Journal of Computers Vol. 36 No. 1, February 2025, pp. 143-162
https://doi.org/10.63367/199115992025023601010

Resource Allocation for MEC in Ultra-dense Networks

Huan Ma[1*], Shuyun Li[1], Zhihui Wang[1],
Xubing Dou[1], and Xinchang Zhang[2]

[1] Key Laboratory of Computing Power Network and Information Security, Ministry of Education, Shandong Computer Science Center (National Supercomputer Center in Jinan), Qilu University of Technology (Shandong Academy of Sciences), Shandong Provincial Key Laboratory of Industrial Network and Information System Security, Shandong Fundamental Research Center for Computer Science, Jinan, China
{10431220402, 10431220470, 10431220464, 10431220601}@stu.qlu.edu.cn

[2] Shandong Normal University, Jinan, China.
zhangxc@sdas.org

Received 29 February 2024; Revised 27 June 2024; Accepted 8 August 2024

Abstract. In the complex context of 5G ultra-dense networks, the unprecedented growth in traffic and computing demand has made the interaction between edge devices and edge servers extremely intricate due to the time-varying nature of the network environment and resource heterogeneity. The combination of UDN and Mobile Edge Computing not only significantly improves network performance but also expands application scenarios and enhances user experience, marking it as one of the crucial trends in future network development. In this paper, we review recent results and research progress on MEC-based resource allocation in 5G ultra-dense networks. We analyze resource allocation strategies from multiple dimensions, including user association, interference management, energy efficiency optimization, delay reduction, access selection, and user mobility. Finally, we discuss the challenges and open issues that need to be addressed for future development in this area.

Keywords: UDN, 5G, MEC, user mobility, resource allocation, load balance

1 Introduction

1.1 5G

The fifth generation of wireless networks will include extreme base stations with unprecedented device density and number of antennas, and ultra-high carrier frequencies with huge bandwidth, realizing ubiquitous and seamless communication between humans, humans and machines, or machines and machines, and thus interconnecting the entire world. 5G technologies include multi-RAT, Wireless Gigabit Alliance (Wi-Gig), Filter Bank Multi-Carrier (FBMC), Beam Divide Multiple Access (BDMA), and Non-Orthogonal Multiple Access (NOMA). Reference [1] reviews emerging and enabling technologies relevant to 5G systems supporting the IoT, such as 5G new air ports (NR), multiple-input multiple-output antennas with beamforming, millimeter-wave commutation, HetNets, and the role of AR in the IoT. Three use case scenarios in 5G, namely eMBB, mMTC, and URLLC.

As more and more heterogeneous provisioning meets increasingly unpredictable demand, increasingly ubiquitous consumer devices such as cars and drones are becoming the next generation of functional user devices when connected to cellular broadband. As agreed by 3GPP, AP or communication link densification is one of the promising approaches to address the growing capacity demand and exploding data traffic. Driven by continued network densification, base stations are becoming smaller and smaller, and the original functional differences between the network and the UE are becoming blurred. Intelligent UEs can help provide wireless connectivity to nearby relevant devices, e.g., providing D2D based data relay, proximity gaming, content distribution, and caching. However, the unpredictable and heterogeneous mobility characterizing user devices, considering the realistic

* Corresponding Author

sources of motivation, can lead to problems such as end-users not offering their personal devices voluntarily.

Several densification processes in 5G networks coexist harmoniously: HetNets, C-RANs, Large Scale/Massive Scale MIMO and Millimeter Wave Networks, D2D Networks, and Large Scale IoT Environments, etc. Reference [2] explores the question of what are the fundamental limitations of network densification. However, some of the idealizations of the model need to be further investigated. Network densification cannot continue indefinitely and extensive research on realistic system models is needed to capture the reality of dense networks.

1.2 Evolution of 5G Communications

In order to support massive device connectivity and provide ubiquitous IoT services, a wide range of efficient communication infrastructures should be involved, including H2H, H2M, M2H and M2M communications.

M2M communication is characterized by low mobility, low transmission frequency, and small data volumes. M2M communication poses a serious challenge of connecting a large number of MTC devices (MTCDs) to the Internet. mMTC encompasses the conditions required to facilitate MTCD connectivity with minimal or no human intervention. However, mMTC presents unique technical challenges to support a large number of MTC devices, and related challenges include QoS configuration, handling highly dynamic and sporadic MTC traffic, significant signaling overhead, and RAN congestion.

The unique traffic characteristics of M2M communications make it a waste of resources if M2M communications are implemented in the same way as H2H communications. At the same time, a large number of MTCDs can exacerbate overload control problems and affect the performance of H2H communications. Both H2H and M2M communications require the support of mobile cellular systems, and the large amount of traffic generated by a massive number of devices requires a major shift in cellular and non-cellular wireless technologies. Reference [3] discusses the role of M2M communication in UDNs and the methods to realize M2M communication in UDNs, including at the PHY, MAC, network, and application layers, in terms of relevant standards, application scenarios, traffic characteristics, and related wireless technologies. Reference [4] introduces the network architecture of heterogeneous MTC networks and proposes an intelligent hybrid random access scheme based on 5G/6G for smart cities, which improves the probability of successful access and meets the quality of service requirements of diverse URLLC and mMTC devices.

Most current studies usually consider UDN and M2M communications as two separate topics. Reference [5] considered a two-tier architecture for clustered packet MTCD and proposed a resource allocation problem to minimize the packet loss probability and maximize the spectral efficiency. Reference [6] proposed hybrid non-orthogonal random access and data transfer to overcome the signaling overhead and resource allocation problems of M2M communications. Reference [7] proposes the use of unused PUSCH resource reallocation algorithm to significantly improve the performance of static and dynamic systems in terms of the number of successful communications. Existing resource allocation schemes in MTC usually consider the signal-to-noise ratio (SNR) rather than its QoS requirements to prioritize MTC devices. Reference [8] proposed a resource allocation scheme with dynamic prioritization for MTC devices with multiple RATs, including two main components: medium access and resource allocation. Reference [9] proposed a new framework customized for large-scale MTC services, which includes joint control of dynamic resource allocation between physical random access channels and physical uplink shared channels. Reference [10] investigated a hybrid network model considering the coexistence of M2M cellular and D2D communications. Reference [11] proposes to share the same spectrum with H2H communications via D2D communications, modeled the resource sharing problem as a bipartite graph, and proposed an interference-aware graph-based resource sharing scheme based on fixed M2M transmit power. In order to further enhance the protection of H2H services, an adaptive power control mechanism is introduced, but only single-cell scenarios are mentioned in the simulation, and future work is suggested to consider multi-cell complex scenarios. Reference [12] proposes a two-phase resource optimization allocation algorithm for H2H or M2M coexistence, where joint power resource allocation is performed in the first phase to meet the QoS requirements of H2H traffic, and the second phase focuses on meeting the QoS requirements of M2M traffic. Reference [13] presents a dynamic borrowed scheduler for multi-service communication in 5G systems, which proposes a shared allocation policy that weighs between H2H and M2M traffic to maximize bandwidth utilization to meet QoS requirements. Congestion between network devices in busy traffic is further considered in the future.

The spectrum sharing cost of H2H users is private information, and traditional incentive mechanisms usually assume that the BS is fully aware of the information on the H2H side, how to design the incentive mechanism with information asymmetry is another key challenge. Reference [14] proposes a blockchain-based framework for privacy preservation, incentive compatibility and spectrum efficiency. First, H2H users sign a spectrum sharing contract with the base station. Next, the shared spectrum is allocated to M2M devices to maximize the total throughput, and the operational details such as secure spectrum sharing, incentive mechanism design, and spectrum efficient allocation are elaborated. Future work should delve into the design of incentive-compatible contracts in complex scenarios with multiple service providers to solve the spectrum allocation problem with information uncertainty. Meanwhile, virtualization is also a promising technical direction for future UDN development. Different M2M services are owned by different providers, trying to separate different functions or services from different providers, the failure of one service will not affect the other service, and the privacy of different services can be protected.

1.3 UDN

HUDN. There are some types of cells in network centric HUDN. Table 1 illustrates a comparison of the different types of APs in HUDNs.

Table 1. Compare different types of APs in HUDN

Type	Features	Deployment scenario	Coverage range	Power
Macro Cell	High power, wide coverage	Outdoor	Several kilometers	43-46dBm
Microcell	Low power, suitable for small coverage areas	Indoor/Outdoor	Up to 300m	23-30dBm
Picocell	Lower power, suitable for very small coverage areas	Indoor	10-15m	Up to 23dBm
Relay	Enhances network signal, extends coverage	Indoor/Outdoor	Up to 300m	30dBm
Remote radio head	Deployed away from the core base station	Outdoor	300-500m	≥30dB

Reference [15] studies 5G HUDN and provided a system architecture consisting of a core network and a virtualized integrated ground-space radio access network. Reference [16] investigates the downlink precoding method to overcome the same-layer and cross-layer interference problems of MUEs and FUEs, and provided a theoretical framework for the deployment of UDNs and HetNets. Reference [17] focuses on interference management for UDNs using dynamic resource allocation. Due to the random deployment of small base stations, the mobility of user devices, and the preference of small base stations in the selection or re-selection process, reference [18] developes a fiber-radio based load balancing method in UDN hotspots. Reference [19] proposes a joint strategy of SBS dormancy and spectrum allocation to address the huge power consumption and spectrum resource constraints in heterogeneous UDNs. Reference [20] summarizes the spatio-temporal arrival characteristics of different service volumes in UDNs, and optimized several promising techniques such as dynamic time-division duplex and full-duplex radio to adapt network services to service volumes.

Network heterogeneity and densification exacerbate interference between multiple network layers, and network-centric HUDNs introduce new problems for 5G and beyond, all leading to traffic load imbalance between SBSs, and an inevitable increase in the complexity and overhead of network coordination. Future work will require the design and implementation of intelligent methods and algorithms for interference cancellation, mitigation and management. In addition, with the emergence of airborne base stations, such as drones and satellites, intelligent connectivity should be considered so that redundant connections can be intelligently released to ensure reliable wireless communications and service continuity, and network heterogeneity makes network management and planning more challenging.

UUDN. In addition to small cell base stations, relay stations, and distributed RRHs in UDNs, UEs themselves can act as APs in UUDNs. Reference [21] mentions that UUDNs are wireless networks in which the density of APs is comparable to the density of users. The network will organize a dynamic AP group (APG) to seamlessly serve each user without user participation.The network should intelligently identify the user’s radio communication environment, and then flexibly organize the required APGs and resources to serve the user. The main feature of the UUDN is that the network will be more intelligent, and it can automatically detect the terminal's capability, the user's demand, and the radio environment. In particular, when a user moves, its Dynamic AP Group will dynamically adjust to support its movement compared to traditional switching and migration. The members of the dynamic AP group will adaptively adjust to jointly and collaboratively transmit data streams to improve spectrum efficiency and user experience. In addition, when an AP joins an APG, the network will provide security through the AP's authentication and UE network authentication.

