International Conference on "The impact of emerging technologies on global societies: Environment, Ethics, Innovation and sustainability (IETGS 2025)" organized by Government Polytechnic College, Dewas, Madhya Pradesh on 27 Feb 2025.

Estimation of Channel State Information and Serving Cluster in User-Centric Cell-Free MIMO Systems: A Comprehensive Survey

Syed Tariq Ali[1], Dr. Anamika Singh [2]
[1]Research Scholar, Department of Electronic & Communication LNCT University Bhopal, Madhya Pradesh, India and Lecturer, Government Women's Polytechnic College, Bhopal, Madhya Pradesh, India
[2]Prof. Department of Electronic & Communication LNCT University Bhopal, Madhya Pradesh, India

Abstract: **User-centric cell-free massive MIMO (Multiple-Input Multiple-Output) systems have emerged as a transformative architecture for future wireless networks, offering uniform coverage, enhanced spectral efficiency, and improved energy efficiency. Two critical components of this architecture are the Estimation of Channel State Information (CSI) and the Formation of Serving Clusters. Accurate CSI estimation is essential for efficient beamforming, interference management, and resource allocation, while the formation of serving clusters ensures that each user is served by an optimal set of distributed units (DUs), minimizing interference and maximizing performance. This survey paper provides a comprehensive review of the state-of-the-art techniques, challenges, and solutions related to CSI estimation and serving cluster formation in user-centric cell-free MIMO systems. We discuss the impact of channel hardening, favorable propagation, pilot assignment strategies, and mobility on CSI estimation. Additionally, we explore various clustering algorithms, utility-based approaches, and dynamic clustering techniques for serving cluster formation. The paper concludes with a discussion on open research challenges and future directions, providing insights into the scalability, latency, and energy efficiency of these systems.**

Keywords: Information security, clustering, MIMO systems, distributed units, algorithms.

1. INTRODUCTION

User-centric cell-free MIMO systems represent a paradigm shift from traditional cellular networks, where each user is served by cluster of Access Points (APs) rather than a single base station. This architecture eliminates cell edges, providing uniform coverage and improved spectral efficiency. However, the success of this architecture heavily relies on various parameters and out of these two key aspects are- **Channel State Information (CSI) Estimation** and **Serving Cluster Formation**.

- **CSI Estimation**: Accurate CSI is crucial for beamforming, interference management, and resource allocation. In user-centric cell-free MIMO systems, the distributed nature of APs and the need for coordination among them make CSI estimation particularly challenging. The use of Time Division Duplex (TDD) systems, which exploit channel reciprocity, is common, but Frequency Division Duplex (FDD) systems require more sophisticated techniques due to the lack of reciprocity.

- **Serving Cluster Formation**: The formation of serving clusters is essential to ensure that each user is served by an optimal set of APs. This process involves selecting APs based on metrics such as signal strength, large-scale fading, and network performance. The dynamic nature of user mobility and network conditions further complicates this process.

This survey paper focuses on these two critical aspects, providing a detailed review of existing techniques, challenges, and future research directions.

2. CONCEPT OF CELL-FREE MASSIVE MIMO

a. Architecture of Cell-Free Massive MIMO Systems

Cell-free massive MIMO systems as shown in Figure.1 are characterized by a large number of distributed access points (APs) that collaboratively serve a smaller number of users. Unlike traditional cellular networks, where each user is served by a single base station, in cell-free systems, each user is served by multiple APs, forming a **serving**

cluster. The key components of the architecture include:

1. **Access Points (APs)**: These are the access points distributed across the coverage area, each equipped with multiple antennas.
2. **Central Units (CUs)**: These are the centralized processing units that coordinate the APs and perform tasks such as beamforming, resource allocation, and CSI estimation.
3. **Fronthaul Links**: These are the communication links between the APs and CUs, which carry user data and control information.

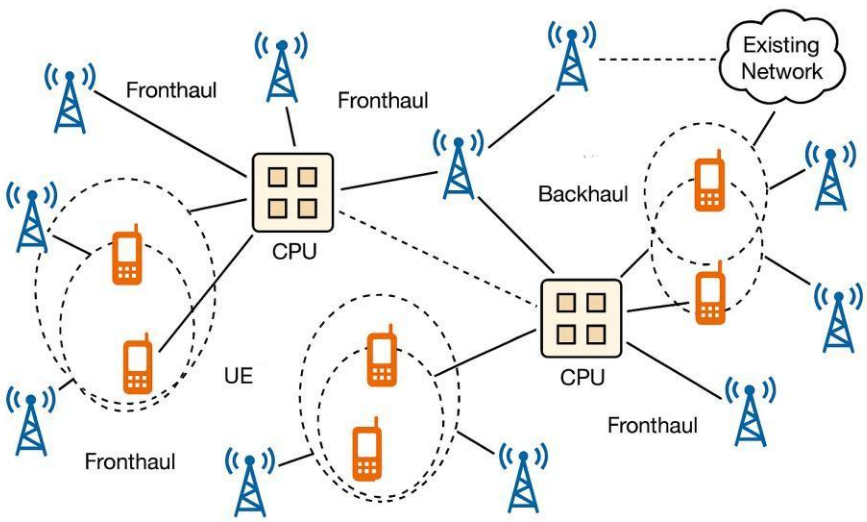


Figure.1: A cell-free network architecture

The architecture is designed to provide **uniform coverage** and **high spectral efficiency** by eliminating cell edges and ensuring that each user is served by the best possible set of APs.

b. Advantages Over Traditional Cellular Networks [2]

- **Uniform Coverage**: In traditional cellular networks, users at the cell edge experience poor signal quality due to interference from neighboring cells. In cell-free systems, each user is served by multiple APs, ensuring uniform coverage across the entire network.
- **Enhanced Spectral Efficiency**: By leveraging the spatial diversity provided by multiple APs, cell-free systems can achieve higher spectral efficiency compared to traditional cellular networks.
- **Improved Energy Efficiency**: The distributed nature of cell-free systems allows for more efficient use of power, as APs can be dynamically activated or deactivated based on user demand.
- **Scalability**: Cell-free systems can scale more effectively to accommodate a large number of users, as the load is distributed across multiple APs rather than concentrated at a single base station.

