Reinforcement Learning-Based Backhaul Routing for APs in Cell-Free Massive MIMO Networks

Junho Seo, Malik Muhammad Saad, Mahnoor Ajmal, Ayesha Siddiqa,
Bomi Jeong, Su Kim, and Dongkyun Kim*
School of Computer Science and Engineering, Kyungpook National University, Daegu, Republic of Korea
Emails: {jhseo, maliksaad, mahnoor.ajmal, asiddiqa, jbm, kimsu, dongkyun}@knu.ac.kr

***Abstract*—The integration of Distributed Access Points (APs) in cell-free massive MIMO networks poses challenges for efficient routing and parent selection, particularly in scenarios involving dense AP deployment. This paper introduces a reinforcement learning-based approach for optimal parent selection in Destination Oriented Directed Acyclic Graph (DODAG) based routing within cell-free massive MIMO networks. Unlike traditional routing protocols that rely heavily on static or computationally intensive methods, our framework enables each AP to act as an autonomous agent, dynamically selecting parent nodes based on throughput, congestion, hop count, and link stability. We model the problem using Q-learning, where each AP learns an optimal parent selection policy through interactions with the network. Simulation results demonstrate the proposed scheme's ability to maintain high throughput and packet delivery ratios even in dense network environments. The framework achieves an average throughput of 14.9 Mbps and a packet delivery ratio of 79% for a 50-node network, effectively addressing the challenges of interference and congestion. This work highlights the potential of reinforcement learning to enhance the scalability and performance of cell-free networks, paving the way for their application in future 6G communication systems.**

***Index Terms*—Cell-Free Massive MIMO, DODAG Routing, Backhaul, Reinforcement Learning, Parent Selection**

I. Introduction

Cell-free massive MIMO (CF-mMIMO) is an emerging technology poised to play a key role in future 6G telecommunication systems due to its ability to enhance spectral efficiency, expand coverage, and reduce user interference [1]. A typical CF-mMIMO system comprises a Central Processing Unit (CPU), Access Points (APs), and User Equipment (UEs). The CPU manages multiple APs, which are distributed across the network and connected to the CPU via wireless backhaul links. APs provide connectivity to UEs, with each UE potentially connecting to multiple APs. Since its inception, research on CF-mMIMO has addressed various topics, including AP selection, clustering, and beam combining [2] [3] [4].

However, most existing studies assume simplistic CF-mMIMO architectures, such as the one illustrated in Fig. 1, without considering scenarios involving a large number of APs. This limitation ignores the need for data forwarding between APs in densely deployed networks. In such cases, multi-hop data forwarding becomes critical, highlighting the need for research on routing protocols tailored to CF-mMIMO networks.

A dense CF-mMIMO architecture resembles the Integrated Access and Backhaul (IAB) architecture standardized in 3GPP TS 38.401 [5]. An IAB network consists of two main entities: the IAB-donor node, which manages IAB nodes via wireless links and forwards data to the core network, and the IAB nodes, which provide connectivity to UEs within their range and forward data to the IAB-donor node. Fig. 2 depicts a simplified IAB architecture. One study highlights the use of wideband mmWave access spectrum and resource sharing between wireless backhaul and access links, emphasizing the potential of IAB [6]. While CF-mMIMO lacks standardized routing protocols, the Backhaul Adaptation Protocol (BAP), under development for IAB, offers insights for CF-mMIMO routing [7]. However, BAP is still evolving, and details on inter-node routing remain undefined, necessitating novel approaches for CF-mMIMO routing.

