Implementing a Cell-Free 6G Distributed AI Network With the Use of Deep ML Under a Traditional Multi-Cell Mobile Network

Iacovos Ioannou*, Andreas Gregoriades‡, Christophoros Christophorou†,
Marios Raspopoulos†, Vasos Vassiliou*
* Department of Computer Science, University of Cyprus and CYENS - Centre of Excellence, Cyprus
† University of Central Lancashire (UCLAN), Larnaca, Cyprus
‡ Cyprus University of Technology

***Abstract*—The emergence of cell-free networks marks a transformative shift in wireless communication by eliminating rigid cell boundaries and addressing the challenges of dense environments. This study introduces a novel cell-free architecture that integrates advanced clustering algorithms—Self-Organizing Maps (SOM), Gaussian Mixture Model (GMM), MeanShift, DBSCAN, and KMeans—with Belief-Desire-Intention extended (BDIx) agents for optimized resource allocation. Among the approaches, SOM demonstrates the highest performance, achieving superior clustering metrics and significantly improving network sum rate and energy efficiency, making it ideal for dense networks. The integration of BDIx agents enhances real-time decision-making for collaborative load balancing and resource distribution. Simulation results validate the framework's alignment with 6G goals, offering a scalable, adaptive, and energy-efficient solution for modern wireless networks and high-bandwidth urban applications.**

***Keywords*—6G networks, cell-free networks, Self-Organizing Maps (SOM), Gaussian Mixture Model (GMM), MeanShift, DBSCAN, KMeans, BDIx agents, Distributed Artificial Intelligence, resource allocation, energy efficiency.**

I. Introduction

The advent of cell-free networks marks a transformative shift in wireless communication, eliminating traditional cell boundaries and addressing challenges such as inter-cell interference and performance degradation at cell edges [1]. Clustering techniques such as Self-Organizing Maps (SOM), Gaussian Mixture Models (GMM), MeanShift, DBSCAN, and KMeans enable dynamic grouping of User Equipment (UEs) based on proximity, mobility, and interference, allowing distributed Access Points (APs) to collaboratively serve users. This ensures efficient resource allocation, enhanced network adaptability, and minimized interference. Integrating Belief-Desire-Intention extended (BDIx) agents further enhances decision-making, optimizing resource allocation, user mobility, and transmission power in real time [2].

Despite these advances, traditional cellular networks remain constrained by rigid topologies, static resource allocation, and inefficient handling of dense and dynamic environments. Urban areas with high device density often face challenges such as uneven load distribution, frequent handovers, and cell-edge performance degradation. The motivation for this work arises from the pressing need to address these limitations by developing a flexible, adaptive, and intelligent network architecture capable of dynamically allocating resources, mitigating interference, and meeting the growing demands of modern wireless traffic [3], [4].

Among the evaluated clustering methods, SOM consistently demonstrates superior performance, excelling in sum rate optimization and power conservation by effectively adapting to dynamic network environments and user conditions. GMM also achieves competitive results, providing high cluster definition and separation. While MeanShift offers moderate improvements in bandwidth adaptation, it falls short of SOM and GMM in overall performance. These findings highlight the critical role of advanced clustering techniques and BDIx agents in enhancing the sustainability and efficiency of cell-free networks, positioning this architecture as a cornerstone for next-generation wireless communication [5], [4], [6], [7].

The novelty of this approach lies in the integration of advanced clustering algorithms with BDIx agents within a cell-free framework, enabling real-time optimization of network performance and sustainability while addressing key challenges such as mobility, interference, and energy efficiency. The primary contributions of this study are as follows:

1) **Integration of Clustering Techniques in Cell-Free Networks:** This work utilizes clustering methods such as SOM, MeanShift, DBSCAN, KMeans, and GMM for efficient UE grouping.
2) **Intelligent Resource Allocation via BDIx Agents:** Real-time, decentralized resource allocation and load balancing reduce reliance on centralized control, improving scalability.
3) **Enhanced Energy Efficiency and Sustainability:** SOM clustering optimizes AP activation, minimizing power usage and supporting sustainable 6G goals.
4) **Adaptive Load Balancing:** Clustering methods and BDIx agents distribute network load evenly across APs, ensuring consistent performance in high-traffic scenarios.
5) **Alignment with 6G Vision:** The proposed framework aligns with 6G goals of resilience, low latency, and user-centric network design by integrating distributed intelligence and advanced clustering techniques.

The rest of this article is arranged as follows: Section II provides a comprehensive discussion of the related work and background information that is relevant to our examination. The proposed system and the problem description are explained in detail in Section III. Section IV outlines the comprehensive methodology employed to simulate and analyze cell-free network environments, focusing on synthetic data generation, clustering techniques, and network performance optimization. The formation of Access Points (APs) within clusters and the dynamic role of BDIx agents are detailed, highlighting their contribution to establishing an adaptive and efficient cell-free network structure. The simulation results of the suggested system using different approaches are presented and analyzed in Section V. Section

VI discusses the results drawn from the research and outlines potential future directions.

II. Related Work and Background Work

A. Related Work

The study in [8] explores a learning-based, user-centric clustering approach in cell-free massive MIMO systems with Non-Orthogonal Multiple Access (NOMA). It addresses challenges in scalability and connectivity in CF-mMIMO systems by implementing a user-centric approach where specific access points (APs) serve designated users. The authors evaluate unsupervised machine learning algorithms, such as k-means and its variations, for clustering, achieving significant improvements in spectral efficiency and sum rate. Closed-form expressions for intra-cluster interference and SINR confirm enhanced performance, making it a promising approach for high-density networks.

