Resource Allocation leveraging Dual Traffic and Channel Predictions in Open-RAN Architecture

Yeonghun Jeong*, Dongho Ham†, Minhyun Kim‡, Jungmo Moon‡, and Jeongho Kwak*
*Dept. of Computer Science and Engineering, Korea University, Seoul, Korea
†Dept. of Electrical Engineering and Computer Science, DGIST, Daegu, Korea
‡Dept. of Intelligent RAN SW Research Section, Electronics and Telecommunications Research Institute, Daejeon, Korea
Email: {dudgns6648, jeonghokwak}@korea.ac.kr, dhham97@dgist.ac.kr, {minhyun.kim, jmmoon}@etri.re.kr

***Abstract*—In this paper, we first envision an agentic-AI framework for open-RAN architecture that enables autonomous and context-aware network control by integrating predictive intelligence into key resource-management loops. Building upon this architectural vision, we develop a dynamic resource-allocation mechanism that leverages dual traffic and channel predictions to guide power-budget decisions across cells. By jointly forecasting cell-level traffic demand and channel conditions using transformer-based AI models, the proposed mechanism acquires predictive knowledge of spatial–temporal cell environments, allowing power resources to be proactively aligned with expected demand. Within this framework, we design an adaptive power budget allocation algorithm that dynamically distributes the total transmit-power budget across cells based on anticipated traffic load and channel quality. Simulation results demonstrate that this predictive strategy achieves a 27.1% improvement in Geometric Average Throughput (GAT) compared to a baseline scheme without predictive information, confirming the effectiveness of embedding predictive and agentic intelligence into open RAN resource control.**

***Index Terms*—Open RAN, resource allocation, power budget, transformer**

I. Introduction

Cellular operators are already confronted with increasingly dense network deployments as they strive to accommodate the relentless growth in high data rate mobile services. While a large number of base stations (BSs) have been deployed across heterogeneous multi-cell networks, the infrastructure is rapidly approaching saturation. Simply deploying additional BSs is no longer a sustainable long-term solution due to physical site constraints, infrastructure costs, backhaul bottlenecks, and the escalating energy footprint associated with dense deployments.

As a software-centric alternative, inter-cell resource sharing and traffic-aware load balancing have been extensively studied [1]–[3], with the goal of directing additional resources toward congested cells to alleviate load imbalance. In parallel, advanced interference management techniques, such as coordinated multipoint (CoMP) transmission and interference-aware power control, have been developed to mitigate inter-cell interference [7]. Building on these ideas, a number of studies have explored dynamic resource-sharing mechanisms designed to maximize network utility or enhance spectral and energy efficiency [5], [6].

Despite this progress, existing approaches generally overlook the explicit characterization and exploitation of cell environments, such as spatial–temporal traffic variations or channel-quality trends. Without incorporating these structural properties, they fail to harness valuable predictive cues that could significantly enhance network-wide performance. Moreover, the fundamental trade-off between throughput and fairness is often treated implicitly, leaving room for more principled solutions that can balance competing objectives based on environmental knowledge.

Motivated by these gaps, we take a fundamentally different perspective: we aim to characterize cell environments and leverage them within an agentic-AI–driven open-RAN architecture that enables predictive and autonomous resource control. Within this framework, open-RAN components continuously observe, forecast, and reason about per-cell dynamics, creating a closed-loop system in which policies can adapt proactively to future conditions rather than reactively to immediate states.

As a concrete realization of this framework, we develop a prediction-assisted power-budget allocation mechanism in which the maximum transmit power of each cell serves as the primary control variable. By jointly forecasting cell traffic demand and channel conditions using transformer-based AI models, the controller gains predictive knowledge of how loads and channel qualities evolve across neighboring cells. This predictive information enables the coordinated distribution of per-cell power budgets under a fixed total power budget, thereby improving both user fairness and overall throughput.

In summary, our work moves beyond conventional static or myopic resource-allocation schemes by embedding predictive intelligence into an agentic-AI open-RAN control loop. Through this integration, the proposed framework captures structural differences across cells, exploits cross-cell dependencies, and systematically aligns resource supply with anticipated demand, ultimately achieving significant network-wide performance gains.

The main contributions of this paper are as follows.

