A survey on fronthaul signaling of user-centric cell-free massive MIMO networks

Syed Tariq Ali, Anamika Singh
Department of Electronic and Communication, LNCT University Bhopal, Madhya Pradesh, India

Article Info

Article history:

Received Feb 15, 2025
Revised Jul 8, 2025
Accepted Oct 7, 2025

Keywords:

5G/6G network
CoMP-JT
Fronthaul signalling
mMIMO
Software defined networks
Ultra-dense network
User-centric cell-free network

ABSTRACT

The mandate for high data rates in mobile communication is increasing and will continue to do so in the future. Although the latest network technologies can meet this demand, they result in more-dense networks. Networks like ultra-dense networks and massive multiple-input multiple-output provide very high data rates, but they cannot meet the future demand. The main issue with existing networks is inter-cell interference and variations in quality of service esp. at the cell edges, leading to research on new network architectures that offer intelligent coordination and collaboration capabilities are being researched, like user-centric cell-free (UC-CF) massive-multiple-input-multiple-output (mMIMO). This network combines the best of ultra-dense networks and mMIMO and eliminates cell edge problems. It is served by access points that cooperate and coordinate with each other. This paper reviews the challenges and opportunities in physical layer parameter-fronthaul signaling for UC-CF mMIMO. We discuss the basics of the network, the importance of fronthaul signaling, and propose various approaches in the literature to address challenges and identify research gaps and provide future directions. Our aims to provide a comprehensive overview of the current state of fronthaul signaling and highlight the key issues that need to be addressed to realize its full potential.

Corresponding Author:

Syed Tariq Ali
Department of Electronic and Communication, LNCT University Bhopal
Madhya Pradesh, India
Email: tariq.ali1990@gmail.com

1. INTRODUCTION

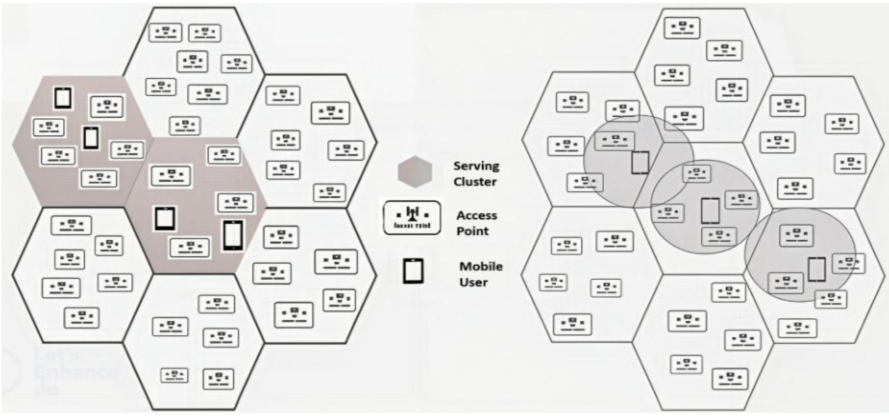
The evolution of wireless communication networks has seen a shift towards user-centric architectures to address the limitations inherent in traditional cellular networks. One such paradigm is the user-centric cell-free (UC-CF) network, characterized by the clustering of distributed access points or antennas around each user. By removing fixed cell boundaries, this architecture enables uninterrupted connectivity for users throughout the coverage area [1], [2]. In the pursuit of practical CF communication, UC-CF networks have emerged as a promising solution, particularly in mitigating inter-cell interference, a significant drawback of traditional cellular architectures.

The integration of UC-CF networks with advanced multiple-input-multiple-output (MIMO) concepts, particularly massive-MIMO (mMIMO), further enhances their capabilities and deployment feasibility [3]-[8]. MIMO technology, which enables the simultaneous transmission of multiple signals through beamforming, revolutionizes wireless communication by optimizing signal transmission across diverse channels [9]. MMIMO enhances traditional MIMO systems by utilizing a significantly larger antenna array at the transmitter, which leads to improved spectral efficiency and higher data throughput [10].

In conventional cellular networks utilizing mMIMO, users at cell edges often experience suboptimal performance due to uneven coverage. To overcome these limitations, CF mMIMO distributes numerous antennas over a wide area, providing uniform coverage and enhancing spectral efficiency [11], [12]. In the context of UC-CF networks, each user dynamically establishes a serving cluster of antennas based on various criteria, thereby optimizing network performance [13]-[15].

The emergence of UC-CF mMIMO networks represents a significant advancement in mobile communication technology, with implications for future generations of wireless networks. The CF system has a lot in common with already existing technologies such as “network-MIMO [16], coordinated multi-point with joint transmission (CoMP-JT) [17], cloud radio access network (CRAN) [18], [19], multi-cell MIMO cooperative network [20], small-cell networks, remote radio heads (RRHs) in distributed antenna systems, and C-RAN [18], [19]”, with slight changes that aim to eliminate the cell-edge problems and minimize fronthaul overhead. For instance, in network-MIMO systems [16], multiple access points coordinate by sharing user data and channel state information (CSI) to perform joint signal processing. In contrast, in a CF mMIMO system, the transmitter acquires and processes CSI locally, resulting in greater scalability than network MIMO, owing to CSI’s lower data exchange rate. Similarly, CRAN shares the same network architecture as the CF mMIMO system but differs in the clustering of users. The primary limitation of CRAN lies in its fully centralized baseband processing, which demands substantial fronthaul bandwidth [18].

In a CF mMIMO network, single or multiple control units (CUs) can control an access point or antenna with each CU potentially differing in terms of centralization, processing capacity, and operational load. Figure 1 illustrates network architecture of both a cell-centric in Figure 1(a) and a UC scheme in Figure 1(b) for a mMIMO system. As stated earlier, the cell-centric approach has a high delay in signal spread when compared to a UC system [13], does not provide considerable gain, and are unable to effectively serve users located near cell boundaries, where signal strength tends to be low [21].



