https://dx.doi.org/10.4314/dujopas.v10i3c.28
ISSN (Print): 2476-8316
ISSN (Online): 2635-3490
Dutse Journal of Pure and Applied Sciences (DUJOPAS), *Vol. 10 No. 3c September 2024*

Review of QoS-Aware Resource Allocation Schemes for 5g Networks

[1]Maniru Malami Umar, [1]Aminu Mohammed, [1]Abdulhakeem Abdulazeez

[1]Department of Computer Science,
Usmanu Danfodiyo University,
Sokoto

Email: manirutambuwal@gmail.com

Abstract
Resource management in wireless networks refers to the processes and techniques used to allocate, control, and optimize the various resources in a wireless communication system to ensure efficient and reliable operation. These resources typically include spectrum (frequency bands), power, time slots, and network infrastructure (e.g., base stations, access points). Effective resource management is crucial for maximizing network performance, ensuring Quality of Service (QoS), and providing a good user experience, especially in environments with limited resources and high demand. Quality of Service (QoS) refers to the ability of the network to provide different levels of service quality to various types of traffic, applications, and users. QoS in 5G is essential due to the diverse range of services the network is designed to support, each with its own specific requirements for performance metrics such as latency, reliability, bandwidth, and availability. In this paper, we study how the QoS-aware resource allocation procedure is handled in wireless networks. Different resource management techniques were studied and also the paper looked at some of the existing QoS-aware resources allocation schemes by highlighting the operation, strength and weakness and the QoS parameters considered by each of the schemes.

Keywords: QoS, Wireless Networks, 5G, resource Management, Resource allocation

INTRODUCTION

Wireless broadband (WiBB) technologies have evolved to meet the growing demand for quality of service (QoS). The main WiBB technologies, including WiMAX and LTE networks, have evolved into fourth-generation (4G) technology, aiming to provide high data rates for bandwidth-demanding applications while enhancing flexibility and spectral efficiency. (Mamman *et al.,* 2018). 4G technology offers high-speed network connectivity of 300Mbps and 75Mbps for downlink and uplink, improving multimedia services like VoIP, live video chats, and video conferencing. By 2025, there are expected to be over 75.4 billion linked devices due to the rapid growth of internet users and connected gadgets. In order to communicate with one another, all of these linked devices will require mobile cellular networks, such as sensor networks and machine-to-machine (or machine-type) communication (MTC) (Wortmann & Fluchter, 2015).

The high connectivity needs for communication between such linked devices are above the capabilities of current 4G technologies (Zhang, Wang & Weihua, 2018). Thus, attempts have been made to standardise fifth-generation (5G) technology, which is anticipated to meet these connection requirements, by specialists from academia and industry. This paper gives an review of QoS-aware resource allocation schemes for 5G networks by studying some of the existing schemes developed. Section II gives and overview of 5G network while section III

Author for Correspondence

discuses the Radio resource Management techniques for 5G network. Then, section IV looks into the QoS-aware resource allocation procedure while section V reviews some of the existing QoS-aware schemes. Finally, section VI concludes the paper.

Overview of 5G Networks

Many base stations (BS) connected to numerous user equipment (UEs) make up a typical 5G network. Mesh and Point to Multiple Points (PMP) modes are the two ways in which the UEs and BS communicate with each other. The UEs' ability to communicate is restricted in mesh mode. In the PMP communication mode, data is received from the BS in the downlink (DL) direction and sent by all linked UEs in the uplink (UL) direction. Three (3) logical groups, or use cases, are used by the third-generation partnership project (3GPP) standard (Release 17, 2020) to classify different types of traffic: enhanced Mobile Broadband (eMBB), massive Machine Type Communication (mMTC), and Ultra-reliable low latency communication (URLLC) (Gupta & Jha, 2015). These scenarios are occasionally referred to as 5G technology features.

