Nigerian Journal of Technology (NIJOTECH)
Vol. 45 No. 1, March 2026, ppp. 1- 23
www.nijotech.com

Print ISSN: 0331-8443
Electronic ISSN: 2467-8821
https://doi.org/10.4314/njt.2026.5706

A SYSTEMATIC REVIEW OF RESOURCE ALLOCATION FOR 5G NETWORKS: CHALLENGES, METHODS AND FUTURE RESEARCH DIRECTIONS

AUTHORS:
*M. M. Umar[1], A. B. Garko[1], H. U. Suru[1] and M. N. Sarki[2]

AFFILIATIONS:
[1]Department of Computer Science, Abdullahi Fodio University of Science and Technology, Aliero, Kebbi State. NIGERIA
[2]Department of Mathematics, Abdullahi Fodio University of Science and Technology, Aliero, Kebbi State. NIGERIA

***CORRESPONDING AUTHOR:**
Email: manirutambuwal@gmail.com

ARTICLE HISTORY:
Received: October 05, 2025.
Revised: February 18, 2026.
Accepted: March 17, 2026.
Published: Early view

KEYWORDS:
5G, Resource allocation, eMBB, mMTC, URLLC.

ARTICLE INCLUDES:
Peer review

DATA AVAILABILITY:
On request from author(s)

EDITORS:
Ike Mowete

FUNDING:
This work is supported by the Petroleum Technology Development Trust Fund (PTDF) through the In-Country Scholarship Scheme (ISS).

Abstract
Fifth-generation (5G) networks are designed to support high data rates, low latency, and reliable communication for different service use cases, including enhanced Mobile Broadband (eMBB), massive Machine-Type Communication (mMTC), and ultra-Reliable Low-Latency Communication (URLLC). Each of these use cases imposes different Quality of Service (QoS) requirements, making efficient resource allocation (RA) a critical challenge in the dynamic and heterogeneous 5G environment. This paper presents a systematic review of existing RA schemes, grouping them into eMBB-based, mMTC-based, URLLC-based, and multi-use-case approaches. The review evaluates the techniques used in each category, their performance, and their limitations in meeting specific service demands. Some of the challenges identified after reviewing these existing RA schemes include unfair resource distribution, inflexible allocation strategies, energy inefficiency, and high computational complexity, particularly in learning-based models. The coexistence of multiple 5G use cases remains a significant unresolved issue. To address these gaps, the paper highlights future research directions that include adaptive and hybrid RA frameworks, finer intra-use-case differentiation, intelligent resource reclamation, and lightweight learning-based solutions suitable for real-time operation. This review also provides a useful reference for improving scalable, efficient, and fair resource allocation in 5G and beyond networks.

1.0 INTRODUCTION
Wireless broadband (WiBB) technologies have evolved to meet the growing demand for quality of service. Long-Term Evolution (LTE) networks and Worldwide Interoperability of Microwave Access (WiMAX) are the most prevalent WiBB technologies, developed as fourth-generation (4G) technology [1]. These technologies offer high-speed connectivity, providing users with better multimedia experiences. It has been predicted that by 2025, there will be over 75.4 billion connected devices due to the rapid growth of internet users and connected devices [2].

HOW TO CITE:
Umar, M. M., Garko, A. B., Suru, M. N. and Sarki, M. N. "A Systematic Review of Resource Allocation for 5G Networks: Challenges, Methods and Future Research Directions", *Nigerian Journal of Technology*, 2026. 45(1), ppp. 1 - 23. https://doi.org/10.4314/njt.2026.5706

The rise in connected devices is driven by the global desire for a smarter society. This requires excellent connectivity and massive cellular traffic for various network applications. Current 4G technology cannot meet these high connectivity needs, leading to the standardization of fifth-generation (5G) technology by professionals from academia and business. This technology will meet the high connectivity needs of these devices [3].

The 5G network has been designed as a heterogeneous and highly flexible system with the capabilities of supporting diverse service categories with different Quality of Service (QoS) requirements. This was achieved with the aid of key enabling technologies, such as network densification**,** massive multiple-input multiple-output (mMIMO) antenna systems**,** and millimeter-wave (mmWave) communications [4]. These enabling technologies jointly form the physical and architectural foundation for enhanced Mobile Broadband (eMBB), massive Machine-Type Communication (mMTC), and ultra-Reliable Low-Latency Communication (URLLC) [5].

Network densification is a key strategy for meeting the exponential growth in mobile data traffic and the stringent quality-of-service (QoS) requirements of emerging 5G applications. Densification is achieved by deploying a multi-tier cellular architecture that combines traditional macro base stations with a large number of low-power small cells, such as micro, pico, and femto base stations. While macro base stations provide wide-area coverage and mobility support, small cells enhance capacity in traffic hotspots, improve indoor coverage, and reduce transmission distances between user equipment (UE) and access points. [6] By shortening link distances and increasing spatial frequency reuse, network densification significantly improves spectral efficiency and system capacity. This is particularly important for eMBB services, which demand high throughput, and for URLLC applications, where reduced transmission distance contributes to lower latency and improved reliability. In mMTC scenarios, densification enables efficient access for a massive number of devices by distributing the traffic load across multiple access points [7]

Massive MIMO is another cornerstone technology of 5G networks, in which base stations are equipped with tens or hundreds of antennas to simultaneously serve multiple users over the same time-frequency resources. By exploiting spatial multiplexing and advanced beamforming techniques, mMIMO substantially increases spectral efficiency while suppressing inter-user and inter-cell interference. For eMBB services, massive MIMO enables high peak data rates and improved cell-edge performance through spatial diversity and beamforming gains. In URLLC scenarios, the enhanced reliability and robustness provided by mMIMO are critical for meeting strict latency and reliability constraints. Furthermore, the ability of mMIMO systems to focus energy toward intended users improves energy efficiency, which is particularly beneficial for battery-constrained mMTC devices [8]

The introduction of millimetre-wave (mmWave) communication is driven by the scarcity of available spectrum in sub-6 GHz bands and the need for extremely wide bandwidths to support multi-gigabit data rates. Operating at frequencies typically above 24 GHz, mmWave bands offer abundant spectrum resources that are well-suited for high-capacity eMBB applications. However, mmWave signals suffer from high path loss, susceptibility to blockage, and limited coverage range [9]. These challenges are mitigated through the combined use of network densification and directional beamforming enabled by massive MIMO antenna arrays. As a result, mmWave communications are primarily deployed in dense urban environments and hotspot areas, where they complement sub-6 GHz bands to provide both coverage and capacity. Although mmWave is mainly associated with eMBB, its integration with advanced beam management and scheduling techniques can also support low-latency URLLC services in localized scenarios [10].

The third-generation partnership project (3GPP) standard divides traffic types into three use cases: Enhanced Mobile Broadband (eMBB), Massive Machine-Type Communication (mMTC), and Ultra-Reliable Low-Latency Communication (URLLC) [11]. These applications are sometimes referred to as 5G technology features. Figure 1 shows an example of a 5G heterogeneous environment with all three use cases.

The eMBB technology enables 5G to offer increased speed and network capacity, with peak data throughput of 10Gbps for UL and 20Gbps for DL, making it 100 times faster than a 4G network. It also boosts energy efficiency and enables up to 500 kph mobility, supporting services like workplace collaboration and VR [11]. Massive-machine-type communication (mMTC) is a data transmission method that connects numerous devices, including sensors and IoT devices, without human interaction. With 5G, it can support services like smart cities, homes, agriculture, asset tracking, remote

https://doi.org/10.4314/njt.2026.5706

monitoring, and IoT, enabling up to 1 million MTC devices/km2 [12]. The URLLC use case aims to provide low-latency and highly responsive services using 5G technology, improving dependability and lowering latency by up to 1ms compared to 4G technology, supporting various services like industrial automation, driverless cars, and remote patient monitoring [12].

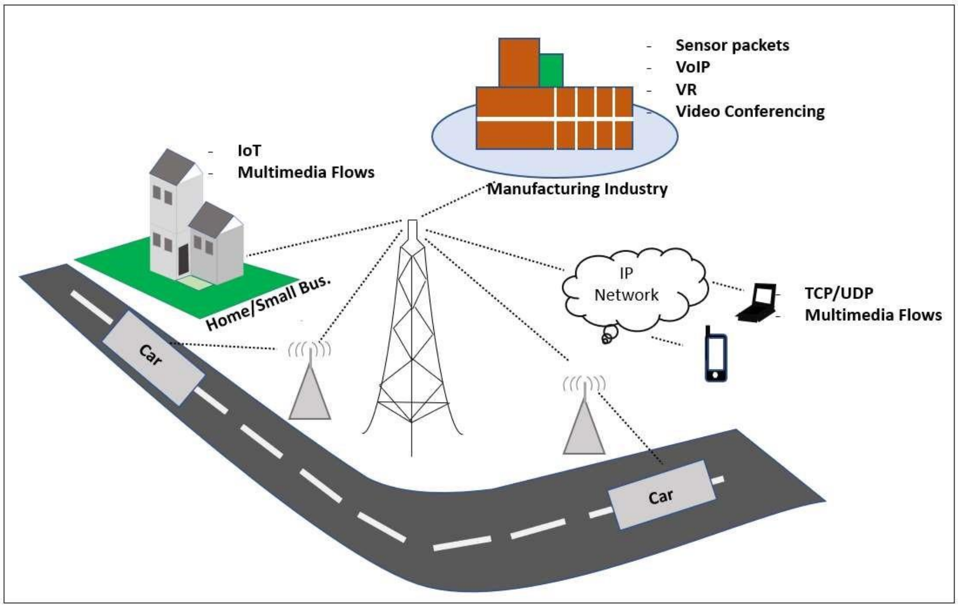

Figure 1: Example of a 5G heterogeneous network with all three use cases [11]

The fifth-generation (5G) network infrastructure integrates various wireless access technologies, network topologies, and communication devices to create a 5G heterogeneous environment. It aims to offer massive machine-type communications, ultra-reliable low-latency communications, and improved mobile broadband. This requires a network composed of macro cells, small cells, Wi-Fi access points, and D2D communications for seamless connectivity and performance. Support for a large number of linked devices with different service needs is another aspect of heterogeneity that makes network management more difficult. Maintaining QoS, reducing interference, and guaranteeing effective network performance become extremely difficult in a 5G heterogeneous environment due to the coexistence of different radio access technologies and varying service demands [13]. Advanced resource management techniques such as network slicing, dynamic spectrum sharing, and intelligent traffic management are employed to overcome obstacles in the 5G ecosystem, focusing on effective resource allocation for maximum network utilization and service-level agreements [14].

In order to provide scalable, adaptable, and high-performance communication services, 5G network resources are essential. These consist of energy resources, network resources, radio resources, and computational resources. In Internet of Things contexts, these resources are necessary for infrastructure, data processing, wireless data transmission, and maintaining network operations. For 5G use cases like eMBB, URLLC, and mMTC to meet their diverse requirements, these resources must be allocated and managed effectively [15]. In 5G networks, resource allocation (RA) is essential for effectively managing scarce network resources to satisfy the various performance needs of applications including huge machine-type communications, ultra-reliable low-latency communications, and improved mobile broadband. In diverse and dynamic situations, it guarantees efficient network utilisation, reduces latency and interference, and upholds constant Quality of Service (QoS). RA in 5G is facing numerous challenges due to the network's heterogeneous nature, diverse service requirements, and the need to support massive device connectivity with varying QoS demands. Additionally, managing interference, dynamic spectrum availability, and real-time adaptation across multiple access technologies adds complexity to efficient resource distribution [15].

