Reinforcement Learning Methods for Adaptive Cloud Resource Allocation

Hiba Qasim[1], Huda Abdulaali[1] and Soukaena Hasan[2]
[1]*Department of Computer Science, College of Science, Al-Mustansiriyah University, 10052 Baghdad, Iraq*
[2]*Department of Computer Science, College of Science, University of Technology, 10066 Baghdad, Iraq*
hibaqasimjaber@uomustansiriyah.edu.iq, huda_it@uomustansiriyah.edu.iq, figSoukaena.h.hashem@uotechnology.edu.iq

Keywords: Cloud Computing, Reinforcement Learning, Resource Allocation, Load Balance.

Abstract: Because cloud workloads and service demand shift continuously, cloud platforms need resource-allocation mechanisms that adjust on the fly to keep quality of service (QoS) high, contain operating expenses, and avoid the common failure mode in which some servers sit idle while others are pushed past capacity. What makes this hard is that cloud environments are dynamic and their states keep changing, so allocation decisions cannot rely on fixed rules. Conventional approaches typically require expert tuning or repeated manual iteration, which makes them costly and difficult to adapt. A central difficulty is therefore assigning resources efficiently enough to satisfy many concurrent users and applications with competing requirements. Reinforcement learning (RL), through its exploration-exploitation mechanism, offers a way to keep adapting as conditions evolve. This paper's central argument is that dynamic-workload resource allocation in cloud settings can be handled effectively with RL-based methods, which let systems learn from ongoing changes, adjust performance in real time, and cut costs accordingly. The review further finds that hybrid schemes pairing meta-heuristic search with reinforcement learning surpass conventional RL algorithms once tested on real-world data.

1 INTRODUCTION

Cloud computing has become indispensable across academia, business, and industry because it consolidates computing resources drawn from data centers spread throughout the world [1], [2]. The service models differ in what they deliver to the customer: software as a service hands the user a ready application, platform as a service supplies a preconfigured development or runtime environment, and infrastructure as a service (IaaS) gives the customer direct control over provisioned processing power, storage, networking, and other core computing building blocks. Today, mobile devices and computers routinely serve as the endpoints through which images, audio, and video are exchanged over communication networks [3]. As adoption of the internet of things continues to rise, real-time monitoring driven by networked sensors is increasingly common [4]. Because cloud servers are typically located far from where data originates, purely cloud-centric designs struggle to meet such demanding requirements, producing latency that is unacceptable for time-sensitive applications. Fog computing (FC) emerged to close this gap by positioning processing capability at the network edge [5], pushing data handling and storage physically closer to its source so that response times shrink [6]. Edge computing (EC) goes a step further than FC: instead of routing complex processing back toward the network core before reaching the cloud, EC resolves requests directly on end devices or nearby gateways. The relationship between edge devices, fog computing, and centralized cloud infrastructure is illustrated in Figure 1. [7]. This edge layer is inherently dynamic, characterized by large device populations, frequent user movement, varied application types, and traffic that comes and goes [8]. Virtualization remains the defining property of the cloud environment itself.

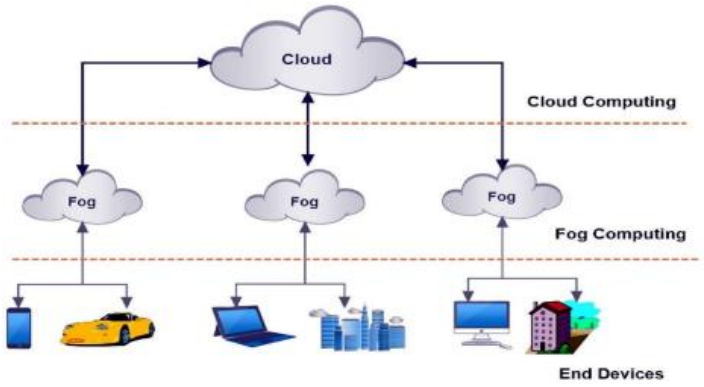

Figure 1: Fog computing, edge devices and cloud computing architecture.

Among the advantages this brings are the ability to consolidate servers and push utilization rates higher. Large technology providers, including Microsoft, Amazon, and Google, run massive data centers whose resource management is inherently complex [9], with each physical host hosting several virtual machines (VMs), each carrying its own size and workload profile.

