Innovation and Technology Studies
Print ISSN: 3007-6919
Online ISSN: 3007-6927
DOI: https://doi.org/10.61784/its3016

A REINFORCEMENT LEARNING-BASED ARCHITECTURE FOR HIERARCHICAL CONTROL OF API WORKFLOWS IN ENERGY-CONSTRAINED AD SYSTEMS

Kelvin Wong[1], Mei-Ling Chan[2], Victor Leung[2*]
[1]Department of Computer Science, City University of Hong Kong, Hong Kong region 999077, China.
[2]Department of Information Systems, City University of Hong Kong, Hong Kong region 999077, China.
Corresponding Author: Victor Leung, Email: 3830204@my.cityu.edu.hk

Abstract: Modern advertising systems face increasing complexity in API workflow management due to interconnected service dependencies, dynamic resource requirements, and stringent energy efficiency constraints. Traditional workflow orchestration approaches struggle to optimize complex API execution sequences while maintaining energy consumption within operational limits. The heterogeneous nature of advertising workflows, including real-time bidding pipelines, content personalization processes, and analytics aggregation tasks, requires sophisticated control mechanisms that can adapt to varying performance requirements and energy availability.

This study proposes a Reinforcement Learning (RL)-based architecture for hierarchical control of API workflows in energy-constrained advertising systems. The framework employs a multi-tier control structure where high-level workflow coordinators manage execution strategies while low-level API controllers optimize individual service performance within energy budgets. Deep Deterministic Policy Gradient (DDPG) and Twin Delayed Deep Deterministic Policy Gradient (TD3) algorithms enable adaptive workflow control policies that balance execution efficiency with energy consumption across distributed advertising infrastructure.

Experimental evaluation using enterprise advertising system traces demonstrates that the proposed architecture achieves 41% improvement in workflow completion rates while reducing energy consumption by 37% compared to traditional orchestration methods. The hierarchical approach successfully manages complex workflow dependencies and energy constraints, resulting in 33% better resource utilization efficiency and 29% reduction in workflow execution latency.

Keywords: Reinforcement learning; API workflow management; Hierarchical control; Energy-constrained systems; Deep deterministic policy gradient; Twin delayed DDPG; Advertising systems; Workflow orchestration

1 INTRODUCTION

Contemporary advertising systems have evolved into complex distributed architectures that orchestrate hundreds of Application Programming Interface (API) services through intricate workflow pipelines designed to deliver personalized advertising experiences to millions of users simultaneously[1]. These systems must efficiently manage complex workflow execution sequences that span multiple service dependencies, data processing stages, and decision points while operating within strict energy consumption constraints imposed by operational cost considerations and environmental sustainability requirements. The challenge lies in optimizing workflow performance across diverse execution patterns while maintaining energy efficiency and ensuring reliable service delivery[2].

Traditional API workflow orchestration approaches rely on static execution plans and rule-based scheduling policies that cannot adapt effectively to dynamic service performance variations or changing energy availability conditions[3]. Workflow engines typically employ predefined execution sequences that fail to consider real-time system conditions, service load variations, or energy consumption patterns[4]. These static approaches often result in suboptimal resource utilization, unnecessarily high energy consumption during low-demand periods, and potential service failures during energy-constrained operations.

The complexity of advertising system workflows stems from several interconnected factors including diverse service types with varying computational requirements, complex dependency relationships between workflow stages, dynamic user request patterns that create unpredictable load distributions, and energy constraints that fluctuate based on power availability and cost considerations. Real-time bidding workflows require rapid execution of multiple API calls within millisecond timeframes, while content personalization processes involve computationally intensive machine learning inference tasks that consume substantial energy resources[5]. Analytics workflows aggregate large volumes of data through sequential API operations that can tolerate higher latency but demand consistent throughput[6].

Energy-constrained operations introduce additional complexity to workflow management by requiring simultaneous consideration of execution performance and power consumption across all workflow stages[7]. Traditional optimization approaches focus primarily on execution time and throughput without considering energy efficiency implications, missing opportunities for sustainable operation that could reduce both environmental impact and operational costs[8]. Energy-aware workflow management requires sophisticated control mechanisms that can balance immediate performance requirements with longer-term energy efficiency objectives.