Although numerous review and survey articles have examined resource allocation (RA) in 5G networks, most existing studies focus on either a specific service category (e.g., eMBB, mMTC, or URLLC), a particular enabling technology (such as massive MIMO, mmWave communications, or machine

https://doi.org/10.4314/njt.2026.5706

learning), or a single class of optimization techniques. As a result, there is limited consolidation of RA schemes across heterogeneous use cases under a unified framework that reflects the diverse performance objectives and constraints of practical 5G deployments.

In particular, the distinct and often conflicting requirements of eMBB, mMTC, and URLLC, high throughput, massive connectivity, and ultra-low latency with high reliability, respectively, lead to fundamentally different RA problem formulations. Existing surveys rarely provide a side-by-side comparison of how RA strategies are designed, evaluated, and constrained within each of these use cases, nor do they consistently examine schemes that support the coexistence of multiple service types within the same network. Furthermore, while many studies categorize RA methods based on algorithmic approaches, fewer systematically analyze the network scenarios (e.g., dense small cells, massive MIMO, mmWave, D2D, and MEC-enabled architectures) under which these schemes are proposed.

Motivated by these gaps, this review aims to provide a use-case–driven and scenario-aware synthesis of RA schemes in 5G networks. The paper systematically classifies existing works according to their targeted service category, eMBB, mMTC, URLLC, and multi-service coexistence, and summarizes their resource allocation strategies, key performance achievements, and inherent limitations. By organizing the literature in this manner and consolidating insights through comparative tables, the review enables a clearer understanding of the trade-offs involved in RA design across different 5G use cases. Additionally, by identifying recurring challenges and limitations reported in the literature, the paper establishes a foundation for discussing open research issues and future directions toward beyond-5G and emerging 6G communication systems.

2.0 METHODOLOGY

This study aimed to conduct a systematic review of resource allocation (RA) in 5G networks. The review aimed to identify and classify existing RA approaches for 5G use cases (eMBB, mMTC, URLLC, and multiple use-case scenarios), the challenges and limitations associated with these approaches, and the future research directions. The process was carried out in three primary stages: (1) planning the review, (2) conducting the review, and (3) reporting the review results. Each stage was further broken down into steps, as illustrated in Figure 2.

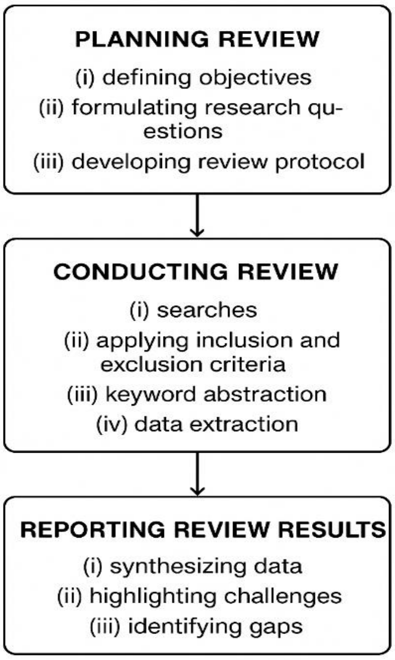

Figure 2: Systematic review process

i. Planning phase

This phase involved (i) defining the objectives of the study, (ii) formulating research questions to guide the review, and (iii) developing the review protocol, including inclusion and exclusion criteria.

ii. Conducting phase:

In this phase, (i) searches were performed in scholarly databases for studies on resource allocation in 5G networks; (ii) inclusion and exclusion criteria were applied to select relevant studies; (iii) keyword abstraction was undertaken to identify core concepts such as "resource allocation," "5G," "eMBB," "mMTC," and "URLLC"; and (iv) data extraction was carried out to collect information on schemes, performance metrics, strengths, and weaknesses.

iii. Reporting phase:

This final phase involved synthesizing and analyzing the data extracted from the selected studies to classify RA schemes by use-case category, highlight key challenges and limitations, and identify gaps to be addressed in future research.

2.1 Research Questions

The following research questions (RQs) guided the study

RQ1: What is the publication trend of resource allocation studies in 5G networks across different use

https://doi.org/10.4314/njt.2026.5706

cases (eMBB, mMTC, URLLC, and multiple-use-case scenarios)?

RQ2: What are the most common resource allocation techniques and strategies proposed for each 5G use case, and how are they classified?

RQ3: What strengths and weaknesses are associated with existing resource allocation schemes in terms of performance, scalability, fairness, and energy efficiency?

RQ4: How do resource allocation approaches address coexistence among heterogeneous traffic types and multiple use cases within 5G networks?

RQ5: What key challenges, limitations, and research gaps persist in current 5G resource allocation studies?

RQ6: What future research directions and emerging approaches are proposed to overcome the identified limitations and improve 5G resource allocation?

3.0 RELATED WORKS

For this paper, we classify RA schemes based on 5G use-cases, which are: eMBB-based RA schemes, mMTC-based RA schemes, URLLC-based RA schemes, and multiple use-case RA schemes.

3.1 EMBB-based RA Schemes

The eMBB use case will support up to 500 kph of mobility while increasing energy efficiency by 100%. Virtual reality (VR), augmented reality (AR), video monitoring, enterprise collaboration, mobile cloud computing, video calls, internet communications, live video games, and more are among the many services that the eMBB enables [16]. Some of the current eMBB-based RA schemes are discussed in this subsection.

The work in [17] developed an energy-efficient resource allocation algorithm for mobile cloud computing. The algorithm uses foglets clustering to group user requests with similar characteristics or traffic types into one groupp. The Order of Preference by Similarity to the Ideal Solution (TOPSIS) technique is used to ensure all requests are placed in appropriate cluster groups. The algorithm reduces the average queuing delay and increases mobile devices' battery life. However, it reduces throughput when clusters with fewer requests require more resources. In [18], the researchers developed a multi-user algorithm for mobile cloud systems to increase throughput, reduce battery power consumption, and delay devices. The algorithm classifies requests based on traffic type, device energy level, bandwidth needed, and delay budget. Devices with the lowest energy level are served first, followed by those with minimal delay budget. However, the algorithm considers bandwidth requirement as the last criterion for resource allocation, ensuring high-quality service for some requests. A Quality of Experience (QoE)-driven resource allocation scheme for live video streaming in 5G networks was proposed in [19]. The scheme uses priority-based video transmission, flexible communication mode switching for UE's, and subset relay assignment. It considers higher bandwidth requests and allows flexible handover of UE's when the network status is low or when UE's exit coverage. The scheme improves QoS and QoE for users, but may waste reserved resources when no QoE-driven requests exist. The work in [20] developed a traffic type-based RA algorithm to enhance network resource utilization and throughput for eMBB-based traffic in a 5G network. The algorithm categorizes user requests based on traffic types and assigns each request to its type queue. It calculates a priority weight based on each request's QoS requirements, assigning higher weights to resources first. However, this may cause packet loss when requests with higher QoS demand arrive frequently. Authors of [21] presented a latency-away dynamic RA algorithm to improve the Quality of Service (QoS) of eMBB traffic in a 5G network. The algorithm uses a genetic algorithm to compute queue weight based on latency and channel quality. It swaps queue requests, putting the lowest latency request at the top and allocating resources to good channel quality requests. This improves QoS, but delays and throughput decrease for high latency budget requests and poor channel conditions. According to researchers in [22], a QoS-based optimization and RA scheme to improve the throughput of eMBB service in 5G satellite networks. The scheme uses the shortest path strategy to determine which packets will be served first. It scans through all packets in different queues and identifies the shortest path to the gNB. It then serves packets with the highest QoS requirements first. Simulation results show that this reduces latency for shortest paths but increases delay for longer paths. The work in [23] proposed a RA scheme to reduce packet transmission delay in VR systems. They created two queues for user requests, serving 75% in the excellent queue and 25% in the bad queue. Simulation results showed the scheme reduces transmission delay and increases packet throughput with good channel conditions. However, overloaded networks and unsuccessful attempts to access resources increase battery power consumption. In order to improve QoE and resource utilisation while lowering battery power consumption in VR devices,

https://doi.org/10.4314/njt.2026.5706

[24] suggested a dynamic resource allocation strategy. Prioritising user requests according to latency, migration time, and resource utilisation rate is done by the scheme using a QoE-driven method. Delay-intolerant queries can be prioritised thanks to this sequential order. However, the plan might not satisfy the high-resource user requests' QoE and QoS criteria, which could result in packet drop and additional delay. To improve the Quality of Experience (QoE) and lower power consumption in VR systems, researchers in [25] created a resolution control and RA algorithm. They employed an adaptive resolution control technique to modify resolution in response to network circumstances and resource availability, as well as the Outer Approximation mechanism to distribute resources to users with greater priority. The algorithm's power consumption performance is identical to that of the benchmark algorithm. A delay-aware resource allocation technique was introduced in [26] to enhance the Quality of Service (QoS) of RT services in eMBB 5G networks. The method works in two stages: a channel-aware resource allocation plan and a delay-based assessment mechanism. While the second phase gives priority to traffic flows with good channel quality, the first step distributes resources among various traffic flows. However, because of their lower priority, NRT traffic flows suffer from greater delays and decreased throughput. In [27], a dynamic scheduling algorithm using Deep Reinforcement Learning to improve QoS and enable efficient routing for eMBB slices in 5G networks. The algorithm selects optimal downlink scheduling rules based on real-time network states like packet delay, fairness, throughput, and packet loss rate. The Deep Q-Network algorithm is used to approximate Q-values for state-action pairs. The system uses a custom reward function to prioritize high throughput and fairness while minimizing delay and packet loss. Simulations show the algorithm significantly reduces packet loss ratio and head-of-line

3.2 MMTC-based RA Schemes

Massive machine type communication (mMTC) is expected to play a key role in 5G networks. It is known as a technique that makes it possible for several low-complexity, energy-constrained MTC devices to send packets fast and with little latency. This procedure involves little to no human intervention [28]. Among the services offered by the mMTC use case for industrial and smart societies are asset tracking, remote monitoring, smart cities, smart homes, smart agriculture, and the Internet of Things (IoT). This subsection reviews a few of the existing mMTC-based RA techniques.

In [29], a data aggregation scheme for mobile traffic control (MTC) devices was proposed to enhance efficiency and reduce congestion. The scheme operates in two phases: aggregation and relaying. MTC devices transmit data to aggregators, while aggregated data is sent to the base station. Random resource scheduling and channel-aware resource scheduling algorithms are used. The scheme improves system performance but reduces resource utilization. The work in [30] proposed a hybrid OMA-NOMA data aggregation mechanism, improving the work in [31]. The scheme allows multiple MTC devices to share the same orthogonal channel while transmitting data, allocating resources based on success probability and the average number of simultaneously served devices. This improves system performance by serving more devices and increasing network resource utilization. However, it increases the delay of distant devices and increases power consumption. Researchers in [32] introduced a traffic-aware RA scheme to increase successful communications and reduce power consumption of MTC devices. They classified incoming traffic into two groups: traffic-aware and traffic-unaware, created queues for each, and used a stochastic geometry method to calculate spectral efficiency. The sub-gradient descent method allocated resources to devices, serving the queue with the highest number of requests first. However, traffic-unaware experienced significant delay, leading to some requests dropping. In [33], a cooperative data aggregation and dynamic resource allocation scheme for mMTC traffic was presented to improve the Quality of Service (QoS) of MTC devices. The scheme uses a fixed data aggregator and multiple data aggregator to address varying QoS requirements and prioritizes transmission requests from delay-intolerant devices. However, delay-tolerant devices will incur more delay and packet drop, reducing throughput and reducing the number of succeeded communications. The work in [34] proposed a dynamic RA scheme for mMTC in 5G networks to increase throughput and reduce energy consumption.