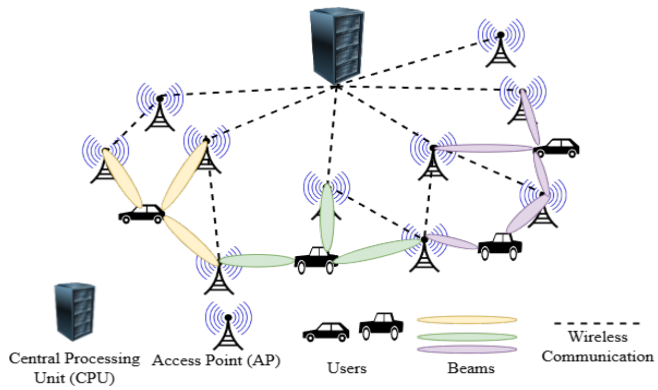

Fig. 1. CF-mMIMO Network Architecture

Routing in CF-mMIMO networks presents several challenges. Flooding-based routing protocols, commonly used in wireless networks, are resource-intensive and unsuitable for CF-mMIMO, where access and backhaul links share the frequency spectrum. Similarly, mesh routing protocols involve high computational and network overhead due to frequent control message exchanges, reducing application-layer throughput. A promising alternative is the use of DODAG (Destination Oriented Directed Acyclic Graph) topology-based routing, as employed in protocols like RPL (Routing Protocol for Low-Power and Lossy Networks) [8].

DODAG topology construction requires only two types of control messages: DODAG request messages, flooded by the root node (CPU in this case), and DODAG reply messages,

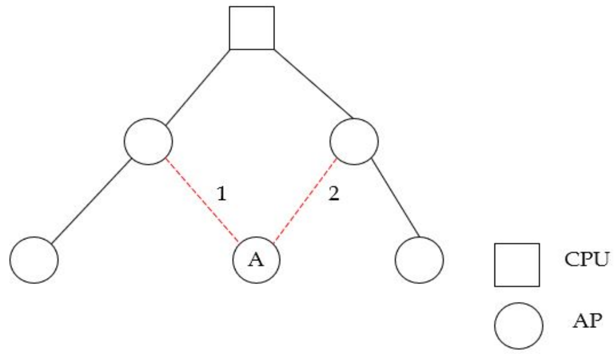

Fig. 2. Example of a problem scenario with DODAG applied in CF-mMIMO Networks

sent by nodes wishing to join the DODAG. The CPU floods request messages, and APs receiving them decide whether to join the DODAG. If an AP chooses to participate, it sends a reply to the upper node; otherwise, it ignores the request. However, in CF-mMIMO networks, single links between APs can limit throughput. When an AP receives multiple DODAG request messages from different parent candidates, it must select one parent based on specific routing metrics. Fig. 2 illustrates such a scenario, where AP_A receives two DODAG requests and must decide between Link 1 and Link 2 to construct the DODAG.

Given the complexity and dynamic nature of CF-mMIMO networks, traditional equation-based parent selection methods may not suffice to achieve optimal throughput. Machine learning approaches offer a viable solution by identifying optimal routing metrics to ensure high network performance and scalability. This study focuses on addressing the parent selection problem in CF-mMIMO networks, exploring machine learning-based approaches to optimize routing and guarantee network throughput for future applications.

In the remainder of this paper, Section II provides a comprehensive review of the relevant literature. Section III describes the system model in detail. Sections IV and V present the proposed scheme and the corresponding performance evaluation, respectively. Finally, Section VI concludes the paper with key findings and future research directions.

II. Literature Review

This section discusses recent research contributions and their limitations, highlighting the requirement for further advancement.

To address challenges in CF-mMIMO networks, researchers have explored a subgroup-centric multicast framework for mitigating pilot contamination [9] and a finite blocklength (FBL) transmission scheme to address traditional infinite blocklength assumptions [10]. The former enhances spectral efficiency through optimized pilot sharing and precoding techniques, while the latter develops low-complexity majorization-minimization algorithms and asymptotically optimal beamforming structures, achieving computational efficiency improvements. Despite their performance gains, both approaches require perfect channel state information (CSI) and specific channel models, limiting their applicability in realistic environments.

The federated learning-enhanced QoS multicast routing protocol (FLQMR) [11] combines federated learning (FL), reconfigurable intelligent surfaces (RIS), and edge computing into a cross-layer design to optimize multicast routing in IoT-enabled MANETs. By integrating physical layer data (e.g., spectral efficiency) with network layer information (e.g., route stability), this approach establishes more stable multicast trees. Simulation results indicate improvements in packet delivery ratios and reduced routing delays, but the protocol introduces significant computational complexity for participating devices.