Machine learning techniques, particularly Reinforcement Learning (RL), offer promising solutions for adaptive workflow control in complex energy-constrained advertising systems[9]. RL agents can learn optimal workflow orchestration policies through continuous interaction with system environments while adapting to changing service performance characteristics and energy availability conditions[10]. The ability to balance multiple competing objectives including execution efficiency, energy consumption, and service reliability makes RL particularly suitable for complex workflow optimization challenges[11].

Deep reinforcement learning algorithms extend traditional RL capabilities by incorporating neural networks to handle high-dimensional state spaces representing complex workflow execution states, service performance metrics, and energy consumption patterns. Deep Deterministic Policy Gradient (DDPG) algorithms enable stable policy learning for continuous control problems including resource allocation, execution timing, and energy distribution across workflow stages. Twin Delayed Deep Deterministic Policy Gradient (TD3) algorithms provide enhanced stability and reduced overestimation bias in complex multi-objective optimization environments.

This research proposes a novel RL-based architecture specifically designed for hierarchical control of API workflows in energy-constrained advertising systems. The architecture employs a multi-tier control structure where high-level workflow coordinators manage strategic execution decisions including service selection, routing optimization, and energy allocation while low-level API controllers focus on tactical service performance optimization within allocated energy budgets.

The framework integrates comprehensive state representations including current workflow execution status, service performance indicators, energy consumption metrics, and resource availability measurements across distributed advertising infrastructure. Action spaces encompass both strategic workflow decisions including execution path selection and service prioritization as well as tactical resource allocation decisions including CPU allocation ratios and energy consumption limits for individual API services.

2 LITERATURE REVIEW

API workflow management in distributed systems has been extensively studied as service-oriented architectures have become dominant paradigms for complex application development[12]. Early research focused on basic workflow orchestration techniques including sequential execution planning, parallel processing coordination, and error handling mechanisms[13]. These foundational studies established principles for workflow management but were limited by static execution models that could not adapt to dynamic system conditions or varying service performance characteristics.

Cloud-based workflow orchestration research evolved to address the unique challenges of distributed computing environments including dynamic resource allocation, service scaling, and fault tolerance mechanisms[14]. Studies examined various approaches for optimizing workflow execution in cloud platforms including intelligent service selection, adaptive routing strategies, and resource optimization techniques[15]. However, most research focused on performance optimization without considering energy consumption or sustainability implications.

Energy-aware computing research has gained significant attention as organizations seek to reduce operational costs and environmental impact while maintaining system performance[16]. Studies examined various approaches for incorporating energy considerations into system optimization including dynamic voltage scaling, workload consolidation, and intelligent resource provisioning[17]. However, most research focused on computational workloads rather than complex workflow orchestration challenges.

RL applications to distributed system management began with simple resource allocation and scheduling problems in relatively homogeneous computing environments[18]. Early studies demonstrated that RL agents could learn effective system management policies through interaction with simulation environments. However, these applications were limited to small-scale systems and single-objective optimization scenarios that did not capture the complexity of modern distributed workflow systems[19].

Deep reinforcement learning research in distributed systems expanded the applicability of RL to more complex optimization problems by incorporating neural networks to handle high-dimensional state spaces and complex decision environments[20]. Studies showed that DDPG could effectively learn resource allocation policies while policy gradient methods proved valuable for continuous parameter optimization[21]. However, most research remained focused on traditional distributed system scenarios rather than specialized workflow orchestration challenges.

Hierarchical reinforcement learning emerged as a solution to scalability challenges in complex distributed systems by decomposing optimization problems into multiple levels of abstraction. Research demonstrated that hierarchical approaches could achieve better learning efficiency and policy performance in large-scale systems compared to monolithic RL methods[22]. However, applications to workflow orchestration in energy-constrained environments remained largely unexplored.

Workflow optimization research in advertising systems has examined specialized techniques for the unique requirements of advertising platforms including real-time bidding optimization, content personalization workflows, and analytics processing pipelines[23]. Studies demonstrated that advertising workflows exhibit distinct execution patterns and performance requirements that differ significantly from general-purpose distributed applications. However, most research focused on individual workflow types rather than comprehensive orchestration strategies.

Recent studies have begun exploring the integration of energy considerations into workflow management, particularly in the context of green computing and sustainable operations[24]. Research has examined approaches for reducing workflow energy consumption through intelligent service placement, execution scheduling, and resource optimization[25]. However, applications to advertising systems with their unique performance and energy requirements remained limited.

