

# Resource Allocation using Reinforcement Learning

Priyanka M E<sup>1</sup>, K Kruthi<sup>2</sup>, Krishna S bhat<sup>3</sup>, Rashmitha N<sup>4</sup>, Tejaswini Sharanappa Patil<sup>5</sup>

<sup>1</sup>Asst.Prof, Department of CSE, Bahubali College of Engineering, Shravanabelagola-573135

<sup>2,3,4,5</sup>Department of ISE, Bahubali College of Engineering, Shravanabelagola-573135

**Abstract**—Resource allocation is a critical challenge in modern computing environments where limited system resources must be shared among multiple competing tasks. Traditional static allocation methods fail to adapt to dynamic workloads and varying demand patterns. This paper presents an intelligent resource allocation system inspired by Reinforcement Learning (RL) principles for efficient distribution of system resources. The proposed system analyzes multimedia inputs such as documents, images, audio, and video files and allocates CPU, memory, and bandwidth based on file size, type, and processing complexity. The allocation logic ensures fairness, minimizes resource wastage, and improves overall system efficiency. Experimental results demonstrate effective resource utilization and adaptability, making the system suitable for real-world applications such as cloud computing and multimedia processing platforms.

**Index Terms**—Reinforcement Learning, Resource Allocation, CPU Scheduling, Bandwidth Allocation, Multimedia Processing.

## I. INTRODUCTION

In recent years, the rapid growth of multimedia data has significantly increased the demand for efficient and intelligent resource management in modern computing systems. Applications that process images, videos, audio files, and large documents require dynamic and continuous allocation of critical system resources such as CPU, memory, and network bandwidth. The heterogeneous nature of these multimedia workloads introduces varying computational requirements, making resource management a complex and challenging task. Traditional resource allocation techniques are generally based on static rules, fixed priorities, or predefined thresholds, which fail to adapt to real-time workload variations. As a result, these methods often lead to inefficient resource utilization, increased

latency, performance bottlenecks, and resource wastage under dynamic operating conditions.

Reinforcement Learning (RL) provides a promising and adaptive approach to overcome these challenges by enabling systems to learn optimal decision-making strategies through continuous interaction with the environment. In an RL-based framework, the system observes the current state of resource demand, takes allocation actions, and receives feedback in the form of rewards that reflect the efficiency and fairness of the allocation. Over time, the system improves its decision-making capability by learning from past experiences. This paper proposes a smart resource allocation system inspired by Reinforcement Learning concepts that dynamically distributes system resources based on file characteristics, demand patterns, and processing requirements. The proposed approach aims to maximize overall system efficiency, ensure fair resource distribution, and reduce wastage, thereby addressing the key limitations of conventional static allocation methods.

## II. LITERATURE SURVEY

The literature survey is done before the formulation of the research aims and objectives, because we have to check if same research problem has been addressed. It involves a systematic and comprehensive analysis of books, scholarly articles and other sources relevant to a specific topic providing a base of knowledge on a topic. The literature survey is important and should be done at the beginning of any project. Writing the literature survey shows the reader how our work relates to existing research and what new insights it will contribute.

Z. Ji, Z. Qin. [1]

This paper proposes a Meta Federate Reinforcement Learning (MFRL) framework to enhance resource

allocation in cellular networks. Unlike traditional centralized RL approaches, the proposed model enables user devices to perform local training, thereby reducing communication overhead and improving data privacy. The base station applies meta-learning to generate an effective global model initialization that can quickly adapt to new network conditions. This approach allows the system to learn how to learn, resulting in faster convergence and improved flexibility. By sharing only model updates instead of raw data, the framework significantly minimizes transmission costs. The model also supports asynchronous communication and heterogeneous devices, making it suitable for practical 5G and 6G deployments. Simulation results show that MFRL outperforms centralized and decentralized RL methods in terms of energy efficiency, convergence speed, and overall performance. The proposed framework demonstrates strong generalization capability in dynamic and unseen environments.

Y. Cao. [2]

This paper investigates resource allocation in non-terrestrial networks (NTNs) with a focus on Low Earth Orbit (LEO) satellite communication systems. Due to satellite mobility, fluctuating link quality, and limited onboard processing power, traditional allocation methods are inefficient in such environments. To overcome these challenges, the authors propose a collaborative deep reinforcement learning (DRL) framework involving both user equipment and satellites. The system adopts a two-time-scale learning approach, where user equipment learns long-term value functions while satellites perform fast, rollout-based scheduling decisions. This design balances computational load and enables near real-time resource optimization. The collaborative framework adapts effectively to dynamic network conditions, including frequent user mobility and changing satellite positions. Unlike greedy or rule-based methods, the DRL agents continuously learn and adjust resource decisions. Simulation results show improvements in throughput, fairness, and response time compared to baseline approaches. The framework also reduces satellite computation overhead by offloading learning tasks to user equipment. Overall, the proposed solution provides a scalable and adaptive approach for future space-based internet and satellite communication systems.