CF-mMIMO research has also focused on two critical challenges: joint fronthaul load balancing with computation resource allocation under the O-RAN paradigm [12] and serving highly mobile IoT devices under imperfect CSI conditions [13]. The former optimizes fronthaul topology through quantization bit optimization at remote units (RUs) while satisfying computation constraints, whereas the latter develops advanced channel models incorporating line-of-sight (LoS) knowledge, non-isotropic scattering, and channel aging effects. These approaches assume perfect synchronization among access points and depend on idealized channel models, which may not hold in complex, dynamic network scenarios.

Recent advancements in CF-mMIMO systems address joint unicast and multi-group multicast transmissions [14] and energy-efficient beamforming with STAR-RIS integration [15]. The first utilizes an APG-based algorithm for joint access point selection and power optimization, achieving significant improvements in spectral efficiency. The second employs an AO-GIPG framework to jointly optimize active and passive beamforming with STAR-RIS, providing 360-degree coverage. Both solutions showcase significant enhancements in network performance but rely on perfect CSI, indicating opportunities for research in more robust, real-world scenarios with advanced precoding schemes.

The integrated access and backhaul (IAB) technique [6] addresses the challenges of mm-wave CF-mMIMO networks by optimizing bandwidth allocation and beamforming matrices to minimize end-to-end delays. While achieving notable improvements in network performance, this approach significantly increases the computational demands on access points (APs). Similarly, fronthaul queuing delay analysis [16] provides a scalable model and closed-form expressions for delay probabilities. The simulation results highlight lower queuing delays compared to small cell networks, but this comes at the cost of higher signaling overhead.

Full-duplex CF-mMIMO systems with practical limited-capacity fronthaul links [17] integrate full-duplex APs with randomly distributed half-duplex users in uplink and downlink communication to enhance throughput and energy efficiency. While this approach achieves throughput improvements and minimizes energy consumption, it imposes increased computational complexity on APs. Spatially correlated Rician fading channel analysis [18] develops power control algorithms for downlink CF-mMIMO systems, achieving significant gains in achievable rates through phase-aware and non-phase-aware minimum mean square error (MMSE) methods. However, the approach struggles with dense user deployments, where computational overhead grows rapidly.

Finally, integrated sensing and communication (ISAC) systems [19] propose a joint beamforming and power allocation framework, balancing sensing and communication performance. The setup achieves better signal-to-noise and interference-plus-noise ratios (SNR and SNIR) but introduces additional computational complexity on the fronthaul. These studies collectively highlight the need for further exploration of CF-mMIMO systems under realistic conditions, incorporating dynamic environments, advanced channel models, and efficient resource allocation schemes.

III. System Model

Consider a cell-free massive MIMO network consisting of M CPUs and N distributed access points (APs) deployed across a geographical area $\mathcal{A} \in \mathbb{R}^2$. Let $\mathcal{M} = \{1, 2, \ldots, M\}$ denote the set of CPUs and $\mathcal{N} = \{1, 2, \ldots, N\}$ represent the set of APs. Each AP can establish a connection either directly with a CPU or through other APs in a multi-hop configuration, forming a DODAG topology $\mathcal{G} = (\mathcal{V}, \mathcal{E})$, where $\mathcal{V} = \mathcal{M} \cup \mathcal{N}$ represents the set of all nodes and $\mathcal{E}$ denotes the set of established connections.

The wireless channel between any two nodes (i, j) experiences Rayleigh fading, characterized by the channel coefficient h_{ij} which is given by:

$$h_{ij} = \sqrt{\beta_{ij}} g_{ij} \tag{1}$$

where β_{ij} represents the large-scale fading coefficient including path loss and shadowing, and $g_{ij} \sim \mathcal{CN}(0, 1)$ denotes the small-scale Rayleigh fading component. The Signal-to-Noise Ratio (SNR) for a link between nodes i and j can be expressed as:

$$\gamma_{ij} = \frac{P_t |h_{ij}|^2}{\sigma^2} \tag{2}$$

where P_t is the transmit power and σ^2 is the noise variance. The achievable capacity C_{ij} for a link between nodes i and j is determined by Shannon's capacity formula:

$$C_{ij} = B \cdot \log_2(1 + \gamma_{ij}) \tag{3}$$

where B is the channel bandwidth and γ_{ij} is the SNR for the link.

