of primary user emulation attack, IET Commun. 11 (15) (2017) 2370–2379.
[59] H. Chen, M. Zhou, L. Xie, K. Wang, J. Li, Joint spectrum sensing and resource allocation scheme in cognitive radio networks with spectrum sensing data falsification attack, IEEE Trans. Veh. Technol. 65 (11) (2016) 9181–9191.
[60] N. Mokari, S. Parsaeefard, H. Saeedi, P. Azmi, Cooperative secure resource allocation in cognitive radio networks with guaranteed secrecy rate for primary users, IEEE Trans. Wirel. Commun. 13 (2) (2014) 1058–1073.
[61] S. Zhang, X. Zhao, Distributed power allocation based on LQG regulator with adaptive weight and switching scheme for cognitive radio networks, IEEE Access 6 (2018) 39180–39196.
[62] S. Zhang, X. Zhao, Distributed power allocation for cognitive radio networks with time varying channel and delay: $H\infty$ state feedback control approach, IEEE Access 6 (2018) 56893–56910.
[63] N.U. Hassan, M. Assaad, H. Tembine, Distributed $h\infty$-based power control in a dynamic wireless network environment, IEEE Commun. Lett. 17 (6) (2013) 1124–1127.
[64] S. Pan, X. Zhao, Y. Liang, Robust power allocation for OFDM-based cognitive radio networks: A switched affine based control approach, IEEE Access 5 (2017) 18778–18792.
[65] S. Mallick, R. Devarajan, Robust resource optimization for cooperative cognitive radio networks with imperfect CSI, IEEE Trans. Wirel. Commun. 14 (2) (2015) 907–920.
[66] Y. Wang, P. Ren, Q. Du, L. Sun, Optimal power allocation for underlay-based cognitive radio networks with primary user's statistical delay qos provisioning, IEEE Trans. Wirel. Commun. 14 (12) (2015) 6896–6910.
[67] Z. Bai, L. Ma, Y. Dong, P. Ma, Energy-efficient resource allocation for secure cognitive radio network with delay qos guarantee, IEEE Syst. J. 13 (3) (2019) 2795–2805.
[68] S.R. Sonti, M. Siva, G. Prasad, Enhanced fuzzy C-means clustering based cooperative spectrum sensing combined with multi-objective resource allocation approach for delay-aware CRNs, IET Commun. 14 (4) (2020) 619–626.
[69] R. Xie, H. Ji, P. Si, Y. Li, Dynamic channel and power allocation in cognitive radio networks supporting heterogeneous services, in: Proc. IEEE Global Telecommunications Conference, 2010, pp. 1–5.
[70] R. Xie, F.R. Yu, H. Ji, Dynamic resource allocation for heterogeneous services in cognitive radio networks with imperfect channel sensing, IEEE Trans. Veh. Technol. 61 (2) (2012) 770–780.
[71] G. Sun others, Coalitional double auction for spatial spectrum allocation in cognitive radio networks, IEEE Trans. Wirel. Commun. 13 (6) (2014) 3196–3206.
[72] T.M. Ho, T. Leanh, S.M.A. Kazmi, C.S. Hong, Opportunistic resource allocation via stochastic network optimization in cognitive radio networks, in: Proc. 16th Asia-Pacific Network Operation and Management Symposium, 2014, pp. 1–4.
[73] H. Xu, B. Li, Resource allocation with flexible channel cooperation in cognitive radio networks, IEEE Trans. Mob. Comput. 12 (5) (2013) 957–970.
[74] A. Zafar, M. Alouini, Y. Chen, R.M. Radaydeh, New resource allocation scheme for cognitive relay networks with opportunistic access, in: Proc. IEEE International Conference on Communications, 2012, pp. 5603–5607.