M. A. Hady. [3]

This survey paper presents a comprehensive review of Multi-Agent Reinforcement Learning (MARL) techniques applied to resource allocation problems in decentralized systems. As modern networks such as IoT, vehicular systems, and cloud infrastructures grow in scale and complexity, centralized control becomes inefficient. MARL enables multiple agents to learn and make decisions simultaneously while interacting in a shared environment. The survey categorizes existing research based on application domains including wireless networks, cloud and edge computing, and energy systems. It highlights how MARL improves resource distribution, task scheduling, bandwidth management, and power control through decentralized learning. The paper also discusses key technical challenges such as scalability, coordination, partial observability, and non-stationarity. Various solutions including communication mechanisms, parameter sharing, and centralized training with decentralized execution (CTDE) are analyzed. The authors further examine open issues like credit assignment, robustness, and safety constraints. Overall, the survey serves as a valuable roadmap for applying MARL to dynamic and distributed resource allocation problems.

P. Li. [4]

This paper presents a reinforcement learning-based framework using Q-learning for adaptive resource scheduling in dynamic computing environments. Traditional scheduling techniques rely on static rules and heuristics, which are ineffective under changing workloads and system conditions. To overcome this, the authors model the scheduling problem as a Markov Decision Process (MDP), where the RL agent learns optimal task-to-resource allocation policies through interaction with the environment. The Q-learning agent evaluates state-action pairs to balance task completion time, resource utilization, and energy consumption. The approach is well-suited for discrete action spaces, enabling precise scheduling decisions. Experimental evaluations under fluctuating workloads and heterogeneous resources show significant improvements over classical methods such as FCFS and Round Robin. The RL-based scheduler achieves higher throughput, reduced idle time, and lower energy usage. A key advantage of the approach is its continuous adaptability to system changes. The paper

also discusses extending the framework using Deep Q-Networks for larger and more complex systems. Overall, the study demonstrates the effectiveness of RL for intelligent and adaptive resource management. A. Lim. [5]

This paper explores the application of reinforcement learning (RL) for dynamic memory allocation, addressing the limitations of traditional static strategies such as first-fit, best-fit, and worst-fit. Conventional memory allocators rely on fixed rules and lack adaptability, often resulting in fragmentation and inefficient memory utilization. The proposed approach models memory allocation as an RL problem, where system memory conditions form the state, allocation decisions represent actions, and rewards capture allocation success, efficiency, and fragmentation reduction. Through continuous interaction with the environment, the RL agent learns an adaptive allocation policy that optimizes long-term memory performance. The model is evaluated under realistic and adversarial memory request patterns, including unpredictable workloads and varying block lifetimes. Experimental results show that the RL-based allocator consistently outperforms traditional methods in reducing fragmentation and allocation failures. A key strength of the approach is its **online learning capability**, allowing real-time adaptation to changing memory demands. The paper highlights the potential of extending the framework using deep RL techniques for large-scale systems. Overall, the study demonstrates the effectiveness of RL in intelligent memory management and its relevance to modern computing environments.

S. Malhotra. [6]

This paper investigates the application of Deep Reinforcement Learning (DRL) for dynamic resource allocation in wireless networks. Traditional heuristic-based allocation methods fail to adapt effectively to increasing user demand and heterogeneous wireless environments. To address this, the authors employ Deep Q-Network (DQN) and Proximal Policy Optimization (PPO) to allocate resources such as bandwidth, power, and antenna usage in real time. The system models a multi-antenna base station serving multiple mobile users competing for limited resources. DQN estimates action values based on system states, while PPO directly learns optimal policies under uncertainty. Both approaches enable adaptive

decision-making in non-stationary wireless conditions. Performance evaluation shows that DQN and PPO outperform traditional schemes like round-robin and equal allocation in terms of throughput, fairness, and convergence speed. The study also analyzes the impact of learning rates on stability and efficiency.