- We introduce an agentic-AI system for open-RAN architecture that enables autonomous, predictive, and closed-

This work was supported by Institute of Information & Communications Technology Planning & Evaluation (IITP) grant funded by the Korea government (MSIT) (RS-2023-00225468, Development of RAN Intelligent Controller for O-RAN intelligence).

loop resource control by continuously observing and forecasting per-cell environments.

- Building upon this framework, we design a power budget allocation algorithm that jointly leverages transformer-based traffic and channel predictions to proactively align per-cell power budgets with anticipated network conditions.
- Through extensive simulations, we show that the proposed predictive strategy *(i)* adapts power budgets to heterogeneous cell environments, *(ii)* consistently outperforms baseline non-predictive approaches in terms of throughput and fairness, and *(iii)* benefits significantly from improved prediction accuracy.

In the rest of this paper, we begin with introducing the agentic-AI system for open-RAN architecture in Section II. Then, we define system model and problem statement in Section III. Next, we provide the power budget allocation algorithm in Section IV. Then, in Section V, we provide simulation results of the proposed algorithm. Finally, we conclude this paper in Section VI.

II. Agentic-AI System for Open-RAN Architecture

In this section, we propose an agentic-AI framework for open-RAN architecture as follows.

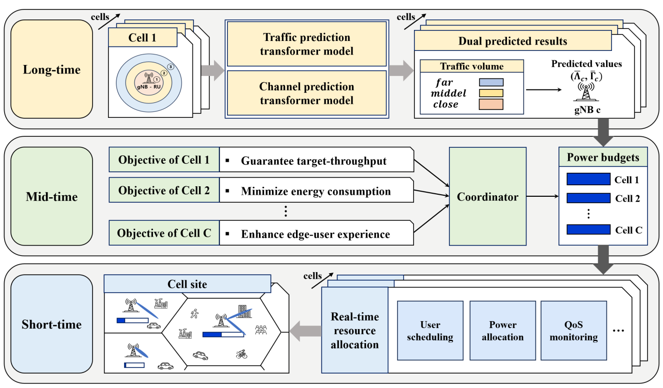

Fig. 1: A proposed hierarchical system in open-RAN architecture.

A. Agentic-AI System for Open RAN architecture

First, we define an agentic-AI system as an AI-native controller that exhibits a higher degree of autonomy, situational awareness, and adaptability than conventional rule-based network management. Each agent is capable of setting its own goals and making context-aware decisions while coordinating with other agents through common goal.

We consider a hierarchical control framework, as shown in Fig. 1, aligned with the non-RT RIC, near-RT RIC, and O-DU components. High-level policies and long-term learning are performed by rApps running on the non-RT RIC. Meanwhile, xApps, running on the near-RT RIC, operate as agents who own their cell-specific objectives and common goal at the same time. The O-DU executes slot-level scheduling and power allocation under the decisions received from the near-RT RIC, and provides feedback for both RICs.

B. Transformer-Based Prediction at the Non-RT RIC

At the non-RT RIC, we exploit abundant computational resources to train transformer-based prediction models for both channel and traffic states. Past time-series samples of user channels and traffic loads which are collected from the network are used to learn models that predict future cell-level environments over a given horizon.

For channel prediction, we distinguish between small-scale channel fading and large-scale channel fading components.

- Small-scale channel fading prediction is used to capture rapid channel fluctuations in interference-limited regions, especially for users far from their serving BS where the instantaneous gain is small and highly variable. Accurate prediction in such regimes enables more effective resource allocation when inter-cell interference is severe and channel variation is large.
- Large-scale channel fading prediction corresponds to forecasting the spatial region where users will remain over a given time horizon. The large-scale channel gain is a key factor that mainly characterizes the cell environment.

For traffic prediction, transformer predicts the future traffic level for each cell, enabling the controller to make decisions by comparing the relative traffic load across cells. Traffic prediction data helps mitigate capacity imbalance and improve per-user throughput, satisfying user demand better.

Dual traffic and channel prediction results are periodically delivered to the near-RT RIC. During operation, the xApps on near-RT RIC utilize these predictions to determine multi-step-ahead per-cell power budgets.

C. Near-RT RIC and O-DU Operation

Given the predicted cell environments, each xApp selects an appropriate power budget for each objective under a common goal and installs it at the O-DUs via the E2 interface. Each O-DU then performs slot-level user scheduling and power allocation within the maximum power constraints.

