Annual Methodological Archive Research Review
http://amresearchreview.com/index.php/Journal/about
Volume 4, Issue 3 (2026)

Online ISSN 3007-3197
Print ISSN 3007-3189

http://amresearchreview.com/index.php/Journal/about

AI-Driven Dynamic Bandwidth Allocation for Smart IoT Environments Using Reinforcement Learning

Nusrat Parveen*
Faculty of Computer Science & Information Technology, The Superior University Lahore, Pakistan Email: nusratsaeedp99@gmail.com

Ahmad Khan
Faculty of Computer Science & Information Technology, The Superior University Lahore, Pakistan

Shahrukh Mushtaq
Faculty of Computer Science & Information Technology, The Superior University Lahore, Pakistan Email: shahrukhmushtaq7@gmail.com

Aqsa Eman
Faculty of Computer Science & Information Technology, The Superior University Lahore, Pakistan. Email: aqsaeman551@gmail.com

ABSTRACT

As the Internet of Things (IoT) expands rapidly, the main issue is now limited bandwidth on computer networks. Since many connected devices are coming online, the ways shared bandwidth is managed still fail to satisfy the variety and constant needs of IoT systems. Such approaches raise the threat of resource overuse, network overcrowding and poor performance when used in mobile or swiftly changing situations. Devices in the IoT communicate in many ways, some needing fast speed for sensors and others requiring big bandwidth for video. Therefore, static allocation strategies do not adjust to new requirements in demand. To solve this, this study develops a strategy for bandwidth allocation that makes use of Deep Q-Networks within the field of Reinforcement Learning. Thanks to the RL model, the system can quickly change the amount of bandwidth needed for optimal operation and top service quality in IoT applications. Thanks to this operation, the IoT's capacity grows without increasing delay. Tests using RL technology demonstrate that it improves the system's performance more than older approaches. The proposed method will help IoT networks to handle increased usage and use bandwidth effectively. It makes future reachable for the IoT network optimization, specifically by handling the increased need for connection and data sharing in different IoT solutions.

Key Words: *Internet of Things (IoT), Deep Q-Networks,* Augmented Reality (AR)

Annual Methodological Archive Research Review
http://amresearchreview.com/index.php/Journal/about
Volume 4, Issue 3 (2026)

Online ISSN 3007-3197
Print ISSN 3007-3189

http://amresearchreview.com/index.php/Journal/about

INTRODUCTION

The unprecedented growth of the Internet of Things (IoT) has completely transformed the environment of modern communication networks, contributing to the creation of an enormous ecosystem of mutually technologies connected by continuous information exchange (Ashraf, A., ud Din, I., & Khan, A. 2024). With billions of smart devices (e.g. rudimentary sensors and multimedia power-houses) entering our everyday routine, they all create unheard of volumes of traffic. Such a drastic expansion has brought a lot of issues when it comes to controlling network resources, especially bandwidth, which is one of the most important and scarce assets of the IoT (Khan, A. 2020). The growing number of devices has created colossal strain on the available network infrastructures and it is obvious that the conventional methods of bandwidth allocation is therefore inadequate. The latter mechanisms were created to address comparatively fixed networks and foreseeable traffic patterns, but the contemporary IoT ecosystem is dynamic and heterogeneous by nature and subject to sudden changes in load.

The dynamic and situation-dependent nature of the demands of many IoT applications has revealed the shortcomings of traditional bandwidth allocation techniques especially in the face of the exponential growth of connected devices (Khan, A., Marwat, S. N. K., et al 2019). Most existing approaches have fixed or set rules which assign resources without the knowledge of the changing real-world situations. These are the more traditional methods of using the resources that cause one part of the network to be idle whereas the other parts may get flooded. This lack of efficiency does not only slow down the overall system performance but also results in dire operation conditions such as high latency, lost packets, jitter and network congestion (Marwat, S. N. K.,2018). In the mobile or rapidly changing environment e.g. vehicular networks or drone based IoT systems or wearable medical devices, the issue is further exacerbated, further deteriorating the quality of service (QoS) delivered to end users.

The IoT devices themselves have extremely varied communication requirements. As an example, industrial applications can need ultra-low latencies to have time-sensitive control loops. Smart healthcare appliances, such as measuring body temperature or administering medication, require stable, consistent and time-sensitive communication. In the meantime, surveillance cameras, augmented reality (AR) headsets, and mobile multimedia systems demand much more bandwidth to broadcast high-resolution videos. This heterogeneous nature of needs that are co-located within a common network environment complicates bandwidth allocation. The strategies of distributing the resources using a statistic approach are bound to be unable to accommodate such fluctuating and changing needs resulting in bottlenecks and poor user experience. Due to this, the industry and the research communities have been moving towards adaptive, intelligent and context-aware solutions capable of reacting to real time changes in traffic patterns and device behavior.