[75] H. Khan, S.-J. Yoo, Multi-objective optimal Resource allocation using particle swarm optimization in cognitive radio, in: Proc. IEEE 7th International Conference on Communications and Electronics, 2018, pp. 44–48.
[76] M.N. Pavan, S. Kumar, G. Nayak, Interference aware resource allocation (IARA) in cognitive radio networks, in: Proc. 13th IEEE International Conference on Industrial and Information Systems, 2018, pp. 202–206.
[77] W. Lee, W. Huang, Optimal resource allocation in underlay cognitive radio networks based on partial CSI, in: Proc. IEEE International Symposium on Broadband Multimedia Systems and Broadcasting, 2016, pp. 1–4.
[78] S. Wang, Efficient resource allocation algorithm for cognitive OFDM systems, IEEE Commun. Lett. 14 (8) (2010) 725–727.
[79] H.B. Salameh, Efficient resource allocation for multicell heterogeneous cognitive networks with varying spectrum availability, IEEE Trans. Veh. Technol. 65 (8) (2016) 6628–6635.
[80] Y. Wei, F.R. Yu, M. Song, Z. Han, User scheduling and resource allocation in hetnets with hybrid energy supply: An actor-critic reinforcement learning approach, IEEE Trans. Wirel. Commun. 17 (1) (2018) 680–692.
[81] Y. Jiang others, Joint power and bandwidth allocation for energy-efficient heterogeneous cellular networks, IEEE Trans. Commun. 67 (9) (2019) 6168–6178.
[82] H. Liu, S. Xu, K.S. Kwak, Geometric programming based distributed resource allocation in ultra dense Hetnets, in: Proc. IEEE Vehicular Technology Conference, 2016, pp. 1–5.
[83] A. Feki, V. Capdevielle, Autonomous resource allocation for dense LTE networks: A multi armed bandit formulation, in: Proc. IEEE 22nd International Symposium on Personal, Indoor and Mobile Radio Communications, 2011, pp. 66–70.
[84] A. Liu, V.K.N. Lau, Joint.BS-user. association, power. allocation, Joint BS-user association power allocation and user-side interference cancellation in cell-free heterogeneous networks, IEEE Trans. Signal Process. 65 (2) (2017) 335–345.
[85] L. Liu, V. Garcia, L. Tian, Z. Pan, J. Shi, Joint clustering and inter-cell resource allocation for CoMP in ultra dense cellular networks, in: Proc. IEEE ICC Wireless Communications Symposium, 2015, pp. 2560–2564.
[86] N. Trabelsi, C.S. Chen, R. El Azouzi, L. Roullet, E. Altman, User association and resource allocation optimization in LTE cellular networks, IEEE Trans. Netw. Serv. Manag. 14 (2) (2017) 429–440.
[87] F. Parzysz, M. Di Renzo, C. Verikoukis, Power-availability-aware cell association for energy-harvesting small-cell base stations, IEEE Trans. Wirel. Commun. 16 (4) (2017) 2409–2422.
[88] C.C. Coskun, E. Ayanoglu, Energy-spectral efficiency tradeoff for heterogeneous networks with QoS constraints, in: Proc. IEEE ICC Green Communications Systems and Networks Symposium, 2017, pp. 1–7.
[89] Y. Hao, Q. Ni, H. Li, S. Hou, On the energy and spectral efficiency tradeoff in massive MIMO-enabled hetnets with capacity-constrained backhaul links, IEEE Trans. Commun. 65 (11) (2017) 4720–4733.