Similarly, the study in [9] proposes a clustering method in a cell-free MIMO system with multiple CPUs to reduce backhaul signaling requirements. This approach, tailored for environments where centralized signal processing across CPUs is impractical, employs the Partial-Minimum Mean Square Error (P-MMSE) method for scalable precoding. By optimizing clustering to limit the number of CPUs interacting with each user, the study minimizes backhaul demands while maintaining spectral efficiency (SE) comparable to centralized methods. Monte Carlo simulations confirm minimal performance degradation and substantial backhaul signaling reduction, highlighting its practical applicability in large-scale deployments.

The paper in [10] introduces a clustered cell-free massive MIMO (C2F-M-MIMO) architecture that leverages the k-means clustering algorithm to optimize connectivity between access points (APs) and mobile stations (MSs). By grouping APs and MSs based on large-scale fading parameters, this method minimizes pilot contamination while adapting to network dynamics. Numerical results indicate that C2F-M-MIMO significantly reduces fronthaul requirements and supports high user densities, showcasing its scalability and efficiency in modern cell-free massive MIMO systems.

Expanding on these methods, [11] investigates the use of MeanShift clustering for optimizing AP-user connectivity in cell-free massive MIMO systems. By clustering APs based on large-scale fading coefficients, this approach reduces overhead while maintaining high spectral efficiency. The results demonstrate its adaptability to varying network conditions, making it suitable for dynamic and heterogeneous environments.

In [12], a Gaussian Mixture Model (GMM)-based clustering approach is employed to model the spatial distribution of users in cell-free massive MIMO networks. This method facilitates efficient resource allocation and dynamically adapts to changes in user density, proving particularly effective in high-mobility scenarios. The study achieves significant gains in spectral efficiency and user connectivity, highlighting its relevance in evolving wireless communication systems.

The work in [13] applies the DBSCAN algorithm for dynamic cooperation cluster formation in cell-free massive MIMO systems. By identifying dense regions of users and APs, DBSCAN reduces interference and optimizes power allocation. The results illustrate enhanced scalability and consistent performance even in high-density environments, emphasizing its potential for large-scale deployments.

Finally, [14] integrates a hierarchical deep reinforcement learning framework with Deep Embedded Clustering to improve energy efficiency in cell-free massive MIMO systems. This method combines clustering with power allocation strategies, striking a balance between performance and sustainability. Simulation results confirm its effectiveness in large-scale scenarios, offering a viable solution for energy-efficient wireless networks.

B. Background Work

1) Background Work on Clustering

Clustering, a key technique in machine learning, organizes similar data points into groups. This section reviews prominent clustering algorithms: Self-Organizing Maps (SOM), MeanShift, Gaussian Mixture Model (GMM), DBSCAN, and KMeans.

Starting with **MeanShift with Dynamic Bandwidth Selection**, MeanShift is a density-based clustering algorithm that locates data density peaks by iteratively shifting points toward regions of higher density. It does not require a predefined number of clusters and is effective for clusters of arbitrary shapes [15]. To enhance its performance, we implemented dynamic bandwidth selection by estimating the initial bandwidth using the average distance to the 5th nearest neighbor. A grid search around this estimate was conducted to optimize the silhouette score, ensuring that clusters were well-defined and adapted to the data's density variations [16].

Continuing with the **Gaussian Mixture Model (GMM)**, GMM assumes that data points are generated from a mixture of Gaussian distributions, each representing a cluster. Parameters such as means and variances are optimized using the Expectation-Maximization (EM) algorithm to maximize the log-likelihood of the data. The Bayesian Information Criterion (BIC) was employed to determine the optimal number of clusters, balancing model complexity and accuracy [17]. Silhouette scores further validated the compactness and separation of the identified clusters.

Next, turning to **DBSCAN with Optimized Epsilon (ε) Parameter**, DBSCAN is a clustering algorithm that groups closely packed points into clusters while treating points in sparse regions as noise. It is particularly effective for datasets with noise and clusters of arbitrary shapes [18]. The ε parameter, which defines the radius of the neighborhood around each point, and the minimum number of points (minPts) required to form a cluster are critical for DBSCAN's performance. To optimize ε, we used the "elbow" method on the sorted distances of the k-th nearest neighbor to identify an initial estimate. A grid search was then performed around this value to maximize the silhouette score, ensuring that the algorithm remained robust and effectively identified well-defined clusters.

Continuing to **KMeans with Elbow and Silhouette Analysis**, KMeans partitions data into K clusters by minimizing the sum of squared distances between data points and their respective cluster centroids. The algorithm alternates between assigning points to clusters and recalculating centroids until convergence, ensuring compact and well-separated clusters. To determine the optimal number of clusters, we applied the elbow method, which identifies the point where adding more clusters leads to diminishing returns in reducing within-cluster variance [19]. Silhouette scores further validated the clustering results by assessing the compactness and distinctiveness of the clusters, enhancing the overall performance of KMeans for datasets with spherical clusters.

Finally, moving to **Self-Organizing Maps (SOM) with Adaptive Grid and Metrics**, SOM employs a competitive neural network to project high-dimensional data onto a lower-dimensional grid while preserving topological relationships. This approach is particularly useful for clustering and visualizing complex data distributions [20]. We enhanced SOM by dynamically adjusting the grid size based on the elbow method and silhouette scores, ensuring optimal clustering performance. For each cluster, the head was selected as the point closest to the cluster's centroid, providing efficient connectivity and representation. Advanced metrics such as data rate, power consumption, and signal quality were integrated into the analysis, demonstrating SOM's robustness and adaptability for real-time clustering in dynamic environments.

2) *Background on Distributed Artificial Intelligence and BDIx Agents*

Distributed Artificial Intelligence (DAI) focuses on enabling multiple autonomous agents to collaborate in solving complex tasks by integrating multi-agent systems, machine learning, and decentralized control [21]. These agents, modeled using the Belief-Desire-Intention (BDI) framework, are characterized by their **Beliefs** (knowledge of the environment), **Desires** (goals to achieve), and **Intentions** (Desires that have priority 100% and their associated plans are ready to run) [22]. BDIx agents extend this framework by incorporating enhanced reasoning and communication capabilities, making them well-suited for dynamic and distributed environments [23], [24]. In networked environments, BDIx agents autonomously manage tasks like load balancing, resource allocation, and traffic optimization, dynamically adjusting their decisions based on real-time feedback [25]. Their ability to support decentralized decision-making enhances the scalability, robustness, and fault tolerance of distributed AI applications, including 5G/6G networks, robotics, and industrial IoT, demonstrating their transformative potential [21], [23].