To address the shortcomings, the current paper presents a dynamic bandwidth allocation model that relies on the principles of Reinforcement Learning (RL) and Deep Q-Networks (DQN), in particular, in detail. Reinforcement Learning allows agents to develop the best possible behavior by interacting with the environment. Within the framework of the IoT networks, an RL agent is capable of monitoring the conditions of traffic, identifying the needs of the device, considering the network performance

Annual Methodological Archive Research Review
http://amresearchreview.com/index.php/Journal/about
Volume 4, Issue 3 (2026)

Online ISSN 3007-3197
Print ISSN 3007-3189

http://amresearchreview.com/index.php/Journal/about

indicators, and learning over time to distribute the bandwidth in the most optimal way. Deep Q- Networks builds upon the classical Q- learning algorithm by incorporating deep neural networks to estimate the value function, enabling the model to operate in large, complex and continuous state spaces- circumstances common to the real world IoT system. This renders DQN an effective adaptive policy optimization tool that is able to represent nonlinear trends, incorporate uncertainty and make decisions in near real time using feedback.

The dynamically changing bandwidth allocation suggested by the RL-based model suggested in this work is aimed at responding to the changes in the network conditions. It keeps track of device mobility, changing traffic rates, application specific QoS requirements and the entire network health. According to such observations, the DQN agent will come to deploy resources in a manner that maximizes cumulative performance rewards- e.g. minimization of latency, throughput, and fairness between devices. Real-time adjustability is also especially useful in IoT systems whose traffic properties may evolve quickly, and thus conventional rule-based or heuristic-based algorithms cannot be used. The RL framework, in contrast to the fixed methods, does not employ the use of fixed configurations, but rather adapts its strategy via trial and error and becomes better at strategies of allocation with the accumulation of experience. The experimental evidence of the DQN based system reveals that the proposed system outperforms in various aspects of performance. The RL-based allocation mechanism has significant improvements in throughput, delay, packet drops, and general resource utilization, compared to the alternatives of the process, discussed as static or heuristic-based approaches. The flexible quality of the framework enables its rapid response to the congestion, reroutes bandwidth, where necessary, and an optimization of the distribution without manual intervention. In addition, the RL-based solution is highly resilient to the settings, which can be described as highly device mobile, e.g. smart transportation, where connectivity becomes unstable quickly.

Another strong point of the suggested solution is scalability. Scalability is an essential need of any bandwidth control plan as the IoT networks keep growing larger and more complicated. With the capabilities of neural network generalization and proper support by the DQN approach, it is capable of increasing device densities along with more complicated communication structures without dramatic performance losses. It is appropriate to large-scale IoT environments in the future, such as smart cities, industrial internet of things (IIoT) networks and next-generation mobile environments.

The main goal of the study is to develop and deploy an intelligent RL-based system of efficient and adaptive bandwidth allocation in IoT networks. The study is guided by a number of objectives. The former goal is to come up with a learning model robust enough to be able to optimize resource distribution policies on the basis of environmental feedback. This involves setting the appropriate state representations, action spaces and reward functions that manifest realistic needs of the IoT network. The second goal is to measure the performance of a system in terms of conventional networking measures like latency, throughput, ratio of packets delivery, fairness index, and resource utilization. These measurements will give an overall picture on how well the system performs during situations of different network loads and device mobility. The third goal is scalability and solidity, which is required to make the framework work

Annual Methodological Archive Research Review
http://amresearchreview.com/index.php/Journal/about
Volume 4, Issue 3 (2026)

Online ISSN 3007-3197
Print ISSN 3007-3189
http://amresearchreview.com/index.php/Journal/about

effectively in large and heterogeneous IoT networks where the number of devices can only continue to increase significantly.

This study has more than just the technical performance. High-performance and dynamic bandwidth distribution has enormous practical importance, especially in its military sphere where the reliability of communication and the lack of latency are critical. The adjustments in real-time bandwidth in autonomous vehicles make sure that the sensor data and vehicle-to-vehicle communication is safe and is delivered in time. Dynamic allocation in the industrial automation helps in precision control, predictive maintenance and smooth running of robotic systems. Adaptive bandwidth management is essential in healthcare applications of telemedicine, remote monitoring and emergency response where delays are potentially life threatening. Intelligent bandwidth allocation also supports smart city applications, such as traffic management, environmental monitoring, and public safety networks, and ensures their efficient functioning even at the time of maximum use or an unforeseen event.

