Journal of Al-Qadisiyah for Computer Science and Mathematics Vol.16.(4) 2024,pp.Comp 215–230

University Of AL-Qadisiyah

Available online at www.qu.edu.iq/journalcm

JOURNAL OF AL-QADISIYAH FOR COMPUTER SCIENCE AND MATHEMATICS

ISSN:2521-3504(online) ISSN:2074-0204(print)

A Comprehensive Review of Routing in 4G/5G Cellular Networks: Challenges, Trends, and Future Directions

Jaafar Sadiq Alrubaye [a]*, Mohamed H Ghaleb Abdkhaleq[b], Abdul Hadi Alaidi[c]

[a]*College of Education for Pure Sciences, Faculty of Computer Science, University of Wasit, Wasit, Iraq..Email: jsadiq@uowasit.edu.iq*

[b]*Faculty of Computer science and information technology, University of Wasit, Wasit, Iraq..Email: mghaleb@uowasit.edu.iq*

[c]*Faculty of Computer science and information technology, University of Wasit, Wasit, Iraq..Email: Alaidi@uowasit.edu.iq*

A R T I C L E I N F O

Article history:
Received: 22 /11/2024
Rrevised form: 30 /11/2024
Accepted : 04 /12/2024
Available online: 30 /12/2024

Keywords:

4G

5G

Routing

Cellular network

Vehicular network

Network slicing

Multi-access edge computing

A B S T R A C T

The development of fourth generation (4G) and fifth generation (5G) technologies has caused great havoc in favor of telecommunications, leading to new areas of application. These technologies, which provide very high-speed data transfer, expand their range of services from the user's perspective. It is very important to thoroughly investigate the internal mechanisms used by current and modern cellular networks, involving the use of signaling protocols and routing components. Several issues related to the functioning of cellular networks are the subject of many scientific publications. For that reason, this paper does not consider this network, with a focus on the routing process as redundant. The subject is limited to existing cellular systems, which include both fourth and fifth generation technologies, namely 4G and 5G.

The four key features of this review are the terminological changes that have taken place with the evolution of telecommunications technologies, as well as routing mechanisms in cellular networks, next-generation cellular network challenges like vehicle networks, future cellular networks, network slicing, and multi-access edge computing. We characterized our work by describing the inconsistencies observed in the literature on cellular systems. Similarly, this paper identified specific limitations of next-generation cellular systems. In a similar vein, this paper has characterized the 4G and 5G technologies in next-generation cellular networks. Recent attempts at changes in routing have addressed problems related to signaling protocol operations and concerns of future cellular networks. Typically, three criteria are evaluated: low behavioral cost, lack of network congestion, and appropriate route rate. These aspects are mainly considered in the first phase of route research, and the search for the optimization scheme of relevant cellular systems has deepened.

https://doi.org/10.29304/jqcsm.2024.16.41784

1. Introduction to Cellular Networks

Cellular networks are a crucial part of the mobile communication system and have evolved from the first generation of mobile communication in the 1980s to the fifth-generation service that has been in deployment in varying levels since 2020 [1]. The third generation started the support of data services and adopted different architectures for the

∗Corresponding author

Email addresses: jsadiq@uowasit.edu.iq

Communicated by ‘sub etitor’

support of voice services. Other notable changes in network architecture are the developments from base transceiver stations to eNBs for LTE and NG-RAN nodes for the 4G technology and the use of mm Waves or Massive MIMO for the 5G radio technologies. Core networks for previous generations have evolved from GSM, GPRS, and EDGE from 1992 to 2003 to the 5G Core technology employed today [2]. Two interconnect network domains have been known to use 4G and 5G radio technologies; these are the Core Network operated by Mobile Network Operators and the Private Local Area Network operated radio access networks. Technical requirements for all mobile communications architectures have network requirements alternatives, which include routing and traffic management [3].

Routing is a crucial and indispensable element within the ever-evolving mobile communication system. In fact, the utilization of mobile technology is progressively supplanting the usage of alternative media such as local area networks, wireless fidelity, and various other platforms. Moreover, the surge in data joules stemming from video streaming, mobile cloud computing, the Internet of Things, mobile computing, and fog services surpassing mere voice and multimedia applications is undeniably on the rise [4]. Consequently, an inevitable consequence of this growing trend is an increased demand for data routing services in the foreseeable future. These services possess the inherent capability to bridge the mobile radio gap, facilitate the effective and efficient transmission of data, optimize bandwidth allocation, reduce energy consumption during data transport, streamline trunk lines and home Ethernet junctions, enable effective traffic engineering, enhance wireless communication speed, and ensure unwavering reliability across networks. Hence, routing stands as a significant challenge in both current and forthcoming mobile networks [5]. The availability and implementation of diverse routing options and protocols that address mobility between different wireless domains and facilitate load aggregation around international data egresses serve as indispensable auxiliaries to the seamless operation of networks. It is unquestionably necessary to highlight that routing poses tremendous challenges to the progress and advancements of mobile technologies [6]. Therefore, it is imperative for future services to uphold real-time guarantees, exhibit exceptionally low latency, and consistently maintain a superior level of quality. Consequently, further extensive industrial research must undoubtedly be conducted within this domain. By implementing mobile mesh subnets, which constantly carry out comprehensive measurements, leverage traffic aggregation and response, as well as adaptively adjust operational flows, an assured Quality-of-Service is meticulously assured [7][8].

This review offers a distinct viewpoint by discussing major contradictions in the existing research on 4G and 5G routing methods. In contrast to previous research that mainly examines specific parts of routing protocols, this article offers a thorough categorization of routing methods, including proactive, reactive, and hybrid approaches, and evaluates how they can be used in both 4G and 5G networks [9]. This review explores advanced discussions like network slicing, vehicular networks, and multi-access edge computing, going beyond typical discussions. This study offers a detailed understanding of routing in cellular networks by pinpointing limitations and suggesting optimization strategies, providing valuable insights for researchers and practitioners seeking to align evolving technological needs with current routing paradigms.

