Preparation of Articles for IEEE TRANSACTIONS and JOURNALS (2022)

First A. Author, *Fellow, IEEE*, Second B. Author, and Third C. Author Jr., *Member, IEEE*

***Abstract*—This document provides a guide for preparing articles for IEEE Transactions, Journals, and Letters. Use this document as a template if you are using Microsoft *Word*. Otherwise, use this as an instruction set. The electronic file of your article will be formatted further at IEEE. Titles should be written in uppercase and lowercase letters, not all uppercase. Avoid writing long formulas with subscripts in the title; short formulas that identify the elements are fine (e.g., "Nd–Fe–B"). Do not write "(Invited)" in the title. Full names of authors are preferred in the author field but are not required. Put a space between authors' initials. ORCIDs can be provided here as well. In the title, all variables should appear lightface italic; numbers and units will remain bold. Abstracts must be a single paragraph. In order for an Abstract to be effective when displayed in IEEE *Xplore* as well as through indexing services such as Compendex, INSPEC, Medline, ProQuest, and Web of Science, it must be an accurate, stand-alone reflection of the contents of the article. They shall not contain displayed mathematical equations, numbered reference citations, nor footnotes. They should include three or four different keywords or phrases, as this will help readers to find it. It is important to avoid over-repetition of such phrases as this can result in a page being rejected by search engines. Ensure that your abstract reads well and is grammatically correct.**

Index Terms*—Enter keywords or phrases in alphabetical order, separated by commas. Using the *IEEE Thesaurus* can help you find the best standardized keywords to fit your article. Use the thesaurus access request form for free access to the *IEEE Thesaurus https://www.ieee.org/publications/services/thesaurus.html.

I. INTRODUCTION

THE rapid growth of wireless communication technologies and the emergence of diverse applications have led to an unprecedented increase in mobile data traffic. The fifth generation (5G) and beyond wireless networks are expected to support a wide range of services with heterogeneous requirements, including enhanced mobile broadband (eMBB), ultra-reliable low-latency communications (URLLC), and massive machine-type communications (mMTC) [1], [2]. To meet these diverse demands, efficient and dynamic resource allocation techniques are crucial for optimizing network performance and ensuring quality of service (QoS) [3].

In 5G and beyond networks, physical resource blocks (PRBs) are the fundamental unit of resource allocation, representing a specific time-frequency resource in the orthogonal frequency-division multiple access (OFDMA) framework [4]. Traditional resource allocation schemes, such as proportional fairness and max-min fairness, often fail to adapt to the dynamic nature of heterogeneous traffic demands and varying channel conditions [5]. Moreover, these schemes typically focus on a single objective, such as throughput maximization or fairness enhancement, without considering the trade-offs between multiple conflicting objectives [6]. To address these challenges, researchers have explored various approaches for dynamic PRB allocation in heterogeneous networks. Reinforcement learning (RL) has emerged as a promising technique for adaptive resource allocation, enabling the system to learn optimal allocation policies through interaction with the environment [7]. Deep reinforcement learning (DRL) algorithms, such as deep Q-networks (DQNs), have been applied to PRB allocation problems, demonstrating improved performance compared to traditional schemes [8], [9]. However, most existing DRL-based approaches focus on single-objective optimization and may not effectively handle the complex trade-offs between multiple objectives in heterogeneous networks [10].

Another critical aspect of resource allocation in 5G and beyond networks is the management of inter-cell and intra-cell interference [11]. Interference can significantly degrade the performance of wireless systems, particularly in dense deployment scenarios [12]. Effective interference management techniques, such as coordinated multipoint (CoMP) transmission and reception, have been proposed to mitigate interference and improve spectral efficiency [13]. However, the

This paragraph of the first footnote will contain the date on which you submitted your paper for review, which is populated by IEEE. It is IEEE style to display support information, including sponsor and financial support acknowledgment, here and not in an acknowledgment section at the end of the article. For example, "This work was supported in part by the U.S. Department of Commerce under Grant 123456." The name of the corresponding author appears after the financial information, e.g. *(Corresponding author: Second B. Author).* Here you may also indicate if authors contributed equally or if there are co-first authors.

The next few paragraphs should contain the authors' current affiliations, including current address and e-mail. For example, First A. Author is with the National Institute of Standards and Technology, Boulder, CO 80305 USA (e-mail: author@ boulder.nist.gov).

Second B. Author Jr. was with Rice University, Houston, TX 77005 USA. He is now with the Department of Physics, Colorado State University, Fort Collins, CO 80523 USA (e-mail: author@lamar.colostate.edu).

Third C. Author is with the Electrical Engineering Department, University of Colorado, Boulder, CO 80309 USA, on leave from the National Research Institute for Metals, Tsukuba 305-0047, Japan (e-mail: author@nrim.go.jp).

Mentions of supplemental materials and animal/human rights statements can be included here.

Color versions of one or more of the figures in this article are available online at http://ieeexplore.ieee.org complexity and overhead associated with these techniques can hinder their practical implementation [14]. Scalability and complexity are also major concerns in the design and optimization of resource allocation algorithms for 5G and beyond networks [15]. As the network size and the number of connected devices grow, the computational complexity of resource allocation algorithms can become a bottleneck, limiting their real-time performance [16]. Therefore, it is essential to develop scalable and computationally efficient algorithms that can adapt to the dynamic nature of heterogeneous networks while maintaining acceptable performance [17].

A. Motivation and Problem Statement

The motivation behind this work stems from the need for efficient and dynamic resource allocation techniques that can adapt to the diverse requirements of heterogeneous 5G and beyond networks. The existing resource allocation schemes often suffer from limitations such as lack of adaptability, single-objective optimization, and high computational complexity [5], [6], [15]. These limitations hinder the ability of 5G and beyond networks to effectively support the wide range of services and applications envisioned for the future wireless ecosystem. The problem addressed in this paper is the development of a comprehensive resource allocation framework that can efficiently allocate PRBs to heterogeneous networks while considering multiple objectives, such as PRB utilization, spectral efficiency, fairness, and energy efficiency. The proposed framework aims to overcome the limitations of existing approaches by incorporating dynamic traffic prediction, adaptive slicing, and interference management techniques. Additionally, the framework focuses on achieving a balance between performance optimization and computational complexity, ensuring scalability and real-time operation in large-scale networks.

B. Overview of Dynamic PRB Allocation in Heterogeneous Networks

Dynamic PRB allocation in heterogeneous networks involves the real-time assignment of time-frequency resources to different users and services based on their QoS requirements, channel conditions, and network state [3]. The main objectives of dynamic PRB allocation are to maximize the utilization of available resources, ensure fairness among users, and satisfy the diverse QoS demands of different services [6]. Several approaches have been proposed for dynamic PRB allocation in heterogeneous networks, including heuristic algorithms, optimization-based techniques, and machine learning-based methods [5], [7], [8]. Heuristic algorithms, such as round-robin and proportional fair scheduling, are simple and computationally efficient but may not adapt well to the dynamic nature of heterogeneous traffic [5]. Optimization-based techniques, such as integer linear programming and convex optimization, can provide optimal solutions but often suffer from high computational complexity and may not scale well to large networks [6]. Machine learning-based approaches, particularly reinforcement learning (RL), have gained significant attention in recent years due to their ability to learn and adapt to the dynamic environment [7]. RL-based methods, such as Q-learning and deep Q-networks (DQNs), have been applied to PRB allocation problems, showing promising results in terms of improved network performance and adaptability [8], [9]. However, most existing RL-based approaches focus on single-objective optimization and may not effectively capture the complex trade-offs between multiple objectives in heterogeneous networks [10].