A. DODAG Formation

The network topology formation follows a DODAG-based routing protocol, initiated through control message exchanges between CPUs and APs. CPUs, acting as root nodes, flood DODAG Information Request (DIR) messages throughout the network. Upon receiving these requests, APs make informed decisions about joining the topology based on three critical

Algorithm 1 DODAG Initialization and Parent Selection

Require: $\mathcal{N}$: Network nodes, $\mathcal{C}$: CPU nodes, $\mathcal{A}$: AP nodes, C_{max} : Max children per node, H_{max} : Max hop count, C_{th} : Min capacity threshold

procedure INITIALIZEDODAG($\mathcal{N}, \mathcal{C}$)
    for each $c \in \mathcal{C}$ **do**
        $c.hopCount \leftarrow 0$, $c.childCount \leftarrow 0$, $c.seqNum \leftarrow 0$
        BroadcastDIR(c)
    end for
    for each $a \in \mathcal{A}$ **do**
        $a.parentNode \leftarrow$ NULL, $a.hopCount \leftarrow \infty$, $a.childCount \leftarrow 0$
        $a.potentialParents \leftarrow \emptyset$
    end for
end procedure
procedure PROCESSDIRMESSAGE($node, dir$)
    $C_{np} \leftarrow B \cdot \log_2(1 + \text{SNR}(node, dir))$
    if $dir.seqNum \geq node.seqNum$ **and** $C_{np} \geq C_{th}$ **then**
        $\Phi \leftarrow$ Calculate-($C_{np}, dir_{childCount}, dir_{hopCount}$)
        $node.potentialParents \leftarrow node.potentialParents \cup \{dir.sourceID\}$
        UpdateParentMetrics($node, dir.sourceID, \Phi$)
    end if
end procedure
procedure SELECTPARENT(node)
    $bestParent, bestMetric \leftarrow$ NULL, $-\infty$
    for each $p \in node.potentialParents$ **do**
        if $p.childCount < X_{max}$ **and** Metric($node, p$) $> bestMetric$ **then**
            $bestParent, bestMetric \leftarrow p$, Metric($node, p$)
        end if
    end for
    if $bestParent \neq node.parentNode$ **and** $\Delta\Phi > \Delta\Phi_{th}$ **then**

        SendDIPMessage($node, bestParent$)
        UpdateNodeStatus($node, bestParent$)
    end if
end procedure
procedure CALCULATEMETRIC(C_{np}, X_p, H_p)
    return $w_1 \cdot \frac{C_{np}}{C_{\max}} + w_2 \cdot \left(1 - \frac{X_p}{X_{max}}\right) + w_3 \cdot \left(1 - \frac{H_p}{H_{max}}\right)$
end procedure
procedure UPDATENODESTATUS($node, newParent$)
    $node.parentNode \leftarrow newParent$, $node.hopCount \leftarrow newParent.hopCount + 1$
    $node.seqNum \leftarrow newParent.seqNum$
    BroadcastDIR($node$)
end procedure
procedure BROADCASTDIR(node)
    Broadcast($\{node.ID, node.hopCount,$
$node.childCount, node.seqNum\}$)
end procedure

parameters: achievable link capacity (C_{ij}), congestion levels (number of existing child nodes at potential parent nodes), and hop count (distance from the CPU).