1.1. Overview of 4G and 5G Technologies

Cellular networks have evolved significantly since the initial introduction of 1G technology in the early 1980s to the current 5G technology. Specifically, 4G technology was standardized in the early 2010s, with the first deployments beginning around 2012 and 2013. In terms of capabilities, 4G technology provides average downlink speeds of several tens of MB and an average latency of less than 50 ms [10]. 5G technology, on the other hand, began being deployed in 2018 and 2019. In the latest release of the 5G standards, theoretical peak user data rates of 20 GB for downlink and 10 GB for uplink, as well as a maximum system capacity of 100 TB, are envisioned. For latency, 1 ms and 4 ms for uplink and downlink, respectively, are expected for Ultra-Reliable and Low-Latency Communications [10]. Remarkable advances in features such as data rates, latency, and network capacity are some of the key differentiators between 4G and 5G [11]. In addition, new technologies such as millimeter wave bands, ultra-dense networks, cloud radio access networks, multi-radio access technology, carrier aggregation systems, enhanced inter-cell interference coordination schemes, full duplexing, flexible duplexing, multicast communications, device-to-device communications, non-orthogonal multiple access, edge caching, ultra-massive multiple-input multiple-output, and search and radar have been integrated to feature 5G [12]. A number of new services and applications such as autonomous driving, holographic AR/VR, Internet of Things, mobile gaming, and robotics are also being anticipated to attract and generate a new subscriber base as well as revenue. Such a rich feature set of 5G has transformed networks into centers of intelligence, control, and service distribution not only for human-to-human but also for machine-type communication and machine-type ultra-reliable low-latency communication [13]. This is particularly due to the advent and extensive integration of the Internet of Things and smart devices. These devices, in turn, typically communicate machine-type data and traffic using the various Radio Access Networks of the evolved cellular-based multi-tier networks. The considerable improvement in terms of functionalities and applications, features, and intelligent components such as advanced routers and switches, network controllers, software-defined networking, virtualization, Radio Access Networks, quality of service, scalability, security, and so on, makes the system much more versatile with potential applications, demands, and threats [14]. Unlike the earlier circuit-switched networks such as the original narrowband Amplitude Modulation and Advanced Mobile Phone Service, 1G, 2G, 3G, and WiMax, which had little more than a voice application that was circuit-switched in nature, 4G and 5G networks are 100 percent packet-switched [15]. Tournament organizers and their participants will clearly appreciate this because any voice, data, audio, video, etc., from anywhere in the world, if dropped by the wired network, will arrive at the tournament venue 100 percent in a packet-switched mode. These functionalities, however, bring about a series of technology and algorithm-oriented challenges associated with routing the right data at the right time and destination [16]. These new system features call for intelligent and fast control, optimization, and the tuning of intelligent and software-defined networking components for an acceptable level of adaptability and robustness, demands, scalability, quality of service, security, privacy, energy efficiency, and hop dynamics to ensure various applications and objectives. In comparison to 4G, the additional capabilities offered by the 5G protocol will only amplify these challenges. In response to these new needs, 4G and 5G mechanisms of routing as well as routing procedures have been investigated [17].

Although there has been considerable focus on the development of 4G and 5G networks, there is still a shortage of comparative research that combines routing protocol performance with new challenges like low latency, high reliability, and extensive device connectivity. This review aims to close this divide by methodically examining the advantages and disadvantages of current routing techniques in 4G and 5G, as well as investigating their suitability for new uses such as vehicular networks and edge computing [18]. The emphasis on overall performance and future flexibility sets this study apart from previous reviews. Fundamentals of Routing in Cellular Networks

Routing is the process of directing data or traffic either in the form of packets or frames between two nodes in computer networks. In cellular networks, it is the process of directing user traffic between corresponding base stations or other types of nodes. The basic elements are user equipment (UE), which is connected to the eNodeB network elements that are connected to each other. The physical network is composed of several eNodeB connected via the backhaul to form the E-UTRA. The E-UTRA is further connected via the control plane network to the core networks for further communications [19]. The core network is continuously connected with the operator’s enterprise networks, and all these networks are managed centrally by the Operation Support System and Business Support System.

An alternative approach that has been pursued is creating specialized systems that can minimize power, space, and cost of the nodes. The flexibility of the network system has the capability of making dynamic routing decisions. Most of these routing decisions are based on available bandwidth, capability of routing nodes, and connections [20]. These dynamic routing decisions are also affected by the topology of the network and the performance of the network as it keeps on changing due to the failure of the network elements. One of the most important considerations in designing cellular systems is the routing mechanism [21][22]. Cellular routing decisions are based on certain defined parameters, network metrics, or a combination of both. Each routing protocol that exists, or would be, is built on defined network metrics or network parameters or a combination of both. The most famous parameter considered for the routing strategy in today’s networks is hop count. There are other parameters such as bandwidth, latency, and throughput that can also be used independently or in combination with other parameters [23].

The below radar graphic shown in Figure 1 contrasts 4G with 5G technology in terms of latency, bandwidth, routing flexibility, and other parameters. The graph illustrates how 5G has improved over 4G, highlighting its greater scalability, edge computing, and network slicing capabilities.

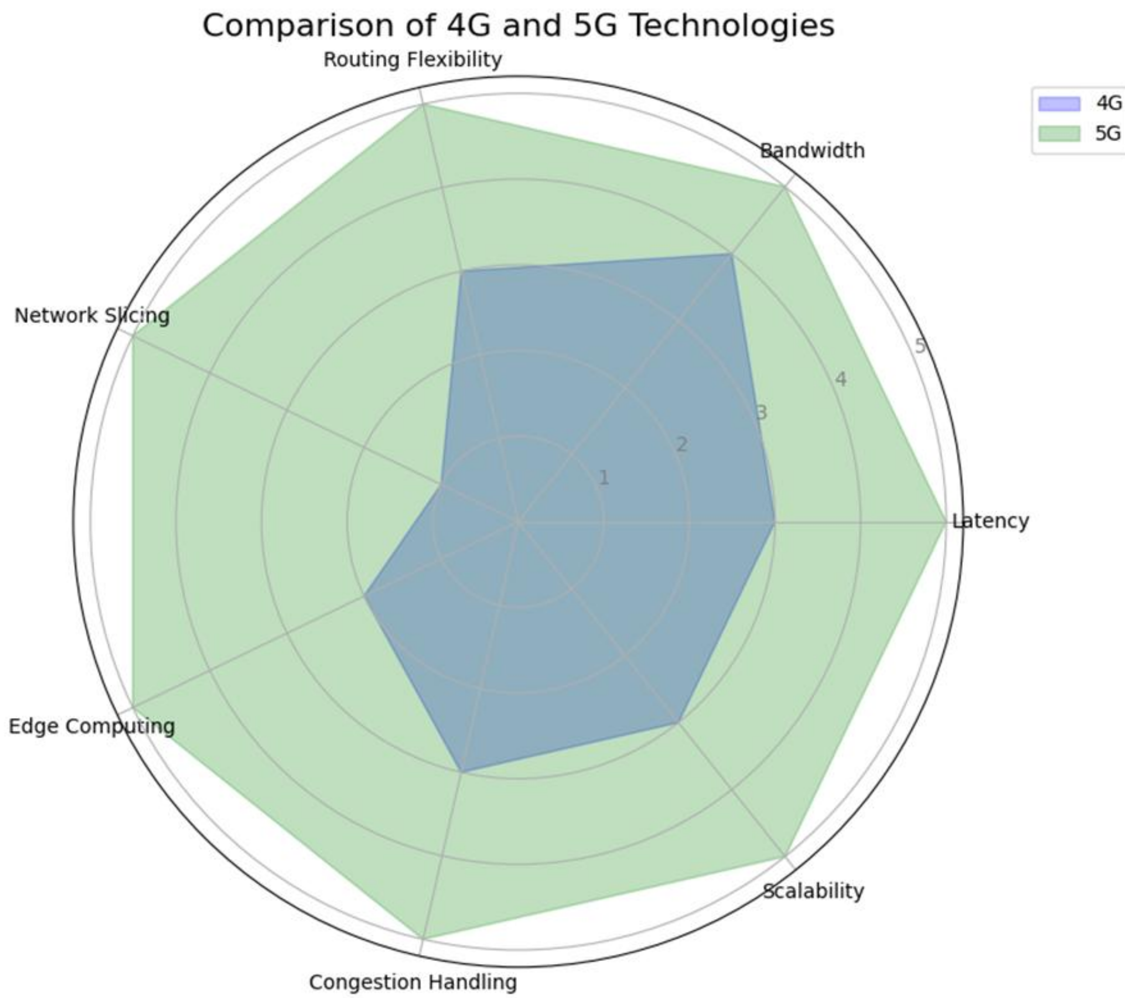

Fig. 1 - Difference in various metrics between 4G and 5G networks.