C. Key Contributions

The main contributions of this paper are as follows:

- We propose a novel Dynamic Adaptive Resource Tracing for Physical Resource Block (DART-PRB) framework for efficient PRB allocation in heterogeneous 5G and beyond networks. The framework incorporates dynamic traffic prediction, adaptive slicing, and interference management techniques to optimize network performance while considering multiple objectives.
- We introduce a hierarchical multi-resolution allocation scheme that adapts the resource allocation granularity based on the service requirements. The scheme employs micro-level, meso-level, and macro-level allocation to effectively serve URLLC, eMBB, and mMTC services, respectively.
- We develop an enhanced DQN engine with adaptive state compression, action masking, and twin Q-networks to enable efficient learning and decision-making in the dynamic network environment. The DQN engine incorporates service-specific constraints and fairness-utilization balancing to achieve a comprehensive optimization of network performance.
- We propose a proactive traffic prediction module that combines ensemble forecasting techniques, including recurrent neural networks (RNNs), autoregressive integrated moving average (ARIMA) models, and seasonal pattern analysis. The prediction module enables accurate estimation of future traffic demands, facilitating efficient resource allocation and proactive network optimization.
- We introduce a channel-aware interference management scheme that leverages spatial correlation-aware clustering, inter-slice interference mapping, and intra-slice interference control. The scheme effectively mitigates interference and improves the overall network capacity and spectral efficiency.
- We conduct extensive simulations and performance evaluations to demonstrate the effectiveness of the proposed DART-PRB framework. The results show significant improvements in PRB utilization, spectral efficiency, fairness, and energy efficiency compared to existing baseline schemes.
- We perform a comprehensive complexity and scalability analysis of the DART-PRB framework, including component-level complexity profiling and

scalability testing across different network sizes. The analysis provides insights into the computational trade-offs and the framework's ability to scale efficiently in large-scale networks.

II. Related Works

Traditional resource allocation schemes in wireless networks often rely on fixed or semi-dynamic scheduling techniques, such as round-robin, water-filling, and proportional fairness [18], [19]. These methods assume complete channel state information (CSI) and fixed scheduling rules, which can be effective in low-traffic environments. However, they tend to incur high signaling overhead and fail to adapt to the rapidly varying conditions of 5G and beyond networks, where heterogeneous services such as V2X, eMBB, and mMTC coexist [7]. Even when enhanced by combinatorial optimization for user grouping and beamforming [20], these traditional approaches remain rigid and computationally intensive under high load.

To address dynamic network conditions, reinforcement learning (RL) has been increasingly adopted for resource allocation. Deep reinforcement learning (DRL) algorithms, such as deep Q-networks (DQNs), have been applied to learn adaptive mappings from network states to allocation decisions [21], [22]. Centralized DRL techniques [21], [22] improve adaptability but are often hindered by slow convergence and the need for extensive CSI feedback. Multi-agent DRL frameworks [23], [24], [25] further distribute the decision-making process, yet still struggle with computational overhead in dense deployments. The proposed DART-PRB framework mitigates these issues by employing a hierarchical multi-agent DRL approach that segments the resource allocation problem across micro (V2X), meso (eMBB), and macro (mMTC) time scales while incorporating opportunistic resource sharing to reduce complexity.

In addition to DRL, several machine learning (ML) techniques have been employed for resource allocation. Supervised learning approaches, such as deep neural networks (DNNs), have been used to predict modulation and coding schemes (MCS) based on CSI and user location [26]. Moreover, decision tree-based methods, including random forest algorithms, have been applied to classify users and segment resources effectively [27], [28]. Although these techniques offer robust performance under relatively static scenarios, they often require large amounts of labeled data and struggle to adapt to real-time dynamics. In contrast, the DART-PRB framework uniquely integrates ML-based traffic prediction (RNNs) with a hierarchical DRL allocation strategy. This hybrid design reduces reliance on extensive training data and achieves rapid adaptation to network fluctuations with lower computational overhead.

Scalability remains a crucial challenge as network sizes grow. Centralized optimization methods, such as those based on semidefinite relaxation (SDR) or exhaustive search [29], [30], often exhibit complexity that scales with the number of PRBs [7]. Although these methods can achieve near-optimal performance, their heavy computational burden renders them impractical for dense 5G and beyond networks. Recent studies [26], [27], [31] have attempted to lower complexity via distributed and machine learning-based approaches. The DART-PRB system addresses scalability by incorporating a dedicated Complexity Analyzer module, which partitions overall execution time into components (traffic prediction, resource allocation, interference management, learning updates, and performance measurement). This modular approach enables adaptive tuning of computational resources, ensuring better scalability compared to traditional centralized methods.

III. System Model and Architecture

The DART-PRB framework aims to optimize resource allocation in heterogeneous 5G and beyond networks by dynamically adapting to traffic demands, channel conditions, and service requirements. This section presents the system model and architecture, including the network configuration, traffic models, channel and interference modeling, performance metrics, and an overview of the DART-PRB framework.

A. Network Configuration and Traffic Models

The DART-PRB framework considers a heterogeneous network configuration consisting of a set of PRBs $N = \{1,2,\dots,NRB\}$ and three distinct sets of UEs: $UV2X = \{1,\dots,M_{veh}\}$ for V2X services, $UeMBB = \{1,\dots,M_{eMBB}\}$ for eMBB service, and $UmMTC = \{1,\dots,M_{mMTC}\}$ for mMTC services. The resources are allocated to each service type based on the service slicing ratios:

$$p = \begin{bmatrix} p_{V2X} \\ p_{eMBB} \\ p_{mMTC} \end{bmatrix} \quad (1)$$

s.t:

$$\sum_{s\in\{V2X,eMBB,mMTC\}} p_s = 1 \ and \ p_s \geq p_s^{min}$$

ensuring a balanced distribution of resources while guaranteeing minimum allocations for each service. The traffic arrival processes for each service type are modeled as follows:

- V2X traffic follows a Poisson process with rate λ_{niu}, where the probability of n packet arrivals in time interval t is given by $\Pr(N_{V2X}(t) = n) = \frac{(\lambda_{niu}.t)^n e^{-\lambda_{niu}.t}}{n!}$.
- eMBB traffic also follows a Poisson process with rate λ_e, where the probability of n session arrivals in time interval t is given by $\Pr(N_{eMBB}(t) = n) = \frac{(\lambda_e.t)^n e^{-\lambda_e.t}}{n!}$.
- mMTC traffic is modeled by the probability of an mMTC UE transmitting in a given time slot, denoted by p_{mMTC_tx}.

B. Channel and Interference Modeling

The wireless channel between the base station and UEs is characterized using the WINNER II path loss model, given by:

$$PL(d) = 10\alpha_{PL}log_{10}(d) + X_\sigma \quad (2)$$

where d is the distance between the transmitter and receiver, α_{PL} is the path loss exponent, and X_σ is a log-normal shadowing term with zero mean and standard deviation σ. This model captures the large-scale fading effects in the network.