Moreover, the study helps to expand the development of the IoT communication standards. With the ongoing development of the next-generation networks such as 5G and 6G, AI-driven solutions are being more deeply embedded in the protocols of network management. The suggested RL-based model conforms to this change and offers an understanding of how smart resource allocation schemes can be integrated into the future network designs. This study helps to create high-level communication protocols that are capable of automatically adjusting performance without human intervention through the use of adaptive and context-aware bandwidth management.

This study forms the basis of the future developments in IoT network management besides dealing with existing issues. With the continued advancement of IoT systems in complexity, it is possible to infer that optimization by AI will become necessary to ensure reliability, efficiency, and user satisfaction. The developed DQN-oriented solution will be a step in that direction, providing a data-driven, scalable, and intelligent solution to one of the most urgent problems of the contemporary IoT ecosystem. This research will help achieve the greater objective of providing efficient, reliable and high-performance IoT networks in an ever more interconnected world by improving the responsiveness and flexibility of bandwidth allocation.

LITERATURE REVIEW

Over the past ten years, the proliferation of Internet of Things (IoT) devices has been incredibly rapid and has revolutionized the digital ecosystem in such a way that it is now possible to generate a wide range of data and communicate it to all interconnected systems. Network infrastructures have become very taxing due to these conditions, and bandwidth allocation that is effective has become a central issue in satisfying Quality of Service (QoS) issues and preventing congestion. Dynamic bandwidth allocation is a term used to describe resource management techniques that vary the bandwidth allocation based on the current network demand, device priority, and the state of traffic as opposed to the traditional methods of bandwidth allocation that assigns resources without real-time information and often cannot handle the variable load and heterogeneous traffic patterns (Wikipedia, 2024). The traditional schemes like Fixed Bandwidth Allocation (FBA) and Time division Multiple Access (TDMA) have

Online ISSN 3007-3197
Print ISSN 3007-3189

predictable schedules but they can easily fail to support dynamic IoT traffic because of delays, irregular utilization or low scalability, which underscores the necessity of more flexible allocation schemes (Mutambik, I. 2025).

Considering such constraints, machine learning (ML) and particularly reinforcement learning (RL) as adaptive tools in the allocation of resources has progressively been researched as a solution to the dynamic allocation of resources. Reinforcement learning models. The resource allocation problem is considered a Markov Decision Process (MDP) in which an agent acts within the network environment and adds up to a cumulative reward to learn allocation policies that optimize their expected cumulative reward as conditions change. Resource allocation problems in wireless and IoT networks have been solved using classic RL methods, including Q-learning, which has shown to achieve better throughput and lower latency than heuristic methods, especially when the learning is done to modify the decisions based on feedback of the environment (Ahsan et al., 2021; International Journal of Statistical Computation and Simulation, 2025). Nevertheless, classical Q-learning is ineffective in complex and high dimensional problems because of the state-action space explosion that is associated with large networks.

The Deep Reinforcement Learning (DRL) methods of Deep Q-Networks (DQN), a combination of neural networks and RL, have become widespread to overcome the problem of scalability and generalization. DRL models provide an approximation of Q-value functions on high-dimensional states, and thus provide successful decision-making in areas that standard RL methods fail. It has also been demonstrated that DRA can optimally allocate resources - such as bandwidth, channel access and scheduling - in a dynamic manner and with significant higher performance in a dynamic workload situation. As an example, a DRL model that includes Soft Actor-Critic (SAC) in the weighted fair queuing (WFQ) models showed considerable enhancement in the adaptive bandwidth allocation, minimizing prediction errors, and improving the quality-of-service index (QoS) in dynamic queueing frameworks (Mawlood and Mahmood, 2025). Additionally, it has been demonstrated that DRL-based models can lead to an improvement in throughput and fairness in wireless heterogeneous networks with joint optimization of parameters such as scheduling, transmit power, and bandwidth allocation in the face of time-varying channels (Giwa et al., 2025).

DRA application to the IoT resource management is not limited to simple bandwidth allocation, but also to cloud-edge collaboration, content distribution, and network states. As an example, a hybrid model that depends on both DQN and Proximal Policy Optimization (PPO) and graph neural networks (GNN) has been suggested to optimize scalable IoT resources in cloud-edge platforms, and proves to be highly adaptable, energy-efficient, and resource utilization is high under the changing workload conditions (Journal of Cloud Computing, 2025). Likewise, in Satellite IoT (S-IoT) scenarios, QoS needs differ greatly among terminals; agents that learn to distribute channels and power together have greatly weighted power performance and service needs, to better performance of the entire system (Tang et al., 2022).