[90] N. Wang, E. Hossain, V.K. Bhargava, Joint downlink cell association and bandwidth allocation for wireless backhauling in two-tier hetnets with large-scale antenna arrays, IEEE Trans. Wirel. Commun. 15 (5) (2016) 3251–3268.

[91] R.G. Stephen, R. Zhang, Joint millimeter-wave fronthaul and OFDMA resource allocation in ultra-dense CRAN, IEEE Trans. Commun. 65 (3) (2017) 1411–1423.

[92] C. Pan, H. Ren, M. Elkashlan, A. Nallanathan, L. Hanzo, Robust beamforming design for ultra-dense user-centric C-RAN in the face of realistic pilot contamination and limited feedback, IEEE Trans. Wirel. Commun. 18 (2) (2019) 780–795.

[93] J. Tang, W.P. Tay, T.Q.S. Quek, Cross-layer resource allocation eith elastic service scaling in cloud radio access network, IEEE Trans. Wirel. Commun. 14 (9) (2015) 5068–5081.

[94] Y.N.R. Li, J. Li, H. Wu, W. Zhang, Energy efficient small cell operation under ultra dense cloud radio access networks, in: Proc. IEEE Globecom Workshops, 2014, pp. 1120–1125.

[95] Y. Shi, J. Zhang, K.B. Letaief, Group sparse beamforming for green cloud-RAN, IEEE Trans. Wirel. Commun. 13 (5) (2014) 2809–2823.

[96] S.-H. Park, O. Simeone, S.S. Shitz, Joint optimization of cloud and edge processing for fog radio access networks, in: Proc. IEEE International Symposium on Information Theory, 2016, pp. 315–319.

[97] M.A. Hasabelnaby, A. Chaaban, End to end rate enhancement in C-RAN using multi pair two way computation, Comput. Res. Repos. (2021) 1–6.

[98] I. Al-Samman, M. Artuso, H. Christiansen, A. Doufexi, M. Beach, A framework for resources allocation in virtualised C-RAN, in: Proc. 27th IEEE International Symposium on Personal, Indoor and Mobile Radio Communications, 2016, pp. 1–7.

[99] A. Alabbasi, M. Berg, C. Cavdar, Delay constrained hybrid CRAN: A functional split optimization framework, in: Proc. IEEE Globecom Workshops, 2019, pp. 1–7.

[100] S.Y. Yerima, K. Al-Begain, End-to-end QoS improvement of HSDPA end-user multi-flow traffic using RAN buffer management, in: Proc. New Technologies, Mobility and Security Conference and Workshops, 2008, pp. 1–5.

[101] N. Reyhanian, H. Farmanbar, Z.Q. Luo, Resource reservation in backhaul and radio access network with uncertain user demands, in: Proc. 21st IEEE international Workshop on Signal Processing Advances in Wireless Communications, 2020, pp. 1–5.

[102] Q. Han, B. Yang, G. Miao, C. Chen, X. Wang, X. Guan, Backhaul-aware user association and resource allocation for energy-constrained hetnets, IEEE Trans. Veh. Technol. 66 (1) (2017) 580–593.

[103] A. Manzoor, D.H. Kim, C.S. Hong, Energy efficient resource allocation in UAV-based heterogeneous networks, in: Proc. 20th Asia-Pacific Network Operations and Management Symposium, 2019, pp. 1–4.

[104] F. Ye, S. Shao, J. Dai, Y. Tian, A Joint resource allocation algorithm in HetNet based on game theory, in: Proc. USNC-URSI Radio Science Meeting (Joint with AP-S Symposium), 2018, pp. 167–168.

[105] W. Alsobhi, A.H. Aghvami, QoS-aware resource allocation of two-tier HetNet: A Q-learning approach, in: Proc. 26th International Conference on Telecommunications, 2019, pp. 330–334.

[106] J. Zhao, Y. Liu, T. Mahmoodi, K.K. Chai, Y. Chen, Z. Han, Resource allocation in cache-enabled CRAN with non-orthogonal multiple access, in: Proc. IEEE International Conference on Communications, 2018, pp. 1–6.

[107] P. Lin, P. Ma, Y. Ma, W. Han, Cost efficient power allocation, user association and energy management in H-CRAN with hybrid energy sources, in: Proc. International Conference on Wireless Communications and Signal Processing, 2019, pp. 1–7.

[108] M. Peng, K. Zhang, J. Jiang, J. Wang, W. Wang, Energy-efficient resource assignment and power allocation in heterogeneous cloud radio access networks, IEEE Trans. Veh. Technol. 64 (11) (2015) 5275–5287.

[109] M. Gerasimenko others, Cooperative radio resource management in heterogeneous cloud radio access networks, IEEE Access 3 (2015) 397–406.

[110] H.W. Tseng, Y.J. Yu, B.-S. Wu, C.-F. Kuo, P.-S. Chen, A resource allocation scheme for device-To-device communication over ultra-dense 5G cellular networks, in: Proc. IEEE International Conference on Applied System Innovation, 2017, pp. 80–83.

[111] S. Dominic, L. Jacob, Fully distributed joint resource allocation in ultra-dense D2d networks: A utility-based learning approach, IET Commun. 12 (19) (2018) 2393–2400.

[112] A. Algedir, H.H. Refai, Adaptive D2D resources allocation underlaying (2-tier) heterogeneous cellular networks, in: Proc. IEEE International Symposium on Personal, Indoor and Mobile Radio Communications, 2017, pp. 1–6.

