ABUAD Journal of Engineering Research and Development (AJERD)
ISSN (online): 2645-2685; ISSN (print): 2756-6811

Volume 8, Issue 2, 52-65

Development of Adaptive Resource Allocation and Interference Mitigation for Spectrum Sharing in D2D-Enabled 5G Heterogeneous Networks: A Case Study of Urban Microcell Environments

Ashraf Adam AHMAD[1], Usman Bem ABUBAKAR[1,2], Fatai Olatunde ADUNOLA[1], Amina JIBRIL[1], Kulu Ahmad AMALO[3]

[1]*Department of Electrical/Electronic Engineering, Faculty of Engineering and Technology, Nigerian Defence Academy, Kaduna, Nigeria*
aaashraf@nda.edu.ng/usmanbem.abubakar2021@nda.edu.ng/foadunola@nda.edu.ng/amina.jibril2021@nda.edu.ng

[2]*Playout Center, Engineering Unit, Nigerian Television Authority, Abuja, Nigeria*
usmanbem.abubakar2021@nda.edu.ng

[3]*Electrical Technology Education Department, College of Technical Education, Kaduna Polytechnic, Kaduna, Nigeria*
kuluamalo@gmail.com

Corresponding Author: aaashraf@nda.edu.ng
Date Submitted: 17/02/2025
Date Accepted: 15/05/2025
Date Published: 16/05/2025

Abstract: *Device-to-device (D2D) communication in heterogeneous networks (HetNets) poses significant challenges in resource allocation and interference management, especially within 5G networks where spectrum sharing between cellular users (CUEs) and D2D user equipment (DUEs) is critical. This study developed an adaptive resource allocation framework using Long Short-Term Reinforcement Learning (LSRL), which integrated Long Short-Term Memory (LSTM) networks with Deep Reinforcement Learning (DRL) technique. The proposed approach addressed the dynamic nature of interference in urban microcell environments by leveraging a Hierarchical Data Format (HDF5) dataset generated from network simulations. These simulations incorporate diverse scenarios, including varying user densities, transmission power levels, and interference conditions. The LSRL-based scheme was evaluated against conventional DRL methods, demonstrating notable improvements in network performance. Specifically, the proposed framework achieved up to a 6.67% increase in sum throughput and an 8.2% enhancement in power efficiency, even under dense user conditions. Additionally, the LSRL model proved resilient to variations in D2D pair distances, maintaining robust spectral efficiency and quality of service (QoS). These findings underscore the potential of the LSRL-based adaptive approach for improving resource management in 5G HetNets, particularly in dense urban deployments, and provide valuable insights for optimizing next-generation wireless communication systems.*

Keywords: *Long Short-Term Reinforcement Learning (LSRL), Deep Reinforcement Learning (DRL), Long Short-Term Memory (LSTM), Device-to-Device (D2D) Communication, Heterogeneous Networks (HetNets)*

1. INTRODUCTION

The advent of 5G networks promises to revolutionize wireless communication by offering faster data rates, lower latency, and increased connectivity. With the rapid proliferation of devices and networks, 5G is anticipated to enable a variety of applications, such as improved mobile broadband, large-scale machine-type communications, and ultra-reliable low-latency communications [1]. The transition from 4G to 5G brings significant architectural changes, such as small cells, massive MIMO, and beamforming, to improve network performance [2]. However, despite these advancements, the problem of increasing network density and high frequency reuse in 5G networks introduce significant challenges related to interference management [1,2]. Effectively addressing these challenges is essential to maintain optimal network performance and ensure reliable communication.

Interference has long been a challenge in wireless communication, but its impact is even more pronounced in 5G due to the denser deployment of network elements and the variety of services being supported [1]. While traditional techniques such as frequency planning, power control, and interference cancellation were effective in 4G networks, they may not be sufficient to handle the complexities of 5G [3]. Additionally, new types of interference emerge in 5G networks, including inter-cell, co-channel, and adjacent channel interference, which require more sophisticated solutions [1,3]. The adoption of millimeter-wave (mmWave) frequencies further complicates interference management, as factors like atmospheric absorption and scattering create unpredictable interference patterns [3]. Addressing these challenges necessitates innovative approaches that can dynamically adapt to network conditions.

Machine learning (ML) is a promising approach for mitigating interference in 5G networks. By utilizing ML algorithms, networks can predict and adjust to interference patterns, optimize resource distribution, and enhance overall performance. ML-driven methods can process large volumes of data from different sources, such as sensors, user devices, and network infrastructure, to detect and address interference in real time. Training ML models on large datasets of network activity enables them to identify patterns and anomalies, helping to prevent interference and enhance performance. Additionally, ML can be utilized to optimize network parameters such as antenna tilt and azimuth, ensuring better coverage and minimizing interference [3]. These capabilities make ML a vital tool in the effort to enhance 5G network reliability and efficiency.

The integration of ML for interference mitigation in 5G networks is crucial for achieving the full potential of this technology. As 5G continues to evolve, ML-driven solutions can enable intelligent interference management, leading to improved spectral efficiency, latency, and user experience. Additionally, ML-enabled 5G networks can support the creation of new applications like intelligent transportation systems, smart cities, and industrial automation, all of which depend on ultra-reliable and low-latency communication [1,2]. Further research is needed to explore and optimize ML-based interference mitigation strategies, ensuring that 5G networks can meet the demands of future communication systems. By addressing these challenges, ML can contribute significantly to the realization of a highly efficient and resilient 5G ecosystem.

The aim of this research is to develop an improved adaptive 5G network intelligent agent for interference mitigation using Long Short-Term Reinforcement Learning (LSRL). To achieve this, the study focused on constructing a Hierarchical Data Format (HDF5) dataset from 5G network sensors, capturing key parameters such as signal strength, frequency, and noise levels. Additionally, an LSRL-based intelligent agent was designed to optimize resource allocation and mitigate interference in heterogeneous networks (HetNets). The developed agent was evaluated against existing Deep Reinforcement Learning (DRL) approaches using key performance metrics, including scalability, energy efficiency, throughput, and quality of service (QoS). The findings highlight the LSRL agent's superior ability to enhance network performance while reducing interference, demonstrating its potential for real-world 5G deployment.

The paper was structured into five main sections. The Introduction outlined the study's background, problem statement, and research objectives. The Review of Related Researches discussed existing works relevant to adaptive resource allocation and interference mitigation in 5G heterogeneous networks. The Methodology section detailed the development of the Hierarchical Data Format (HDF5) dataset, the adaptive 5G network intelligent agent, and the performance evaluation setup. In the Results and Discussion section, the outcomes of the simulations were presented and analyzed, highlighting the effectiveness of the proposed LSRL-based approach. Finally, the Conclusion summarized the key findings, implications, and potential future research directions. The paper concluded with References that supported the study.

2. REVIEW OF RELATED RESEARCHES

Resource allocation and interference mitigation are critical issues in 5G and beyond-5G networks, as increasing network density and dynamic user environments exacerbate signal degradation and performance issues. Recent research has explored various AI-driven approaches, including deep reinforcement learning, neural networks, and optimization-based techniques, to enhance interference management and improve key performance metrics such as SINR, throughput, and energy efficiency. The following summaries highlight notable contributions in this domain, outlining their methodologies, key findings, and existing limitations.

Fan et al., [4] in their work explored EMI and IEMI classification for High-Speed Rail (HSR) wireless communications using Long Short-Term Memory (LSTM) deep learning. They proposed a dynamic system model that allows for real-time detection based on a large dataset gathered from various scenarios. While achieving effective classification, the study does not extensively discuss its generalizability to other transportation systems, and practical implementation may face challenges related to data collection, storage, and processing. Similarly, Sejan et al., [5] in their research explored Recurrent Neural Network (RNN) approaches for co-channel interference mitigation, demonstrating that in a network with 10 connected devices, Bidirectional LSTM (Bi-LSTM) outperforms LSTM and Gated Recurrent Unit (GRU) in terms of mean squared error and sum rate, though the study's focus on networks with only 10 and 20 devices limits its applicability to larger-scale deployments. Farhan and Aal-nouman [6] introduced a Reinforcement Learning-based strategy for joint interference mitigation and resource allocation in dense beyond-5G heterogeneous networks. This approach achieved a sum-rate improvement of at least 10% and 25% over other existing matching and power allocation methods, respectively. However, the dependence on Q-learning, which involves managing a large number of states and actions, along with the computational overhead of a two-step process, presents challenges in highly dynamic environments. The study by Alruwaili et al., [7] introduced an Incremental Radial Basis Function (RBF)-based approach for cross-tier interference mitigation in resource-constrained dense IoT networks, optimizing real-time interference patterns, yet challenges in scalability and implementation efficiency persist due to the complexity of multi-tier networks. Lastly, Irkicatal et al., [8] in their work applied deep reinforcement learning in Rate-Splitting Multiple Access (RSMA) to optimize precoders and power allocation in multi-antenna interference channels, but limitations arise from imperfect channel side information at the transmitter, and the study does not extensively explore implementation challenges in dynamic environments.