(a) (b)

Figure 1. Network architecture of (a) cell-centric and (b) UC network

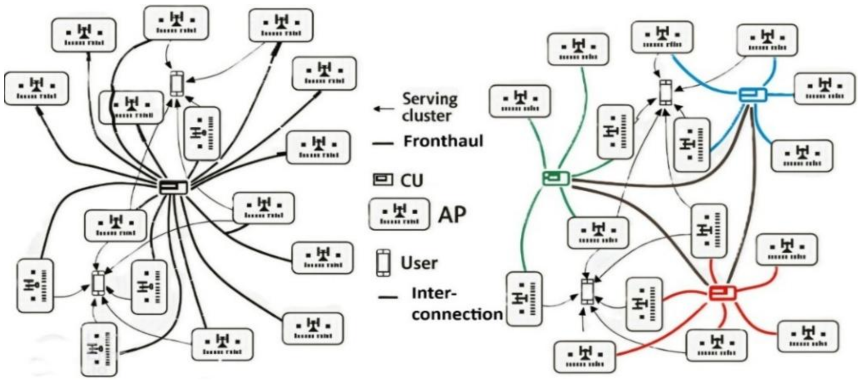
This survey paper will serve as a reference for the various fronthaul load or capacity methods, challenges, and solutions, especially in the UC-CF-mMIMO system. We investigate some of the major issues confronting the physical layer parameters. We also explore some issues regarding the future of research and anticipate some of the challenges that real-time deployment may pose. We also investigate how the UC-CF mMIMO network architecture can benefit specifically from the standpoint of fronthaul load.

Unlike existing studies focusing on network architectures, this paper delves into the UC methodology, offering insights into its unique characteristics and potential benefits. The most closely connected to this analysis are those focusing on combined physical layer parameters of coordinated distributed setups and clustering strategies. Siddique *et al.* [22] provide a review of the reviews heterogeneous wired and wireless backhaul architectures in CRAN, but does not address UC-CF systems.

This survey paper focuses particular on fronthaul signaling in a CF-UC system. In section 2, provides an overview of UC-CF mMIMO network architecture and the significance of fronthaul in systems with typical transmission modes. In section 3, investigate how fronthaul load influences the deployment of UC-CF networks. It also explores potential strategies and highlights emerging technologies aimed at overcoming this limitation. Section 4 presents the results and discussion, summarizing several open-research problems identified throughout the investigation. The final section 5 concludes the paper, summarizing the main findings, essential components, and discussing possible research prospects.

2. THEORY: NETWORK ARCHITECTURE

In Figure 2, we present two common types of architectures for the UC-CF mMIMO network, i.e., with a single control unit (CU) in Figure 2(a) and with multiple CUs in Figure 2(b). The selection of serving cluster access points for each user is critical in both systems because it defines the load and signaling at each user. This serving cluster can be formed in many ways, as simple as distance based or power-level-based and as complex as multistage approaches. In terms of scalability, a network with multiple CUs is both more scalable and practical than one with a single CU [23]. It is primarily due to the hierarchical design of the architecture of multiple CUs; it provides relatively low fronthaul traffic flow as well as lower delay.



(a) (b)

Figure 2. UC-CF mMIMO network of (a) with single CU (b) with multiple CU

In single-CU architecture, fronthaul capacity is severely limited unless signals are processed at the access point. A serial fronthaul link between the access points is employed to practically overcome this problem, similar to the most recent radio stripes system technology [24], in which numerous access points are combined into a single, readily deployable adhesive tape. Despite research on single-CU networks [1], multiple CUs are still required for implementation in the actual world. The multiple-CU architecture is compatible with the already existing cellular structure because at the backhaul of CUs, cells already exist. This network system is generally implemented using technologies like distributed SDN [25] to enable better coordination and dynamic AP-to-CU mapping. This setup also enhances load distribution and mitigates the risk of single points of failure [26]. As previously stated, cell structure exists on the CU backhaul, resulting in fewer connections in a UC approach because each user's data must be found in its neighboring CUs.

As seen in Figure 2, fronthaul represents the connection between a CU and an AP. The quantity of signaling data that the APs communicate with the CU depends on how much cooperation or coordination there is in the system. Sharing data often consists of channel status information (CSI), power coefficients, beamforming vectors, scheduling data, mobility management functions, and any other information that is useful or required to run the network. Although a lot of information travels from CU to AP, it does not follow that information travelling in other directions is irrelevant. Actually, CU also requires communicating the user's data to and from the APs continuously. The insufficient capacity of the fronthaul poses significant challenges in the system's actual implementation; it is primarily determined by the transmission mode and the level of coordination and cooperation among the APs. This problem is obvious in the system having wireless fronthaul [27]. On the other hand, the UC-CF network levies much lower weight on the fronthaul compared to others, mainly due to limited and optimized serving clusters.

The transmission mode in Figure 3 of a UC-CF mMIMO network is typically defined by the CoMP standard [17]. Within joint processing (JP), the system operates in either joint transmission (JT) or dynamic point selection (DPS) mode, with coordinated scheduling and beamforming (CS/CB) also being applied. In JT, user data is simultaneously available across several cooperating APs for concurrent transmission. In DPS, although multiple APs within the cluster have access to the user's data, only one—referred to as the transmission point (TP)—actively transmits at any given time. As instantaneous channel conditions change, the TP also changes from one sub-frame to another [17]. Both of these modes are combined to select multiple APs from serving clusters for data transmission. In CS/CB configurations, a single AP handles data transmission, while coordination for scheduling and beamforming involves all APs within the user's cluster. This approach requires the user's data to be present at just one transmitting AP, thereby lowering the load on fronthaul resources.