Interference is a critical factor affecting the performance of the DART-PRB system. The SINR for user u on PRB n is calculated as:

$$\gamma_{u,n} = \frac{P_u G_u |h_{u,n}|^2}{\sum_{u' \neq u} P_{u'} G_{u'} |h_{u',n}|^2 I_{u',n} + N_0} \quad (3)$$

where P_u is the transmit power of user u, G_u is the channel gain, $h_{u,n}$ is the channel coefficient, $I_{u',n}$ is the interference from other users u' on PRB n, and N_0 is the noise power. The noise-plus-interference term $\eta_n = \sum_{u' \neq u} P_{u'} G_{u'} |h_{u',n}|^2 I_{u',n} + N_0$ represents the aggregate interference and noise on PRB n. The DART-PRB framework employs advanced interference management techniques, such as coordinated multipoint (CoMP) transmission and reception, to mitigate the impact of interference and improve the overall network performance. These techniques, along with the dynamic adaptation of slicing ratios based on the optimization objectives, enable the framework to efficiently allocate resources and meet the diverse requirements of different service types in the heterogeneous network.

C. Performance Metrics and Optimization Objectives

The DART-PRB framework considers several key performance metrics and optimization objectives to ensure efficient resource allocation and satisfaction of diverse service requirements. The main performance metrics include:

- PRB Utilization: The PRB utilization metric $U_{PRB} = \frac{\sum_{b \in B} \sum_{n=1}^{N} 1(a_{b,n}=1)}{B.N} .100\%$ measures the percentage of allocated PRBs out of the total available PRBs, where $a_{b,n} \in \{0,1\}$ indicates PRB n is allocated in cell b, and $1(\cdot)$ is the indicator function.
- Spectral Efficiency: The spectral efficiency metric quantifies the average data rate per unit bandwidth. It is calculated as: $\eta = \frac{\sum_{u \in U} R_u}{B.W}$ where R_u is the achieved data rate of UE u, and W is the total system bandwidth.
- Fairness: The fairness metric ensures a balanced distribution of resources among UEs. The Jain's fairness index [32] is used to evaluate the fairness of resource allocation: $J = \frac{(\sum_{u \in U} R_u)^2}{|u| . \sum_{u \in U} {R_u}^2}$, where $|U|$ denotes the cardinality of set U. The fairness index ranges from $\frac{1}{|U|}$ (worst case) to 1 (best case).
- Energy Efficiency: The energy efficiency metric measures the ratio of the total achieved data rate to the total power consumption. It is defined as: $\in = \frac{\sum_{u \in U} R_u}{\sum_{b \in B} P_b}$, where P_b is the total power consumption of BS b, including transmission power and circuit power.
- SLA Satisfaction: The SLA satisfaction metric evaluates the percentage of UEs that meet their respective SLA requirements. For each service type $s \in S$, the SLA satisfaction metric is defined as: $S_s = \frac{|\{u \in U_s : SLA_u \text{ is satisfied}\}|}{|U_s|} .100\%$ where SLA_u represents the SLA requirements of UE u.

The DART-PRB framework optimizes performance by dynamically allocating PRBs to UEs based on their service types and traffic demands, with the optimization objectives including: maximizing PRB utilization $\left(max_{a_{b,n}} U_{PRB}\right)$ spectral efficiency $\left(max_{a_{b,n}}, P_{b,u,n} \eta\right)$; energy efficiency $\left(max_{a_{b,n}}, P_{b,u,n} \in\right)$; SLA satisfaction $\left(max_{a_{b,n}} \sum_{s \in S} \omega_s . S_{s'}\right)$, where ω_s represents the priority weight for each service type; and ensuring fairness $\left(max_{a_{b,n}} J\right)$. These objectives are subject to constraints including the PRB allocation rule (each PRB can be assigned to at most one UE per cell), power limits (ensuring that the total transmit power of each BS does not exceed its maximum budget), and SLA requirements (the achieved performance must meet each service types of SLA). The framework leverages a mix of hierarchical resource allocation, reinforcement learning, and heuristic algorithms to efficiently tackle this multi-objective optimization problem, with detailed algorithmic solutions discussed in Section 4.

D. Overview of the DART-PRB Framework

The DART-PRB framework operates on heterogeneous network deployments supporting multiple network slices. Figure 1 presents the physical network topology that serves as the foundation for our resource allocation system, showing both downlink and uplink communication patterns. The left panel illustrates downlink communications where all UEs connect directly to the base station with varying SINR values, while the right panel demonstrates uplink communications, highlighting the hybrid connectivity approach where vehicular UEs can communicate either directly with the base station (cellular mode) or through Roadside Units (RSUs) in sidelink mode. This heterogeneous deployment directly informs multiple components of our DART-PRB framework. The spatial distribution of UEs and their SINR values serve as critical inputs for the Channel-Aware Interference Management module, particularly for Spatial Correlation-Aware Clustering. The diverse service types (eMBB, mMTC, vehicular) correspond to our Hierarchical Multi-Resolution Allocation approach, which assigns resources at different temporal resolutions. The network state, including node positions, connection types, and measured SINR values, forms the initial state representation that our Constrained DQN Engine processes, transforming this complex spatial representation into a manageable state space for efficient learning.

Figure 2 presents a high-level overview of the DART-PRB framework, showcasing the key components and their interactions. The framework consists of four main modules: Network Model, Hierarchical Resource Allocation, Learning & Optimization, and Performance Evaluation. The Network Model module encompasses the heterogeneous network setup, providing the foundation for the resource allocation problem and capturing the diverse requirements of different applications. The Hierarchical Resource Allocation module is responsible for assigning PRBs to UEs based on their service types and requirements, operating at micro, meso, and macro levels. The Learning & Optimization module employs advanced techniques to dynamically adapt the allocation strategy, including a reinforcement learning component with a constrained DQN engine that learns the optimal policy while considering fairness-utility balancing and interference management constraints. The Performance Evaluation module assesses the effectiveness of the DART-PRB framework using various metrics, providing insights into the system's performance and helping identify areas for further optimization. The arrows in Figure 1 indicate the flow of information and control between the modules, enabling the framework to achieve efficient and adaptive resource allocation in heterogeneous 5G and beyond networks, while its modular design allows for flexibility and extensibility to accommodate future advancements.

IV. Proposed Resource allocation framework

The DART-PRB framework introduces a novel resource allocation approach that combines hierarchical multi-resolution allocation, dynamic traffic prediction, constrained deep reinforcement learning, and interference-aware techniques to efficiently meet the diverse requirements of 5G and beyond networks. The proposed framework aims to maximize resource utilization, ensure fairness among different service types, and satisfy the stringent Quality of Service (QoS) constraints of heterogeneous applications. By leveraging the strengths of machine learning and optimization techniques, DART-PRB adapts to the dynamic nature of traffic demands and network conditions, enabling intelligent and efficient resource allocation decisions.