Because demand is hard to predict, individual servers can end up over- or under-loaded, producing an imbalance in how resources are distributed among the VMs on a given host. That imbalance, in turn, can drive up energy consumption, destabilize QoS, and trigger breaches of service level agreements [10].

Resource allocation sits at the core of cloud architecture: it determines how well a system can scale and automatically match resources to shifting workload and demand, thereby keeping applications available, responsive, and efficient in their resource use [11]. Task allocation, more specifically, aims to cut processing time while spreading the workload across available resources as evenly as possible, a process depicted for cloud computing in Figure 2 [12].

Many classic techniques to allocate resources in cloud computing depend on heuristics, rules and control heory. A promising technique for solving resource allocation problems with minimal complexity and excellent adaptability is reinforcement learning (RL). However, in complex cloud systems, the issue of high-dimensional state space concerns traditional techniques [13].

Figure 2: Resource allocation system in cloud computing.

RL's main advantage is its ability to cope with dynamic, unpredictable conditions even when they are difficult [14]; Figure 3 gives a visual depiction of an RL system.

Figure 3: RL- Agent and environment interaction.

It falls to the agent to choose an appropriate action at each step, and within the RL framework the correct choice is whichever action yields the highest reward [15]; separately, identity authentication remains a core security mechanism for verifying users before system access is granted [16]. This paper's contributions can be summarized as follows:

1) This paper reviews the recent methods in the clustering, optimization and energy management for resource allocation in the cloud.

2) This paper focuses not only on energy, but also on SLA violation and QoS.
3) Limitations in the existing methods are discussed in this study.
4) This study's sections are organized as follows: A thorough examination of a number of RL techniques is given in Section 2, which evaluates them according to their implementation, application domain, focus, strengths, and limits. The technique used in this comprehensive evaluation will be discussed in Section 3. The article is concluded, and future work is discussed in Section 4.

2 RELATED WORK

This section describes the summary of the existing research on resource allocation in cloud computing based on reinforcement learning.

One study combined fuzzy logic with reinforcement learning to steer cloud infrastructure toward greener resource choices, reporting strong results in cutting SLA breaches while keeping data-center energy efficiency high; the approach was validated against PlanetLab virtualization traces, where it achieved favorable PUE and CPU-consumption figures [11]. A related line of work built a prediction-enabled feedback-control allocator (PCRA) around a purpose-built Q-value prediction model that forecasts management-operation outcomes under varying system states, drawing on several prediction learners trained via Q-learning; this PCRA scheme was reported to beat conventional ML-based and rule-based baselines by 7% and 13%, respectively [17]. Another architecture targeted latency reduction, workload diversity, and cost-effectiveness simultaneously, with experiments showing that its deep-reinforcement-learning (DRL) core - implemented as an Asynchronous Advantage Actor-Critic (A3C) agent - improved the scalability, adaptability, and overall efficiency of dynamic cloud systems, the primary aim being joint gains in energy efficiency and QoS [13], [18].

One implementation combined a brute-force search strategy with an LSTM (long short-term memory) network, and evaluation across several metrics - service quality, resource usage, cost, and response time - showed clear gains relative to conventional
baseline models [19]. Elsewhere, a collaborative cloud-edge scheme (CERAI/CERAU) targeting both private and public clouds built its allocation logic around P-DQN and deterministic policy gradients within a deep-RL framework, tested on simulated data as well as real Google traces [20]. An approach called IMARM was designed to raise performance, cut execution time, strengthen fault tolerance, lower energy consumption, and balance load, achieving this by managing VM start/stop operations and enabling task migration [21]. A separate multi-objective scheduler (MOABCQ) for heterogeneous cloud settings was shown to shrink makespan and cost, boost throughput, and support load balancing using whatever resources the system currently has available [22]. The AC-CIWAS model takes a different tack, continuously tracking server state and the energy impact of incoming workloads; it relies on an actor-critic (AC) method within a DRL framework to optimize task placement for energy efficiency, computing expected cumulative reward over time [23]. Load-balancing behavior has also been examined via Workflowsim simulations on the Sipht job dataset, comparing Round-Robin, Minimum-Minimum (Min-Min), Minimum Completion Time (MCT), Maximum-Minimum (Max-Min), and First-Come-First-Served (FCFS) strategies [24]. A further collaborative scheme for cloud content delivery was built to make the best use of processing, caching, and communication resources while cutting network latency in asymmetric IoV settings [25]. Another method clusters VMs into overloaded and underloaded groups based on measured load, then applies a Hybrid Lyrebird Falcon Optimization (HLFO) algorithm - merging Falcon Optimization (FOA) with Lyrebird Optimization (LOA) - together with RL to sharpen clustering and improve load-balancing outcomes [26]. Placement and allocation of VMs within a cloud system forms the central concern of yet another study [27].