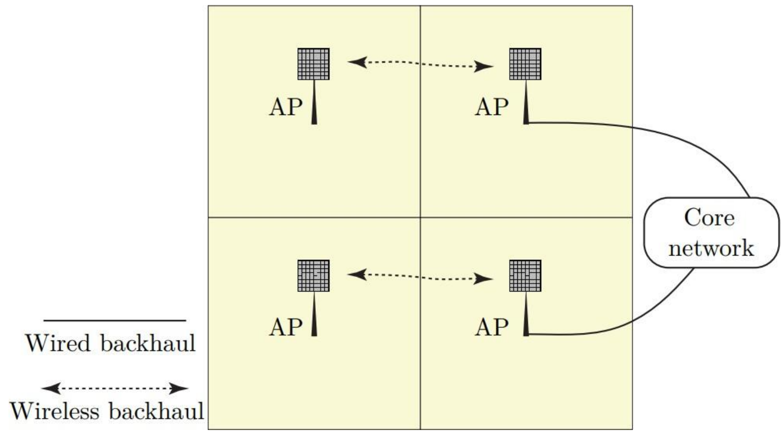
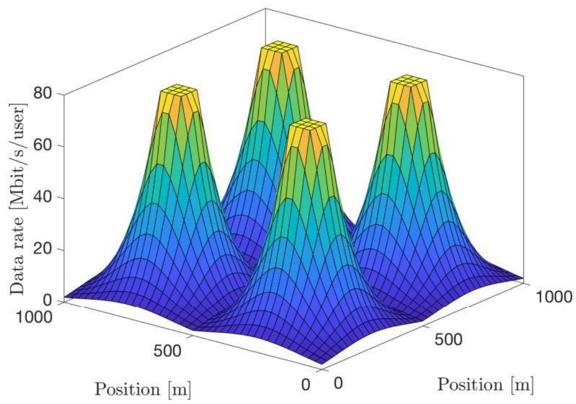
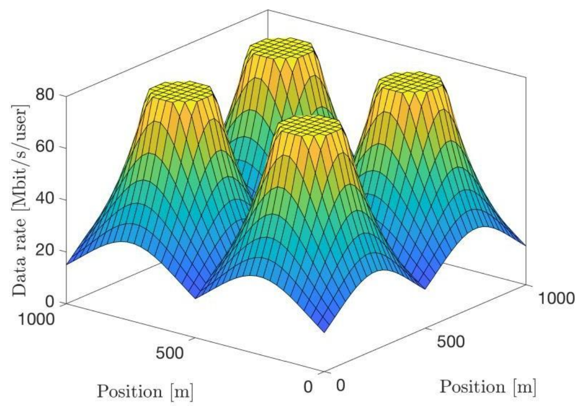


Figure.2: Cellular network with four APs [1]

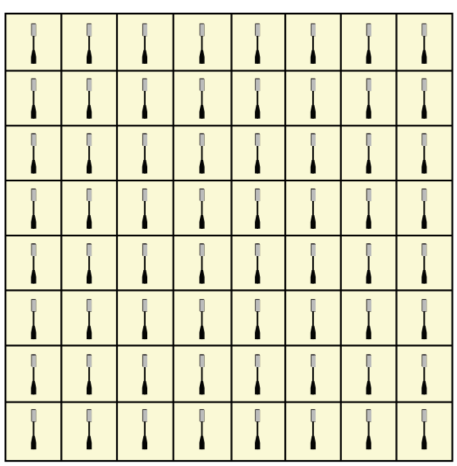


(a)

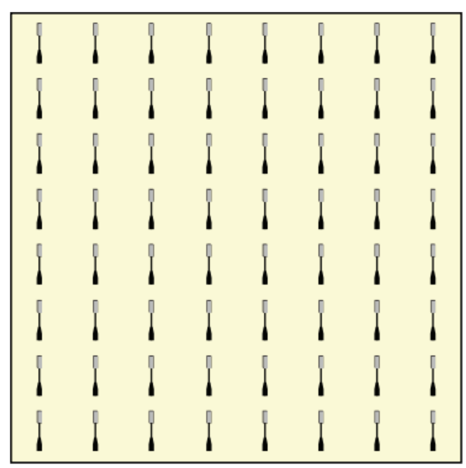


(b)

Figure.3: Downlink data rate achieved by a UE at different locations in the cellular network for different antennas (a) 9 dBi fixed-gain (b) 64 omni-directional [1]