In UUDN, Dynamic AP Grouping (DAPGing) is introduced, in which four functional entities, Local Service Center (LSC), Local Data Center (LDC), Network Service Center (NSC), and Network Data Center (NDC), are used to provide user-centric services, and each registered user in the UUDN has a unique Dynamic APG-ID. Reference [22] presents a general tutorial on user-centric clustering. Reference [23] proposes a dynamic AP grouping based mobility management strategy for UCWA users. Reference [24] discusses a user-centered dynamic DAPGing approach in which each UE can design its own APG based on local measurements only, regardless of the decisions of other UEs. Reference [25] proposes a maximum data rate oriented dynamic AP grouping (MDTR-DAPGing) system. Reference [26] investigates collaborative APs to provide access services to users in NOMA-based UUDNs, thus improving the system EE performance.

Several research works [27-32] have been devoted to address user association and resource allocation in UDN. Reference [27] proposes a new MUC clustering for resource allocation in UDN to maximize the sum rate of each resource block. Reference [28] modeles the joint user association and resource allocation problem as a mixed integer planning problem, considering multiple factors such as load balancing, user quality of service, EE, and cross-layer interference. Reference [29] considers the joint user association and resource allocation problem in millimeter-wave self-return UDN, and proposed a network and rate maximization algorithm based on coalition game. Reference [30] focuses on user association and resource allocation and proposes a unified NOMA in UDN. Reference [31] investigates the joint optimization of user association and dynamic TDD in UDN. Reference [32] proposes a new method for CoMP based user association and resource allocation. A clustering method based on fiducial filtering and location loading is used to reduce the network complexity. Then, a competition-based resource allocation scheme is proposed based on the clustering results.

In UUDN, mobile behavior is an important factor that should be considered when optimizing the overall system performance. Heterogeneous and cooperative networks are another issue for further research. Future work can try to support UUDN through complex multilayer scenarios, multiple RATs and irregular coverage. Finally, user authentication methods are crucial to protect user privacy and providing end-to-end security is important for UUDN. Table 2 summarizes the comparison between HUDN and UUDN in detail.

Table 2. Comparison of HUDN and UUDN

Aspect	HUDN	UUDN
Network architecture	Centralized structure	Distributed architecture
Mobility management	Services centralized at the network core, provided by central nodes	Decentralized services, possibly provided on user devices or edge nodes
Mobility management	Switching with terminal involvement	AP dynamic grouping, no terminal involvement
Connection	Symmetric	Asymmetric
Network sensing	Focus on network sensing and data collection	Focus on sensing devices and environmental awareness
Wireless resource management	Independent units	Cooperative, user-centric
Interference management	Independent units	Cooperative, user-centric

Although UDN and MEC technologies have shown great potential in theoretical and experimental studies, practical deployment still faces many challenges. Existing studies often fail to provide comprehensive solutions for resource allocation, quality of service assurance, interference control, and user mobility management. In this paper, we review related works to summarize the main techniques and methods in the field of UDN and MEC. Additionally, we provide an in-depth analysis of various optimization problems encountered in resource allocation and propose strategies and techniques to address these challenges. The rest of the paper is organized as follows. Section II reviews the research results and technical approaches in the field of UDN and MEC. Section III provides an in-depth discussion on resource allocation in MEC-based UDNs, focusing on load balancing, access selection, and user mobility issues in dynamic environments. Section IV discusses future directions and some of the challenges faced in dynamic scenarios. Finally, Section V summarizes the main ideas of this paper.

2 Related Work

A comprehensive survey of wireless networks across generations from 1G to 5G, presenting intelligent management techniques and backhaul solutions for a mix of UDN and other enabling technologies [19]. Reference [33] emphasizes the importance of user association in 5G networks and summarizes algorithms related to user association designed for HetNets, massive MIMO networks, millimeter wave scenarios and energy harvesting networks. Reference [34] investigates 5G+ wireless systems MANs, discussing the appropriate adoption of scenarios for engaging users through adaptive and flexible network infrastructures. Reference [35] introduces resource allocation methods for different scenarios of UDNs in 5G and beyond networks, involving LTE-U, CRNs, HETNets, C-RANs, D2D networks and millimeter wave networks. Reference [36] presents several effective solutions for denser scenarios, including new schemes and theories such as large-scale convex optimization, MFG), stochastic geometry-assisted methods, stochastic optimization, game theory and clustering techniques. Reference [37] proposes a switching parameter optimization method based on the DDPG algorithm, which ensures seamless connectivity and maintains high-quality service for small and medium-sized base stations in UDNs, with greatly improved mobility performance at different speeds.

Reference [38] validates the theoretical analysis of Successive Interference Cancellation (SIC) in a multi-tier 5G Ultra-Dense Network (UDN), an enhanced SIC technique is developed. This method employs a stochastic geometry model that accurately simulates real-world conditions by accounting for the random distribution of users and base stations within the UDN. Reference [39] proposes a scaling law for the AN density required to support a given UE density at a given rate based on UE requirements, and discusses the possibility of increasing system capacity due to the utilization of network densification. Reference [40] explores DenseNets from the perspective of different deployment strategies, covering the densification of classical macro-layers, extremely dense indoor femtosecond layers, and outdoor Distributed Antenna Systems (DAS), which can be dynamically configured as either single microcells or multiple independent microcells. Reference [41] presents an architecture based on mobility separation, coordination of radio resources among multiple nodes, data plane processing, and integration with the WAN, proposing multi-hop wireless self-return to enhance deployment flexibility. Reference [42] considers multiple correlation modeling in a dense network environment where base stations have idle mode capability, overcoming the capacity constraints imposed by backhaul links and greatly simplifying the computation of the average downlink rate for individual connections with multi-cell cells. However, the average rate decreases when connecting to more distant cells, and future work can compensate for this by expanding beyond bottlenecks with limited backhaul capacity. In ultra-dense deployments, where the network consists of a large number of small base stations that consume high amounts of energy, the power control of the base stations is equivalent to controlling the power-saving modes of the base stations by switching between sleep and wake-up modes.

Several research works [43-51] aim to optimize the delay problem for users in MEC. Reference [43] exploits the RSMA technique to minimize the sum of the user's maximum experiential latency due to offloading and processing between different MEC servers by co-optimizing the proportion of the user's computational tasks assigned to the servers, and the rate of the assigned public messages. Reference [44] studies the joint task, power and spectrum allocation problems in MEC-based SCNs, and the authors developed a distributed multi-stack Q-learning method that uses multiple stacks to record the wireless environment and user information, thus improving the learning efficiency. Reference [45] investigates a novel D2D multi-assistant MEC system based on D2D, in which local users request their nearby WDs to act as assistants for collaborative computation, minimizing computational latency by optimizing the task allocation to local users. A wireless-powered multi-user MEC system is considered in [46], where joint computation and wireless resource allocation optimization for dynamic task arrivals over a finite time range consisting of multiple time slots are investigated to minimize the system energy consumption over the entire range, including the transmission energy of the APs and the computation energy of the MEC server. Reference [47] envisions a multi-user WiFi based MEC architecture that jointly solves the problem of task allocation and resource allocation to reduce the energy consumption on the mobile terminal side under the application delay constraint. Reference [48] considered a NOMA assisted MEC system to reduce energy consumption and delay. Reference [49] applies both NOMA uplink and downlink transmissions to MEC, and the analysis results showed that the use of NOMA can effectively reduce the delay and energy consumption of MEC offloading. Reference [50] presentes a heterogeneous complex multi-MEC system consisting of ground-based and UAV airborne MEC servers, which jointly considered the risk-perceived behavior of individual users and the risk of failure of shared computing resources. Reference [51] proposes a stochastic optimization problem involving dynamic offloading and resource scheduling between local devices, BSs, and back-end clouds to minimize energy and computational resource consumption in a MEC system through EH devices, while meeting the QoS requirements of IoT devices.

Highly dynamic task arrivals and wireless channel states pose a great challenge for RA in MEC. Reference [52] investigates stochastic task allocation and CPU cycle frequency scaling for MEC systems, and proposes a stochastic optimization-based algorithm, TOFFEE, which does not require prior statistical information related to task arrivals or wireless network states. Literature [53] proposes a mobile-aware service adaptive deployment method, constructs a mobility model based on the probability density function of dwell time and a service deployment overhead model based on user mobility, and adopts a quantum ant colony algorithm to solve the service adaptive deployment scheme. User mobility is further investigated in [54], where a mobility-aware offloading algorithm is proposed to obtain an approximate optimal offloading scheme. Reference [55] combines edge caching with mobility management for dynamic service placement and solves the task offloading problem by a Thompson sampling-based algorithm. However, the MEC servers in these works are relatively static, the mobility trajectories are usually considered known, and the optimization problem is solved by the MEC servers. Reference [56] investigates the IoT task offloading problem in MEC-based Hetworks, and makes a first attempt to thoroughly explore the task offloading problem in a heterogeneous network that supports MEC and wireless backhaul, jointly optimizing the offloading decisions, transmission power, and the allocation of radio and computational resources, with the aim of minimizing the energy consumption of the devices, while adhering to their latency deadlines.

With the development of the IoT, a large number of smart devices are connected together and connected to the Internet. IoT-based smart cities can provide a variety of smart services to citizens and managers, thereby improving the utilization of public resources such as transportation, healthcare, environment, entertainment, and energy. The convergence of transmission, computing, and caching is having a profound impact on the development of flexible and efficient IoT in smart cities. Recently more and more smart IoT devices are being created for consumer use, including home automation, smart cars, smart wearables, etc., thus generating large amounts of local data. At the same time, higher requirements for QoE and QoS have been imposed. The integration of heterogeneous IoT networks with 5G networks provides a framework for deploying the three main communication paradigms of MEC, D3D, and SDN, which takes into account the dense deployments announced by the ITU and the 5GPP as well as the requirements for system scalability and availability, and improves the availability and scalability of the overall system [57]. For a system as complex and inclusive as UDN, effective RA mechanisms are very important, and unlike previous work, this survey focuses on organizing and summarizing RA in the context of mec-based UDN. The third part of the article focuses on different optimization criteria, server load balancing, server access selection and user mobility aspects considered in a dynamic environment.

3 MEC-based UDN Resource Allocation

3.1 Different Optimization Criteria

In a UDEC environment, there are a large number of edge devices and edge servers with heterogeneous computing resources and communication links. The system resources in UDEC are usually diverse, e.g., a variety of service cache placement policies, different CPU frequencies for mobile devices and central cloud or edge servers, and memory. The distributed and heterogeneous nature of the resources further complicates the scheduling process with time-varying states such as network disconnections, channel states, and backhaul delays. Therefore, it is necessary to fully and efficiently utilize system resources to improve UDEC performance.