A method named RLPRAF pursues optimal cloud resource allocation by blending reinforcement learning with linear regression and auto-computation; its authors report cost savings of roughly 30% and note applicability to real-world deployments [28]. Similarly, a D4PG-based reinforcement-learning approach targets dynamic allocation together with cost reduction in cloud computing settings [29]. The comparison of these previously surveyed approaches is summarized in Table 1.

3 RESULTS AND DISCUSSION

Across the surveyed papers, the common thread is using reinforcement learning to allocate cloud resources more effectively and efficiently. The works in [13], [17], [18], [24], [25], [26] in particular target QoS and energy-efficiency gains through better allocation strategies. Study [13] is scalable and converges quickly while handling heterogeneous resources well, though at the price of implementation complexity and sensitivity to hyperparameter choices. Study [17] delivered strongly validated

results but proved difficult to implement in practice. The model in [18] adapts to diverse workloads, cuts latency, and improves cost-efficiency, yet carries added model complexity and computational overhead. Study [24] has been applied to real cloud platforms such as AWS and GCP but is limited by a thin set of evaluation metrics. Study [25] adds metrics such as latency reduction, execution-time minimization, fast convergence, and responsiveness to changing network conditions, but its evaluation is not comprehensive and would need further refinement before it could be deployed more broadly.

Several papers [11], [13], [19], [20], [22], [23], [27], [28] draw on real-world traces from providers such as Google Cloud and Alibaba Cloud, which strengthens confidence in their results, and most rely on comparable metrics - response time, throughput, resource efficiency - to judge algorithm performance, even though their specific emphases differ. Study [11] pushes CPU utilization up and Power Usage Effectiveness down while safeguarding SLAs, but is scoped narrowly to CPUs and data centers. Study [20] adapts to varying user demand and workload but was built for smaller systems and may falter as user counts or resource pools grow. Study [22] balances workload and lowers cost while paying less attention to energy considerations. Study [23] adapts dynamically and improves energy efficiency at the cost of heavier training computation and time. Study [27] targets dynamic VM allocation and reduces wasted resources but itself demands substantial computational resources. Study [28] reduces SLA violations and manages resources proactively, remaining robust to workload swings, though it is difficult to generalize and implement.
Study [29] improves efficiency, scalability, and cost, but relies on a single-objective allocation formulation and does not specify a dataset. Some studies rely on synthetic experimental setups [17], [26], while others draw on data from genuine cloud environments - a distinction that affects both the performance figures obtained and how much confidence can be placed in them.

Table 1: Comparison of the existing review papers.

ID	Year	Algorithm	Metric	Advantage	Disadvantage	Results
11	2018	Fuzzy Logic and Q-learning.	Energy	-maximize CPU utilization	- Focus on CPUs only	Resource load balance
17	2020	PCRA	-QoS - cost	-Integration of RL and ML	-Complex implementation	Efficient and enhance cost
18	2020	DRL	-QoS	-minimize latency	- Model Complexity	Efficient resource use
13	2021	actor-critic DRL (A3C)	-QoS	-Scalable	Complexity of Implement	Outperformed baseline strategies
19	2021	Implement an LSTM model	Cost	-Reducing waste -enhances efficiency	-complicated and requires special skills.	-model effectively predicts and optimizes cloud resource Allocation
20	2022	(DDPG) and (P-DQN)	- Cost per User	-adapt to different user demands and workloads	Works for Small Systems	- address the difficulty of resource Allocation
21	2022	(IMARM) using the multi-agent	Energy Time	Enhanced Fault Tolerance	Overhead from Checkpointing	Reduction Energy 30%
22	2022	ABC	-cost -RU	-Balanced Workload	Complexity in Dynamic	Reduce SLA Violations
23	2023	actor-critic	Time QoS	-Improved Energy Efficiency	-Requires computation for training	Reduce the job's workload
24	2023	Qlearning (FCF) (Min), and (RR)	- (QoS) - LB	model has practical applications	lack in evaluations metric	Reduce load balance
25	2023	- DQN	QoS	- adapt to the changes Of network	- Lack of Comprehensive evaluation	Optimize edge computing
26	2024	- Combining RNNs, CNNs	Energy	- dynamic load	-complex for real time workload	faster task completion
27	2024	Asynchronous Actor-Critic	LB	discrete event- Sim	- Dynamic VM Allocation	- getting stuck in suboptimal solutions
28	2024	Q-learning	CPU	CloudSim	-Reduction in SLA	-Complex to generalize
29	2024	(D4PG)	RU	--------	-improve efficiency	-Single objective