[113] K. Wen, Y. Chen, Y. Hu, A resource allocation method for D2D and small cellular users in HetNet, in: Proc. 3rd IEEE International Conference on Computer and Communications, 2017, pp. 628–632.

[114] M.A. Elshatshat, S. Papadakis, V. Angelakis, Improving the spectral efficiency in dense heterogeneous networks using D2D-assisted eICIC, in: Proc. IEEE 23rd International Workshop on Computer Aided Modeling and Design of Communication Links and Networks, 2018, pp. 1–6.

[115] A.N. Uwaechia, N.M. Mahyuddin, A comprehensive survey on millimeter wave communications for fifth-generation wireless networks: Feasibility and challenges, IEEE Access 8 (2020) 62367–62414.

[116] N. Eshraghi, V. Shah-Mansouri, B. Maham, Fair beamwidth selection and resource allocation for indoor millimeter-wave networks, in: Proc. IEEE ICC Mobile and Wireless Networking, 2017, pp. 1–6.

[117] G. Kwon, H. Park, Joint user association and beamforming design for millimeter wave UDN with wireless backhaul, IEEE J. Sel. Areas Commun. 37 (12) (2019) 2653–2668.

[118] B. Soleimani, M. Sabbaghian, Cluster-based resource allocation and user association in mmwave femtocell networks, IEEE Trans. Commun. 68 (3) (2020) 1746–1759.

[119] Z. Zhao, J. Shi, Z. Li, L. Yang, Y. Zhao, W. Liang, Matching theory assisted resource allocation in millimeter wave ultra dense small cell networks, in: Proc. IEEE International Conference on Communications, 2019, pp. 1–6.

[120] I. Al Qerm, B. Shihada, Enhanced machine learning scheme for energy efficient resource allocation in 5G heterogeneous cloud radio access networks, in: Proc. IEEE International Symposium on Personal, Indoor and Mobile Radio Communications, 2018, pp. 1–7.

[121] A.Y. Al-Zahrani, F.R. Yu, An energy-efficient resource allocation and interference management scheme in green heterogeneous networks using game theory, IEEE Trans. Veh. Technol. 65 (7) (2016) 5384–5396.

[122] M. Liu, F.R. Yu, Y. Teng, V.C.M. Leung, M. Song, Joint computation offloading and content caching for wireless blockchain networks, in: Proc. IEEE Conference on Computer Communications Workshops, 2018, pp. 517–522.

[123] Z. Allybokus, K. Avrachenkov, J. Leguay, L. Maggi, Multi-path alpha-fair resource allocation at scale in distributed software-defined networks, IEEE J. Sel. Areas Commun. 36 (12) (2018) 2655–2666.

[124] I. Alqerm, B. Shihada, A cooperative online learning scheme for resource allocation in 5G systems, in: Proc. IEEE ICC Wireless Communications Symposium, 2016, pp. 1–7.

[125] S. Sadr, R.S. Adve, Partially-distributed resource allocation in small-cell networks, IEEE Trans. Wirel. Commun. 13 (12) (2014) 6851–6862.

[126] I. AlQerm, B. Shihada, Sophisticated online learning scheme for green resource allocation in 5g heterogeneous cloud radio access networks, IEEE Trans. Mob. Comput. 17 (10) (2018) 2423–2437.

[127] N. Singh, S. Agarwal, T. Agarwal, P.K. Mishra, RBF-SVM based resource allocation scheme for 5G CRAN networks, in: Proc. 3rd International Conference and Workshops on Recent Advances and Innovations in Engineering, 2018, pp. 22–25.

[128] F. Pace, D. Milios, D. Carra, D. Venzano, P. Michiardi, A data-driven approach to dynamically adjust resource allocation for compute clusters, Comput. Res. Repos. (2018) 1–15.

[129] G. Xu, C.Z. Xu, S. Jiang, Prophet: Scheduling executors with time-varying resource demands on data-parallel computation frameworks, in: Proc. IEEE International Conference on Autonomic Computing, 2016, pp. 45–54.

[130] M. Chen, Y. Miao, H. Gharavi, L. Hu, I. Humar, Intelligent traffic adaptive resource allocation for edge computing-based 5g networks, IEEE Trans. Cogn. Commun. Netw. 6 (2) (2020) 499–508.