Anand et al, [9] in their work introduced the Machine Learning Multi-Classification and Offloading Scheme (MLMCOS) to mitigate co-tier interference in 5G HetNets. This scheme outperformed the Proportional Fair (PF) scheduling algorithm, Variable Radius and Proportional Fair scheduling (VR+PF), and a Cognitive Approach (CA) in terms of throughput, delay, and Packet Loss Ratio (PLR). However, it does not address issues related to cross-tier interference or energy consumption. Al-Jumaily et al., [10] in their work explored interference analysis and protection distance optimization for 5G Base Stations and Fixed Satellite Services Earth Stations using Artificial Neural Network Learning Models (ANN-LMs) in the 3.4-4.2 GHz frequency range, highlighting the need for better models but lacking discussion on economic feasibility, scalability, and adaptability to evolving network conditions. Udoh et al [11] in their research compared frequency planning and artificial neural networks (ANN) for 5G interference mitigation using SIMULINK, finding ANN to be more effective with a mitigation level of 23.89% compared to 21.69% for frequency planning, though real-world applicability was not extensively analyzed. Anand et al., [12] in their work investigated Open-Radio Access Network (O-RAN) in 5G, detailing the split into Remote Radio Unit (RRU), Distributed Unit (DU), and Centralized Unit (CU) for improved QoS in HetNets but did not discuss energy efficiency or specific QoS optimization mechanisms. Warrier et al., [13] in their study proposed power control and interference mitigation for 5G-connected UAVs were formulated as an optimization problem to maximize Signal to Interference and Noise Ratio (SINR), introducing a deep Q-learning (DQL) algorithm to manage air-ground interference, though it did not address latency, energy efficiency, or scalability concerns. Similarly, Elsayed et al., [14] in their work developed an AI-enabled interference mitigation system for UAVs in urban 5G networks, achieving a 41.66% performance improvement over traditional techniques using deep reinforcement learning (DRL) for centralized power control, but it did not extend findings to terrestrial networks, where interference dynamics differ due to mobility and altitude variations.

Ahmad and Hussain [15] in their research addressed co-channel interference in heterogeneous networks, particularly in Public Safety Networks (PSN), LTE-based Railway Networks (LRN), and Unmanned Aerial Vehicles (UAVs) by leveraging deep learning (DL)-based Coordinated Multipoint (CoMP) techniques to enhance Mobile User (MU) performance. While the approach improves resource utilization through offloading PSN User Equipment (UEs) to LRN or UAVs, it does not extensively explore potential drawbacks such as signaling overhead and handover complexities. Similarly, Konan et al., [16] in their work, proposed a supervised learning-based approach using logistic regression and Batch Gradient Descent for predicting Signal-to-Interference-plus-Noise Ratio (SINR) in 5G networks, achieving an accuracy of 0.90 and an error of 0.1. However, scalability to larger datasets and alternative ML techniques are not extensively analyzed. Iqbal et al., [17] in their study introduced a self-adaptive resource allocation scheme for ultra-dense 5G heterogeneous networks (HetNets), utilizing independent and cooperative learning techniques to optimize small cell transmit power, significantly improving QoS for Macro and Small User Equipments (MUEs, SUEs). Despite its effectiveness, the study lacks an in-depth discussion on reinforcement learning (RL) implementation challenges. A related study by Iqbal et al., [18] proposed a Q-learning-based adaptive power allocation algorithm for multi-tiered 5G small cell (SC) HetNets, introducing a reward function to ensure minimum SINR requirements; however, real-world validation and alternative interference mitigation techniques are not explored. Yang et al., [19] in their work proposed an RL-assisted full dynamic beamforming method for inter-cell interference (ICI) mitigation in 5G downlink, combining beamforming and dynamic Q-learning, enhancing SINR while reducing computational complexity, though its applicability to uplink and varying network conditions remains underexplored. Also, Yadav and Tripathi [20] in their research presented a machine learning-based 3D MIMO beamforming system for interference mitigation in 5G networks, improving device-to-device (D2D) communication, but without addressing the implementation challenges or limitations of its spatial distance SVM algorithm.

Osman et al, [21] in their work proposed an interference avoidance distributed deep learning model for IoT and device-to-any-destination communication by leveraging Lagrange optimization to predict optimal distances for uplink and downlink communication, improving system throughput and energy efficiency; however, it lacks scalability analysis for large-scale networks and adaptability to dynamic conditions. Afolabi et al., [22] in their work explored a reinforcement learning-based Q-learning model to manage interference in heterogeneous networks with macrocell and picocell base stations, demonstrating improved throughput but without addressing real-world deployment challenges, including varying user densities and interference patterns. Similarly, Guo et al., [23] in their research presented a regression-based algorithm using a multilayer perceptron (NN-MLP) for uplink interference identification and SINR prediction in 5G ultra-dense networks, improving interference modeling but lacking insights into model complexity and interpretability. Wu et al., [24] in their study introduced an Adaptive Deep Learning-based Autoencoder (AE) for beyond-5G multi-user interference channels, outperforming traditional equalizers in interference mitigation but without a detailed analysis of computational complexity and security vulnerabilities. Wang et al., [25] in their work proposed a Self-Learning Interference Mitigation (SLIM) scheme using Multi-Agent Reinforcement Learning (MARL) for autonomous networks, reducing computational complexity via Mean Field Theory but focusing only on downlink interference and being limited to a specific urban model (3GPP dual-stripe). Also, Elsayed et al., [26] in their work applied transfer reinforcement learning to 5G-NR mm-Wave networks for joint user-cell association and beam selection, showing a 12% rate improvement under mobility and a 29% convergence speedup with TQL, yet it overlooked broader reinforcement learning variations and additional network optimizations.

Barazideh et al., [27] in their work proposed a reinforcement learning-based interference mitigation framework for terahertz communication networks, employing a multi-armed bandit approach to dynamically adjust thresholds and reduce interference power, achieving superior bit-error-rate performance compared to traditional methods; however, its reliance on large labeled datasets for training deep neural network models presents a limitation. Finally, Mismar et al., [28] in their research formulated the joint beamforming, power control, and interference coordination problem in the downlink direction as an optimization task to maximize users' received SINR, introducing a deep reinforcement learning solution that enables multiple simultaneous actions using binary encoding at base stations. However, the exhaustive search approach required for solving the non-convex optimization problem results in exponential runtime complexity, making it computationally intensive and time-consuming.

Despite significant advancements in interference mitigation for 5G networks through AI-driven techniques like deep reinforcement learning, neural networks, and Q-learning, several research gaps remain. Many existing studies focus on specific network conditions, limiting their applicability to real-world, heterogeneous environments with varying user densities, mobility patterns, and interference types. Scalability challenges, particularly in large-scale networks, are often overlooked, as is the computational complexity of implementing these models in practical deployments. Additionally, while some studies propose adaptive learning-based interference mitigation, their adaptability to dynamic and rapidly changing network conditions remains insufficiently explored. Furthermore, the integration of AI-driven interference mitigation with other key 5G technologies, such as massive MIMO, network slicing, and ultra-reliable low-latency communication (URLLC), is not completely addressed. Addressing these gaps is essential to enhancing the reliability, efficiency, and real-world applicability of interference mitigation strategies, ensuring 5G networks achieve their full potential in supporting diverse applications across different environments. As such, this research proposed a machine learning algorithm in order to mitigate interference among others in 5G in order to meet its full potential.

3. METHODOLOGY

This section outlines the methodology for developing an enhanced adaptive intelligent agent for interference mitigation in 5G networks. It covers the modeling of 5G heterogeneous networks, random deployment of user equipment (UEs), analysis of cross-tier and co-tier interference, and the roles of macrocell and femtocell base stations. Additionally, it details the calculation of data rates, quality of service (QoS) assessment, and the generation of a synthetic 5G network dataset. The development process of the Long Short-Term Memory (LSTM) model for temporal data prediction, along with its integration with deep reinforcement learning to create an adaptive recurrent deep learning agent for efficient QoS distribution, is discussed.

The system works by employing an adaptive resource allocation framework designed to optimize spectrum usage and mitigate interference in D2D-enabled 5G heterogeneous networks. The framework integrates Long Short-Term Memory (LSTM) networks with Deep Reinforcement Learning (DRL) to form the Long Short-Term Reinforcement Learning (LSRL) model, which dynamically adapts to varying network conditions, such as user density and transmission power. The methodology involves generating a Hierarchical Data Format (HDF5) dataset through network simulations, capturing metrics like throughput, transmission power, interference levels, and quality of service (QoS) under different scenarios. MATLAB 2023b was used as the simulation tool to implement the LSRL model and evaluate its performance. The system's performance was assessed using key metrics, including sum throughput, power efficiency, scalability, and QoS maintenance, with comparative analysis against existing DRL-based approaches to demonstrate the effectiveness of the proposed LSRL model.

3.1 Development of Hierarchical Data Format (HDF5) Dataset

The conventional dense cell deployment in 5G heterogeneous networks often leads to intra-cell (co-tier) and inter-cell (cross-tier) interference, which remains a major factor contributing to network performance degradation. Therefore, minimizing interference is essential to enhance the quality of service (QoS) experienced by user equipment. In a 5G HetNet, where a fixed macrocell base station (eNB) is positioned at the center and multiple small cell evolved node Bs (SCeNBs) and device user equipment (DUEs) are randomly distributed, the coverage areas of the two-tier base stations—macrocell and small cell—may overlap. As a result, each DUE can fall within the range of either the eNB or an SCeNB, influencing the interference levels and overall network performance.