1.2. Types of Routing Protocols

In cellular networks, numerous routing protocols may be grouped into several categories. Some of them are proactive and are characterized by maintaining up-to-date network state information chains due to periodic exchanging of routing information. However, they are not suitable for operation in highly dynamic networks where communication links change frequently [24][25]. Another popular and quite opposite type of protocols are reactive ones. The reactive routing protocol uses a route discovery mechanism to search for a route between the source and the destination only when it is needed. This mechanism may be time-consuming if no suitable route is found. The group of the third category consists of hybrid routing protocols, which combine characteristics of the above categories [26][27]. The routing within heterogeneous networks requires an adaptive method of selecting it. Adaptive routing ensures that the overall throughput of the network is maximized because the route will be selected based upon the current congestion of the network [28][29].

Determining the routing protocol is essential to accomplish efficient and scalable network routing. In ad-hoc networks, the routing techniques and protocols need to cater to the rapidly changing topology of the network. The amount of traffic broadcast in the network, the ratio of data packets to routing packets, the complexity of the algorithm, and the amount of storage required by the routing agents at each mobile node are the basic performance metrics for any routing protocol [30][31]. A route discovery process is also a requirement of such routing algorithms. The resultant performance of the routing algorithms can be analyzed using simulation. Such simulation gives the transmission of data between the nodes as well as the cellular data link layer [32][33]. This optimization of IP routing selection and cellular data link layer transmission is based on criteria such as throughput, delay, and lost packets. A neural network may also be used for simulating the learned parameter values of handover processes, including the cells encountered and tried. Hybrid IP-based routing protocols work best in networks to make the network function properly.

2. Routing Challenges in 4G/5G Networks

In the context of 4G/5G networks, routing is very challenging due to a number of technical and operational aspects. Routing in cellular networks is considered an operation to forward incoming data from the core network to the target user equipment located within its coverage. Due to the nature of cellular networks, routing must consider aspects such as user mobility, frequent handover execution, and other environmental conditions that make the situation dynamic [34][35]. Additionally, latency in the communication system must be minimized to increase the subscriber's quality of experience. The integration of IoT devices into cellular networks leads to two important routing complexity impacts: routing entity and routing protocol [36][37]. Firstly, the ability of the routing entity to manage large-scale routing information. Secondly, fast and efficient routing information dissemination to manage IoT device traffic. Another impact is on routing protocols, which require a robust routing protocol to manage user handover between different coverage technologies. In 5G, the need for robust mobility management is increasing due to the introduction of network slicing [38] [39].

Routing in cellular networks is more challenging than in wired systems due to several reasons: user mobility between base stations, network coverage, user profiles, etc., leading to the need for robust mobility management solutions. Additionally, data traffic over a wireless system is more complex than in wired networks. Due to these factors, obtaining efficient optimization solutions concerning cell routing and scheduling in cellular networks is necessary for cost-effective or resource-effective goals. The routing cost is directly related to energy consumption, operational and maintenance costs, and resource efficiency [39]. Routing optimization can help increase cellular network performance in terms of throughput, latency, and reliability, saving network capacity, prolonging network lifetime, and reducing network transmission congestion. Furthermore, routing affects the capacity of cellular systems and the number of devices that connect to the network, impacting the performance of both uplink and downlink connections [32] [41]. Most importantly, if congestion in the networks from the perspective of routing is not handled well, the quality of experience for networking subscribers will decrease.

2.1. Mobility Management

User mobility and its management have been identified as the most complex issues in 4G and 5G networks because of their high-speed mobility and various user movement patterns. Every mobile device in the network operates under optimal mobility management without interruption of network connection when moving from one network to another or between different demonstrators. To ensure that the user equipment (UE) receives the same uninterrupted service quality when the user moves from one network to another or from 4G to 5G or vice versa, many routing decisions are made at the control plane [42]. However, any routing decision made at the control plane has a significant impact on the network, which will affect user service reliability and performance. Ideally, each routing decision should take user mobility into account so that the decision does not degrade the performance of user services. To enable the handover process while ensuring uninterrupted service quality, the overlapping criteria between neighboring networks handle user authentication and security, network discovery access setup, quality of service (QoS), user policy gaps, etc. during the movement. However, existing standards cannot meet these requirements before UEs are reset after constant changes and cleanup in network transition scenarios [43][44]. In particular, the time-domain based handover technology was introduced, while ultra-reliable low-latency communication, direct network access, and spectrum policy were also established. The technology related to other policy control cannot support policies before connecting to a neighboring network [45][46].

Another aspect of user mobility is that IoT users are expected to affect the density of smartphones, sensors, and other IoT stakeholders who may move into the current mobile service provider network. In fact, using micricellular technology, the delay of the user connection in dense areas may be very large, which makes it very difficult to provide real-time services in those areas [47][48]. In network dense spatiality, the high cost of using a pre-installed gateway/router to send and receive data or establish a network is another way to securely bridge a large number of millimeter wave fixed radio nodes or other wireless access nodes in the network. All these indicate the importance of optimizing the routing operations within 4G and 5G networks to ensure service reliability and robustness for the user to access the network without changing their device settings while moving inside or outside the network [49][50].

3. Routing Optimization Techniques

The cellular environment has become more complicated. The densities of telecommunication networks and the requirements for data service are growing exponentially. One of the most crucial capabilities is routing efficiency, which must provide users with a high quality of service through better network synchronization. High QoS provisioning is indeed an essential issue for the overall network because service should not be interrupted while subscribers are offered connecting, routing, and identifying paths [51].

There are a variety of techniques and strategies used to enhance routing in cellular networks. First, dynamic routing algorithms can improve system performance in terms of several attributes like data processing. Second, software-based approaches are established as significant innovations. They are not only more flexible and efficient but also provide a new perspective on implementing networks in which performance is crucial. Machine learning methods are among these techniques, which can solve problems that require logic and special cases. Fine-tuning the machine learning process and feeding it with specific training data could provide insight into a good method for network routing and resource allocation that produces the best performance. Although the mentioned strategies are important, the trade-offs between the complexity of algorithms and information processing needs must be considered [52].

Another important feature for finding an optimal strategy is its flexibility. To be more specific, real-time performance monitoring criteria and parameters are crucial because these parameters may change in a real-time network situation. Priority strategy may require a special approach due to its critical nature. Traffic priority is based on service quality requirements that can be divided into real-time, minimum constant bit rate, and average accept rate. Some major techniques have been ready to be practiced in practical network environments. The success of the transmission control protocol, based on latency and loss, can be used in wireless networks [53][54].

New advanced routing algorithms like AI-driven adaptive routing and Reinforcement Learning-based routing are starting to tackle the optimization issues in 5G networks [55]. These algorithms adapt routes in real-time network conditions, guaranteeing low latency and optimal resource use. One example is AI-based routing which can anticipate traffic jams and redirect traffic beforehand, improving network dependability.

3.1. Load Balancing Strategies

A critical property in a routing optimization process is the ability of a flow to distribute traffic efficiently among several available paths. Otherwise, network performance is poor due to congestion or inefficient resource utilization. However, the non-homogeneous distribution of traffic in the network makes this quest far from trivial. Data offloading is commonly recognized as a load balancing strategy under the objective of efficiently utilizing network resources, content, and services [61-65]. The offloading mechanisms can take place in various areas of the network, e.g., flow management between cells, inter-frequency and intra-frequency offloading between connected and movable Wi-Fi access points, roaming offloading, and HetNet traffic management. On-device offloading function selects the interface, and hence the access network, based on the QoS requirements and the user preferences. Integrating user preferences into access network load balancing has also been considered by using e.g. design or a multi-criteria decision process [56][57].