Latency. In [58-60], considering the problem of minimizing the overall delay in multiuser scenarios in UDNs with MEC. Reference [58] proposes a low-complexity heuristic algorithm for clustering small cells by jointly considering small-cell clustering, transmission power allocation, and computational task allocation in UDNs. For a given set of clusters, a genetic algorithm-based resource allocation algorithm is derived, which in turn minimizes the overall delay for all users. In [59], deploying multiple SC associations to the same user while partitioning and offloading their compute-intensive tasks, a novel offline task partitioning method is proposed, and the results show that there exists an optimal ordering of multiple associations that significantly reduces the overall latency. Reference [60] integrates blockchain and AI into 5G UDEC networks, and proposes a novel dual-timescale deep reinforcement learning (2Ts-DRL) approach to minimize the total offloading latency and network resource usage by jointly optimizing computation offloading, resource allocation, and service cache placement. In practical scenarios, SCceNBs in different regions may have different service cache placement states. Future work can consider more complex heterogeneous service cache placement scenarios, where each SCceNB has a personalized service cache model for different EDs to meet its computational offloading requirements.

Distributed computing resources in edge clouds and energy dynamics in mobile device batteries make offloading tasks for users in UDNs challenging. Aiming to minimize the delay while saving the battery life of user devices. Reference [61] proposes an innovative framework for MEC task offloading in SD-UDN using the idea of software-defined networking, including three planes: user, data, and control. The SD-UDN controller updates all the information tables based on the delay and energy consumption of the tasks to give the task offloading policy for the mobile devices and edge cloud resource allocation policy for the edge cloud. However, the latency and energy consumption of mobile devices sending requests to the SD-UDN controller are ignored. Future work considers task offloading in more complex deployments with user mobility. Reference [62] investigates the MEC-based task offloading and channel resource allocation problem in 5G UDN, and proposes a channel RA algorithm based on differential evolutionary algorithm considering the complexity of the problem and the coupling of decision variables. Future work considers more complex resource allocation problems under multilateral servers in 5G UDN. Reference [63] proposes an online task offloading algorithm based on state-of-the-art DRL techniques: A3C, which focuses on the case where a complete a priori knowledge of the instantaneous channel state information CSI and BS computational capabilities is lacking, with the aim of minimizing the task duration while satisfying the energy budget constraint. Notably, A3C-MEC fails to satisfy the energy budget constraint for non-UDNs. Future work will further explore multi-server offloading scenarios, i.e., allowing a single mobile user to simultaneously offload subtasks to different BSs to balance task duration and energy consumption. Reference [64] explores the spectrum sharing and edge computation offloading problem for SDN-based ultra-dense networks, and designed a bivalent auction-assisted spectrum sharing scheme, which, with the assistance of an SDN controller, can obtain global information about channel quality on the MBS edge cloud, the SBS edge cloud, and the UEs, but only the communication resources in computation offloading are considered.

Reference [65] theoretically analyzes the trade-off between energy efficiency and service latency, jointly optimizing task offloading decisions, elastic computational resource scheduling, and radio resource allocation in a dense MEC-based C-RAN. Reference [66] investigates the resource allocation problem of NOMA-MEC systems in UDNs, where all the users served by each SBS are divided into different clusters, where the users in each cluster use the NOMA transmission scheme. In addition, a TDMA transmission scheme is used between different clusters to avoid NOMA inter-cluster interference. The authors propose the mean field DDPG algorithm, which solves the problem of DQN's inability to handle continuous actions. Future work extends it to various wireless communication scenarios, such as data caching, task migration, etc. In addition, considering that CSI is not perfect in practice, channel estimation modeling is also one of the research directions to improve the practicality of the proposed algorithm. Reference [67] proposes a heuristic task offloading algorithm HTOA, which alternately updates the radio resource management in each iteration, and solves the joint channel allocation and user upload power control problems by using the greedy strategy and the golden split method. Simulations show that the algorithm can effectively reduce the delay and energy consumption of task offloading, and its performance is better as the number of users increases. Future work will further consider the complex scenarios of user mobility and the energy consumption of the MEC server.

UDEC has been shown to be the most promising network architecture for the 5G and even 6G era. Surveying the current research works in the field of MEC, most of them assume that ESs are already deployed. Reference [68] proposes an optimal deployment and allocation strategy for deploying ESs in UDNs, based on queuing theory and vector quantization techniques, which can optimize the number and location of ESs as well as the allocation of mobile subscribers for a given UDN environment in order to minimize the cost of the service provider and ensure the service completion time. In addition, most of the literature is based on the assumption that users are absolutely rational and design a series of methods to achieve the desired goals. However, due to the subjectivity of users, the actual results will deviate from the real situation. Reference [69] investigates the computational task offloading problem in real-world scenarios using the prospect theory framework. For the task offloading problem in small cellular network scenarios, an artificial fish swarm algorithm is used to optimize the system energy with finite delay. It remains challenging for future work to explore more for practical real-world scenarios.

Jointly Optimizing the Total Cost. The problem of task offloading and resource scheduling in MEC-based UDNs, several literatures aim to minimize the system cost. Reference [70] firstly constructes a system model for MEC-assisted UDN and constructed a system overhead minimization problem. The problem is transformed into three subproblems: offloading strategy, channel allocation and power allocation. Reference [71] studies the UDRN problem of deploying MECs in relay networks and decomposed the problem into three subproblems: task caching, partial offloading and stable matching. A stable matching algorithm was obtained using the GS algorithm with low complexity to obtain a match between the user and the relay. However, only two-hop transmissions were considered. Reference [72] proposes a more flexible and hierarchical matching relay selection algorithm. For vulnerable users, a weighted relay selection algorithm is proposed to maximize the system performance. However, it requires more resource investment and higher latency, and needs to be selected according to the actual situation. Reference [73] proposes a Newton-IPM-based computational RA algorithm and a genetic algorithm-based BS selection and resource scheduling algorithm. Reference [74] considers common issues of task offloading, uplink transmission power control, communication and computational resource allocation in heterogeneous UDNs to maximize the long-term utility of the system. In particular, each user has the flexibility to partition its task into several subtasks, and the authors propose a Multi-Intelligent Depth Deterministic Policy Gradient (MADDPG) approach for dealing with task partitioning and power control in a continuous action space. Reference [75] develops a novel Calibrated Contextual Bandit Learning algorithm where users are able to make local task offloading decisions independently in order to minimize the long term average task latency among all users. However, it is only applicable to static scenarios and requires re-learning when the network topology changes. Future work can consider additional change detection algorithms to monitor different topologies in combination with utilizing active adaptive methods. To cope with SD diversity and resource constraints, collaborative offloading has emerged as a viable technique to improve the ability to schedule independent tasks. Existing work often ignores the impact of insufficient capacity of edge servers and network congestion. Reference [76] implements a collaborative offloading scheme between MEC servers, sets a threshold to alleviate congestion, and provides additional computational resources to nearby edge servers to reduce total execution time, task failures, and network congestion. Table 3 summarizes and compares the literature mentioned above.

Table 3. Comparison of the literature

Characteristic	Reference
Studied overall delay minimization in multi-user scenarios in UDNs.	[58, 59, 60]
Aims at delay constraints while reducing energy consumption.	[62, 63, 108]
Explored udn spectrum sharing and edge computing offloading issues.	[64]
Analyzed theoretically the trade-off between energy efficiency and service delay.	[65]
Joint allocation	[66, 67]
Optimizes the number and location of servers based on queuing theory and VQ techniques	[68]
Minimizing system cost	[70]
The UDRN problem of deploying MECs in relay networks is considered.	[71, 72]
Proposed BS selection and resource scheduling algorithm based on genetic algorithm.	[73]
The problem of resource allocation in heterogeneous UDNs is considered.	[74]
Users can independently make local task offloading decisions through forecasting.	[75]
Implement a collaborative offloading scheme between MEC servers to reduce network congestion.	[76]

3.2 Dynamic Environment

Load Balancing. Due to the highly dynamic mobility of mobile devices and the randomness of work requests, offloading too many tasks to the same MEC server can lead to server overload and congestion, which is a huge waste of computational resources when the server operates under low load, resulting in a sharp drop in QoS and QoE. Fair resource sharing and load balancing among MEC-BSs is an essential issue due to the clustering effect of users and the heterogeneity of MEC-BSs in terms of computation and caching capacity. Reference [77] investigates the cellular association problem between SBS and users in dense wireless networks using artificial intelligence, introducing a new imitation learning mechanism that allows them to collaboratively improve their cellular association policies with minimal computation, helping users to select their serving BSs faster. Reference [78] provides a collaborative distributed computing framework consisting of UEs, edge cloud nodes, and cloud centers that Breaking the original standard MEC developed by ETSI. Reference [79] envisioned a MEC collaborative architecture to achieve resource sharing among MEC-BSs in UDNs, considering factors such as latency, resource consumption, and radio channel state, where multiple weights and numbers of tasks can be randomly delivered to different servers. However, considering simple computing task scenarios, the MD randomly requesting various computing tasks and the dynamic change of ES computing resources are ignored.

Meanwhile, due to the dynamic nature of MEC environments, if the load policy only considers the tasks of the current application and the load of each MEC server, it will result in additional load overhead between MEC servers, which is known as the ping-pong effect in load balancing. Reference [80] applies SDN to task allocation in MEC scenarios with UDN, and proposes a load balancing algorithm based on user load prediction load estimation based on the routing of the corresponding information between MEC servers, which is able to effectively reduce the impact of the ping-pong effect. In addition, subtask offloading of divisible tasks is discussed. Reference [81] proposes a user-centric SDN edge resource sharing model, specifically around a SBS where multiple MEC servers can share their 3C resources via OpenFlow-enabled switches that can dynamically routed to the appropriate SBS for MEC processing. While the average latency emphasized in the literature is meaningful, it is more reasonable to characterize the latency with a certain level of confidence, i.e., the probability of exceeding the maximum tolerance latency. Future work focuses on analyzing the delay in SDN-UDN from a probabilistic perspective.

In the actual MEC-enabled UDN, for the diverse offloading requirements of IoT applications in dynamic network environments, reference [82] investigates a distributed delay-constrained computational offloading strategy based on UDN computation and network coordination, and designed an extended game-theoretic approach based on the Lyapunov optimization theory, taking into account the uncertainties of users' mobility, the diversity of demands, and the limited edge resources. Distributed two-stage and multi-stage stochastic planning algorithms under various uncertainties are proposed. It should be noted that there is an additional cost in case of offloading failure. Reference [83] investigates computational offloading decisions in a hierarchical network architecture and proposes a joint channel allocation and resource management scheme, taking into account stochastic task arrivals/scheduling and dynamic changes in available resources. Future work could investigate the energy harvesting capability of mBS and other efficient server scaling algorithms. Reference [84] considers not only the case of MDs randomly requesting various computational tasks, but also the case of tasks randomly arriving at the edge servers and the dynamically changing computational resources of the edge services, and proposes a more computationally efficient two-tier game-theoretic greedy approximation offloading scheme, i.e., collaborative computational offloading among IoT MDs, edge servers connected to the MBS, and SCs.