This research generated a synthetic dataset, comprising around 1000 diverse seed values, to simulate the behavior of a 5G network under various conditions, mimicking real-world data where complex problem datasets are often unavailable. The dataset includes network states (representing conditions such as throughput, interference, and SINR), interference levels (caused by both device-to-device (D2D) and cellular users), and resource allocation parameters (capturing power and spectrum allocation strategies). Given the dynamic nature of network behavior, where interference and throughput fluctuate over time, historical network states were recorded to create time-series data for the synthetic dataset. The generated dataset was then converted into a hierarchical format using the .hdf5 extension to reduce storage size, enhance querying efficiency, enable faster data access (random access and optimized I/O), and improve data loading and saving across multiple platforms. For optimal model performance, the dataset underwent preprocessing, including normalization of interference and signal-to-noise ratio (SNR) values and splitting into training, validation, and testing sets. The flowchart for the creation of the 5GNetwork.hdf5 dataset is presented in Figure 1.

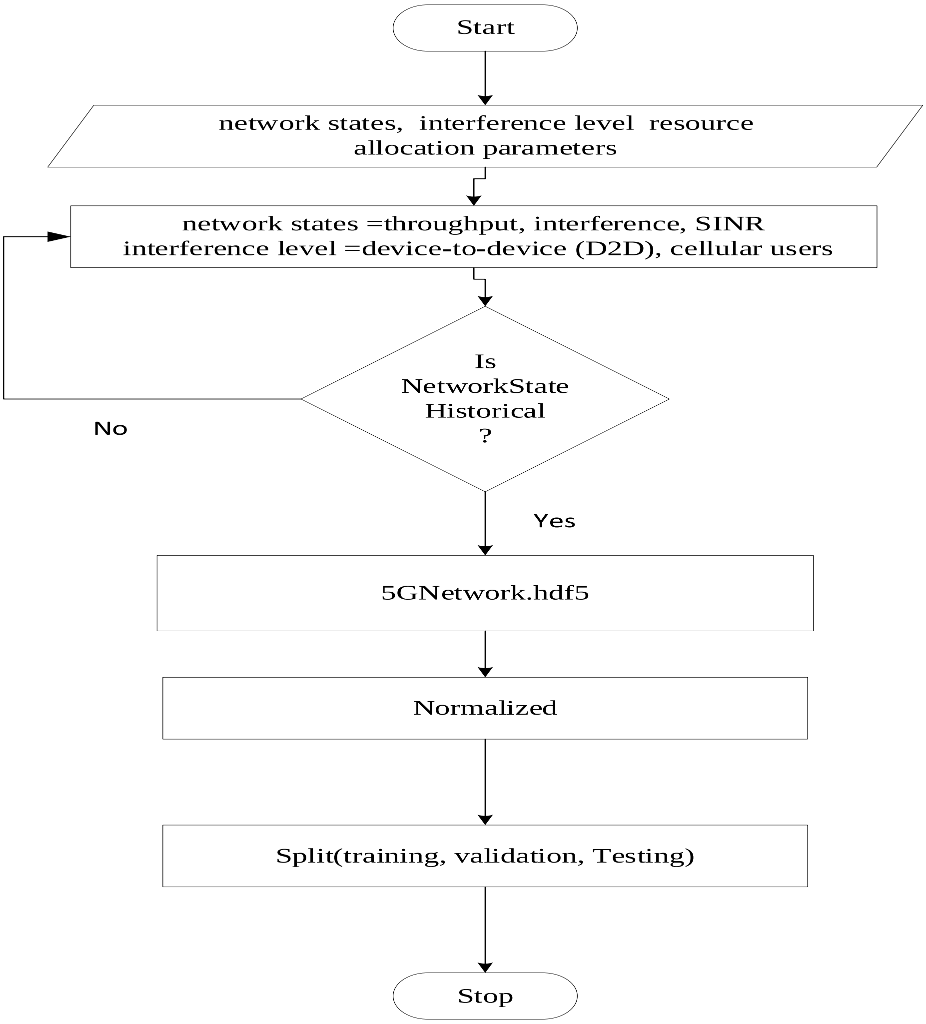


Figure 1: 5GNetwork.hdf5 flowchart

The implementation process of Figure 1 begins by initializing variables and configurations for network simulation, defining key metrics such as throughput, interference levels, SINR, and resource allocation parameters that characterize the network's dynamic state. Network states and interference levels are captured dynamically over time using loops or simulation tools like MATLAB's random data generation functions to mimic realistic network variations. At a decision point, the program evaluates whether enough historical network states have been recorded. If the data is insufficient, the process loops back to collect more; otherwise, it proceeds to data organization and conversion. The collected data is then structured and saved in the HDF5 format using MATLAB's h5create and h5write functions. Data preprocessing involves normalization, ensuring consistency across metrics by scaling SINR and interference values to a standard range (e.g., 0 to 1) for improved efficiency in machine learning models. The dataset is then split into training, validation, and testing subsets (e.g., 70:20:10 ratio) to facilitate the model development. Finally, the processed data is stored in HDF5 format, optimizing storage efficiency, retrieval speed, and cross-platform compatibility. The program terminates once the dataset is fully prepared and ready for use in LSTM training and reinforcement learning simulations.

3.2 Development of the Adaptive 5G Network Intelligent Agent

5G heterogeneous networks exhibit time-dependent behaviors such as channel state variations, user mobility, and traffic fluctuations, which can be effectively modeled using Long Short-Term Memory (LSTM) networks. LSTMs, a specialized form of Recurrent Neural Networks, are designed to capture long-term dependencies while mitigating the vanishing gradient problem. The hierarchical network dataset serves as input to the LSTM network, comprising sequences of historical network states such as interference, power levels, and SINR values. This data is structured in a sliding window format, where each window represents a snapshot of network conditions over time. The LSTM layer captures temporal dependencies, leveraging its hidden state to retain relevant information, which is crucial for predicting future network conditions. The processed sequence is then passed through a fully connected layer that maps LSTM outputs to predicted future network states, such as interference levels and throughput [28]. Traditional deep reinforcement learning (DRL) agents, which serve as device-to-device (D2D) pairs in this context, rely on the current state for decision-making. Each agent interacts with the environment to learn optimal strategies based on the Markov Decision Process, where actions result in rewards or penalties, influencing state transitions. Under varying network conditions, the agent dynamically selects appropriate communication modes and optimal resource blocks from available radio resources. In D2D-enabled cellular HetNets, each D2D user's transmitter acts as an agent, evaluating network states within a given time slot based on environmental conditions. The state observed by any agent, 's_l' can be expressed by Equation (1) [29]:

$$s_l = \left(A_{RB}, A_{CSI}, A_{QoS}\right) \in S \tag{1}$$

where: A_{RB} represents Resource block (RB) occupancy status, A_{CSI} is Channel status Information, A_{Qos} id Quality of Service for all users. In a D2D-enabled HetNet, the agent operates in different modes and has a specific number of resource blocks for communication. Depending on the current state conditions, the agent will determine the optimal mode selection, transmit power, and RB allocation. So, the action of an agent 'a_l' is mathematically expressed by Equation (2) [29].

$$a_l = (A_{MS}, A_{PC}, A_{RB}) \in A \tag{2}$$

where A_{MS} referred to mode selection, A_{PC} referred to transmit power control, A_{RB} referred to Resource Block assignment $a_l \in A$ under the current state s_l. In a device-to-device communication, the DUEs are further separated into a homogenous independent density. Where $k\varepsilon\{1,2,\dots,NC\}$, $l\varepsilon\{1,2,\dots,NC\}$, and $m\varepsilon\{1,2,\dots,NC\}$ represent the set of CUEs, DUEs, and SUEs respectively. In a heterogeneous cellular network the DUEs, SUEs, and CUEs can communicate simultaneously by sharing the same radio resources. By utilizing the device-to-device communication mode, DUEs can communicate with each other directly, and CUEs and SUEs can communicate by using cellular and small-cell communication modes. The interference scenarios for this system include: i) eNB receiving signals from the transmitter of SUE and DUE; ii) DUE receiver receiving signals from the transmitter of CUE and SUE; iii) DUE receiver receiving signals from the transmitters of other D2D pairs; and iv) SCeNB receiving signals from the transmitter of DUE and CUE.