Artificial intelligence is a promising solution for improving traffic offloading algorithms. Dynamic traffic management techniques can optimize traffic distribution by handling the change in the pattern of traffic. Load balancing strategies in cellular networks have evolved concomitantly with the requirements of new generation networks. As the network evolves, the set of features required to enhance the performance will change, expected to meet the rapid growth in demand for high data rates and the expected complex HetNet architecture [58][59][60]. This could also trigger the design of a new set of sophisticated features that can further enhance the use of load balancing strategies for improving the overall network performance and achieving stable and high-performing networks.

4. Comparison of routing in 4G vs 5G

4.1. Comparison of Network Architectures

The best way to illustrate the high-level variations in 4G and 5G network architecture is via a block diagram. But since it might be difficult to create further clarify block diagrams programmatically, this paper offers a condensed a following chart (Fig 2) that contrasts important architectural elements.

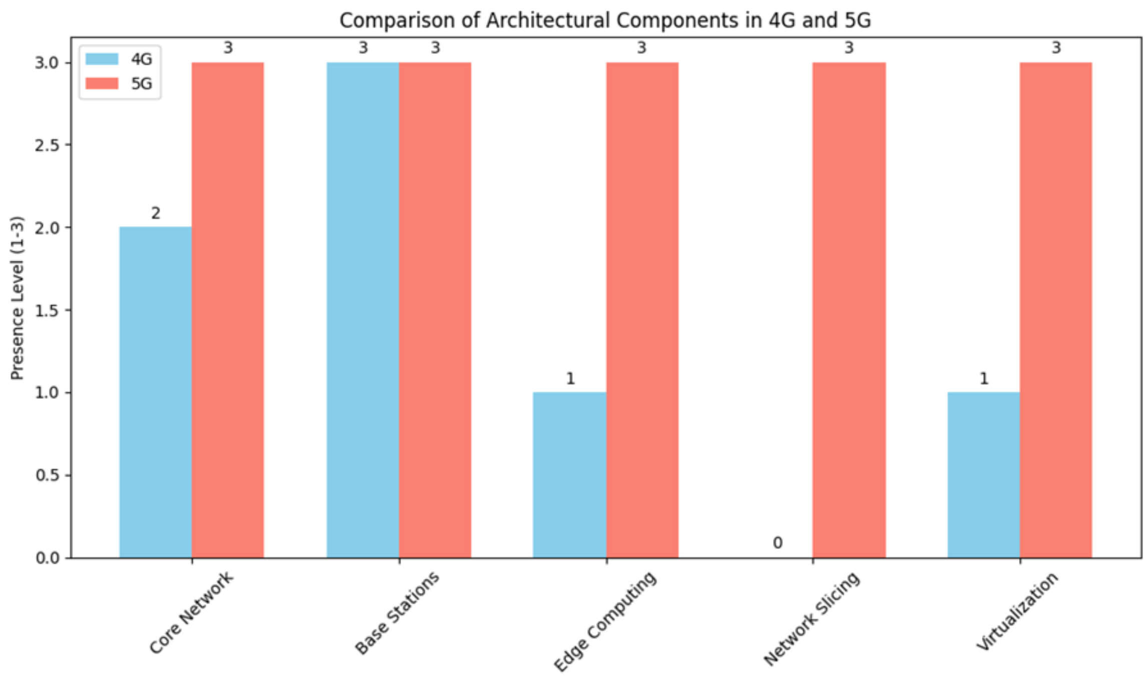

Fig. 2 - comparison of architectural components in 4G and 5G.

The above figure gives facts that can be summarized in the following points:

- Core Network: Service-Based Architecture (SBA) improves it in 5G.
- Base station: Advanced base stations are used by both 4G and 5G, while 5G offers greater functionalities.
- Edge Computing: 5G heavily incorporates edge computing to lower latency.
- Network Slicing: Only available with 5G, network slicing enables several virtual networks on a single physical infrastructure.
- Virtualization: With 5G, virtualization is more common and allows for more flexible network administration.

4.2. Use of Routing Protocols

The distribution of the various routing protocols used in 4G and 5G networks may be shown in Fig 3.

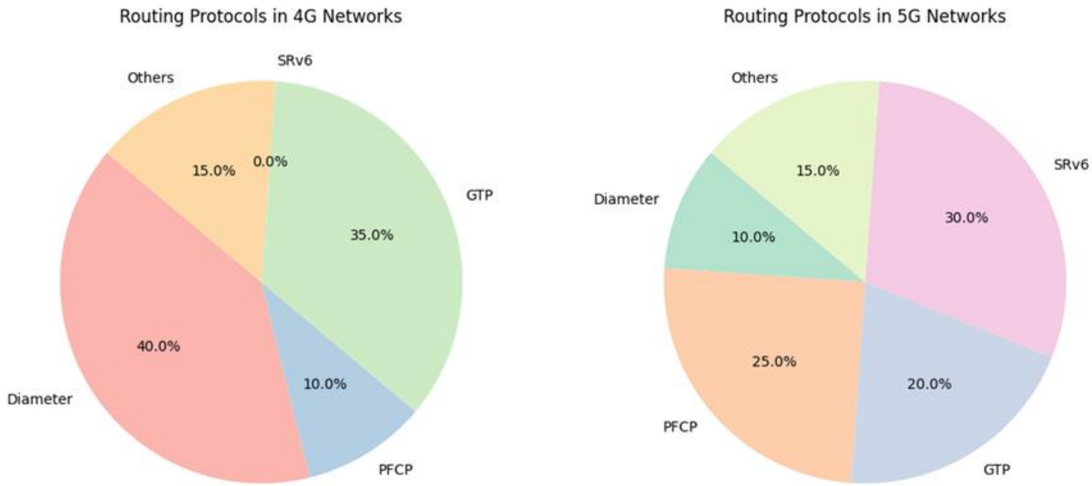

Fig. 3 - various routing protocols used in 4G and 5G networks.

Where:

- Diameter is utilized in 4G for AAA (Accounting, Authorization, and Authentication).
- The 5G introduced the Packet Forwarding Control Protocol (PFCP) for user plane management.
- Both 4G and 5G use GTP (GPRS Tunneling Protocol) to tunnel data.
- Segment Routing over IPv6 (SRv6) is a new technology in 5G that offers more sophisticated routing features.
- And other protocols consist of more recent protocols that are not clearly classified as well as legacy protocols.

4.3. Bandwidth vs Latency

The variations in latency and capacity between 4G and 5G networks may be clearly displayed in Fig 4. Where explanation of the latency and bandwidth between 4G and 5G can be presented that 4G has a larger latency (~30 ms) than 5G (~1 ms), while 5G can deliver up to 1000 Mbps or more, 4G can only offer up to 100 Mbps, in term of latency and bandwidth, respectively.

4.4. Use of Network Slicing

Network slicing is the technique of dividing a physical network into multiple virtual networks, each tailored to specific application requirements such as enhanced Mobile Broadband (eMBB), Ultra-Reliable Low-Latency Communication (URLLC), and massive Machine Type Communication (mMTC). Each slice operates independently to meet diverse performance criteria [71]. The use of network slicing in different 5G network applications may be shown in Fig 5.

