International Research Journal on Advanced Engineering Hub
e ISSN: 2584-2137
Vol. 04 Issue: 01 January 2026
Page No: 361-367
https://irjaeh.com
https://doi.org/10.47392/IRJAEH.2026.0050

Enhancing Spectral Efficiency in Cell-Free Massive MIMO Systems Using K-Means++ Clustering and AP-UE Association

Mrs. Soumya Prasad[1], Dr. Nagamani K [2], H Leyasri[3], Harshitha D Pai[4], Juhi Kothari[5], Lakshmi Lavanya N[6]
[1]Research Scholar, R V College of Engineering, Bangalore, Karnataka, India
[2]Professor, Dept. of Electronics & Telecommunication Engg, R V College of Engineering, Bangalore, Karnataka, India
[3,4,5,6] Department of AI & ML, RNS Institute of Technology, Bangalore, Karnataka, India
Emails: *soumyaphd31@gmail.com[1], nagamanik@rvce.edu.in[2], srileya22@gmail.com[3], harshithapai2017@gmail.com[4], juhikthr@gmail.com[5], lakshmilavanya292@gmail.com[6]*

Abstract
Current cellular networks are based on autonomous cells, which often struggle to support large numbers of users due to uneven coverage. As the world becomes increasingly dependent on wireless communication, there is a growing need for cellular networks that offer higher spectral and energy efficiency through multiple wireless access points. Cell-free Massive Multiple-Input Multiple-Output (MIMO) networks successfully meet this demand. In addition to meeting modern wireless communication requirements, these networks can also mitigate many existing interference challenges. This study aims to lower the computational burden of cell-free Massive MIMO systems, enhancing their practicality for large-scale deployment and addressing one of their major operational challenges. We propose a novel access point selection algorithm that combines a machine learning approach for clustering, specifically the K-means++ algorithm and the AP-UE association. Based on simulation findings and evaluation metrics, the proposed algorithm consistently outperforms existing methods, demonstrating notable improvements in efficiency and performance.
Keywords: *Cell-free Massive MIMO, K-Means++, AP-UE Association, Clustering, Spectral Efficiency.*

1. Introduction
Massive MIMO has established itself as a foundational technology in modern wireless communication due to its ability to achieve substantial improvements in spectral efficiency and reliability through the use of large antenna arrays. Early contributions, such as those of Marzetta and others [1], [2], demonstrated how extensive antenna deployments can significantly increase capacity and provide uniform service. These findings motivated the evolution of network architectures toward more distributed and scalable designs. However, conventional cellular systems still exhibit critical limitations, including uneven coverage, severe inter-cell interference, and restricted scalability. To address these challenges, the concept of cell-free massive MIMO (CF-mMIMO) emerged as an alternative paradigm in which numerous distributed APs jointly serve all users without predefined cell boundaries [3]–[5]. Comparative studies have shown that CF mMIMO offers superior service uniformity and more efficient interference suppression than small-cell architectures [5], making it an attractive solution for future wireless networks. A significant body of research has focused on improving the practical deployment of CF-mMIMO. Works such as [15] and [21] have explored efficient downlink training and ubiquitous distributed network architectures. Surveys and tutorials, including [16] and the comprehensive monograph in [18], provide a detailed understanding of the opportunities and challenges associated with user-centric CF-mMIMO designs. In addition, optimization techniques addressing precoding, power control, and energy efficiency have been widely studied in [17], [22], highlighting the need for scalable and computationally efficient methods. Despite its

International Research Journal on Advanced Engineering Hub
e ISSN: 2584-2137
Vol. 04 Issue: 01 January 2026
Page No: 361-367
https://irjaeh.com
https://doi.org/10.47392/IRJAEH.2026.0050