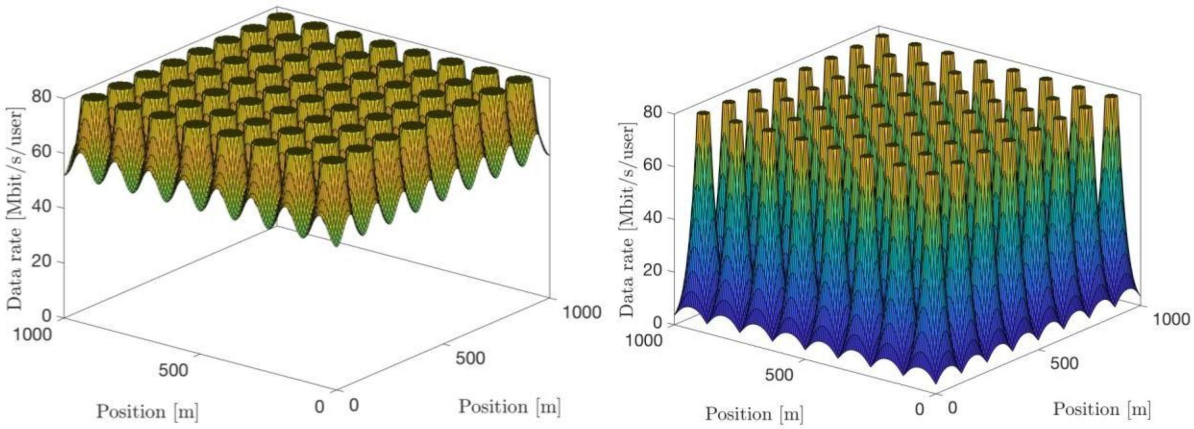


(a)



(b)

Figure.4: (a) Cellular setup and (b) Cell-free setup with 64 single-antenna APs [1]



(a) (b)

Figure.5: Downlink data rate achieved by a UE at different locations in a network with 64 APs for (a) Cell-free network (b) Cellular network with same AP locations [1]

3. ESTIMATION OF CHANNEL STATE INFORMATION (CSI)

Channel State Information (CSI) estimation is a critical component in user-centric cell-free massive MIMO systems, as it directly impacts the system's performance, including spectral efficiency, interference mitigation, and overall network reliability. This section provides an overview of the current techniques used for CSI estimation in such systems, along with their limitations.

1. Pilot-Based Channel Estimation Techniques

Pilot-based channel estimation is the most commonly used technique in user-centric cell-free massive MIMO systems. In this approach, users transmit pilot sequences during the training phase, and the access points (APs) use these pilots to estimate the channel state information.

a. Uplink Pilot Transmissions

- **Description:** In user-centric cell-free systems, uplink pilot transmissions are used to estimate the CSI at the APs. The APs then forward the estimated CSI to a central processing unit (CPU) for data recovery [3].
- **Limitations:**
    - **Pilot Contamination:** The primary limitation of pilot-based techniques is pilot contamination, which occurs when multiple users share the same pilot sequence, leading to interference and inaccurate CSI estimation [4] [6] [8].
    - **Hardware Impairments:** Non-ideal hardware at the APs and user equipment (UEs) can distort the received pilot signals, further degrading the accuracy of CSI estimation [3].

b. Advanced Pilot Assignment Algorithms

- **Description:** To mitigate pilot contamination, several advanced pilot assignment algorithms have been proposed. These algorithms aim to assign pilots in a way that minimizes interference and improves the accuracy of CSI estimation [12].
    - **Genetic Algorithm (GA):** A GA-based pilot assignment scheme has been proposed to find the optimal pilot sequence for each user, reducing pilot contamination [8].
    - **Hungarian Matching Algorithm:** This algorithm has been used to solve pilot assignment problems by matching users to pilots in a way that minimizes interference [8] [9].

- **Auction Algorithm:** An auction algorithm-based pilot assignment scheme has been proposed to iteratively assign pilots to users, ensuring that the assignment process is fair and efficient [13].

- **Limitations:**

  - **Computational Complexity:** These algorithms can be computationally intensive, especially in large-scale systems with many users and APs [8] [13].

  - **Scalability:** As the number of users increases, the number of available pilots becomes limited, making it challenging to assign unique pilots to all users [6] [8].

c. Pilot Power Allocation

- **Description:** Pilot power allocation schemes have been proposed to reduce pilot contamination by optimizing the power allocated to each pilot. For example, a pilot power allocation scheme that minimizes the channel estimation error has been proposed for user-centric cell-free systems [14].

- **Limitations:**

  - **Optimization Challenges:** Pilot power allocation requires solving complex optimization problems, which can be challenging in real-time systems [14].

  - **Interference Trade-off:** Reducing pilot power to minimize contamination can also reduce the signal-to-noise ratio (SNR), leading to less accurate CSI estimates [14].

2. Machine Learning-Aided Channel Estimation

Machine learning techniques have been increasingly used to improve the accuracy of CSI estimation in user-centric cell-free massive MIMO systems.

a. Deep Learning-Aided Channel Estimation

- **Description**: Deep learning techniques, such as convolutional neural networks (CNNs) and recurrent neural networks (RNNs), have been used to estimate CSI in cell-free systems. These techniques can learn the underlying patterns in the channel and improve estimation accuracy [6] [7].

- **Limitations:**

  - **Training Complexity:** Training deep learning models requires large amounts of data and computational resources, which can be challenging in real-time systems [6] [7].

  - **Generalization:** Deep learning models may not generalize well to new environments or channel conditions, requiring frequent retraining [6] [7].

b. Multitask Learning Framework

- **Description:** A multitask learning framework has been proposed to dynamically select from multiple deep learning models for pilot decontamination. This approach has been shown to outperform single-task learning models in terms of channel estimation accuracy [4].

- **Limitations:**
  - **Model Complexity:** The framework requires training and maintaining multiple models, which can increase the overall complexity of the system [4].

  - **Dynamic Selection:** Dynamically selecting the best model in real-time can be challenging and may require additional computational resources [4].