The emergence of microservices architectures and containerized applications has created new opportunities and challenges for workflow orchestration in distributed systems[26]. Studies have examined distributed orchestration approaches for managing complex service dependencies while maintaining loose coupling and scalability benefits. However, most research focused on general microservice optimization rather than the specific requirements of advertising systems with their unique performance and energy constraints[27-29].

Multi-objective optimization in workflow management has been studied as researchers recognized the need to balance competing goals including execution time, resource consumption, cost, and reliability. Studies explored various approaches for incorporating multiple objectives into workflow optimization algorithms including weighted scoring functions and Pareto optimization techniques. However, most research focused on static optimization methods rather than adaptive learning approaches that could respond to changing system conditions.

3 METHODOLOGY

3.1 System Architecture and Hierarchical Control Framework

The proposed RL-based architecture addresses API workflow control through a multi-tier hierarchical structure that separates strategic workflow orchestration from tactical service optimization while maintaining coordination mechanisms that ensure system-wide efficiency and energy compliance. The system architecture incorporates high-level workflow coordinators that manage execution strategies, service selection decisions, and energy allocation policies alongside low-level API controllers that optimize individual service performance within allocated energy budgets and resource constraints.

The hierarchical control framework models workflow orchestration as a multi-level decision process where high-level decisions influence execution strategies and resource allocation while low-level decisions focus on service-specific performance optimization. High-level state representations include workflow execution progress, overall energy consumption trends, service availability indicators, and system-wide performance metrics. Low-level states encompass individual service performance characteristics, resource utilization patterns, and energy consumption measurements for specific API operations[30].

Control hierarchies are designed to balance autonomy and coordination across different abstraction levels. High-level coordinators operate on longer time horizons to make strategic decisions about workflow execution paths, service prioritization, and energy allocation strategies. Low-level controllers respond rapidly to immediate service performance requirements while respecting energy constraints and coordination signals from higher-level controllers.

3.2 Deep Deterministic Policy Gradient for Strategic Control

The high-level workflow coordinator employs DDPG algorithms to learn optimal strategic control policies for workflow orchestration including execution path selection, service prioritization, and energy allocation decisions across different workflow categories. The actor-critic architecture enables stable policy learning in continuous action spaces while handling the complex multi-objective optimization requirements typical of energy-constrained advertising systems.

Actor networks generate continuous action distributions that specify strategic workflow parameters including execution timing, service selection priorities, and energy allocation ratios across different workflow types. The neural network architecture processes high-level state information including workflow queue status, system-wide energy availability, service performance indicators, and historical execution patterns. Multiple fully connected layers with batch normalization learn complex relationships between system conditions and optimal strategic decisions.

Critic networks evaluate strategic policy performance across multiple objectives including workflow completion rates, energy consumption efficiency, and system-wide resource utilization. The multi-objective evaluation provides comprehensive feedback for policy improvement while ensuring balanced consideration of execution efficiency and energy constraints. Experience replay mechanisms store strategic decision transitions to enable stable learning across diverse workflow scenarios and system conditions[31].

3.3 Twin Delayed Deep Deterministic Policy Gradient for Tactical Control

Low-level API controllers utilize TD3 algorithms to optimize tactical service performance within individual workflow stages while respecting energy budgets and coordination signals from high-level controllers. TD3 provides enhanced stability and reduced overestimation bias compared to standard DDPG, making it particularly suitable for the complex tactical optimization challenges in energy-constrained environments.

The TD3 architecture incorporates twin critic networks that provide more stable value estimation and delayed policy updates that reduce the overestimation bias common in actor-critic algorithms. Actor networks generate tactical control actions including CPU allocation ratios, memory allocation levels, and energy consumption limits for individual API services. The delayed update mechanism ensures more stable policy learning in the dynamic advertising system environment.

Target networks and noise injection mechanisms further enhance learning stability and exploration effectiveness in the complex tactical control environment. Clipped double Q-learning provides more conservative value estimates that improve policy performance in the multi-objective optimization context. The tactical controllers learn to balance immediate service performance requirements with energy constraints while maintaining coordination with higher-level strategic decisions[32].

3.4 Energy-Aware Workflow Coordination

The energy-aware coordination framework integrates power consumption considerations into workflow control decisions through comprehensive energy modeling and constraint enforcement mechanisms. Energy budget allocation algorithms distribute available power across different workflow categories based on priority levels, execution requirements, and efficiency considerations. Dynamic energy management adapts power allocation based on real-time availability and consumption patterns.