5G, powered by eMBB and a larger electromagnetic spectrum, is expected to offer faster speeds and more network capacity, with peak data throughput of 10Gbps for UL and 20Gbps for DL. It will enable 500 mph mobility and 100% energy efficiency, supporting services like mobile cloud computing, enterprise collaboration, VR, AR, and video monitoring. (Zhang, Wang & Weihua, 2018)

Massive-Machine-type communication (mMTC) is an automated data transmission method that connects numerous networked devices, including sensors and IoT devices. With 5G, it is expected to support up to 1 million MTC devices/km2 with 500 kmph of mobility, supporting IoT, asset tracking, remote monitoring, smart cities, smart farms, and other industrial and smart society services. (Miuccio, Panno & Riolo, 2021).

The URLLC use case aims to provide low latency and high responsive services, utilizing 5G technology to improve dependability and reduce latency by up to 1ms compared to 4G technology, supporting various services like industrial automation, autonomous vehicles, remote patient monitoring, and site monitoring. (Gao, Yang, Hong & Chen, 2022). Several users with various QoS requirements are connected in a 5G network environment with the intention of communicating.

Radio Resource Management (RRM) Techniques

Efficient radio resource management (RRM) techniques are essential for meeting users' Quality of Service (QoS) demands. These strategies divide available resources across rival users, including scheduling, call admission control, congestion management, and power conservation. Power-saving algorithms aim to maximize battery life by minimizing power usage. (Mohammed, *et al.*, 2019). Congestion control algorithms aim to reduce user waste and delays by lowering the contention rate among UEs. Call admission control techniques regulate the number of users allowed access to a network, deciding whether to approve or deny access. (Mohammed *et al.*, 2018).

Scheduling is an RRM technique that determines the sequence of packets or users served based on network resource availability. It allows competing UEs to share available resources. Scheduling involves both uplink and downlink directions, with schedulers controlling resource distribution for uplink transmission. Downlink schedulers gather packets for specific users and assign resource blocks (RBs) to each user, ensuring efficient network resource sharing. (Mamane *et al*., 2022). Figure 1 shows how resource management process is handled in wireless broadband networks

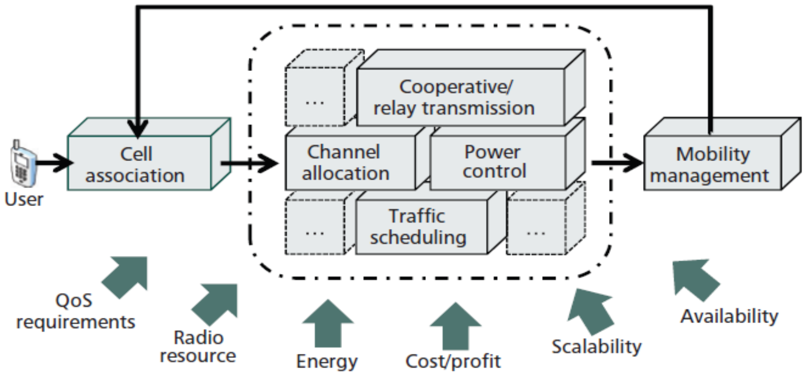

Figure 1: Resource Management in wireless networks Wang, Song, Niyato & Xiao (2015).

Resource allocation in a 5G heterogeneous network is crucial for ensuring peak performance and service quality. This involves effectively managing network resources to various devices and services, including macrocells, small cells, and additional access points. Factors such as frequency spectrum allocation, cell affiliation, load balancing, interference control, network slicing, and QoS differentiation are essential for effective allocation. (Kamal *et al*, 2021).

QoS-Aware Resource Allocation Techniques

QoS-aware resource allocation techniques distribute network resources based on users' unique QoS needs, ensuring access to bandwidth, spectrum, RAN resources, edge computing resources, power, and other resources for optimal performance. These techniques are particularly useful in 5G networks, allowing for dynamic allocation and alteration of resources based on current circumstances, thereby ensuring optimal network performance and a positive user experience by addressing the diverse QoS needs of users and services (Beshley, 2021). Below is the procedure of how a QoS-aware algorithm works:

Step 1: QoS Classification & Prioritization: 5G networks classify traffic based on QoS needs, with different priority levels for applications like voice calls, gaming, video streaming, and IoT data. They also introduce various QoS classes, such as mMTC for Internet of Things, eMBB for high-speed internet, and URLLC for essential applications.