The scheme uses Sparse Code Multiple Access (SCMA) multiplexing technique and dynamic load-aware Physical Random-Access Channel (PRACH) and Physical Uplink Shared Channel (PUSCH) for RA. This increases throughput and reduces energy consumption for MTC devices. However, the scheme wastes resources by assuming frequent traffic arrival in specific directions. [35] presented a joint congestion control and RA algorithm to enhance resource utilization and transmission efficiency for mMTC in 5G networks. The algorithm uses a dynamic resource allocation technique and random-

https://doi.org/10.4314/njt.2026.5706

access procedure based on adaptive access class bearing (ACB). It increases successful transmissions and lowers energy consumption for MTC traffic. However, the QoS of some MTC traffic may not be guaranteed when the network is overloaded, leading to unsuccessful trials. The work in [36] proposed an improved RA algorithm for 5G MTC networks to improve traffic quality of service and resource utilization efficiency. The algorithm works in two stages: medium access and resource allocation. Medium access uses wireless signals and wait time for MTC devices, while resource allocation considers received signal strength, total induced transmission delay, and transmission-waiting time. The scheme distributes resources equally, increasing priority for devices with higher priority but increasing battery consumption due to frequent retrying. In [37], a contention-based RA scheme to enhance transmission efficiency and decrease battery consumption in mMTC networks. The scheme uses Spare Code Multiple Access (SCMA) techniques to allocate resources among contending devices, allowing multiple MTC devices to transmit data on the same resource element. It operates in two modes: collision-free and contention-based. While the scheme increases transmission efficiency, resource utilization decreases, especially when no collision occurs, and energy consumption increases in congested networks. In [38], a joint access control and resource allocation (RA) algorithm was proposed to enhance resource utilization and decrease battery power consumption for mMTC traffic. The algorithm checks MTC devices' priority, channel availability, access level, and data transmission time before serving or waiting for transmission. This increases network resource utilization but causes congestion at gNodeB and increases battery power consumption due to the process of access control and resource allocation. The work of [39] presented a dynamic slot and RA scheme to reduce congestion at gNodeB and increase resource utilization for MTC devices. The scheme considers access level, network resource availability, and priority level of MTC requests. In network-congested situations, higher-priority devices are treated without considering access level. This improves network resource utilization and transmission efficiency, but lower priority traffic experiences more delay, exceeding their delay budget. Authors in [40] developed an advanced reinforcement learning algorithm for UAV-assisted 5G mMTC slicing networks in emergency scenarios. They used UAVs as flying base stations to maintain connectivity and QoS. The challenge was to maximize energy efficiency while minimizing transmission power. They introduced the Dueling Deep Q-Network (DDQN) algorithm, which separated state-value and advantage functions for more accurate policy evaluation. Simulations showed DDQN significantly outperformed Q-Learning, DQN, and random allocation strategies, achieving higher energy efficiency, throughput, and lower latency. The researchers in [41] proposed a dynamic beam splitting and merging algorithm to enhance random-access performance in 5G network slices. The method reduces collisions caused by multiple devices accessing limited preamble resources.

The system uses directional beams in massive MIMO systems to split overloaded beams into narrower ones or merge underloaded ones. The system adjusts beam numbers and widths based on real-time random-access intensity, ensuring Quality of Service (QoS) and minimizing signaling overhead. Simulation results show significant improvements in random-access throughput and success probability. The work in [42] presented a joint task replication and resource allocation framework to improve service reliability and reduce latency in 5G mMTC networks. The algorithm targets Mobile Edge Computing networks, which face congestion due to simultaneous task offloading by IoT devices. Task replication ensures completion even if some nodes fail. The authors used the Lyapunov optimization framework to stabilize task queues, manage massive user interactions efficiently, and handle reliability and delay constraints. Simulation results showed significant improvements in service reliability, latency reduction, and overall operational cost

3.3 URLLC-based RA Schemes

Offering services that will enable applications with exceptionally low latency and excellent responsiveness is the aim of the URLLC use-case. Compared to 4G technology, 5G technology is expected to increase reliability while lowering latency by up to 1ms. Services including remote patient monitoring, remote site monitoring, autonomous autos, and industrial automation, to name a few. A couple of the existing Ratechniques for this use case are listed in this subsection.

In [43] a periodic RA scheme was proposed to improve data transmission reliability for URLLC traffic in 5G networks. The scheme allocates resources for packets for the first transmission, reserving bandwidth for retransmissions if not successfully transmitted. The Hybrid Automatic Repeat Request (HARQ) scheme uses a chase combination between packets to retransmit untransmitted packets. This reduces latency and increases reliability, but it wastes resources reserved for retransmission, especially when packets are

https://doi.org/10.4314/njt.2026.5706

delivered at the first transmission. Additionally, the scheme only considers one retransmission, increasing packet drops. The authors of [44] developed a Retransmission Algorithm (RA) to reduce packet drop rate and battery power consumption for mission-critical devices in 5G networks. The algorithm considers two devices, an actuator and a robot, and uses Orthogonal Multiple Access (OMA) techniques to allow multiple packet transmissions on the same frequency channel. If a packet is not successfully transmitted, the algorithm uses available resources to retransmit it until it is delivered or when network resources are exhausted. However, resource utilization should be increased to ensure successful packet delivery. In [45] a Resource RA algorithm was presented to enhance resource utilization and throughput in 5G URLLC use cases. The algorithm considers three parameters: packet error probability, channel condition, and delay requirement. It checks the channel condition of each queued packet, separating good from bad ones. It also considers the maximum delay for each packet, considering the system's maximum delay budget. The algorithm improves network throughput but does not guarantee message reliability, necessitating a strategy to check message reliability. Authors of [46] proposed a RA scheme to enhance message reliability and QoS in a mission-critical URLLC system. The scheme uses a block coordinate descent (BCD) strategy to check bandwidth availability before transmission and successive convex approximation (SCA) techniques to ensure successful message reception. Simulation results show that the scheme improves device reliability in normal network scenarios. However, failure to transmit messages increases battery power consumption, delay, and packet dropp. The work in [47] improved vehicle-to-everything (V2X) communications by creating a Reliability-Assisted (RA) algorithm. The technique prioritises packets with favourable channel circumstances, high QoS requirements, low delay budget, and low energy level using a bipartite graph mechanism and a non-cooperative game. By giving priority to packets that meet these criteria, more packets are sent. However, because of their lower priority, high delay budget packets have higher latency and do not guarantee QoS. The method seeks to increase V2X networks' power consumption and dependability.

A resource allocation plan for multiple users was put up by researchers in [48] in order to enhance QoS for URLLC systems in 5G networks. Three QoS criteria are considered by the scheme: channel quality, maximum transmission time, maximum packet error probability, and delay requirement. It considers each user's channel quality information and gives priority to packets with low delay budget and error probability. When no frequent traffic type arrives, the strategy wastes the resources that were reserved for each type of traffic. In order to maximise resource utilisation and guarantee user quality of service in a Coordinated Multipoint (CoMP)-enabled URLLC system, the work in [49] introduced a RA algorithm. The algorithm groups packets into a cluster with greater QoS needs based on channel information and each packet's QoS requirements. It breaks up pre-coded data into its component parts, using a zero-forcing linear precoding technique, and verifies the delivery status of every packet that is sent. By guaranteeing that no packet is served when there is no resource available, the method enhances resource utilisation. Congested networks, on the other hand, serve fewer packets from the cluster with lower QoS requirements, which lowers the cluster's throughput. A resource allocation technique for multi-user single-input, single-output (SISO) URLLC systems in 5G networks was created by the authors in [50]. The algorithm rearranges all queued packets, using a Gaussian-based channel selection and allocation process, and chooses packets with the shortest distance to the gNB using a path-following technique.

According to simulation results, the algorithm serves more packets, increasing throughput, but it also causes packets that are far from the gNB to wait more. A hybrid resource allocation strategy was put up in [51] to increase the dependability of packet delivery for URLLC traffic in 5G systems. To cut down on latency and boost throughput, the technique makes use of a queuing module, coding mechanism, and packet-duplication mechanism. Although simulation results indicate enhanced throughput and improved dependability, sending a single packet across many channels requires more resources. In order to enhance resource utilisation and quality of service for vehicular users in URLLC traffic, the work in [52] offers a fair resource allocation mechanism. Physical resource blocks (PRBs) are assigned to the radio access network (RAN) portion of the system, while edge computing resources are assigned to the edge portion. In order to improve QoS and increase fairness, the scheme gives preference to users with low delay budgets and high channel quality. However, when computing resources or PRBs are unavailable, the technique increases latency and lacks a retransmission strategy for retransmitted packets. For Vehicular Edge Computing systems, a URLLC-aware resource allocation method was created in [53] with the goal of minimizing system utility while meeting URLLC limits for a variety of workloads, including 3D

https://doi.org/10.4314/njt.2026.5706

gaming and AR/VR. They learnt the best communication and computation resource allocation policies using the Soft Actor-Critic algorithm and a Lyapunov-guided Deep Reinforcement Learning framework. But the approach necessitates real-time feedback and a high level of computational complexity.

3.4 Multiple Use Case – based RA Schemes

Because of the heterogeneous nature of the 5G network, it is expected that in some scenarios, different use cases will coexist in the same user environment. In these circumstances, an efficient RA strategy that can handle the several use cases and appropriately address each one is needed. This section looks at many resource allocation algorithms that consider multiple use-cases to show the different use cases that are taken into consideration, how they operate, and the benefits and drawbacks of each scheme.