Energy constraint enforcement mechanisms ensure that workflow execution remains within available power budgets through dynamic resource allocation and execution scheduling. Constraint violation detection algorithms monitor energy consumption patterns and trigger corrective actions when consumption approaches budget limits. Predictive energy management uses historical consumption patterns and workload forecasts to optimize energy allocation proactively.

Coordination protocols between hierarchical levels incorporate energy-aware communication that enables effective resource management while maintaining workflow performance objectives. High-level coordinators provide energy allocation targets and constraint signals to low-level controllers, while tactical controllers report energy consumption measurements and performance metrics to support strategic decision-making. The coordination framework adapts energy allocation based on changing workflow patterns and power availability conditions.

4 RESULTS AND DISCUSSION

4.1 Workflow Completion and Execution Efficiency

The RL-based hierarchical control architecture demonstrated substantial improvements in workflow completion rates when evaluated using enterprise advertising system traces spanning multiple workflow categories and operational conditions. Overall workflow completion rates increased by 41% compared to traditional orchestration methods, with particularly significant improvements for complex multi-stage workflows that benefited from intelligent execution scheduling and resource optimization. The hierarchical approach enabled strategic coordination of workflow execution while maintaining tactical optimization within individual services.

Workflow-specific performance analysis revealed consistently positive results across different execution patterns and service categories. Real-time bidding workflows achieved 47% improvement in completion rates while maintaining strict latency requirements through optimized execution sequencing and intelligent resource allocation. Content personalization workflows showed 39% better completion rates through predictive resource provisioning and adaptive execution strategies. Analytics workflows experienced 35% improvement in batch processing efficiency through intelligent workload scheduling and resource optimization.

The hierarchical control structure successfully balanced strategic workflow orchestration with tactical service optimization, preventing resource conflicts and ensuring optimal execution across all workflow categories. High-level coordinators learned to prioritize workflows based on business value and energy constraints while low-level controllers optimized individual service performance within allocated resources. The framework avoided the over-provisioning problems common in traditional approaches by dynamically adjusting resource allocation based on real-time workflow demands and energy availability.

4.2 Energy Consumption Optimization

Energy consumption reduction achieved 37% improvement compared to traditional workflow orchestration methods that focus solely on execution performance without considering power efficiency. The energy-aware optimization learned to balance computational requirements with power consumption across different workflow types and execution stages. During low-demand periods, the framework achieved up to 52% energy savings through intelligent service consolidation and dynamic resource scaling strategies.

Service-specific energy optimization showed significant benefits across different workflow components. CPU-intensive personalization services achieved 43% energy reduction through intelligent workload distribution and dynamic frequency scaling. Memory-intensive analytics operations improved energy efficiency by 34% through optimized data placement and processing scheduling. Network-intensive bidding communications showed 29% energy savings through intelligent routing and bandwidth optimization.

The multi-objective optimization successfully balanced energy efficiency with workflow performance requirements across all evaluation scenarios. Energy savings were achieved without compromising completion rates or execution latency, demonstrating the effectiveness of the energy-aware approach in identifying optimization opportunities that benefit both performance and sustainability objectives. The framework learned to exploit natural variations in workflow demand patterns to optimize energy consumption during predictable low-utilization periods.

4.3 Resource Utilization and System Efficiency

Resource utilization efficiency improved by 33% through intelligent allocation and coordination strategies that maximized hardware utilization while minimizing energy consumption. The hierarchical control architecture enabled more effective resource sharing across different workflow types while maintaining appropriate isolation and performance guarantees. Dynamic resource allocation based on real-time demand patterns and energy constraints resulted in more balanced system utilization.

CPU utilization optimization achieved 38% improvement through intelligent workload distribution that considered both performance requirements and energy efficiency objectives. Memory utilization showed 31% improvement through predictive allocation strategies that anticipated workflow resource needs based on historical patterns and real-time system conditions. Network resource utilization improved by 26% through optimized communication patterns and intelligent routing decisions.

The framework successfully eliminated resource waste and over-provisioning scenarios common in traditional workflow orchestration systems. Predictive resource allocation enabled proactive provisioning that met workflow requirements without excessive resource allocation. Dynamic scaling capabilities adapted resource allocation based on changing workflow demands while maintaining energy efficiency objectives.