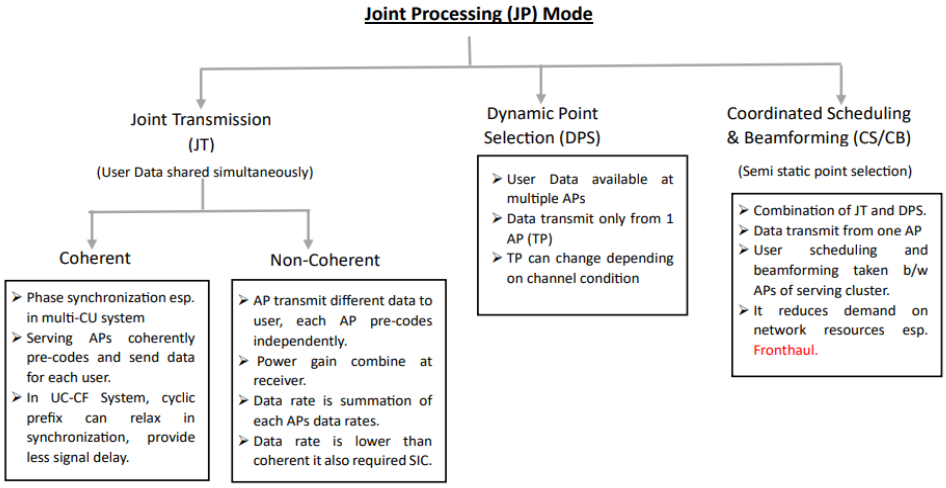


Figure 3. Flow chart of transmission mode

In JT mode, there are two transmission modes, namely coherent and non-coherent transmission, that deliver different achievable data rates for the user. In coherent transmission mode, each serving AP coherently pre-codes data and sends it for each user, and this method requires strict phase-synchronization among APs, which is critical, especially in multi-CU systems where the APs associated with a user are managed by multiple CUs. Each user of the UC-CF system is technically at the centre of the serving cluster in this mode, resulting in a smaller serving distance than a cell-edge user in an existing cell-centric system and, consequently, a much smaller spread of signal delay [13]. This gives the UC-CF system a notable advantage over the cell-centric system. In contrast, non-coherent transmission involves each AP sending distinct data streams to the user, each using its own precoding, which contributes to power accumulation at the receiver [28]. The overall throughput is the aggregate of individual AP rates, although it typically falls short of the performance achieved with coherent transmission [14], mainly due to successive interference cancellation (SIC), which requires the user to decode data. In this mode, severe phase-synchronization requirements are not required, making the transmitter end easier but the receiver end more complex due to SIC. But in a UC system, due to the smaller cluster size, the complexity of SIC can be affordable. Furthermore, each of these modes necessitated a different fronthaul capacity and model.

3. METHOD: FRONTHAUL ISSUES

There are many advanced algorithms and schemes that have been proposed for CF mMIMO systems; many of which are also applicable to UC-CF mMIMO frameworks. All these are based on the simple assumption of an error-less fronthaul links and perfect hardware. In practical deployment, all these assumptions are unlikely to hold; therefore, we address the practical challenges faced by UC-CF system has to encounter in implementation and then review existing solutions designed to tackle them. The inadequate capacity of fronthaul connections between access points and control units is the main issue for a UC-CF system [29]. When optical fiber cable is used as a fronthaul link [30], it causes huge power loss (e.g., 0.25 W/Gbps), and since each AP hosts many antennas, this results in a high volume of signal transmissions. When a signal is converted to digital form, a fronthaul link's capacity needs to be multiplied to match user data rates, especially in the uplink, to ensure that the signal is transmitted accurately. Reducing fronthaul signaling remains a major hurdle in the practical deployment of UC-CF mMIMO networks [29], [31]. Several strategies can be employed to mitigate the limitations posed by constrained fronthaul capacity.

3.1. Software defined networks

In this direction, to alleviate fronthaul load, distributed software defined networks (SDN) [25] can also be used in the management of the CUs. Typically, SDN architecture includes 3-layers: application, control, and infrastructure (or data). It is suited for handling complex network scenarios. One commonly used protocol is OpenFlow [32], which facilitates remote control, is one of a number of preset protocols that it uses to carry out its role. As illustrated in Figure 4, the southbound interface connects the control and infrastructure layers for coordination and management. The northbound interface offers APIs and functions for monitoring and application-level communication. SDN controllers also issue directives to the network elements they manage.

In distributed SDN, we can avoid potential single-point failures by logically centralizing the system's physically distributed controllers. If a controller becomes non-operational, others can seamlessly assume its responsibilities. Deploying SDN controllers—such as the Floodlight controller [33]— at the CUs, SDN can integrate into the UC-CF mMIMO system. As indicated earlier, it will enable APs to operate the SDN southbound interface. SDN will enable dynamic assignment of CUs and APs in a UC-CF system, which will result in automatic migration of AP from overloaded CUs to underloaded ones. This automatic migration will handle CU failure circumstances and reduce signaling on some overloaded fronthauls [26]. This distributed SDN is generally used in the computer network but can also be applied to UC systems.

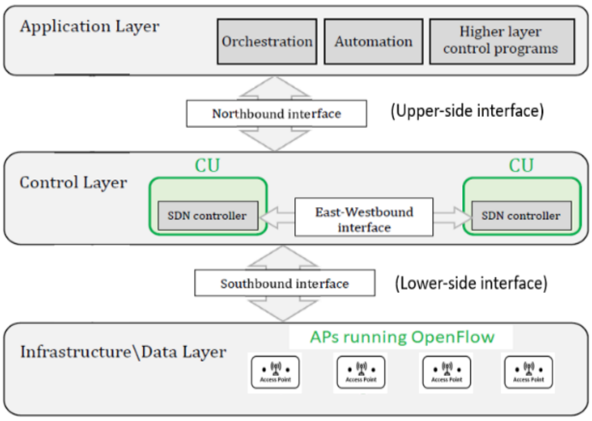


Figure 4. 3-layer architecture of SDN [25]

3.2. Quantizing

Minimizing fronthaul signaling can be achieved through source coding and fronthaul data compression [34]. The data compression is achieved by quantizing the data to reduce the fronthaul load. The load on the fronthaul of centralized CUs is also dependent on different types of quantization [35]. In the first case, quantized CSI and data signals are forwarded, while in the second, only the quantized versions of the weighted signals are forwarded. Bashar *et al.* [35] compares these schemes using the "use-and-then-forget (UatF) bounding technique", which concluding that the first scheme yields slightly better uplink throughput. This advantage depends on the user count and the number of antennas per AP, but the performance gap narrows as antenna count grows. The findings also indicate that performance becomes close to perfect if quantization bits exceed seven. Maryopi *et al.* [36], Bashar *et al.* [37] study Bussgang decomposition to model the effect of quantization where mid-rise uniform quantizer function is used and Bashar *et al.* [38] investigates the uplink for different scenarios of channel estimation and information sent back to CU with optimal uniform quantization. Also, cooperative access networks is used for optimum fronthaul quantization [39] in distributed mMIMO network.

The low-resolution ADCs [40] can be employed at APs for a practical solution towards minimizing load, and based on the location of channel estimation and the transmission of information to the CUs as shown in Figure 5, they can be further classified as:

- Compress-forward-estimate (CFE) [7]: in this scenario as shown in Figure 5(a), each AP compresses both the received pilot and data signals through quantizing and then forwards it to CUs through the fronthaul link. The CUs then carry out channel estimation, determine the combining weights, and decode the data. In this process, centralized combining is performed even with limited fronthaul capacity link, by shifting the quantization process to the APs.
- Estimate-compress-forward (ECF) [7]: in this method as shown in Figure 5(b), begins with the channel estimation performed locally at each APs, after which each AP individually quantized estimated channels and data signals and sent this compressed data in the direction of the CU. The CU recovers the CSI after receiving it and uses centralized combining to accomplish data detection. This method reduces fronthaul signaling in a distributed manner by performing quantization at the APs.
- Estimate-multiply-compress-forward (EMCF) [41]: in this scheme as shown in Figure 5(c), the AP initially estimates the channels, followed by multiplying the data received from user equipment with the local vector computed from local channel estimation at each AP. After multiply, data compression is performed by quantizing the results and forwarding them to CU, followed by data detection at CU. In this scheme, combining vectors is design, as well as compression is performed at APs.