3. Distributed Channel Estimation Techniques

Distributed channel estimation techniques have been proposed to reduce the computational complexity and improve the scalability of user-centric cell-free massive MIMO systems.

a. Distributed User-Centric Channel Estimation

- **Description:** In this approach, each AP estimates the CSI for its associated users using local signals and compressed versions of signals from other APs. This reduces the need for centralized processing and improves scalability [15].
- **Limitations:**
  - **Communication Overhead:** While this approach reduces the need for centralized processing, it still requires communication between APs, which can lead to overhead [15].
  - **Accuracy Trade-off:** Distributed estimation may lead to less accurate CSI estimates compared to centralized approaches [15].

b. Decentralized Massive Access Random Scheme

- **Description:** A decentralized massive random-access scheme has been proposed for user- centric cell-free systems. In this scheme, each AP independently performs maximum likelihood (ML) estimation of the large-scale fading coefficients (LSFC) for the associated users [19].
- **Limitations:**
  - **Interference Management:** Decentralized estimation can lead to interference from neighbouring APs, which can degrade the accuracy of CSI estimates [19].
  - **Signaling Overhead:** The scheme requires signaling between APs to assess the activity of associated users, which can lead to overhead [19].

4. Subspace-Based Semi-Blind Channel Estimation

Subspace-based semi-blind channel estimation techniques have been proposed to improve the accuracy of CSI estimation in user-centric cell-free massive MIMO systems.

a. Subspace-Based Channel Estimation

- **Description:** This approach exploits the covariance matrix of the received signal to obtain the signal subspace. Based on the channel gains and pilot assignment results, a subspace selection scheme is developed to separate the target signal from interference [18].
- **Limitations:**
  - **Computational Complexity:** The eigenvalue decomposition (EVD) required for subspace estimation can be computationally intensive, especially in large-scale systems [18].
  - **Interference Mitigation:** While this approach can mitigate pilot contamination, it may not be effective in scenarios with high levels of interference [18].

5. Joint Active User Detection and Channel Estimation

Joint active user detection and channel estimation techniques have been proposed to improve the efficiency of CSI estimation in grant-free random-access scenarios.

a. Variational Bayesian Inference

- **Description:** A variational Bayesian inference-based algorithm has been proposed for joint active user detection and channel estimation. This approach exploits the block sparsity of the effective channel matrix to improve the accuracy of CSI estimation [17].

- **Limitations:**
    - **Computational Complexity**: The algorithm requires solving complex optimization problems, which can be computationally intensive [17].
    - **Convergence:** The convergence of the algorithm may be slow in scenarios with a large number of users and APs [17].

b. Adaptive Federated Learning

- **Description:** An adaptive federated learning-based approach has been proposed for joint pilot design and active user detection. This approach uses a distributed edge computing system to train local detection networks and achieve globally optimal performance [16].
- **Limitations:**
    - **Communication Overhead:** The approach requires communication between edge distributed units (EDUs), which can lead to overhead [16].
    - **Delay Constraints:** The algorithm may not be suitable for scenarios with strict delay constraints due to the need for iterative training [16].

6. Pilot Reuse and Clustering Techniques

Pilot reuse and clustering techniques have been proposed to mitigate pilot contamination and improve the efficiency of CSI estimation in user-centric cell-free massive MIMO systems.

a. Repulsive Clustering-Based Pilot Assignment

- **Description:** A repulsive clustering-based pilot assignment scheme has been proposed to mitigate pilot contamination by forming clusters of users with diverse pilot sequences [20].
- **Limitations:**
    - **Cluster Formation:** The formation of clusters can be computationally intensive, especially in large-scale systems [20].
    - **Interference:** While this approach reduces pilot contamination, it may not be effective in scenarios with high levels of interference [20].

b. Diverse Clustering Approach

- **Description:** A diverse clustering approach has been proposed to maximize intra-cluster diversity and assign the same pilots to users within the same cluster [21].
- **Limitations:**
    - **Cluster Formation:** The formation of clusters can be computationally intensive, especially in large-scale systems [21].
    - **Interference:** While this approach reduces pilot contamination, it may not be effective in scenarios with high levels of interference [21].

7. Rate-Splitting Assisted Channel Estimation

Rate-splitting assisted channel estimation techniques have been proposed to improve the spectral efficiency of user-centric cell-free massive MIMO systems.

a. Rate-Splitting Multiple Access (RSMA)

- **Description:** RSMA has been proposed as a superior approach to conventional space division multiple access (SDMA) for improving the spectral efficiency of cell-free systems. Closed- form spectral efficiency expressions have been derived for the common and private streams, considering imperfect CSI due to channel aging and pilot contamination [22].
- **Limitations:**
  - **Complexity:** The derivation of closed-form expressions requires complex mathematical modeling, which can be challenging in practice [22].
  - **Channel Aging:** The approach assumes imperfect CSI due to channel aging, which can degrade the accuracy of CSI estimates [22].

4. FORMATION OF SERVING CLUSTERS

User-centric cell-free massive MIMO systems provide promising solutions to enhance wireless communication and offer improved spectral efficiency, fairness, and reduced interference compared to traditional cellular systems. But in practicality these systems need the formation of clusters of access points (APs) that serve users, ensuring efficient resource allocation and interference management. This section delves into the current techniques used for cluster formation, supported by insights from the research papers provided.

1. Machine Learning Approaches for Cluster Formation

Machine learning has been widely adopted to optimize cluster formation in user-centric cell-free massive MIMO systems. These techniques leverage data-driven models to dynamically adjust clusters dynamically, ensuring optimal performance.

Deep Learning for User Clustering

- **Long Short-Term Memory (LSTM) Networks**: Author in [23], proposes a deep learning approach using LSTM to solve the user clustering problem. The model aims to maximize the sum spectral efficiency while controlling the number of active connections. The solution is scalable and does not require retraining, making it effective even with imperfect channel state information due to pilot contamination.
- **Deep Reinforcement Learning (DRL)**: Researchers also explore DRL for user-centric AP clustering. In Paper [28], a distributed DRL framework is proposed, where each user has an actor, and the action design focuses on the difference in cluster size over time. This approach reduces the neural network size and improves scalability. Author in [29] introduces a DRL framework that allows clusters to vary in size based on propagation conditions, optimizing sum spectral efficiency, fronthaul capacity, and power consumption.