In ultra-dense networks, when large-scale SMDs simultaneously offload computational tasks in dynamic wireless environments, the joint optimization of their offloading decisions becomes very complex. Reference [85] proposes an MFG-based distributed strategy to optimize the computation offloading decision of each SMD, so as to effectively utilize the transmission energy of SMDs and the computational resources of MECs. In which each SMD is able to dynamically adjust the offloading rate according to different channel conditions and the demand/competition of other SMDs for MEC computational resources. However, the authors only considered a simple channel-varying, generic traffic model, and future work could consider the computational tasks of dynamic SMDs and include MEC processing delays and queuing delays in the optimization problem.

The dense deployment of SBSs in UDNs results in a large amount of coverage overlap and leads to complex multi-cell networks with highly coupled resources. The SI placement problem becomes a challenging combinatorial optimization problem [86, 87]. As a prerequisite for computational offloading, most existing work has investigated centralized SI placement models, where a centralized manager exists to collect large amounts of network information in real time [88, 89]. In addition, the user is forced to reveal all local information to the centralized scheduler, which may not be desirable in a given application due to privacy concerns. Reference [90] investigates the SI placement problem in dynamic MEC-enabled UDNs and proposes a distributed strategy for multi-user multi-server edge systems in UDNs that can make SI placement decisions quickly and efficiently. The model focuses on the key practical issues of dynamic arrival tasks, network congestion effects and decentralization. Further research is conducted in the future on how to utilize reinforcement learning to empower edge devices and enhance distributed intelligence at the network edge. Table 4 summarizes and compares the literature mentioned above.

Table 4. Comparison of the literature

Characteristic	Reference
A load balancing algorithm based on user load prediction load estimation is proposed.	[80]
Aims to achieve resource sharing load balancing between MEC-BS in UDNs.	[77-79]
Aiming at the diverse offloading requirements of IoT applications in dynamic network environments.	[82-84]
A user-centric edge resource sharing model for SD-UDN is proposed.	[81]
Solving the dynamic uplink computing offloading decision problem.	[85]
SI placement in dynamic MEC-enabled UDNs is studied.	[90]

Allocation of Access Options. When considering transmission and computation requirements in MEC-based UDNs, differences in access channel quality, MEC server status, service requirements, and user experience should be taken into account when choosing access among many available APs in the neighborhood. It is necessary to consider access selection and RA in conjunction. Reference [91] proposes a joint scheme of AP grouping and RA in NOMA-based user-centric UDN, aiming to maximize the system EE. Reference [92] proposes a joint access selection AS and heterogeneous resource allocation scheme for NOMA-based UDN with MEC, in order to satisfy the different service requirements of the densely deployed UEs, and to combine the proposed AS algorithms with the UE's resource requirements, the AP's remaining resources, and the AP's and UE's remaining resources. s remaining resources and the distance between APs and UEs to maximize the system EE.

Reference [93] concentrates on selecting suitable edge nodes for 5G data offloading with minimum cost and energy using PSO algorithms. However, time constraints are not considered, and various population intelligence algorithms, including energy optimization based on time consumption, can be implemented and tested in 5G networks in the future to improve the overall network efficiency. In [94], a DQN based AC algorithm is proposed for selecting the most suitable edge servers to improve the learning efficiency. In [95], mobile users are grouped into a base station considering physical distance and workload in the server selection phase. After grouping, the original problem is divided into a parallel multiuser-to-one-server offloading decision subproblem, and a distributed offloading strategy based on binary coded genetic algorithm is designed to obtain adaptive offloading decisions. In the UDN architecture, the simultaneous deployment of macro eNodeB and small eNodeB MEC servers can effectively reduce the latency and energy consumption and improve the quality of user experience. For the access eNodeB selection problem, a user access selection algorithm considering MEC is proposed in UDN, which can effectively reduce latency and energy consumption [96]. When processing data in edge nodes with centralized units or distributed cloud units, the load of each edge node needs to be considered.

The following literature addresses the requirements of multi-user, multi-server and multi-base station scenarios. Reference [97] explores scheduling algorithms for multi-access edge computing in UDN, jointly considering transmission scheduling and computational resource scheduling, and designing a scalable model adapted to existing cellular networks in UDN scenarios. Future work could try joint scheduling, providing multiple scheduling options to improve efficiency. Reference [98] proposed a heuristic greedy computation offloading algorithm that performs computation offloading across multiple MEC servers, which can significantly reduce the overall energy consumption of MDs in UDNs, and this trend scales well as the number of task types increases. Reference [99] considers a multi-user multiple MEC scenario in an ultra-dense heterogeneous cellular network, assuming that each type of task is requested by a certain MD with a specified probability, aiming to minimize the power consumption and task execution latency by jointly optimizing the task allocation decision and CPU frequency. However, only the dual-connectivity case was considered, and future work could consider extending to other scenarios using the CoMP technique to study more MECO scenarios with mobile users and dynamic user states.

In UDN, users can be covered by multiple base stations, but the users themselves are not aware of the loads of the BSs. Reference [100] applies the idea of SDNs and proposes an edge-collaboration architecture based on global awareness to realize resource sharing and efficient offloading of tasks. In particular, considering the high load of local servers and the idle resources of remote servers, game theory is used to obtain the optimal offloading strategy for users. Reference [101] proposes a blockchain-based decentralized computing offloading coordination platform to develop an improved GS-based user matching algorithm to find the matching relationship between the offloading requester's computing tasks and the edge servers/UEs. Reference [102] proposes a scalable and sustainable IoT framework that integrates UDN-based hierarchical multiple access and computation offloading between MECs and the cloud, which can dramatically reduce the end-to-end latency and energy consumption of computing data for massive IoT devices. Furthermore, in a MEC-enabled UDN, in addition to the

UEs being able to offload computational tasks to edge servers, the UEs are able to offload computations to other UEs via D2D links in their neighborhood [103]. In [104], considering the use of non-orthogonal multiple access so that many UEs are able to transmit data at the same frequency and offload computation tasks to the edge server simultaneously through the corresponding BSs. Therefore, it is very challenging to coordinate the computation offloading between UEs and edge servers in UDN. Table 5 summarizes and compares the literature mentioned above.

Table 5. Comparison of the literature

Characteristic	Reference
A joint scheme for AP grouping and RA is proposed.	[91]
Proposed scheme for access selection AS and heterogeneous RA	[92]
Aims at selecting the most appropriate edge server.	[93, 94]
Performs computation offloading between multiple MEC servers.	[95, 96, 98]
For multi-user, multi-server and multi-base station scenarios.	[97, 99]
Coordinates computation offloading between UEs and edge servers.	[101, 104]

User Movement. In 5G UDN, APs are independent and equal, and the coverage area of APs is smaller. When the UE is moving, the interaction between the UE and the AP will be more frequent. Especially for high mobility users, the UE requires more energy consumption during switching, and the network performance will be degraded by high latency and high computational complexity in order to ensure the continuity of data transmission and reception during this back-and-forth switching process, which may lead to service disruption and extra cost [105]. Reference [106] describes a developed in-vehicle network and offloading mechanism that introduces a resource management model with real-time allocation and load balancing, and the proposed approach integrates task prioritization, multi-agent collaboration, context-aware decision making, and distributed learning to optimize network performance. Reference [107] studies the user-centric AP clustering problem in MEC-based UDN, in which the services provided by APs to users are modeled as M/M/1/L queuing processes, and mobility prediction techniques are utilized to achieve a dynamic AP clustering scheme, in which the clustering structure automatically adapts to the dynamic distribution of user traffic in a specific region. However, it ignores the fact that the density of APs, the prediction duration and the cluster size all affect the performance of AP clustering. Reference [108] proposes a mobile-aware adaptive flow rule placement scheme consisting of two parts: a path estimator and a flow manager. The path estimator predicts the future locations of end-users present in the network based on their historical location sets, thus minimizing the total cost associated with flow rule placement. However, its control overhead increases due to incorrect location predictions. Reference [109] investigateds how to perform intelligent routing based on the diverse service requirements of crowds to effectively control network traffic, and designed a DRL smart routing algorithm that learns the request distribution patterns of crowds and makes decisions accordingly, and simultaneously meets the latency constraints of various service requests from crowds, which can significantly improve the utilization of network infrastructure and avoid the delay in the presence of large numbers of network congestion when crowds are performing their daily activities.

Some studies related to service migration in MEC environments. In order to support service migration of MEC servers, a MEC platform was developed to implement mobility management requirements using an IP Mobility Support Gateway, applying the extended virtualized Mobility Support Gatewa, but ignoring the case of mobility for mobile users who need service migration [110]. Reference [111] developes a novel user-centric energy-aware mobility management (EMM) scheme to optimize the latency of radio access and computation under the constraint of long-term user energy consumption. Reference [112] proposes a solution to decompose data in the cloud into a set of files and services, and perform caching on the user's mobile device to quickly provide frequently requested files and services, taking into account the mobility of mobile users who need task migration. Reference [113] focuses on the combination of Virtual Network Function (VNF)/Service Function Chain (SFC) and MEC, and proposes an online algorithm, called Follow-Me Chain, to prevent unacceptably long service delays during inter-MEC switching for VNFs with the corresponding SFC. Future work investigates MEC server switching in high-density environments to enable high mobility users and regular users to migrate services together without sufficient resources. Reference [114] investigates the use of service-specific overlays in high-density environments such as stadiums or subway stations where users need data or services at the same time to provide faster data access to MDs from the cloud to the edge in congested environments by using data replication methods. Although these approaches provide effective solutions for service migration in UDNs, they lack consideration of high mobility scenarios and details of migrating services in MEC servers, and do not take into account the frequent movement and generation of tasks by mobile users. Reference [115] proposes a new mechanism for high-mobility users in ultra-dense MEC networks to help select appropriate MEC servers for migrated services, schedule service migration efficiently enough, and recover services faster even in blocking situations. Finally, a system prototype is implemented using "networkx" to perform real service migration and saturation tests are performed in UDNs.