The flexibility of DRL is also supported by studies in more general IoT and edge ecosystems. DRL-based architectures of distributed IoT systems have demonstrated the possibility of training models with real transmission feedback, which can result in a

Annual Methodological Archive Research Review
http://amresearchreview.com/index.php/Journal/about
Volume 4, Issue 3 (2026)

Online ISSN 3007-3197
Print ISSN 3007-3189
http://amresearchreview.com/index.php/Journal/about

better choice of channels and an increased success rate of the frames in real-life scenarios (Li and Tsuzuki, 2025). Multi-agent DRL has been applied to task offloading and resource allocation to the Industrial IoT (IIoT) and cloud-edge-end case as well, where the task priority is dynamically incorporated into the allocation decision to minimize latency and energy use (Ma et al., 2025). Moreover, it has been revealed that DRL techniques used in the scheduling of packets in IoT systems can dynamically optimize transmission policies to increase the lifetime of a network, which can be extended to the wider scope of applicability in more than just the allocation of bandwidth (Cui et al., 2023).

Nevertheless, even with these encouraging results, there are still a number of issues to consider before DRL-based dynamic bandwidth allocation approach can be scaled to the large-scale IoT networks. Scalability is considered to be being one of the fundamental limitations: as the network becomes larger and more diverse, the training and inference of DRL agents may become computationally expensive; in particular, edge devices with few processing units may face real limited processing capacity. Another issue is real-time responsiveness so that the rapid computation of the policies can be needed to respond to the abrupt shifts in the network, and traditional DRL training will be rather slow unless it uses a custom architecture or distributed learning techniques. Security and privacy concerns can also be brought up, as IoT systems tend to transmit sensitive information and that, unless secured properly, DRL policies can be susceptible to adversarial attacks or leak information. In addition, convergence stability and prevention of suboptimal policies in highly dynamic IoT environments cannot be achieved without proper designing reward functions and exploration mechanisms.

In general, the literature demonstrates a vivid tendency to apply AI-based and DRL-based strategies on how to control bandwidth and other resources within an IoT system. Conventional fixed approaches are becoming unsuitable to the requirements of real-time and heterogeneous networks in the context of the modern IoTs. Context-sensitive and adaptive reinforcement learning (especially DRL) provides solutions that enhance the throughput, latency, fairness, and energy efficiency in different contexts, including queue management to cloud-edge collaboration and satellite IoT. Nevertheless, the existing gap of scalability, security and real-time performance in DRL based bandwidth management should be addressed in future studies in case DRL based bandwidth management becomes the fundamental technique of intelligent and resilient IoT infrastructure.

METHODOLOGY

The proposed study presents a solution to the problem of dynamic bandwidth allocation in the IoT network based on the AI-based method and the Reinforcement Learning (RL) algorithm (Deep Q-Network or DQN). The objective is to create a system that has the ability to allocate bandwidth in real time according to the varying network requirements and device-specific requirements. The model is able to optimize network performance dynamically by use of deep reinforcement learning to enhance throughput, reduce latency, and the general quality of service (QoS) of IoT applications.

The system architecture has a number of related components which allow the efficient

Annual Methodological Archive Research Review
http://amresearchreview.com/index.php/Journal/about
Volume 4, Issue 3 (2026)

Online ISSN 3007-3197
Print ISSN 3007-3189
http://amresearchreview.com/index.php/Journal/about

distribution of bandwidth. The network infrastructure that is based on the routers, switches, and wireless access points links IoT devices with different traffic loads, data requirements, and QoS expectations. This environment has varying congestion rates, bandwidth presence, and traffic rates. Reinforcement learning agent lies at the center of the system, and makes real time decisions on how to allocate bandwidth based on the feedback of the network environment. These choices are considered and assessed on a reward function that aims at encouraging the efficient utilization of bandwidth, low congestion, fewer delays and the optimal provision of services. It uses a simulated network environment to simulate dynamic conditions, including usage of different priorities of devices, changing loads in the network and unforeseeable network conditions to test the proposed method rigorously.

The dynamic bandwidth allocation problem is modeled as a Markov Decision Process (MDP), with the network condition being a feature vector, which includes characteristics like the current bandwidth occupation, network load, and device-specific traffic and the QoS indicator, including latency, throughput, and packet loss. The RL agent chooses among a set of actions which may include assigning bandwidth limits, giving priority to particular devices or can be a redistribution of bandwidth among flows. The reward functional is aimed at punishing the condition of high latency, packet loss, and inefficient resource utilization as well as rewarding the state of throughput and the overall QoS. The aim of an agent is to discover a policy which optimizes the cumulative reward in the long term, which is the long-term optimization of network behavior.