Step 2: Resource Allocation Mechanism: The network uses dynamic resource allocation techniques, such as bandwidth and power distribution based on QoS needs and network conditions, to prioritize low latency for video calls. Scheduling algorithms control resource distribution across users and applications, using strategies like Round Robin, Max-Min justice, and Proportional Fair Scheduling to balance efficiency and justice.

Step 3: Adapting to Network: The network adapts to its network by continuously monitoring user needs, network conditions, and resource allocation. If the network becomes congested, resources may need to be transferred from lower-priority services. To maintain quality of service, the network constantly modifies its coding and modulation schemes based on channel conditions. More resilient but less effective coding algorithms may be employed to maintain reliable connections during bad channel conditions.

Step 4: QoS Enforcement: The network enforces Quality of Service (QoS) through two processes: admission control, where the network checks if resources are available to meet QoS requirements, and resource reservation, where the network can reserve resources in advance for high-priority or time-sensitive services to ensure QoS requirements are met even under high load conditions

Step 5: Interference Management: In dense 5G deployments, managing interference between small cells is crucial. QoS-aware schemes like power control, beamforming, and frequency reuse minimize interference, ensuring QoS for all users.

Any QoS-aware resource allocation scheme should follow the above steps to ensure QoS needs of traffic classes is met for efficient allocation of network resources. The next section will give us a brief on some of the existing schemes developed for 5G networks

Existing QoS-aware resource allocation schemes

This section presents a review of some existing QoS-aware resource allocation scheme for 5G networks. The review gives an insight of the problem solved, objective of the scheme, QoS parameters considered and some of the technical limitations of the schemes.

A QoS-aware scheme for 5G networks is proposed by Wang *et al.,* 2015, which was based on evolutionary game theory, to efficiently manage cell association and antenna allocation. The scheme maximizes data rates and network revenue by considering users' data requirements and available resources. It achieves equilibrium solutions, ensuring neither users nor cells can improve performance by changing their allocation. However, the complexity of managing multiple small cells and users may pose scalability challenges as the network size increases.

Datsika *et al.,* (2018) presented a QoS-aware medium-transparent (QMT) resource management scheme to efficiently allocate resources in optical and wireless domains. The QMT scheme dynamically allocates bandwidth among different Access Points (APs) based on their QoS class, incorporating QoS differentiation by adjusting time slot allocations. Simulation results show that the QMT scheme significantly reduces delay for high-priority traffic classes compared to the FIFO scheme. However, the scheme's reliance on accurate channel state information and potential overhead may pose challenges in practical deployments.

AlSobhi and Aghvami (2019) proposed a QoS-Aware Resource Allocation using Q-learning control algorithms in 5G networks. They introduced two power allocation algorithms based on the Q-learning approach: the Distributed Q-learning Algorithm, where each Femto Access Point (FAP) acts as an agent, and the Formulated Q-learning Algorithm, which redefined states and reward functions based on proximity of MUEs and FAPs. The Cooperative Approach was used to explore its impact on resource allocation and QoS, but it is complex and requires further improvement. The scheme also uses Reward Functions (RFs) to optimize QoS for both MUEs and FUEs, considering interference levels and distances. Simulation results show that the Distributed Q-learning algorithm outperforms the Formulated and Cooperative approaches in power allocation and QoS maintenance.

Bruzgiene, Narbutaite, and Adomkus (2019) developed a QoS-aware resource allocation scheme to meet different traffic's QoS requirements. They used mathematical models and network simulations to understand the relationship between QoS metrics like Bit Error Rate and data loss and their impact on QoE. The study assessed the impact of scheduling strategies on network resilience and service reliability. Results showed that the choice of scheduling algorithm significantly affects both QoS and QoE. The Best CQI algorithm offers high throughput but fairness, while the Round Robin algorithm offers fairer resource distribution but lower overall throughput.