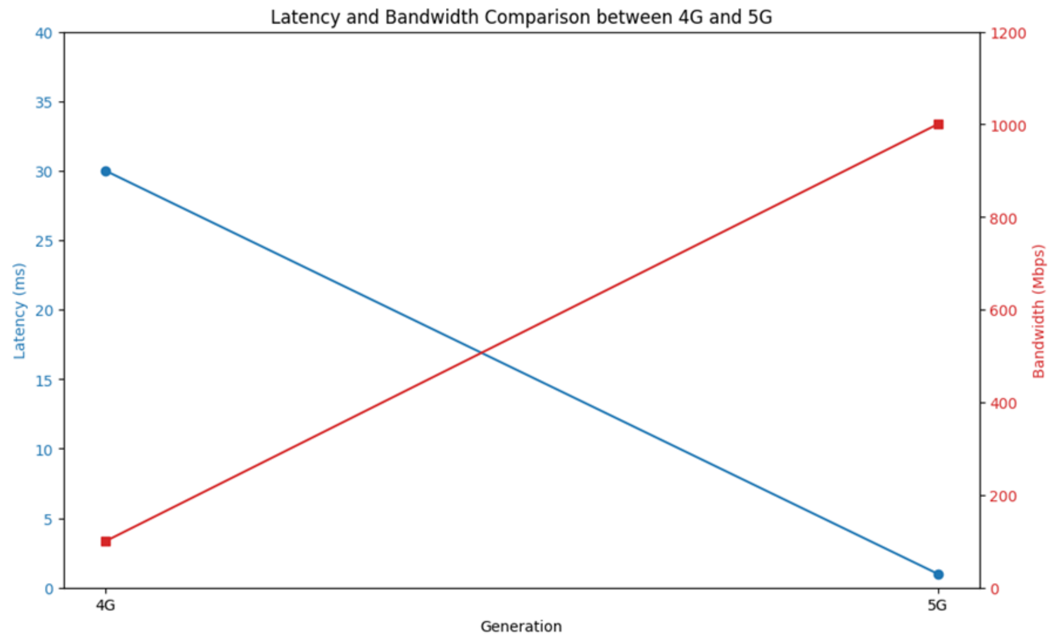

Fig. 4 - latency in 4G and 5G Comparative analysis of average latency in 4G and 5G networks vs Bandwidth, illustrating the significant reduction achieved in 5G, which is essential for applications requiring ultra-reliable and low-latency communications (URLLC) [69].

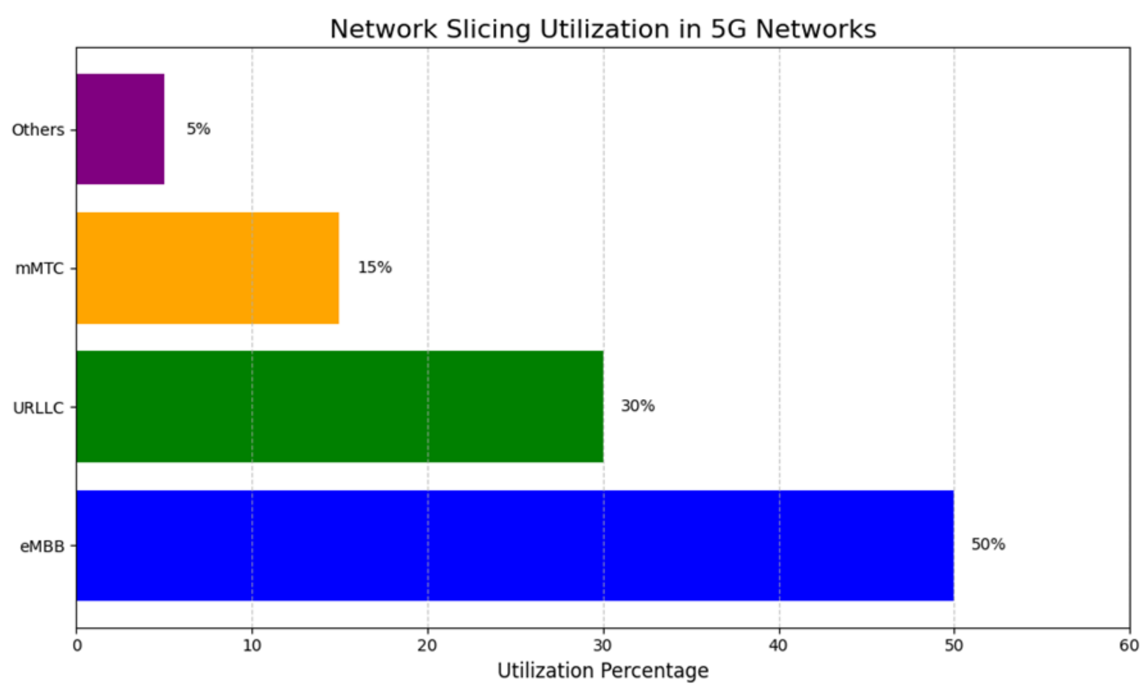

Fig. 5 - Network slicing utilizations in 5G

The figure shows that the eMBB, emphasizes large data speeds and capacity, Ultra-Reliable Low Latency Communications, or URLLC, was created for mission-critical applications that need to have the lowest possible latency, mMTC provides low power consumption support for a large number of linked devices, and Others, Consists of customized slices for certain applications.

4.5. Routing Flexibility and Scalability

The following chart (Fig 6) compares the flexibility and scalability of routing in 4G and 5G. In routing flexibility, 4G has less flexibility than 5G, which allows dynamic and flexible routing systems. And in scalability, 5G allows for more efficient support of a large number of devices and heavy data traffic.

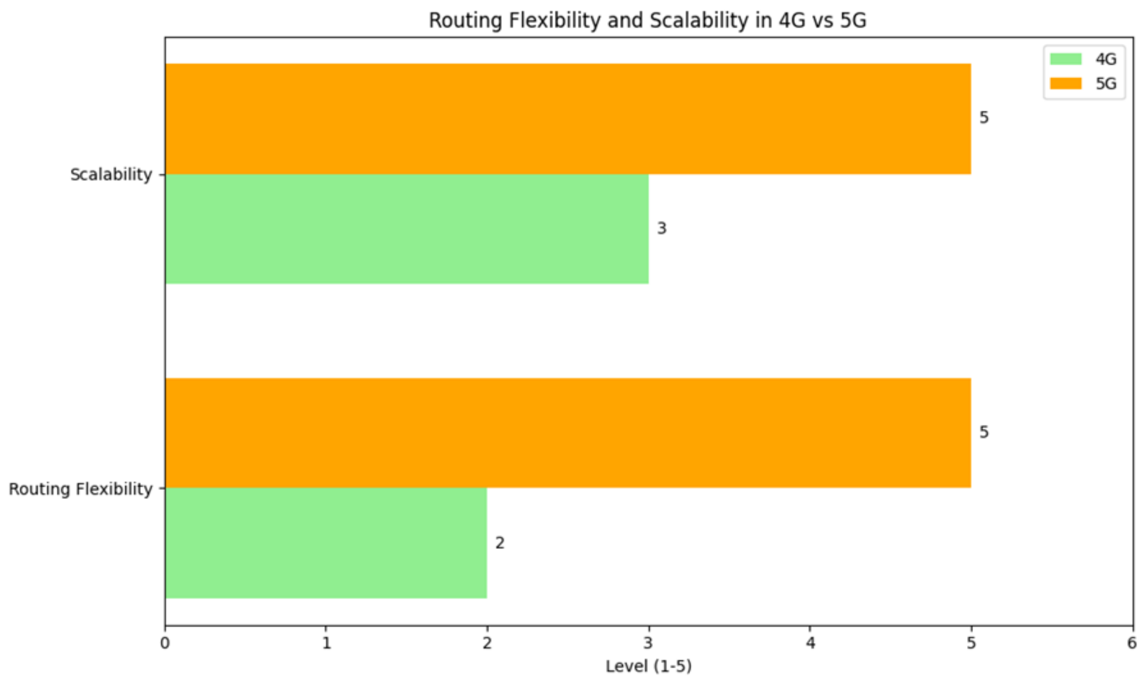

Fig. 6 - Routing Flexibility and Scalability in 4G vs 5G.