In [54], a side channel attack-aware RA scheme was proposed to reduce blocking rates and improve packet resource utilization in a 5G-sliced network. The scheme operates in three phases: allocation of Radio Resource (Active antenna unit) in the first phase, allocation of Virtual Distributed Units (vDU) to successful requests, and routing and transport resource assignment in the third phase. The scheme reduced the blocking rate of new requests and improved resource utilization, but transmission reliability is not guaranteed due to the lack of a transmission reliability strategy. Authors in [55] proposed a multi-objective resource allocation scheme to optimize the throughput and transmission reliability of eMBB and URLLC-based traffic in 5G networks. The scheme creates two waiting queues, one for eMBB traffic and one for URLLC traffic. If traffic intensity exceeds a threshold, URLLC traffic is given higher priority, serving before eMBB traffic. However, normal traffic is allocated resources to both queues, considering QoS needs and delay budget. Packets with low delay budgets are served before exceeding their budget. The simulation results show that the scheme improves transmission reliability but increases eMBB traffic delay, especially when the network is congested. Researchers in [56] developed a slicing-based resource allocation algorithm to improve packet throughput and reduce end-to-end delay. The algorithm uses adaptive modulation coding to dynamically multiplex eMBB and URLLC services, creating separate queues for each. It estimates the achievable data rate for each packet and checks the Signal to Noise Ratio (SNR) of all queued packets. The algorithm improves throughput and transmission reliability with good channel conditions. In [57], a slicing resource allocation scheme to enhance spectral efficiency and QoS for eMBB and URLLC traffic in 5G networks. The scheme divides available network resources into two, with eMBB traffic focusing on throughput and URLLC traffic on delay budget. The delay-aware mechanism reduces battery power consumption for URLLC devices. Simulation results show the scheme improves spectral efficiency and increases device battery life. However, it wastes resources when frequent arrival of only one traffic type and resources are already reserved for another. The work of [58] presents an efficient resource allocation scheme for 5G heterogeneous networks. The scheme prioritizes URLLC traffic and allocates resources to queued packets, swapping priority after the first transmission. Dropped URLLC traffic is prioritized using the HARQ strategy for improved transmission reliability. eMBB traffic is scheduled at the time slot boundary using a Proportional Fair (PF) resource allocation strategy. This scheme improves resource utilization and ensures transmission reliability, but may increase packet delay due to too many dropped packets being retransmitted. In [59], a dynamic resource allocation scheme was proposed to improve the throughput of 5G traffic types. The scheme uses a forward induction process to select the most suitable UE, considering QoS need and payload. It allocates the minimum required resources to each UE to avoid starvation. The scheme also considers data rates and channel quality, calculating the Channel Quality Index (CQI) of each UE. Simulation results show that the scheme reduces delay and increases overall system throughput. However, resources are wasted when no available resources are available to serve packets already allocated their minimum resource requirements. Researchers in [60] proposed a priority-based resource allocation scheme to improve resource utilization and QoS in network-sliced environments. The scheme considers slice information, such as priority level and demand profile, and reserves resources for the slice with higher priority and demand profile. It uses a constrained Markov Decision Process (CMDP) strategy to select packets with the highest priority. The simulation experiment shows that the scheme increases throughput for higher priority packets, but wastes resources when a specific slice has minimal or no frequent packet arrival. In [61], a dynamic multiplexing resource allocation scheme was presented for eMBB and URLLC-based traffic to improve QoS and transmission reliability. The scheme operates in two phases: first, using a greedy allocation strategy to prioritize eMBB users, and second, allocating mini-

https://doi.org/10.4314/njt.2026.5706

slots to URLLC requests based on transmission power. This improves eMBB traffic throughput and reduces battery power consumption in URLLC systems. However, the system does not guarantee QoS or throughput for URLLC-based systems, and eMBB traffic is given higher priority in resource allocation.

The researchers in [62] proposed a QoS-Aware optimal resource allocation scheme for Machine-to-machine and human-to-machine communication. The scheme classifies incoming traffic into M2M or H2H, assigns minimum bandwidth requirements, and categorizes traffic types into real-time and unreal-time streams. This increases the QoS of real-time traffic and radio resource efficiency, but increases the delay of non-real-time traffic when the network is congested. This increases the packet drop rate and degradation in performance of these traffic types. In order to enhance system performance and lower battery power consumption, the work of [63] suggested a dynamic resource allocation mechanism for eMBB and URLLC services in 5G networks. To modify power control parameters for UEs, the approach makes use of a resource-based power control (ROPC) mechanism in conjunction with a dynamic pattern cancellation indication method (DPCI). Based on current channel conditions, a dynamic selection mechanism (DSM) chooses the optimal mechanism between DPCI and ROPC.

According to simulation results, the technique increases throughput and saves power for UEs running URLLC; nevertheless, resources are squandered when they are assigned to use cases that do not have any packets in their waiting queue. Additionally, by increasing delay-intolerant traffic, the strategy lowers system throughput and causes packet drops. In order to address the cohabitation of eMBB and URLLC services in 5G networks, an intelligent resource management system was introduced in [64]. The framework optimises resource block allocation and scheduling by semi-supervised learning and Deep Reinforcement Learning, while prioritising URLLC packets through puncturing. Additionally, it integrates a Double Deep Q-Network for URLLC and a Co-training DRL algorithm for eMBB traffic. High eMBB sum rate and dependability under varied traffic loads are demonstrated by simulations. In order to optimise resource allocation for Vehicle-to-Everything (V2X) and eMBB services in 5G networks, the study of [65] suggested a real-time radio access network (RAN) slicing technique based on reinforcement learning (RL). The strategy determines the best downlink resource block (PRB) allocation using a low-complexity heuristic approach and a Q-learning algorithm. Through online network interaction, the system learns the best slicing ratios and dynamically modifies them in response to service utility and traffic demands. In comparison to conventional fixed-ratio slicing, simulation results demonstrate enhanced PRB utilisation and decreased outage likelihood. The work in [66] presents a flexible resource allocation framework for enhanced eMBB and mMTC services in dynamic Low Earth Orbit satellite networks. The approach addresses network slicing challenges due to LEO's time-varying topology and heterogeneous service requirements. The scheme introduced a sub-slot concept for modeling mMTC traffic, allowing finer-grained allocation of satellite link resources. A Mixed Integer Linear Programming model was developed to jointly allocate and route resources, resulting in improved performance over benchmark schemes. In [67], a hybrid Non-Orthogonal Multiple Access (NOMA)-based resource allocation method was presented for 5G networks to support the coexistence of eMBB and uRLLC.

The method addresses limitations in traditional OMA and SC-NOMA, such as poor spectrum efficiency and decoding complexity. The scheme uses two user-pairing techniques: near-far/far-near (NF-FN) and near-near/far-far (NN-FF) to minimize interference and complexity. Simulation results show that hybrid NOMA outperforms OMA and SC-NOMA in spectral efficiency, throughput, and latency, offering a balanced trade-off. [68] developed a Continuous-Time Markov Chain (CTMC)-based RA model to support the coexistence of eMBB and URLLC services in 5G MEC-NFV systems. The model considers factors like container setup time, failures, and repair events, and prioritizes URLLC traffic. Simulations validated the model across various load, reliability, and configuration scenarios. Results showed that increasing container number and improving setup rates improved availability and reduced latency, but increased energy consumption, especially under heavier loads. The model assumes homogeneous virtual environments and may need adaptation for real-world heterogeneous MEC deployments.

3.5 Summary of Related Reviews

This sub-section presents a summary of reviewed existing RA schemes by presenting the use case covered, network scenario, resource allocation strategy employed, key findings, and limitations of each scheme.

https://doi.org/10.4314/njt.2026.5706

Table 1: Summary of reviewed literatures

Ref	Use case covered	Network Scenario	Resource Allocation Strategy	Key Findings	Limitation
[17]	eMBB	Mobile cloud / fog-enabled 5G	Foglets clustering + TOPSIS (MCDM heuristic)	Reduces queuing delay and improves battery life	Throughput drops when small clusters demand more resources
[18]	eMBB	Mobile cloud system	Rule-based prioritization (energy + delay aware)	Improves throughput and device battery performance	Bandwidth requirement treated as the last criterion
[19]	eMBB	D2D underlaid 5G	QoE-driven priority scheduling + flexible handover	Enhances QoE/QoS for live streaming users	Reserved resources are wasted when QoE traffic is absent
[20]	eMBB	Heterogeneous 5G	Traffic-type weighted queue scheduling	Improves throughput and utilization	Low-QoS traffic suffers from delay and possible packet loss
[21]	eMBB	Multi-tier slicing network	Genetic algorithm-based queue weighting	Improves QoS for strong-channel, low-latency users	Weak-channel and high-latency users degraded
[22]	eMBB	5G-satellite integrated network	Shortest-path routing-based RA	Reduces delay for the shortest path/high-QoS packets	Long-path traffic experiences increased delay
[23]	eMBB	VR over 5G	Queue splitting (75% excellent / 25% bad)	Reduces transmission delay and increases throughput	Increased battery drain under congestion and retries
[24]	eMBB	VR microservice distribution	QoE-based dynamic prioritization	Improves QoE and reduces battery use	High-resource requests risk packet drops and delay
[25]	eMBB	Metaverse / mobile AR	Resolution control + optimization	Maintains power efficiency and throughput	Low-priority users experience QoE degradation
[26]	eMBB	Heterogeneous RAN	Hybrid: delay-aware + channel-aware RA	Improves RT QoS performance	NRT traffic suffers higher delay and reduced throughput
[27]	eMBB	RAN slicing	DRL scheduling (Deep Q-Network)	Reduces delay and packet loss; adapts to real-time states	High complexity and fairness degradation
[29]	mMTC	Aggregation-based cellular	Two-phase aggregation + random/channel-aware scheduling	Reduces congestion and improves efficiency	Low resource utilization due to early aggregation
[30]	mMTC	Hybrid OMA–NOMA uplink	OMA-NOMA aggregation allocation	Improves capacity and resource utilization	Increases delay and power use for distant devices
[31]	mMTC	Stochastic geometry-based MTC	Traffic-aware stochastic optimization	Increases success probability and saves energy	Traffic-unaware group delayed and may drop packets
[32]	mMTC	Dynamic TDD system	Traffic classification + sub-gradient RA	Improves spectral efficiency	Low-priority traffic experiences delay and drops
[33]	mMTC	Cooperative aggregation	Fixed + multi-aggregator prioritization	Improves QoS for delay-intolerant devices	Delay-tolerant devices degraded (delay/throughput loss)
[34]	mMTC	SCMA-based uplink 5G	Joint congestion control + dynamic RA	Improves throughput and energy efficiency	QoS is not guaranteed under overload
[35]	mMTC	5G NR contention-based access	SCMA + dual-mode contention RA	Improves efficiency and supports more devices	Higher energy consumption under congestion
[36]	mMTC	Metric-based access + allocation	Scheduling using SNR + delay + waiting time	Improves utilization and QoS	Retries increase battery drain

https://doi.org/10.4314/njt.2026.5706

Table 1 contd.

Ref	Use case covered	Network Scenario	Resource Allocation Strategy	Key Findings	Limitation
[37]	mMTC	Short-packet status updates	Joint access control + RA	Improves resource utilization	Causes gNB congestion and energy overhead
[38]	mMTC	Heterogeneous traffic grant-free	Priority-enabled dynamic slot allocation	Improves transmission success and efficiency	Low-priority traffic delayed/dropped
[39]	mMTC	UAV-assisted slicing	DDQN-based learning	Improves energy efficiency and throughput	High training complexity; UAV energy overhead
[40]	mMTC	Massive MIMO slicing	Beam splitting/merging	Improves random access throughput and success probability	Requires accurate traffic estimation; overlapping beams
[41]	mMTC	MEC-enabled mMTC	Task replication + Lyapunov optimization	Improves reliability and reduces latency	High system complexity; real-time feedback dependency
[42]	URLLC	Mission-critical IoT	Periodic RA + HARQ retransmission reservation	Improves reliability and reduces latency	Resource wastage; limited retransmission count
[43]	URLLC	Mission-critical devices	Retransmission-based RA	Reduces packet drops and supports delivery	Increased power consumption and resource inefficiency
[44]	URLLC	Multi-user MISO OFDMA	Channel + delay-based prioritization	Improves throughput and utilization	Reliability is not explicitly guaranteed
[45]	URLLC	Mission-critical URLLC	BCD + SCA optimization	Enhances reliability and QoS	Delay and battery use increase when retransmissions occur
[46]	URLLC	V2X URLLC	Graph/game-based prioritization	Reduces delay and power; improves delivery	High-delay packets are penalized and may violate QoS
[47]	URLLC	Multi-user URLLC	QoS-aware reservation scheduling	Reduces latency and improves QoS	Reserved resources are wasted when traffic absent
[48]	URLLC	CoMP-enabled URLLC	ZF precoding + clustering heuristic	Improves utilization and QoS	Lower-QoS clusters lose throughput under congestion
[49]	URLLC	Multi-user SISO	Queue reordering + Gaussian selection	Improves throughput and packets served	Distant packets delayed
[50]	URLLC	Short block length redundancy	Queuing + coding + packet duplication	Improves reliability and throughput	High resource consumption due to duplication
[51]	URLLC	Vehicular URLLC + edge	Joint RAN PRB + edge compute allocation	Improves fairness and QoS	No delivery guarantee; higher delay when PRBs scarce
[52]	URLLC	Vehicular edge computing	Lyapunov-guided DRL (LySAC)	Low latency and high reliability	High training complexity; requires real-time feedback
[53]	eMBB & URLLC	5G RAN slicing	Multi-stage allocation (AAU + vDU + routing)	Reduces blocking and improves utilization	No explicit transmission reliability strategy
[54]	eMBB & URLLC	Joint coexistence scheduling	Multi-objective priority switching	Improves reliability and URLLC performance	eMBB delay increases under congestion
[55]	eMBB & URLLC	RAN slicing multiplexing	AMC-based multiplexing + SNR filtering	Improves throughput and reliability	Poor-channel users degraded; delay increases

https://doi.org/10.4314/njt.2026.5706

Table 1 contd.