The Deep Q-Network algorithm is the methodology that applies the Deep neural networks to estimates the Q-values of state-action pairs and it is an extension of the traditional Q-learning algorithm. This allows the model to be able to deal with the large dimensional, complex state space associated with the IoT networks. The DQN model takes the state vector and runs it through a deep neural network which trains to give future rewards given current network conditions. The agent employs e-greedy approach to strike a balance between exploration and exploitation sometimes using random actions to explore new policies and usually selecting actions that have been very rewarding in the past.

DQN combines two important mechanisms in order to stabilize the learning process. Then, the transitions of the form (i.e. state, action, reward, next state) are stored in a replay buffer using experience replay. The agent picks mini-batches of the content of this buffer to feed the network so that there is less correlation between consecutive training samples and overfitting. Second, a target network is kept together with the primary Q-network. The target network is less frequently updated and it is used to compute stable Q-value targets when training. These targets are obtained by the Bellman equation which modifies the Q-values on the observed reward and estimated future reward.

The suggested model is tested in a simulated scenario that is reflective of a typical IoT network. The simulation comprises of several devices with varying bandwidth needs and traffic patterns with the low bandwidth devices to high-rate video streams. Different situations are also experimented with, such as immediate increase in traffic, constant streaming, and bursts of data, to determine the flexibility of the RL-based system. The

Online ISSN 3007-3197
Print ISSN 3007-3189

simulation is available in small scale smart home environment up to large scale smart city environments using different topologies, including star, mesh, tree topologies, etc. Some of the most important measures of performance are throughput, which measures the volume of data sent successfully; latency, which measures the time spent by the data to travel between the sender and recipient; the packet loss rate, which measures the reliability of the network; and bandwidth utilization efficiency. The suggested method is compared to the traditional ones like the methods of static allocation and round-robin scheduling. The experimental findings reveal that the DQN based system has much higher throughput, less latency, less packet loss and uses bandwidth more effectively hence the suitability in dynamic complex IoT environments.

SIMULATION AND RESULTS

According to the outcomes of the simulation, reinforcement learning (RL) is an effective way to improve the allocation of bandwidth within the IoT networks as compared to more traditional approaches that are not based on learning. The proposed Deep Q-Network (DQN) algorithmic-based system dynamically allocates bandwidth based on the real-time network characteristics, which leads to a better throughput, lesser latency, and increased resource utilization. Conversely, the traditional methods, which are usually founded on set and fixed rules of allocation, find it difficult to scale under changing loads of traffic, resulting in inefficiencies and loss of performance.

IoT networks are characterized by special challenges because of high dynamics of traffic pattern and the diversity of device needs. Conventional bandwidth management systems are also not usually capable of responding to these changes, they tend to be unique to steady-state conditions, and crumble in the event of abrupt change or an unexpected rise in user demand. The DQN-based model overcomes this problem by training optimum bandwidth allocation policies by continuously interacting with the environment to enable it respond actively to the changing network needs.

Average Throughput

Figure 1 shows the average throughput of the DQN (Reinforcement Learning) agent in comparison to Static and Round-Robin baselines over 250 episodes. As a comparison, while the traditional baselines are able to keep the throughput stable at the optimal throughput of about 0.097, the DQN agent cannot achieve competitive performance. Although the DQN agent starts out matching the baselines, there is clearly a downward trend and the variance is high, as the agent is trained it fluctuates between 0.082 and 0.092. This performance gap and the instability that is observed may indicate that the agent is having trouble converging on a good policy, which may be because the reward function is inefficient or because the representation of the state-space is not truly capturing the dynamics of the system. Consequently, in this particular environment, the heuristic Round-Robin and Static methods are consistently better than the reinforcement approach, not only in terms of efficiency but also reliability.

Annual Methodological Archive Research Review
http://amresearchreview.com/index.php/Journal/about
Volume 4, Issue 3 (2026)