4 CONCLUSIONS

This review surveyed a range of reinforcement-learning-based techniques for cloud resource allocation.

The analysis shows that k-means underlies most of the clustering methods examined, though its dependence on a randomly chosen starting cluster undermines model effectiveness. Optimization-based methods such as HLFO and ABC deliver stronger QoS, but they are constrained by slow convergence and a tendency to settle into local optima.

For cloud resource allocation and prediction, reinforcement learning techniques including Q-learning, D4PG, Actor-Critic, and DQN were used. Several of the previous research that demonstrated better resource management performance employed the DQN and Q-learning techniques. Additionally, the hybrid strategy uses multi-objective optimization to optimize resource use with great success. In the future, it is recommended to employ hybrid approaches for meta-heuristics optimization strategies in cloud management for resource allocation.

ACKNOWLEDGEMENT

The Authors would like to thank Mustansiriyah University (https://uomustansiriyah.edu.iq/) Baghdad -Iraq for its support in the present work.

REFERENCES

[1] Shahidinejad, M. Ghobaei-Arani, and M. Masdari, "Resource provisioning using workload clustering in cloud computing environment: a hybrid approach," Cluster Comput., vol. 24, no. 1, pp. 319-342, 2021.

[2] Z. Chen, J. Hu, G. Min, C. Luo, and T. El-Ghazawi, "Adaptive and efficient resource allocation in cloud datacenters using actor-critic deep reinforcement learning," IEEE Trans. Parallel Distrib. Syst., vol. 33, no. 8, pp. 1911-1923, 2021.

[3] X. Chen, Y. Zhang, H. Wang, and L. Liu, "Resource allocation for cloud-based software services using prediction-enabled feedback control with reinforcement learning," IEEE Trans. Cloud Comput., vol. 10, no. 2, pp. 1117-1129, 2020.

[4] P. S. Prabha, "Adaptive cloud resource allocation using attention-driven deep reinforcement learning," Eng. Technol. Appl. Sci. Res., vol. 14, no. 3, pp. 15234-15241, 2024.

[5] K. Lone and S. A. Sofi, "A review on offloading in fog-based internet of things: Architecture, machine learning approaches, and open issues," High-Confidence Comput., vol. 3, no. 2, p. 100124, 2023.

[6] R. Liu, R. Piplani, and C. Toro, "Deep reinforcement learning for dynamic scheduling of a flexible job shop," Int. J. Prod. Res., vol. 60, no. 13, pp. 4049-4069, 2022.

[7] Alsadie, "A comprehensive review of AI techniques for resource management in fog computing: Trends, challenges and future directions," IEEE Access, vol. 12, pp. 1-20, 2024.

[8] Sonmez, A. Ozgovde, and C. Ersoy, "Fuzzy workload orchestration for edge computing," IEEE Trans. Netw. Serv. Manag., vol. 16, no. 2, pp. 769-782, 2019.

[9] R. Bianchini et al., "Toward ML-centric cloud platforms," Commun. ACM, vol. 63, no. 2, pp. 50-59, 2020.

[10] T. Khan et al., "Machine learning (ML)-centric resource management in cloud computing: A review and future directions," J. Netw. Comput. Appl., vol. 204, p. 103405, 2022.

[11] T. Thein, M. M. Myo, S. Parvin, and A. Gawanmeh, "Reinforcement learning based methodology for energy-efficient resource allocation in cloud data centers," J. King Saud Univ. Comput. Inf. Sci., vol. 32, no. 10, pp. 1127-1139, 2020.