Ref	Use case covered	Network Scenario	Resource Allocation Strategy	Key Findings	Limitation
[56]	eMBB & URLLC	Slicing-based partitioning	Resource partition + delay-aware mechanism	Improves spectral efficiency and battery life	Resource wastage when one traffic dominates
[57]	eMBB & URLLC	Coexistence scheduler	URLLC priority + HARQ + PF scheduling	Improves reliability and utilization	Retransmissions increase the delay for other traffic
[58]	eMBB & URLLC	Dynamic channel-aware RA	Forward induction + CQI-based selection	Reduces delay and improves throughput	Wastes resources when minimum allocations are unused
[59]	eMBB & URLLC	Network slicing	CMDP-based allocation	Improves throughput for high-priority packets	Resources wasted for inactive slices
[60]	eMBB & URLLC	Multiplexing coexistence	Greedy + mini-slot allocation	Improves eMBB throughput and URLLC power use	URLLC QoS not guaranteed; eMBB overly prioritized
[61]	mMTC & URLLC	M2M/H2H coexistence	QoS-aware "optimal" RA	Improves real-time QoS and utilization	Un-real-time traffic is delayed/dropped under congestion
[62]	eMBB & URLLC	Adaptive control coexistence	ROPC + DPCI adaptive switching	Improves throughput and saves power	Delay and packet drops increase for delay-sensitive traffic
[63]	eMBB & URLLC	Intelligent RAN slicing	Semi-supervised learning + DRL	Achieves balanced URLLC and eMBB performance	High computational cost and training complexity
[64]	eMBB & URLLC	V2X slicing	RL-based slicing (SoftMax policy)	Improves PRB utilization and reduces outage	Requires accurate state estimation; training overhead
[65]	eMBB & mMTC	LEO satellite slicing	MILP + sub-slot slicing	Improves resource reuse and routing	Complex and traffic-dependent
[66]	eMBB & URLLC	Hybrid NOMA slicing	Hybrid NOMA pairing (NF-FN, NN-FF)	Improves spectral efficiency and fairness	High system complexity and sensitivity to user distribution

3.6 Novelty and Contributions

The novelty of this review lies in its unified and systematic synthesis of 5G RA approaches across major service categories and network scenarios, together with an explicit analysis of limitations and future research opportunities. Specifically, the main contributions are summarized as follows:

i. Unified cross–use-case perspective

The review organizes and compares RA schemes across eMBB, mMTC, URLLC, and multi-service coexistence settings, highlighting how differing service requirements lead to distinct optimization objectives and constraints.

ii. Scenario-aware classification of RA schemes

The review explicitly maps each RA approach to the assumed network scenario (e.g., dense small cells, mmWave, massive MIMO, D2D, MEC, and their combinations), enabling a clearer understanding of applicability and design trade-offs.

iii. Method-centric synthesis with limitations

Beyond summarizing algorithms, the review consolidates key limitations, such as computational complexity, scalability, signalling overhead, robustness to imperfect CSI, energy constraints, and fairness, thereby emphasizing deployability rather than only theoretical performance.

iv. Structured future directions toward beyond 5G/6G

The review identifies promising research directions, including AI-native RA, joint communication-and-sensing-assisted scheduling, learning under uncertainty, and cross-layer RA for extreme reliability and latency, thereby connecting current 5G RA challenges to emerging 6G requirements

https://doi.org/10.4314/njt.2026.5706

4.0 RESULTS AND DISCUSSIONS

This section presents the findings of the systematic review and discusses how the research questions (RQs) formulated in Section 3.0 have been addressed. The extracted studies were classified into four categories: eMBB-based, mMTC-based, URLLC-based, and multiple use-case RA schemes, and evaluated according to their techniques, strengths, and weaknesses.

Table 2: Summary of how research questions (RQs) were addressed

Research Question (RQ)	**Findings / How the RQ Was Answered**
RQ1: Publication trends across 5G RA studies	Publications on resource allocation in 5G networks show an upward trajectory from 2017 to 2025, reflecting global research interest (Asia, Europe, North America, Africa). This indicates the growing maturity and importance of RA research in 5G use cases.
RQ2: Resource allocation techniques and strategies	Identified a wide range of techniques: heuristic and priority-based (shortest path, queuing), optimization-based (genetic algorithms, MILP), learning-based (DRL, DQN, Lyapunov-guided RL), and hybrid schemes (NOMA + slicing). Each target specific performance metrics such as latency, throughput, fairness, or energy efficiency.
RQ3: Strengths and weaknesses of existing schemes	Tables 1–4 detail the strengths (improved throughput, latency, QoE) and weaknesses (resource wastage, high complexity, fairness issues) across eMBB, mMTC, URLLC, and multiple use-case RA schemes. Trade-offs remain a central challenge.
RQ4: Addressing coexistence among heterogeneous traffic types	Multiple-use-case RA schemes attempt joint handling of eMBB, mMTC, and URLLC via slicing, hybrid NOMA, DRL-based joint optimization, and priority mechanisms. These improve fairness and throughput but still risk uneven QoS under congestion or imbalanced traffic.
RQ5: Key challenges, limitations, and research gaps	Persistent issues identified include static RA policies, unfair resource distribution, inefficient energy usage, limited intra-use-case differentiation, high computational complexity, and a lack of cross-domain coordination. These reflect gaps in current RA models.
RQ6: Future research directions and emerging approaches	Suggested hybrid/adaptive frameworks, fine-grained QoS differentiation, intelligent resource reclamation, lightweight/federated learning models, and cross-layer integration. These directly address the gaps in RQs 3–5 and form a roadmap for next-generation RA research.

RQ1: Publication trends across 5G resource allocation studies

The review revealed an increasing trend of publications between 2017 and 2025, with a marked acceleration from 2020 onward, reflecting the maturity of 5G deployments. Studies originated from multiple regions, including Asia, Europe, North America, and Africa, showing a globally distributed research interest. This confirms the growing importance of RA research across different use-case categories.

RQ2: Resource allocation techniques and strategies across all categories, diverse RA methods were identified

Heuristic-based (e.g., shortest path, priority queuing), optimization-based (e.g., genetic algorithms, mixed integer programming), learning-based (e.g., DRL, DQN, Lyapunov-guided RL), and hybrid mechanisms (e.g., NOMA + slicing). Each approach targets different performance metrics, latency reduction, throughput maximization, energy savings, or fairness improvement, showing that RA research has diversified to meet heterogeneous 5G requirements.

RQ3: Strengths and weaknesses of existing schemes

Table 1 shows that while each RA scheme improves at least one performance metric, trade-offs persist. For example, QoE-driven schemes increase user experience but risk underutilizing reserved resources; DRL-based approaches adapt in real-time but have high computational cost; and hybrid approaches improve coexistence but introduce complexity. These patterns clearly identify where existing RA strategies succeed and where they fall short.

RQ4: Addressing coexistence among heterogeneous traffic types

The review found multiple-use-case RA schemes attempting to jointly handle eMBB, mMTC, and URLLC traffic. Techniques like slicing, hybrid

https://doi.org/10.4314/njt.2026.5706

NOMA, DRL-based joint optimization, and priority-based mechanisms were prevalent. Although these schemes improve fairness and throughput under mixed-traffic scenarios, resource wastage and uneven QoS persist, especially under congestion or when some traffic types dominate.

RQ5: Key challenges, limitations, and research gaps

After reviewing some existing RA schemes, some of the unresolved challenges are:

i. Static or semi-dynamic RA policies that fail to adapt to real-time fluctuations

ii. Unfair resource distribution for low-priority or remote devices

iii. Inefficient energy usage from repeated retransmissions or over-provisioning

iv. Scalability issues of centralized or computationally intensive models

v. Lack of intra-use-case differentiation and cross-domain coordination

vi. This shows that while the state of the art has progressed, many schemes still fall short of fully addressing 5G heterogeneity and real-time dynamics.

RQ6: Future research directions and emerging approaches

Section 6.0 synthesized future opportunities. Hybrid and adaptive frameworks, intelligent resource reclamation, fine-grained QoS differentiation, lightweight/federated learning-based RA models, and cross-layer integration were identified as promising directions. These directly stem from gaps highlighted in RQs 3–5 and provide a roadmap for next-generation RA studies. Table 5 summarizes how the research questions were addressed.

4.1 Summary of Findings

The research questions have been systematically addressed by (i) mapping the publication trends and regional spread (RQ1), (ii) classifying RA techniques and strategies (RQ2), (iii) evaluating their strengths and weaknesses (RQ3), (iv) examining coexistence approaches (RQ4), (v) identifying key challenges and limitations (RQ5), and (vi) synthesizing future research opportunities (RQ6). The findings confirm that although significant progress has been made in resource allocation for 5G networks, open challenges remain in fairness, energy efficiency, and scalability, requiring new, adaptive, and integrated RA paradigms.

5.0 CHALLENGES AND LIMITATIONS

This sub-section will summarize the challenges and limitations of the reviewed existing RA schemes. The challenges and limitations for the eMBB-based schemes were first discussed, followed by mMTC-based schemes, then URLLC-based schemes and finally these from multiple use case-based schemes.

Despite extensive efforts to optimize RA for eMBB in 5G networks, current algorithms and frameworks exhibit several persistent limitations that highlight critical research gaps. A major concern is the trade-off in QoS for low-priority or high-demand users, as many schemes, such as those in [25] and [26] – deprioritize bandwidth-intensive or lower-priority requests, resulting in increased delays, packet loss, and unmet QoS requirements. Furthermore, unfair resource distribution is evident in approaches like those of [20] and [27] which favors high-priority or high-bandwidth traffic at the expense of other service categories. Some QoE-driven or traffic-reserved strategies, like the one proposed by [12] often lead to resource underutilization when specific traffic types are inactive. Additionally, heavy reliance on channel quality, as seen in the methods of [21] results in poor performance for users in degraded radio environments. Latency sensitivity is also inadequately addressed in some fixed-priority schemes, notably those by [22] and [26] which fail to meet the demands of latency-critical applications such as VR and real-time video streaming. Although successful, intelligent learning-based models, such as the DRL-based scheduler [27], have limited scalability and a high computational overhead, which makes them impractical for real-time implementation in large-scale networks. Lastly, a lot of the current approaches rely on static or semi-dynamic policies that are unable to handle real-time variations in network load and user behaviour, rather than dynamic adaptability across traffic patterns. All of these issues highlight the necessity of more flexible, equitable, and effective RA solutions in next 5G and beyond eMBB systems.