Recall that the sharing of same radio resources by multiple SUEs and DUEs, eNB and SCeNB causes mutual interference. Therefore, it becomes eminent to maximize the overall system throughput under the minimum quality of service requirements of all users. Hence, the objective function and constraints for resource allocation, mode selection, and power control of resource management challenges withing the cellular, small cell, and device to device users can be expressed mathematically as equation (3) [29]:

$$P_1 = \max_{\psi_l^f, \psi_m^f, \psi_l^m, P_k^C, P_l^D, P_m^S} R_{overall} = \sum_{i\epsilon C}[Blog_2(1+\gamma_k^m) + \sum_{l\in D}\gamma_l^f[Blog_2(1+\gamma_k^m) + \sum_{k\epsilon S}\gamma_l^f[Blog_2(1+\gamma_k^m)] \tag{3}$$

Subject to;

$$\begin{aligned}
&C1: \gamma_k^M \geq \gamma_{k,th}^M \quad \forall_k \in C\\
&C2: \gamma_l^D \geq \gamma_{l,th}^D \quad \forall_k \in D\\
&C3: \gamma_m^S \geq \gamma_{m,th}^S \quad \forall_k \in S\\
&C4: 0 \leq P_k^C \leq P_{max,}^C \quad \forall_k \in C\\
&C5: 0 \leq P_l^D \leq P_{max,}^D \quad \forall_l \in D\\
&C6: 0 \leq P_m^D \leq P_{max,}^D \quad \forall_m \in S\\
&C7: \textstyle\sum_{i\epsilon C}\psi_l^f \leq 1, \psi_l^f \in \{0,1\}, \forall l \in D,\\
&C8: \textstyle\sum_{i\epsilon C}\psi_m^f \leq 1, \psi_m^f \in \{0,1\}, \forall m \in \mathrm{S}
\end{aligned}$$

As the agent interacts with the environment, the changes in the environmental status is described by transition probability. For instance, an agent executes an action, a, in a time slot, t, and state, s, and gains a reward, r. The probability that the agent will move to the next state, s′, at time t + 1 is defined by the transition probability, $P_r(s'|s,a)$ Pr (s′|s, a). Therefore, the transition probability, $P_r(s'|s,a)$ is the probability that an agent executes the action $a \in A$ under the state $s \in S$ and moves into a new state $s' \in S$ expressed by Equation (4) [29]:

$$P_r(s'|s,a) = \begin{Bmatrix} 1, & S' = state(a), \\ 0, & otherwise. \end{Bmatrix} \tag{4}$$

In reinforcement learning, the optimization problem is achieved using a reward function. Thus, to maximize the general system sum rate, every agent makes a joint decision for resource allocation, power control and mode selection. The reward for the gent is given by Equation (5) [29].

$$R_l(s,a) = \begin{Bmatrix} R_j^D, & if\ condition\ C2, C5, and\ C7\ are\ met \\ -1, & otherwise. \end{Bmatrix} \tag{5}$$

where C1 is the minimum signal to noise ratio, C2 is Resources Block availability, and C3 is Data rates. The objective of the agent is to learn the strategy for maximizing the rewards, the action is selected based on the quality of service (signal-to-noise ratio, resources block availability, and data rates). The decision-making rule of the agent is defined by the policy. In every state, s, the strategy of the agent to execute an action a, is represented by policy π(s, a), where we have a ∈ A, s ∈ S, and ∑a∈A π(s, a) = 1, ∀s ∈ S. Integrating the LSTM-based predictions to predefine the network, the agent can anticipate future states and optimize its actions in advance. In D2D-enabled networks, interference patterns and user behaviours change dynamically. The LSTM can track these changes over time, providing a better input for the DRL agent.

The LSRL agent uses an ϵ-greedy strategy to avoid local optimal solutions and receive a new experience when it interacts with the environment. During the training period, the agent will use the ϵ-greedy strategy to select the next action, a, from Q(s, a; θ), obtain reward rt immediately and move into the new state s_{t+1} at time, t + 1 and finally store the quadruple information (s_t, a_t, r_t, s_{t+1}) in replay memory. During network training, the agent leverages sample experience data stored in the replay memory to enhance the convergence speed and stability of the algorithm. After each iteration, the main LSTM parameters (θ) of the network model in the action-value function are synchronized with the target network. The algorithm for the LSRL agent is highlighted.

LSRL Agent Algorithm

1. **Input:** Trained Synthetic hierarchical data 5G HetNets cellular network environmental parameters with D2D parameters;
2. **Output:** LSRL model training and decision results in the testing phase, power control, and Effective resource allocation;
3. **Initialization:** target network from LSTM parameters, Set θ = θ′, Q(s, a: θ) and Q(s, a: θ′) target policy, π∗, Experience replay memory D, and capacity N
4. Define learning rate η, discount factor γ, batch size B, and exploration rate ϵ.
5. Initialize state sequence S_t with an initial observed state S_0.

Training Phase:

6. **for** episode e=1, 2,...,Ee:
7. Reset the environment and observe the initial state S_0.
8. Initialize the state sequence $S_t = [S_0]$.
9. **for** each time step t=1, 2 ,...,T:
10. Predict action probabilities A_t using the LSTM model:
11. $A_t = \text{LSTM}(S_t;\theta)$
12. Select an action a_t using ϵ-greedy strategy:
13. With probability ϵ, select a random action.
14. Otherwise, select $a_t=\text{argmax}(A_t)$
15. Execute action a_t and observe reward r_t and next state s_{t+1}.
16. Append s_{t+1} to the state sequence S_t, maintaining a fixed sequence length.
17. Store transition (S_t, a_t, r_t, S_{t+1}) in replay memory D.
18. Sample a random minibatch of transitions (S_b, a_b, r_b, S_{b+1}) from D.
    Compute the target y_b:
    $y_b = r_b+\gamma \max_{a'} \text{LSTM}(S_{b+1},a';\theta')$
    Update the LSTM model by minimizing the loss function:
    $L(\theta) = \frac{1}{B}\sum_b \big(y_b - LSTM(S_{b+1}, \grave{a}; \grave{\theta})\big)$
19. Periodically update the target network parameters:
    $\theta' \leftarrow \theta$
20. **end for**
21. **end for**

Window programming editor of MATLAB was then used to codify the LDRQ Network algorithm. The main file that controls the entire program was named LSRL Network. Equally, a standalone user-friendly graphical user interface simulator was developed to assist those with little or no programming knowledge in using the developed LSRL Network MATLAB simulator. The interface of the developed LSRL Network MATLAB simulator is shown in Figure 2.

As shown in Figure 2, the developed LSRL Network Simulator interface consists of two main panels: the network parameters panel and the algorithms qualitative metrics panel. The algorithms panel enables users to conduct a comparative analysis between the LSRL Network Simulator and the results obtained from the DRL agent implementation in [30]. Additionally, users can simulate various network density scenarios using the LSRL Network Simulator by adjusting parameters within the network parameters panel, which includes settings such as network size and the total number of DUEs

3.3 Performance Evaluation Setup

Evaluating the performance of any developed system is essential. Therefore, the proposed system was assessed using key metrics such as throughput, transmit power, and channel gain while analyzing their variations as the distance and number of DUEs changed within the heterogeneous network. The comparison in this study was conducted with similar research that employed the same dataset and evaluation metrics, ensuring a fair and accurate performance assessment. Specifically, the performance of the proposed system was compared with the results from [30], which utilized a similar interference mitigation approach based on a deep reinforcement learning agent. To maintain consistency, the default simulation parameters from [30], as outlined in Table 1, were adopted and explanations/justification given afterwards.

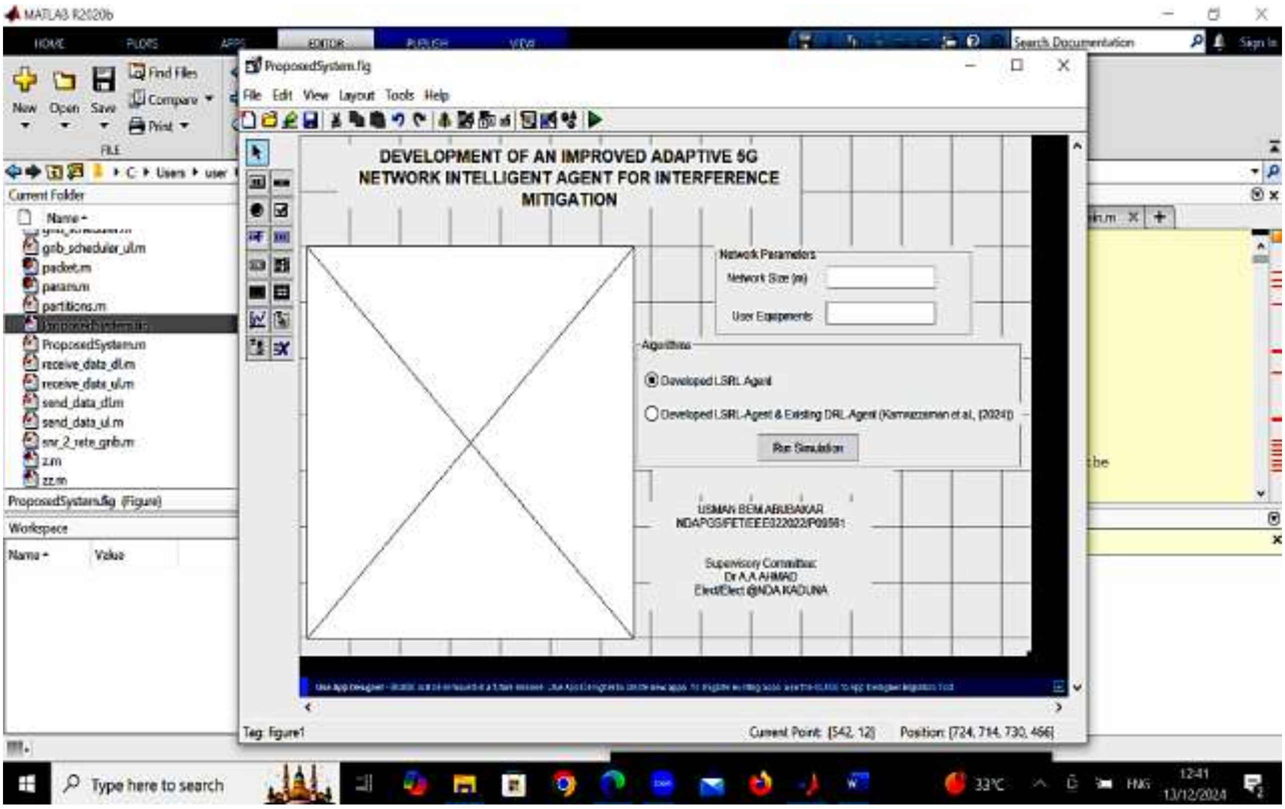


Figure 2: LSRL network simulator

Table 1: Simulation parameters (adopted from [30])