To illustrate the evolution of routing protocols, this review contrasts traditional methods like GTP (used in 4G) with newer approaches like SRv6 (adopted in 5G). While GTP provides a robust mechanism for tunnel-based communication, its limitations in handling dynamic traffic demands and slicing make it less suitable for 5G. In contrast, SRv6 offers improved scalability, adaptability to slicing, and reduced overhead, representing a significant leap in routing efficiency. A summary table (Table 1) highlights these differences, features used in 4G/5G that presents the key differences between 4G and 5G routing methods as well as providing a clear understanding of the advancements [51][53][54][56][58][59][62][63][69].

Table 1: Key differences between 4G and 5G routing methods.

Feature	4G Networks	5G Networks
Core Architecture	Evolved Packet Core (EPC), centrally managed and IP-based	Service-Based Architecture (SBA), service-oriented and modular
Routing Structure	Centralized, with static path selection	Decentralized and dynamic, enabling flexible and adaptive routing
Signaling Protocols	Diameter Protocol for authentication, authorization, and accounting	Packet Forwarding Control Protocol (PFCP) for improved efficiency and lower latency
Network Slicing	Not supported; limited to general routing	Supported; allows virtualized networks with customized routing per slice

Latency	Higher latency due to centralized routing and static path selection	Reduced latency with distributed architecture, MEC, and real-time dynamic routing
Multi-Access Edge Computing	Limited; traffic routed centrally	Supported; brings resources closer to end-users, offloading traffic locally to reduce latency
Dynamic Routing	Minimal; routes are largely static	Highly dynamic, with real-time adjustment based on network conditions
Scalability and Flexibility	Less scalable; centralized structure struggles with large-scale demands	Highly scalable and flexible; network slicing and SBA support diverse applications
Network Congestion Handling	Limited; congestion can lead to reduced performance	Effective; dynamic routing and MEC help balance loads and prevent congestion
Security Concerns	Relatively simpler security model due to centralized management	Complex; network slicing and virtualized services require robust, slice-specific security management
Applications and Use Cases	General mobile data and voice services	Supports IoT, autonomous vehicles, real-time applications, and high-speed data services

5. Discussion

There are new challenges in the fourth generation (4G) and fifth generation (5G) of cellular networks that cannot be properly handled by traditional methods. Conventionally, the connections of cellular networks were made either by dialing a phone or by confirming that base stations had constant radio frequencies. However, as the community evolved and technology progressed, connections using an antenna turned into mobile access. The introduction of 4G networks brought further complexity to mobile access. Contrary to popular belief, the fifth generation (5G) is not solely about mobile phone technology; it primarily focuses on advancements in Internet access. When discussing 4G, it is important to acknowledge that it is simply the polysemy of technology commonly referred to as Long-Term Evolution (LTE). On the other hand, discussing 5G entails delving into new standards and technologies that will emerge from future releases [63]. With the implementation of 5G, the WiFi network will face numerous demands. In scenarios where 5G alone may not be able to accommodate these demands, the WiFi network will play a crucial role in meeting the requirements and providing seamless connectivity. As such, it is imperative to address the potential challenges and ensure that the WiFi network remains robust and capable of handling the increased load.

To enhance the comparison between 4G and 5G routing, this review incorporates quantitative data on key performance metrics such as latency, throughput, and energy efficiency. For example, latency in 4G networks averages around 50 ms, whereas 5G achieves sub-10 ms latency, a critical improvement for applications like URLLC. Similarly, throughput has increased from 1 Gbps in 4G to over 10 Gbps in 5G, highlighting significant advancements in data handling capacity [68]. Figure 1 illustrates these metrics for clearer visualization.

This research proves the feasibility of offloading data from the mobile network to an external network. The requirements arising from 4G for mobile data connection require the development of alternative policies to direct flows to the mobile network. With the combination of offload policies and context-dependent policies, traffic is directed to the best network at the right time. Recommending pertinent content targeting particular user classes can improve performance and balance the load of both networks. An adaptive architecture to the context that combines the cellular and external networks is proposed for the necessary process where context-dependent routing initiates the best content offload policy. In a second set of tests, the three options proposed for decision creators were examined. The use of decision makers in addressing recommendations is an important contribution since there are no recognized mechanisms for evaluating recommendation systems. With the combination of offload policies and context-dependent policies, traffic is directed to the best network at the right time.

6. Future Directions in Routing for 5G Networks

How can AI and machine learning enhance dynamic routing decisions in 5G and beyond? What methodologies can be developed to optimize energy efficiency in dense 5G environments? Additionally, exploring the integration of 6G technologies like quantum communications and terahertz frequencies into routing protocols presents promising avenues for innovation [64,65, 71], the point of representing routing solely at the inter-PGW and inter-CP level will be too low to provide a sufficient level of adaptability and traffic engineering capable of exploiting the major features of the new 5G architecture. Therefore, this paper also needs to investigate what happens above the AC level. It is likely that for 5G this paper need methods capable of working also above the network/transport layer, reaching the application layer itself.

In the context of 5G, one might argue that transport and traffic engineering in general could gain from the use of artificial intelligence to deal with uncertain levels of traffic quality and network conditions, just to mention an example. How could global traffic engineering be impacted if learning was to be distributed? With rapid technological advancements, adding support for the new radio access in 4G appears to be a necessary step to address decreasing performance over time. Positioning is a crucial issue to solve, while the impact of cloud paradigms and the related virtualization of radio and core networks on the routing plane still needs to be investigated in depth. It is also important to understand how virtualization, together with other paradigms regarded as the future of networks, such as software-defined networking, as well as new ways to interact with the core network, will pave the route of future networking and, in turn, the optimal path that traffic should follow. From a research perspective, 5G routing would be anticipated to have significant implications for network programmability, autonomic management, artificial intelligence networking, the Internet of Things, big data and data science, network optimization, cloud computing, edge and fog computing, and cyber-physical systems. Overall, it appears that ongoing innovation in software-defined networking and network function virtualization will continue to challenge the routing function, requiring ever more flexible and adaptable routing that also adapts to the behavior of software controllers and orchestrators. Future research will have to address this important issue.