Unsupervised Machine Learning Algorithms

- **K-Means and Variants**: Researcher also investigates the use of unsupervised ML algorithms like k-means, k-means++, and improved k-means++ for user clustering in NOMA-aided CF- mMIMO systems. These algorithms group users to maximize sum-rate and reduce bit error rate (BER). The proposed system with UC and ML algorithms demonstrates significant improvements in achievable sum-rate and BER compared to non-UC approaches [24].

2. Clustering Algorithms for User-Centric Systems

Clustering algorithms play a pivotal role in forming user-centric clusters, ensuring that each user is served by a subset of APs. These algorithms are designed to balance computational complexity and performance.

User-Centric (UC) Clustering

- **Genetic Algorithm for Fairness**: Author in [30], proposes a UC clustering algorithm using a genetic algorithm to solve the max-min rate optimization problem. This approach ensures fairness among users by deriving the lower bound of ergodic rates for CF mMIMO URLLC systems.

- **Dynamic Clustering for Mobility**: Author in [31], introduces a dynamic clustering algorithm (UC-kmeans++) to address user mobility. The algorithm reduces the inter-AP switching rate and improves throughput under high mobility conditions, providing a more effective solution for mobile users.

Scalable Clustering Methods

- **Distributed Processing**: Author in [25] presents a scalable UC framework with two methods to limit the number of radio units (RUs) serving each user. These methods reduce computational complexity and backhaul signaling demands, improving energy efficiency (EE) and spectral efficiency (SE).

- **Hybrid Network- and User-Centric Clustering**: Author in [26] proposes a hybrid approach combining network- and user-centric clustering to minimize fronthaul signaling. The method achieves up to 94% reduction in fronthaul load while maintaining spectral efficiency.

3. Pilot Assignment and Clustering Techniques

Pilot assignment is critical in cell-free massive MIMO systems to mitigate pilot contamination and improve channel estimation accuracy. Clustering-based pilot assignment techniques have been developed to address these challenges.

Clustering-Based Pilot Assignment

- **K-Means Clustering for Pilot Assignment**: Various authors propose clustering-based pilot assignment schemes, like in [35] author introduces a low-complexity clustering approach to refine user clustering and pilot allocation, enhancing data rates while maintaining low complexity. Similarly, in [37] author devises a k-means clustering-based pilot assignment (KCPA) and a Tabu search-based scheme (KCTSPA) to reduce pilot contamination in mmWave systems.

- **Graph-Based Clustering**: Author in [36], uses a graph representation of AP positions to cluster APs into groups, reducing pilot contamination by reusing pilots within groups.

Pilot Assignment in mmWave Systems

- **Structured Pilot Assignment**: Author in [37] extends the KCPA and KCTSPA schemes to mmWave systems, demonstrating reduced-uplink channel estimation errors with low complexity.

4. Resource Optimization and Load Balancing

Resource optimization is essential for maintaining the efficiency and scalability of user-centric cell-free massive MIMO systems. Techniques focus on balancing computational and communication resources.

Backhaul and Fronthaul Load Balancing

- **Dynamic Clustering and Resource Allocation**: Author in [27] proposes a dynamic clustering method using the Kuhn-Munkres algorithm for backhaul and Particle Swarm Optimization (PSO) for fronthaul clustering. This approach improves sum rates and load balancing, achieving an 18.23% improvement in sum rate and a 30% enhancement in Load Balancing Index (LBI).

- **Fronthaul Load Balancing**: Author in [5], addresses fronthaul load balancing and computation resource allocation, considering fronthaul topology and limited capacity. The framework highlights the importance of optimizing quantization bits in analog-to-digital conversion for improved performance.

Energy Efficiency and Computational Resource Management

- **Energy Efficiency Optimization**: Author in [25], emphasizes the importance of managing inter-CPU communication to control backhaul traffic and reduce computational complexity. The proposed techniques achieve a 98% reduction in computational complexity while maintaining spectral efficiency.

5. Other Techniques for Cluster Formation

Several other techniques have been explored to enhance cluster formation in user-centric cell-free massive MIMO systems.

Multi-Agent Reinforcement Learning (MARL)

- **MARL for AP Selection**: Author in [33] proposes a MARL algorithm where each AP is an agent trained to select which users to serve. The approach outperforms greedy selection and achieves performance comparable to the ideal canonical case.

Layered Partially Non-Orthogonal ZF-Based Beamforming

- **AP Selection for Beamforming**: Authors in [32] & [34] proposes a clustering method using layered partially non-orthogonal ZF-based beamforming. The method initializes AP clusters based on signal-to-leakage-and-noise-ratio and iteratively updates clusters to maximize throughput.

Matched-Decision AP Selection

- **Competitive Mechanism for AP Clustering**: Author in [38] introduces a matched-decision AP selection framework that connects users to intermediate AP clusters and expands clusters dynamically. The method improves spectral efficiency and reduces the number of UEs per AP.