This is because joint decision making among multiple MEC servers becomes difficult due to time synchronization and message exchange overhead. Reference [116] considers multiple offload scenarios where multiple MEC servers can be selected in each offload round, where MUs move randomly throughout the network. When tasks are migrated across multiple MEC system regions, or when task processing results are migrated due to the mobility of UEs, network connectivity may be interrupted, which can trigger unnecessary data retransmissions and signal interference. Therefore, effective management of network connectivity within different MEC systems is a pressing issue. Especially in the case of a large number of UEs traversing the MEC system area, it is crucial to coordinate SDN and multiple MEC systems to optimize computation offloading and resource allocation to reduce network energy consumption. Future research must consider coordination strategies for the management space and policies between SDN and MEC systems to develop excellent network connectivity management schemes for superior energy efficiency and network performance.

In addition, the architecture of the ultra-dense network creates more difficulties in information collection and delivery, and tends to saturate the backhaul network, leading to insufficient resources. It is difficult to synchronize accurate and complete system information between BSs and UE for making mobility management decisions [117]. Reference [118] proposes a cloud system architecture for mobile devices to address the buffering of mobile devices in dense and congested environments, proposing an OS-side architecture that can manage traffic from different selves and efficiently utilize cloud resources, but lacks consideration of MEC environments. Reference [119] proposes a 5G UDN security authentication scheme based on blockchain technology, and proposed a blockchain APG-PBFT algorithm based on the Byzantine Fault Tolerant consensus algorithm, which can be used with APs to generate a trusted chain APG and share the authentication results in the APG using the blockchain message propagation mechanism, which can reduce the frequency of authentication of the UEs when they are moving among the APs and improve the access efficiency. Therefore, it is a good solution idea for the UE to move smoothly in the trusted APG by reducing the frequent authentication. Table 6 summarizes and compares the literature mentioned above.

Table 6. Comparison of the literature

Characteristic	Reference
Reduce the authentication frequency when UE moves between APs and improve access efficiency.	[119]
Dynamic AP clustering using mobility prediction technique	[107]
A mobility-aware adaptive flow rule placement scheme is proposed.	[108]
Intelligent routing based on the diverse service requirements of the crowd for effective control of network traffic.	[109]
Research related to service migration schemes in MEC environments.	[110, 111, 112, 114, 115]
Extended to multiple offload scenarios where multiple MEC servers can be selected.	[116]
Prevents unacceptably long service delays when switching between MECs.	[113]

3.3 Interference

SBSs are deployed at a very high density in 5G UDNs, which, on the one hand, shorten the distance between users and interfering APs, generating more and stronger sources of interference. On the other hand, there are many idle SBSs in the UDN that do not serve any UE , and many base stations become major sources of interference to each other, and their coordination requires complex mechanisms [120]. In densely deployed small base stations, the idle cells of unconnected users can be turned off, which can partially mitigate their interference to neighboring cells. In many cases, turning on and using these dormant SBSs to simultaneously serve any nearby UEs and their written SBSs not only increases the required signal power, but also reduces the number of switching processes for high-speed mobile UEs. In dense small cell networks, activating the smart idle mode feature can similarly suppress interference and can minimize its power consumption without having users connected to it. RA also performs multidomain interference management in the frequency, time, space, cache, and computational resource domains simultaneously.

Reference [121] investigates the characteristics of interference in UDN and summarizes some of the mainstream IM techniques. Due to the complexity of interference, most IM techniques designed for sparse networks lose their advantages when applied in UDN. An IM entity is designed for UDN and an adaptive switching power control method is proposed, which can significantly improve the network capacity. Inter-cell interference (ICI) caused by different cells using the same spectrum at the same time. The problem is compounded if the femtocells are located in the cell edge region of the macrocells.ICI reduces system throughput and network capacity and negatively affects the cell edge users.Reference [122] provides a comprehensive survey of methods to reduce femtocell UDN interference. Reference [123] considers user association and power allocation in millimeter-wave based UDN with load balancing constraints, energy harvesting at the base station, user quality of service requirements, energy efficiency, and cross-layer interference limitations. An effective interference coordination mechanism is proposed in conjunction with the energy harvesting at the base station to cognitively limit the interference between base stations and users in ultra-dense networks. However, the density of APs, the prediction duration and the cluster size can significantly affect the performance of AP clustering in the actual UDN deployment, which needs to be carefully designed to be realized.

Inter-subdivision interference may jeopardize the gains of network densification, CoMP transmission is very important in UDNs. MEC and CoMP are different technologies, and CoMP and MEC work in tandem to provide latency-sensitive services in UDNs. Also, the storage capacity of each SBS in UDN is limited and the cache hit probability can be significantly improved with the help of CoMP. Reference [124] proposes a CoMP-based uplink transmission scheme to collaborate on receiving tasks from tagged users, however, due to the user-centered coordination, it cannot solve the scheduling problem of multi-user and multi-base station schemes. Reference [125] aims to bridge the gap between MEC and CoMP based on three MEC functions, demonstrating that CoMP transport reduces UDN complexity and transmission delay through collaboration between MEC servers in UND. At the same time, it can reduce the backhaul pressure of CoMP transport and make it possible to realize scalable CoMP. All three CoMP approaches can effectively eliminate interference in the same collaboration set, but only the salient issues of interference in static UEs. UDNs are considered: First, when SBSs sharing the same spectrum are located in dense locations, the overlapping regions between their coverage areas expand simultaneously, leading to an increased likelihood of UEs experiencing extremely severe interference in these areas.Second, with the deployment of small size of small base stations, the interference level is enhanced with theLoS propagation from neighboring cells. Third, is the fact that SBSs in UDNs are deployed randomly based on the random distribution of UEs, and interference also has a random nature that is difficult to manage.

The direction of interference management in future wireless networks remains an open question, and the key is how to design self-organizing interference management techniques to mitigate the impact of channel and user behavioral correlations on self-organizing IM. In addition to scalability, stability, and agility, density conditions should be considered, which means that the overhead and complexity of IM solutions should not increase infinitely with network densification. In order to take full advantage of network densification, it is necessary to characterize the interference in UDNs, which is no longer satisfied by the traditional implementation of IM in a centralized manner. In particular, the network capacity may even approach zero when the network infrastructure is sufficiently dense in UDNs. Future research should aim at designing efficient multilayer structural power control and cell association algorithms to control interference, while supporting the association of users with different base stations and analyzing the trade-off between reducing the interference level and increasing the resource utilization, among others.

4 Challenges and Future Work

4.1 Comprehensive Consideration of Configuration Elements for Dynamic Scenarios

Handling dynamic environments is crucial in modern MEC applications, particularly in highly mobile scenarios such as UAV-assisted MEC networks and Telematics. Existing research often provides simple dynamic configurations, but these strategies frequently lack detailed considerations of task size, computational requirements, and processing time. Additionally, these approaches often rely on unrealistic assumptions about the availability of CSI in real-world deployments, which is especially problematic in high-speed mobile environments.

More complex dynamic scenarios require addressing not only transient performance objectives (e.g., latency, reliability, and energy efficiency) but also long-term system performance and stability. For instance, in MEC-enabled vehicular networks, existing approaches to computational offloading and resource allocation focus on minimizing transient latency or reliability, or maximizing transient energy efficiency [126, 127]. While these goals can be achieved, the practical implementation of such approaches can be computationally intensive. This is unrealistic due to the high mobility of connected vehicles (CVs) and the assumption that ideal CSI is available [128]. Furthermore, it is necessary to consider optimizing dynamic elements such as trajectory, velocity, and acceleration to adapt to changing network conditions and external environments.

In UAV-assisted MEC networks, it is crucial to dynamically optimize realistic factors such as drive trajectories, speeds, and turning angles [129]. UAV-assisted missions, such as civil infrastructure inspection, precision agriculture, and search and rescue operations, require not only low latency and high reliability but also energy efficiency and mission execution accuracy. Therefore, future research needs to develop more sophisticated and comprehensive dynamic configuration strategies that satisfy immediate computational and communication needs while adapting to long-term system development and changes. This includes, but is not limited to, using advanced machine learning and artificial intelligence techniques to adapt policies in real-time and developing new algorithms to handle large-scale, highly dynamic data streams and task offloading requirements.

4.2 Considering Interactions between Dynamic Elements

In dynamic and complex MEC networks, the simultaneous occurrence of multiple interdependent operations presents several challenges. For instance, dynamic changes in the number of tasks offloaded to the MEC server directly affect the server's CPU processing power and speed. Additionally, when the MEC server returns a large amount of result data to the UE, frequent data reception and download operations can significantly drain the UE's battery. Fluctuations in the UE's battery power also impact the efficiency of task generation, offloading, and local execution. These factors necessitate precise tuning and optimization of computational offloading and RA strategies to ensure network performance and user experience.

Literature [130] proposes methods for deciding whether a task should be executed locally or offloaded to a nearby server at runtime based on server availability and estimated invocation latency, effectively balancing response time and device energy consumption. Literature [131] further investigates the joint impact of task prioritization and mobile computing load on MEC network performance by analyzing the wireless bandwidth of the network and the computational power of UAVs to improve overall system utility. Literature [132] explores modeling tasks as directed acyclic graphs (DAGs) and using priority-based linear transformations for task offloading to adapt to multi-tasking scenarios with dependencies in a satellite IoT environment, achieving efficient resource utilization and low-latency services.

Therefore, future work needs to comprehensively consider the interactions between these dynamic factors and design more intelligent and adaptive computation offloading and resource allocation strategies to minimize network costs and improve service quality.

4.3 Consider Strategies for Task Migration, Task Discarding and Task Retransmission

In dynamic and complex MEC networks, optimizing task migration, discarding and retransmission policies becomes critical to ensure network QoS in the face of network congestion, high mobility of UE, and limitations of MEC server resources. The following strategies suggest ways to address these challenges:

Predictive task migration strategies: where network performance is predictable, RL can be used to develop strategies that leverage historical and real-time data to assess when to migrate tasks to maximize QoS.

Dynamic Migration Decision Making: When network performance is unstable or unpredictable, real-time analysis of network state and resource availability is required to make decisions between migrating immediately and waiting for a better time.

Resource and Cost Evaluation: Before deciding where to migrate, it is important to assess the resources available for the target servers, the bandwidth requirements during the migration, and the cost of the migration.

Managing migration task dependencies and complexity: Given that network conditions may change during task migration, multiple related migration tasks must be managed. This includes setting constraints to account for inter-task dependencies, ensuring effective task execution and maintaining system stability even in a dynamically changing network environment.

4.4 Deployment of Backhaul Networks in UDNs

With the increasing convergence of mobile networks and the Internet, customer demand for augmented experiences continues to rise. Emerging services such as virtual reality (VR), augmented reality (AR), user-centered computing, and telemedicine all require high QoS and QoE with tactile-level latency, ultra-high reliability, security, and privacy. For example, the model proposed in the literature [133] aims to optimize the delivery of scalable 360-degree video content for mobile VR users by intelligently allocating ground- and air-edge computation/communication resources, as well as locating drone BSs. However, the architectural characteristics of UDNs often lead to resource constraints in backhaul networks. Synchronizing accurate and comprehensive system information between BSs and UEs faces many challenges, especially when dealing with frequent service migration and switching of high-mobility users, which makes it crucial to ensure network reliability and quality of service. In the face of these challenges, future work should focus on the following key areas:

Optimize backhaul network architecture: Develop backhaul network solutions that can adapt to the demands of high-density traffic to improve network flexibility and scalability.