- Estimate–multiply-compress-forward-weight (EMCFW) [41]: this scheme as shown in Figure 5(d), is similar to the above-mentioned EMCF scheme except a filter is used at CU to boost the overall detection performance. The signal is multiplied by receiver filter coefficients at the CU, and all other steps similar to EMCF, such as combining vector and compression, are performed at APs in a distributed manner.

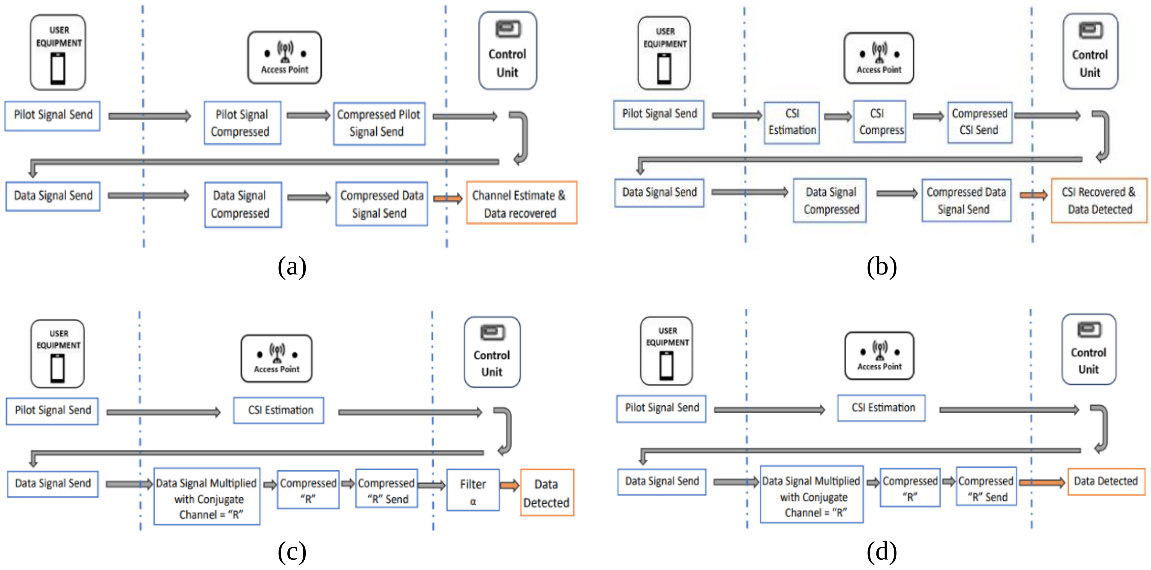


Figure 5. Signal compression on fronthaul in various quantizing method for uplink transmission: (a) CFE, (b) ECF, (c) EMCF, and (d) EMCFW

In downlink transmission, two-approaches are considered:

- Compress-after-precoding (CAP) [42]: this scheme is especially suitable for the downlink transmission of data in centralized precoding. In this scheme, the CU first applies the precoding to the data signal and subsequently compresses it before forwarding the signal to the APs for transmission.
- Precoding-after-compress (PAC) [43]: unlike the CAP this scheme is suitable for the distributed precoding. This method quantizes the signal at the CU level for each UE individually to compress it. After that, each APs then receives these compressed symbols and independently generate the corresponding precoding vectors for each user on their own.

As it is clear, compression through quantization is a direct and efficient method to minimize the fronthaul load. However, this quantization (especially uniform) causes the quantization error to be immunized and causes a considerable performance loss. As a result, different types of transmission schemes have different performance because they require different fronthaul rate allocation for different parameters. Therefore, to compare these six uplink and downlink transmission schemes on the basis of required processing at AP and CU, see Table 1.

Table 1. Comparison of various quantization schemes

Scheme	Processing at		Overall design and compression
	APs	CU	
CFE	Both CSI and data quantization	Channel estimation and design combining	Centralized-distributed
ECF	Channel estimation, both CSI and data quantization	Design combining	Centralized-distributed
EMCF	Channel estimation, design combining	-	Distributed-distributed
EM-CFW	Channel estimation, design combining	Receiver filter design	Distributed-distributed
CAP	Pre-coding design quantization	-	Centralized-centralized
PAC	Quantization	Pre-coding design	Distributed-distributed

3.3. Compute and forward

Another method for reducing fronthaul burden is the compute-and-forward (C&F) scheme [44], in which the cardinality of the transmitted symbol is reduced. This system combines the signals from UEs with the same cardinality into a single integer that is forwarded by each AP. The computation rate determines how quickly UEs can be identified by choosing a coefficient vector. Theoretically, this technique fulfils the fronthaul requirement for a CF mMIMO system under lossless transmission, however it is only effective when the UEs transmit with identical power. In practical scenarios, each UE transmits varying power levels to counteract pathloss effects. To address this, the expanded compute-and-forward (E-C&F) scheme [45] was introduced. It accounts for the unequal power distribution of UEs and also tolerates different types of noise at the AP to ensure minimum loss in computation, ultimately resulting in an overall performance gain over the standard C&F scheme.

3.4. Beamforming

The beamforming can also be used to minimize the fronthaul load, especially under the preset fronthaul capacity [46] condition. In this scheme, the total transmission power of the APs is minimized — often via optical fiber links—while maintaining quality-of-service (QoS) and fronthaul limitations. Similarly, for CF-mMIMO networks with constraints, beamforming with group sparse structure is used. This group's sparse structure outperformed the zero-forcing (ZF) precoding schemes which is some benchmark solutions itself. Using the instantaneous or statistical CSI instead of sharing actual CSI between APs and CUs can reduce traffic on the fronthaul. Interdonato *et al.* [47], two distributed beamforming strategies are introduced that avoid real-time CSI exchange, while power control remains centralized and is optimized based on long-term channel statistics.