4.4 Workflow Execution Latency

Average workflow execution latency decreased by 29% across all workflow categories through intelligent scheduling and resource optimization strategies that minimized execution delays and resource contention. The framework achieved particularly significant improvements for latency-sensitive bidding workflows, which experienced 34% reduction in end-to-end execution times through optimized service sequencing and dedicated resource allocation.

Latency variability reduction proved equally important for workflow reliability and predictability. The framework reduced 95th percentile latency by 41% for bidding workflows and 37% for personalization processes through consistent resource allocation and proactive performance optimization. Predictive execution planning eliminated latency spikes during resource transitions and system load variations.

The hierarchical control structure enabled effective latency optimization through coordinated resource management across all workflow stages. Strategic execution planning minimized dependencies and critical path delays while tactical service optimization ensured optimal performance within individual workflow components. The integrated approach achieved better latency performance than systems that optimize individual services independently.

4.5 Learning Efficiency and Adaptation

The hierarchical RL architecture demonstrated superior learning efficiency compared to monolithic approaches, achieving stable policy convergence within 78,000 training episodes compared to over 140,000 episodes required by non-hierarchical methods. The decomposition of complex workflow optimization into strategic and tactical control levels enabled more focused learning and reduced exploration requirements for individual agents.

Strategic control learning showed rapid convergence to effective workflow orchestration policies, with high-level coordinators achieving stable performance within 35,000 training episodes. The DDPG agents successfully learned to balance workflow prioritization with energy constraints while maintaining system-wide efficiency objectives. Experience sharing across different workflow scenarios accelerated learning convergence and improved policy generalization.

Tactical control adaptation demonstrated effective specialization within individual workflow domains. TD3 agents learned service-specific optimization strategies that maximized performance within allocated energy budgets while respecting coordination signals from strategic controllers. The enhanced stability of TD3 algorithms proved particularly valuable for tactical optimization in the dynamic advertising system environment.

Continuous learning capabilities enabled ongoing adaptation to changing workflow patterns and system conditions without requiring complete retraining. The framework successfully adapted to new workflow types, changing service performance characteristics, and evolving energy availability patterns through incremental policy updates. Online learning mechanisms maintained optimization effectiveness as advertising system requirements evolved over time.

Scalability analysis revealed robust performance across different system scales and workflow complexities. The hierarchical architecture effectively managed complexity through distributed decision-making while maintaining coordination effectiveness. Performance improvements remained consistent as workflow diversity and system scale increased, confirming the scalability advantages of the hierarchical approach.

5 CONCLUSION

The development and successful evaluation of the RL-based architecture for hierarchical control of API workflows in energy-constrained advertising systems represents a significant advancement in workflow orchestration technology for complex distributed advertising platforms. The research demonstrates that sophisticated reinforcement learning techniques can effectively address the complex challenges of balancing workflow execution efficiency with energy consumption constraints while maintaining service quality and system reliability. The architecture's achievement of 41% improvement in workflow completion rates, 37% energy consumption reduction, and 33% better resource