[131] Y. He, L. Tang, Z. Zhou, S. Mumtaz, K.M.S. Huq, J. Rodriguez, Two time-scale resource allocation in hybrid energy powering 5G wireless system, in: Proc. IEEE Global Communications Conference, 2019, pp. 1–6.

[132] D. Zhang, Y. Qiao, L. She, R. Shen, J. Ren, Y. Zhang, Two time-scale resource management for green internet of things networks, IEEE Internet Things J. 6 (1) (2019) 545–556.

[133] M. Qin, J. Li, Q. Yang, N. Cheng, K. Kwak, X. Shen, Multiple time-scale SON function coordination in ultra-dense small cell networks, in: Proc. IEEE Global Communications Conference, 2018, pp. 1–6.

[134] M. Qin others, Learning-aided multiple time-scale SON function coordination in ultra-dense small-cell networks, IEEE Trans. Wirel. Commun. 18 (4) (2019) 2080–2092.

[135] R. Thakur, V.J. Kotagi, C.S.R. Murthy, Resource allocation and cell selection framework for LTE-unlicensed femtocell networks, Comput. Networks 129 (2017) 273–283.

[136] A. Shahid, V. Maglogiannis, I. Ahmed, K.S. Kim, E. De Poorter, Energy-efficient resource allocation for ultra-dense licensed and unlicensed dual-access small cell networks, IEEE Trans. Mob. Comput. 1 (1) (2019) 1–20.

[137] M. Tao, E. Chen, H. Zhou, W. Yu, Content-centric sparse multicast beamforming for cache-enabled cloud RAN, IEEE Trans. Wirel. Commun. 15 (9) (2016) 6118–6131.

[138] A. Shahid, K.S. Kim, E. De Poorter, I. Moerman, Self-organized energy-efficient cross-layer optimization for device to device communication in heterogeneous cellular networks, IEEE Access 5 (2017) 1117–1128.

[139] S. Zhang, N. Zhang, S. Zhou, J. Gong, Z. Niu, X. Shen, Energy-aware traffic offloading for green heterogeneous networks, IEEE J. Sel. Areas Commun. 34 (5) (2016) 1116–1129.

[140] R. Ramamonjison, V.K. Bhargava, Energy allocation and cooperation for energy-efficient wireless two-tier networks, IEEE Trans. Wirel. Commun. 15 (9) (2016) 6434–6448.

[141] M.R. Khawer, J. Tang, F. Han, UsICIC — A proactive small cell interference mitigation strategy for improving spectral efficiency of LTE networks in the unlicensed spectrum, IEEE Trans. Wirel. Commun. 15 (3) (2016) 2303–2311.

[142] L. Xu, Y. Mao, S. Leng, G. Qiao, Q. Zhao, Energy-efficient resource allocation strategy in ultra dense small-cell networks: A Stackelberg game approach, in: Proc. IEEE ICC Next Generation Networking and Internet Symposium, 2017, pp. 1–6.

[143] A. Kaur, K. Kumar, Intelligent spectrum management based on reinforcement learning schemes in cooperative cognitive radio networks, Phys. Commun. 43 (2020) 101226.

[144] W. Yang, X. Zhao, Robust resource allocation for orthogonal frequency division multiplexing-based cooperative cognitive radio networks with imperfect channel state information, IET Commun. 11 (2) (2017) 273–281.

[145] X. Huang, D. Zhang, S. Tang, Q. Chen, J. Zhang, Fairness-based distributed resource allocation in two-tier heterogeneous networks, IEEE Access 7 (2019) 40000–40012.

[146] S. Fan, H. Tian, Cooperative resource allocation for self-healing in small cell networks, IEEE Commun. Lett. 19 (7) (2015) 1221–1224.

[147] R. Wei, Y. Wang, Y. Zhang, A two-stage cluster-based resource management scheme in ultra-dense networks, in: Proc. IEEE/CIC ICCC Symposium on Wireless Communications Systems, 2014, pp. 738–742.

[148] W.C. Cheung, T.Q.S. Quek, M. Kountouris, Throughput. optimization, spectrum. allocation, Throughput optimization spectrum allocation and access control in two-tier femtocell networks, IEEE J. Sel. Areas Commun. 30 (3) (2012) 561–574.