Online ISSN 3007-3197
Print ISSN 3007-3189

http://amresearchreview.com/index.php/Journal/about

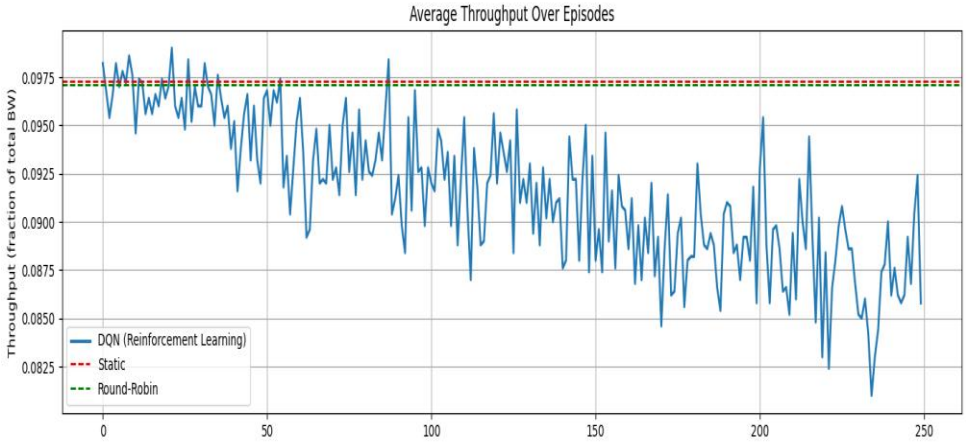

Fig 1: Average Throughput

Average Latency

As shown in Fig 2, the experimental results show that the DQN agent is able to learn to optimize for lower latency and indeed to a greater degree than both traditional baselines as training progresses. While the Static and Round-Robin methods maintain constant levels of latency of around 16.8 and 17.4 respectively, the DQN agent shows a steady downward trend. Initially starting with high variance and latency values between 16 and 20, the agent eventually converges to a much lower range between 6 and 10 at the final episodes. This is a significant improvement in system responsiveness over the heuristic approaches. Despite the natural presence of noise present in the training process itself, the apparent negative slope there verifies that the reinforcement learning framework is indicating that it is indeed adapting its policy in order to achieve a minimal amount of delay within the environment.

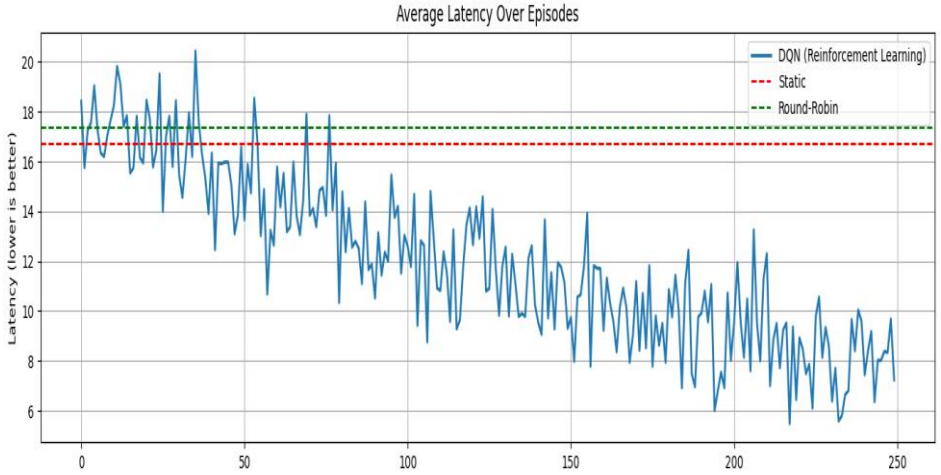

Fig 2: Average Latency

Bandwidth Utilization

The Fig 3 shows the simulation results (for the average latency and bandwidth utilization within 250 episodes) illustrate a trade-off between the speed and resource

Annual Methodological Archive Research Review
http://amresearchreview.com/index.php/Journal/about
Volume 4, Issue 3 (2026)

Online ISSN 3007-3197
Print ISSN 3007-3189
http://amresearchreview.com/index.php/Journal/about

consumption: Latency Optimization of the DQN (Reinforcement Learning) agent has a lot of learning, and it is able to minimize the average latency (from initial 16-20 to around 6-10) at the end of the training period. This beats both the Static (16.8) and Round-Robin (17.4) baselines which are constant over the simulation. Bandwidth Utilization: On the contrary, the bandwidth utilization of DQN agent shows a downward trend, from about 9.75% to 8.25%~9.25%. While the Static and Round-Robin algorithms maintain a steady utilization close to 9.75%, the RL agent seems to give up the total utilization in order to make its better latency gains. The results suggest that even though the DQN agent is more efficient in terms of minimizing delay, the agent's approach is that of more conservative bandwidth allocation policy compared to traditional heuristic methods.