utilization efficiency provides compelling evidence for the practical value of hierarchical RL approaches in energy-constrained advertising system management.
The hierarchical control structure successfully addresses the scalability and coordination challenges inherent in optimizing complex workflows with diverse performance requirements and energy constraints. The combination of strategic workflow orchestration through DDPG agents with tactical service optimization using TD3 algorithms enables effective multi-level optimization while maintaining system-wide coordination. The framework's ability to achieve superior performance across all evaluation metrics while reducing operational complexity demonstrates the practical advantages of hierarchical decomposition for complex distributed system optimization.
The energy-aware optimization framework successfully integrates power consumption considerations into workflow control decisions without compromising execution performance or service reliability. The multi-objective approach identifies optimization opportunities that simultaneously improve workflow completion rates, reduce execution latency, and decrease energy consumption. The framework's ability to adapt energy allocation based on workflow demand patterns and power availability enables significant operational cost savings while maintaining strict performance requirements.
The comprehensive performance improvements across all workflow categories demonstrate the architecture's effectiveness in handling the heterogeneous requirements typical of advertising system operations. The ability to achieve 47% completion rate improvement for bidding workflows while maintaining sub-millisecond latency requirements, alongside 39% enhancement for content personalization processes, confirms the framework's capability to optimize diverse workflow characteristics within unified control strategies.
The substantial improvements in resource utilization efficiency and latency reduction provide significant operational benefits beyond pure workflow completion metrics. The 33% improvement in resource utilization enables more efficient hardware utilization while the 29% latency reduction enhances user experience and system responsiveness. These comprehensive improvements demonstrate the value of integrated optimization approaches that consider multiple system objectives simultaneously.
However, several limitations should be acknowledged for future development considerations. The framework's effectiveness depends on accurate workflow modeling and energy consumption prediction, which may be challenging in highly dynamic advertising environments with rapidly changing campaign characteristics and service requirements. The complexity of coordinating multiple hierarchical controllers while maintaining global optimization objectives may require additional mechanisms for handling conflicting goals or resource constraints during peak demand periods.
Future research should explore the integration of additional optimization objectives including service reliability, data privacy compliance, and regulatory requirements into the hierarchical control framework. The incorporation of advanced prediction techniques including real-time campaign analysis, user behavior forecasting, and system performance modeling could improve control effectiveness through better anticipation of workflow demand patterns and energy requirements.
The development of specialized modules for emerging advertising technologies including programmatic creative optimization, cross-device attribution workflows, and privacy-preserving analytics could extend the architecture's applicability to next-generation advertising platforms. Integration with edge computing infrastructure and distributed workflow execution environments could create comprehensive solutions for globally distributed advertising system architectures.
This research contributes to the broader understanding of how hierarchical reinforcement learning can address complex distributed system control challenges while incorporating energy efficiency as a first-class optimization objective. The architecture demonstrates that advanced machine learning techniques can successfully balance multiple competing goals including performance, sustainability, and resource efficiency while adapting to dynamic operational conditions.
The implications extend beyond advertising systems to other domains requiring sophisticated workflow orchestration across distributed infrastructure with energy constraints. The framework's approach to balancing strategic coordination with tactical optimization while incorporating energy considerations offers valuable insights for developing intelligent control solutions across various distributed computing environments. As workflow complexity continues to increase and energy efficiency becomes increasingly critical, hierarchical RL architectures that integrate performance and environmental objectives will likely play essential roles in sustainable distributed system management and optimization.

COMPETING INTERESTS

The authors have no relevant financial or non-financial interests to disclose.

REFERENCES

[1] Anitha K, Anitha A, Preetha S, et al. Seamless data flow: Constructing end-to-end data pipelines for real-time marketing analytics. In Data Engineering for Data-driven Marketing, 2025: 73-90.
[2] Xing S, Wang Y. Proactive data placement in heterogeneous storage systems via predictive multi-objective reinforcement learning. IEEE Access, 2025.
[3] Maddukuri N. Workflow optimization for mobile computing. International Journal of Science and Research Archive, 2025, 14(01): 340-346.