[149] W.S. Jeon, D.G. Jeong, Energy-efficient distributed resource allocation with low overhead in relay cellular networks, IEEE Trans. Veh. Technol. 66 (12) (2017) 11137–11150.

[150] I. Alqerm, B. Shihada, Energy-efficient power allocation in multitier 5g networks using enhanced online learning, IEEE Trans. Veh. Technol. 66 (12) (2017) 11086–11097.

[151] J. Tan, S. Xiao, S. Han, Y. Liang, V.C.M. Leung, Qos-aware user association and resource allocation in LAA-LTE / WiFi coexistence systems, IEEE Trans. Wirel. Commun. 18 (4) (2019) 2415–2430.

[152] D.W.K. Ng, E.S. Lo, R. Schober, Energy-efficient resource allocation in multi-cell OFDMA systems with limited backhaul capacity, IEEE Trans. Wirel. Commun. 11 (10) (2012) 3618–3631.

[153] J.B. Rao, A.O. Fapojuwo, Analysis of spectrum efficiency and energy efficiency of heterogeneous wireless networks with intra-/inter-RAT offloading, IEEE Trans. Veh. Technol. 64 (7) (2015) 3120–3139.

[154] J. Tang, D.K.C. So, E. Alsusa, K.A. Hamdi, A. Shojaeifard, K.-K. Wong, Energy-efficient heterogeneous cellular networks with spectrum underlay and overlay access, IEEE Trans. Veh. Technol. 67 (3) (2018) 2439–2453.

[155] I. Mondal, A. Neogi, P. Chaporkar, A. Karandikar, Bipartite graph based proportional fair resource allocation for D2D communication, in: Proc. IEEE Wireless Communications and Networking Conference, 2017, pp. 1–6.

[156] P. Xu, K. Cumanan, Optimal power allocation scheme for non-orthogonal multiple access with α-fairness, IEEE J. Sel. Areas Commun. 35 (10) (2017) 2357–2369.

[157] F. Fossati, S. Moretti, S. Secci, A mood value for fair resource allocations, in: Proc. IFIP Networking Conference, IFIP Networking and Workshops, 2017, pp. 1–9.

[158] L. Xu, Y. Mao, S. Leng, G. Qiao, Q. Zhao, A cluster-based resource allocation strategy with energy harvesting in dense small-cell networks, in: Proc. International Conference on Cyber-Enabled Distributed Computing and Knowledge Discovery, 2017, pp. 303–310.

[159] H. Zhang, H. Liu, J. Cheng, V.C.M. Leung, Downlink energy efficiency of power allocation and wireless backhaul bandwidth allocation in heterogeneous small cell networks, IEEE Trans. Commun. 66 (4) (2018) 1705–1716.

[160] S. Boyd, V. Lieven, Convex Optimization, Cambridge Univ. Press, Cambridge, U.K., 2004.

[161] M.V. Nguyen, T.Q. Duong, C.S. Hong, S. Lee, Y. Zhang, Optimal and sub-optimal resource allocation in multi-hop cognitive radio networks with primary user outage constraint, IET Networks 1 (2) (2012) 47–57.

[162] D. Tsilimantos, J.-M. Gorce, K. Jaffrès-Runser, H. Vincent Poor, Spectral and energy efficiency trade-offs in cellular networks, IEEE Trans. Wirel. Commun. 15 (1) (2016) 54–66.

[163] Y. Teng, W. Zhao, Robust group sparse beamforming for dense C-RANs with probabilistic SINR constraints, in: Proc. IEEE Wireless Communications and Networking Conference, 2017, pp. 1–6.

[164] B. Dai, W. Yu, Sparse beamforming and user-centric clustering for downlink cloud radio access network, IEEE Access 2 (2014) 1326–1339.

[165] W. Wang, X. Liu, List-coloring based channel allocation for open-spectrum wireless networks, in: Proc. IEEE Vehicular Technology Conference, 2005, 1, pp. 690–694.

[166] Y. Yangt, B. Bai, W. Chen, Achieving high frequency reuse in dense cellular networks: A matrix graph approach, in: Proc. IEEE Global Conference on Signal and Information Processing, 2014, pp. 937–941.