3.5. Level of cooperation or coordination

The degree of coordination among the APs and their respective CUs also defined the fronthaul load. The effect of different levels of cooperation in a CF mMIMO network is presented in [1]. We will investigate this from the UC perspective. It is obvious that as the level of cooperation decreases and it lead to lower fronthaul signaling. As a result, APs are required to include as many baseband functions as possible, resulting in a less cost-effective system. Bjornson and Sanguinetti [1], Buzzi *et al.* [2] categorize these cooperation strategies into four distinct levels, assuming one CU per network: i) fully centralized—all estimation is performed at CU. ii) Two-stage, partially centralized, with CSI estimated locally at each AP. iii) The same as the second, except that CU performs detection. iv) A fully distributed small cell network. These schemes provided the highest to lowest fronthaul usage, respectively. In terms of spectral efficiency, the highest performance is typically achieved under the first cooperation level, while the lowest is seen in the fourth. However, exceptions exist; for instance, under high AP density conditions, the fourth scheme may surpass the third in performance.

3.6. Using deep convolution neural network

The study of the effect on fronthaul load using different power control strategies in which a deep convolutional neural network (DCNN) is employed [48] as shown in Figure 6. Using DCNN, large-scale fading (LSF) is mapped to the assigned power coefficients in this study. At APs, the combine-quantize-and-forward technique is employed to deliver quantized versions of combined signal data multiplied by the channel conjugate to the CU. In other techniques, such quantize-and-forward, APs can send quantized data to the central unit, which will then estimate receive data and a channel Similarly, ZF detection is carried out at the CPU, and to manage the associated complexity, a heuristic sub-optimal method is proposed in [49]. This method leverages DCNN to convert the problem into a geometric programming framework. This reduces the fronthaul load while increasing the complexity of the system.

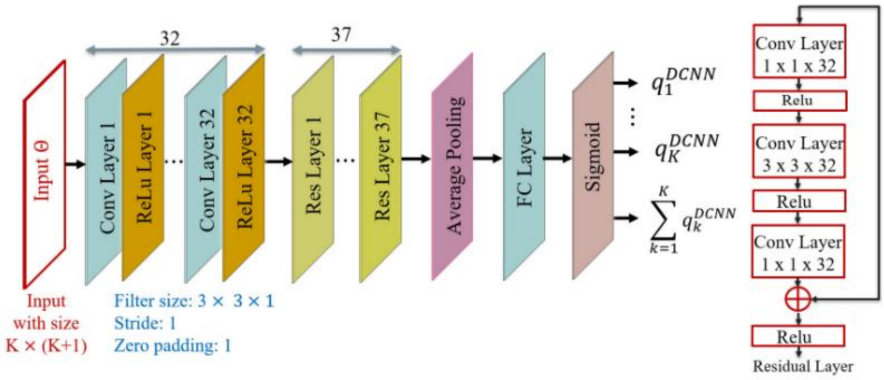


Figure 6. The DL-based power scheme for CF mMIMO system using DCNN [48]

3.7. Wired and wireless fronthaul

One more technique to alleviate the load on the fronthaul is by implementing a serial fronthaul in the network. In this method, APs are allowed to relay data between other APs, especially in the network where some APs function as a proxy for a CUs. However, a until robust and scalable wired connection solution is implemented in the system, this is not a scalable solution for actual deployment. These wired bus network connections among APs are such that APs do not intervene in relaying signals to other APs. The radio-stripe system [24] provides the solution to actually fulfil all these conditions. In this system, radio frequency elements are embedded along the transmission medium, while signal processing tasks are executed directly within the cable infrastructure. One of the most important advantages of this is that it can be deployed easily with no impact on the surroundings or an existing structure. The wireless fronthaul provides more flexibility in deploying the APs; its location also plays a significantly influences in maximizing the data rate [27]. One more method exists for wireless backhaul: "integrated access and backhaul (IAB) technology". It allows using wireless spectrum for backhaul or fronthaul.

The overall performance of the uplink for CF-mMIMO as well as CF mMIMO is investigated in [50], where a two-level design for the fronthaul is used under different conditions. Due to low interference, the results demonstrate that the UC method performs significantly better than the CF scheme, especially in sum SE.

4. RESULT AND DISCUSSION

Our investigation revealed that various approaches, including SDN-based dynamic CU management and quantization-based fronthaul load reduction, offer viable solutions to mitigate fronthaul limitations in UC-CF mMIMO systems. By dynamically assigning CUs and optimizing AP-CU communication, SDN architectures enhance system scalability and resilience to failures [25], [33]. Quantization techniques, such as CFE and EMCF, demonstrate promising fronthaul load reduction capabilities, albeit with different performance trade-offs [35].

Moreover, C&F schemes and advanced beamforming techniques provide efficient means to minimize fronthaul load while preserving QoS factors and system performance [45], [46]. Leveraging DCNNs for power control enables sophisticated fronthaul optimization strategies, enhancing system efficiency and reducing complexity [48]. Additionally, incorporating both wired and wireless fronthaul solutions, such as serial fronthaul networks like radio stripes and IAB technologies, presents scalable alternatives for mitigating fronthaul congestion [24]. Table 2 list out advantages and drawbacks of various solution selected in our survey paper to tackle issue of fronthaul in UC-CF mMIMO system.

Table 2. Advantages and drawbacks of various solution for lowering fronthaul load

Scheme	Advantages	Drawbacks
SDN	– Dynamic CU management. – Enhanced system scalability and resilience. – Optimized AP-CU communication.	– Implementation complexity.
Quantization	– Fronthaul load reduction through source coding and compression. – Minimized fronthaul signaling.	– Performance trade-offs between different quantization schemes.
C&F	– Reduced fronthaul load by combining signals from multiple UEs. – Expanded C&F scheme accommodates unequal power distributions and noise.	– Effectiveness dependent on power equality among UEs.
Beamforming	– Minimized fronthaul load while preserving QoS factors. – Group sparse beamforming structures outperform benchmark precoding schemes.	– System complexity.
Level of cooperation	– Flexible adaptation to fronthaul limitations by varying cooperation levels. – Optimized fronthaul usage patterns.	– Cost-effectiveness trade-offs.
DCNN	– Advanced fronthaul optimization strategies using DCNNs. – Enhanced system efficiency.	– Increased system complexity. – Performance trade-offs between different power control schemes.
Wireless fronthaul	– Scalable solutions for mitigating fronthaul congestion. – Enhanced system resilience. – Flexibility in deployment.	– Deployment challenges in practical scenarios. – Limited scalability.