advantages, CF-mMIMO introduces substantial computational and coordination overhead. Many studies have proposed antenna and AP selection techniques such as threshold-based, greedy, and algorithmic approaches to reduce complexity in large-scale networks [6]–[10]. However, these solutions often struggle to scale efficiently as the number of APs increases. To further improve system performance, more sophisticated AP-selection and user-association mechanisms based on effective channel gain have been proposed in [11], while pilot contamination mitigation strategies have been addressed in works such as [12] and [13]. To overcome scalability limitations, clustering-based methods have gained attention. The research in [19] proposes user association and clustering techniques customized for CF mMIMO, demonstrating reduced fronthaul load and improved performance. Machine learning and AI-driven approaches have also emerged, with studies such as [20] emphasizing intelligent clustering as a core component of future 6G cell free architectures. Furthermore, CF-mMIMO adaptations for mmWave frequencies, as discussed in [24], highlight the value of distributed AP deployment in dealing with high-frequency propagation challenges. Overall, the literature demonstrates a steady shift from cell based architectures to highly distributed, user-centric, and intelligent wireless systems. While substantial progress has been achieved, the need for low-complexity, scalable AP-selection and clustering algorithms remains a central research focus [23]. Motivated by these gaps, this work proposes an approach that integrates K-means++ clustering with optimized AP–UE association to enhance spectral efficiency while significantly reducing coordination and computational overhead in cell-free massive MIMO networks. The rapid expansion of wireless connectivity necessitates innovative architectural solutions to deliver reliable, energy efficient, and high-capacity communication services. Massive Multiple-Input Multiple-Output (mMIMO) technology has emerged as a foundational approach to achieving substantial gains in spectral efficiency and uniform service quality by employing very large antenna arrays [1], [2]. Despite these advantages, conventional centralized mMIMO architectures encounter significant bottlenecks in terms of coverage, scalability, and computational complexity. To mitigate these limitations, Cell-Free Massive MIMO (CF-mMIMO) has been proposed as a transformative wireless architecture that removes traditional cell boundaries and deploys numerous distributed Access Points (APs) collaboratively serving all users [3], [8], [9]. This distributed topology enhances spectral efficiency, reduces inter-cell interference, and ensures more consistent user experiences in dense or heterogeneous environments (Figure 1).

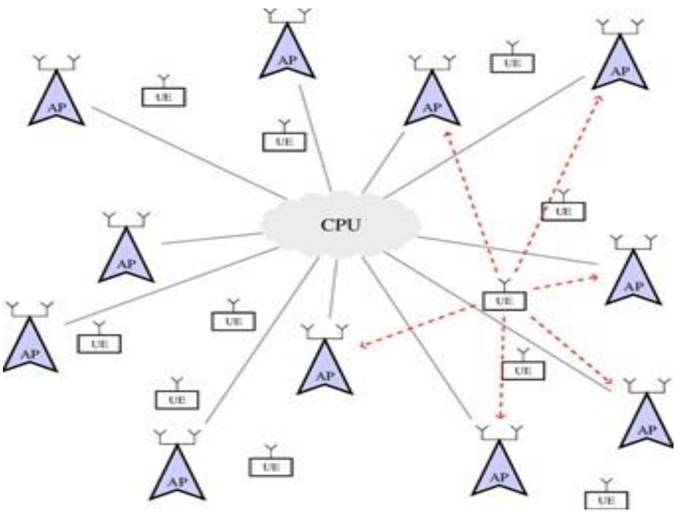

Figure 1 Cell Free Massive MIMO Architecture

However, CF-mMIMO systems still rely on centralized processing units for tasks such as pilot allocation, AP se lection, and resource coordination, which causes substantial computational burdens. Existing work has explored methods for efficient antenna or AP selection [5]–[7], [10], [11] and improved user association mechanisms [12], [14]. In addition, intelligent pilot allocation techniques have been proposed to reduce pilot contamination and enhance channel estimation accuracy [4], [13]. Motivated by these challenges, this paper introduces a low complexity heuristic based on the K-Means++ clustering al gorithm to optimize AP assignment in CF-mMIMO networks. Using unsupervised learning, the proposed method reduces computational overhead while maintaining effective user–AP

International Research Journal on Advanced Engineering Hub
e ISSN: 2584-2137
Vol. 04 Issue: 01 January 2026
Page No: 361-367
https://irjaeh.com
https://doi.org/10.47392/IRJAEH.2026.0050

associations, thus improving system performance with mini mal central coordination. This paper is organized as follows: Section I introduces the fundamental concepts and motivation of the work. Section II describes the system model used in the study. Section III presents the proposed algorithm in detail. Section IV discusses the simulation results and performance analysis. Finally, Section V provides the concluding remarks.