[4] Dakić V. Machine learning and dynamic parameters-based system architecture for automatic workload scheduling in heterogenous high-performance computing. Doctoral dissertation, University of Zagreb, Faculty of Electrical Engineering and Computing, Department of Control and Computer Engineering, 2025.
[5] Wang Z. Analysis of the shared energy storage business model for building clusters in commercial pedestrian blocks. Economic Management and Practice, 2025, 3(3): 1-21. DOI: 10.61784/emp2001.
[6] Menaka M, Kumar K S. Workflow scheduling in cloud environment—Challenges, tools, limitations & methodologies: A review. Measurement: Sensors, 2022, 24: 100436.
[7] Sheta S V. A comprehensive analysis of real-time data processing architectures for high-throughput applications, 2022.
[8] Choudhary A, Govil M C, Singh G, et al. Energy-aware scientific workflow scheduling in cloud environment. Cluster Computing, 2022, 25(6): 3845-3874.
[9] Ramachandran K K. Optimizing IT performance: A comprehensive analysis of resource efficiency. International Journal of Marketing and Human Resource Management (IJMHRM), 2023, 14(3): 12-29.
[10] Peddinti S R, Pandey B K, Tanikonda A, et al. Optimizing microservice orchestration using reinforcement learning for enhanced system efficiency. Distributed Learning and Broad Applications in Scientific Research| Annual, 2021, 7.
[11] Reddy P V, Reddy K G. An energy efficient RL based workflow scheduling in cloud computing. Expert Systems with Applications, 2023, 234: 121038.
[12] Ji E, Wang Y, Xing S, et al. Hierarchical reinforcement learning for energy-efficient API traffic optimization in large-scale advertising systems. IEEE Access, 2025.
[13] Verma R, Rane D, Jaswal S. Service-oriented computing paradigms: Opportunities and challenges. Soft Computing Principles and Integration for Real-Time Service-Oriented Computing, 2024: 154-173.
[14] Barika M, Garg S, Zomaya A Y, et al. Orchestrating big data analysis workflows in the cloud: Research challenges, survey, and future directions. ACM Computing Surveys (CSUR), 2019, 52(5): 1-41.
[15] Rajesh S C, Goel L. Architecting distributed systems for real-time data processing in multi-cloud environments. International Journal of Emerging Technology and Innovative Research, 2025, 12: b623-b640.
[16] Kocot B, Czarnul P, Proficz J. Energy-aware scheduling for high-performance computing systems: A survey. Energies, 2023, 16(2): 890.
[17] Hathwar D K, Bharadwaj S R, Basha S M. Power-aware virtualization: Dynamic voltage frequency scaling insights and communication-aware request stacking. In Computational Intelligence for Green Cloud Computing and Digital Waste Management, 2024: 84-108. IGI Global Scientific Publishing.
[18] Lu Y, Chen C, Mei Y. Evaluation of the vertical synergy of science and technology financial policies from a structural-functional perspective: Based on the experience of Guangdong Province and cities. Economic Management and Practice, 2025, 3(3): 22-34. DOI: 10.61784/emp2002.
[19] Chen Y, Chan W H, Su E L M, et al. Multi-objective optimization for smart cities: A systematic review of algorithms, challenges, and future directions. PeerJ Computer Science, 2025, 11: e3042.
[20] Nguyen T T, Nguyen N D, Nahavandi S. Deep reinforcement learning for multiagent systems: A review of challenges, solutions, and applications. IEEE Transactions on Cybernetics, 2020, 50(9): 3826-3839.
[21] Qadeer A, Lee M J. Deep-deterministic policy gradient based multi-resource allocation in edge-cloud system: A distributed approach. IEEE Access, 2023, 11: 20381-20398.
[22] Pateria S, Subagdja B, Tan AH, et al. Hierarchical reinforcement learning: A comprehensive survey. ACM Computing Surveys (CSUR), 2021, 54(5): 1-35.
[23] Adepoju AH, Eweje A, Collins A, et al. Automated offer creation pipelines: An innovative approach to improving publishing timelines in digital media platforms. International Journal of Multidisciplinary Research and Growth Evaluation, 2024, 5(6): 1475-1489.
[24] Buyya R, Ilager S, Arroba P. Energy-efficiency and sustainability in new generation cloud computing: a vision and directions for integrated management of data centre resources and workloads. Software: Practice and Experience, 2024, 54(1): 24-38.
[25] LuoLe Z, ZuChang Z, XiaoMin L, et al. The dual effects of a country's overseas patent network layout on its export: scale-up or quality improvement. Social Science and Management, 2025, 2(2): 12-29. DOI: 10.61784/ssm3046.
[26] Verma P, Maurya AK, Yadav RS. A survey on energy-efficient workflow scheduling algorithms in cloud computing. Software: Practice and Experience, 2024, 54(5): 637-682.
[27] Usman M, Ferlin S, Brunstrom A, et al. A survey on observability of distributed edge & container-based microservices. IEEE Access, 2022, 10: 86904-86919.
[28] Cao J, Zheng W, Ge Y, et al. DriftShield: Autonomous fraud detection via actor-critic reinforcement learning with dynamic feature reweighting. IEEE Open Journal of the Computer Society, 2025.
[29] Araújo G, Barbosa V, Lima LN, et al. Energy consumption in microservices architectures: a systematic literature review. IEEE Access, 2024, 12: 186710-186729.
[30] Mai N, Cao W. Personalized learning and adaptive systems: AI-driven educational innovation and student outcome enhancement. International Journal of Education and Humanities, 2025.
[31] Wang M, Zhang X, Yang Y, et al. Explainable machine learning in risk management: balancing accuracy and interpretability. Journal of Financial Risk Management, 2025, 14(3): 185-198.

[32] Cao W, Mai N, Liu W. Adaptive knowledge assessment via symmetric hierarchical Bayesian neural networks with graph symmetry-aware concept dependencies. Symmetry, 2025.