5. COMPARATIVE ANALYSIS OF CSI ESTIMATION AND SERVING CLUSTER FORMATION

In this section, we provide a detailed comparative analysis of the two critical components of user-centric cell-free massive MIMO systems: **Channel State Information (CSI) Estimation** and **Serving Cluster Formation**. Both aspects are essential for the efficient operation of these systems, but they address different challenges and employ distinct techniques. The following table summarizes the key differences and similarities between these two components:

Aspect	CSI Estimation	Serving Cluster Formation
Objective	Accurately estimate the channel state between users and access points (APs) to enable efficient beamforming, interference management, and resource allocation.	Form optimal clusters of APs to serve each user, ensuring uniform coverage, reduced interference, and improved spectral efficiency.
Key Challenges	- Pilot contamination - Hardware impairments - Scalability - Channel aging - Inter-cluster interference - Computational complexity	- Dynamic user mobility - Scalability - Interference management - Resource allocation - Energy efficiency
Techniques	- Pilot-based channel estimation - Machine learning (e.g., deep learning, multitask learning) - Distributed channel estimation - Subspace-based semi-blind estimation - Joint active user detection and channel estimation - Pilot reuse and clustering techniques	- Machine learning (e.g., LSTM, DRL, k-means) - Clustering algorithms (e.g., genetic algorithms, dynamic clustering) - Pilot assignment and clustering techniques - Resource optimization and load balancing - Multi-agent reinforcement learning (MARL)

Performance Metrics	- Accuracy of CSI estimates - Spectral efficiency - Interference mitigation - Latency - Energy efficiency	- Sum spectral efficiency - Fairness among users - Interference reduction - Energy efficiency - Scalability
Limitations	- Pilot contamination remains a significant issue - Hardware impairments degrade accuracy - Scalability issues in large-scale deployments - Channel aging affects accuracy in high-mobility scenarios - High computational complexity	- Dynamic user mobility complicates cluster formation - Scalability challenges in large networks - Interference from neighboring clusters - Energy efficiency concerns - High computational complexity in real-time systems
Future Directions	- Advanced pilot assignment algorithms - Machine learning for adaptive CSI estimation - Distributed processing to reduce complexity - Mitigation of hardware impairments - Techniques to address channel aging	- Development of more robust and scalable clustering algorithms - Improved interference management techniques - Enhanced integration of machine learning for real-time cluster formation - Energy-efficient clustering methods

6. DISCUSSION ON LIMITATIONS

Despite the significant advancements in both CSI estimation and serving cluster formation, several limitations persist in user-centric cell-free massive MIMO systems:

1. **Pilot Contamination**: This remains a critical issue in CSI estimation, especially in scenarios with many users and limited pilot sequences. Pilot contamination leads to inaccurate CSI estimates, which degrade system performance. [4] [6] [8].
2. **Hardware Impairments**: Non-ideal hardware at both the APs and user equipment (UEs) can distort signals, leading to inaccurate CSI estimates and suboptimal cluster formation [3].
3. **Scalability**: As the number of users and APs increases, the computational complexity and communication overhead of both CSI estimation and cluster formation become overwhelming, limiting the scalability of these systems. [5] [10] [11].
4. **Channel Aging**: In high-mobility scenarios, channel aging can lead to outdated CSI, which degrades the accuracy of both CSI estimation and cluster formation. [22]
5. **Inter-Cluster Interference:** Users at the edge of clusters may experience interference from neighboring clusters, which can degrade the accuracy of CSI estimates and reduce the effectiveness of cluster formation.[9], [11].
6. **Energy Efficiency:** Both CSI estimation and cluster formation techniques need to be energy- efficient, especially in scenarios with limited power resources. Inefficient energy usage can lead to increased power consumption and reduced battery life for user devices.[3] [15].
7. **Computational Complexity**: Many of the proposed techniques require complex computations, which can be challenging in real-time systems, leading to increased latency and reduced system performance. [5] [8] [10].

7. FUTURE DIRECTIONS

To address the limitations of current techniques, future research in user-centric cell-free massive MIMO systems should focus on the following directions:

1. **Advanced Pilot Assignment Algorithms:** Developing more sophisticated pilot assignment algorithms that

can dynamically adapt to changing channel conditions and user densities will help mitigate pilot contamination and improve CSI estimation accuracy.

2. **Machine Learning Techniques:** Leveraging machine learning, such as deep learning and federated learning, can improve the accuracy and efficiency of both CSI estimation and cluster formation. These techniques can adapt to complex channel conditions and user mobility patterns.
3. **Distributed Processing:** Developing distributed processing techniques can reduce the computational complexity and communication overhead of both CSI estimation and cluster formation, improving scalability and efficiency.
4. **Hardware Impairment Mitigation:** Techniques to mitigate the impact of hardware impairments on CSI estimation and cluster formation will improve the overall system performance, especially in large- scale deployments.
5. **Channel Aging Mitigation:** Developing techniques to address channel aging will improve the accuracy of CSI estimates and cluster formation in high-mobility scenarios.
6. **Robust and Scalable Clustering Algorithms:** Future research should focus on developing more robust and scalable clustering algorithms that can handle dynamic user mobility and large-scale deployments.
7. **Energy-Efficient Techniques:** Energy efficiency should be a key consideration in the design of future CSI estimation and cluster formation techniques, especially for battery-powered user devices.

8. CONCLUSION

User-centric cell-free massive MIMO systems represent a transformative architecture for future wireless networks, offering uniform coverage, enhanced spectral efficiency, and improved energy efficiency. However, the success of these systems heavily relies on accurate **Channel State Information (CSI) Estimation** and efficient **Serving Cluster Formation**. While significant progress has been made in both areas, several challenges remain, including pilot contamination, hardware impairments, scalability, channel aging, and energy efficiency.

This survey has provided a comprehensive review of the state-of-the-art techniques, challenges, and solutions related to CSI estimation and serving cluster formation in user-centric cell-free MIMO systems. The comparative analysis highlights the key differences and similarities between these two critical components, while the discussion on limitations and future directions provides insights into the ongoing research efforts needed to address these challenges.