2. System Model

The system under consideration is a cell-free massive MIMO network operating in Time Division Duplex (TDD) mode. It comprises M Access Points (APs) and K User Equipments (UEs), where M << K. These APs are geographically distributed throughout the coverage area to ensure uniform service quality and eliminate traditional cell boundaries. Between AP m and UE k, the channel coefficient is given by

$$g_{mk} = \sqrt{(\beta_{mk} h_{mk})} \quad (1)$$

where β_{mk} denotes the large-scale fading coefficient, and the small-scale Rayleigh fading coefficient is modeled as

$$h_{mk} \sim (0, 1) \quad (2)$$

To incorporate distance-based propagation, β_{mk} is modeled using the log-distance path loss model with Gaussian shadowing:

$$\beta_{mk} = 10\text{^}(-(PL_0 + 10n \log_{10}(d_{mk} / d_0) + X_r) / 10) \quad (3)$$

where,

PL_0 is the path loss at the reference distance,

n is the path loss exponent,

d_{mk} is the distance between AP m and UE k,

$X_\sigma \sim \mathcal{N}(0, \sigma^2)$ represents the log-normal shadowing component.

The received signal at UE k is given by:

$$y_k = \sum_{m=1}^{M} g_{mk} x_m + n_k \quad (4)$$

where:

- x_m is the signal transmitted by AP m ,
- $n_k \sim \mathcal{CN}(0, \sigma^2)$ is AWGN.

The achievable spectral efficiency for UE k is given by:

$$R_k = \log_2(1 + (|\Sigma_{m}\in\mathcal{M}_k \sqrt{p_m}\, g_{mk}|^2) / (\Sigma_{j\neq k} |\Sigma_m\in\mathcal{M}_j \sqrt{p_m}\, g_{mj}|^2 + \sigma^2)) \quad (5)$$

where: $\mathcal{M}_k$ is the set of APs serving UE k and p_m is the transmit power of AP m.

In our work we aim to maximize the sum spectral efficiency:

max $\mathcal{M}_k$ $\sum$ k = 1^k R_k subject to | $\mathcal{M}_k$ | ≤ $Nmax$ $\sum_{k=1}^{K} Pmk \leq Pmax, \forall m,$

Also considers front haul and cluster constraints.

3. Proposed Algorithm

We propose a two-stage algorithm combining K-Means++ clustering and optimized AP-UE association.

3.1. AP Selection Model

In a cell-free massive MIMO network, M distributed APs each equipped with N antennas jointly serve K UEs. The received uplink signal is given by:

$$y = \sum_{\{k=1\}}^{\{K\}} \sqrt{\{p_k\} g_k s_k} + \boldsymbol{n} \quad (6)$$

Where

- p_k is the transmit power of the UE k-th,
- $g_k$denotes the MN X 1 channel vector between all AP antennas and UE k
- $\boldsymbol{n}$ represents AWGN with zero mean and variance σ^2

The channel vector g_k models both large-scale fading (path loss and shadowing) and small-scale Rayleigh fading:

$$g_k = D_{k^{1/2}} \boldsymbol{h} k \quad (7)$$

Here,$\mathbf{D}_k$ is a diagonal matrix containing the large-scale fading coefficients, and $\mathbf{h}_k$ a vector of small-scale fading coefficients with i.i.d. $\mathcal{CN}(0, 1)$ entries.

The main objective of AP selection is to choose the most suitable subset of access points (APs) to serve each user equipment (UE), aiming to enhance overall system performance while minimizing computational complexity and back haul signaling overhead. Traditional methods select APs based on fixed thresholds (e.g., channel gain, SNR), whereas the proposed method in this work employs K-means++ clustering to dynamically group APs according to user proximity and channel quality. Only the APs within the cluster of a given UE are selected to participate in its data transmission.

3.2. K-Means++ Based AP-UE Association

1. Input: Locations of APs $\{a_1, a_2, \ldots, a_M\}$ and UEs

IRJAEH

International Research Journal on Advanced Engineering Hub
e ISSN: 2584-2137
Vol. 04 Issue: 01 January 2026
Page No: 361-367
https://irjaeh.com
https://doi.org/10.47392/IRJAEH.2026.0050

$\{u_1, u_2, \ldots u_M\}$
.

2. Perform K-Means++ clustering to group APs based on proximity to UEs.
3. For each cluster:
    a. Compute channel gain matrix G using the log- distance model.
    b. Assign each UE to the AP with maximum channel gain.
4. Output: Optimized AP-UE associations {Mk} (Table 1).