Let $\mathcal{P}_n$ denote the set of potential parent nodes for AP n. For each AP $n \in \mathcal{N}$, the parent selection metric $\Phi_{n,p}$ for a potential parent $p \in \mathcal{P}_n$ is defined as:

$$\Phi_{n,p} = w_1 \frac{C_{np}}{C_{\max}} + w_2 \left(1 - \frac{X_p}{X_{\max}}\right) + w_3 \left(1 - \frac{H_p}{H_{\max}}\right) \quad (4)$$

where:

- C_{np} is the achievable capacity between AP n and parent p, normalized by $C_{\max}$, the maximum system-defined capacity.
- X_p represents the current number of child nodes connected to parent p, normalized by $X_{\max}$.
- H_p denotes the hop count from parent p to its serving CPU, normalized by $H_{\max}$.
- $\{w_1, w_2, w_3\}$ are non-negative weights satisfying $\sum_{i=1}^{3} w_i = 1$.

The topology optimization problem for each AP n can be formulated as:

$$\begin{aligned} \mathcal{P}_n^* = \arg\max_{p \in \mathcal{P}_n} \quad & \Phi_{n,p} \\ \text{subject to:} \quad & C_{np} \geq C_{\text{th}}, \\ & X_p \leq X_{\max}, \\ & H_p \leq H_{\max} \\ & |\mathcal{P}_n^*| = 1 \end{aligned} \quad (5)$$

where C_{th} is the minimum capacity threshold required to maintain reliable communication, and the constraint $|\mathcal{P}_n^*| = 1$ ensures that each AP selects exactly one parent, maintaining the DODAG property. The selected parent $p^* \in \mathcal{P}_n^*$ updates its child count as $X_{p^*} = X_{p^*} + 1$, and the joining AP updates its hop count as $H_n = H_{p^*} + 1$. This process continues iteratively until all APs establish their optimal connections, resulting in a topology that maximizes network throughput while satisfying the system constraints.

The detailed algorithm for DODAG initial connection setup and Parent Selection topology is summarized in Algorithm 1. The algorithm initializes with CPUs at hop count 0 and APs at infinity, where each node maintains sequence numbers for message freshness and monitors parameters including SNR, hop count, and congestion levels. Nodes establish and maintain parent-child relationships through DIR message processing and parent selection metrics.

IV. Proposed Scheme

This paper presents a reinforcement learning (RL) based approach for parent selection in DODAGs to enhance overall network throughput. The network topology consists of a CPU as the root node and multiple connected nodes forming an acyclic graph. Each node in the network is modeled as an RL agent, learning to optimize its parent selection autonomously based on local observations and network-wide metrics. The primary goal of this approach is to address challenges such as congestion, link instability, and suboptimal routing by enabling nodes to make dynamic and distributed decisions that maximize the overall network performance.

A. State Representation

Each node captures the environment's state based on its observations of the network. The state vector for a node i, denoted as S_i, includes several parameters for each potential parent node j. These parameters are the ($\text{SNR}_{i,j}$) between node i and potential parent node j, which represents the channel quality; the hop count (H_j) from the parent node j to the CPU, which reflects the distance of the parent from the root; and the congestion level (C_j) at the parent node j, measured in terms of queue length or load. Additionally, the state includes the stability ($S_{i,j}$) of the link between node i and parent node j, which is derived from historical success rates of the connection, and the achievable throughput ($T_{i,j}$) between node i and parent node j. The throughput is calculated as:

$$T_{i,j} = B \cdot \log_2(1 + \text{SNR}_{i,j}), \quad (6)$$

where B is the bandwidth available for communication. This comprehensive state representation enables each node to evaluate its parent selection options effectively and dynamically.

B. Action Space

The action space for each node represents the selection of one parent from its set of potential parent nodes. For a node i with n potential parents, the action space is:

$$A_i = \{0, 1, ..., n-1\} \quad (7)$$

where each action $a \in A_i$ corresponds to choosing a specific parent node j.

C. Reward Function

The reward function is designed to guide the RL agents toward actions that maximize throughput while minimizing hops and avoiding highly congested or unstable parents. For a child node i selecting a parent node j, the reward $R_{i,j}$ is calculated as:

$$R_{i,j} = \alpha \cdot T_{i,j} - \beta \cdot H_j - \gamma \cdot C_j + \delta \cdot S_{i,j} \quad (8)$$

where $T_{i,j}$ represents the throughput between node i and parent j as defined above, H_j is the hop count from the parent j to the CPU, C_j represents the congestion level at parent j, and $S_{i,j}$ is the stability of the link between i and j. The terms α, β, γ, and δ are weighting factors that balance the impact of each term, satisfying $\alpha + \beta + \gamma + \delta = 1$. This reward structure incentivizes nodes to select parents that provide high throughput, fewer hops, less congestion, and greater link stability, aligning with the overall goal of maximizing network throughput.