As the field evolves, advancements in machine learning, distributed processing, and energy-efficient techniques will be crucial for meeting the demands of future wireless communication systems. By addressing the current limitations and exploring new research directions, user-centric cell-free massive MIMO systems can achieve their full potential, enabling next-generation wireless networks with unprecedented performance and scalability.

REFERENCES

[1]. Özlem Tugfe Demir; Emil Björnson; Luca Sanguinetti, Foundations of User-Centric Cell-Free Massive MIMO, now, 2021.

[2]. 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, 2017.

[3]. 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: Information Theory, Nov. 2020, [Online]. Available: https://dblp.uni-trier.de/db/journals/corr/corr2011.html#abs-2011-06680

[4]. C. Victor, S. I. Mrutu, and A. Mvuma, "A Multitask Learning Framework for Pilot Decontamination in 5G Massive MIMO," Tanzania Journal of Engeering and Technology, Sep. 2023, doi: 10.52339/tjet.v42i3.910.

[5]. Z.-Y. Li, F. Göttsch, S. Li, M. Chen, and G. Caire, "Joint Fronthaul Load Balancing and Computation Resource Allocation in Cell-Free User-Centric Massive MIMO Networks," arXiv.org, vol. abs/2310.14911, Oct. 2023, doi: 10.48550/arxiv.2310.14911.

[6]. Deshpande, S., Aggarwal, M., Sabherwal, P., & Ahuja, S. "Machine Learning aided Channel Estimation for Cell-Free Networks using a novel pilot assignment algorithm," Feb. 2023, doi: 10.22541/au.167705647.78023463/v1.

[7]. Deshpande, S., Aggarwal, M., Sabherwal, P., & Ahuja, S. (2024). Deep Learning-aided Channel Estimation Combined with Advanced Pilot Assignment Algorithm to Mitigate Pilot Contamination for Cell-Free

Networks. Indian Journal of Science and Technology, 17(5), 465–477. https://doi.org/10.17485/ijst/v17i5.2961

[8]. Al Ayidh, Y. A. Sambo, and M. Imran, “Mitigation pilot contamination based on matching technique for uplink cell-free massive MIMO systems,” Dental science reports, vol. 12, no. 1, Oct. 2022, doi: 10.1038/s41598-022-21241-0

[9]. Mussbah, M., Schwarz, S., & Rupp, M. (2023). Pilot Contamination Reduction for Access Point Clustering-based Pilot Assignment. 1–6. https://doi.org/10.1109/pimrc56721.2023.10293898

[10]. C. Pan, H. Mehrpouyan, Y. Liu, M. Elkashlan, and N. Arumugam, “Joint Pilot Allocation and Robust Transmission Design for Ultra-Dense User-Centric TDD C-RAN With Imperfect CSI,” IEEE Transactions on Wireless Communications, vol. 17, no. 3, pp. 2038–2053, Jan. 2018, doi: 10.1109/TWC.2017.2788001.

[11]. Wei, C., Xu, K., Shen, Z. et al. Location-aided uplink transmission for user-centric cell-free massive MIMO systems: a fairness priority perspective. J Wireless Com Network 2022, 84 (2022). https://doi.org/10.1186/s13638-022-02171-x

[12]. M. Sarker and A. O. Fapojuwo, "Suppressing Pilot Contamination for Massive Access in User- centric Cell-free Massive MIMO Systems," 2022 IEEE 95th Vehicular Technology Conference: (VTC2022-Spring), Helsinki, Finland, 2022, pp. 1-6, doi: 10.1109/VTC2022- pring54318.2022.9860397.

[13]. M. Sarker and A. O. Fapojuwo, “Access Point-User Association and Auction Algorithm-Based Pilot Assignment Schemes for Cell-Free Massive MIMO Systems,” IEEE Systems Journal, pp. 1–12, Jan. 2023, doi: 10.1109/jsyst.2023.3255865.

[14]. M. Sarker and A. O. Fapojuwo, "Pilot Power Allocation Scheme for User-Centric Cell-free Massive MIMO Systems," 2023 IEEE 20th Consumer Communications & Networking Conference (CCNC), Las Vegas, NV, USA, 2023, pp. 763-768, doi: 10.1109/CCNC51644.2023.10059832.

[15]. R. Rompaey and M. Moonen, “Distributed Combined Channel Estimation and Optimal Uplink Receive Combining for User-Centric Cell-free Massive MIMO Systems,” IEEE open journal of signal processing,2024 doi: 10.1109/ojsp.2024.3377098.

[16]. Diao, Lei et al. “Adaptive Federated Learning-based Joint Pilot Design and Active User Detection in Scalable Cell-free Massive MIMO Systems.” 2023 4th Information Communication Technologies Conference (ICTC) (2023): 232-236. doi: 10.1109/ictc57116.2023.10154724.

[17]. J. Liu, S. Zhu, M. Zhao, and W. Zhou, “Variational Bayesian Inference-Based Joint Active User Detection and Channel Estimation in Cell-Free Massive MIMO,” pp. 1–5, Jun. 2024, doi: 10.1109/vtc2024-spring62846.2024.10683590.

[18]. B. Zhong, X. Zhu, and E. G. Lim, “Subspace-Based Semi-Blind Channel Estimation for User- Centric Cell-Free Massive MIMO Systems,” pp. 1–5, Jun. 2024, doi: 10.1109/vtc2024- spring62846.2024.10683682.

[19]. Y. Hu, D. Wang, X. Xia, and X. You, “Decentralized Massive Access Random Scheme in User- Centric Cell-Free Massive MIMO System,” Apr. 2024, doi: 10.1109/wcnc57260.2024.10570945.

