RADemics

Hybrid Swarm Intelligence Models with Deep Q Learning for Dynamic Resource Allocation in Cloud Computing

Sowmiya.S.M, M. Hema, Gopika.G. S
VELS INSTITUTE OF SCIENCE, TECHNOLOGY AND ADVANCED STUDIES, EASWARI ENGINEERING COLLEGE, SATHYABAMA INSTITUTE OF SCIENCE AND TECHNOLOGY

Hybrid Swarm Intelligence Models with Deep Q Learning for Dynamic Resource Allocation in Cloud Computing

[1]Sowmiya.S.M, Assistant Professor, Department of Computer, Science and Engineering, Vels Institute of Science, Technology and advanced studies, Velan Nagar, P.V. Vaithiyalingam Road, Pallavaram, Chennai 600117, Mail ID: sowmiyasm20@gmail.com

[2]M. Hema, Assistant Professor, Information technology, Easwari Engineering College, Chennai, Mail ID: hema.m@eec.srmrmp.edu.in

[3]Gopika.G. S, Assistant Professor, Computer Science and Engineering, Sathyabama Institute of Science and Technology, Chennai, Mail ID: gopikags41@gmail.com /gopika.cse@sathyabama.ac.in

Abstract

Efficient and intelligent resource allocation remains a fundamental challenge in cloud computing due to the dynamic, heterogeneous, and high-dimensional nature of cloud environments. This chapter presents an adaptive hybrid optimization framework that integrates Swarm Intelligence (SI) algorithms with Deep Q Learning (DQL) to address the complexity of real-time decision-making in large-scale cloud systems. The proposed architecture leverages the global exploration capabilities of swarm-based models and the policy-learning strengths of reinforcement learning to enable adaptive and SLA-aware resource management. By combining collective search behavior with neural network-driven policy refinement, the hybrid system dynamically allocates computational resources, minimizes latency, and optimizes long-term performance under varying workload conditions. Extensive analysis demonstrates that the model efficiently navigates high-dimensional state spaces, balances exploration and exploitation, and improves convergence speed while reducing SLA violations. The chapter also includes complexity analysis and performance evaluations, confirming the scalability and responsiveness of the hybrid framework in distributed cloud infrastructures. This work contributes to the development of next-generation autonomous systems for cloud optimization by bridging bio-inspired computation and deep learning.

Keywords:

Swarm Intelligence, Deep Q Learning, Cloud Resource Allocation, SLA Awareness, Hybrid Optimization, Dynamic Scheduling

Introduction

The rise of cloud computing has revolutionized modern IT infrastructure by offering on-demand access to scalable computing resources over virtualized platforms [1]. This paradigm enables organizations to provision resources dynamically, reduce capital expenditure, and deliver services globally with minimal latency [2]. As cloud adoption increases, the demand for efficient resource management strategies becomes paramount [3]. Cloud service providers are tasked with allocating CPU cycles, memory, storage, and network bandwidth to meet service-level agreements (SLAs) while minimizing operational costs and energy consumption [4]. due to the unpredictable nature of user requests, heterogeneous resource configurations, and dynamic workloads, maintaining optimal performance under constrained conditions poses a complex and multi-objective optimization problem [5].

Traditional resource allocation methods, including static policies and deterministic heuristics, often lack the flexibility to handle real-time variability in workload patterns [6]. These approaches are typically reactive and rule-based, offering limited scope for continuous learning or adaptation in large-scale cloud environments [7]. As the scale and complexity of data centers grow, the computational overhead associated with these conventional techniques increases, rendering them inefficient [8]. Many of these methods operate under ideal assumptions and cannot capture the stochastic behavior inherent in real-world cloud systems [9]. This has led to an increased interest in intelligent optimization techniques capable of learning and adapting to uncertain, dynamic environments. Solutions based on machine learning, metaheuristics, and reinforcement learning are being explored to develop more autonomous, scalable, and SLA-compliant resource provisioning frameworks [10].

Swarm intelligence (SI) has emerged as a powerful alternative for solving optimization problems in distributed systems [11]. Inspired by the collective behavior of biological agents such as ants, birds, and bees, SI algorithms use decentralized control and simple communication rules to explore large solution spaces efficiently [12]. Algorithms like Particle Swarm Optimization (PSO), Ant Colony Optimization (ACO), and Artificial Bee Colony (ABC) have been successfully applied in task scheduling, VM migration, and energy-aware resource management in cloud environments [13]. These models are especially well-suited for high-dimensional optimization tasks due to their ability to escape local minima and converge toward globally optimal solutions. Their parallelizable and fault-tolerant nature also makes them attractive for deployment in cloud-based systems [14]. The absence of memory retention and long-term learning mechanisms restricts their adaptability in non-stationary environments, where system states and user demands evolve rapidly over time [15].