The O-DU continuously monitors QoS metrics (e.g., per-user throughput, delay, outage probability) and reports aggregated KPIs back to the RIC. These measurements are used to *(i)* evaluate the accuracy of the channel and traffic predictions, and *(ii)* assess the effectiveness of the current power-budget allocation method. Based on this feedback, the controllers at the non-RT RIC and near-RT RIC update model parameters.

Through this closed-loop prediction–control-feedback cycle, the proposed agentic-AI in open RAN framework dynamically reconfigures per-cell power budgets and intra-cell resource allocation, thereby improving long-term utility in cellular networks beyond what static resource-sharing schemes can achieve.

III. System Model and Problem Statement

A. Network and Resource Allocation Model

We consider a wireless cellular network in the downlink. Denote by $\mathcal{C} \triangleq \{1, \ldots, C\}$, and $\mathcal{K} \triangleq \{1, \ldots, K\}$ the set of BSs and users, respectively. Each BS c serves $\mathcal{K}_c$ number of associated users. Each user is connected to a single BS. We consider that BSs share a common pool of available power resources, and the total transmit power averaged over time is limited by a network-wide budget, assuming that an instantaneous network-wide total power constraint to be:

$$\sum_{c\in\mathcal{C}}\sum_{s\in\mathcal{S}} p_s^c(t) \ \le\ P_{\rm tot}(t), \qquad \forall t, \tag{1}$$

where $P_{\rm tot}(t)$ is the total available power per time-slot over the network.

We further define longer time scale parameter, T, to be fixed as per-cell power budget during T slots. The parameter T should be determined according to specific metric, such as GAT. We denote per-cell time-slot constraint as

$$\sum_{s\in\mathcal{S}} p_s^c(uT+\tau) \le P^{c,\max}(uT), \tag{2}$$
$$\forall u,\ \forall \tau \in \{uT, (u+1)T-1\}.$$

We illustrate the algorithms that find the appropriate upper bound, $P^{c,max}(uT)$, in Section IV.

In each slot t, each BS c schedules its associated users on the available subchannels and allocates the corresponding transmit power. We denote by $x_s^{c,k}(t) \in \{0,1\}$ the scheduling indicator, which satisfies

$$\sum_{k\in\mathcal{K}_c} x_s^{c,k}(t) \le 1, \quad \forall c,\ \forall s, \tag{3}$$

where $x_s^{c,k}(t)$ equals 1 when user k in cell c is scheduled on subchannel s at slot t, and 0 otherwise. In other words, each subchannel can be assigned to at most one user per slot.

B. Problem Statement

We define our long-term objective as a network utility function that jointly captures throughput and fairness among UEs as

$$\max_{\mathbf{p},\mathbf{x},\mathbf{P}} \qquad \sum_{k\in\mathcal{K}} U_k\Big(\lim_{B\to\infty}\frac{1}{B}\sum_{t=0}^{B-1} r_k(\mathbf{p}(t),\mathbf{x}(t),\mathbf{P}^{c,\max}(t))\Big), \tag{4a}$$

$$\text{subject to} \quad (1),(2),(3), \tag{4b}$$

where $U_k(\cdot)$, $r_k(\cdot)$, B, $\mathbf{p} = \{\mathbf{p}(t)\}_{t=0}^{\infty}$, $\mathbf{x} = \{\mathbf{x}(t)\}_{t=0}^{\infty}$, and $\mathbf{P}^{c,max}(t) = \{\mathbf{P}^{c,max}(t)\}_{t=0}^{\infty}$ denote that utility function, instantaneous user rate, number of time slots, $\mathbf{p}(t) = \{p_s^c(t), s\in\mathcal{S}, c\in\mathcal{C}\}$, $\mathbf{x}(t) = \{x_s^{c,k}(t), s\in\mathcal{S}, c\in\mathcal{C}\}$, and $\mathbf{P}^{c,max}(t) = \{P^{c,max}(t), c\in\mathcal{C}\}$.

To this end, we seek to maximize the network utility by determining the per-cell power budgets, $\{P^{c,\max}(t)\}$, while relying on the existing interference-management algorithm in [4] to find the instantaneous transmit power $\{p_s^c(t)\}$ and scheduling decisions $\{x_s^{c,k}(t)\}$.