[20]. S. Mohebi, A. Zanella, and M. Zorzi, “Repulsive Clustering Based Pilot Assignment for Cell-Free Massive MIMO Systems,” European Signal Processing Conference, pp. 717–721, Mar. 2022, doi: 10.48550/arXiv.2203.12403.

[21]. S. Mohebi, A. Zanella, and M. Zorzi, “Pilot Reuse in Cell-Free Massive MIMO Systems: A Diverse Clustering Approach,” Dec. 2022, doi: 10.48550/arxiv.2212.08872.

[22]. F. Odoom, J. O. Agyapong, Y. Bunyaminu, and K. Lee, “Rate-Splitting Assisted Cell-Free Massive MIMO Systems with Channel Aging and Pilot Contamination,” pp. 619–620, Oct. 2024, doi: 10.1109/ictc62082.2024.10827274.

[23]. G. Di Gennaro, A. Buonanno, G. Romano, S. Buzzi, and F. Palmieri, “A Deep Learning Approach for User-Centric Clustering in Cell-Free Massive MIMO Systems,” pp. 661–665, Sep. 2024, doi: 10.1109/spawc60668.2024.10694529

[24]. R. Arshad, S. Baig, and S. Aslam, “User clustering in cell-free massive MIMO NOMA system: A learning based and user centric approach,” alexandria engineering journal, Mar. 2024, doi: 10.1016/j.aej.2024.01.064

[25]. M. M. M. Freitas et al., “Scalable User-Centric Distributed Massive MIMO Systems with Restricted Processing Capacity,” IEEE Transactions on Wireless Communications, p. 1, Jan. 2024, doi: 10.1109/twc.2024.3491153.

[26]. P. Lai et al., “Hybrid Network- and User-Centric Scalable Cell-Free Massive MIMO for Fronthaul

Signaling Minimization,” IEEE Transactions on Vehicular Technology, pp. 1–6, Jan. 2024, doi: 10.1109/tvt.2024.3456114.

[27]. M. Ajmal, M. A. Tariq, M. M. Saad, S. Kim, and D. Kim, “Scalable Cell-Free Massive MIMO Networks Using Resource-Optimized Backhaul and PSO-Driven Fronthaul Clustering,” IEEE Transactions on Vehicular Technology, pp. 1–16, Jan. 2024, doi: 10.1109/tvt.2024.3465458.

[28]. Y. Tsukamoto, A. Ikami, N. Aihara, T. Murakami, H. Shinbo, and Y. Amano, User-centric AP Clustering with Deep Reinforcement Learning for Cell-Free Massive MIMO. 2023. doi: 10.1145/3616390.3618291.

[29]. C. F. Mendoza, S. Schwarz, and M. Rupp, “User-Centric Clustering in Cell-Free MIMO Networks using Deep Reinforcement Learning,” pp. 1036–1041, May 2023, doi: 10.1109/icc45041.2023.10279626.

[30]. J. Wang, J. Fang, H. Chen, H. Guo, F. Shu, and P. Zhu, “User-Centric Clustering for Uplink Cell- Free Massive MIMO URLLC Systems,” pp. 780–785, Aug. 2024, doi: 10.1109/iccc62479.2024.10681698.

[31]. Y. Liu and B. Li, “User Mobility-Based Dynamic Clustering Algorithm for Cell-Free Massive MIMO Systems,” pp. 778–783, May 2024, doi: 10.1109/icet61945.2024.10672692.

[32]. H. Kato, T. Hara, S. Suyama, S. Nagata, and K. Higuchi, “Low-Complexity User-Centric AP Clustering Method in Downlink Cell-Free MIMO with Regularized ZF-Based Beamforming,” pp. 1–6, Oct. 2023, doi: 10.1109/vtc2023-fall60731.2023.10333421.

[33]. B. Banerjee, R. C. Elliott, W. A. Krzyrnień and H. Farmanbar, "Access Point Clustering in Cell- Free Massive MIMO Using Multi-Agent Reinforcement Learning," 2022 IEEE 33rd Annual International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC), Kyoto, Japan, 2022, pp. 1086-1092, doi: 10.1109/PIMRC54779.2022.9977980.

[34]. D. Ishii, T. Hara, N. Nonaka, and K. Higuchi, “Clustering Method in Downlink Cell-Free MIMO Using Layered Partially Non-orthogonal ZF-Based Beamforming,” pp. 1–5, Jun. 2023, doi: 10.1109/vtc2023-spring57618.2023.10199306.

[35]. F. Gamal, M. Elsaadany, M. Abdel-Raheem, and O. A. M. Aly, “Pilot Assignment in Cell-Free Massive MIMO: A Low-Complexity Clusterization Technique,” pp. 434–439, Jul. 2024, doi: 10.1109/itc-egypt61547.2024.10620505.

[36]. M. Mussbah, S. Schwarz, and M. Rupp, “Access Point Clustering-based Pilot Assignment for Cell- free Massive MIMO,” Asilomar Conference on Signals, Systems and Computers, pp. 722–726, Oct. 2022, doi: 10.1109/IEEECONF56349.2022.10052089.

[37]. B. Zhong, X. Zhu and E. G. Lim, "Clustering-based Pilot Assignment for User-Centric Cell-Free mmWave Massive MIMO Systems," 2022 IEEE 96th Vehicular Technology Conference (VTC2022- Fall), London, United Kingdom, 2022, pp. 1-5, doi: 10.1109/VTC2022-Fall57202.2022.10012778.

[38]. M. Freitas et al., "Matched-Decision AP Selection for User-Centric Cell-Free Massive MIMO Networks," in IEEE Transactions on Vehicular Technology, vol. 72, no. 5, pp. 6375-6391, May 2023, doi: 10.1109/TVT.2023.3235980.