Arisrei Lim. [7]

This paper addresses resource allocation challenges in the Internet of Vehicles (IoV), which involves high mobility, dense communication, and strict latency requirements. The authors propose a Multi-Agent Reinforcement Learning (MARL) framework to support efficient Vehicle-to-Infrastructure (V2I) and Vehicle-to-Vehicle (V2V) communications. Each vehicle and roadside unit is modeled as an autonomous learning agent using an actor-critic architecture to separate policy learning from value estimation. The framework is enhanced with prioritized experience replay, allowing agents to learn more effectively from important interactions. Agents make decisions based on local observations such as channel conditions, traffic density, and task queues. Simulation results in high-density urban scenarios show significant improvements over baseline methods, including higher communication capacity and successful transmission rates. The decentralized nature of MARL suits IoV environments where centralized control is impractical. The system adapts quickly to network changes as vehicles join or leave. Overall, the study demonstrates that MARL provides a scalable, adaptive, and efficient solution for real-time resource allocation in vehicular networks and future smart transportation systems

S. Gracla. [8]

This paper proposes a multi-task deep reinforcement learning (DRL) framework for robust resource allocation in environments affected by rare and extreme disruptions, known as Black Swan events. Traditional allocation methods and standard DRL models often fail under such anomalies due to poor generalization and lack of robustness. To address this, the authors formulate resource allocation as a multi-task learning problem, enabling the DRL agent to learn across both normal and rare scenarios simultaneously. The framework integrates Elastic Weight Consolidation (EWC) and Gradient Episodic Memory (GEM) into an actor-critic architecture to prevent

catastrophic forgetting. EWC preserves previously learned knowledge, while GEM allows rehearsal of past experiences. Experimental results under simulated disruptions such as sudden demand spikes and link failures show improved stability, faster recovery, and lower reward variance compared to single-task DRL models. The approach demonstrates strong generalization and resilience in unpredictable environments. Overall, the study highlights the importance of robustness in RL-based resource management and provides a reliable framework for critical systems such as cloud, wireless, and edge computing infrastructures.

D. Zhang. [9]

This survey paper reviews the application of reinforcement learning (RL) techniques to spatial resource allocation problems across domains such as transportation, logistics, and industrial systems. Spatial resource allocation involves distributing resources over geographical or networked spaces where location and spatial relationships are critical. The authors explain the fundamentals of RL and highlight its advantages over traditional optimization methods in handling dynamic and spatially dependent environments. The survey categorizes RL applications including vehicle routing, traffic signal control, warehouse management, and industrial scheduling. Algorithms such as Q-learning, policy gradient methods, and deep RL are discussed in detail. The paper also examines key challenges such as scalability, multi-agent coordination, and large state spaces. To address these issues, approaches like decentralized learning, hierarchical RL, and multi-agent RL are reviewed. The survey identifies open research problems including real-world deployment and simulation-to-reality transfer. Overall, the paper serves as a comprehensive reference, emphasizing RL's potential for efficient, adaptive, and real-time spatial resource management.

Z. Zheng. [10]

This paper investigates the use of reinforcement learning (RL) for efficient resource allocation in smart grid communication networks, where both power and network resources must be managed under strict constraints. The authors model the allocation problem as a Markov Decision Process (MDP) in a multi-domain virtual network environment. A policy-based

RL framework is proposed to dynamically allocate resources while considering bandwidth, latency, and energy consumption limits. Unlike traditional rule-based methods, the RL agent adapts continuously to changing network conditions. The policy network enables fast decision-making and improves the acceptance ratio of network requests. Experimental results show notable improvements in CPU utilization, latency control, and operational cost reduction. The approach performs particularly well under fluctuating demand and multi-domain coordination challenges. The framework also enhances system resilience by handling overloads and failures effectively. Overall, the study demonstrates that RL can significantly improve efficiency, scalability, and reliability in smart grid communication systems.

### III. FRAMEWORK

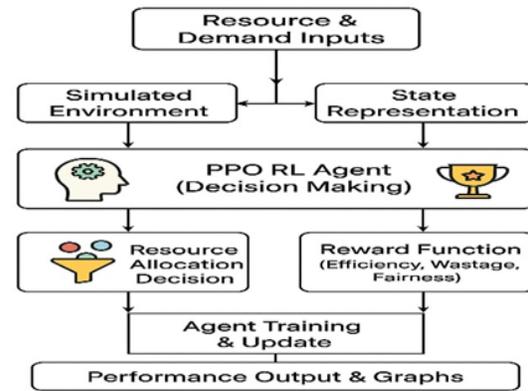


Fig. 1. Overall system architecture

Figure 1 This diagram illustrates an intelligent task and resource allocation system designed for distributed computing environments. Users submit input tasks, which are first collected by the Task Receiver. These tasks are then passed to the Intelligent Agent, the central decision-making unit, which determines the most suitable physical servers for execution. The Resource Optimizer, aided by a Mapping Table that records previous task-to-server allocations, evaluates the agent's decisions and provides rewards to improve future task assignments. Once tasks are processed on the physical servers, the completed results are returned to the users. Users can also provide Quality of Experience (QoE) feedback, which, along with internal rewards from the optimizer,

is fed back to the Intelligent Agent to enhance its learning and optimize subsequent resource allocation. This feedback-driven loop ensures dynamic, efficient task distribution while maximizing system performance and user satisfaction.