Algorithm 1 Score-Based Power-Budget Allocation

Step 1: Normalize predicted values
Compute $\tilde{\Lambda}_c(uT)$ and $\tilde{\Gamma}_c(uT)$ according to (7) and (8).
Step 2: Score relative cell value
Compute $\alpha_c(uT)$ according to (10).
Step 3: Determine cell power budget
Compute $P^{c,\max}(uT)$ according to (10).

In this work, we develop algorithms that determine per-cell maximum power budgets which will be described in Section IV.

IV. Power Budget Allocation Algorithm

In this section, we propose two heuristic algorithms for determining the per-cell maximum power budgets. Both algorithms utilize cell-level traffic and channel predictions, either an offline-simulation–based algorithm or a score-based algorithm.

First, we define dual predicted results of cell as the average value of cell which are evaluated every T time slots. For each cell $c \in \mathcal{C}$, we assume that prediction modules provide

- the predicted traffic level (e.g., the number of active UEs or the aggregate arrival rate) over T as

$$\bar{\Lambda}_c(uT) \triangleq \frac{1}{T}\sum_{t=uT}^{uT+(u+1)T-1} \lambda_c(t), \qquad \forall u, \tag{5}$$

- the predicted effective channel quality (e.g., an average or percentile of the predicted channel gains) over T as

$$\bar{\Gamma}_c(uT) \triangleq \frac{1}{T}\sum_{t=uT}^{uT+(u+1)T-1} \gamma_c(t), \qquad \forall u, \tag{6}$$

where $\lambda_c(t)$ and $\gamma_c(t)$ are predicted values of traffic and channel respectively. Dual traffic and channel predicted values can be obtained from various transformer-based models that provide highly accurate estimates of cell-level traffic and channel conditions [8], [9].

A. Offline Simulation-Based Power-Budget Allocation

We propose running offline simulations to identify GAT-maximizing per-cell power budget. We conduct a lot of offline simulations for various cell environment cases in advance. Then, we build a mapping book in which we record pairs of cell environment and corresponding GAT-maximizing power budget. When a similar cell environment case is predicted online, we can easily choose the power budgets by mapping between cell environment and GAT-maximizing power budgets.

B. Score-Based Power-Budget Allocation

In order to derive a general power-allocation rule, we assign a numerical score to each cell based on its predicted traffic and channel quality. We denote $w_\lambda, w_\gamma \in [0,1]$ as non-negative weights that leverage trade-off between traffic and channel, respectively. For example, we can assign more weight to w_λ

when per-user throughput is prioritized or more weight to w_γ when edge-user experience is of primary concern.

We first normalize the predicted traffic and channel values of cell c as

$$\tilde{\Lambda}_c(uT) = \frac{\bar{\Lambda}_c(uT)}{\sum_{c\in\mathcal{C}} \bar{\Lambda}_c(uT)}, \qquad \forall u, \tag{7}$$

$$\tilde{\Gamma}_c(uT) = \frac{\bar{\Gamma}_c(uT)}{\sum_{c\in\mathcal{C}} \bar{\Gamma}_c(uT)}, \qquad \forall u. \tag{8}$$

Then, the score of cell c is defined as

$$\mathrm{score}_c(uT) = w_\lambda\,\tilde{\Lambda}_c(uT) + w_\gamma\,\tilde{\Gamma}_c(uT), \qquad \forall u. \tag{9}$$

The score reflects both the relative traffic load and the relative channel quality of the cell. Next, we convert the scores into power-splitting coefficients $\alpha_c(uT)$ by

$$\alpha_c(uT) = \frac{\mathrm{score}_c(uT)}{\sum_{c\in\mathcal{C}} \mathrm{score}_c(uT)}, \qquad \forall u, \tag{10}$$

which satisfies

$$\sum_{c\in\mathcal{C}} \alpha_c(uT) = 1, \qquad \forall u. \tag{11}$$

Finally, the per-cell maximum power budgets are determined as

$$P^{c,\max}(uT) = \alpha_c(uT)\,P_{\mathrm{tot}}(t), \qquad \forall u, \tag{12}$$

where $P_{\mathrm{tot}}(t)$ denotes the total power budget defined in (1). The entire process of scored-based power-budget allocation is described in Algorithm 1.