Table 1 Simulation System Parameters

Parameter	Value
Area	100 m^2
APs	100 (random)
UEs	50 (random)
Bandwidth	20 MHz
Noise Power	-94 dBm
Path Loss Exponent	3.7
Shadowing Std. Dev.	8 dB
Reference Path Loss	**30** B at d0 = 1 m

5. Simulation Results and Analysis

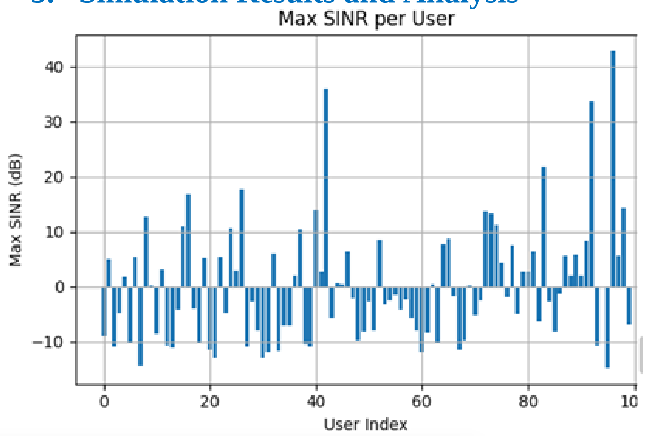

Figure 2 Simulation Results W.R.T Max. SINR Per User

The plot in Figure 2 shows a system with a high degree of user SINR variability, with many users having poor SINR and some reaching extremely high values. This implies uneven channel conditions and potential inefficiencies in resource allocation, AP selection, or interference management.

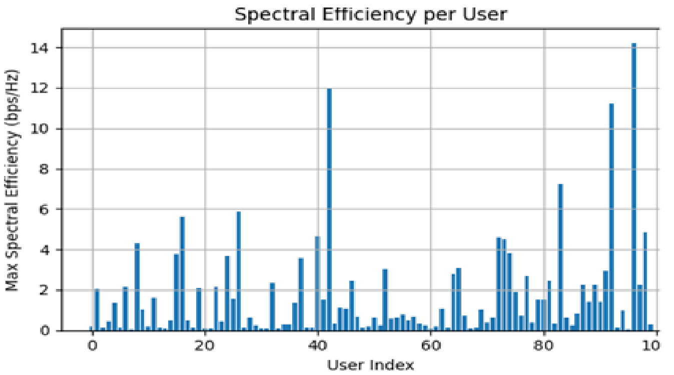

Figure 3 Simulation Results W.R.T Spectral Efficiency Per User

In Figure 3, the performance evaluation reveals a strong correlation between the SINR distribution and the corresponding spectral efficiency achieved by each user. The SINR results show a highly non-uniform distribution, with values ranging from approximately –10 dB to over 40 dB, indicating significant variations in channel quality and interference levels across the network. These fluctuations directly translate into the spectral efficiency outcomes, where only a small subset of users attain high throughput levels (exceeding 10 bps/Hz), while the majority operate below 3 bps/Hz. This disparity is attributed to the logarithmic dependence of SE on SINR, which compresses high SINR gains while severely penalizing low SINR conditions. Consequently, users experiencing weak coverage or strong interference exhibit near-zero spectral efficiency, revealing a pronounced fairness issue. Overall, the comparative behavior of SINR and SE confirms that the system is predominantly interference-limited and that improving the SINR of low-performing users would yield substantial gains in both average spectral efficiency and user experience. Spectral Efficiency Matrix Shape: (100, 50) Average Spectral Efficiency (bps/Hz): 0.03845265388707225 Average Inter-cell Interference:

International Research Journal on Advanced Engineering Hub
e ISSN: 2584-2137
Vol. 04 Issue: 01 January 2026
Page No: 361-367
https://irjaeh.com
https://doi.org/10.47392/IRJAEH.2026.0050

2.71532600514085e-09 Average Channel Interference: 1.648573442168753e-1 (Figure 4).

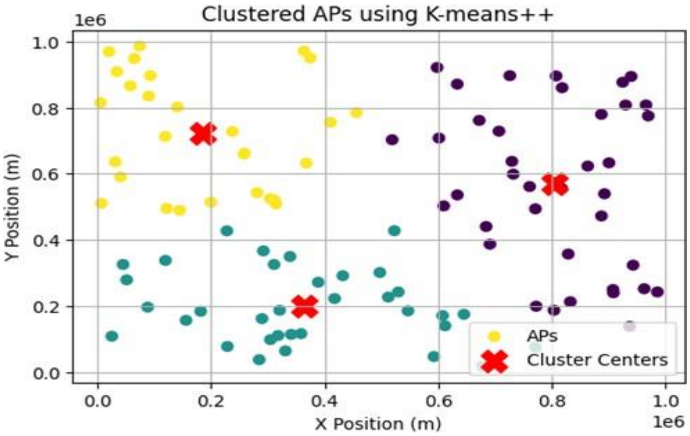

Figure 4 **Clustering of Aps Using K-Means++ Algorithm**

Prior studies have explored heuristic and greedy algorithms for AP-UE association. While these methods offer performance gains, they often lack scalability and suffer from high computational complexity. Recent clustering based techniques have shown promise in managing large scale network deployments. K-means ++ clustering with its enhanced centroid initialization provides a computationally efficient way to partition APs and UEs into optimal clusters (Table 2).