Parameters	**Values**
Cell Radius	500
SCeNBs density	10-5
DUEs density	10-3
CUEs maximum transmission power	23 (dBm)
DUEs maximum transmission power	10-20 (dBm)
SUEs maximum transmission power	10-20 (dBm)
Path loss exponent	3 and 4
SINR threshold for CUE, SUE, and DUE	8 dB
Noise power	-118 dBm
Shadowing standard deviation	8 dB
RB Bandwidth	180 KHz
Learning rate	0.01
Discount factor	0.9
e-greedy	0.1
Replay memory capacity	2000
Mini-batch size	32
Number of steps in each epoch	20

The cell radius from Table 1 defines the coverage area of macrocell base stations (eNBs), with a 500 m radius representing a typical urban microcell size. This allows for realistic modeling of interference and resource sharing in dense environments, ensuring the system is tested under practical network scenarios. The density of small-cell evolved Node B (SCeNB) deployment within the macrocell coverage simulates a sparse small-cell environment, enabling the evaluation of cross-tier interference mitigation techniques crucial for 5G heterogeneous networks. Similarly, the density of device-to-device (DUE) communications reflects the number of direct communication devices per unit area. A higher density ensures rigorous testing of the interference management capabilities of the LRLS agent, particularly for DUEs sharing spectrum resources.

Cellular user equipment (CUE) maximum transmission power represents the typical power output of mobile devices in 5G networks. This parameter directly impacts SINR calculations and ensures a realistic evaluation of power control mechanisms. The range of transmission power for DUEs and small-cell user equipment (SUEs) accounts for varying power levels in D2D communication and small-cell operations, enabling a comprehensive assessment of power control strategies under different transmission conditions. The path loss exponent models signal attenuation over distance, distinguishing between urban (higher exponent) and suburban (lower exponent) environments. These values test the system's adaptability to varying propagation conditions and interference scenarios.

The minimum SINR threshold defines the quality of service (QoS) requirements for maintaining reliable communication, ensuring the developed system meets realistic user performance expectations while effectively managing interference. Noise power, representing thermal noise in the system, is a critical component in SINR calculation, providing a baseline for assessing signal quality under different interference scenarios. The bandwidth per resource block reflects standard allocation in 5G networks, directly affecting achievable data rates and ensuring compatibility with real-world 5G deployments.

The learning rate controls the pace of updates to the reinforcement learning model, balancing stability and convergence speed to ensure efficient training without overshooting optimal solutions. The discount factor, set at 0.9, ensures the agent considers both immediate and long-term performance, which is crucial for sustained interference mitigation. The exploration rate, set at 0.1, balances the agent's exploration of new actions with the exploitation of learned strategies, promoting effective learning while avoiding suboptimal local solutions. Replay memory, with a capacity of 2000, stores past experiences for training the reinforcement learning agent, ensuring sufficient diversity in training data while managing memory usage efficiently. The mini-batch size of 32 balances computational efficiency and gradient stability, ensuring better training. Finally, the number of steps per epoch determines the frequency of policy updates, facilitating iterative refinement of the agent's performance while maintaining computational resource efficiency.

Finally, the user-friendly interface created using the MATLAB guide command can easily be used to simulate the comparative analysis. The computing system used for the development of the dashboard was configured with an Intel Core i5 processor at 2.4GHz (4 CPUs), 8GB RAM, and a Windows 10 Pro 64-bit operating system to ensure efficient processing of simulation data. The dashboard itself was developed using MATLAB 2023b, leveraging its robust visualization and data processing capabilities.

4. RESULTS AND DISCUSSION

This section presents the results obtained from simulating the developed LSRL Network Simulator. When the algorithm is set to the developed LSRL agent and the user initiates the simulation, the network simulator enables visualization of the network scenario and various interference patterns as shown in Figure 3 based on aforementioned simulation setup.

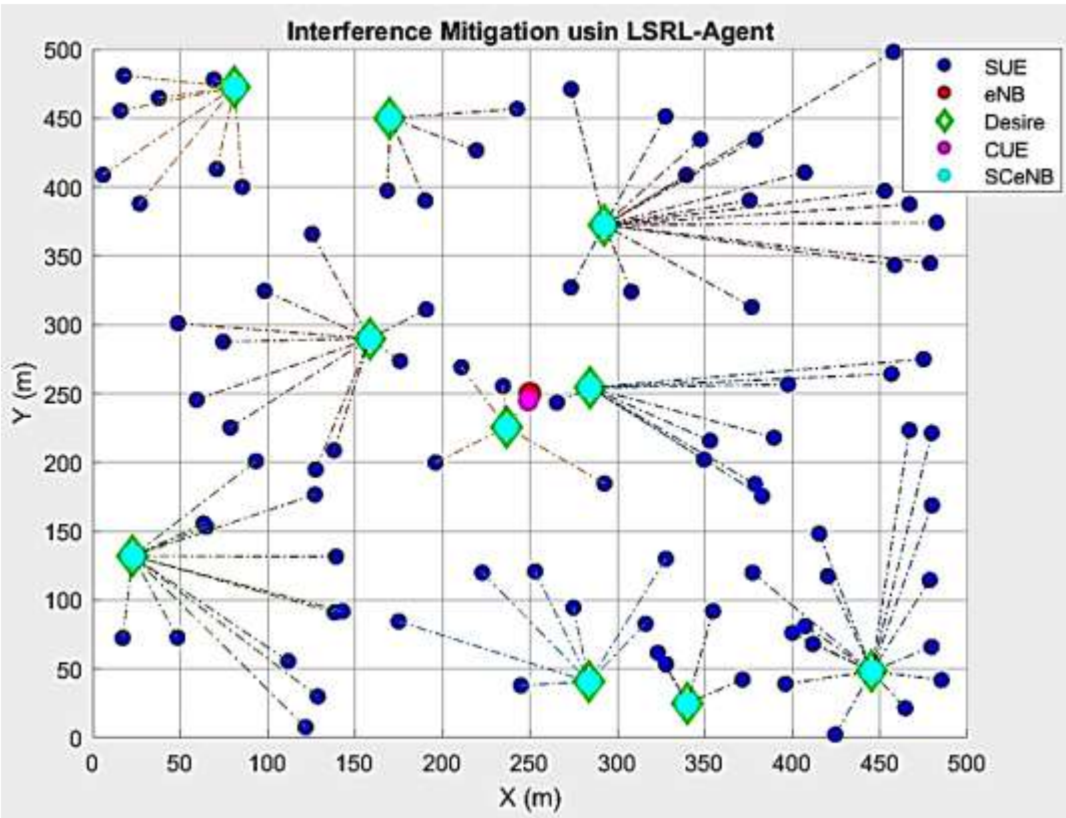
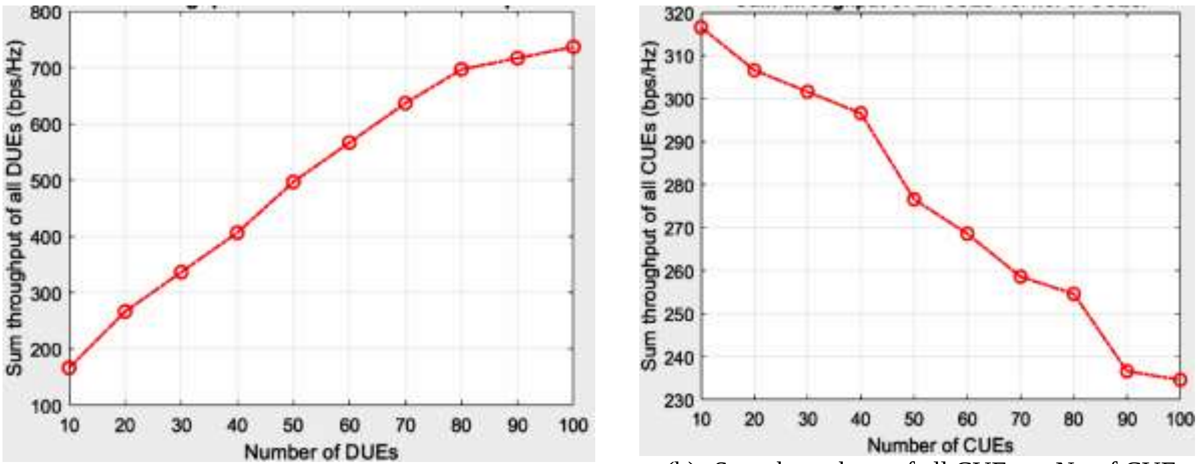


Figure 3: Network visualization

In Figure 3, SUEs are represented in blue and communicate directly with the SCeNB. These users operate within the small-cell coverage area and may also engage in D2D communication under specific conditions. CUEs, shown in magenta, do not typically participate in direct D2D communication but share network resources with both DUEs and SUEs. As conventional cellular users, CUEs communicate through either the macrocell base station (eNB) or small cells (SCeNBs). Additionally, DUEs utilize their transmitters (DUE_t) and receivers (DUE_r) to share radio resources allocated for cellular users without routing through the central base station, enabling underlay communication. In this heterogeneous network (HetNet), SUEs, CUEs, and DUEs coexist, sharing communication resources while contending with mutual interference, which is managed through the advanced resource allocation techniques developed by the LSRL agent.