Therefore, key design tasks are to choose the weight parameters in a way that reflects the desired performance objective. Once the maximum power budgets $\{P^{c,\max}(t)\}$ are determined, each BS performs user scheduling and instantaneous power allocation within its own budget following the existing algorithm such as [4].

Under the framework described in Section II, our proposed dual prediction-based per-cell power-budget allocation algorithm can be easily incorporated into agentic-AI system for open-RAN architecture. It is because proposed algorithm is a two-level decision process: the upper level dynamically allocates per-cell power budgets using prediction-based methods, while the lower level optimizes user scheduling and power allocation.

V. Performance Evaluation

In our simulations, we focus on three key observations as follows. *(i)* We observe that adopting different power budgets is beneficial as cell environments vary, illustrating the need for environment-aware power-budget adaptation. *(ii)* We find that the power budgets selected by the proposed algorithm consistently yield superior throughput compared to a baseline strategy, demonstrating the benefit of incorporating predictive information into power-allocation decisions. *(iii)* Our results reveal that improvements in prediction accuracy translate directly into higher network performance, highlighting the central role of reliable traffic and channel forecasting in enabling effective resource control.

We utilize transformer-based prediction schemes to obtain dual prediction results in prior studies and implement the open source codes provided by [8], [9]. Then, we apply offline simulation-based power-budget allocation algorithm for given prediction results. We conduct two sets of simulations corresponding to two scenarios. In the first scenario (i.e., Sections V-B and V-C), we assume perfect predictions for both traffic and channel conditions. In the second scenario (i.e., Section V-C), we evaluate performance under varying levels of prediction accuracy.

A. Simulation Setup

We consider two small cells with a radius of 50 m where BSs are located on a line in a distance of 100 m. We adopt a SISO system as a simplified model.[1] The number of users per cell is set to 10 users in Cell 1 and 2 users in Cell 2 in order to capture unbalanced cell environments. Each user is assumed to move around the network with a constant speed of 40 km/h to capture dynamic network variations. The system bandwidth is set to be 80 MHz. The total simulation duration is 1000 slots. At every time slot, the sum of the maximum available transmit powers of the two cells is constrained by

$$P^{1,\max} + P^{2,\max} = 2.0\ \mathrm{W}. \tag{13}$$

Users in both cells are placed and move in regions where they experience inter-cell interference from the neighboring BS. This setting will demonstrate that our algorithm works well under interference-limited region.

As an evaluation metric, Geometric Average Throughput (GAT) which results in geometric average of user throughput is used. We evaluated our algorithm compared to an equal-power resource allocation (EQ-RA) as a baseline, where two cells always transmit within the fixed power budget under the above constraint (13).

In modeling the propagation environment, we consider a carrier frequency of $f_c = 3.5$ GHz, and 3D distance between BS n and UE k at slot t is given by

$$d^{\mathrm{3D}}_{c,k}(t) = \sqrt{d^2_{c,k}(t) + (h_{\mathrm{BS}} - h_{\mathrm{UE}})^2}, \tag{14}$$

where the BS and UE heights are set to $h_{\mathrm{BS}} = 10$ m and $h_{\mathrm{UE}} = 2$ m, respectively. The large-scale path loss (in dB) follows the 3GPP TR 38.901 model and is given by

$$\mathrm{PL}_{c,k}(t) = 32.4 + 20\log_{10}(f_c) + 31.9\log_{10}\big(d^{\mathrm{3D}}_{c,k}(t)\big), \tag{15}$$

where $d^{\mathrm{3D}}_{c,k}(t)$ is measured in meters. The corresponding large-scale channel power gain is

$$G^{\mathrm{LS}}_{c,k}(t) = 10^{-\mathrm{PL}_{c,k}(t)/10}. \tag{16}$$

On top of this, time-varying small-scale fading is applied by associating each UE with a generated complex channel fading

[1] It can be easily extended to the MIMO system framework.