Improve data synchronization techniques: Develop more efficient data synchronization methods to ensure fast and accurate exchange of information between BSs and UEs, especially in high user density environments, in order to maintain continuity and accuracy of network operations.

Research on QoS for Service Migration: In-depth research on QoS issues during service migration, especially on how to manage service migration and switching in high mobility environments to minimize service disruptions and delays.

Jointly Optimize Backhaul Capacity and Power Consumption: Considering the capacity limitations and power consumption constraints of backhaul networks, develop joint optimization strategies to balance network performance and energy efficiency, especially in resource-constrained ultra-dense network environments.

5 Conclusion

In this paper, we systematically explore the resource allocation issues for MEC-based UDNs, focusing on key aspects such as optimization criteria, dynamic environments, user mobility, and interference management. In addition, we discuss load balancing, access selection and interference control strategies in dynamic environments. As UDN density continues to evolve, future work can extend to more complex scenarios, fully multidimensionalize the configuration elements of real dynamic scenarios, and explore the load balancing problem among cloudlet nodes in Hetworks.

6 Acknowledgement

This paper is sponsored by “Key Laboratory of Computing Power Network and Information Security, Ministry of Education”.

References

[1] F. Salahdine, J. Opadere, Q. Liu, T. Han, N. Zhang, S. Wu, A survey on sleep mode techniques for ultra-dense networks in 5G and beyond, Computer Networks 201(2021) 108567.
[2] M. Kamel, W. Hamouda, A. Youssef, Ultra-dense networks: A survey, IEEE Communications surveys & tutorials 18(4) (2016) 2522-2545.
[3] S. Chen, R. Ma, H. Chen, H. Zhang, W. Meng, J. Liu, Machine-to-machine communications in ultra-dense networks—A survey, IEEE Communications Surveys & Tutorials 19(3)(2017) 1478-1503.
[4] H. Han, W. Zhai, J. Zhao, Smart city enabled by 5G/6G networks: An intelligent hybrid random access scheme, arXiv

preprint arXiv(2021) 2101.06421.
[5] G. Bartoli, R. Fantacci, D. Marabissi, Resource allocation approaches for two-tiers machine-to-machine communications in an interference limited environment, IEEE Internet of Things Journal 6(5)(2019) 9112-9122.
[6] Y. Wu, N. Zhang, K. Rong, Non-orthogonal random access and data transmission scheme for machine-to-machine communications in cellular networks, IEEE Access 8(2020) 27687-27704.
[7] L. Miuccio, D. Panno, S. Riolo, A new contention-based PUSCH resource allocation in 5G NR for mMTC scenarios, IEEE Communications Letters 25(3)(2021) 802-806.
[8] W.U. Rehman, T. Salam, A. Almogren, K. Haseeb, I.U. Din, S.H. Bouk, Improved resource allocation in 5G MTC networks, IEEE Access 8(2020) 49187-49197.
[9] L. Miuccio, D. Panno, S. Riolo, Joint control of random access and dynamic uplink resource dimensioning for massive MTC in 5G NR based on SCMA, IEEE Internet of Things Journal 7(6)(2020) 5042-5063.
[10] A.M. Atchome, H.H.R. Sherazi, R. Alahassa, F. Tossa, T.O. Edoh, L.A. Grieco, A.C. Vianou, A hybrid network model embracing nb-iot and d2d communications: Stochastic geometry analysis, in: Proc. 11th EAI International Conference, AFRICOMM 2019, 2019.
[11] S. Hamdoun, A. Rachedi, Y. Ghamri-Doudane, Graph-based radio resource sharing schemes for MTC in D2D-based 5G networks, Mobile Networks and Applications 25(3)(2020) 1095-1113.
[12] N. Alhussien, T.A. Gulliver, Optimal resource allocation in cellular networks with H2H/M2M coexistence, IEEE Transactions on Vehicular Technology 69(11)(2020) 12951-12962.
[13] A. Saddoud, W. Doghri, E. Charfi, L.C. Fourati, 5G dynamic borrowing scheduler for IoT communications, in: Proc. 15th international wireless communications & mobile computing conference, 2019.
[14] Z. Zhou, X. Chen, Y. Zhang, S. Mumtaz, Blockchain-empowered secure spectrum sharing for 5G heterogeneous networks, IEEE Network 34(1)(2020) 24-31.
[15] Z. Zhang, G. Yang, Z. Ma, M. Xiao, Z. Ding, P. Fan, Heterogeneous ultradense networks with NOMA: System architecture, coordination framework, and performance evaluation, IEEE vehicular technology magazine 13(2)(2018) 110-120.
[16] H.H. He, J. Jiang, R. Jin, Research on Downlink Precoding for Interference Cancellation in Massive MIMO Heterogeneous UDN, Journal of Applied Mathematics and Physics 6(1)(2018) 283-289.
[17] M. Susanto, S.N. Hasim, H. Fitriawan, Interference Management with Dynamic Resource Allocation Method on Ultra-Dense Networks in Femto-Macrocellular Network, Jurnal Rekayasa Elektrika 17(4)(2021) 230-238.
[18] M. Salhani, M. Liinaharja, Load Balancing Algorithm within the Small Cells of Heterogeneous UDN Networks: Mathematical Proofs, Journal of Communications 13(11)(2018) 627-634.
[19] Q. Liu, J. Shi, Base station sleep and spectrum allocation in heterogeneous ultra-dense networks, Wireless Personal Communications 98(4)(2018) 3611-3627.
[20] Y. Zhong, X. Ge, H.H. Yang, T. Han, Q. Li, Traffic matching in 5G ultra-dense networks, IEEE Communications Magazine 56(8)(2018) 100-105.
[21] S. Chen, F. Qin, B. Hu, X. Li, Z. Chen, User-centric ultra-dense networks for 5G: Challenges, methodologies, and directions, IEEE Wireless Communications 23(2)(2016) 78-85.
[22] Y. Lin, R. Zhang, L. Yang, C. Li, L. Hanzo, User-centric clustering for designing ultradense networks: Architecture, objective functions, and design guidelines, IEEE Vehicular Technology Magazine 14(3)(2019) 107-114.
[23] P. Koleva, V. Poulkov, Heuristic Access Points Grouping for Mobility Driven User-Centric Ultra Dense Networks, Wireless Personal Communications 126(1)(2022) 1-24.
[24] C. Wang, B. Hu, S. Chen, Y. Wang, Joint dynamic access points grouping and resource allocation for coordinated transmission in user-centric UDN, Transactions on Emerging Telecommunications Technologies 29(3)(2018) e3265.
[25] B. Hu, Y. Wang, C. Wang, A maximum data transmission rate oriented dynamic APs grouping scheme in user-centric UDN, in: Proc. 2017 International Symposium on Intelligent Signal Processing and Communication Systems, 2017.
[26] Y. Liu, X. Li, H. Ji, H. Zhang, A multiple APs cooperation access scheme for energy efficiency in UUDN with NOMA, in: Proc. IEEE Conference on Computer Communications Workshops, 2017.
[27] Y. Lin, R. Zhang, L. Yang, L. Hanzo, Modularity-based user-centric clustering and resource allocation for ultra dense networks, IEEE Transactions on Vehicular Technology 67(12)(2018) 12457-12461.
[28] H. Zhang, S. Huang, C. Jiang, K. Long, V.C.M. Leung, H.V. Poor, Energy efficient user association and power allocation in millimeter-wave-based ultra dense networks with energy harvesting base stations, IEEE Journal on Selected Areas in Communications 35(9)(2017) 1936-1947.
[29] Y. Liu, X. Fang, P. Zhou, K. Cheng, Coalition game for user association and bandwidth allocation in ultra-dense mmWave networks, in: Proc. 2017 IEEE/CIC International Conference on Communications in China, 2017.
[30] Z. Qin, X. Yue, Y. Liu, Z. Ding, A. Nallanathan, User association and resource allocation in unified NOMA enabled heterogeneous ultra dense networks, IEEE Communications Magazine 56(6)(2018) 86-92.
[31] N. Sapountzis, T. Spyropoulos, N. Nikaein, U. Salim, Joint optimization of user association and dynamic TDD for ultra-dense networks, in: Proc. IEEE INFOCOM 2018-IEEE Conference on Computer Communications, 2018.
[32] W.K. Lai, C.Y. Hsu, Y.P. Kuo, QoS-guaranteed User Association and Resource Allocation with CoMP JT in Ultra-Dense Networks, in: Proc. 20th Asia-Pacific Network Operations and Management Symposium, 2019.
[33] D. Liu, L. Wang, Y. Chen, M. Elkashlan, K.K. Wong, R. Schober, L. Hanzo, User association in 5G networks: A survey and an outlook, IEEE Communications Surveys & Tutorials 18(2)(2016) 1018-1044.