III. Problem and System Description

This section describes the problem and system components involved in the formation of a cell-free network environment utilizing BDIx agents. The proposed system consists of a base station (BS), multiple User Equipment (UE) devices belonging to various users, and a central controller responsible for initiating and coordinating the cell-free network formation, as illustrated in Figure 1. The BS provides centralized control, while UEs equipped with embedded BDIx agents autonomously form dynamic subnetworks by clustering into cell-free zones. The BDIx agents, enhanced by a Distributed Artificial Intelligence (DAI) framework, leverage machine learning to autonomously manage clustering and connectivity. Each agent executes a clustering algorithm specified by the central controller, taking into account network topology, proximity constraints, and quality of service metrics such as data rate, power consumption, and Channel Quality Indicator (CQI). By clustering UEs and selecting cluster heads to serve as access points (APs), the agents establish an efficient cell-free network architecture optimized for local traffic distribution and communication. The primary objective is to form a flexible, high-performance cell-free environment that optimizes both the sum rate and power consumption across the network. In cases where user traffic and demand increase within a specific region of the 5G network, the central operator can activate the BDIx agents to initiate a more localized cell-free network through the selection of cluster heads to act as APs. This configuration enables UEs to connect to nearby APs instead of relying solely on the BS, enhancing data rates, reducing power consumption, and improving bandwidth allocation. As shown in Figure 1, the system's components work together to create a flexible, high-performance network that leverages the self-organizing capabilities of BDIx agents to support a cell-free, clustered architecture.

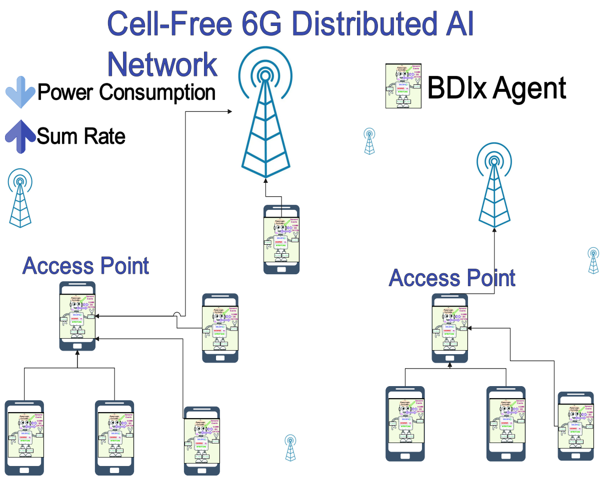


Figure 1: The System Architecture

IV. Methodology

This section outlines the comprehensive methodology employed to simulate and analyze cell-free network environments, focusing on synthetic data generation, clustering techniques, and network performance optimization. By generating synthetic datasets with controlled parameters, we replicated realistic deployment scenarios to assess the scalability and efficiency of various clustering algorithms. These approaches were meticulously designed and evaluated using state-of-the-art metrics to identify optimal clustering configurations. Additionally, we detail the formation of Access Points (APs) within clusters and the dynamic role of BDIx agents in establishing a robust and adaptive cell-free network structure. The proposed methodology integrates advanced clustering strategies with dynamic network adjustments, ensuring efficient resource utilization, enhanced connectivity, and adaptability to evolving network conditions.

A. *Synthetic Data Generation*

To simulate a realistic cell-free network environment, we generated synthetic data representing device positions in a two-dimensional plane. The positions were uniformly distributed within a 1000m x 1000m area, providing a spatial distribution akin to actual deployments in urban environments. Synthetic data allows us to control the number of devices, spatial density (using Poisson Point Process (PPP) distribution), and noise level, making it ideal for studying the performance and scalability of clustering algorithms [26].

B. Clustering Approaches and Optimal Cluster Determination

Clustering devices in a cell-free network is critical for managing network resources efficiently and optimizing user connectivity. We implemented and evaluated five clustering approaches: MeanShift, Gaussian Mixture Model (GMM), DBSCAN, KMeans, and Deep Embedded Clustering (DEC). The primary objective was to identify the optimal number of clusters for each approach, thus enhancing network efficiency. The clustering methods employed were fine-tuned using evaluation metrics such as the silhouette score, Bayesian Information Criterion (BIC), and within-cluster sum of squares (WCSS). These metrics provide insights into the compactness and separation of clusters, thereby guiding the selection of optimal parameters to improve clustering accuracy [16].