sequence $\{h_{c,k}(t)\}$, normalized such that $\mathbb{E}[|h_{c,k}(t)|^2] = 1$. Hence, the instantaneous channel power gain of the link between BS c and UE k is modeled as

$$g_{c,k}(t) = G_{c,k}^{\mathrm{LS}}(t)\,|h_{c,k}(t)|^2. \quad (17)$$

B. Power Budgets according to Cell Environments

Fig. 2 illustrates the power budget at cell 1 that maximizes the performance metric under various cell environments. The letters D, G, and B in front of the parentheses denote the channel environment of the users belonging to each cell: D stands for *diverse* (heterogeneous channel qualities), G for *good* (users located within half of the cell radius), and B for *bad* (users located farther from the serving BS and thus experiencing weak channels). This classification is based on the user mobility range relative to the cell radius R; users moving mostly closer to the serving BS than a reference distance, R/2, are labeled as G, those moving farther are labeled as B, and the remaining mixed cases are labeled as D. The number in parentheses indicates the cell identifier.

As shown in Fig. 2, power budget at Cell 1 that maximizes the GAT differs significantly across cell channel environments. It is observed that power budget at Cell 1 tends to be higher than Cell 2, where Cell 1 is connected to a lot users than Cell 2. It implies that the traffic load in some case, 10 users in Cell 1 and 2 users in Cell 2, is more of importance rather than cell channel channel conditions in terms of fairness and throughput. In addition, the result highlights the importance of dynamically adjusting power budgets among cells to maximize overall network performance.

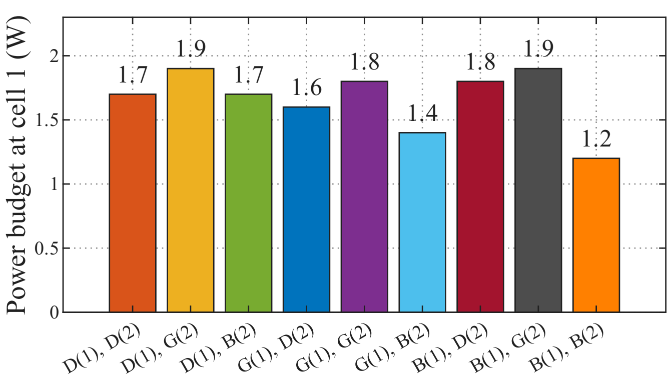

Fig. 2: Power budget at cell 1 that maximizes GAT under diverse cell environments.

C. Performance Comparison

Table I compares the GAT performance and target-throughput satisfaction ratio of the proposed prediction-based power allocation algorithm (P-RA) and the equal-power baseline (EQ-RA) in the $B(1), B(2)$ environment as shown in Fig 2. Here, $B(1), B(2)$ indicates that users in both cells are located in regions where they experience strong inter-cell interference from the neighboring BS. Power budgets of both cells are set as $(1.0\ \mathrm{W}, 1.0\ \mathrm{W})$ for EQ-RA, whereas $(1.2\ \mathrm{W}, 0.8\ \mathrm{W})$ for P-RA based on offline simulation.

As shown in Table I, the proposed P-RA algorithm achieves up to 27.1% higher GAT compared to EQ-RA. It demonstrates that the proposed algorithm can provide substantial gains in overall GAT performance. The gain comes from situation in highly interference-limited scenarios. It is often beneficial for the two BSs to operate at highly asymmetric power levels trying not to interrupt a scheduled user associated with the other BS. Moreover, P-RA achieves 83.3% of target-throughput (60Mbps) satisfaction ratio implying that most of users are satisfied with fine throughput. On the other hand, EQ-RA achieves 16.7% at both of target-throughput (60Mbps, 70Mbps) satisfaction ratios, implying that only some of users are satisfied with high throughput, resulting low GAT performance.

TABLE I: Performance comparison between P-RA and EQ-RA

Scheme	P-RA	EQ-RA
GAT [Mbps/user]	62.17	48.91
GAT Improvement Rate [%]	27.1	-
Target-throughput (60 Mbps) satisfaction ratio [%]	83.3	16.7
Target-throughput (70 Mbps) satisfaction ratio [%]	0.0	16.7

D. Impact of Prediction Accuracy

Fig. 3 illustrates performance comparison according to the prediction accuracy in the $B(1), B(2)$ environment in Fig 2. Prediction error describes the case where Cell 1 actually has a higher traffic load, but the controller mis-predicted that Cell 2 has higher traffic with a certain probability. It is observed that GAT decreases as the prediction accuracy decreases from 100% down to 10%. In particular, GAT is reduced by 20.2% with 10% prediction accuracy compared to GAT of 100% prediction accuracy. This result clearly shows that prediction accuracy has a significant impact on the overall network performance.