Algorithm 1 DART-PRB Resource Allocation

Initialize:
  Network parameters: $N_{RB}, N_s, M_s, \lambda_s, \Delta t$
  DQN parameters: $\gamma, \alpha, \epsilon, K, L, B$
  Traffic prediction parameters: T, δ, η
  Resource sharing parameters: ρ, θ
  Fairness-utilization balancing parameters: ω, β
for each time step t=1,2…,T **do**
  Observe network state $s_t = \{N_{RB}^{(s)}, U_s, D_s, H_s\}_{s=1}^{N_s}$
  Predict traffic demands $\widehat{D}_{s,t+1} = f_s(D_s, t, \delta, \eta), \forall_s \in (1, \dots, N_s\}$
  Compute resource requirements $R_{s,t+1} = g_s(\widehat{D}_{s,t+1}, H_s, N_{RB}^{(s)}), \forall_s \in (1, \dots, N_s\}$
  Obtain DQN action $a_t = DQN(s_t, \gamma, \alpha, \epsilon, K, L, B)$
  Map action a_t to resource allocation $\{N_{RB,t+1}^{(s)}\}_{s=1}^{N_s}$
  Perform opportunistic resource sharing:
    Cluster UEs based on spatial correlation
    Map inter-slice interference
    Borrow resources from underutilized slices
  Optimize fairness-utilization trade-off:
    Compute fairness metric F_t and utilization metric U_t
    Adjust allocation $\{N_{RB,t+1}^{(s)}\}_{s=1}^{N_s}$ based on ω and β
  Allocate PRBs $\{N_{RB,t+1}^{(s)}\}_{s=1}^{N_s}$ to UEs
  Observe reward r_t and next state s_{t+1}
  Update DQN with experience tuple s_t, a_t, r_t, s_{t+1}
end for

Algorithm 1 iteratively implements the DART-PRB methodology. Each time step integrates traffic prediction for proactive demand estimation, a constrained DQN for QoS-based action learning, opportunistic resource sharing to exploit underutilized spectrum, and fairness–utilization balancing for equitable distribution. This structured flow ensures that outputs (e.g., predicted traffic or updated Q-values) feed subsequent decisions, enabling DART-PRB to adapt dynamically under changing conditions while meeting service-specific performance requirements.

A. Hierarchical Multi-Resolution Allocation

The hierarchical multi-resolution allocation module in the DART-PRB framework addresses the distinct QoS requirements of heterogeneous services by operating at three resolution levels. At the micro-level V2X/URLLC services, which demand ultra-low latency and high reliability for mission-critical applications, the resource allocation problem is formulated as:

$$\min_{x_{u,n}} \sum_{u \in U_{V2X/URLLC}} (\alpha_L L_u + \alpha_R (1 - R_u)) \tag{4}$$
s.t:
$$\sum_{u \in U_{\frac{V2X}{URLLC}}} x_{u,n} \leq 1,\ L_u \leq L_{max}^{\frac{V2X}{URLLC}},\ L_u \leq L_{max}^{\frac{V2X}{URLLC}},\ \forall_n \in N$$

where $x_{u,n}$ is a binary indicator for PRB allocation, and α_L and α_R balance latency L_u and reliability R_u. For the meso-level targeting eMBB services that require high throughput, the problem is posed as:

$$\max_{y_{u,n}} \sum_{u \in U_{eMBB}} \log(1 + R_u) \tag{5}$$
s.t:
$$\sum_{u \in U_{eMBB}} y_{u,n} \leq 1,\ R_u \geq R_{min}^{eMBB},\ \forall_n \in N$$

with $y_{(u,n)} \in [0,1]$ representing the fraction of resource allocated to each user. At the macro-level for mMTC services, which serve a massive number of devices with sporadic traffic and strict energy constraints, the optimization is formulated as:

$$\max_{Z_{u,n}} \sum_{u \in U_{mMTC}} \mathbb{I}(u\ served) - \beta \sum_{u \in U_{mMTC}} E_u \tag{6}$$
s.t:
$$\sum_{u \in U_{mMTC}} Z_{u,n} \leq 1,\ \Pr(\text{collision}) \leq P_{max}^{collision},\ E_u \leq E_{max}^{mMTC},\ \forall_n \in N$$

where $Z_{u,n}$ is binary, $\mathbb{I}(u\ served)$ is an indicator function, E_u denotes energy consumption, and β is a weighting factor. This hierarchical approach enables targeted resource allocation for each service type while optimizing overall network performance and user experience, thereby overcoming the limitations of static schemes in dynamic 5G environments.

B. Dynamic Traffic Prediction Module

The dynamic traffic prediction module of DART-PRB leverages advanced machine learning techniques to forecast future traffic demands and enable proactive resource allocation.

In this module, historical traffic data $X = \{x1, x2, \ldots, x_T\}$ is first modeled using an LSTM network, where the future traffic demand is predicted as $\hat{x}_{t+1} = f(x_t, h_t)$ with h_t representing the hidden state at time t, and the model is trained by minimizing the mean squared error $\mathcal{L}_{MSE} = \frac{1}{T}\sum_{t=1}^{T}(x_t - \hat{x}_t)^2$; in parallel, an ARIMA model characterized by $\emptyset(B)(1-B)^d x_t = \theta(B)\varepsilon_t$ is employed to capture linear trends and seasonal effects, where B denotes the backshift operator, and seasonal pattern analysis is performed by estimating the seasonal component as $S_t = \frac{1}{N}\sum_{i=1}^{N} x_t^{(i)}$; these individual forecasts, denoted as $\hat{x}_{t+1}^{(RNN)}$, $\hat{x}_{t+1}^{(ARIMA)}$,and $\hat{x}_{t+1}^{(Seasonal)}$ respectively, are then combined via an ensemble approach to form the final prediction $\hat{x}_{t+1}^{(Ensemble)} = w_{RNN}\hat{x}_{t+1}^{(RNN)} + w_{ARIMA}\hat{x}_{t+1}^{(ARIMA)} + w_{Seasonal}\hat{x}_{t+1}^{(Seasonal)}$ with weights w_{RNN} , w_{ARIMA} , $w_{Seasonal}$ summing to unity; by continuously updating these models with recent data and dynamically adjusting the confidence intervals based on historical errors, the module achieves robust and adaptive traffic forecasting that underpins the intelligent resource allocation decisions in the DART-PRB framework.

C. Constrained DQN Engine

The constrained DQN engine in the DART-PRB framework is designed to learn optimal resource allocation policies while satisfying service-specific QoS constraints. In this approach, the resource allocation problem is modeled as a Markov Decision Process (MDP), where the state space S encompasses current network conditions (such as traffic demand, channel quality, and resource utilization), and the action space A consists of possible allocation decisions (e.g., PRB assignments, power levels). The objective is to learn an optimal policy π^* that maximizes the expected cumulative discounted reward, expressed as:

$$\pi^* = arg\max_{\pi} \mathbb{E}\left[\sum_{t=0}^{\infty}\gamma^t r(s_t, a_t)|s_0, \pi\right] \tag{7}$$

where $\gamma \in [0,1]$ is the discount factor. To approximate the optimal Q-function $Q^*(s,a)$, a deep neural network $Q_\theta(s,a)$ is employed and trained by minimizing the mean squared temporal difference error:

$$\mathcal{L}(\theta) = \mathbb{E}_{(s,a,r,s')\sim D}\left[\left(r + \gamma \max_{a'} Q_{\theta^-}(s',a') - Q_\theta(s,a)\right)^2\right] \tag{8}$$

with θ^- representing the periodically updated target network parameters. To incorporate resource and QoS constraints, the problem is extended to a constrained MDP formulation:

$$\max_{\pi} \mathbb{E}\left[\sum_{t=0}^{\infty}\gamma^t r(s_t, a_t)|s_0, \pi\right] \tag{9}$$

s.t:

$$\mathbb{E}\left[\sum_{t=0}^{\infty}\gamma^t c_i(s_t, a_t)|s_0, \pi\right] \le d_i, \forall_i$$

where $c_i(s_t, a_t)$ is the $i-th$ constraint function and d_i its threshold. Techniques such as Lagrangian relaxation and constrained policy optimization are employed to integrate these constraints into the learning process. Additionally, multi-agent extensions allow distributed agents to learn in a decentralized fashion while sharing experiences during centralized training, thus enhancing the systems adaptability and scalability. This constrained DQN engine thereby enables adaptive, constraint-aware resource allocation decisions that meet the stringent requirements of 5G and beyond networks.