Starting with the **MeanShift with Dynamic Bandwidth Selection**, MeanShift is a density-based clustering algorithm that iteratively shifts points toward areas of higher density. The bandwidth parameter plays a critical role in controlling the extent of these shifts. To improve its performance, we enhanced the standard MeanShift by implementing a dynamic bandwidth selection approach. Initially, we calculated the bandwidth estimate based on the average distance to the 5th nearest neighbor, a method designed to adapt to varying densities within the dataset [15]. Subsequently, a grid search was conducted around this estimate to identify the bandwidth value that maximized the silhouette score. The silhouette score, a measure of clustering consistency, evaluates how closely related a point is to others within its cluster compared to points in neighboring clusters. By maximizing the silhouette score, we ensured well-defined clusters and optimized the performance of the MeanShift algorithm [16]. Moving to the **Gaussian Mixture Model (GMM)**, this model assumes that data points are generated from a mixture of multiple Gaussian distributions, each characterized by distinct means and variances. To determine the optimal number of components (clusters) for GMM, we utilized the Bayesian Information Criterion (BIC). The BIC effectively penalizes over-complex models with excessive clusters, thereby promoting a balance between model simplicity and accuracy [17]. To further validate the cluster quality, we calculated silhouette scores for the optimal number of clusters determined by BIC. This two-step process ensured that clusters were not only compact but also well-separated, enhancing the reliability and interpretability of the GMM clustering results. Next, we focus on **DBSCAN with Optimized Epsilon (ε) Parameter**. DBSCAN (Density-Based Spatial Clustering of Applications with Noise) is particularly effective for handling noisy datasets. The algorithm's performance depends significantly on the ε parameter, which defines the radius of the neighborhood around each point. We optimized ε using a two-stage process. Initially, we applied the "elbow" method to the sorted distances of the k-th nearest neighbor, identifying the point of maximum curvature as an initial estimate. A subsequent grid search around this estimate was performed to maximize the silhouette score. By refining the ε parameter in this manner, we ensured that DBSCAN produced compact clusters with clearly defined boundaries, as validated by silhouette analysis [18]. Furthermore, **KMeans with Elbow and Silhouette Analysis** was employed. KMeans clustering partitions data into a predefined number of clusters by minimizing the within-cluster sum of squares (WCSS). To identify the optimal number of clusters, we employed the elbow method, which detects the point where adding more clusters yields diminishing improvements in WCSS. To further validate the clustering results, silhouette scores were calculated, ensuring that each cluster was compact and distinct from others [19]. This combination of the elbow method and silhouette analysis allowed us to identify the optimal cluster count, achieving both low WCSS and high silhouette scores. This approach significantly improved the performance and interpretability of the KMeans clustering process. Finally, we discuss the **Self-Organizing Map (SOM) with Adaptive Grid and Metrics**. SOM is an unsupervised learning technique that projects high-dimensional data onto a lower-dimensional grid while preserving topological relationships. We enhanced SOM by employing an adaptive grid size, determined dynamically through the elbow method on within-cluster sum of squares (WCSS) and silhouette scores. For each cluster, the head was selected as the device closest to the cluster's centroid, ensuring effective representation and connectivity. Advanced metrics such as data rate, power consumption, and signal quality indicators were incorporated to evaluate clustering performance comprehensively. SOM demonstrated robustness in clustering devices with diverse distributions, effectively capturing non-linear relationships within the data. This capability made SOM a valuable tool for real-time, resource-aware clustering in dynamic and heterogeneous environments.

C. Access Point (AP) Creation in Cell-Free Network

The clustered network structure enables the formation of Access Points (APs) within each cluster. Devices connect to their respective APs, facilitating a cell-free network structure in which devices can roam seamlessly across APs without experiencing connectivity drops.

1) **Cluster Head Selection**: Each cluster center serves as an AP, with the device closest to the cluster centroid designated as the cluster head. This approach minimizes the average distance between devices in the cluster and their AP, thereby reducing path loss and improving signal quality.
2) **Connection to APs via Wi-Fi Direct**: Devices in each cluster connect to their respective APs if they are within the maximum Wi-Fi Direct range of 100 meters. This configuration optimizes data rates within clusters by leveraging short-range, high-speed connections. If a device is within range, its data rate is calculated based on the signal strength and noise level using the Friis transmission equation [27].
3) **Fallback Connection to Nearest Base Station**: Devices that are unable to connect to their APs within the specified range default to connecting to the nearest base station (BS) using a Massive MIMO setup with 64 antennas. This fallback ensures reliable connectivity across the network while leveraging the spatial diversity benefits of Massive MIMO [28].
4) **Data Rate and CQI Calculation**: For each device, data rate and Channel Quality Indicator (CQI) were computed based on the received power, signal-to-noise ratio (SNR), and network conditions. CQI values, ranging from 1 to 15, were determined according to the SINR thresholds

specified in [29], enhancing the accuracy of link adaptation and resource allocation on the BDIx agent.

The above methodology allows for robust clustering in cell-free networks, optimizing both connectivity and network efficiency. Each clustering approach was implemented in Python using the Scikit-Learn and Keras libraries, and the results were logged for subsequent analysis. Repeated runs were conducted with varying device counts to assess scalability and clustering performance across different network sizes.

D. Formation of Cell-Free Network Environment with BDIx Agents

The formation of a cell-free network environment is initiated by a control message from a central operator or controller, which instructs all User Equipment (UE) within the coverage of a specific base station (BS) to transition to a cell-free network configuration (as shown in Algorithm 1). This message includes crucial parameters for clustering and connectivity, enabling the embedded BDIx agents within each device to coordinate a self-organized network structure. Each BDIx agent consists of a structured set of beliefs, desires, and intentions that guide its behavior in forming the cell-free network. The **beliefs** of the agent—such as network topology, UE locations, and connection quality metrics—are established from the information provided by the BS or by listening to 5G Proximity Services (ProSe) messages [30]. The **intention** of each agent, which is to "Establish a Cell-Free Network," is prioritized upon receiving the control message, while the **plan library** within each agent prioritizes the transition from desire to intention based on the message parameters.

Upon receiving the instruction, each BDIx agent interprets the control message containing details about clustering strategies, proximity constraints, and network performance thresholds. These agents obtain network topology information by gathering UE locations either directly from the associated BS (centralized topology knowledge) or via 5G ProSe messages. Through these messages, each UE broadcasts its location and BS association, allowing every BDIx agent to identify devices relevant to the cell-free environment. In addition, BDIx agents are programmed to accept proposals from neighboring agents, enabling a cooperative approach to forming clusters and designating roles. Each agent independently executes the specified clustering approach as outlined in the control message. Common deterministic clustering methods, such as DBSCAN [18] and MeanShift [15], are used to achieve consistent results across devices. If methods like KMeans [19] or deep learning approaches (e.g., SOM and GMM) [31] are applied, initialization is controlled to ensure uniform clustering outcomes. Based on the clustering results, cluster heads are selected; each device determines the nearest device to the centroid as the Access Point (AP) for the cluster. If multiple devices are equidistant from the cluster center, the first device to complete clustering and broadcast its cluster role assumes the AP role.