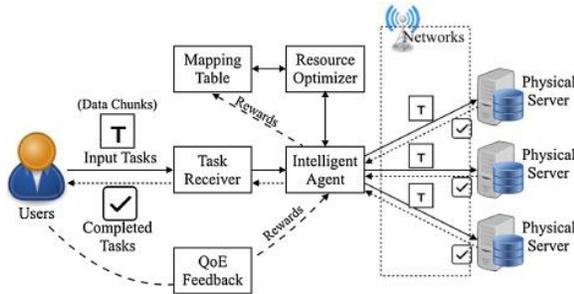


Fig. 2. Use case diagram for users

Figure 2 The diagram illustrates a reinforcement learning-based intelligent resource allocation system for distributed computing environments. Users submit input tasks, which are received by the Task Receiver. These tasks are then managed by an Intelligent Agent, which is responsible for assigning the tasks to appropriate Physical Servers for execution. The Intelligent Agent interacts with a Resource Optimizer and a Mapping Table to determine the most efficient allocation of tasks based on available resources and previous performance data. Once the tasks are processed by the physical servers, the results are sent back to the users as Completed Tasks. Additionally, users provide Quality of Experience (QoE) Feedback, which, along with the system’s own reward signals, is fed back to the Intelligent Agent to continuously improve the task allocation strategy. The system thus forms a closed-loop framework, leveraging feedback and rewards to optimize resource utilization and enhance overall performance.

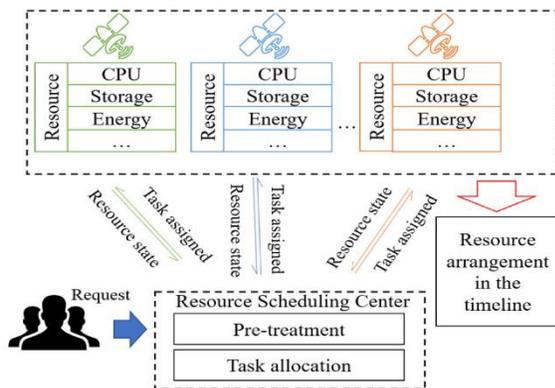


Fig. 3. Resource Demand and Allocation Hierarchy

Figure 3 The Resource Demand and Allocation Hierarchy is a critical component of the proposed Reinforcement Learning-Based Smart Resource Allocation System, providing a structured framework for requesting, prioritizing, and distributing limited system resources among competing tasks. In modern computing environments like cloud platforms, data centers, and edge systems, CPU, memory, and bandwidth are shared by multiple applications with varying requirements. At the first level of the hierarchy, task demands are generated, specifying the amount of CPU, memory, and bandwidth needed for execution. Capturing these requirements accurately is essential, as under-allocation slows tasks and over-allocation wastes resources. The next level involves classifying and prioritizing tasks based on deadlines, criticality, or system-defined policies, ensuring high-priority tasks receive sufficient resources and preventing starvation or unfair allocation. Once task priorities are established, the system evaluates resource availability and constraints in real time, comparing incoming demands against available CPU, memory, and bandwidth. The PPO-based Reinforcement Learning (RL) agent then observes this hierarchical structure and system state to make optimal allocation decisions. Unlike static or rule-based methods, the RL agent learns from experience, balancing objectives such as maximizing utilization, minimizing wastage, ensuring fairness, and meeting performance targets. Resource allocation execution applies these decisions by distributing CPU cores, memory blocks, and network bandwidth according to priority and demand, while performance metrics like utilization, task completion time, fairness, and resource wastage are calculated to provide feedback for continuous learning and adaptation.

The hierarchical model also enhances fairness, load balancing, scalability, and robustness. By organizing tasks and allocations hierarchically, resources are distributed proportionally, avoiding monopolization and maintaining system stability. Reinforcement learning further improves adaptability, enabling the system to handle dynamic workloads, demand spikes, and varying configurations without manual intervention. The clear structure simplifies monitoring and analysis, allowing administrators to track performance at different hierarchy levels. Overall, this approach ensures efficient utilization of CPU, memory, and bandwidth, reduces bottlenecks and

wastage, and provides intelligent, automated resource management suitable for large-scale, dynamic computing environments.