D. Opportunistic Resource Sharing and Adaptive Slicing

Opportunistic resource sharing leverages the temporal and spatial variations in traffic and channel conditions to improve spectrum utilization. In DART-PRB, spectrum sensing techniques are employed to identify underutilized or vacant PRBs that can be opportunistically accessed without causing harmful interference to primary users. Advanced spectrum sensing methods—such as energy detection, cyclostationary feature detection, and compressed sensing—are used to accurately detect the presence of primary users. Based on these sensing outcomes, dynamic spectrum access techniques (e.g., overlay and underlay sharing) are applied. In overlay sharing, secondary users access the spectrum only when primary users are inactive, whereas underlay sharing allows simultaneous transmissions provided that the interference remains below a predefined threshold. This opportunistic mechanism represented by an interference constraint:

$$\sum_{u\in U} x_{u,n}.I_{u,n} \le I_{th}, \forall_n \in N \tag{10}$$

where $x_{u,n}$ indicates the allocation of PRB n to secondary user u, $I_{u,n}$ is the interference generated, and I_{th} is the maximum allowable interference level. The opportunistic resource sharing module integrates with the hierarchical multi-resolution allocation and constrained DQN engine, utilizing real-time spectrum sensing data and interference constraints to dynamically adjust resource sharing among users.

Adaptive slicing enables the dynamic adjustment of resource allocation ratios among different service types (e.g., V2X/URLLC, eMBB, and mMTC) based on real-time traffic demands and performance requirements. This module formulates the slicing problem as an MDP, where the state space S represents the current network conditions (traffic load, channel quality, and resource utilization), and the action space A corresponds to the possible slicing ratios $\rho = [\rho_{V2X}, \rho_{eMBB}, \rho_{mMTC}]^T$. The reward function $r(s,a)$ is designed to reflect network utility while penalizing QoS violations. The optimal slicing policy obtained by equation (7) and then enforce fairness among service types, fairness metrics such as Jain's fairness index are incorporated into the reward function:

$$J = \frac{\left(\sum_{s\in\{V2X,eMBB,mMTC\}}\rho_s\right)^2}{3.\sum_{s\in\{V2X,eMBB,mMTC\}}\rho_s^2)} \tag{11}$$

The adaptive slicing module continuously monitors network performance and updates slicing ratios via a constrained DQN variant. These learned slicing ratios serve as high-level guidelines for the hierarchical allocation and opportunistic

TABLE I
SYSTEM SIMULATION PARAMETERS

Parameter	Symbol	Value	Descriptions
Number of PRBs	N_{RB}	200	Total number of PRBs available in the system
Subcarrier bandwidth	f_{sc}	30 KHz	Bandwidth of each subcarrier
Number of Subcarriers per PRB	N_{sc}^{RB}	12	Number of subcarriers constituting a PRB
Bandwidth per PRB	B	360 KHz	Total bandwidth of each PRB
Simulation step size	T_{drop}	0.5 ms	Time granularity of the simulation
TTI duration	F_d	0.5 ms	Transmission Time Interval (TTI) duration
Number of service types	N_r^s	3	Total number of service types considered
Number of V2X UEs	M_{veh}	8	Number of UEs requesting V2X services
Number of eMBB UEs	M_{eMBB}	4	Number of UEs requesting eMBB services
Number of mMTC UEs	M_{mMTC}	12	Number of UEs requesting mMTC services
V2X packet arrival rate	λ_v	200	Average arrival rate of V2X packets
eMBB session arrival rate	λ_e	300	Average arrival rate of eMBB sessions
mMTC transmission probability	p_{mMTC}^{tx}	0.75	Probability of mMTC UEs transmitting packets
Maximum mMTC packet size	-	128 bytes	Maximum size of mMTC packets
V2X packet size	S_m	300 bytes	Fixed size of V2X packets
eMBB session bit rate	$R_b^{session}$	1 Mbps	Constant bit rate of eMBB sessions
Vehicle velocity	-	22.22 m/s (80 km/h)	Average velocity of vehicles in the network
Simulation steps per round	-	10	Number of simulation steps in each round
Number of simulation rounds	-	10	Total number of simulation rounds
Traffic prediction window	-	5	Number of future rounds predicted for traffic
Minimum V2X resource ratio	-	0.2	Minimum guaranteed resource ratio for V2X
DQN batch size	-	32	Batch size for DQN training
DQN replay buffer capacity	-	10,000	Capacity of the experience replay buffer
DQN target network update frequency	-	10	Frequency of updating the target DQN network
DQN discount factor	γ	0.95	Discount factor for future rewards in DQN
DQN learning rate	-	0.001	Learning rate for DQN optimization
RL learning rate	α	0.1	Learning rate for reinforcement learning
RL action space dimensions	A_r, A_x	4 , 20	Dimensions of the RL action space
Number of RL episodes	-	100	Number of episodes for RL training
Temperature parameter	τ	1	Temperature for exploration in RL
Averaging window	Avg_T	20	Window size for averaging in RL

sharing modules, which then perform fine-grained resource allocation within each slice.

E. Fairness–Utilization Balanced Optimization

The DART-PRB framework seeks to achieve an optimal trade-off between fairness and resource utilization while meeting the diverse QoS requirements of heterogeneous service types. To this end, the framework employs multi-objective optimization techniques that jointly maximize resource utilization and ensure fairness among users.

Let $x = x_{u,n}$ denote the resource allocation matrix, where $x_{u,n}$ represents the fraction of PRB n allocated to user u. The overall optimization problem is formulated with two conflicting objectives:

$$\max_x U(x), and \ \max_x F(x) \tag{12}$$

s.t:

$$\sum_{u\in U} x_{u,n} \leq 1, \forall_n \in N,$$
$$x_{u,n} \geq 0, \forall_u \in U, n \in N,$$
$$R_u(x) \geq R_u^{min}, \forall_u \in U$$
$$D_u(x) \leq D_u^{max}, \forall_u \in U$$

where:

- $U(x) = \sum_{u\in U} R_u(x)$ denotes the overall resource utilization (i.e., the sum throughput across all users), with $R_u(x)$ calculated based on allocated resources and channel conditions.
- $F(x)$ represents the fairness metric. For example, using Jain's fairness index, we have:

$$F(x) = \frac{(\sum_{u\in U} R_u(x))^2}{|U| \sum_{u\in U} R_u(x)^2} \tag{13}$$

where is the total number of users. The constraints ensure that resource allocations remain feasible, and that QoS requirements (minimum data rate RR_u^{min} and maximum delay D_u^{max}) are satisfied.

To solve the above multi-objective problem, DART-PRB employs a scalarization technique, converting the two objectives into a single objective function via a weighted sum:

$$\max_x \alpha U(x) + (1-\alpha)F(x) \tag{14}$$

where $\alpha \in [0,1]$ determines the trade-off between throughput and fairness. To adaptively balance this trade-off under varying network conditions, the framework employs a reinforcement learning-based weight adjustment module that formulates the selection of α as a markov decision process and optimizes it via constrained policy optimization techniques. This integrated approach enables DART-PRB to dynamically adjust resource allocations, thereby enhancing overall network performance while satisfying diverse QoS requirements.