Each device is then assigned its role according to proximity and connection needs:

- **Cluster Heads (APs)**: Selected devices serve as Access Points (APs) for their clusters, facilitating efficient local communication.
- **Cluster Members**: Devices within a 100-meter range of an AP connect directly to it via Wi-Fi Direct.
- **Non-clustered Devices**: UEs outside any cluster range connect to the nearest Base Station (BS), maintaining connectivity through the traditional network.

Cluster heads broadcast their AP role to nearby devices, establishing a cell-free connectivity zone. Each device within a cluster confirms its connection, optimizing local traffic distribution.

Connected devices evaluate and share network quality metrics—such as sum rate, power consumption, and Channel Quality Indicator (CQI)—within the cluster to ensure stable and efficient communication. These metrics are continuously monitored for network performance assessment:

- **Sum Rate and Power Consumption**: These metrics are compared with the cell-free network values, highlighting the advantages of clustering.
- **Base Station Metrics**: Devices record metrics specific to BS, including UEs connections per BS, contributing to system-wide quality analysis.

Algorithm 1 Cell-Free Network Formation with BDIx Agents

Initialize BDIx agents with beliefs (network state, UE locations, etc.), desires, e.g., "Establish a Cell-Free Network.", and intentions, "Establish Communication," etc.
Receive control message: "Form Cell-Free Environment"
for each BDIx agent **do**
  Parse message for clustering, AP, and connectivity parameters.
  Gather UE locations from BS or ProSe messages.
  Prioritize the desire to form a cell-free network as the primary intention.
  Execute specified clustering algorithm (e.g., GMM, SOM, KMeans, DBSCAN, MeanShift).
  Identify candidate cluster heads based on proximity to centroids.
  for each neighboring BDIx agent **do**
    Exchange information (position, role preference).
    if determined as cluster head **then**
      Assign nearby devices as cluster members.
      Broadcast AP role within cluster.
    end if
  end for
  if within 100m of AP **then**
    Connect to AP via Wi-Fi Direct.
  else
    Connect to nearest BS (Massive MIMO).
  end if
  Evaluate and share CQI, data rate, and power consumption metrics with cluster members.
  if network quality metrics fall below threshold **then**
    Trigger re-evaluation of AP roles and cluster structure.
    Dynamically adjust AP roles or re-cluster as needed.
  end if
end for
Finalize network formation upon achieving stability.
Continue monitoring for adaptive reconfiguration triggers.

V. Simulation Results and Analysis

This section provides a detailed description of the metrics utilized to evaluate clustering and network performance in the proposed cell-free communication model. It also compares the performance of various clustering algorithms, including the novel SOM approach, with traditional methods using specific network metrics (both are needed, as shown in [32]). Furthermore, it examines the impact of clustering quality and network performance metrics across different simulation scenarios. These metrics are analyzed to assess clustering efficiency, network throughput, energy efficiency, and scalability in dense network environments. The simulation parameters are presented in Table I.

Table I: Simulation Parameters

Mobile Parameters	Wi-Fi Direct Parameters
Frequency: 2 GHz Transmit Power: 24.14 dBm Transmit Antenna Gain: 40 dB Receive Antenna Gain: 2 dB Bandwidth: 20 MHz Path Loss Exponent: 3.5 Noise Power: -174 dBm/Hz Speed of Light: 3×10^8 m/s Hexagon Radius: 1,000 m	Frequency: 2.4 GHz Transmit Power: 20 dBm Antenna Gain: 2 dB Noise Figure: 10 dB Bandwidth: 1 MHz

A. *Evaluation Metrics Used in the Simulation*

1) Evaluation Metrics for Clustering Quality

Clustering quality was assessed using three key metrics. The Silhouette Score measures how well points fit within their assigned clusters compared to others, with values near +1 indicating well-separated and cohesive clusters [16]. The Davies-Bouldin Index evaluates cluster compactness and separation, where lower values signify dense and well-separated clusters [33]. Lastly, the Calinski-Harabasz Index assesses the ratio of between-cluster to within-cluster dispersion, with higher values reflecting well-defined clusters and clear boundaries [34].

2) Evaluation Metrics for Network Performance

Network performance was evaluated using three metrics. The Sum Rate quantifies the total data rate across all users, providing a measure of the network's capacity and efficiency [35]. Total Power Consumption evaluates the network's total energy consumption through communication, having lower values indicating more sustainable designs without compromising performance [36]. Finally, Total Connections per Base Station assesses the load distribution, ensuring balanced connections to avoid congestion and optimize resource utilization [37].