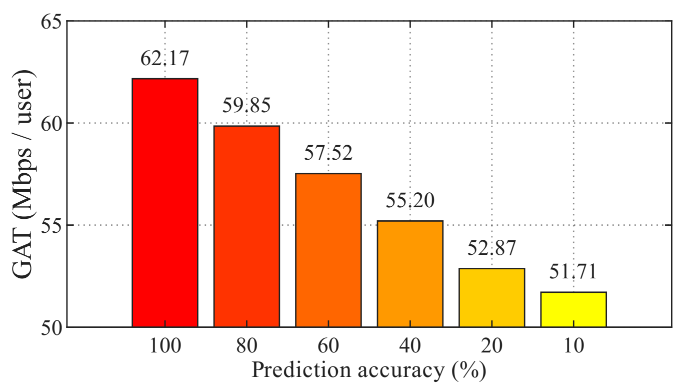

Fig. 3: GAT comparison across prediction accuracies.

VI. Conclusion

In this paper, we first envisioned an agentic-AI framework for open-RAN architecture. On top of this framework, we proposed a prediction-based power–budget allocation framework in open-RAN architecture. The main idea is to utilize dual traffic and channel predictions in order to get an insight of cell environment, and then adjust the per-cell power budget. Through simulations, we verified that our proposed algorithm achieves significant performance improvement compared to baseline algorithm. Finally, we showed the impact of prediction accuracy implying that prediction accuracy directly leads to higher network-wide utility.

REFERENCES

[1] T. S. Ukyab, L. Suzuki, D. Davis, Z. Luo, S. Fu, S. Hasan, S. Ratnasamy, and S. Shenker, "Making cellular networks more efficient by roaming-in-place," in *Proceedings of ACM SIGCOMM*, 2025, pp. 979–993.

[2] K. Attiah, K. Banawan, A. Gaber, A. Elezabi, K. Seddik, Y. Gadallah, and K. Abdullah, "Load balancing in cellular networks: A reinforcement learning approach," in *Proceedings IEEE Consumer Communications & Networking Conference (CCNC)*, 2020, pp. 1–6.

[3] E. Gures, I. Shayea, M. Ergen, M. H. Azmi, and A. A. El-Saleh, "Machine learning-based load balancing algorithms in future heterogeneous networks: A survey," *IEEE Access*, vol. 10, pp. 37689–37717, 2022.

[4] K. Son, S. Lee, Y. Yi, and S. Chong, "REFIM: A practical interference management in heterogeneous wireless access networks," *IEEE Journal on Selected Areas in Communications*, vol. 29, no. 6, pp. 1260–1272, 2011.

[5] J. Kwak, K. Son, Y. Yi, and S. Chong, "Greening effect of spatio-temporal power sharing policies in cellular networks with energy constraints," *IEEE Transactions on Wireless Communications*, vol. 11, no. 12, pp. 4405–4415, Dec. 2012.

[6] K. Son, R. Guruprasad, S. Nagaraj, M. Sarkar, and S. Dey, "Dynamic cell reconfiguration framework for energy conservation in cellular wireless networks," *Journal of Communications and Networks*, vol. 18, no. 4, pp. 567–579, 2016.

[7] Y. Kim, J. Jeong, S. Ahn, J. Kwak, and S. Chong, "Energy and delay guaranteed joint beam and user scheduling policy in 5G CoMP networks," *IEEE Transactions on Wireless Communications*, vol. 21, no. 4, pp. 2742–2756, 2021.

[8] H. Jiang, M. Cui, D. W. K. Ng, and L. Dai, "Accurate channel prediction based on transformer: Making mobility negligible," *IEEE Journal on Selected Areas in Communications*, vol. 40, no. 9, pp. 2717–2732, Sep. 2022.

[9] B. Gu, J. Zhan, S. Gong, W. Liu, Z. Su, and M. Guizani, "A spatial-temporal transformer network for city-level cellular traffic analysis and prediction," *IEEE Transactions on Wireless Communications*, vol. 22, no. 12, pp. 9412–9423, Dec. 2023.