The simulation results, analyzed using sum throughput, scalability, energy efficiency, and outage probability as key quality-of-service metrics, are further discussed. Additionally, the adaptability of the developed algorithm to network dynamics, particularly in response to an increasing number of user equipment or network size, is examined. To assess the performance of the developed system against existing approaches, network throughput and scalability were evaluated, with the comparative analysis results presented in Figure 4.



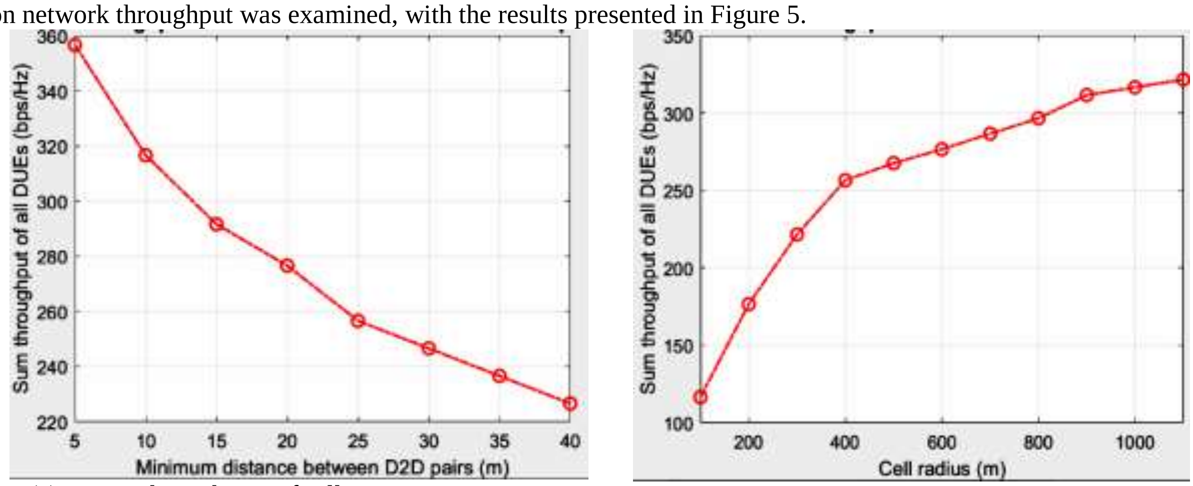
(a) Sum throughput of all DUEs vs No of D2D pairs

(b) Sum throughput of all CUEs vs No of CUEs pairs

Figure 4: Network throughput and scalability for DUEs and CUEs

Figure 4(a) illustrates the performance of the developed LSRL agent in efficiently allocating resources while mitigating interference as the number of active D2D pairs increases. Since DUEs share resources with cellular users without routing their communication through the eNB or SCeNB, an increase in D2D pairs leads to a corresponding rise in the total data rate successfully delivered by all active pairs, measured in bits per second per Hz (bps/Hz). This indicates that the LSRL scheme is scalable, maintaining high throughput even as D2D density grows.

In contrast, Figure 4(b) shows that an increase in the number of cellular users (CUEs) results in a decline in throughput. This decrease is due to heightened resource contention and increased interference from both DUEs and SUEs as the number of CUEs grows. To further investigate interference mitigation, the impact of increasing the distance between DUE pairs on network throughput was examined, with the results presented in Figure 5.



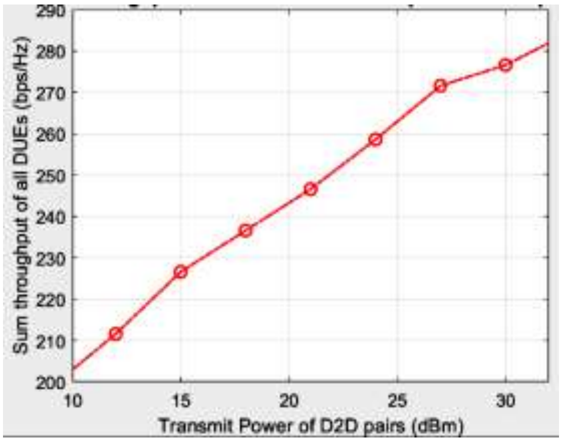
(a) Sum throughput of all DUEs vs Distance between D2D pairs

(b) Sum throughput of all Radius Vs all Radius

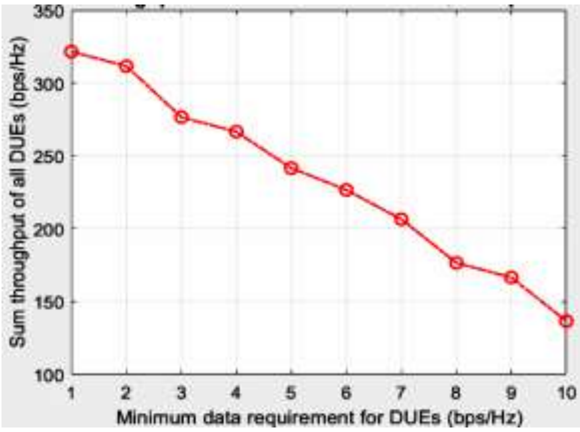
Figure 5: Distance and radius effect on throughput

Figure 5(a) shows that as the distance between DUEs increases, the sum throughput decreases. This decline is attributed to reduced channel gain and the need for higher transmission power, which in turn generates additional interference. However, Figure 5(b) demonstrates that increasing the cell radius reduces interference from the macro base station (eNB) to D2D communications. The LSRL algorithm effectively adapts to larger cell radii, mitigating inter-tier interference and enhancing overall throughput.

To assess the effectiveness of the developed scheme in optimizing power allocation for the eNB, SCeNB, and user equipment, the network's energy efficiency was evaluated by gradually increasing the transmit power of D2D pairs while maintaining baseline performance metrics. The results of this evaluation are presented in Figure 6.



(a) Sum throughput of all DUEs vs transmit power of D2D pairs

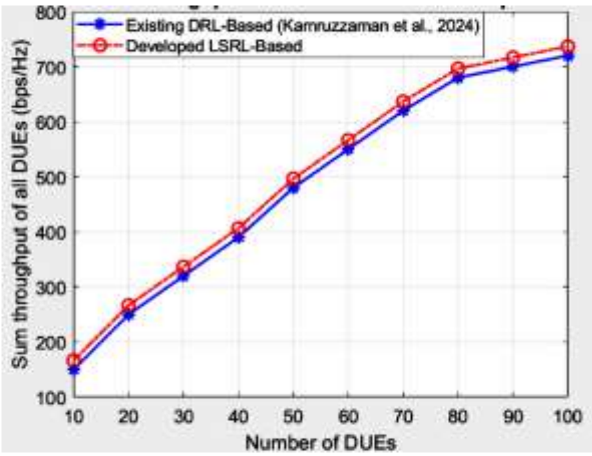


(b) Sum throughput of all DUEs Vs Minimum QoS Requirement

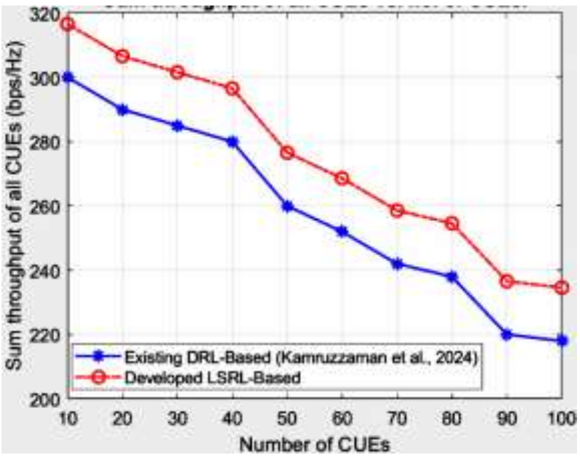
Figure 6: Transmit power and minimum QoS

Figure 6(a) illustrates that as transmission power increases, the network's sum throughput improves due to the establishment of stronger communication links. However, beyond an optimal threshold, the throughput saturates, highlighting the need to balance transmission power with QoS constraints to prevent excessive interference and energy consumption. Figure 6(b) further demonstrates that the LSRL algorithm effectively optimizes power allocation while maintaining QoS requirements, making it well-suited for energy-sensitive applications.

Validating the performance of a newly developed algorithm is crucial; therefore, the system's performance was compared with existing work. The evaluation involved selecting the appropriate radio button on the user interface (Developed LSRL-Agent & Existing DRL-Agent [30]), configuring key network parameters such as network size and user equipment, and initiating the simulation using the run simulation button. To assess how effectively resources are allocated and interference is managed, the throughput and scalability of both the proposed LSRL-based scheme and the existing DRL-based algorithm were analyzed. The results of this comparative analysis are presented in Figure 7.