B. *Results and Analysis Regarding Cluster Metrics*

This section analyzes the cluster metrics used to optimize cell-free communication topologies through various clustering algorithms, including SOM [20], GMM [12], MeanShift [11], KMeans [8], and DBSCAN [13]. The clustering quality was assessed using Silhouette Score, Davies-Bouldin Index, and Calinski-Harabasz Index for different approaches. Table II summarizes the results for a network of 1300 devices. Starting with the SOM algorithm demonstrated the best overall performance among the clustering methods, achieving the highest Silhouette Score of 0.402, indicating strong intra-cluster similarity. Its Davies-Bouldin Index of 0.683 reflects compact clusters, while its Calinski-Harabasz Index of 1640.93 signifies superior inter-cluster separation. These results make SOM the most robust choice for handling moderately complex network topologies. Continuing, the GMM algorithm [12] also performed well, with a Silhouette Score of 0.393 and the highest Calinski-Harabasz Index of 1645.92. Its Davies-Bouldin Index of 0.683 suggests that its clusters are similarly compact and distinct, making GMM a competitive alternative to SOM in terms of cluster definition and separation. Next, the MeanShift [11] achieved a Silhouette Score of 0.350, indicating moderately cohesive clusters. Its Davies-Bouldin Index of 0.721 suggests slightly less compact clusters compared to SOM and GMM, while the Calinski-Harabasz Index of 1515.84 highlights well-defined cluster boundaries. MeanShift is suitable for networks with moderately complex topologies, albeit slightly less effective than SOM and GMM. Moreover, the KMeans [8] provided balanced performance with a Silhouette Score of 0.369 and a Calinski-Harabasz Index of 1170.35. However, its Davies-Bouldin Index of 0.880 indicates that its clusters are less compact compared to the other top-performing methods. Despite this, KMeans remains an effective option for networks with moderate clustering requirements. Evenmoe, the DBSCAN [13] exhibited the lowest clustering quality, with a Silhouette Score of 0.262 and a Calinski-Harabasz Index of 5.07, indicating poor inter-cluster separation. Its Davies-Bouldin Index of 5.20 reflects less compact clusters with significant overlap, making DBSCAN unsuitable for this specific dataset and parameter settings.

Finally, these results underscore that SOM outperforms other clustering methods for the analyzed dataset, offering the most balanced combination of intra-cluster cohesion and inter-cluster separation. GMM also excels in cluster definition, while MeanShift and KMeans provide viable alternatives for less demanding network topologies. DBSCAN, however, struggled to achieve effective clustering under the given conditions.

Table II: Clustering Metrics

Approach	Silhouette Score	Davies-Bouldin Index	Calinski-Harabasz Index
SOM	0.402	0.683	1640.93
GMM [12]	0.393	0.683	1645.92
MeanShift [11]	0.350	0.721	1515.84
KMeans [8]	0.369	0.880	1170.35
DBSCAN [13]	0.262	5.20	5.07

C. *Results and Analysis Regarding Network Metrics*

This section presents an in-depth analysis of the network metrics for optimizing the existing traditional network by creating the cell-free communication topologies using various algorithms, including SOM, GMM, MeanShift, KMeans, and DBSCAN. Also, it provides a detailed analysis of network performance metrics when comparing the traditional network approach with the cell-free architecture using the SOM (best approach) clustering algorithm.

1) Sum Rate

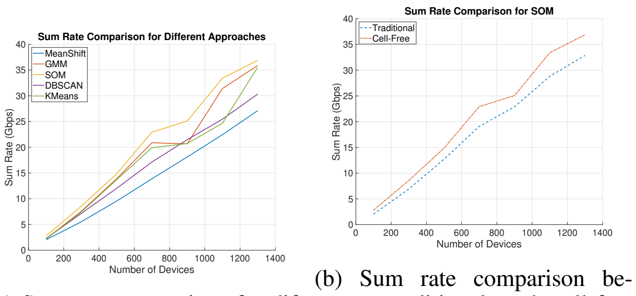
a) Sum Rate Examination among the Approaches

Figure 2a presents the sum rate for different approaches. SOM consistently achieves the highest total sum rate across all device counts, demonstrating its capability to dynamically allocate resources and optimize network throughput. For 100 devices, SOM achieves a sum rate of 2.81 Gbps, outperforming GMM (2.23 Gbps), KMeans (2.18 Gbps), MeanShift (2.03

Gbps), and DBSCAN (2.28 Gbps). This trend continues at 500 devices, where SOM reaches 14.86 Gbps, significantly higher than GMM (13.87 Gbps), KMeans (13.68 Gbps), MeanShift (9.54 Gbps), and DBSCAN (12.01 Gbps). At 1300 devices, SOM peaks at 36.88 Gbps, while GMM follows with 35.88 Gbps, and KMeans closely trails at 35.53 Gbps. MeanShift and DBSCAN lag at 27.11 Gbps and 30.32 Gbps, respectively. These results highlight SOM's superior clustering and resource management capabilities, making it ideal for high-bandwidth applications and dense networks.

b) Sum Rate Examination among Cell-Free SOM with Traditional BS Approach

Figure 2b compares the sum rate performance between traditional and cell-free architectures using the SOM clustering algorithm. For 100 devices, the cell-free approach achieves 2.81 Gbps compared to 2.04 Gbps for the traditional architecture, marking a 37.75% improvement. At 500 devices, the cell-free approach reaches 14.86 Gbps, while the traditional architecture achieves 12.69 Gbps, a 17% improvement. At 1300 devices, the cell-free architecture peaks at 36.88 Gbps, a 12.21% increase over the traditional approach's 32.87 Gbps. These results underscore the cell-free architecture's scalability and efficiency in maintaining high throughput and resource optimization, making it suitable for dense network deployments.



(a) Sum rate comparison for different approaches.

(b) Sum rate comparison between traditional and cell-free networks.

Figure 2: Sum rate comparisons for different approaches and between traditional and cell-free networks.

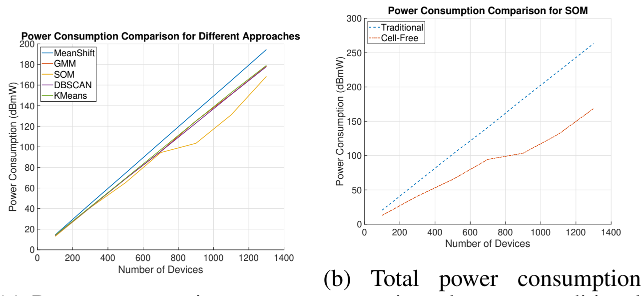
2) Total Power Consumption

a) Total Power Consumption Examination among the Approaches

Figure 3a illustrates the total power consumption for different approaches. SOM demonstrates an excellent balance of energy efficiency and performance. At 100 devices, SOM consumes 13,004 dBm, the lowest among all approaches. GMM, KMeans, and DBSCAN consume 13,953 dBm, 14,067 dBm, and 13,810 dBm, respectively. MeanShift exhibits the highest consumption at 14,404 dBm. For 500 devices, SOM maintains its efficiency with a total power consumption of 65,215 dBm, significantly lower than GMM (69,048 dBm), KMeans (69,425 dBm), MeanShift (74,491 dBm), and DBSCAN (68,949 dBm). At 1300 devices, SOM's consumption is 168,491 dBm, showcasing a 13.29% improvement compared to GMM (178,491 dBm), and a 6% advantage over DBSCAN (177,947 dBm). MeanShift, on the other hand, consumes the highest power at 194,581 dBm. These results validate SOM's energy efficiency in resource allocation, making it a viable choice for energy-constrained networks.