Several significant issues remain after examining some of the current RA systems suggested for mMTC in 5G networks, indicating significant research gaps. As demonstrated in [28, 29], many current schemes have poor resource utilisation, particularly when allocations are made too soon or based on erroneous traffic estimates. In systems where prioritisation is determined by proximity or delay sensitivity, like in [30] and [32], distant or low-

https://doi.org/10.4314/njt.2026.5706

priority MTC devices frequently suffer from increased latency, packet losses, or decreased QoS. Some customers experience poor service reliability as a result of congestion and collisions that are not resolved when the network is heavily loaded [34, 37]. Furthermore, many models' energy efficiency is harmed by overhead from intricate control systems or repeated access attempts [34], 36]. Although promising, learning-based and dynamic techniques are limited in their scalability and flexibility in dynamic contexts due to their significant computing expense and requirement for precise real-time information [38, 40]. Furthermore, interference and management complexity may be introduced by overlapping beam and contention mechanisms [39]. These drawbacks highlight the necessity for RA methods that are lightweight, flexible, and fair in order to function well in real-time 5G mMTC deployments, sustain QoS across a range of device classes, and function well under fluctuating traffic densities.

Even while URLLC-based RA systems have advanced significantly, there are still significant drawbacks that draw attention to unmet research needs. As demonstrated in [41] and [43], many methods either waste resources by reserving bandwidth needlessly or are not flexible enough to accommodate changing traffic loads. Particularly in crowded situations, packet drops and delay violations result from the frequent limitation or ineffective management of reliability mechanisms such as retransmissions [42, 44]. Furthermore, unlimited or repetitive retransmissions reduce battery efficiency in a number of ways [45] and [48]. High-priority traffic is served at the expense of lower-priority or remote users, resulting in an uneven application of fairness and QoS guarantees that lowers system-wide throughput and fairness [47] and [50]. Furthermore, despite their promise, learning-based approaches struggle with scalability because of their high computational complexity and reliance on accurate feedback and system state models [51]. All of these issues point to the necessity of resource allocation models that are reliable, scalable, and energy-efficient in order to provide end-to-end fairness and dependability under a variety of changing URLLC service demands. Numerous enduring restrictions draw attention to significant drawbacks and difficulties with the current RA methods that have been suggested for various use cases. As demonstrated in the works of [58, 59, and 67], many schemes prioritise particular traffic types (e.g., URLLC or eMBB), which degrades QoS and increases delays for lower-priority or unclassified traffic during congestion. Underutilization of resources is still a prevalent problem, especially when traffic-specific or static reservations are implemented, and the anticipated traffic is not there, as these solutions suggested by [55, 57, and 58] demonstrate.

Furthermore, under-optimized transmission patterns or numerous retransmissions result in energy inefficiencies [56 and 66]. The schemes in [62, 63, and 64] are examples of advanced models that use reinforcement learning or optimization algorithms, which involve considerable computational cost and require accurate traffic forecasts or real-time network state awareness to function reliably. Furthermore, there is either a lack of detail or an oversimplification of the cross-domain coordination between the satellite, RAN, and edge computing layers. These drawbacks highlight the necessity of cross-layer, lightweight, and adaptable resource allocation systems that guarantee efficiency, scalability, and equity in dynamic, diverse service contexts.

6.0 FUTURE RESEARCH DIRECTIONS

The discussion in section 5.0 highlights critical limitations and challenges in existing RA schemes for 5G networks, particularly in handling diverse use cases (eMBB, URLLC, mMTC) and dynamic network conditions. For researchers, practitioners, telecommunication experts, and policy makers, the following research directions should be considered:

i. The lack of hybrid RA frameworks and schemes that can simultaneously manage the heterogeneous QoS requirements of all three use cases. Future work should focus on developing adaptive algorithms that dynamically balance resources across eMBB (bandwidth-intensive), URLLC (low-latency), and mMTC (massive connectivity) traffic, ensuring fairness and efficiency without degrading performance for any service category.

ii. Existing schemes often treat each use case as monolithic, ignoring intra-use case variations, for example, delay-sensitive eMBB applications (e.g., real-time video streaming) versus delay-tolerant ones (e.g., file downloads). Addressing this requires fine-grained QoS differentiation mechanisms that optimize resource distribution based on real-time traffic subtypes, preventing inefficiencies in scenarios with mixed priority demands.

iii. The static or semi-dynamic nature of many RA schemes, which fail to adapt to fluctuating network loads and user behaviour. Future research should explore lightweight, learning-based models that leverage real-time data (e.g., channel conditions,

https://doi.org/10.4314/njt.2026.5706

traffic patterns) to dynamically adjust resource allocation, minimizing underutilization and over-reservation. This is particularly crucial for URLLC and mMTC, where premature resource reservations or rigid scheduling lead to inefficiencies.

iv. Many existing algorithms allocate fixed resources or reserve bandwidth for inactive traffic, resulting in wastage. Novel approaches should incorporate intelligent resource reclaiming mechanisms, where unused allocations are dynamically reassigned to active traffic, improving overall network efficiency.

v. The need for broader performance evaluation metrics. Most studies compare proposed schemes against a limited set of benchmarks, often neglecting system-wide factors like energy efficiency, fairness, and scalability. Future research should adopt comprehensive evaluation frameworks that assess RA algorithms not only on throughput and latency but also on energy consumption (e.g., reducing retransmissions in URLLC), fairness (e.g., equitable treatment of distant vs. Proximate users), and scalability (e.g., performance in ultra-dense mMTC deployments).

vi. Moreover, although learning-based models (like DRL) exhibit potential, their substantial computational overhead restricts their practicality. Decentralised or federated learning architectures should be the main focus of research in order to minimise complexity while preserving flexibility.

vii. Finally, cross-layer and cross-domain coordination remains underexplored. Many schemes oversimplify interactions between RAN, edge computing, and satellite layers, leading to suboptimal end-to-end performance. Future work should investigate integrated RA frameworks that enable seamless resource sharing across network layers, ensuring reliability and low latency in multi-service environments. Addressing these gaps will require collaboration between academia and industry to validate solutions in real-world testbeds, bridging the gap between theoretical advancements and practical deployment in 5G and beyond networks.

7.0 RESOURCE ALLOCATION RECOMMENDATIONS TOWARD 6G COMMUNICATION SYSTEMS

Sixth-generation (6G) communication systems are envisioned to support extreme performance targets, including terabit-per-second data rates, sub-millisecond end-to-end latency, ultra-high reliability, massive device connectivity, and seamless integration of communication, sensing, and computing. Achieving these goals will require a fundamental evolution of resource allocation (RA) mechanisms beyond those currently employed in 5G networks [70]. Based on the limitations and challenges identified in the reviewed literature, several key recommendations for 6G-oriented RA design are discussed below:

7.1 AI-native Resource Allocation

Unlike 5G systems, where artificial intelligence (AI) is often applied as an add-on optimization tool, 6G networks are expected to adopt AI-native architectures in which learning-based decision-making is deeply integrated into the radio access network. Resource allocation should increasingly rely on advanced machine learning and deep reinforcement learning techniques capable of operating under dynamic, uncertain, and high-dimensional environments. These approaches can enable real-time adaptation to traffic variations, mobility patterns, and heterogeneous service demands. To ensure practical deployability, future research should emphasize lightweight learning models, online and federated learning frameworks, and robustness against imperfect or delayed information [71].

7.2 Integrated Sensing and Communication–Assisted Resource Allocation

Integrated sensing and communication (ISAC) is a defining feature of 6G systems, enabling networks to jointly sense the environment and deliver data services. Sensing information, such as user location, mobility trajectories, obstacle detection, and environmental awareness, can be exploited to enhance resource allocation decisions. For example, sensing-assisted channel prediction and mobility-aware scheduling can reduce uncertainty in channel state information, thereby improving reliability and latency performance for mission-critical services. Incorporating sensing feedback into RA frameworks offers new opportunities for proactive and context-aware resource management [71].

7.3 Context-Aware and Semantic Resource Allocation

6G networks are expected to support context-aware and semantic communications, where the objective shifts from maximizing raw data throughput to ensuring task-oriented or goal-oriented performance. Resource allocation strategies should therefore account for application-level context, user intent, and data relevance. By prioritizing semantically meaningful information, future RA mechanisms can reduce unnecessary transmissions and improve

https://doi.org/10.4314/njt.2026.5706

spectral and energy efficiency, particularly for ultra-dense and massive IoT scenarios [70].

7.4 Joint Communication–Computing–Sensing Resource Allocation

The convergence of communication, computing, and sensing in 6G necessitates joint multi-domain resource allocation. Future RA frameworks should simultaneously consider radio resources, edge computing capacity, sensing accuracy, and energy constraints. Such holistic allocation is particularly important for latency-sensitive and computation-intensive applications such as extended reality, autonomous systems, and industrial automation. Cross-layer and cross-domain optimization will be essential to meet stringent quality-of-service requirements [70].

7.5 Reliability and Sustainability-Aware Resource Allocation

Extreme reliability requirements and sustainability considerations will play a central role in 6G system design. Resource allocation schemes should incorporate reliability guarantees under uncertainty while minimizing energy consumption and operational costs. This includes adaptive redundancy, intelligent packet duplication, and energy-aware scheduling strategies. AI-driven predictive mechanisms can further support sustainable operation by anticipating traffic and channel conditions, enabling proactive resource provisioning [71].

8.0 CONCLUSION

This paper has provided a comprehensive review of existing resource allocation (RA) schemes in 5G networks, covering the three core service categories: eMBB, mMTC, and URLLC. The analysis examined RA approaches designed for individual use cases as well as schemes that aim to support multiple services simultaneously. Although substantial progress has been made in aligning resource allocation strategies with specific traffic requirements, many existing solutions still face notable challenges. These include high computational overhead, limited fairness, inefficient resource utilization in certain scenarios, and difficulties in scaling to dense or highly dynamic network conditions. Moreover, most approaches are tailored to isolated use cases and offer limited flexibility when applied to mixed traffic environments. The review also highlights several open research gaps, particularly in areas such as cross-slice optimization, adaptive and real-time scheduling, and energy-efficient resource management. Looking ahead, the adoption of AI-driven methods, context-aware decision-making, and real-time analytics presents a promising path toward more flexible, robust, and fair resource allocation frameworks for future 5G networks.