(a) Sum throughput of all DUEs vs No of D2D pairs

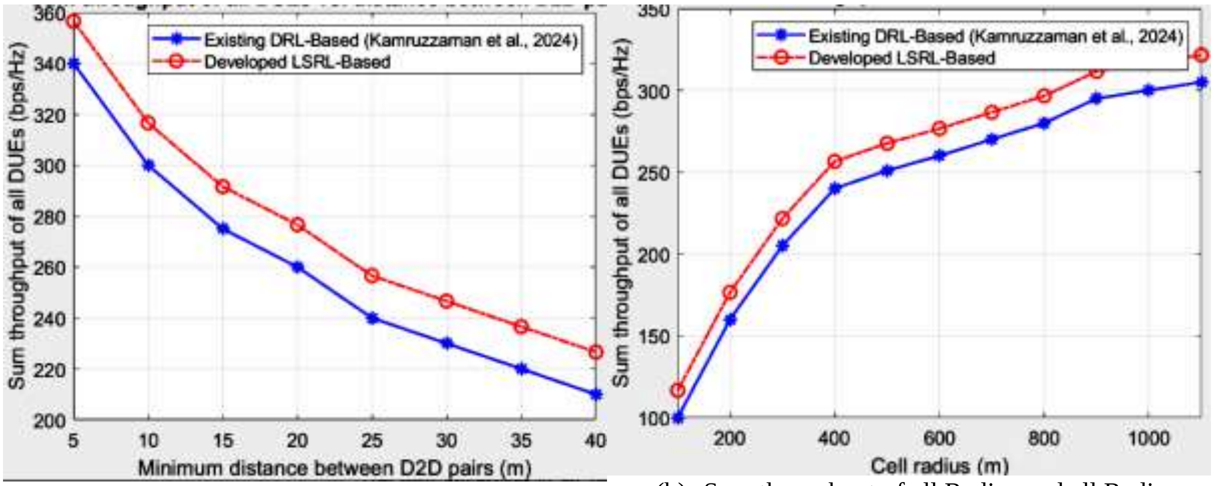


(b) Sum throughput of all CUEs vs No of CUEs pairs

Figure 7: Network throughput and scalability for DUEs and CUEs

The sum throughput of the DUEs and the number of active users in the system under different conditions and outputs were evaluated in Figure 7. It is observed that an increment in the number of DUEs equally leads to an increase in the overall system throughput. Nevertheless, the increases in DUEs also introduce more interference between DUEs and CUEs link, which decreases the sum throughput of CUEs, as seen in Figure 7(b) for both algorithms. The developed LSRL scheme mitigated the interference effectively with a higher sum throughput of DUEs compared to the approach in [30] with a 3.08% higher sum throughput than the DRL approach.

To effectively mitigate network interference, the distance between device user equipment pairs was increased, and the impact on network throughput was analyzed. The results of this investigation are presented in Figure 8.



(a) Sum throughput of all DUEs and Distance between D2D pairs

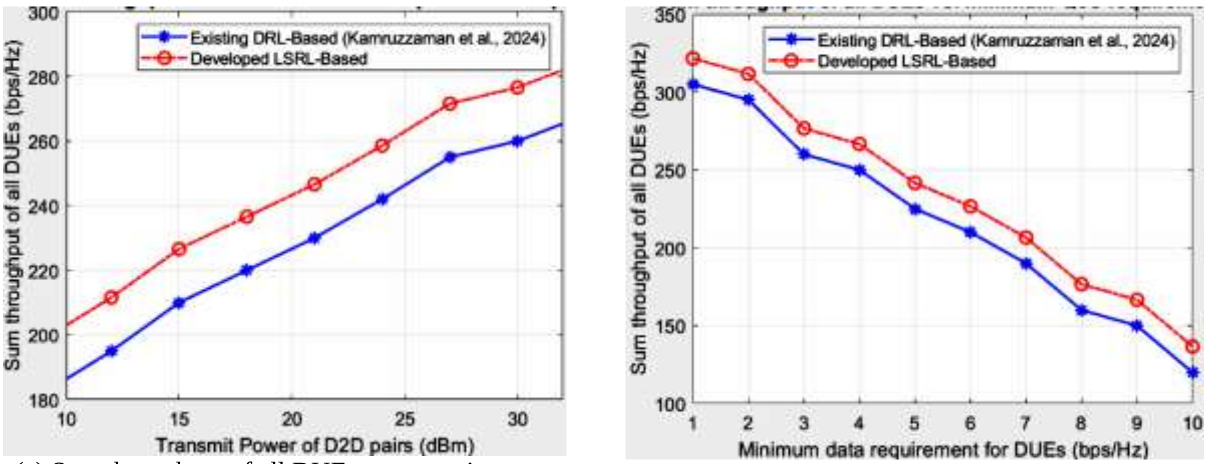
(b) Sum throughput of all Radius and all Radius

Figure 8. Distance and radius effect on throughput

In Figure 8(a), as the distance between D2D pairs increases, the sum throughput of DUEs decreases for both approaches. Specifically, when the distance expands from 10 m to 20 m, the sum throughput declines by 12.66% in the proposed LSRL scheme and by 14.7% in the DRL scheme. The lower reduction in throughput for the LSRL scheme demonstrates its superior performance over longer distances between D2D pairs. This advantage is primarily due to two factors: first, the decrease in average channel gain between the transmitter and receiver as distance increases, and second, the need for higher transmit power to maintain QoS, which leads to greater interference between DUEs and CUEs. Additionally, the LSRL scheme achieves a 6.38% improvement over the DRL scheme.

Figure 8(b) further examines the impact of the eNB/SCeNB cell radius on the sum throughput of D2D pairs. As the cell radius increases, mutual interference among users and interference from eNBs on D2D pairs are reduced, leading to a progressive improvement in sum throughput across all schemes. The proposed LSRL algorithm consistently outperforms the DRL scheme, as depicted in Figure 8. Specifically, when the cell radius expands from 200 m to 400 m, the sum throughput increases by 45.45% for the LSRL scheme and 40.2% for the DRL scheme. This demonstrates that the LSRL scheme surpasses the DRL scheme, achieving a 6.67% improvement.

To assess the effectiveness of the developed scheme in optimizing power allocation for the eNB, SCeNB, and user equipment, the network's energy efficiency was evaluated by increasing the transmit power of D2D pairs while maintaining baseline performance metrics. The simulation results are presented in Figure 9.



(a) Sum throughput of all DUEs vs transmit power of D2D pairs

(b) Sum throughput of all DUEs Vs Minimum QoS Requirement

Figure 9. Transmit power and minimum QoS

Figure 9(a) illustrates the relationship between the sum throughput of all DUEs and their transmission power. In both schemes, increasing the maximum transmission power leads to a noticeable improvement in DUE throughput. When the transmission power of D2D pairs increased from 10 dBm to 15 dBm, the sum throughput improved by 11.85% in the proposed LSRL scheme and 10.5% in the DRL scheme. The LSRL scheme demonstrated superior performance, outperforming the DRL scheme by 8.2%.

Figure 9(b) presents the sum throughput of all D2D pairs concerning the minimum QoS requirements, specifically the minimum throughput needed to establish a successful link between users. As the minimum throughput requirement increases from 2 bps/Hz to 9 bps/Hz, the sum throughput decreases significantly in both schemes. For the proposed LSRL scheme, the sum throughput drops from 321 bps/Hz to 136 bps/Hz, while for the DRL scheme, it decreases to 300 bps/Hz. This decline occurs due to stricter admission constraints imposed by higher QoS requirements. Without effective power control, mode selection, and resource allocation, the system struggles to meet these requirements, as reflected in the simulation results.

Overall, the comparative analysis shows that the LSRL scheme achieves an average improvement of 7.0% over the DRL scheme, highlighting its superior efficiency in managing transmission power while maintaining QoS constraints. This demonstrates the effectiveness of the LSRL approach in optimizing resource allocation and mitigating interference in D2D communications. One limitation of this study is the reliance on simulated data rather than real-world deployment scenarios, which may limit the generalizability of the results. Additionally, the study primarily focuses on urban environments with high user density, leaving the performance in rural or less dense networks unexamined. Furthermore, while the proposed LSRL-based approach demonstrates significant improvements in throughput and energy efficiency compared to traditional DRL methods, the computational complexity associated with training the LSTM network may limit its real-time applicability in large-scale networks. Future studies would focus on optimizing the computational aspects of the model and validating the framework using real-world data from 5G networks.

5. CONCLUSION

This study evaluated the performance of the proposed LSRL-based scheme in mitigating interference and optimizing resource allocation in a heterogeneous network, comparing it with an existing DRL-based approach. The results demonstrated that the LSRL scheme effectively maintained high throughput and energy efficiency, particularly as the number of DUEs increased and network conditions changed. The scheme exhibited superior adaptability, reducing performance degradation under varying distances between DUE pairs and increasing cell radius. Additionally, the LSRL approach achieved better power optimization while meeting minimum QoS requirements, making it suitable for energy-sensitive applications. Overall, the proposed scheme consistently outperformed the DRL approach, highlighting its potential for improving network efficiency. Future research could explore integrating adaptive learning mechanisms to further enhance real-time decision-making and extending the model to accommodate dynamic traffic loads in large-scale deployments.