[167] Y. Meng, J. Li, H. Li, M. Pan, A transformed conflict graph-based resource-allocation scheme combining interference alignment in OFDMA femtocell networks, IEEE Trans. Veh. Technol. 64 (10) (2015) 4728–4737.

[168] L. Liang, W. Wang, Y. Jia, S. Fu, A cluster-based energy-efficient resource management scheme for ultra-dense networks, IEEE Access 4 (2016) 6823–6832.

[169] H. Han, Y. Li, X. Guo, A graph-based random access protocol for crowded massive MIMO systems, IEEE Trans. Wirel. Commun. 16 (11) (2017) 7348–7361.

[170] Y. Zhang, S. Wang, J. Guo, Clustering-based interference management in densely deployed femtocell networks, in: Proc. IEEE/CIC ICCC Symposium on Wireless Networking and Multimedia, 2015, pp. 478–483.

[171] H. Huang, M. Trivellato, A. Hottinen, M. Shafi, P.J. Smith, R. Valenzuela, Increasing downlink cellular throughput with limited network MIMO coordination, IEEE Trans. Wirel. Commun. 8 (6) (2009) 2983–2989.

[172] S. Basloom, A. Nazar, G. Aldabbagh, M. Abdullah, N. DImitriou, Resource allocation using graph coloring for dense cellular networks, in: Proc. International Conference on Computing, Networking and Communications, 2016, pp. 1–5.

[173] Y. Yang, C. Lai, L. Wang, X. Wang, A energy-aware clustering algorithm via game theory for wireless sensor networks, in: Proc. 12th International Conference on Control, Automation and Systems, 2012, pp. 261–266.

[174] Y. Liu, Y. Wang, Y. Zhang, R. Sun, L. Jiang, Game-theoretic hierarchical resource allocation in ultra-dense networks, in: Proc. 27th IEEE International Symposium on Personal, Indoor and Mobile Radio Communications, 2016, pp. 1–6.

[175] E. Pateromichelakis, M. Shariat, A. Quddus, M. Dianati, R. Tafazolli, Dynamic clustering framework for multi-cell scheduling in dense small cell networks, IEEE Commun. Lett. 17 (9) (2013) 1802–1805.

[176] C. Zhao, X. Xu, Z. Gao, L. Huang, A coloring-based cluster resource allocation for ultra dense network, in: Proc. IEEE International Conference on Signal Processing, Communications and Computing, 2016, pp. 1–5.

[177] Y. Guo, J. Ma, C. Wang, K. Yang, Incentive-based optimal nodes selection mechanism for threshold key management in MANETs with selfish nodes, Int. J. Distrib. Sens. Netw. (2013) 1–14.

[178] L. Blumrosen, N. Nisan, Combinatorial auctions, in: Algorithmic Game Theory, Cambridge Univ. Press, Cambridge, U.K., 2007, pp. 267–298.

[179] J.-H. Cho, A. Swami, T. Cook, Combinatorial auction-based multiple dynamic mission assignment, in: Proc. IEEE Military Communications Conference, 2011, pp. 1327–1332.

[180] L. Ferdouse, O. Das, A. Anpalagan, Auction based distributed resource allocation for delay aware OFDM based cloud-RAN system, in: Proc. IEEE Global Communications Conference, 2017, pp. 1–6.

[181] I. Alqerm, B. Shihada, Energy efficient traffic offloading in multi-tier heterogeneous 5g networks using intuitive online reinforcement learning, IEEE Trans. Green Commun. Netw. 3 (3) (2019) 691–702.

[182] A. Kaur, K. Kumar, A reinforcemnt learning based evolutionary multi-objective optimization algorithm for spectrum allocation in cognitive radio networks, Phys. Commun. 43 (2020) 101196.

[183] Y. Zhao, H. Xia, Z. Zeng, S. Wu, Joint clustering-based resource allocation and power control in dense small cell networks, in: Proc. IEEE/CIC ICCC Symposium on Wireless Communications Systems, 2016, pp. 1–5.

[184] M. Kalil, A. Al-Dweik, M.F. Abu Sharkh, A. Shami, A. Refaey, A framework for joint wireless network virtualization and cloud radio access networks for next generation wireless networks, IEEE Access 5 (2017) 20814–20827.