IV. Simulation Setup and Performance Evaluation

A. Simulation Environment and Parameter Settings

To evaluate the performance of the DART-PRB framework, we conduct extensive simulations using a comprehensive set of parameters that closely mimic real-world heterogeneous network scenarios. The simulation environment is designed to assess the efficiency and adaptability of the proposed resource allocation mechanism under various traffic conditions and network configurations.

The simulation environment in **Table 1** emulates a heterogeneous network scenario with a single base station serving V2X, eMBB, and mMTC services using N_{RB}=200 PRBs, where each PRB spans 360 kHz (derived from a subcarrier bandwidth f_{sc}=30 kHz and 12 subcarriers per PRB). The simulation runs for 10 rounds with 10 simulation steps per round, each step lasting T_{drop}=0.5 ms (equal to the TTI duration F_d), allowing fine-grained capture of network dynamics. Traffic models are defined such that V2X services have a packet arrival rate of λ_v=200 packets per second and a fixed packet size of 300 bytes, eMBB services have a session arrival rate of λ_e=300 sessions per second at a constant bit rate of 1Mbps, and mMTC services are characterized by a transmission probability p_{mMTC}^{tx}=0.75 with a maximum packet size of 128 bytes; the number of UEs is set as 8 for V2X, 4 for eMBB, and 12 for mMTC, with positions updated based on an average vehicle velocity of 22.22 m/s (80 km/h). A traffic prediction mechanism with a 5-round prediction window is implemented to anticipate future demands, while a minimum resource ratio of 0.2 is maintained for V2X services to guarantee QoS. The RL component employs a DQN trained with a batch size of 32 and a replay buffer capacity of 10,000, using a discount factor γ=0.95 and learning rate 0.001; the RL action space is defined by dimensions A_r=4 and A_x=20, and the agent is trained over 100 episodes with a temperature parameter τ =1. This comprehensive setup enables rigorous evaluation of the DART-PRB frameworks ability to dynamically allocate resources while ensuring fairness, high utilization, and adherence to QoS constraints in realistic network conditions.

B. Comparative Analysis of Resource Allocation Schemes

To demonstrate the superiority and novelty of the proposed DART-PRB framework, we conduct a comprehensive comparative analysis with state-of-the-art resource allocation schemes. **Table 2** presents a detailed comparison, highlighting the key features, capabilities, and limitations of each scheme.

The Advanced DART-PRB framework introduces a hierarchical multi-resolution allocation strategy that efficiently allocates resources at micro, meso, and macro levels based on the specific requirements of URLLC, eMBB, and mMTC services, thereby ensuring optimal resource utilization and adherence to diverse QoS constraints in heterogeneous networks. Unlike the Basic DART-PRB scheme, which relies on a heuristic allocation strategy based solely on service types and may lead to suboptimal performance and limited adaptability, Advanced DART-PRB leverages a proactive traffic prediction module that employs ensemble forecasting techniques—including RNNs, ARIMA models, and seasonal pattern analysis—to accurately estimate future traffic demands and preemptively optimize resource allocation, reducing the risk of over- or under-provisioning compared to reactive approaches like Traffic-based or RL-based Allocation. Furthermore, its multi-objective optimization approach simultaneously considers PRB utilization, fairness, and energy efficiency through an enhanced DQN engine that incorporates adaptive state compression, action masking, and twin Q-networks, while integrating service-specific constraints and fairness-utilization balancing comprehensively optimize network performance. Additionally, Advanced DART-PRB employs a channel-aware interference management scheme that uses spatial correlation-aware clustering, inter-slice interference mapping, and intra-slice interference control to effectively mitigate interference and improve network capacity and spectral efficiency, an aspect often overlooked by Basic DART-PRB and other existing schemes. Although these advanced techniques result in increased computational complexity compared to simpler schemes like Static Equal Allocation and Basic DART-PRB, the benefits of enhanced resource utilization, fairness, and energy efficiency outweigh the complexity trade-off, establishing Advanced DART-PRB

TABLE II
BENCHMARK SCHEME ANALYSIS

Scheme	Allocation Strategy	Traffic Awareness	Optimization Objective	Limitations
Advanced DART-PRB (Proposed)	Hierarchical multi-resolution allocation	Proactive ensemble forecasting	Multi-objective: PRB utilization, fairness, energy efficiency	Increased complexity due to advanced techniques
Basic DART-PRB	Heuristic allocation based on service types	Limited traffic awareness	Single objective: Maximize PRB utilization	Lack of fairness consideration and suboptimal performance
RL-based Allocation	Reinforcement learning-based allocation	Reactive to traffic patterns	Single objective: Maximize cumulative rewards	Limited scalability and adaptability to dynamic environments
Static Equal Allocation	Equal allocation among service types	No traffic awareness	Fairness-oriented	Inefficient resource utilization and inability to adapt to traffic variations
Traffic-based Allocation	Allocation proportional to traffic demands	Short-term traffic prediction	Throughput maximization	Neglects fairness and energy efficiency aspects

as a superior solution for efficient and adaptive resource allocation in next-generation 5G and beyond networks.

C. Discussion of Simulation Results and Insights

The simulation results provide a comprehensive assessment of the DART-PRB frameworks performance and its effectiveness in optimizing resource allocation in heterogeneous 5G and beyond networks. The proposed framework is evaluated using several critical performance metrics, including PRB utilization, spectral efficiency, fairness, energy efficiency, and SLA violation rates. The results demonstrate the superiority of the Advanced DART-PRB framework compared to the Basic DART-PRB and other state-of-the-art resource allocation schemes.

TABLE III
BENCHMARK SCHEME PERFORMANCE

Metrics	Advanced DART-PRB	Basic DART-PRB	RL-based	Static Equal	Traffic-based
DL PRB (%)	74.51	48.89	71.83	54.33	71.19
UL PRB (%)	71.57	46.34	73.82	55.19	79.57
DL V2X Outage (%)	0.67	0.11	1.68	1.20	1.80
UL V2X Outage (%)	1.20	1.80	1.70	2.00	5.98
DL Spectral Efficiency (bps/Hz)	5.55	2.25	3.01	2.24	2.66
UL Spectral Efficiency (bps/Hz)	3.51	2.33	2.66	1.94	2.32
DL Fairness Index	0.8011	0.8231	0.7446	0.9926	0.7759
UL Fairness Index	0.8172	0.8493	0.7272	0.9998	0.6493
SLA Violation Rate (%)	10.35	36.67	16.17	21.79	19.51
Energy Efficiency (Mbps/J)	0.25	0.03	0.04	0.01	0.02