[34] S. Andreev, V. Petrov, M. Dohler, H. Yanikomeroglu, Future of ultra-dense networks beyond 5G: Harnessing heterogeneous moving cells, IEEE Communications Magazine 57(6)(2019) 86-92.
[35] N. Sharma, K. Kumar, Resource allocation trends for ultra dense networks in 5G and beyond networks: A classification and comprehensive survey, Physical Communication 48(2021) 101415.
[36] 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 Communications Surveys & Tutorials 21(3)(2019) 2134-2168.
[37] C.F. Kwong, C. Shi, Q. Liu, S. Yang, D. Chieng, P. Kar, Autonomous handover parameter optimisation for 5G cellular networks using deep deterministic policy gradient, Expert Systems with Applications 246 (2024) 122871.
[38] F. Qamar, S. H. Ali Kazmi, M. N. Hindia, R. Hassan and D. P. Dahnil, Enhancing 5G UDN Performance with Adaptive Successive Interference Cancellation Technique for Optimizing Success Probability and Ergodic Capacity,2024 IEEE 7th International Symposium on Telecommunication Technologies (ISTT), Langkawi Island, Malaysia, 2024, pp. 96-101.
[39] A. Gotsis, S. Stefanatos, A. Alexiou, UltraDense networks: The new wireless frontier for enabling 5G access, IEEE Vehicular Technology Magazine 11(2)(2016) 71-78.
[40] S.F. Yunas, M. Valkama, J. Niemelä, Spectral and energy efficiency of ultra-dense networks under different deployment strategies, IEEE Communications Magazine 53(1)(2015) 90-100.
[41] R. Baldemair, T. Irnich, K. Balachandran, E. Dahlman, G. Mildh, Y. Selén, S. Parkvall, M. Meyer, A. Osseiran, Ultra-dense networks in millimeter-wave frequencies, IEEE Communications Magazine 53(1)(2015) 202-208.
[42] M. Kamel, W. Hamouda, A. Youssef, Performance analysis of multiple association in ultra-dense networks, IEEE Transactions on Communications 65(9)(2017) 3818-3831.
[43] M. Diamanti, C. Pelekis, E.E. Tsiropoulou, S. Papavassiliou, Delay Minimization for Rate-Splitting Multiple Access-Based Multi-Server MEC Offloading, IEEE/ACM Transactions on Networking 32(2)(2024) 1035-1047.
[44] S. Wang, M. Chen, X. Liu, C. Yin, Task and resource allocation in mobile edge computing: An improved reinforcement learning approach, in: Proc. 2019 IEEE Globecom Workshops, 2019.
[45] H. Xing, L. Liu, J. Xu, A. Nallanathan, Joint task assignment and resource allocation for D2D-enabled mobile-edge computing, IEEE Transactions on Communications 67(6)(2019) 4193-4207.
[46] F. Wang, H. Xing, J. Xu, Optimal resource allocation for wireless powered mobile edge computing with dynamic task arrivals, in: Proc. ICC 2019-2019 IEEE International Conference on Communications, 2019.
[47] B. Dab, N. Aitsaadi, R. Langar, A novel joint offloading and resource allocation scheme for mobile edge computing, in: Proc. 16th IEEE Annual Consumer Communications & Networking Conference, 2019.
[48] Z. Ding, J. Xu, O.A. Dobre, H.V. Poor, Joint power and time allocation for NOMA–MEC offloading, IEEE Transactions on Vehicular Technology 68(6)(2019) 6207-6211.
[49] Z. Ding, P. Fan, H.V. Poor, Impact of non-orthogonal multiple access on the offloading of mobile edge computing, IEEE Transactions on Communications 67(1)(2018) 375-390.
[50] P.A. Apostolopoulos, G. Fragkos, E.E. Tsiropoulou, S. Papavassiliou, Data offloading in UAV-assisted multi-access edge computing systems under resource uncertainty, IEEE Transactions on Mobile Computing 22(1)(2023) 175-190.
[51] F. Zhao, Y. Chen, Y. Zhang, Z. Liu, X. Chen, Dynamic offloading and resource scheduling for mobile-edge computing with energy harvesting devices, IEEE Transactions on Network and Service Management 18(2)(2021) 2154-2165.
[52] Y. Chen, N. Zhang, Y. Zhang, X. Chen, Wen Wu, X.S. Shen, TOFFEE: Task offloading and frequency scaling for energy efficiency of mobile devices in mobile edge computing, IEEE Transactions on Cloud Computing 9(4)(2021) 1634-1644.
[53] Zhang, Q., Li, C., Huang, Y., & Luo, Y, Effective multi-controller management and adaptive service deployment strategy in multi-access edge computing environment, Ad Hoc Networks 138 (2023): 103020.
[54] W. Zhan, C. Luo, G. Min, C. Wang, Q. Zhu, H. Duan, Mobility-aware multi-user offloading optimization for mobile edge computing, IEEE Transactions on Vehicular Technology 69(3)(2020) 3341-3356.
[55] T. Ouyang, X. Chen, Z. Zhou, R. Li, X. Tang, Adaptive user-managed service placement for mobile edge computing via contextual multi-armed bandit learning, IEEE Transactions on Mobile Computing 22(3)(2023) 1313-1326.
[56] E.E. Haber, T.M. Nguyen, C. Assi, W. Ajib, An energy-efficient task offloading solution for MEC-based IoT in ultra-dense networks, in: Proc. IEEE Wireless Communications and Networking Conference, 2019.
[57] A.A. Ateya, A.D. Algarni, M. Hamdi, A. Koucheryavy, N.F. Soliman, Enabling heterogeneous IoT networks over 5G networks with ultra-dense deployment—using MEC/SDN, Electronics 10(8)(2021) 910.
[58] P. Sun, H. Zhang, H. Ji, L. Xi, Small cells clustering and resource allocation in dense network with mobile edge computing, in: Proc. 2019 IEEE Wireless Communications and Networking Conference, 2019.
[59] M. Elbayoumi, W. Hamouda, A. Youssef, Edge computing and multiple-association in ultra-dense networks: Performance analysis, IEEE Transactions on Communications 70(8)(2022) 5098-5112.
[60] S. Yu, X. Chen, Z. Zhou, X. Gong, D. Wu, When deep reinforcement learning meets federated learning: Intelligent multitimescale resource management for multiaccess edge computing in 5G ultradense network, IEEE Internet of Things Journal 8(4)(2021) 2238-2251.
[61] M. Chen, Y. Hao, Task offloading for mobile edge computing in software defined ultra-dense network, IEEE Journal on Selected Areas in Communications 36(3)(2018) 587-597.
[62] 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.
[63] B. Gu, M. Alazab, Z. Lin, X. Zhang, J. Huang, AI-enabled task offloading for improving quality of computational expe-

rience in ultra dense networks, ACM Transactions on Internet Technology (TOIT) 22(3)(2022) 1-17.
[64] F. Li, H. Yao, J. Du, C. Jiang, Z. Han, Y. Liu, Auction design for edge computation offloading in SDN-based ultra dense networks, IEEE Transactions on Mobile Computing 21(5)(2022) 1580-1595.
[65] Q. Zhang, L. Gui, F. Hou, J. Chen, S. Zhu, F. Tian, Dynamic task offloading and resource allocation for mobile-edge computing in dense cloud RAN, IEEE Internet of Things Journal 7(4)(2020) 3282-3299.
[66] L. Li, Q. Cheng, X. Tang, T. Bai, W. Chen, Z. Ding, Resource allocation for NOMA-MEC systems in ultra-dense networks: A learning aided mean-field game approach, IEEE Transactions on Wireless Communications 20(3)(2021) 1487-1500.
[67] S. Pang, S. Wang, Joint wireless source management and task offloading in ultra-dense network, IEEE Access 8(2020) 52917-52926.
[68] B. Li, P. Hou, H. Wu, F. Hou, Optimal edge server deployment and allocation strategy in 5G ultra-dense networking environments, Pervasive and Mobile Computing 72(2021) 101312
[69] D.G. Zhang, W.M. Dong, T. Zhang, J. Zhang, P. Zhang, G.-X. Sun, New computing tasks offloading method for mec based on prospect theory framework, IEEE Transactions on Computational Social Systems 11(1)(2024) 770-781.
[70] Y. Gao, H. Zhang, F. Yu, Y. Xia, Y. Shi, Joint computation offloading and resource allocation for mobile-edge computing assisted ultra-dense networks, Journal of Communications and Information Networks 7(1)(2022) 96-106.
[71] F. Ke, Y. Lin, Y. Liu, H. Zhou, M. Wen, Q. Zhang, Task Offloading, Caching and Matching in Ultra-Dense Relay Networks, IEEE Transactions on Vehicular Technology 72(3)(2023) 4010-4025.
[72] W. Liang, S. Ma, S. Yang, B. Zhang, A. Gao, Hierarchical Matching Algorithm for Relay Selection in MEC-Aided Ultra-Dense UAV Networks, Drones 7(9)(2023) 579.
[73] Y. Lu, X. Chen, Y. Zhang, Y. Chen, Cost-efficient resources scheduling for mobile edge computing in ultra-dense networks, IEEE Transactions on Network and Service Management 19(3)(2022) 3163-3173.
[74] Z. Cheng, M. Min, Z. Gao, L. Huang, Joint task offloading and resource allocation for mobile edge computing in ultra-dense network, GLOBECOM 2020-2020 IEEE Global Communications Conference, 2020.
[75] R. Zhang, P. Cheng, Z. Chen, S. Liu, B. Vucetic, Y. Li, Calibrated bandit learning for decentralized task offloading in ultra-dense networks, IEEE Transactions on Communications 70(4)(2022) 2547-2560.
[76] I. Al-Hammadi, M. Li, S.M.N. Islam, Independent tasks scheduling of collaborative computation offloading for SDN-powered MEC on 6G networks, Soft Computing 27(14)(2023) 9593-9617.
[77] K. Hamidouche, A.T.Z. Kasgari, W. Saad, M. Bennis, M. Debbah, Collaborative artificial intelligence (AI) for user-cell association in ultra-dense cellular systems, in: Proc. 2018 IEEE International Conference on Communications Workshops, 2018.
[78] T.X. Tran, A. Hajisami, P. Pandey, D. Pompili, Collaborative mobile edge computing in 5G networks: New paradigms, scenarios, and challenges, IEEE Communications Magazine 55(4)(2017) 54-61.
[79] T. Yang, H. Zhang, H. Ji, X. Li, Computation collaboration in ultra dense network integrated with mobile edge computing, in: Proc. 2017 IEEE 28th Annual International Symposium on Personal, Indoor, and Mobile Radio Communications, 2017.
[80] H. Guo, W. Huang, J. Liu, Y. Wang, Inter-server collaborative federated learning for ultra-dense edge computing, IEEE Transactions on Wireless Communications 21(7)(2022) 5191-5203.
[81] D. Wu, J. Yan, H. Wang, R. Wang, User-centric edge sharing mechanism in software-defined ultra-dense networks, IEEE Journal on Selected Areas in Communications 38(7)(2020) 1531-1541.
[82] S. Xia, Z. Yao, Y. Li, Z. Xing, S. Mao, Distributed Computing and Networking Coordination for Task Offloading Under Uncertainties, IEEE Transactions on Mobile Computing 23(5)(2024) 5280-5294.
[83] J. Ren, K.M. Mahfujul, F. Lyu, S. Yue, Y. Zhang, Joint channel allocation and resource management for stochastic computation offloading in MEC, IEEE Transactions on Vehicular Technology 69(8)(2020) 8900-8913.
[84] H. Guo, J. Liu, J. Zhang, W. Sun, N. Kato, Mobile-edge computation offloading for ultradense IoT networks, IEEE Internet of Things Journal 5(6)(2018) 4977-4988.
[85] R. Zheng, H. Wang, M. De. Mari, M. Cui, X. Chu, T.Q.S. Quek, Dynamic computation offloading in ultra-dense networks based on mean field games, IEEE Transactions on Wireless Communications 20(10)(2021) 6551-6565.
[86] A. Brogi, S. Forti, C. Guerrero, I. Lera, How to place your apps in the fog: State of the art and open challenges, Software: Practice and Experience 50(5)(2020) 719-740.
[87] F.A. Salaht, F. Desprez, A. Lebre, An overview of service placement problem in fog and edge computing, ACM Computing Surveys 53(3)(2021) 1-35.
[88] T. Ouyang, Z. Zhou, X. Chen, Follow me at the edge: Mobility-aware dynamic service placement for mobile edge computing, IEEE Journal on Selected Areas in Communications 36(10)(2018) 2333-2345.
[89] H. Ma, Z. Zhou, X. Chen, Predictive service placement in mobile edge computing, in: Proc. 2019 IEEE/CIC International Conference on Communications in China (ICCC), 2019.
[90] J. He, D. Zhang, X. Li, Y. Zhou, Y. Zhang, Distributed Service Placement in Ultra-Dense Edge Computing: A Game-Theoretical Approach, in: Proc. 2021 IEEE 23rd Int Conf on High Performance Computing & Communications; 7th Int Conf on Data Science & Systems; 19th Int Conf on Smart City; 7th Int Conf on Dependability in Sensor, Cloud & Big Data Systems & Application, 2021.
[91] Y. Liu, X. Li, F.R. Yu, H. Ji, H. Zhang, V.C.M. Leung, Grouping and cooperating among access points in user-centric ultra-dense networks with non-orthogonal multiple access, IEEE Journal on Selected Areas in Communications 35(10)