[185] X. Wang others, Virtualized cloud radio access network for 5g transport, IEEE Commun. Mag. 55 (9) (2017) 202–209.

[186] W. Lee, T. Na, J. Kim, How to create a network slice? - A 5G core network perspective, in: Proc. International Conference on Advanced Communication Technology, 2019, pp. 616–619.

[187] R. Su others, Resource allocation for network slicing in 5g telecommunication networks: A survey of principles and models, IEEE Netw. 33 (6) (2019) 172–179.

[188] Y. Sun, S. Qin, G. Feng, L. Zhang, M.A. Imran, Service provisioning framework for RAN slicing: User admissibility, slice association and bandwidth allocation, IEEE Trans. Mob. Comput. (2020) 1–15.
[189] M. Amine, A. Kobbane, J. Ben-Othman, New network slicing scheme for UE association solution in 5G ultra dense HetNets, in: Proc. IEEE International Conference on Communications, 2020, pp. 1–6.
[190] F. Wei, G. Feng, Y. Sun, Y. Wang, S. Qin, Y. Liang, Network slice reconfiguration by exploiting deep reinforcement learning with large action space, IEEE Trans. Netw. Serv. Manag. 17 (4) (2020) 2197–2211.
[191] A.S. Charismiadis, D. Tsolkas, N. Passas, L. Merakos, A metaheuristic approach for minimizing service creation time in slice-enabled networks, in: Proc. ICC IEEE International Conference on Communications, 2020, pp. 1–6.
[192] X. Chen, Z. Liu, Y. Chen, Z. Li, Mobile edge computing based task offloading and resource allocation in 5g ultra-dense networks, IEEE Access 7 (2019) 184172–184182.
[193] P. Yang, Y. Zhang, J. Lv, Load optimization based on edge collaboration in software defined ultra-dense networks, IEEE Access 8 (1) (2020) 30664–30674.
[194] J. Zheng others, Joint downlink and uplink edge computing offloading in ultra-dense hetnets, Mob. Networks Appl. 24 (5) (2019) 1452–1460.
[195] C. Zhao, Y. Cai, A. Liu, M. Zhao, L. Hanzo, Mobile edge computing meets mmwave communications: Joint beamforming and resource allocation for system delay minimization, IEEE Trans. Wirel. Commun. 19 (4) (2020) 2382–2396.
[196] M.S. Gupta, K. Kumar, Application aware networks' resource selection decision making technique using group mobility in vehicular cognitive radio networks, Veh. Commun. 26 (2020) 1–24.
[197] A. Kaur, K. Kumar, Imperfect CSI based intelligent dynamic spectrum management using cooperative reinforcement learning framework in cognitive radio networks, IEEE Trans. Mob. Comput. (2020) 1–13.

Nidhi Sharma has received her B.Tech. degree from the Department of Electronics and Communication Engineering, Chitkara University, Himachal Pradesh, India in 2013, and M.Tech. degree in Mobile and Pervasive Computing from Indira Gandhi Delhi Technical University for Women, Delhi, India in 2016. From 2018 to 2019, she was an Assistant Professor at the National Institute of Technology, Hamirpur, India (Institute of National Importance). Presently, she is pursuing her Ph.D. degree in Electronics and Communication Engineering from National Institute of Technology, Hamirpur, India. She has publications in the field of communication on international platform. Her area of interest includes wireless communication, resource management, machine learning especially the application of artificial intelligence in ultra-dense networks.

Krishan Kumar has received his B.E. degree from the Department of Electronics and Communication Engineering, CR State College of Engineering, Murthal, Haryana, India in 2002, and his M.Tech. degree in Electronics and Communication Engineering from the National Institute of Technology, Kurukshetra, India (Institute of National Importance), in 2005. He has received his Ph.D. degree from the Department of Electronics and Communication Engineering, Motilal Nehru National Institute of Technology Allahabad, Prayagraj, India, under Quality Improvement Program scheme. Presently, he has been working as an Assistant Professor at the National Institute of Technology, Hamirpur, India (Institute of National Importance), since 2006. He is the reviewer of various SCI-indexed journals. He is handling various projects sponsored by Department of Science and Technology, Government of India. His area of interest includes wireless communication especially cognitive radio networks, vehicular networks with mobility management issues, and application of artificial intelligence.