Table 1 provides a detailed comparison of the proposed DART-PRB framework with state-of-the-art schemes, showcasing the average values of key performance metrics for each approach. The Advanced DART-PRB framework surpasses other schemes across most evaluated metrics. For DL PRB utilization, it achieves the highest average rate at 74.51%, reflecting its superior resource allocation efficiency. In UL PRB utilization, the framework records a rate of 71.57%, which is slightly below the RL-based Allocation (73.82%) and Traffic-based Allocation (79.57%) schemes. This difference arises because the DART-PRB framework emphasizes a balanced approach, optimizing not just resource utilization but also other vital metrics like fairness, energy efficiency, and SLA satisfaction. By adopting a multi-objective optimization strategy, it ensures robust overall performance rather than focusing solely on maximizing UL PRB utilization at the expense of other factors. Despite the marginally lower UL PRB utilization, the DART-PRB framework excels in minimizing V2X outage probabilities, achieving 0.67% for DL and 1.20% for UL—demonstrating its reliability for critical V2X communications. It also delivers the highest spectral efficiency, with 5.55 bps/Hz for DL and 3.51 bps/Hz for UL, outperforming competing schemes. This enhancement stems from its opportunistic resource sharing and adaptive slicing techniques. Additionally, the framework ensures strong fairness, with fairness index values of 0.8011 for DL and 0.8172 for UL, reflecting equitable resource distribution across service types and users. In contrast, the Static Equal Allocation scheme achieves the highest fairness index by evenly distributing resources, though this comes at the cost of reduced PRB utilization and spectral efficiency. In energy efficiency, the DART-PRB framework stands out with a value of 0.25 Mbps/J, far exceeding other schemes. This gain results from its multi-objective optimization and channel-aware interference management strategies. Furthermore, it maintains the lowest SLA violation rate at 10.35%, ensuring consistent satisfaction of service requirements across diverse service types. Overall, the DART-PRB framework delivers a well-rounded, high-performing solution across all key metrics.

Figures 3 and 4 present a detailed analysis of the DART-PRB frameworks DL and UL performance, respectively, across various metrics. The subplots in each figure showcase the frameworks performance in terms of PRB utilization, V2X outage probability, spectral efficiency, resource allocation fairness, SLA violation rate, energy efficiency, and overall performance comparison. These figures corroborate the findings from **Table 1**, providing a visual representation of the DART-PRB superiority in both DL and UL scenarios.

Table IV presents the average resource allocation ratios for each service type achieved by the different resource allocation schemes, along with a stability metric that indicates the consistency of the allocation ratios across the simulation rounds. The framework maintains a balanced allocation ratio among the service types, with 0.350 for V2X, 0.400 for eMBB, and 0.250 for mMTC. It also achieves a high stability metric of 0.900, indicating consistent allocation ratios throughout the simulation. The Static Equal Allocation scheme has a perfect stability metric of 1.000, as it maintains equal allocation ratios across all rounds, but this comes at the cost of lower overall performance, as demonstrated in **Table III**.

TABLE IV
AVERAGE RESOURCE ALLOCATION

Algorithm	V2X Ratio	eMBB Ratio	mMTC Ratio	Stability
Advanced DART-PRB	0.350	0.400	0.250	0.900
Basic DART-PRB	0.300	0.350	0.350	0.800
RL-based Allocation	0.250	0.300	0.450	0.750
Static Equal Allocation	0.330	0.330	0.340	1.000
Traffic-based Allocation	0.200	0.300	0.500	0.650

Table V summarizes the algorithm adaptation and convergence metrics for the different resource allocation schemes, providing insights into their ability to adapt to changing network conditions and converge to stable allocation strategies. Advanced DART-PRB framework exhibits a high convergence speed of 0.85, indicating its ability to quickly converge to stable allocation strategies. It also achieves the highest traffic adaptation score of 0.95, demonstrating its effectiveness in adapting to changing network conditions and traffic demands. The Static Equal Allocation scheme has a perfect convergence speed of 1.00, as it maintains a fixed allocation strategy throughout the simulation. However, it has the lowest traffic adaptation score of 0.30, highlighting its inability to adapt to varying network conditions.

TABLE V
ALGORITHM ADAPTATION AND CONVERGENCE METRICS

Algorithm	Convergence Speed	Traffic Adaptation
Advanced DART-PRB	0.85	0.95
Basic DART-PRB	0.75	0.80
RL-based Allocation	0.60	0.70
Static Equal Allocation	1.00	0.30
Traffic-based Allocation	0.90	0.60

Figure 12 illustrates the slicing ratio evolution of the Advanced DART-PRB framework over the simulation rounds, along with the average resource allocation by algorithm. **Figure 12 (a)** demonstrates the frameworks ability to dynamically adapt the allocation ratios based on the changing network conditions and traffic demands. This adaptability ensures optimal resource utilization and adherence to the diverse QoS requirements of different service types. **Figure 12 (b)** compares the resource allocation ratios achieved by different algorithms, highlighting the balanced and efficient allocation strategy employed by the Advanced DART-PRB framework.

Figure 13 presents the estimated throughput by service type for different resource allocation schemes. The Advanced DART-PRB framework achieves the highest throughput values for all three service types: V2X, eMBB, and mMTC. This superior performance shows the ability to efficiently allocate resources, maximize spectral efficiency, and adapt to the specific requirements of each service type. The results demonstrate the effectiveness of the Advanced DART-PRB framework in delivering high-quality services to users across different application scenarios.

In summary, the simulation results and analysis demonstrate the significant performance gains achieved by the Advanced DART-PRB framework compared to the Basic DART-PRB and other state-of-the-art resource allocation schemes. The proposed framework provides a well-balanced performance across multiple metrics, ensuring efficient resource allocation, high spectral efficiency, fairness, energy efficiency, and low SLA violation rates. The insights gained from these simulations highlight the effectiveness of the advanced techniques employed in the DART-PRB framework and its ability to adapt to the dynamic nature of heterogeneous 5G and beyond networks. The proposed framework serves as a comprehensive and efficient solution for tackling the challenges of resource allocation in these complex network environments.

REFERENCES

References

[1] M. Shafi et al., "5G: A tutorial overview of standards, trials, challenges, deployment, and practice," IEEE Journal on Selected Areas in Communications, vol. 35, no. 6, pp. 1201-1221, June 2017.

[2] S. Chen, Y.-C. Liang, S. Sun, S. Kang, W. Cheng, and M. Peng, "Vision, requirements, and technology trend of 6G: How to tackle the challenges of system coverage, capacity, user data-rate and movement speed," IEEE Wireless Communications, vol. 27, no. 2, pp. 218-228, April 2020.

[3] M. Alsenwi et al., "A survey on resource allocation techniques in 5G networks," IEEE Access, vol. 8, pp. 95979-96001, 2020.

[4] J. Navarro-Ortiz et al., "A survey on 5G usage scenarios and traffic models," IEEE Communications Surveys & Tutorials, vol. 22, no. 2, pp. 905-929, Secondquarter 2020.

[5] T. O. Olwal, K. Djouani, and A. M. Kurien, "A survey of resource management toward 5G radio access networks," IEEE Communications Surveys & Tutorials, vol. 18, no. 3, pp. 1656-1686, thirdquarter 2016.

[6] S. Zhang et al., "Wireless networks with energy harvesting and information transfer: A survey," IEEE Communications Surveys & Tutorials, vol. 20, no. 3, pp. 1983-2010, thirdquarter 2018.

[7] N. C. Luong et al., "Applications of deep reinforcement learning in communications and networking: A survey," IEEE Communications Surveys & Tutorials, vol. 21, no. 4, pp. 3133-3174, Fourthquarter 2019.

[8] H. Ye, G. Y. Li, and B. F. Juang, "Deep reinforcement learning based resource allocation for V2V communications," IEEE Transactions on Vehicular Technology, vol. 68, no. 4, pp. 3163-3173, April 2019.

[9] Z. Qin et al., "A cooperative deep reinforcement learning framework for resource management in blockchain-enabled industrial IoT networks," IEEE Internet of Things Journal, vol. 8, no. 8, pp. 7081-7093, 15 April15, 2021.