## V. RESULT

The proposed reinforcement learning-based resource allocation system demonstrates effective and intelligent management of CPU, memory, and bandwidth resources. Through the Proximal Policy Optimization (PPO) algorithm, the RL agent learns optimal allocation strategies over multiple training episodes, gradually improving reward values and reducing penalties associated with inefficient resource usage. Experimental results indicate that the system efficiently handles varying workloads, detects and resolves bottlenecks, and ensures fair distribution of resources across tasks. Validation through different test cases, including file uploads of varying sizes, confirms the system's ability to maintain stability, adapt to dynamic demands, and achieve near-optimal performance compared to traditional static allocation methods. The interactive dashboards and real-time monitoring further enhance transparency and usability.

## VI. CONCLUSION

The proposed reinforcement learning-based smart resource allocation system successfully addresses the challenges of efficient resource management. By dynamically allocating CPU, memory, and bandwidth based on real-time demand, the system improves overall resource utilization.

The reinforcement learning agent learns optimal allocation strategies through continuous interaction with the environment. Bottleneck detection and reward-penalty feedback helps prevent resource overloading. Experimental results show improved performance across multiple episodes.

The system adapts well to varying workloads. Validation results confirm stable and reliable operation. Overall, the proposed approach provides an intelligent, scalable, and efficient solution for modern resource allocation problems.

## VII. FUTURE SCOPE

The system can be extended to larger-scale and more complex environments, including cloud data centers, edge computing networks, and IoT infrastructures, where multiple resources must be managed simultaneously. Incorporating advanced deep reinforcement learning techniques and multi-agent frameworks can improve adaptability and scalability in dynamic and high-dimensional scenarios. The model could also integrate predictive analytics to anticipate workload spikes and optimize allocation proactively. Additionally, real-world deployment could enable fully autonomous, self-optimizing systems capable of continuous learning, resource scaling, and fault tolerance. Cross-domain applications in healthcare, transportation, and industrial automation further highlight the potential of reinforcement learning to transform resource management in diverse sectors.

## REFERENCES

- [1] Z. Ji, Z. Qin, and X. Tao. 2023. "Meta Federated Reinforcement Learning for Distributed Resource Allocation." *IEEE Transactions on Neural Networks and Learning Systems*, July 2023, pp. 2112–2124.
- [2] Y. Cao, S.-Y. Lien, Y.-C. Liang, D. Niyato, and X. Shen. 2024. "Terrestrial Networks." *ACM Computing Surveys*, February 2024, pp. 54–76.
- [3] M. A. Hady, S. Hu, M. Pratama, J. Cao, and R. Kowalczyk. 2025. "Multi-Agent Reinforcement Learning for Resource Allocation Optimization: A Survey." *Elsevier Information Fusion*, April 2025, pp. 105–139.
- [4] P. Li, Y. Xiao, J. Yan, X. Li, and X. Wang. 2024. "Reinforcement Learning for Adaptive Resource Scheduling in Complex System Environments." *Springer Journal of Intelligent & Robotic Systems*, November 2024, pp. 322–339.
- [5] A. Lim and A. Maddukuri. 2024. "Reinforcement Learning for Dynamic Memory Allocation." *IEEE Access*, October 2024, pp. 78541–78554.
- [6] S. Malhotra, F. Yashu, M. Saqib, D. Mehta, J. Jangid, and S. Dixit. 2025. "Deep Reinforcement Learning for Dynamic Resource Allocation in Wireless Networks." *Computer Communication (Elsevier)*, February 2025, pp. 101–117.

- [7] Arisrei Lim, Abhiram Maddukuri. 2025. "Multi-Agent Reinforcement Learning for Efficient Resource Allocation in Internet of Vehicles." arXiv preprint arXiv:2501.06789, January 2025, pp. 1–15.
- [8] S. Gracla, C. Bockelmann, and A. Dekorsy. 2023. "A Multi-Task Approach to Robust Deep Reinforcement Learning for Resource Allocation." IEEE Internet of Things Journal, April 2023, pp. 4587–4598.
- [9] D. Zhang, M. Wang, J. Mango, X. Li, and X. Xu. 2024. "A Survey on Applications of Reinforcement Learning in Spatial Resource Allocation." ACM Transactions on Cyber-Physical Systems, March 2024, pp. 89–112.
- [10] Z. Zheng, Y. Han, Y. Chi, F. Yuan, W. Cui, H. Zhu, Y. Zhang, and P. Zhang. 2023. "Network Resource Allocation Algorithm Using Reinforcement Learning Policy-Based Network in a Smart Grid Scenario." IEEE Transactions on Smart Grid, August 2023, pp. 1185–1197.