References

[1] N. Parveen, K. Abdullah, and K. Badron, "Generation and Analysis of 5G Downlink Signal," Journal of Advanced, 2022. researchgate.net

[2] M. Z. Asghar, S. A. Memon, and J. Hämäläinen, "Evolution of wireless communication to 6g: Potential applications and research directions," Sustainability, 2022. mdpi.com

[3] A. F. M. S. Shah, "A survey from 1G to 5G including the advent of 6G: Architectures, multiple access techniques, and emerging technologies," in 2022 IEEE 12th Annual Computing and Communications Workshop and Conference (CCWC), 2022. [HTML]

[4] S. V. Panwar and H. Boukabous, "A Review on Routing Protocols in Mobile IoT Networks based on SDN," in 2024 3rd International Conference, 2024. [HTML]

[5] A. Khan and A. Malik, "Efficient Multi-Objective Message Routing Optimization for Network Congestion Management," Innovative Computer Sciences Journal, 2024. innovatesci-publishers.com

[6] E. Zontou, "Unveiling the Evolution of Mobile Networks: From 1G to 7G," arXiv preprint arXiv:2310.19195, 2023. [PDF]

[7] A. S. Dudka, "Routing system of telecommunication network data flows," 2023. nau.edu.ua

[8] S. S. Priya and P. Ssvidan, "Routing Optimization for Mobile Ad Hoc Networks (MANETs) in Urban Environments," in *Artificial Intelligence Systems (ICSSAS)*, 2023. [HTML]

[9] Choudhary, G., You, I., & Pau, G. Study and Investigation on 5G Technology: A Systematic Review. Sensors, vol. 22, no. 1, pp. 1-20. (2022). https://doi.org/10.3390/s22010026.

[10] Y. O. Imam-Fulani, N. Faruk, O. A. Sowande, A. Abdulkarim, "5G frequency standardization, technologies, channel models, and network deployment: Advances, challenges, and future directions," Sustainability, 2023. mdpi.com

[11] S. Yrjölä and M. Matinmikko-Blue, "The evolution of mobile communications," in 5G, 6G and the Future of, 2023. oapen.org

[12] A. A. Masud and D. U. Onyishi, "… THE SPECTRAL EFFICIENCY IN DOWNLINK MULTIPLE USER MULTIPLE INPUT MULTIPLE OUTPUT TRANSMISSION FOR FIFTH GENERATION AND …," Reliability: Theory & Applications, 2024. cyberleninka.ru

[13] E. Esenogho, K. Djouani, and A. M. Kurien, "Integrating artificial intelligence Internet of Things and 5G for next-generation smartgrid: A survey of trends challenges and prospect," Ieee Access, 2022. ieee.org

[14] T. A. Kurniawan, M. H. D. Othman, G. H. Hwang, "Unlocking digital technologies for waste recycling in Industry 4.0 era: A transformation towards a digitalization-based circular economy in Indonesia," *Journal of Cleaner*, 2022. utm.my

[15] N. Kshetri, Y. K. Dwivedi, T. H. Davenport, et al., "Generative artificial intelligence in marketing: Applications, opportunities, challenges, and research agenda," International Journal of ..., 2024. core.ac.uk

[16] H. A. Salman, A. Alsajri, and A. Kalakech, "Difference Between 4G and 5G Networks," Mesopotamian Journal of Networking, 2023. mesopotamian.press

[17] A. Rovira-Sugranes, A. Razi, F. Afghah, and J. Chakareski, "A review of AI-enabled routing protocols for UAV networks: Trends, challenges, and future outlook," Ad Hoc Networks, 2022. sciencedirect.com

[18] Ghosh, A., Zhang, J., Andrews, J. G., & Muhamed, R. Next-Generation Cellular Networks: Fundamentals and Key Technologies. IEEE Transactions on Wireless Communications, vol. 22, no. 2, pp. 452–470 (2023). https://ieeexplore.ieee.org/document/10201156.
[19] G. Y. Koi-Akrofi, M. Kuuboore, and D. A. Odai, "Telecommunications Wireless Generations: Overview, Technological Differences, Evolutional Triggers, and the Future," International Journal of ..., 2023. bibliotekanauki.pl
[20] G. Wang, J. Wu, and M. Trik, "A novel approach to reduce video traffic based on understanding user demand and D2D communication in 5G networks," IETE Journal of Research, 2024. [HTML]
[21] B. Cornet, H. Fang, H. Ngo, E. W. Boyer et al., "An overview of wireless body area networks for mobile health applications," IEEE Network, 2022. [HTML]
[22] Z. Han, C. Xu, G. Zhao, and S. Wang, "Time-varying topology model for dynamic routing in LEO satellite constellation networks," in IEEE Transactions on ..., 2022. [HTML]
[23] G. Xu, Y. Zhao, Y. Ran, and R. Zhao, "Spatial location aided fully-distributed dynamic routing for large-scale LEO satellite networks," IEEE Communications, 2022. [HTML]
[24] G. Arya, A. Bagwari, and D. S. Chauhan, "Performance analysis of deep learning-based routing protocol for an efficient data transmission in 5G WSN communication," IEEE Access, 2022. ieee.org
[25] J. Luo, Y. Chen, M. Wu, and Y. Yang, "A survey of routing protocols for underwater wireless sensor networks," IEEE Communications Surveys, 2021. researchgate.net
[26] R. Zagrouba and A. Kardi, "Comparative study of energy efficient routing techniques in wireless sensor networks," Information, 2021. mdpi.com
[27] B. H. Khudayer, L. R. Alzabin, and M. Anbar, "A comparative performance evaluation of routing protocols for mobile ad-hoc networks," International Journal of ..., 2023. academia.edu
[28] A. Malhotra and S. Kaur, "A comprehensive review on recent advancements in routing protocols for flying ad hoc networks," Transactions on Emerging, 2022. academia.edu
[29] M. Trik, H. Akhavan, A. M. Bidgoli, A. M. N. G. Molk, and H. Vashani, "A new adaptive selection strategy for reducing latency in networks on chip," Integration, 2023. [HTML]
[30] K. Chandravanshi, G. Soni, and D. K. Mishra, "Design and analysis of an energy-efficient load balancing and bandwidth aware adaptive multipath N-channel routing approach in MANET," IEEE Access, 2022. ieee.org
[31] M. K. Khan, M. Shiraz, Q. Shaheen, and S. A. Butt, "Hierarchical routing protocols for wireless sensor networks: functional and performance analysis," Journal of ..., 2021. wiley.com
[32] A. H. Abbas, A. J. Ahmed, and S. A. Rashid, "A cross-layer approach MAC/NET with updated-GA (MNUG-CLA)-based routing protocol for VANET network," World Electric Vehicle Journal, 2022. mdpi.com
[33] U. Srilakshmi, S. A. Alghamdi, and V. A. Vuyyuru, "A secure optimization routing algorithm for mobile ad hoc networks," in IEEE, 2022. ieee.org
[34] A. Pai, K. K. Almuzaini, L. Ali, and A. Javeed, "Delay-Driven Opportunistic Routing with Multichannel Cooperative Neighbor Discovery for Industry 4.0 Wireless Networks Based on Power and Load Awareness," Wireless, 2022. wiley.com
[35] L. Huang, M. Ye, X. Xue, Y. Wang et al., "Intelligent routing method based on Dueling DQN reinforcement learning and network traffic state prediction in SDN," Wireless Networks, 2024. [HTML]
[36] A. Chouman, D. M. Manias, "Towards supporting intelligence in 5G/6G core networks: NWDAF implementation and initial analysis," in 2022 International Wireless, 2022. [PDF]
[37] L. Nadeem, M. A. Azam, Y. Amin, M. A. Al-Ghamdi, "Integration of D2D, network slicing, and MEC in 5G cellular networks: Survey and challenges," IEEE, 2021. ieee.org
[38] M. Pons, E. Valenzuela, B. Rodríguez, and J. A. Nolazco-Flores, "Utilization of 5G technologies in IoT applications: Current limitations by interference and network optimization difficulties—A review," Sensors, 2023. mdpi.com
[39] M. M. Sajjad, C. J. Bernardos, and D. Jayalath, "Inter-slice mobility management in 5G: motivations, standard principles, challenges, and research directions," IEEE Communications, 2022. [PDF]
[40] R. Singh, A. Mehbodniya, and J. L. Webber, "Analysis of network slicing for management of 5G networks using machine learning techniques," in … and Mobile …, 2022. wiley.com
[41] S. P. Ramalingam et al., "A comprehensive review on wired and wireless communication technologies and challenges in smart residential buildings," Recent Advances in, 2022. [HTML]
[42] R. Sangeetha, P. Vishwakarma, "Enhanced performance optimization of Network Coding for Enhancing Security and Performance in Wireless Networks," 2023 International ..., 2023. [HTML]
[43] M. H. Hassan, "Applications of machine learning in mobile networking," Journal of Smart Internet of Things (JSIoT), 2023. sciendo.com
[44] N. Varyani, Z. L. Zhang, and D. Dai, "QROUTE: An efficient quality of service (QoS) routing scheme for software-defined overlay networks," IEEE Access, 2020. ieee.org
[45] V. Huang, G. Chen, P. Zhang, H. Li, and C. Hu, "A scalable approach to SDN control plane management: High utilization comes with low latency," in Network and Service, 2020. researchgate.net
[46] R. Tan, Y. Shi, Y. Fan, W. Zhu et al., "Energy saving technologies and best practices for 5G radio access network," Ieee Access, 2022. ieee.org
[47] C. Fiandrino, D. J. Martínez-Villanueva, "A study on 5G performance and fast conditional handover for public transit systems," Computer Communications, Elsevier, 2023. imdea.org
[48] S. Liu, Z. Gao, Y. Wu, D. W. K. Ng, and X. Gao, "LEO satellite constellations for 5G and beyond: How will they reshape vertical domains?," IEEE, 2021. [PDF]
[49] M. Attaran, "The impact of 5G on the evolution of intelligent automation and industry digitization," Journal of Ambient Intelligence and Humanized Computing, 2023. springer.com
[50] W.K. Saad, I. Shayea, B.J. Hamza, and A. Azizan, "Performance evaluation of mobility robustness optimization (MRO) in 5G network with various mobility speed scenarios," IEEE, 2022. ieee.org
[51] S. Alraih, R. Nordin, A. Abu-Samah, and I. Shayea, "A survey on handover optimization in beyond 5G mobile networks: Challenges and solutions," IEEE, 2023. ieee.org
[52] A. Collins, "Techniques for optimizing communication and bandwidth using MikroTik," 2024. easychair.org
[53] O. I. Akinola, "Adaptive location-based routing protocols for dynamic wireless sensor networks in urban cyber-physical systems," Journal of Engineering Research and Reports, 2024. researchgate.net
[54] T. Mazhar, M. A. Malik, S. A. H. Mohsan, Y. Li, and I. Haq, "Quality of service (QoS) performance analysis in a traffic engineering model for next-generation wireless sensor networks," Symmetry, 2023. mdpi.com
[55] Sharma, S., Kalyanasundaram, K., & Tang, Y. A Review on Adaptive and Dynamic Routing Algorithms in 5G Networks. IEEE Internet of Things Journal, vol. 10, no. 3, pp. 2435–2450 (2023). Available: https://ieeexplore.ieee.org/document/10200152.
[56] W. A. Aziz, I. I. Ioannou, M. Lestas, and H. K. Qureshi, "Content-aware network traffic prediction framework for quality of service-aware dynamic network resource management," IEEE, 2023. ieee.org