Q-learning, a model-free reinforcement learning algorithm, is used for this scheme. Each node maintains a Q-value table $Q(S, A)$, where S is the current state and A is the action (parent selection). The Q-value represents the expected cumulative reward for taking a specific action in a given state.

During training, the Q-value is updated iteratively using the Bellman equation:

$$Q(S_t, A_t) \leftarrow Q(S_t, A_t) + \eta \cdot \Big[R_t + \gamma \cdot \max_{A} Q(S_{t+1}, A) - Q(S_t, A_t)\Big] \tag{9}$$

where η is the learning rate, γ is the discount factor for future rewards, and R_t is the immediate reward observed after taking action A_t in state S_t. By iteratively refining the Q-values, each node learns an optimal parent selection policy that balances exploration and exploitation, adapting to dynamic network conditions.

V. Performance Evaluation

The simulation is conducted to evaluate the proposed reinforcement learning-based parent selection approach in a DODAG network. The network comprises a single root node (CPU) and multiple connected nodes forming an acyclic graph. The performance of the system is analyzed for varying network sizes, ranging from 10 to 50 nodes.

The nodes are distributed randomly within the network, and their communication follows the DODAG topology. Each node acts as an independent RL agent, using Q-learning to optimize its parent selection policy. The agents interact with the environment by selecting parent nodes based on their local observations, aiming to maximize the overall network throughput. The environment provides feedback in the form of a reward, which considers factors such as signal-to-noise ratio (SNR), hop count, congestion, and link stability.

The machine learning framework for the simulation leverages Q-learning, a model-free reinforcement learning algorithm. The agents are initialized with a learning rate (η) of 0.1, which controls the speed at which the Q-value table is updated during training. A discount factor (γ) of 0.9 is used to balance immediate and future rewards, ensuring the agents consider the long-term impact of their decisions. The exploration-exploitation trade-off is managed through an epsilon-greedy strategy, where the exploration rate (ϵ) starts at 1.0, allowing the agents to explore all possible actions fully, and gradually decays to 0.1 over the episodes, favoring exploitation of learned policies as the training progresses. Each simulation runs for 50 episodes to ensure convergence of the Q-values.

Additionally, the RL environment is structured to reflect realistic network dynamics. The nodes receive state observations that include SNR, hop count, congestion, and stability metrics. The reward function integrates these metrics, with throughput (T) calculated using (6), where B is the bandwidth. Adjustments for interference and congestion are incorporated by penalizing effective SNR and increasing congestion penalties as the number of nodes grows. This ensures that the simulation captures the complexities of real-world network conditions and evaluates the proposed framework under both ideal and congested scenarios.