Zhang *et al.*, (2020) propose a resource management mechanism for a virtualized backhaul network of heterogeneous cloud-radio access networks (H-CRANs) using time and wavelength-division-multiplexed passive optical network (TWDM-PON) technology. The architecture combines H-CRAN and TWDM-PON for 5G backhaul, using network virtualization to securely and efficiently share physical network resources among multiple SPs. The scheme uses a wavelength allocation algorithm based on load balancing and a QoS mapping algorithm to address QoS mismatch.

Huang *et al.*, (2020) developed a Joint Task Offloading and QoS-aware Resource Allocation scheme to optimize resource allocation in fog computing environments. The scheme uses an AHP-based evaluation framework to decompose QoS analysis into hierarchical sub-problems, and an RB allocation algorithm to allocate resources based on IoT device priority and satisfaction. A bilateral matching game optimizes the association between fog nodes and IoT devices under various QoS requirements. Simulation results show the scheme efficiently balances network load, improves resource block utilization, and reduces network overhead. However, it faces scalability issues when the network grows larger with diverse IoT devices and real-time implementation complexity in dynamic and heterogeneous network environments.

Beshley *et al.* (2021) proposed a QoS-aware resource allocation scheme to address the saturation problem in 3GPP LTE cellular systems due to the expected increase in machine-type communication (MTC) devices. The scheme includes an enhanced architecture for 5G LTE networks, integrating M2M gateways that prioritize and aggregate MTC traffic before transmission to base stations. The M2M gateway manages four queues for different types of MTC traffic, prioritizing handling based on QoS needs. The optimal radio resource allocation method optimizes the LTE frame formation process, reducing signal traffic and improving resource utilization. The scheme improved radio resource utilization efficiency by 13%, but its effectiveness in real-world deployments remains to be tested.

Bouaziz, *et al.*, (2021) introduced a Quality of Service (QoS)-aware resource allocation and femtocell selection mechanism for 5G networks. The mechanism aims to improve resource utilization, ensure balanced network load, and allocate resources fairly. The scheme outperforms existing methods, achieving higher resource utilization, lower dropped request probability, improved average throughput, and better fairness in resource allocation. However, it does not address inter-cell interference's potential impact on performance in practical deployments.

Boujelben (2021) introduced a resource allocation scheme for Mobile Edge Networks (MENs) that optimizes user association, precoding, and power allocation using millimeter-wave multiple-input multiple-output (MIMO) technology. The scheme uses a max-k-cut problem to minimize interference between users within a small base station. It then uses an iterative algorithm to jointly optimize precoding and power allocation. The convergence of this algorithm is analytically proven, ensuring a solution that meets optimization criteria. The scheme significantly improves Wireless Sensor Network (WSR) performance, manages interference, and meets QoS constraints. However, the complexity of the problem remains high, especially in large-scale networks.

Niu Cao & Pun (2021) proposed a resource allocation scheme for Mobile Edge Networks (MENs) that optimizes user association, precoding, and power allocation using millimeter-wave multiple-input multiple-output (MIMO) technology. The scheme aims to minimize interference between users within a small base station using a distributed local-search algorithm. The iterative algorithm combines the MkC problem for user association with D.C. programming for joint precoding and power allocation. Simulation results show the scheme improves Wireless Sensor Network (WSR) performance, manages interference, and meets QoS constraints. However, the effectiveness depends on channel conditions.

Bolla *et al*., (2022) proposed a method for efficient resource selection for multi-site deployment of 5G-ready vertical applications (vApps). The scheme focuses on grouping components based on interdependencies and affinities, ensuring optimal positioning of microservices. It considers QoS constraints and guides resource selection. The Resource Selection System (RSO) selects appropriate resources for vApp deployment, ensuring QoS parameters are met while minimizing costs and efficient use across multiple sites. Simulation results show the RSO's ability to manage multi-site vApp deployments efficiently, focusing on cost reduction and QoS.

Joda *et al*., (2022) proposed a method for optimizing Component Carriers (CCs) and resource allocation in 5G networks to maximize user throughput while minimizing latency and overhead. The scheme uses Modified Largest Weighted Delay First (M-LWDF) for RB allocation and Gradient with Minimum/Maximum Rate Constraint (GMR) for CC selection and RB allocation. Joint Optimization is applied to minimize CC activations and deactivations frequency to reduce control channel overhead. The scheme achieves a 95.5% reduction in CC activations and deactivations compared to existing methods. However, the full-buffer traffic model may not accurately represent real-world traffic conditions.