Each of above solution presents unique advantages and challenges for reducing fronthaul load in UC-CF-mMIMO systems. The choice of solution depends on various factors, including system requirements, deployment constraints, and cost considerations. Further research and development efforts are needed to address the challenges and optimize the implementation of these solutions in real-world scenarios. Overall, our findings underscore the critical role of fronthaul optimization in realizing the full potential of UC-CF-mMIMO systems, offering insights into future research directions and practical deployment considerations strategies and address system scalability challenges.

5. CONCLUSION AND FUTURE SCOPE

In conclusion, our survey elucidates the practical challenges associated with fronthaul limitations in UC-CF-mMIMO systems and presents state-of-the-art methodologies to mitigate these challenges. By leveraging dynamic CU management, quantization techniques, compute and forward schemes, advanced beamforming strategies, and DCNNs, fronthaul congestion can be effectively alleviated while optimizing system performance.

Furthermore, incorporating wired and wireless fronthaul solutions offers scalable alternatives for mitigating fronthaul limitations and enhancing system resilience. However, future research endeavors are necessary to address scalability challenges and develop practical deployment strategies for UC-CF-mMIMO systems. In summary, our survey underscores the importance of fronthaul optimization in realizing the full potential of UC-CF mMIMO systems, offering valuable insights for researchers and practitioners in the field.

In addition to this, we outline some of the key issues and difficulties in order to resolve the fronthaul link capacity issue in the future. Future works can look into how non-uniform quantizers can be used to solve the limited fronthaul problem. In general, UC-CF-mMIMO systems can be used to formulate issue statements for limited capacity fronthaul, and millimeter wave is still an uncharted territory that researchers can investigate in the future. Finally, the impact of this multiple-CU architecture for UC networks on fronthaul is crucial; hence, more research is required.

FUNDING INFORMATION

Authors state no funding involved.

AUTHOR CONTRIBUTIONS STATEMENT

This journal uses the Contributor Roles Taxonomy (CRediT) to recognize individual author contributions, reduce authorship disputes, and facilitate collaboration.

Name of Author	C	M	So	Va	Fo	I	R	D	O	E	Vi	Su	P	Fu
Syed Tariq Ali	✓	✓	✓	✓	✓	✓		✓	✓	✓			✓	
Dr. Anamika Singh	✓					✓		✓		✓	✓	✓		

C : **C**onceptualization
M : **M**ethodology
So : **So**ftware
Va : **Va**lidation
Fo : **Fo**rmal analysis
I : **I**nvestigation
R : **R**esources
D : **D**ata Curation
O : Writing - **O**riginal Draft
E : Writing - Review & **E**diting
Vi : **Vi**sualization
Su : **Su**pervision
P : **P**roject administration
Fu : **Fu**nding acquisition

CONFLICT OF INTEREST STATEMENT

Authors state no conflict of interest.

DATA AVAILABILITY

Data availability is not applicable to this paper as no new data were created or analyzed in this study.

REFERENCES

[1] E. Bjornson and L. Sanguinetti, "Making cell-free massive MIMO competitive with MMSE processing and centralized implementation," *IEEE Transactions on Wireless Communications*, vol. 19, no. 1, pp. 77–90, Jan. 2020, doi: 10.1109/TWC.2019.2941478.

[2] S. Buzzi, C. D'Andrea, A. Zappone, and C. D'Elia, "User-centric 5G cellular networks: resource allocation and comparison with the cell-free massive MIMO approach," *IEEE Transactions on Wireless Communications*, vol. 19, no. 2, pp. 1250–1264, Feb. 2020, doi: 10.1109/TWC.2019.2952117.