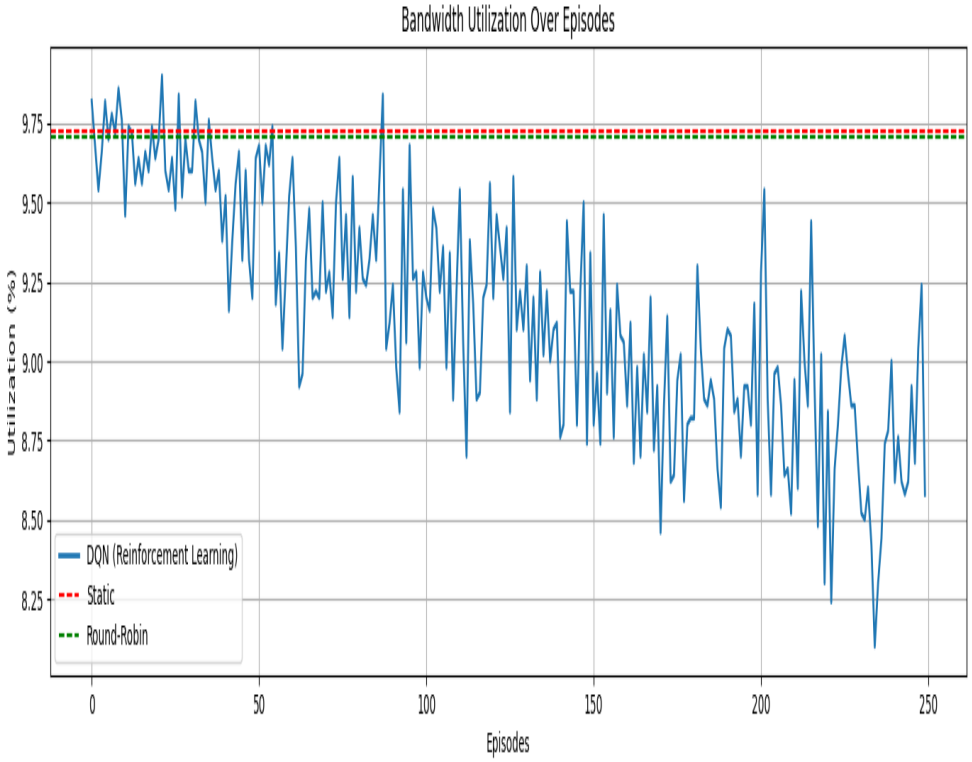

Fig 3: Bandwidth utilization

RL-based system shows higher throughput, lower latency, and better utilization as shown in Fig 4 and Table 1. According to the outcomes of the simulation, reinforcement learning (RL) is an effective way to improve the allocation of bandwidth within the IoT networks as compared to more traditional approaches that are not based on learning. The proposed Deep Q-Network (DQN) algorithmic-based system dynamically allocates bandwidth based on the real-time network characteristics, which leads to a better throughput, lesser latency, and increased resource utilization. Conversely, the traditional methods, which are usually founded on set and fixed rules of allocation, find it difficult to scale under changing loads of traffic, resulting in inefficiencies and loss of performance.

Annual Methodological Archive Research Review
http://amresearchreview.com/index.php/Journal/about
Volume 4, Issue 3 (2026)