Table 1: Summary of Related works identifying the QoS parameter considered by each scheme

Scheme, Year & Reference	QoS Parameters Considered
QoS-Aware Cell Association in 5G Heterogeneous Networks with Massive MIMO - Wang *et al.,* (2015)	Data Rate, Latency & Resource Allocation
QoS-aware resource management for converged Fiber Wireless 5G Fronthaul networks - Datsika *et al.,* (2018)	Delay & Throughput
QoS-Aware Resource Allocation of Two-tier HetNet: A Q-learning Approach - AlSobhi and Aghvami (2019)	Signal to Interference and Noise Ratio (SINR), Shannon Capacity of MUE and FUE & System capacity thresholds for both MUE and FUE
Quality-focused resource allocation for resilient 5G network - Bruzgiene, Narbutaite, and Adomkus (2019)	BER, data loss, and latency
Joint Task Offloading and QoS-aware Resource Allocation in Fog-enabled Internet of Things Networks - Zhang *et al.,* (2020)	Transmission rate & delay
QoS-aware virtualization resource management mechanism in 5G backhaul heterogeneous networks - Huang *et al.,* (2020)	Resource Utilization, Wavelength Tuning Overhead, Load Balancing, Delay & Energy Consumption
QoS-Aware Resource Allocation and Femtocell Selection for 5G Heterogeneous Networks - Beshley *et al.* (2021)	Resource Utilization Ratio, Dropped Request Probability, Total Average Throughput & Fairness Index
QoS-Aware Resource Allocation for Mobile Edge Networks - Bouaziz, *et al.,* (2021)	Weighted Sum-Rate (WSR), Interference Minimization & Power Allocation
Scalable and QoS-aware Resource Allocation to Heterogeneous Traffic Flows in 5G - Boujelben (2021)	Weighted Sum-Rate (WSR), Interference Minimization, Power Allocation
QoS-Aware Optimal Radio Resource Allocation Method for Machine-Type Communications in 5G LTE and beyond Cellular Networks - Niu Cao & Pun (2021)	Priority Traffic Aggregation, Adaptive Channel bandwidth, Latency and Throughput.

QoS-Aware Joint Component Carrier Selection and Resource Allocation for Carrier Aggregation in 5G - Bolla *et al.*, (2022)	Latency, Throughput, Packet Loss Ratio & Minimum/Maximum Rate
Multi-site Resource Allocation in a QoS-Aware 5G Infrastructure - Joda *et al.*, (2022)	Delay, Jitter, Packet loss & Throughput

Conclusion

This paper looked into the Qos-aware resource allocation schemes for 5G networks. Looking at the reviewed schemes, it can be concluded that, the QoS-aware schemes provides several benefits to resource management in wireless networks. These include improved user experience such as reduced latency in video chats and steady speeds in data-intensive apps, by guaranteeing that each service has the resources it needs. Also, by allocating resources in accordance with current requirements, the network's efficiency is increased and waste is decreased. Finally, 5G networks can accommodate a variety of services, from mission-critical apps to light web browsing, and all of them will obtain the required level of service quality thanks to QoS-aware schemes

Acknowledgement

This work is supported by Tertiary Education Trust Fund (TetFund) through The Institution-Based Research (IBR) Grant of Usmanu Danfodiyo University, Sokoto (IBR-UDUS-B8-013)

References

AlSobhi, W., & Aghvami, A. H. (2019). Qos-aware resource allocation of two-tier HetNet: A Q-learning approach. 2019 26th International Conference on Telecommunications (ICT). doi:10.1109/ict.2019.8798829

Beshley, H., Beshley, M., Medvetskyi, M., & Pyrih, J. (2021). 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). doi:10.1155/2021/9966366