[3] H. He, X. Yu, J. Zhang, S. Song, and K. B. Letaief, "Cell-free massive MIMO for 6G wireless communication networks," *Journal of Communications and Information Networks*, vol. 6, no. 4, pp. 321–335, Dec. 2021, doi: 10.23919/jcin.2021.9663100.
[4] C. D'Andrea and E. G. Larsson, "User association in scalable cell-free massive MIMO systems," in *2020 54th Asilomar Conference on Signals, Systems, and Computers*, Nov. 2020, vol. 2020-Novem, pp. 826–830, doi: 10.1109/IEEECONF51394.2020.9443536.
[5] H. D. Tuan, A. A. Nasir, H. Q. Ngo, E. Dutkiewicz, and H. V. Poor, "Scalable user rate and energy-efficiency optimization in cell-free massive MIMO," *IEEE Transactions on Communications*, vol. 70, no. 9, pp. 6050–6065, Sep. 2022, doi: 10.1109/TCOMM.2022.3194046.
[6] H. Q. Ngo, A. Ashikhmin, H. Yang, E. G. Larsson, and T. L. Marzetta, "Cell-free massive MIMO versus small cells," *IEEE Transactions on Wireless Communications*, vol. 16, no. 3, pp. 1834–1850, Mar. 2017, doi: 10.1109/TWC.2017.2655515.
[7] H. Masoumi and M. J. Emadi, "Performance analysis of cell-free massive MIMO system with limited fronthaul capacity and hardware impairments," *IEEE Transactions on Wireless Communications*, vol. 19, no. 2, pp. 1038–1053, Feb. 2020, doi: 10.1109/TWC.2019.2950316.
[8] J. Zhang, E. Bjornson, M. Matthaiou, D. W. K. Ng, H. Yang, and D. J. Love, "Prospective multiple antenna technologies for beyond 5G," *IEEE Journal on Selected Areas in Communications*, vol. 38, no. 8, pp. 1637–1660, Aug. 2020, doi: 10.1109/JSAC.2020.3000826.
[9] R. W. Heath and A. Lozano, *Foundations of MIMO communication*. Cambridge University Press, 2018.
[10] L. Lu, G. Y. Li, A. L. Swindlehurst, A. Ashikhmin, and R. Zhang, "An overview of massive MIMO: benefits and challenges," *IEEE Journal on Selected Topics in Signal Processing*, vol. 8, no. 5, pp. 742–758, Oct. 2014, doi: 10.1109/JSTSP.2014.2317671.
[11] E. Björnson, M. Matthaiou, and M. Debbah, "Massive MIMO with non-ideal arbitrary arrays: hardware scaling laws and circuit-aware design," *IEEE Transactions on Wireless Communications*, vol. 14, no. 8, pp. 4353–4368, Aug. 2015, doi: 10.1109/TWC.2015.2420095.
[12] E. G. Larsson, O. Edfors, F. Tufvesson, and T. L. Marzetta, "Massive MIMO for next generation wireless systems," *IEEE Communications Magazine*, vol. 52, no. 2, pp. 186–195, Feb. 2014, doi: 10.1109/MCOM.2014.6736761.
[13] H. A. Ammar and R. Adve, "Power delay profile in coordinated distributed networks: user-centric v/s disjoint clustering," in *2019 IEEE Global Conference on Signal and Information Processing (GlobalSIP)*, Nov. 2019, pp. 1–5, doi: 10.1109/GlobalSIP45357.2019.8969384.
[14] H. A. Ammar, R. Adve, S. Shahbazpanahi, G. Boudreau, and K. V. Srinivas, "Downlink resource allocation in multiuser cell-free MIMO networks with user-centric clustering," *IEEE Transactions on Wireless Communications*, vol. 21, no. 3, pp. 1482–1497, Mar. 2022, doi: 10.1109/TWC.2021.3104456.
[15] J. Zhang, S. Chen, Y. Lin, J. Zheng, B. Ai, and L. Hanzo, "Cell-free massive MIMO: a new next-generation paradigm," *IEEE Access*, vol. 7, pp. 99878–99888, 2019, doi: 10.1109/ACCESS.2019.2930208.
[16] H. Huh, A. M. Tulino, and G. Caire, "Network MIMO with linear zero-forcing beamforming: large system analysis, impact of channel estimation, and reduced-complexity scheduling," *IEEE Transactions on Information Theory*, vol. 58, no. 5, pp. 2911–2934, May 2012, doi: 10.1109/TIT.2011.2178230.
[17] 3GPP, "Coordinated multi-point operation for LTE physical layer aspects," 2013.
[18] A. Checko *et al.*, "Cloud RAN for mobile networks - a technology overview," *IEEE Communications Surveys and Tutorials*, vol. 17, no. 1, pp. 405–426, 2015, doi: 10.1109/COMST.2014.2355255.
[19] D. Liu, S. Han, C. Yang, and Q. Zhang, "Semi-dynamic user-specific clustering for downlink cloud radio access network," *IEEE Transactions on Vehicular Technology*, vol. 65, no. 4, pp. 2063–2077, Apr. 2016, doi: 10.1109/TVT.2015.2431917.
[20] D. Gesbert, S. Hanly, H. Huang, S. Shamai Shitz, O. Simeone, and W. Yu, "Multi-cell MIMO cooperative networks: a new look at interference," *IEEE Journal on Selected Areas in Communications*, vol. 28, no. 9, pp. 1380–1408, Dec. 2010, doi: 10.1109/JSAC.2010.101202.
[21] C. Zhu and W. Yu, "Stochastic modeling and analysis of user-centric network MIMO systems," *IEEE Transactions on Communications*, vol. 66, no. 12, pp. 6176–6189, Dec. 2018, doi: 10.1109/TCOMM.2018.2867458.
[22] U. Siddique, H. Tabassum, E. Hossain, and D. I. Kim, "Wireless backhauling of 5G small cells: challenges and solution approaches," *IEEE Wireless Communications*, vol. 22, no. 5, pp. 22–31, Oct. 2015, doi: 10.1109/MWC.2015.7306534.
[23] E. Bjornson and L. Sanguinetti, "Scalable cell-free massive MIMO systems," *IEEE Transactions on Communications*, vol. 68, no. 7, pp. 4247–4261, Jul. 2020, doi: 10.1109/TCOMM.2020.2987311.
[24] Z. H. Shaik, E. Bjornson, and E. G. Larsson, "MMSE-optimal sequential processing for cell-free massive MIMO with radio stripes," *IEEE Transactions on Communications*, vol. 69, no. 11, pp. 7775–7789, Nov. 2021, doi: 10.1109/TCOMM.2021.3100619.
[25] F. Bannour, S. Souihi, and A. Mellouk, "Distributed SDN control: survey, taxonomy, and challenges," *IEEE Communications Surveys and Tutorials*, vol. 20, no. 1, pp. 333–354, 2018, doi: 10.1109/COMST.2017.2782482.
[26] H. A. Ammar, Y. Nasser, and A. Kayssi, "Dynamic SDN controllers-switches mapping for load balancing and controller failure handling," in *2017 International Symposium on Wireless Communication Systems (ISWCS)*, Aug. 2017, vol. 2017-Augus, pp. 216–221, doi: 10.1109/ISWCS.2017.8108113.
[27] H. A. Ammar, R. Adve, S. Shahbazpanahi, and G. Boudreau, "Analysis and design of distributed MIMO networks with a wireless fronthaul," *IEEE Transactions on Communications*, vol. 70, no. 2, pp. 980–998, Feb. 2022, doi: 10.1109/TCOMM.2021.3128538.
[28] "3GPP TR 21.915, Release 15 Description, Technical Specification Group Services and System Aspects," 2019.
[29] E. Björnson and L. Sanguinetti, "Cell-free versus cellular massive MIMO: what processing is needed for cell-free to win?," in *2019 IEEE 20th International Workshop on Signal Processing Advances in Wireless Communications (SPAWC)*, Jul. 2019, vol. 2019-July, pp. 1–5, doi: 10.1109/SPAWC.2019.8815488.
[30] H. Q. Ngo, L.-N. Tran, T. Q. Duong, M. Matthaiou, and E. G. Larsson, "On the total energy efficiency of cell-free massive MIMO," *IEEE Transactions on Green Communications and Networking*, vol. 2, no. 1, pp. 25–39, Mar. 2018, doi: 10.1109/TGCN.2017.2770215.
[31] M. Bashar, K. Cumanan, A. G. Burr, M. Debbah, and H. Q. Ngo, "Enhanced max-min SINR for uplink cell-free massive MIMO systems," in *IEEE International Conference on Communications*, May 2018, vol. 2018-May, pp. 1–6, doi: 10.1109/ICC.2018.8422577.
[32] N. McKeown *et al.*, "Openflow: enabling innovation in campus networks," *ACM SIGCOMM Computer Communication Review*, vol. 38, no. 2, pp. 69–74, Mar. 2008, doi: 10.1145/1355734.1355746.
[33] "Project Floodlight, Floodlight Controller," 2021.
[34] S. Das and M. Ruffini, "A variable rate fronthaul scheme for cloud radio access networks," *Journal of Lightwave Technology*, vol. 37, no. 13, pp. 3153–3165, Jul. 2019, doi: 10.1109/JLT.2019.2912127.