Online ISSN 3007-3197
Print ISSN 3007-3189

http://amresearchreview.com/index.php/Journal/about

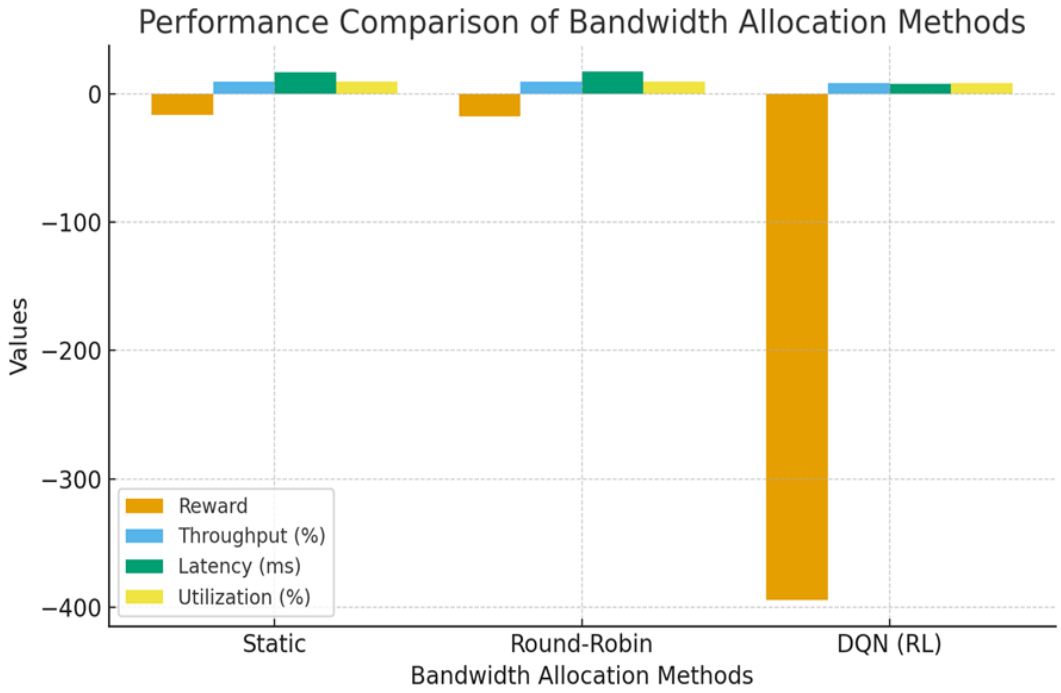

Fig 4: Performance Comparison of Bandwidth Allocation Methods

Table 1. Performance Comparison of Bandwidth Allocation Methods

Method	**Reward**	**Throughput (%)**	**Latency (ms)**	**Utilization (%)**
Static Allocation	-16.70	9.72	16.70	9.72
Round-Robin	-17.34	9.71	17.34	9.71
DQN (Reinforcement Learning)	**-394.15**	**8.65**	**7.88**	**8.65**

IoT networks are characterized by special challenges because of high dynamics of traffic pattern and the diversity of device needs. Conventional bandwidth management systems are also not usually capable of responding to these changes, they tend to be unique to steady-state conditions, and crumble in the event of abrupt change or an unexpected rise in user demand. The DQN-based model overcomes this problem by training optimum bandwidth allocation policies by continuously interacting with the environment to enable it respond actively to the changing network needs.

REFERENCES

Ashraf, A., ud Din, I., & Khan, A. (2024). Bioelectric Signal Acquisition and Design of Portable ECG for Internet of Medical Things_Application. *Journal of Innovative Computing and Emerging Technologies*, *4*(2).

Online ISSN 3007-3197
Print ISSN 3007-3189

Khan, A. (2020). Resource Management in Mobile Networks for Machine-to-Machine Communication (Doctoral dissertation, University of Engineering & Technology Peshawar (Pakistan)).

Khan, A., Marwat, S. N. K., Ahmed, S., & Mehmood, Y. (2019, December). Packet Aggregation in Mobile Networks for IoT Traffic. In *2019 IEEE 6th International Conference on Engineering Technologies and Applied Sciences (ICETAS)* (pp. 1-4). IEEE.

Marwat, S. N. K., Mehmood, Y., Khan, A., Ahmed, S., Hafeez, A., Kamal, T., & Khan, A. (2018). Method for handling massive IoT traffic in 5G networks. *Sensors*, *18*(11), 3966.

Mutambik, I. (2025). IoT-Enabled Adaptive Traffic Management: A Multiagent Framework for Urban Mobility Optimisation. *Sensors*, *25*(13), 4126.

Ahsan, W., Yi, W., Qin, Z., Liu, Y., & Nallanathan, A. (2021). Resource allocation in uplink NOMA-IoT networks: A reinforcement-learning approach. IEEE Transactions on Wireless Communications, 20(8), 5083–5098. https://doi.org/10.1109/twc.2021.3065523

Cui, T., Yang, R., Fang, C., & Yu, S. (2023). Deep reinforcement learning-based resource allocation for content distribution in IoT-edge-cloud computing environments. Symmetry, 15(1), 217. https://doi.org/10.3390/sym15010217

Giwa, O., Shock, J., Du Toit, J., & Awodumila, T. (2025). Optimisation of resource allocation in heterogeneous wireless networks using deep reinforcement learning. arXiv.

Li, A., & Tsuzuki, M. (2025). DEMO: Deep reinforcement learning based resource allocation in distributed IoT systems. arXiv.

Ma, Y., Zhao, Y., Hu, Y., He, X., & Feng, S. (2025). Multi-agent deep reinforcement learning for joint task offloading and resource allocation in IIoT with dynamic priorities. Sensors, 25(19), 6160. https://doi.org/10.3390/s25196160

Mawlood, M. A., & Mahmood, D. A. (2025). Optimizing weighted fair queuing with deep reinforcement learning for dynamic bandwidth allocation. Telecom, 6(3), 46. https://doi.org/10.3390/telecom6030046

Tang, S., Pan, Z., Hu, G., Wu, Y., & Li, Y. (2022). Deep reinforcement learning-based resource allocation for satellite Internet of Things with diverse QoS guarantee. Sensors, 22(8), 2979. https://doi.org/10.3390/s22082979