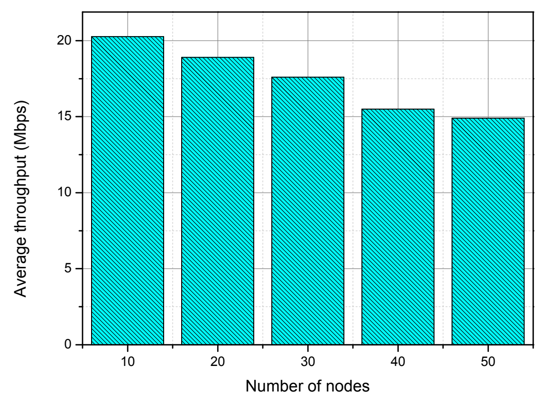

Fig. 3. Average Throughput

Fig. 3 illustrates the average throughput per node as the number of nodes in the network increases. The throughput shows a slight decline as the network size grows, decreasing from 20.26 Mbps for 10 nodes to 14.9 Mbps for 50 nodes. This modest reduction indicates that the proposed reinforcement learning-based scheme effectively adapts to network conditions, maintaining satisfactory performance even as the network becomes denser. The decline in throughput can be attributed primarily to increased interference and congestion in larger networks. As the number of nodes grows, the likelihood of overlapping transmissions increases, reducing the effective SNR for some links. Additionally, with more child nodes competing for limited resources, parent nodes experience higher congestion, which further impacts throughput. The penalty for higher hop counts to the root node also contributes to this reduction in performance. Despite these challenges, the results demonstrate that the proposed scheme successfully mitigates these effects through intelligent parent selection and adaptive learning. The overall performance remains satisfactory, highlighting the robustness of the reinforcement learning framework in handling real-world network complexities.

Fig. 4 illustrates the Packet Delivery Ratio (PDR) as the number of nodes in the network increases. The PDR shows a gradual decline from 92% for 10 nodes to 79% for 50 nodes. This decrease is expected as the network becomes denser, leading to increased interference and resource contention among nodes.

The decline in PDR can be attributed to the higher probability of packet collisions and transmission failures in larger networks. As the number of nodes increases, overlapping transmissions cause interference, which impacts the reliability of links. Additionally, congestion at parent nodes due to a larger number of connected child nodes can result in packet drops. The penalty for increased hop counts also contributes to the slight reduction in PDR, as nodes farther from the root face a higher risk of packet loss.

Despite these challenges, the proposed reinforcement learning-based parent selection scheme demonstrates strong

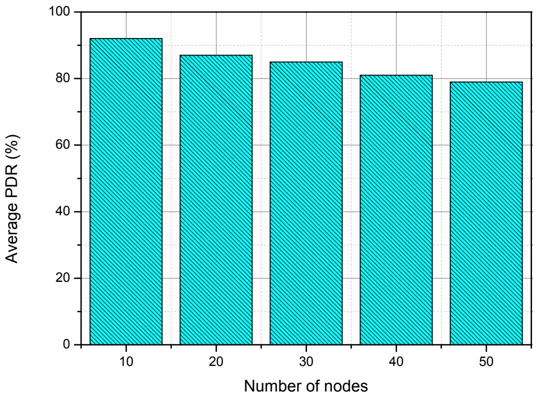

Fig. 4. Average PDR

adaptability, maintaining a high PDR even in larger networks. The system effectively mitigates the impact of interference and congestion, ensuring reliable packet delivery under varying network conditions.

VI. Conclusion

This paper presented a reinforcement learning-based parent selection scheme for DODAG routing in cell-free massive MIMO networks, addressing the challenges of dense AP deployments. By modeling each AP as an autonomous Q-learning agent, the framework optimally balances throughput, congestion, hop count, and link stability in parent selection. Simulations revealed that the proposed method sustains high network performance under varying conditions, achieving an average throughput of 14.9 Mbps and a packet delivery ratio of 79% in a 50-node network. The decline in performance with increased network size is mitigated through adaptive learning, which dynamically optimizes decisions in response to interference and congestion. Compared to traditional static routing protocols, the reinforcement learning-based approach demonstrates superior adaptability and efficiency. This study underscores the importance of machine learning in addressing real-world challenges in next-generation networks, particularly for enabling scalable and robust routing mechanisms. Future work will explore integrating additional metrics, such as energy efficiency and latency, to further optimize network performance in diverse operational scenarios.

Acknowledgment

This work was supported in part by the Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education under Grant NRF-2018R1A6A1A03025109. In addition this research has supported by the National Research Foundation of Korea (NRF) grant funded by the Korea government (MSIT), (NRF-2022R1A2C1003620).

References

[1] T.-N. Tran and B. An, "QoS Multicast Routing Utilizing Cross-Layer Design for IoT-Enabled MANET in RIS-Aided Cell-Free Massive MIMO," *IEEE Internet of Things Journal*, vol. 11, no. 7, pp. 11876–11893, 2024.

[2] H. T. Dao and S. Kim, "Effective Channel Gain-Based Access Point Selection in Cell-Free Massive MIMO Systems," *IEEE Access*, vol. 8, pp. 108127–108132, 2020.