[35] M. Bashar, K. Cumanan, A. G. Burr, H. Q. Ngo, M. Debbah, and P. Xiao, "Max-min rate of cell-free massive MIMO uplink with optimal uniform quantization," *IEEE Transactions on Communications*, vol. 67, no. 10, pp. 6796–6815, Oct. 2019, doi: 10.1109/TCOMM.2019.2926706.

[36] D. Maryopi, M. Bashar, and A. Burr, "On the uplink throughput of zero forcing in cell-free massive MIMO with coarse quantization," *IEEE Transactions on Vehicular Technology*, vol. 68, no. 7, pp. 7220–7224, Jul. 2019, doi: 10.1109/TVT.2019.2920070.

[37] M. Bashar, K. Cumanan, A. G. Burr, H. Q. Ngo, E. G. Larsson, and P. Xiao, "Energy efficiency of the cell-free massive MIMO uplink with optimal uniform quantization," *IEEE Transactions on Green Communications and Networking*, vol. 3, no. 4, pp. 971–987, Dec. 2019, doi: 10.1109/TGCN.2019.2932071.

[38] M. Bashar *et al.*, "Uplink spectral and energy efficiency of cell-free massive MIMO with optimal uniform quantization," *IEEE Transactions on Communications*, vol. 69, no. 1, pp. 223–245, Jan. 2021, doi: 10.1109/TCOMM.2020.3028305.

[39] A. Burr, M. Bashar, and D. Maryopi, "Cooperative access networks: optimum fronthaul quantization in distributed massive MIMO and cloud RAN - invited paper," in *2018 IEEE 87th Vehicular Technology Conference (VTC Spring)*, Jun. 2018, vol. 2018-June, pp. 1–5, doi: 10.1109/VTCSpring.2018.8417560.

[40] G. Femenias and F. Riera-Palou, "Fronthaul-constrained cell-free massive MIMO with low resolution ADCs," *IEEE Access*, vol. 8, pp. 116195–116215, 2020, doi: 10.1109/ACCESS.2020.3004499.

[41] M. Bashar, K. Cumanan, A. G. Burr, H. Q. Ngo, E. G. Larsson, and P. Xiao, "On the energy efficiency of limited-backhaul cell-free massive MIMO," in *IEEE International Conference on Communications*, May 2019, vol. 2019-May, doi: 10.1109/ICC.2019.8761134.

[42] G. Femenias and F. Riera-Palou, "Cell-free millimeter-wave massive MIMO systems with limited fronthaul capacity," *IEEE Access*, vol. 7, pp. 44596–44612, 2019, doi: 10.1109/ACCESS.2019.2908688.

[43] M. N. Boroujerdi, A. Abbasfar, and M. Ghanbari, "Cell free massive MIMO with limited capacity fronthaul," *Wireless Personal Communications*, vol. 104, no. 2, pp. 633–648, Oct. 2019, doi: 10.1007/s11277-018-6038-1.

[44] Q. Huang and A. Burr, "Compute-and-forward in cell-free massive MIMO: great performance with low backhaul load," in *2017 IEEE International Conference on Communications Workshops, ICC Workshops 2017*, May 2017, pp. 601–606, doi: 10.1109/ICCW.2017.7962724.

[45] J. Zhang, J. Zhang, D. W. K. Ng, S. Jin, and B. Ai, "Improving sum-rate of cell-free massive MIMO with expanded compute-and-forward," *IEEE Transactions on Signal Processing*, vol. 70, pp. 202–215, 2022, doi: 10.1109/TSP.2021.3129337.

[46] C. Hua, Y. Luo, and H. Liu, "Wireless backhaul resource allocation and user-centric clustering in ultra-dense wireless networks," *IET Communications*, vol. 10, no. 15, pp. 1858–1864, Oct. 2016, doi: 10.1049/iet-com.2016.0119.

[47] G. Interdonato, M. Karlsson, E. Bjornson, and E. G. Larsson, "Local partial zero-forcing precoding for cell-free massive MIMO," *IEEE Transactions on Wireless Communications*, vol. 19, no. 7, pp. 4758–4774, Jul. 2020, doi: 10.1109/TWC.2020.2987027.

[48] M. Bashar *et al.*, "Exploiting deep learning in limited-fronthaul cell-free massive MIMO uplink," *IEEE Journal on Selected Areas in Communications*, vol. 38, no. 8, pp. 1678–1697, Aug. 2020, doi: 10.1109/JSAC.2020.3000812.

[49] M. Bashar *et al.*, "Deep learning-aided finite-capacity fronthaul cell-free massive MIMO with zero forcing," in *ICC 2020 - 2020 IEEE International Conference on Communications (ICC)*, Jun. 2020, vol. 2020-June, pp. 1–6, doi: 10.1109/ICC40277.2020.9149210.

[50] P. Agheli, M. J. Emadi, and H. Beyranvand, "Performance analysis of cell-free and user-centric MIMO networks with optical fronthaul and backhaul links," *arxiv preprint: 2011.06680*, Mar. 2021, doi: 10.48550/arXiv.2011.06680.

BIOGRAPHIES OF AUTHORS

Syed Tariq Ali is a research scholar pursuing a Ph.D. at LNCT University in Bhopal, India. Alongside his doctoral studies, he serves as a senior lecturer in the ETE Department at Government Women's Polytechnic College, Bhopal. Holding an M.Tech in digital communication from MANIT, Bhopal, he boasts nine years of teaching experience in polytechnic colleges. With a keen focus on MIMO and self-adaptive gaussian mixture model (GMM), Syed Tariq Ali's research endeavors center around user-centric approaches and optimization techniques in wireless communication systems. Currently, he is engaged in an innovative research project in optimization of physical layer parameters of user-centric cell-free mMIMO. He can be contacted at email: tariq.ali1990@gmail.com.

Dr. Anamika Singh is currently serving as assistant professor, research supervisor, and Ph.D. Guide in the Department of Electronics and Communication Engineering at LNCT University, Bhopal. She holds a Ph.D. in electronics and communication engineering and has extensive teaching and research experience at both undergraduate and postgraduate levels. Her interests include digital communication systems, VLSI design, embedded systems, wireless sensor networks (WSN), signal processing, and IoT. She has guided four Ph.D. scholars and currently supervises five doctoral candidates in areas such as low-power VLSI, machine learning in communications, and energy-efficient networks. She is recognized for her innovative and rigorous mentoring approach. She can be contacted at email: anamikasingh723@gmail.com.