REFERENCES

[1] Trabelsi, N., Fourati, L. C., & Chen, C. S. (2023). Interference Management in 5G and Beyond Networks: A Comprehensive Survey. *Computer Networks 239(6):110159.* doi: 10.1016/j.comnet.2023.110159

[2] Siddiqui, M. U. A., Qamar, F., Ahmed, F., Nguyen, Q. N., & Hassan, R. (2021). Interference management in 5G and beyond network: Requirements, challenges and future directions. *IEEE Access*, *9*, 68932-68965, doi: 10.1109/ACCESS.2021.3073543.

[3] Jiya, E. A., Ibikunle, F. A., Oluwafemi, I. B., Adedayo, A. B., Kola-Junior, A., & Babatunde, K. S. (2022). Overview of Interference Management Techniques in 5G Cellular Networks. *2022 5th Information Technology for Education and Development (ITED)*, 1-7, doi: 10.1109/ITED56637.2022.10051504.

[4] Fan, Y., Zhang, L., Li, K., Bao, M., & Lu, M. (2024, June). Deep learning-based EMI and IEMI classification in 5G-R high-speed rail wireless communications. In *2024 IEEE 99th Vehicular Technology Conference (VTC2024-Spring)* (pp. 1-6). IEEE. doi: 10.1109/VTC2024-Spring62846.2024.10683274

[5] Sejan, M. A. S., Rahman, M. H., Aziz, M. A., Tabassum, R., You, Y. H., Hwang, D. D., & Song, H. K. (2024). Interference Management for a Wireless Communication Network Using a Recurrent Neural Network Approach. *Mathematics*, *12*(11), 1755. doi: 10.3390/math12111755.

[6] Farhan, S., & Aal-nouman, M. (2024). Reinforcement Learning-based Joint Interference Mitigation and Resource Allocation in Dense Beyond 5G Heterogeneous Networks. *International Journal of Intelligent Engineering & Systems*, *17*(3)., doi: 10.22266/ijies2024.0630.46.

[7] Alruwaili, O., Logeshwaran, J., Natarajan, Y., Alrowaily, M. A., Patel, S. K., & Armghan, A. (2024). Incremental RBF-based cross-tier interference mitigation for resource-constrained dense IoT networks in 5G communication system. *Heliyon*., doi: 10.1016/j.heliyon.2024.e32849.

[8] Irkicatal, O. N., Ceran, E. T., & Yuksel, M. (2024). Deep Reinforcement Learning Enhanced Rate-Splitting Multiple Access for Interference Mitigation. *arXiv preprint arXiv:2403.05974*. Available: http://arxiv.org/abs/2403.05974

[9] Anand, D., Togou, M. A., & Muntean, G. M. (2023). A Multi-Classification Machine Learning-Based Solution to Mitigate Co-Tier Interference in 5G HetNets. In *ICC 2023-IEEE International Conference on Communications* (pp.

4840-4845). IEEE, doi: 10.1109/ICC45041.2023.10278599.

[10] Al-Jumaily, A., Sali, A., Jiménez, V. P. G., Lagunas, E., Natrah, F. M. I., Fontán, F. P., & Al-Jumeily, D. (2023). Evaluation of 5G and fixed-satellite service earth station (FSS-ES) downlink interference based on artificial neural network learning models (ANN-LMS). *Sensors*, *23*(13), 6175. doi: 10.3390/s23136175

[11] Udoh, R., Ukommi, U., & Ubom, E. (2023). Interference Mitigation In 5G Network Using Frequency Planning and Artificial Neural Network (ANN). *Journal of Multidisciplinary Engineering Science and Technology (JMEST)*, *10*(12), 16534-16540.

[12] D. Anand, G. S. Member, and M. A. Togou, "A Machine Learning Solution for Video Delivery to Mitigate Co-Tier Interference in 5G HetNets," *IEEE Transaction on Multimedia*, *25*, 5117–5129, 2023, doi: 10.1109/TMM.2022.3187607.

[13] Warrier, A., Al-Rubaye, S., Panagiotakopoulos, D., Inalhan, G., & Tsourdos, A. (2022, September). Interference mitigation for 5G-connected UAV using deep Q-learning framework. In *2022 IEEE/AIAA 41st Digital Avionics Systems Conference (DASC)* (pp. 1-8). IEEE. doi: 10.1109/DASC55683.2022.9925817.

[14] Elsayed, M., & Erol-Kantarci, M. (2019). AI-enabled future wireless networks: Challenges, opportunities, and open issues. *IEEE Vehicular Technology Magazine*, *14*(3), 70-77, doi: 10.1109/MVT.2019.2919236.

[15] Ahmad, I., & Hussain, A. (2023). Co-channel Interference Management for the Next-Generation Heterogeneous Networks using Deep Leaning. *arXiv preprint arXiv:2301.10177*. Available: http://arxiv.org/abs/2301.10177

[16] N. F. R. O. Konan, E. Mwangi, & C. Maina (2022). "Enhancement of Signal to Interference plus Noise Ratio Prediction (SINR) in 5G Networks using a Machine Learning Approach," *International Journal of Engineering Trends and Technology*, *70*(10), 319–328, doi: 10.14445/22315381/IJETT-V70I10P231.

[17] Iqbal, M. U., Ansari, E. A., Akhtar, S., Rafiq, M. N., Farooq-i-Azam, M., & Hassan, B. (2023). Hybrid Learning based Radio Resource Management in 5G Heterogeneous Networks. *Pakistan Journal of Engineering and Technology*, *6*(1), 92-97., doi: 10.51846/vol6iss1pp92-97.

[18] Iqbal, M. U., Ansari, E. A., & Akhtar, S. (2021). Interference mitigation in HetNets to improve the QoS using Q-learning. *IEEE Access*, *9*, 32405-32424., doi: 10.1109/ACCESS.2021.3060480.

[19] Yang, A., Yue, X., & Ouyang, Y. (2021). Reinforcement learning assisted beamforming for inter-cell interference mitigation in 5g massive mimo networks. *arXiv preprint*. Available: http://arxiv.org/abs/2103.11782

[20] Yadav, R., & Tripathi, A. (2022). 3D MIMO beamforming using spatial distance SVM algorithm and interference mitigation for 5G wireless communication network. *Journal of Cases on Information Technology (JCIT)*, *24*(4), 1-26., doi: 10.4018/JCIT.296717.

[21] Osman, R. A., Saleh, S. N., & Saleh, Y. N. (2021). A novel interference avoidance based on a distributed deep learning model for 5G-enabled IoT. *Sensors*, *21*(19), 6555., doi: 10.3390/s21196555.

[22] Afolabi, A. S., Ahmed, S., & Akinola, O. A. (2021). A Reinforcement Learning Approach for Interference Management in Heterogeneous Wireless Networks. *International Journal of Interactive Mobile Technologies*, *15*(12), doi: 10.3991/ijim.v15i12.20751.

[23] Guo, Y., Hu, C., Peng, T., Wang, H., & Guo, X. (2020, June). Regression-based uplink interference identification and SINR prediction for 5G ultra-dense network. In *ICC 2020-2020 IEEE International Conference on Communications (ICC)* (1-6). IEEE., doi: 10.1109/ICC40277.2020.9149375.

[24] Wu, D., Nekovee, M., & Wang, Y. (2020). Deep learning-based autoencoder for m-user wireless interference channel physical layer design. *IEEE Access*, *8*, 174679-174691. doi: 10.1109/ACCESS.2020.3025597.

[25] Wang, Y., Feng, G., Sun, Y., Qin, S., & Liang, Y. C. (2020). Decentralized learning based indoor interference mitigation for 5G-and-beyond systems. *IEEE Transactions on Vehicular Technology*, *69*(10), 12124-12135. doi: 10.1109/TVT.2020.3012311.

[26] Elsayed, M., Erol-Kantarci, M., & Yanikomeroglu, H. (2020). Transfer reinforcement learning for 5G new radio mmWave networks. *IEEE Transactions on Wireless Communications*, *20*(5), 2838-2849. doi: 10.1109/TWC.2020.3044597.

[27] Barazideh, R., Semiari, O., Niknam, S., & Natarajan, B. (2020). Reinforcement learning for mitigating intermittent interference in terahertz communication networks. In *2020 IEEE International Conference on Communications Workshops (ICC Workshops)* (1-6). IEEE. doi: 10.1109/ICCWorkshops49005.2020.9145080.

[28] Mismar, F. B., Evans, B. L., & Alkhateeb, A. (2019). Deep reinforcement learning for 5G networks: Joint beamforming, power control, and interference coordination. *IEEE Transactions on Communications*, *68*(3), 1581-1592., doi: 10.1109/TCOMM.2019.2961332.

[29] Mughees, A., Tahir, M., Sheikh, M. A., Amphawan, A., Meng, Y. K., Ahad, A., & Chamran, K. (2023). Energy-efficient joint resource allocation in 5G HetNet using Multi-Agent Parameterized Deep Reinforcement learning. *Physical Communication*, *61*, 102206. doi: 10.1016/j.phycom.2023.102206

[30] Kamruzzaman, M., Sarkar, N. I., & Gutierrez, J. (2024). Machine Learning-Based Resource Allocation Algorithm to Mitigate Interference in D2D-Enabled Cellular Networks. *Future Internet*, *16*(11), 408. doi: 10.3390/fi16110408.