[3] J. Wang, L. Dai, L. Yang, and B. Bai, "Clustered Cell-Free Networking: A Graph Partitioning Approach," *IEEE Transactions on Wireless Communications*, vol. 22, no. 8, pp. 5349–5364, 2023.

[4] S. K. Singh and A. K. Sah, "Distributed Beam combining in Cell-Free Massive MIMO Networks," *IEEE Transactions on Communications*, pp. 1–1, 2024.

[5] 3GPP, "NG-RAN; architecture description," Tech. Rep. 3GPP TS 38.401.

[6] A. H. Jazi, S. M. Razavizadeh, and T. Svensson, "Integrated Access and Backhaul (IAB) in Cell-Free Massive MIMO Systems," *IEEE Access*, vol. 11, pp. 71658–71667, 2023.

[7] 3GPP, "5G; NR; Backhaul Adaptation Protocol (BAP) specification," Tech. Rep. 3GPP TS 38.340 version 18.1.0 Release 18.

[8] RFC, "RPL: IPv6 Routing Protocol for Low-Power and Lossy Networks," Tech. Rep. 6550.

[9] A. De La Fuente, G. Femenias, F. Riera-Palou, and G. Interdonato, "Subgroup-Centric Multicast Cell-Free Massive MIMO," *IEEE Open Journal of the Communications Society*, vol. 5, pp. 6872–6889, 2024.

[10] X. Zhang, L. Xiang, J. Wang, and X. Gao, "Massive MIMO Multicasting With Finite Blocklength," *IEEE Transactions on Wireless Communications*, vol. 23, no. 10, pp. 15018–15034, 2024.

[11] A. Amalia, Y. Pramitarini, R. H. Yoga Perdana, K. Shim, and B. An, "Federated Learning-Enhanced QoS Multicast Routing to Support RIS and Edge Computing in IoT-Enabled MANETs with CF-mMIMO," in *2024 International Technical Conference on Circuits/Systems, Computers, and Communications (ITC-CSCC)*, pp. 1–6, 2024.

[12] Z. 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," *IEEE Transactions on Wireless Communications*, vol. 23, no. 10, pp. 14125–14139, 2024.

[13] H. Li, Y. Wang, C. Sun, and Z. Wang, "User-Centric Cell-Free Massive MIMO for IoT in Highly Dynamic Environments," *IEEE Internet of Things Journal*, vol. 11, no. 5, pp. 8658–8675, 2024.

[14] M. S. Abbas, Z. Mobini, H. Q. Ngo, and M. Matthaiou, "Unicast-multicast cell-free massive mimo: Gradient-based resource allocation," *arXiv preprint arXiv:2410.03506*, 2024.

[15] D. Choudhury and A. K. Sah, "Energy Efficient Joint Beamforming Design for Hardware-Impaired STAR-RIS Assisted CF-mMIMO Systems," *IEEE Transactions on Communications*, pp. 1–1, 2024.

[16] A. V. T and J. Francis, "Analysis of Fronthaul Delays in Uplink of Cell-free Massive MIMO and Small Cell Networks," in *ICC 2024 - IEEE International Conference on Communications*, pp. 49–54, 2024.

[17] S. Kurma, K. Singh, P. K. Sharma, C.-P. Li, and T. A. Tsiftsis, "On the Performance Analysis of Full-Duplex Cell-Free Massive MIMO with User Mobility and Imperfect CSI," *IEEE Transactions on Communications*, pp. 1–1, 2024.

[18] Y. Zhang, W. Xia, H. Zhao, Y. Zhu, W. Xu, and W. Lu, "Performance Analysis of Cell-Free Massive MIMO-URLLC Systems Over Correlated Rician Fading Channels With Phase Shifts," *IEEE Transactions on Wireless Communications*, vol. 23, no. 9, pp. 12206–12222, 2024.

[19] U. Demirhan and A. Alkhateeb, "Cell-free isac mimo systems: Joint sensing and communication beamforming," *IEEE Transactions on Communications*, pp. 1–1, 2024.