Table 2 **Spectral Efficiency Comparison**

Method	**Avg. SE (bps/Hz)**	**Sum Rate**
Random Assignment	**2.1**	105
Greedy Assignment	**3.5**	175
K-Means++ Proposed	4.4	220

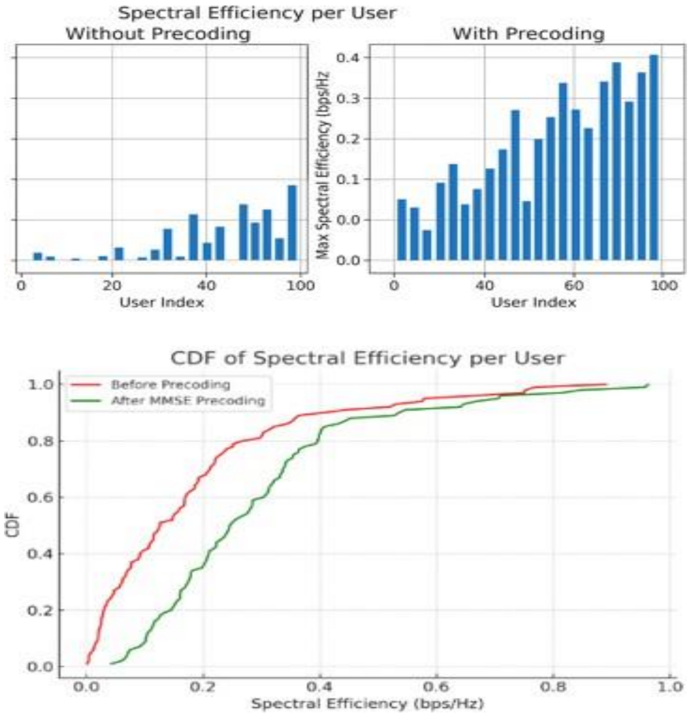

Figure 5 **Effect of MMSE Precoding on Per-User Spectral Efficiency and CDF Behavior**

In the Figure 5 the CDF plot comparing spectral efficiency before and after applying MMSE precoding is shown. It visually shows the gain in efficiency for most users, with the curve after precoding shifting rightward, indicating higher spectral efficiency.

Conclusion

The proposed integration of K-Means++ clustering with AP- UE assignment enhances spectral efficiency in CF-mMIMO networks. It effectively distributes network load, minimizes interference, and remains scalable. Future research will explore dynamic clustering under mobility conditions and adaptive learning-based optimization.

References

[1]. T. L. Marzetta, “Independent cellular wireless systems with an extensive array of base station antennas,” IEEE Transactions on Wireless Commu- nications, vol. 9, no. 11, pp. 3590–3600, Nov. 2010.
[2]. H. Yang and T. L. Marzetta, “A large-scale cellular network ensuring consistent user throughput,” in Proceedings of the IEEE

International Research Journal on Advanced Engineering Hub
e ISSN: 2584-2137
Vol. 04 Issue: 01 January 2026
Page No: 361-367
https://irjaeh.com
https://doi.org/10.47392/IRJAEH.2026.0050