Table 3: List of acronyms and definitions

Acronym	Definition
3GPP	Third Generation Partnership Project
4G	Fourth Generation Mobile Communication System
5G	Fifth Generation Mobile Communication System
6G	Sixth Generation Mobile Communication System
ACB	Access Class Barring
AI	Artificial Intelligence
AMC	Adaptive Modulation and Coding
AP	Access Point
AR	Augmented Reality
BS	Base Station
BCD	Block Coordinate Descent
CAPEX	Capital Expenditure
CoMP	Coordinated Multi-Point Transmission
CQI	Channel Quality Indicator
CSI	Channel State Information
CTMC	Continuous-Time Markov Chain
CMDP	Constrained Markov Decision Process
D2D	Device-to-Device Communication
DPCI	Dynamic Power Control Indicator
DRL	Deep Reinforcement Learning
DQN	Deep Q-Network
DDQN	Double Deep Q-Network
DSM	Decision Selection Mechanism
eMBB	Enhanced Mobile Broadband
EE	Energy Efficiency
gNB	Next-Generation Node B
GA	Genetic Algorithm
HARQ	Hybrid Automatic Repeat Request
H2H	Human-to-Human Communication
IoT	Internet of Things
ISAC	Integrated Sensing and Communication
KPI	Key Performance Indicator
LEO	Low-Earth Orbit
LTE	Long-Term Evolution
Lyapunov	Lyapunov Optimization Framework
mMIMO	Massive Multiple-Input Multiple-Output
mMTC	Massive Machine-Type Communication
MEC	Mobile Edge Computing
MILP	Mixed-Integer Linear Programming
ML	Machine Learning
M2M	Machine-to-Machine Communication
mmWave	Millimetre-Wave Communication
NFV	Network Function Virtualization
NOMA	Non-Orthogonal Multiple Access
OMA	Orthogonal Multiple Access
OPEX	Operational Expenditure

https://doi.org/10.4314/njt.2026.5706

PF	Proportional Fair Scheduling
PHY	Physical Layer
PRACH	Physical Random Access Channel
PRB	Physical Resource Block
PUSCH	Physical Uplink Shared Channel
QoE	Quality of Experience
QoS	Quality of Service
RA	Resource Allocation
RAN	Radio Access Network
RIS	Reconfigurable Intelligent Surface
ROPC	Robust Optimal Power Control
RT	Real-Time
SCA	Successive Convex Approximation
SCMA	Sparse Code Multiple Access
SINR	Signal-to-Interference-plus-Noise Ratio
SISO	Single-Input Single-Output
SLA	Service Level Agreement
SNR	Signal-to-Noise Ratio
TDMA	Time Division Multiple Access
TOPSIS	Technique for Order Preference by Similarity to Ideal Solution
UE	User Equipment
UAV	Unmanned Aerial Vehicle
URLLC	Ultra-Reliable Low-Latency Communication
V2X	Vehicle-to-Everything Communication
VR	Virtual Reality
WiMAX	Worldwide Interoperability for Microwave Access

REFERENCES

[1]. Umar, M. M., Mohammed, A., & Abdulazeez, A. "Review of QoS-aware resource allocation schemes for 5G networks". *Dutse Journal of Pure and Applied Sciences*, 10(3c), pp. 296-303, 2024. doi:10.4314/dujopas.v10i3c.28

[2]. Sufyan, A., Khan, K. B., Khashan, O. A., Mir, T., & Mir, U. "From 5G to beyond 5G: A comprehensive survey of wireless network evolution, challenges, and promising technologies". *Electronics, 12(10),* pp 1 - 25, 2023 doi: 10.3390/electronics12102200

[3]. Shirvani Moghaddam, S. "The past, present, and future of the internet: A statistical, technical, and functional comparison of wired/Wireless fixed/Mobile internet". *Electronics*, 13, pp 1 - 35, 2024. doi:10.20944/preprints202404.1855.v1

[4]. Scalise, PP., Boeding, M., Hempel, M., Sharif, H., Delloiacovo, J., & Reed, J. "A systematic survey on 5G and 6G security considerations, challenges, trends, and research areas". *Future Internet*, 16(3), pp 1 – 38, 2024. doi:10.3390/fi16030067

[5]. Idowu-Bismark, O., Idachaba, F., & Atayero, A. A. "Large-scale parameter modelling for millimetre-wave multiple-input multiple-output channel in 5G ultra-dense network". *Indonesian Journal of Electrical Engineering and Computer Science*, 26(2), pp.794 - 807, 2022. doi:10.11591/ijeecs.v26.i2.pp794-807

[6]. Mchangama, A., Ayadi, J., Jiménez, V. PP., & Consoli, A. "MmWave massive MIMO small cells for 5G and beyond mobile networks: An overview". In *Proc. 12th International Symposium on Communication Systems, Networks and Digital Signal Processing (CSNDSP)*, pp.1-6, 2020, Porto, Portugal. doi:10.1109/csndsp49049.2020.9249602

[7]. Eldowek, B. M., Bauomy, N. A., El-Rabaie, E. M., Abd El-Atty, S. M., & Abd El-Samie, F. E. "A survey of 5G 19millimetre wave, massive multiple-input multiple-output, and vehicle-to-vehicle channel measurements and models. *International Journal of Communication Systems*, 34(16), pp. 13380-13394, 2021. doi:10.1002/dac.4830

[8]. Bogale, T. E., & Le, L. B. "Massive MIMO and mmWave for 5G wireless HetNet: Potential benefits and challenges". *IEEE Vehicular Technology Magazine*, 11(1), pp.64-75, 2026. doi:10.1109/mvt.2015.2496240

[9]. Gao, Z., Dai, L., Mi, D., Wang, Z., Imran, M. A., & Shakir, M. Z. "MmWave massive-MIMO-based wireless backhaul for the 5G ultra-dense network". *IEEE Wireless Communications*, 22(5), pp.13-21, 2015. doi: 10.1109/mwc.2015.7306533

[10]. Shaban, M., "Development and implementation of high-gain, and high-isolation multi-input multi-output antenna for 5G mmWave communications". *Telecom*, 6(1), pp. 1-17, 2025. doi:10.3390/telecom6010014

[11]. Maniru M. U, A., Mohammed, A., Almu, K. O. Aremu. Study on Resource Allocation Schemes for 5G Heterogeneous Networks. International Conference on Technological Solutions for Smart Economy, pp.128-135, 2024

[12]. Kumar, R., Sinwar, D., & Singh, V. "QoS aware resource allocation for coexistence mechanisms between eMBB and URLLC: Issues, challenges, and future directions in 5G". *Computer Communications*, 213, pp.208-235, 2024 doi:10.1016/j.comcom.2023.10.024

[13]. Sudhamani, C., Roslee, M., Tiang, J. J., & Rehman, A. U. "A survey on 5G coverage improvement techniques: Issues and future challenges". *Sensors*, 23(4), pp. 1 – 47, 2023. doi:10.3390/s23042356. 2023

https://doi.org/10.4314/njt.2026.5706

[14]. Khadidos, A. O., Manoharan, H., Selvarajan, S., Khadidos, A. O., Alshareef, A. M., & Altwijri, M. "Distribution of resources beyond 5G networks with heterogeneous parallel processing and graph optimization algorithms". *Cluster Computing*, 27(6), pp.8269-8287, 2024. doi:10.1007/s10586-024-04367-w

[15]. Ibrahim, S., Younis, Y. S., Hamza, K. S., & Ashour, M. M. "Improving resource allocation in 5G networks using traffic segmentation based on machine learning techniques". *International Journal of Telecommunications*, 05 (1), pp.1-15, 2025. doi:10.21608/ijt.2025.372415.1095

[16]. Lee, I., & Kim, D. K. "Decentralized multi-agent DQN-based resource allocation for heterogeneous traffic in V2X communications". *IEEE Access*, 2024 (12), pp.3070-3084, 2024. doi:10.1109/access.2023.3349350

[17]. Boccardi, F., Heath, R. W., Lozano, A., Marzetta, T. L., & Popovski, PP. "Five disruptive technology directions for 5G". *IEEE Communications Magazine*, 52(2), pp.74-80, 2014. doi:10.1109/MCOM.2014.6736746

[18]. Ahmad, A., Paul, A., Khan, M., Jabbar, S., Rathore, M.N., Chilamkurti, N., and Min-Allah, N. "Energy-efficient Hierarchical Management for mobile cloud computing". *IEEE Transactions on Sustainable Computing*, 2 (2), pp. 100-112, 2017. 10.1109/TSUSC.2017.2714344

[19]. Chen, M., Liang, B., and Dong, M. "Multi-user Multi-task offloading and resource allocation in mobile cloud systems". *IEEE Transactions on Wireless Communications.* 17 (10), *pp.* 6790-680. 2018. doi:10.1109/TWC.2018.2864559

[20]. Yun, J., Piran, J. MD., and Suh, D.Y. "QoE-driven resource allocation for live video streaming over D2D underlaid 5G cellular network". *IEEE Access*, 6, pp. 72563-72580. 2018. doi:1109/ACCESS.2018.2882441

[21]. Desogus, C., Anedda, A., Murroni, M., and Muntean, G. "A traffic type-based differentiated reputation algorithm for radio resource allocation during multi-service content delivery in a 5G heterogeneous scenario". *IEEE Access*, 7, pp. 27720-27735, 2019. doi:10.1109/ACCESS.2019.2902190

[22]. Oladejo, S. O and Falowo, O. E. "Latency-aware dynamic resource allocation scheme for multi-tier 5G network: A network slicing multitenancy scenario". *IEEE Access*, 8, pp. 74834-7485, 2020. 10.1109/ACCESS.2020.2988710

[23]. Cola, T. D., and Bisio, I. "QoS Optimization for eMBB services in converged 5G-satellite networks". *IEEE Transactions on Vehicular Technology*, 14 (8), pp. 1-13, 2020.

[24]. Yang, Y., Feng, L., Zhang, C., Ou, Q., and Li, W. "Resource allocation for virtual reality content sharing based on 5G D2D multicast communication". *EURASIP Journal on Wireless Communication and Networking*, 2020(112), pp. 1-12, 2020

[25]. Alencar, D., Both, C., Antunes, R., Olivera, H., Cerqueira, E., and Rosario, D. "Dynamic microservice allocation for virtual reality distribution with QoE support" *IEEE Transactions on Network and Service Management*, 19 (1), pp. 729-740, 2021

[26]. Si, O., Zhao, J., Han, H., Lam, K.Y., and Liu, Y. "Resource Allocation and Resolution Control in the metaverse with mobile Augmented Reality". *Proceedings of IEEE Global Communications Conference (GLOBECOM)*, Rio de Janeiro, Brazil, pp. 1-7, 2022

[27]. Madi, N.K.M., Nasrallah, M.M., and Hanapi, Z.M. "Delay-based resource allocation with fairness guarantee and minimal loss for eMBB in 5G heterogeneous networks". *IEEE Access*, 10, pp. 75619-75636, 2022. doi: 10.1109/ACCESS.2022.3192450

[28]. Benmadani, E. H., Azni, M., Essa Alharbi, T., Alzaidi, M. S., & Tounsi, M. "Deep reinforcement learning-based dynamic scheduling for real-time applications in LTE and RAN slicing for eMBB in 5G". *IEEE Access*, 13, pp.33555-33570, 2025. doi: 10.1109/access.2025.3541531

[29]. Miuccio, L., Panno, D., and Riolo, S. "Dynamic Uplink resource dimensioning for massive MTC in 5G Networks based on SCMA". *Proceedings of the 25th European Wireless Conference, Aarhus, Denmark*, pp. 1-5. 2019

[30]. Guo, J., Durrani, S., Zhou, X., and Yanikomeroglu. "Massive Machine Type Communication with Data Aggregation and Resource Scheduling". *IEEE Transactions on Communication*, 65 (90) pp. 4012-4026, 2017

[31]. [31] Onel, L., Lopez, A., Alves, H., Pedro, H.J., and Nardelli, M.L. "Aggregation and Resource Scheduling in Machine-type Communication Networks: A Stochastic Geometry Approach". *IEEE Transactions on Wireless Communication*, 17 (7), pp. 4750-4765, 2018

https://doi.org/10.4314/njt.2026.5706

[32]. Teng, Y., Liang, W., Zhang, Y., and Yang, R. "Traffic-aware resource allocation scheme for mMTC in dynamic TDD system". *IET Communications,* 12 (15), pp. 1910-1918, 2018.

[33]. Salam, T., Rehman, W. U., and Tao, X. "Cooperative Data Aggregation and Dynamic Resource Allocation for Massive Machine Type Communication". *IEEE Access,* 6, pp. 4145-4158, 2017.