(2017) 2295-2311.
[92] S. Seng, X. Li, H. Ji, H. Zhang, Joint access selection and heterogeneous resources allocation in UDNs with MEC based on non-orthogonal multiple access, in: Proc. 2018 IEEE international conference on communications workshops, 2018.
[93] N. Bacanin, M. Antonijevic, T. Bezdan, M. Zivkovic, K. Venkatachalam, S. Malebary, Energy efficient offloading mechanism using particle swarm optimization in 5G enabled edge nodes, Cluster Computing 26(1)(2023) 587-598.
[94] C. Duo, P. Dong, Q. Gao, B. Li, Y. Li, MEC Computation Offloading-Based Learning Strategy in Ultra-Dense Networks, Information 13(6)(2022) 271.
[95] Z. Liao, J. Peng, B. Xiong, J. Huang, Adaptive offloading in mobile-edge computing for ultra-dense cellular networks based on genetic algorithm, Journal of Cloud Computing 10(1)(2021) 1-16.
[96] J. Zhao, W. Wu, G. Xiao, F. Chao, Y. Zhang, Access selection considering mobile edge computing in ultra dense network, in: Proc. 2017 3rd IEEE International Conference on Computer and Communications, 2017.
[97] V.N. Udupa, V.K. Tumuluru, Clustering and Reinforcement Learning based Multi-Access Edge Computing in Ultra Dense Networks, in: Proc. 2023 International Conference on Artificial Intelligence and Smart Communication, 2023.
[98] H. Guo, J. Liu, J. Zhang, Computation offloading for multi-access mobile edge computing in ultra-dense networks, IEEE Communications Magazine 56(8)(2018) 14-19.
[99] T.Q. Dinh, J. Tang, Q.D. La, T.Q.S. Quek, Offloading in mobile edge computing: Task allocation and computational frequency scaling, IEEE Transactions on Communications 65(8)(2017) 3571-3584.
[100] P. Yang, Y. Zhang, J. Lv, Load optimization based on edge collaboration in software defined ultra-dense networks, IEEE Access 8(2020) 30664-30674.
[101] S. Seng, X. Li, C. Luo, H. Ji, H. Zhang, A D2D-assisted MEC computation offloading in the blockchain-based framework for UDNs, in: Proc. ICC 2019-2019 IEEE International Conference on Communications, 2019.
[102] L.P. Qian, Y. Wu, B. Ji, L. Huang, D.H.K. Tsang, HybridIoT: Integration of hierarchical multiple access and computation offloading for IoT-based smart cities, IEEE network 33(2)(2019) 6-13.
[103] R. Wang, J. Yan, D. Wu, H. Wang, Q. Yang, Knowledge-centric edge computing based on virtualized D2D communication systems, IEEE Communications Magazine 56(5)(2018) 32-38.
[104] A. Kiani, N. Ansari, Edge computing aware NOMA for 5G networks, IEEE Internet of Things Journal 5(2)(2018) 1299-1306.
[105] X. Guan, X. Wan, F. Ye, B.Y. Choi, Handover minimized service region partition for mobile edge computing in wireless metropolitan area networks, in: Proc. 2018 IEEE International Smart Cities Conference, 2018.
[106] H. Min, A.M. Rahmani, P. Ghaderkourehpaz, K. Moghaddasi, M. Hosseinzadeh, A joint optimization of resource allocation management and multi-task offloading in high-mobility vehicular multi-access edge computing networks, Ad Hoc Networks 166 (2025) 103656.
[107] S. He, T. Wang, S. Wang, Mobility-driven user-centric AP clustering in mobile edge computing-based ultra-dense networks, Digital Communications and Networks 6(2)(2020) 210-216.
[108] S. Bera, S. Misra, M.S. Obaidat, Mobi-flow: Mobility-aware adaptive flow-rule placement in software-defined access network, IEEE Transactions on Mobile Computing 18(8)(2019) 1831-1842.
[109] L. Zhao, J. Wang, J. Liu, N. Kato, Routing for crowd management in smart cities: A deep reinforcement learning perspective, IEEE Communications Magazine 57(4)(2019) 88-93.
[110] T. Kondo, K. Isawaki, K. Maeda, Development and evaluation of the MEC platform supporting the edge instance mobility, in: Proc. 2018 IEEE 42nd Annual Computer Software and Applications Conference, 2018.
[111] Y. Sun, S. Zhou, J. Xu, EMM: Energy-aware mobility management for mobile edge computing in ultra dense networks, IEEE Journal on Selected Areas in Communications 35(11)(2017) 2637-2646.
[112] I.A. Ridhawi, M. Aloqaily, Y. Kotb, Y.A. Ridhawi, Y. Jararweh, A collaborative mobile edge computing and user solution for service composition in 5G systems, Transactions on Emerging Telecommunications Technologies 29(11)(2018) e3446.
[113] Y.T. Chen, W. Liao, Mobility-aware service function chaining in 5G wireless networks with mobile edge computing, in: Proc. ICC 2019-2019 IEEE International Conference on Communications, 2019.
[114] M. Aloqaily, I.A. Ridhawi, H.B. Salameh, Y. Jararweh, Data and service management in densely crowded environments: Challenges, opportunities, and recent developments, IEEE Communications Magazine 57(4)(2019) 81-87.
[115] H. Lin, X. Xu, J. Zhao, X. Wang, Dynamic service migration in ultra-dense multi-access edge computing network for high-mobility scenarios, EURASIP Journal on Wireless Communications and Networking 2020(1)(2020) 191.
[116] S. Liu, P. Cheng, Z. Chen, W. Xiang, B. Vucetic, Y. Li, Contextual user-centric task offloading for mobile edge computing in ultra-dense network, IEEE Transactions on Mobile Computing 22(9)(2023) 5092-5108.
[117] J. Wang, K. Liu, M. Ni, J. Pan, Learning based mobility management under uncertainties for mobile edge computing, in: Proc. 2018 IEEE Global Communications Conference, 2018.
[118] V. Balasubramanian, M. Aloqaily, F. Zaman, Y. Jararweh, Exploring computing at the edge: A multi-interface system architecture enabled mobile device cloud, in: Proc. 2018 IEEE 7th International Conference on Cloud Networking, 2018.
[119] Z. Chen, S. Chen, H. Xu, B. Hu, A security authentication scheme of 5G ultra-dense network based on block chain, IEEE Access 6(2018) 55372-55379.
[120] M. Polignano, P. Mogensen, P. Fotiadis, L. Chavarria, I. Viering, P. Zanier, The inter-cell interference dilemma in dense outdoor small cell deployment, in: Proc. 2014 IEEE 79th vehicular technology conference, 2014.
[121] J. Liu, M. Sheng, L. Liu, J. Li, Effect of densification on cellular network performance with bounded pathloss model,

IEEE Communications Letters 21(2)(2017) 346-349.
[122] G.V. Adekogba, K.B. Adedeji, Y.O. Olasoji, Interference Reduction Scheme for Femtocell Ultra-Dense-Network: Concept and Research Challenges, ITEGAM-JETIA 10(49)(2024) 144-158.
[123] H. Zhang, S. Huang, C. Jiang, K. Long, V.C.M. Leung, H.V. Poor, Energy efficient user association and power allocation in millimeter-wave-based ultra dense networks with energy harvesting base stations, IEEE Journal on Selected Areas in Communications 35(9)(2017) 1936-1947.
[124] W. Sun, J. Liu, 2-to-*M* coordinated multipoint-based uplink transmission in ultra-dense cellular networks, IEEE Transactions on Wireless Communications 17(12)(2018) 8342-8356.
[125] S. Chen, T. Zhao, H.H. Chen, W. Meng, Downlink coordinated multi-point transmission in ultra-dense networks with mobile edge computing, IEEE Network 33(2)(2019) 152-159.
[126] X. Kong, G. Duan, M. Hou, G. Shen, H. Wang, X. Yan, M. Collotta, Deep reinforcement learning-based energy-efficient edge computing for internet of vehicles, IEEE Transactions on Industrial Informatics 18(9)(2022) 6308-6316.
[127] N. Waqar, S.A. Hassan, A. Mahmood, K. Dev, D.T. Do, M. Gidlund, Computation offloading and resource allocation in MEC-enabled integrated aerial-terrestrial vehicular networks: A reinforcement learning approach, IEEE Transactions on Intelligent Transportation Systems 23(11)(2022) 21478-21491.
[128] Y. Liu, Y. Ni, H. Zhao, H. Zhang, H. Zhu, H. Cao, K. Yu, An efficient power allocation algorithm for green reconfigurable intelligent surface assisted vehicular network, IEEE Transactions on Intelligent Transportation Systems 23(12)(2022) 23736-23749.
[129] P. McEnroe, S. Wang, M. Liyanage, A survey on the convergence of edge computing and AI for UAVs: Opportunities and challenges, IEEE Internet of Things Journal 9(17)(2022) 15435-15459.
[130] J. Peksa, D. Mamchur, A Review on the State of the Art in Copter Drones and Flight Control Systems, Sensors 24(11)(2024) 3349.
[131] S. Pang, J. Zheng, M. Wang, S. Qiao, X. He, C. Gao, Minimize average tasks processing time in satellite mobile edge computing systems via a deep reinforcement learning method, Journal of Cloud Computing 12(1)(2023) 159.
[132] F. Chai, Q. Zhang, H. Yao, X. Xin, R. Gao, M. Guizani, Joint multi-task offloading and resource allocation for mobile edge computing systems in satellite iot, IEEE Transactions on Vehicular Technology 72(6)(2023) 7783-7795.
[133] L. Zhang, J. Chakareski, UAV-assisted edge computing and streaming for wireless virtual reality: Analysis, algorithm design, and performance guarantees, IEEE Transactions on Vehicular Technology 71(3)(2022) 3267-3275.