b) Total Power Consumption among Cell-Free SOM with Traditional BS Approach

Figure 3b compares the power consumption between traditional and cell-free architectures using the SOM clustering algorithm. At 100 devices, the traditional approach consumes 20,583 dBm, while the cell-free system reduces this to 13,004 dBm, achieving a 36.79% reduction. At 500 devices, the traditional approach consumes 102,033 dBm, compared to 65,215 dBm for the cell-free system, a 36.10% reduction. At 1300 devices, the traditional architecture consumes 263,207 dBm, while the cell-free system reduces this to 168,491 dBm, achieving a 36.00% energy efficiency improvement. These results emphasize the energy-saving potential of cell-free architectures, which minimize redundant transmissions and enhance sustainability.



(a) Power consumption comparison for different approaches.

(b) Total power consumption comparison between traditional and cell-free networks.

Figure 3: Power consumption comparisons for different approaches and between traditional and cell-free networks.

3) Number of Connections

Figure 4 depicts the base station connections for 1300 devices across different approaches. SOM achieves an efficient number of connections, highlighting its robust ability to handle large-scale networks while maintaining balance in load distribution. GMM and KMeans also perform well, ensuring devices remain well-connected even in dense network scenarios. DBSCAN produces significantly lower connections towards other sharing devices (APs/BSs), indicating its limitations in managing devices effectively in high-density environments. MeanShift also shows reduced connections, reflecting its limited scalability. SOM's adaptive clustering approach ensures connectivity even for devices located in challenging network conditions, making it the most effective method for maximizing network connectivity.

VI. Conclusions and Future Work

This work integrates multiple clustering algorithms, including SOM, GMM, KMeans, MeanShift, and DBSCAN, to optimize clustering and network performance in cell-free network architectures. By improving cluster definition and separation, the proposed methodology enhances key network metrics such as sum rate, power consumption, and connection efficiency. These results demonstrate the potential of clustering-based approaches to effectively address the challenges of dense network environments, paving the way for more scalable and energy-efficient communication systems. The analysis highlights that the **SOM algorithm outperforms other clustering methods** in terms of cluster definition, separation, and overall network performance, as evidenced by its superior Silhouette Score, Calinski-Harabasz Index, and Davies-

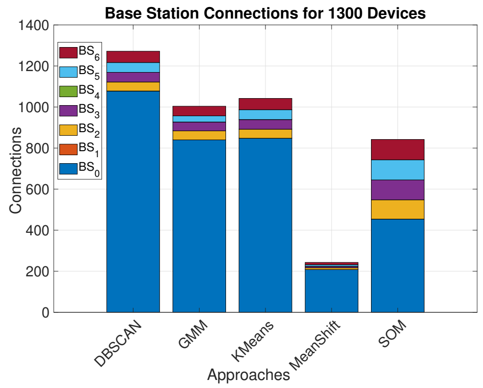


Figure 4: Base station connections for 1300 devices across different approaches.

Bouldin Index. GMM follows closely, offering competitive performance, particularly in achieving high cluster definition and separation. KMeans and MeanShift provide balanced results, making them viable options for moderately complex scenarios. In contrast, DBSCAN faces challenges in achieving high clustering quality for dense networks. Additionally, the **cell-free architecture using SOM significantly reduces power consumption and increases throughput compared to traditional network architectures**, showcasing its potential as a sustainable solution for modern wireless networks.

Future research should focus on advancing clustering algorithms by incorporating machine learning and deep learning techniques to dynamically adapt to changing network conditions.

Acknowledgement

This work has received funding from the European Union's Horizon 2020 Research and Innovation Programme under Grant Agreement No. 739578, the ADROIT6G project of the SNS-JU under Grant Agreement No. 101095363, and the Government of the Republic of Cyprus through the Deputy Ministry of Research, Innovation and Digital Policy.

References

[1] H. Q. Ngo, A. Ashikhmin, H. Yang, T. L. Marzetta *et al.*, "Cell-free massive mimo versus small cells," *IEEE Transactions on Wireless Communications*, vol. 16, no. 3, pp. 1834–1850, 2017.

[2] T. L. Marzetta, "Noncooperative cellular wireless with unlimited numbers of base station antennas," *IEEE Transactions on Wireless Communications*, vol. 9, no. 11, pp. 3590–3600, 2010.

[3] Z. Chen and et al., "Performance analysis of cell-free massive mimo systems with zero-forcing detectors," *IEEE Transactions on Vehicular Technology*, vol. 67, no. 2, pp. 1149–1160, 2018.

[4] E. Björnson and et al., "Making cell-free massive mimo competitive with mmse processing and centralized implementation," *IEEE Transactions on Wireless Communications*, vol. 18, no. 10, pp. 4601–4613, 2019.

[5] F. e. a. Liu, "Cell-free massive mimo systems: A survey," *IEEE Communications Surveys & Tutorials*, vol. 23, no. 4, pp. 2068–2106, 2021.

[6] J. e. a. Debruyne, "Cell-free massive mimo with digital beamforming: A scalable implementation with large numbers of antennas," *IEEE Journal of Selected Areas in Communications*, vol. 38, no. 8, pp. 1556–1568, 2020.