[34]. Miuccio, L., Panno, D., and Riolo, S. "Joint congestion control and resource allocation for massive MTC in 5G network based on SCMA". *Proceedings of the 15th International Conference on Telecommunications, Gras, Austria,* pp. 1-8. 2019.

[35]. Rehman, W. U., Salam, T., Almogren, A., Haseeb, K., Din I, U., and Bouk S, H. "Improved Resource Allocation in 5G MTC Networks". *IEEE Access,* 8, pp. 49187-49197, 2020.

[36]. Miuccio, L., Panno, D., and Riolo, S. "A new contention-based PUSCH resource allocation in 5G NR for mMTC scenarios". *IEEE Communication Letter*, 25 (3), pp. 802-806, 2021.

[37]. Yu, Baoquan., Cai, Yueming., and Wu, Dan. "Joint Access Control and Resource Allocation for Short-Packet-Based mMTC in Status Update Systems". *IEEE Journal on Selected Areas in Communications,* 39 (3), pp. 851-865, 2021.

[38]. Weerasinghe, T. N., Casares-Giner, V., Balapuwaduge, I.A.M., and Li, F.Y. "Priority enabled grant-free access with dynamic slot allocation for heterogeneous mMTC traffic in 5G NR Network". *IEEE Transaction on Communications*, 69 (5), pp. 3192-3206, 2021.

[39]. Gupta, R. K., Kumar, S., & Misra, R. "Resource allocation for UAV-assisted 5G mMTC slicing networks using deep reinforcement learning". *Telecommunication Systems,* 82(1), pp. 141-159, 2022. doi:10.1007/s11235-022-00974-3

[40]. Fu, X., Shen, Q., Yang, B., & Gao, X. "Dynamic provisioning of random-access capacity in mMTC Slice based on beam splitting/Merging". *IEEE Internet of Things Journal*, 11(2), pp. 3331-3347, 2024. doi:10.1109/jiot.2023.3296148

[41]. Huang, R., WushaoW., Zhi, Z., Chongwu D., and Xu C. "MEC-Enabled Task Replication with Resource Allocation for Reliability-Sensitive Services in 5G mMTC Networks". IEEE Transaction on Services Computing, 18 (1), pp. 253-269, 2025.

[42]. Han, Y., Elayoubi, S. E., Serrano, A. G., Varma, V. S., and Messai, M. "Periodic Radio Resource Allocation to meet latency and reliability requirements in 5G Networks". *Proceedings of the 87th IEEE Vehicular Technology Conference (VTC),* Parto, Portugal, pp. 1-6, 2018.

[43]. Ren, H., Pan, C., Deng, Y., Elkashlan, M., and Nallanathan, A. "Resource allocation for URLLC in 5G Mission-critical IoT Networks". *Proceedings of 2019 IEEE International Conference on Communications (ICC),* Shanghai, China, pp. 1-6, 2019.

[44]. Ghanem, W.R., Jamali, V., Sun, Y., and Schober, R. "Resource allocation for multi-user downlink MISO OFDMA-URLLC Systems". *IEEE Transactions on Communications*, 68 (11) pp. 1-18, 2020.

[45]. Ren, H., Pan, C., Deng, Y., Elkashlan, M., and Nallanathan, A. "Resource Allocation for Secure URLLC in Mission-critical IoT Scenarios". *IEEE Transactions on Communications*, 68(9), pp. 5793-5807, 2020.

[46]. Feng, L., Li, W., Lin, Y., Zhu, L., Guo, S., and Zhen, Z. "Joint Computation Offloading and URLLC Resource Allocation for Collaborative MEC Assisted Cellular-V2X Networks". *IEEE Access,* 8, pp. 24914 – 24926, 2020.

[47]. Ghanem, W. R., Jamali, V., Sun, Y., and Schober, R. "Resource Allocation for multi-user URLLC-OFDMA Systems". *Proceedings of the IEEE International Conference on Communications Workshop*, Shanghai, China, pp. 1-6, 2019.

[48]. Khan, J., and Jacob, L. "Resource Allocation for CoMP enabled URLLC in 5G-RAN architecture". *IEEE System Journal,* 15 (4), pp. 4864-4875, 2021.

[49]. Nasir, A. A., Tuan, H. D., Nguyen, H. H., Debbah, M., and Poor, H. V. "Resource allocation and beamforming design in short blocklength regime for URLLC". *IEEE Transactions on Wireless Communications,* 20 (2), pp. 1321-1335, 2021.

[50]. Gao, Y., Yang, H., Hong, X., and Chen, L. "A hybrid scheme for MCS Selection and Spectrum Allocation for URLLC traffic under delay and reliability constraints". *Entropy,* 24 (5), pp. 1-13, 2022.

[51]. Haider, V.T., Mehmeti, F., Cantarero, A., and Kellerer, W. "Joint fair allocation of RAN and computing resources to vehicular users with URLLC traffic". *Proceedings of the IEEE 20th Annual Consumer Communication & Networking Conference,* München, Germany, pp. 1-9, 2023.

https://doi.org/10.4314/njt.2026.5706

[52]. Wu, Q., Wang, W., Fan, PP., Fan, Q., Wang, J., & Letaief, K. B. "URLLC-aware resource allocation for heterogeneous vehicular edge computing". *IEEE Transactions on Vehicular Technology*, 73(8), pp.11789-11805, 2024. doi:10.1109/tvt.2024.3370196. 2024

[53]. Li, Y., Zhao, Y., Li, J., Zhang, J., Yu, X., and Zhang, J. "Side Channel Attack-Aware Resource Allocation for URLLC and eMBB Slices in 5G RAN". *IEEE Access*, 8, pp. 2090-2099. 2019.

[54]. Darabi, M., and Lampe, L. Multi-Objective Resource Allocation for Joint eMBB and URLLC Traffic with Different QoS Requirements. *Proceedings of the IEEE GlobeCom Workshops, Waikoloa*, HI, USA, pp. 1-6, 2019.

[55]. Korrai, PP., Lagunas, E., Sharma, K., Chatzinotas, S., and Ottersten, B. "Slicing-based resource allocation for multiplexing of eMBB and URLLC services in 5G wireless networks". *Proceedings of the 24th International Workshop of Computer Aided Modeling & Design of Communication Links and Networks, 2019*, Limassol, Cyprus, pp. 1-5, 2019.

[56]. Ma, T., Zhang, Y., Wang, F., Wang, D., and Guo, D. "Slicing resource allocation for eMBB and URLLC in 5G RAN". *Wireless Communications and Mobile Computing*, pp. 1-11, 2020.

[57]. Pradhan, A., and Das, S. "Joint Preference Metric for Efficient Resource Allocation in Co-existence of eMBB and URLLC". *Proceedings of the 12th International Conference on Communication Systems & Networks*, Bengaluru, India, pp. 1-3, 2020.

[58]. Al-Ali, M., Yaacoub, E., and Mohamed, A. "Dynamic Resource Allocation of eMBB-URLLC Traffic in 5G New Radio". *Proceedings of the 2020 IEEE International Conference on Advanced Networks and Telecommunications Systems*, New Delhi, India, pp. 1-6, 2020.

[59]. Zhang, X., Guo, X., and Zhang, H. "RB allocation scheme for eMBB and URLLC Coexistence in 5G and beyond". *Wireless Communications and Mobile Computing*, 2021 (1), *pp. 1-7*, 2021. doi: 10.1155/2021/6644323.

[60]. Beshley, H., Beshley, M., Medvestskyi, M., and Pyrih, J. "QoS-Aware Optimal Radio Resource Allocation Method for Machine-Type Communications in 5G LTE and beyond Cellular Networks". *Wireless Communications and Mobile Computing*, 2021 (1), pp. 1-18, 2021. doi: 10.1155/2021/9966366.

[61]. Han, X., Xiao, K., Liu, R., Liu, X., Alexandropoulos, G. C and Jin, S. "Dynamic resource allocation schemes for eMBB and URLLC services in 5G wireless networks". *Intelligent and Converged Networks*, 3 (2), pp. 145-160, 2022.

[62]. Sohaib, R. M., Onireti, O., Sambo, Y., Swash, R., Ansari, S., & Imran, M. A. "Intelligent resource management for eMBB and URLLC in 5G and beyond wireless networks". *IEEE Access*, 11, pp.65205-65221, 2023. Doi:10.1109/access.2023.3288698

[63]. Zhou W, Azharul I, and Kyung Hi. "Real-time RL-based 5G network slicing design and traffic model distribution: Implementation for V2X and eMBB services". *KSII Transactions on Internet and Information Systems*, 17(9), pp. 2573-258, 2023. doi: 10.3837/tiis.2023.09.014

[64]. Ahsan, M., Vu, T. X., & Chatzinatos, S. "Flexible resource allocation for eMBB and mMTC services in a time-varying satellite topology". *EEE International Conference on Communications Workshops (ICC Workshops)*, Denver, CO, USA, pp. 1-6, 2024,. doi:10.1109/iccworkshops59551.2024.10615494

[65]. Shekhar, C., & Singh, PP. "Optimization of resource allocation in 5G networks: A network slicing approach with hybrid NOMA for enhanced uRLLC and eMBB coexistence". *International Journal of Communication Systems*, 37(17), pp.1-24, 2024 doi:10.21203/rs.3.rs-4275233/v1

[66]. Souza, C., Falcão, M., Balieiro, A., Alves, E., & Taleb, T. "Dynamic resource allocation for URLLC and eMBB in MEC-NFV 5G networks". *Computer Networks*, 260 (C), pp.1 – 18, 2025. doi:10.1016/j.comnet.2025.111127. 2025

[67]. Mamane, A., Fattah, M., Ghazi, M., Bekkali, M., Balboul, Y., and Mazer, S. "Scheduling algorithms for 5G networks and beyond: Classification and Survey". *IEEE Access*, 10, pp. 51643-51661, 2022.

[68]. Attar, H., Issa, H., Ababneh, J., Abbasi, M., Solyman, A. A., Khosravi, M., & Said Agieb, R. "5G system overview for ongoing smart applications: Structure, requirements, and specifications". *Computational Intelligence and Neuroscience*, 2022 (1), pp.1-11, 2022. doi: 10.1155/2022/2476841

[69]. Miuccio, L., Panno, D., and Riolo, S. "Dynamic Uplink resource dimensioning for massive MTC in 5G Networks based on

https://doi.org/10.4314/njt.2026.5706

SCMA". *Proceedings of the 5th International Conference on Computer Information and Big Data Applications*, Wuhan, China, pp.490-496, 2024. doi: 10.1145/3671151.3671240

[70]. Kulkani, N. PP., Mantri, D. S., Prasad, N. R., Pawar, PP. M., & Prasad, R. (2022). "6G future vision: Requirements, design issues and applications". *6G Enabling Technologies*, River Publishers, pp.23-43. doi: 10.1201/9781003360889-2

[71]. Tianjiao C, Qinqin T, Guangyi L. "Efficient Task Scheduling and Resource Allocation for AI Training Services in Native AI Wireless Networks 2023." *IEEE International Conference on Communications Workshops (ICC Workshops)*, Rome, Italy, pp. 1342-1348, 2023. Doi: 10.1109/ICCWorkshops57953.2023.10283537

[72]. Ko, H., Lee, J., and Pack, S. "Priority-based dynamic resource allocation scheme in Network slicing". *Proceedings of the 2021 International Conference on Information Networking (ICOIN), 2021,* Juju Island, Korea (South), pp. 62-65, 2021.

https://doi.org/10.4314/njt.2026.5706