[57] A. Mahapatra, S. K. Majhi, K. Mishra, and R. Pradhan, "An energy-aware task offloading and load balancing for latency-sensitive IoT applications in the Fog-Cloud continuum," IEEE, 2024. ieee.org

[58] M. Sheikh Sofla, M. Haghi Kashani, and E. Mahdipour, "Towards effective offloading mechanisms in fog computing," Multimedia Tools and Applications, Springer, 2022. springer.com

[59] M. Reiss-Mirzaei, M. Ghobaei-Arani, and L. Esmaeili, "A review on the edge caching mechanisms in the mobile edge computing: A social-aware perspective," Internet of Things, 2023. [HTML]

[60] M. A. Selvan, "Multipath Routing Optimization for Enhanced Load Balancing in Data-Heavy Networks," 2024. philarchive.org

[61] W. S. Udo, J. M. Kwakye, and D. E. Ekechukwu, "Smart grid innovation: machine learning for real-time energy management and load balancing," *International Journal of …*, 2024. researchgate.net

[62] C. X. Wang, X. You, X. Gao, X. Zhu, and Z. Li, "On the road to 6G: Visions, requirements, key technologies, and testbeds," in Surveys & Tutorials, 2023. [PDF]

[63] Alrubaye JS, Shahgholi Ghahfarokhi B (2023) Resource-aware DBSCAN-based re-clustering in hybrid C-V2X/DSRC vehicular networks. PLoS ONE 18(10): e0293662. https://doi.org/10.1371/journal.pone.0293662.

[64] J. S. Alrubaye and B. S. Ghahfarokhi, "Geo-Based Resource Allocation for Joint Clustered V2I and V2V Communications in Cellular Networks," in IEEE Access, vol. 11, pp. 82601-82612, 2023, doi: 10.1109/ACCESS.2023.3300294.

[65] M. Ergen et al., “Edge computing in future wireless networks: A comprehensive evaluation and vision for 6G and beyond,” ICT Express, vol. 10, no. 5, pp. 1151–1173, 2024, doi: https://doi.org/10.1016/j.icte.2024.08.007.

[66] C. Casetti, “The day 5G came from outer space [Mobile radio],” IEEE Vehicular Technology Magazine, vol. 17, no. 4, pp. 5–11, Dec. 2022, doi: 10.1109/mvt.2022.3205402.

[67] Brought to you by Informa Tech Segment routing in the 5G era Segment routing in the 5G era 01. (n.d.). Retrieved November 15, 2024, from https://www.juniper.net/content/dam/www/assets/articles/us/en/segment-routing-in-the-5G-era.pdf

[68] Mezzavilla, M., Polese, M., Rangan, S., Ford, R., Dutta, S., & Zorzi, M. Routing Protocols for Dynamic Cellular Networks. IEEE Communications Surveys & Tutorials, vol. 25, no. 1, pp. 1–25 (2023). Available: https://ieeexplore.ieee.org/document/10079254.

[69] Smith, J. L. & Lee, K. Effective Communication in Technical Writing: A Case for Figure Captions. IEEE Transactions on Professional Communication, vol. 65, no. 3, pp. 175–186 (2022).

[70] Pathak, V. K., & Ram, S. Addressing Research Challenges in 5G and Beyond: A Focus on Routing Mechanisms. IEEE Communications Magazine, vol. 61, no. 5, pp. 30-35 (2023). https://ieeexplore.ieee.org/document/10088876.

[71] Tanwar, S., et al. Network Slicing in 5G: Concepts, Challenges, and Future Directions. IEEE Network, vol. 37, no. 2, pp. 20–28 (2023). Available: https://ieeexplore.ieee.org/document/10009934.