[10] Y. S. Nasir and D. Guo, "Multi-agent deep reinforcement learning for dynamic power allocation in wireless networks," IEEE Journal on Selected Areas in Communications, vol. 37, no. 10, pp. 2239-2250, Oct. 2019.

[11] A. Naeem et al., "A survey on interference management in 5G cellular networks: Challenges and approaches," IEEE Access, vol. 8, pp. 63770-63793, 2020.

[12] F. Qamar et al., "A survey on interference management in 5G small cell networks," IEEE Transactions on Vehicular Technology, vol. 69, no. 7, pp. 7239-7260, July 2020.

[13] R. Tandon et al., "A survey on coordinated multipoint (CoMP) systems," IEEE Wireless Communications, vol. 24, no. 3, pp. 6-13, June 2017.

[14] M. Feng, S. Mao, and T. Jiang, "Joint frame design, resource allocation and user association for massive MIMO heterogeneous networks with wireless backhaul," IEEE Transactions on Wireless Communications, vol. 17, no. 3, pp. 1937-1950, March 2018.

[15] I. Leyva-Mayorga et al., "Network slicing for 5G with SDN/NFV: Concepts, architectures, and challenges," IEEE

Communications Magazine, vol. 55, no. 5, pp. 80-87, May 2017.
[16] R. Tang et al., "Computing for network function virtualization: A survey," IEEE Access, vol. 6, pp. 33084-33106, 2018.
[17] G. Wang et al., "Resource allocation for network slices in 5G with network resource pricing," in IEEE GLOBECOM, 2017, pp. 1-6.
[18] G. Miao and J. Zander, "Resource allocation in multicell wireless networks," Eurecom Technical Report, 2007.
[19] D. Gesbert, S. Hanly, H. Huang, S. Shamai, and O. Simeone, "Multi-cell MIMO cooperative networks: A new look at interference," IEEE Journal on Selected Areas in Communications, vol. 28, no. 9, pp. 1380-1408, Dec. 2010.
[20] P. Zhao, L. Feng, P. Yu, W. Li, and X. Qiu, "A social-aware resource allocation for 5G device-to-device multicast communication," IEEE Access, vol. 5, pp. 15717-15730, 2017.
[21] S. Malhotra, F. Yashu, M. Saqib, D. Mehta, J. Jangid, and S. Dixit, "Deep reinforcement learning for dynamic resource allocation in wireless networks," arXiv:2502.01129, 2025.
[22] S. Yu and J. W. Lee, "Deep reinforcement learning based resource allocation for D2D communications underlay cellular networks," Sensors, vol. 22, no. 23, p. 9459, 2022.
[23] S. Hwang, H. Kim, H. Lee, and I. Lee, "Multi-agent deep reinforcement learning for distributed resource management in wireless powered communication networks," arXiv:2010.09171, 2020.
[24] H. Zhang, S. Chong, X. Zhang, and N. Lin, "A deep reinforcement learning based D2D relay selection and power level allocation in mmWave vehicular networks," IEEE Wireless Communications Letters, vol. 10, no. 5, pp. 1104-1108, May 2021.
[25] Y.-H. Xu, W. Zhou, Y.-G. Zhang, and G. Yu, "Stochastic game for resource management in cellular zero-touch deterministic industrial M2M networks," IEEE Wireless Communications Letters, vol. 11, no. 5, pp. 1000-1004, May 2022.
[26] W. Lee and K. Lee, "Resource allocation scheme for guarantee of QoS in D2D communications using deep neural network," IEEE Communications Letters, vol. 25, no. 3, pp. 887-891, March 2021.
[27] T. Zhang, H. Wang, X. Chu, and J. He, "A signaling-based incentive mechanism for device-to-device content sharing in cellular networks," IEEE Communications Letters, vol. 21, no. 7, pp. 1629-1632, July 2017.
[28] L. Chen, W. Zhang, and Z. Huang, "Decision tree based resource allocation in 5G networks," IEEE Transactions on Wireless Communications, vol. 22, no. 5, pp. 2891-2907, May 2023.
[29] R. AliHemmati, B. Liang, M. Dong, G. Boudreau, and S. Seyedmehdi, "Power allocation for underlay device-to-device communication over multiple channels," IEEE Transactions on Signal and Information Processing over Networks, vol. 4, no. 3, pp. 467-480, Sept. 2018.
[30] P. Mach and Z. Becvar, "Resource allocation for D2D communication with multiple D2D pairs reusing multiple channels," IEEE Wireless Communications Letters, vol. 8, no. 2, pp. 344-347, April 2019.
[31] L. Tang, Q. Tan, Y. Shi, C. Wang, and Q. Chen, "Adaptive virtual resource allocation in 5G network slicing using constrained Markov decision process," IEEE Access, vol. 6, pp. 61184-61195, 2018.
[32] R. Jain, D.-M. Chiu, and W. R. Hawe, "A quantitative measure of fairness and discrimination for resource allocation in shared computer system," Eastern Research Laboratory, Digital Equipment Corporation, Tech. Rep. DEC-TR-301, 1984.

Azharul Islam received his B.Sc. degree in Computer Science and Software Engineering (CSSE) from the American International University of Bangladesh (AIUB) in 2015 and later earned his M.Sc. degree in Electronics Engineering (EE) at Inha University, South Korea, 2021. He is currently pursuing his Ph.D. degree at the Department of Electrical and Computer Science Engineering (ECE) at Inha University, South Korea. Since 2019, he has been serving as a research assistant at the Mobile Telecommunication Research Lab (MTRL). His research interests include 5G technologies, heterogeneous wireless networks, artificial intelligence, resource allocation management, and vehicular communication networks.

KyungHi Chang (Senior Member, IEEE) received the B.S. and M.S. degrees in electronics engineering from Yonsei University, Seoul, South Korea, in 1985 and 1987, respectively, and the Ph.D. degree in electrical engineering from Texas A&M University, College Station, TX, USA, in 1992. From 1989 to 1990, he was a Member of the Research Staff with the Samsung Advanced Institute of Technology (SAIT), where he was involved in digital signal processing system design. From 1992 to 2003, he was with the Electronics and Telecommunications Research Institute (ETRI) as a Principal Member of the Technical Staff, where he led the design teams involved in the WCDMA UE modem and 4G radio transmission technology (RTT). He is currently with the Electrical and Computer Engineering Department, Inha University. His research interests include radio transmission technology in 3GPP LTE and 5G NR systems, public safety, and mobile ad hoc networks (especially for UAV), cellularV2X technology, non-terrestrial networks (NTNs), network intelligence for 6G, and applications of AI technologies. He was a recipient of the LG Academic Award, in 2006, the Haedong Best Paper Award, in 2007, the IEEE ComSoc Best Paper Award, in 2008, the Haedong Academic Award, in 2010, the SKT SafeNet Best Idea Award, in 2015, the KICS Outstanding

Academic Society Award, in 2018 and 2019, the MSIT Minister’s Commendation and a KICS Fellow, in 2020, and the Presidential Commendation, in 2021. He is currently the Chairman of the 6G Forum, the Executive Committee, the Technology Committee for National Integrated Public Network, and the Expert Committee in SafeNet Forum. He served as the Editor-in-Chief and the Executive Director for the Journal of Korean Institute of Communications and Information Sciences (KICS), from 2010 to 2012 and in 2013, respectively, and the Vice President for KICS, from 2017 to 2018, and since 2021. He also served as the Editor for ITU-R TG8/1 IMT.MOD.