Bolla, R., Bruschi, R., Davoli, F., Lombardo, C., & Pajo, J. F. (2022). Multi-site resource allocation in a qos-aware 5G infrastructure. IEEE Transactions on Network and Service Management, 19(3), 2034-2047. doi:10.1109/tnsm.2022.3151468

Bouaziz, A., Saddoud, A., Chaari, L., & Chaouchi, H. (2021). Qos-aware resource allocation and Femtocell selection for 5G heterogeneous networks. doi:10.21203/rs.3.rs-471274/v1

Boujelben, Y. (2021). Scalable and qos-aware resource allocation to heterogeneous traffic flows in 5G. IEEE Internet of Things Journal, 8(20), 15568-15581. doi:10.1109/jiot.2021.3074111

Bruzgiene, R., Narbutaite, L., & Adomkus, T. (2019). Quality-focused resource allocation for resilient 5G network. Photonic Network Communications, 37(3), 361-375. doi:10.1007/s11107-018-00820-0

Datsika, E., Kartsakli, E., Vardakas, J. S., Antonopoulos, A., Kalfas, G., Maniotis, P. Verikoukis, C. (2018). Qos-aware resource management for converged fiber wireless 5G Fronthaul networks. 2018 IEEE Global Communications Conference (GLOBECOM). doi:10.1109/glocom.2018.8647231

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

Huang, X., Cui, Y., Chen, Q., & Zhang, J. (2020). Joint task offloading and qos-aware resource allocation in fog-enabled internet-of-Things networks. IEEE Internet of Things Journal, 7(8), 7194-7206. https://doi.org/10.1109/jiot.2020.2982670

Joda, R., Elsayed, M., Abou-Zeid, H., Atawia, R., Sediq, A. B., Boudreau, G., & Erol-Kantarci, M. (2022). Qos-aware joint component carrier selection and resource allocation for carrier aggregation in 5G. ICC 2021 - IEEE International Conference on Communications. doi:10.1109/icc42927.2021.9500923

Kamal, M.A., Raza, H. W., Alam, M. M., Su'ud, M. M., Sajak, A. A (2021). Resource Allocation Schemes for 5G Network: A systematic review. *Sensors, Vol. 2021 (21), pp. 1-45*

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

Mamman M., Zurina M.H., Azizol, A., and Abdullah, M. (2018). Call Admission Control for Real-Time and Non-real-time traffic for Vehicular LTE downlink networks. *Lecture Notes in Electrical Engineering. (Mobile and Wireless Technologies 2017)* 46-53

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

Mohammed. A., Olena, O.S., Andriy, O.S., Pavel. I.K., & Volodymyr Y., K. (2018). A Hybrid Approach to Call Admission Control in 5G Networks. *Advances in Fuzzy Systems, Vol. 2018, pp. 1-7.*

Mohammed, H.A., Anabai, H.K., Jeong, R., & Jin. U.K. (2019). Energy Efficiency and Coverage Trade-Off in 5G for Eco-Friendly and Sustainable Cellular Networks. *Symmetry, Vol. 11(3), pp. 1-21*

Niu, G., Cao, Q., & Pun, M. (2021). Qos-aware resource allocation for mobile edge networks: User association, Precoding and power allocation. IEEE Transactions on Vehicular Technology, 70(12), 12617-12630. https://doi.org/10.1109/tvt.2021.3076353

Wang, P., Song, W., Niyato, D., & Xiao, Y. (2015). Qos-aware cell association in 5G heterogeneous networks with massive MIMO. IEEE Network, 29 (6), 76-82. doi:10.1109/mnet.2015.7340428

Wortmann, F., and Fluchter, K., (2015). Internet of Things: Technology and Value Added. *Business & Information System Engineering, Vol. 5, pp. 221-225.*

Zhang, S., Wang, Y., and Zhou, W., (2018). Towards Secure 5G Networks: A survey. *Computer Networks, Vol. 162, pp. 1-37.*

Zhang, H., Huang, C., Zhou, J., & Chen, L. (2020). Qos-aware virtualization resource management mechanism in 5G Backhaul heterogeneous networks. IEEE Access, 8, 19479-19489. doi:10.1109/access.2020.2967101