[12] Pradhan, S. K. Bisoy, and A. Das, "A survey on PSO based meta-heuristic scheduling mechanism in cloud computing environment," J. King Saud Univ. Comput. Inf. Sci., vol. 34, no. 8, pp. 4888-4901, 2022.

[13] Z. Chen, J. Hu, G. Min, C. Luo, and T. El-Ghazawi, "Adaptive and efficient resource allocation in cloud datacenters using actor-critic deep reinforcement learning," IEEE Trans. Parallel Distrib. Syst., vol. 33, no. 8, pp. 1911-1923, 2021.

[14] M. Mousavi, J. Rezazadeh, and O. A. Sianaki, "Machine learning applications for fog computing in IoT: a survey," Int. J. Web Grid Serv., vol. 17, no. 4, pp. 293-320, 2021.

[15] P. Ladosz, L. Weng, M. Kim, and H. Oh, "Exploration in deep reinforcement learning: A survey," Inf. Fusion, vol. 85, pp. 1-22, 2022.

[16] R. M. Abdullah, S. A. H. Alazawi, and P. Ehkan, "SAS-HRM: Secure authentication system for human resource management," Al-Mustansiriyah J. Sci., vol. 34, no. 3, pp. 64-71, 2023.

[17] X. Chen et al., "Resource allocation for cloud-based software services using prediction-enabled feedback control with reinforcement learning," IEEE Trans. Cloud Comput., vol. 10, no. 2, pp. 1117-1129, 2020.

[18] Y. Li, X. Zhang, and R. Wang, "Deep Reinforcement Learning for Cloud Resource Allocation: A Survey," IEEE Trans. Cloud Comput., vol. 10, no. 3, pp. 1560-1578, 2022.

[19] Wu, Y. Gong, H. Zheng, Y. Zhang, J. Huang, and J. Xu, "Enterprise cloud resource optimization and management based on cloud operations," Appl. Comput. Eng., vol. 76, pp. 8-14, 2024.

[20] J. Xu, Z. Xu, and B. Shi, "Deep reinforcement learning based resource allocation strategy in cloud-edge computing system," Front. Bioeng. Biotechnol., vol. 10, p. 908056, 2022.

[21] Belgacem, S. Mahmoudi, and M. Kihl, "Intelligent multi-agent reinforcement learning model for resources allocation in cloud computing," J. King Saud Univ. Comput. Inf. Sci., vol. 34, no. 6, pp. 2391-2404, 2022.

[22] Kruekaew and W. Kimpan, "Multi-objective task scheduling optimization for load balancing in cloud computing environment using hybrid artificial bee colony algorithm with reinforcement learning," IEEE Access, vol. 10, pp. 17803-17818, 2022.

[23] M. Arvindhan and D. R. Kumar, "Adaptive resource allocation in cloud data centers using actor-critical deep reinforcement learning for optimized load balancing," Int. J. Recent Innov. Trends Comput. Commun., vol. 11, no. 5s, pp. 310-318, 2023.

[24] P. V. Lahande et al., "Reinforcement learning approach for optimizing cloud resource utilization with load balancing," IEEE Access, vol. 11, pp. 127567-127577, 2023.

[25] T. Cui, R. Yang, C. Fang, and S. Yu, "Deep reinforcement learning-based resource allocation for content distribution in IoT-edge-cloud computing environments," Symmetry, vol. 15, no. 1, p. 217, 2023.

[26] R. Khan, "Dynamic load balancing in cloud computing: Optimized RL-based clustering with multi-objective optimized task scheduling," Processes, vol. 12, no. 3, p. 519, 2024.

[27] Sundar, "Reinforcement learning techniques for autonomous cloud optimization and adaptive resource management," Int. J. Artif. Intell. Data Sci. Mach. Learn., vol. 6, no. 3, pp. 134-145, 2025.

[28] R. Panwar and M. Supriya, "RLPRAF: Reinforcement learning-based proactive resource allocation framework for resource provisioning in cloud environment," IEEE Access, vol. 12, pp. 1-14, 2024.

[29] S. K. Khanday, "Reinforcement learning strategies for dynamic resource allocation in cloud-based architectures," 2024.