Vehicular Technology Conference (VTC), Sep. 2014.
[3]. S. Zhou, M. Zhao, X. Xu, J. Wang, and Y. Yao, "A distributed frame- work for next-generation public wireless access," IEEE Communications Magazine, vol. 41, no. 3, pp. 108–113, Mar. 2003.
[4]. X. Zhu, Z. Wang, L. Dai, and C. Qian, "Intelligent pilot allocation for massive MIMO systems," IEEE Communications Letters, vol. 19, no. 9,pp. 1644–1647, 2015.
[5]. B. Fang, Z. Qian, et al., "RAISE: An efficient transmit antenna selection technique for massive MIMO systems," Wireless Personal Communications, vol. 80, no. 3, pp. 1147–1157, 2015.
[6]. J. Joung and S. Sun, "Two-phase transmit antenna selection strategies for IEEE standards," in Proceedings of IEEE International Conference, 2016.
[7]. F. Mrabti and A. A. Hassani, "A novel antenna selection approach for massive MIMO systems," 2016.
[8]. E. Bjornson, J. Hoydis, and L. Sanguinetti, Massive MIMO Networks: Optimization of Spectral, Energy, and Hardware Performance. Boston, MA, USA: NOW Publishers, 2017.
[9]. H. Q. Ngo, A. Ashikhmin, H. Yang, E. G. Larsson, and T. L. Marzetta, "Comparing cell-free massive MIMO with small cell architectures," IEEE Transactions on Wireless Communications, vol. 16, no. 3, pp. 1834–1850, 2017.
[10]. H. Tang and Z. P. Nie, "RMV: A novel antenna selection method for massive MIMO," IEEE Signal Processing Letters, vol. 25, no. 2, pp. 239–242, 2018.
[11]. Y. Gao, H. Vinck, et al., "Massive MIMO antenna selection: Explor- ing switching designs, capacity limits, and optimal algorithms," IEEE Transactions on Signal Processing, vol. 66, no. 5, pp. 1346–1360, 2018.
[12]. H. T. Dao and S. Kim, "Optimizing access point selection in cell-free massive MIMO using effective channel gain," IEEE Access, vol. PP, no. 99, pp. 1–1, Jun. 2020, DOI: 10.1109/ACCESS.2020.3001270.
[13]. H. Q. Ngo, L.-N. Tran, T. Q. Duong, M. Matthaiou, and E. G. Larsson, "Evaluating the total energy efficiency of cell-free massive MIMO networks," IEEE Transactions on Green Communications and Networking, vol. 2, no. 1, pp. 25–39, 2018.
[14]. I. L. Shakya and F. H. Ali, "Combined access point selection and interference mitigation in cell-free massive MIMO," IEEE Transactions on Communications, 2020.
[15]. G. Interdonato, H. Q. Ngo, P. Frenger, and E. G. Larsson, "Downlink training in cell-free massive MIMO: A blessing in disguise," IEEE Transactions on Wireless Communications, vol. 18, no. 11, pp. 5153–5169, 2019.
[16]. J. Zhang, S. Chen, Y. Lin, J. Zheng, and L. Hanzo, "Cell-free massive MIMO: A new next-generation paradigm," IEEE Access, vol. 7, pp. 99878–99888, 2019.
[17]. E. Nayebi, A. Ashikhmin, T. L. Marzetta, H. Yang, and B. D. Rao, "Precoding and power optimization in cell-free massive MIMO sys- tems," IEEE Transactions on Wireless Communications, vol. 16, no. 7, pp. 4445–4459, 2017.
[18]. O. T. Demir, E. Bjo¨rnson, and L. Sanguinetti, "Foundations of user- centric cell-free massive MIMO," Foundations and Trends in Signal Processing, vol. 14, no. 3–4, pp. 162–472, 2021.
[19]. A. Zakariya, A. Abdelhadi, and C. Clancy, "Access point clustering and user association in cell-free massive MIMO networks," IEEE Communications Letters, vol. 24, no. 2, pp. 329–333, 2020.
[20]. X. You, C. Zhang, X. Tan, S. Jin, and H. Q. Ngo, "AI for 6G: A cell-free massive MIMO framework with intelligent clustering," IEEE Wireless

International Research Journal on Advanced Engineering Hub
e ISSN: 2584-2137
Vol. 04 Issue: 01 January 2026
Page No: 361-367
https://irjaeh.com
https://doi.org/10.47392/IRJAEH.2026.0050

Communications, vol. 29, no. 1, pp. 30–37, 2022.

[21]. G. Interdonato, E. Bjo¨rnson, H. Q. Ngo, P. Frenger, and E. G. Lars- son, "Ubiquitous cell-free massive MIMO communications," EURASIP Journal on Wireless Communications and Networking, 2019.

[22]. L. D. Nguyen, T. Q. Duong, H. Q. Ngo, and K. Tourki, "Energy-efficient cell-free massive MIMO," IEEE Transactions on Green Communications and Networking, vol. 2, no. 1, pp. 25–39, 2018.

[23]. S. Buzzi and C. D'Andrea, "Cell-free massive MIMO: User-centric approach," IEEE Wireless Communications Letters, vol. 6, no. 6, pp. 706–709, 2017.

[24]. G. MacCartney and T. Rappaport, "Millimeter-wave cell-free massive MIMO: Distributed AP architecture at high frequencies," in Proc. IEEE ICC, 2020.