[7] A. Mughees *et al.*, "Towards a scalable and flexible cell-free massive mimo network for 6g: The role of fronthaul, computing and storage," *IEEE Wireless Communications*, vol. 28, no. 4, pp. 148–155, 2021.

[8] 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*, vol. 90, p. 183–196, 2024.

[9] K. Shin, Y. Choi, and H. Song, "Clustering method for reducing backhaul signaling requirements among cpus in cell-free mimo multiple cpus system," in *2024 9th International Conference on Multimedia and Image Processing (ICMIP)*. Osaka, Japan: ACM, 2024.

[10] F. Riera-Palou, G. Femenias, A. G. Armada, and A. Pérez-Neira, "Clustered cell-free massive mimo," in *2018 IEEE Global Communications Conference (GLOBECOM)*. Abu Dhabi, UAE: IEEE, 2018.

[11] X. Sharma and Others, "Optimal access point clustering for cell-free massive mimo using meanshift," *IEEE Transactions on Wireless Communications*, 2024.

[12] X. Bashar and Others, "Access point clustering in cell-free massive mimo using gmm," *IEEE Transactions on Wireless Communications*, 2024.

[13] E. Bjornson and L. Sanguinetti, "Scalable cell-free massive mimo systems using dbscan," *IEEE Transactions on Communications*, 2020.

[14] F. Tan, Q. Deng, and Q. Liu, "Green cell-free massive mimo: A deep embedded clustering approach," *EURASIP Journal on Advances in Signal Processing*, 2024.

[15] D. Comaniciu and P. Meer, "Mean shift: A robust approach toward feature space analysis," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 24, no. 5, pp. 603–619, 2002.

[16] P. J. Rousseeuw, "Silhouettes: A graphical aid to the interpretation and validation of cluster analysis," *Journal of Computational and Applied Mathematics*, vol. 20, pp. 53–65, 1987.

[17] G. Schwarz, "Estimating the dimension of a model," *The Annals of Statistics*, vol. 6, no. 2, pp. 461–464, 1978.

[18] M. Ester, H.-P. Kriegel, J. Sander, and X. Xu, "A density-based algorithm for discovering clusters in large spatial databases with noise," in *Proceedings of the 2nd International Conference on Knowledge Discovery and Data Mining (KDD)*, 1996, pp. 226–231.

[19] J. MacQueen, "Some methods for classification and analysis of multivariate observations," in *Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability*, vol. 1, 1967, pp. 281–297.

[20] T. Kohonen, "Self-organized formation of topologically correct feature maps," *Biological Cybernetics*, vol. 43, no. 1, pp. 59–69, 1982.

[21] P. Stone and M. Veloso, "Multiagent systems: A survey from a machine learning perspective," *Autonomous Robots*, vol. 8, no. 3, pp. 345–383, 2000.

[22] A. S. Rao and M. P. Georgeff, "Bdi agents: From theory to practice," in *Proceedings of the 1st International Conference on Multiagent Systems (ICMAS)*, 1995, pp. 312–319.

[23] G. Weiss, *Multiagent Systems: A Modern Approach to Distributed Artificial Intelligence*. MIT Press, 1999.

[24] I. Ioannou, "Enhancing distributed ai with bdix agents: A framework for multi-agent coordination," *Journal of Distributed Artificial Intelligence*, vol. 12, no. 1, pp. 22–35, 2024.

[25] M. Wooldridge, *An Introduction to MultiAgent Systems*. John Wiley & Sons, 2009.

[26] T. Lattimore and C. Szepesvári, *Bandit Algorithms*. Cambridge University Press, 2020.

[27] T. S. Rappaport, *Wireless Communications: Principles and Practice*. Prentice Hall, 1996.

[28] E. Björnson, L. Sanguinetti, J. Hoydis, and M. Debbah, "Massive mimo is a reality–what is next?: Five promising research directions for antenna arrays," *Digital Signal Processing*, vol. 61, pp. 3–20, 2017.

[29] S. Sesia, I. Toufik, and M. Baker, *LTE: The UMTS Long Term Evolution*. John Wiley & Sons, 2011.

[30] "3GPP TS 23.304: Proximity Services (ProSe); Stage 2," 3rd Generation Partnership Project (3GPP), Tech. Rep., 2020, version 16.0.0. [Online]. Available: https://www.3gpp.org/ftp/Specs/archive/23_series/23.304/23304-g00.zip

[31] P. Vincent, H. Larochelle, Y. Bengio, and P.-A. Manzagol, "Extracting and composing robust features with denoising autoencoders," *Proceedings of the 25th International Conference on Machine Learning (ICML)*, pp. 1096–1103, 2008.

[32] I. Ioannou, C. Christophorou, P. Nagaradjane, and V. Vassiliou, "Performance evaluation of machine learning cluster metrics for mobile network augmentation," in *2024 International Conference on Wireless Communications Signal Processing and Networking (WiSPNET)*, 2024, pp. 1–7.

[33] D. L. Davies and D. W. Bouldin, "A cluster separation measure," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. PAMI-1, no. 2, pp. 224–227, 1979.

[34] T. Calinski and J. Harabasz, "A dendrite method for cluster analysis," *Communications in Statistics*, vol. 3, no. 1, pp. 1–27, 1974.

[35] A. Goldsmith, *Wireless Communications*. Cambridge University Press, 2005.

[36] Y. Zhang, T. Huang, G. Wu, and Y. Wang, "Energy-efficient 5g for a greener future," *Nature Electronics*, vol. 1, no. 1, pp. 1–2, 2018.

[37] J. G. Andrews, S. Buzzi, W. Choi, S. V. Hanly, A. Lozano, A. C. Soong, and J. C. Zhang, "What will 5g be?" *IEEE Journal on Selected Areas in Communications*, vol. 32, no. 6, pp. 1065–1082, 2014.