International Journal of Advance Scientific Research
(ISSN – 2750-1396)
VOLUME 06 ISSUE 01 Pages: 150-167
OCLC – 1368736135

Journal Website: http://sciencebring.com/index.php/ijasr

Research Article

Deep Reinforcement Learning Oriented Queuing And Swarm Intelligence Framework For Adaptive Task Scheduling In End Edge Cloud Computing Environments

Submission Date: December 19, 2025, **Accepted Date:** January 12, 2026,
Published Date: January 31, 2026

Dr. Leonhard Weiss
Department of Computer Science, KU Leuven, Belgium

Abstract

The unprecedented growth of data intensive applications, cyber physical systems, and latency sensitive services has fundamentally reshaped the operational expectations of cloud, edge, and fog computing ecosystems. Modern computing environments are no longer limited to static cloud data centers but have evolved into complex end edge cloud architectures in which computation is dynamically distributed across geographically dispersed and resource heterogeneous nodes. Within this paradigm, task scheduling emerges as a central and persistent challenge, as the allocation of computational workloads to available resources directly determines system throughput, energy efficiency, delay, fairness, and economic sustainability. Traditional heuristic and metaheuristic approaches, while historically successful, increasingly struggle to cope with the stochasticity, scale, and heterogeneity that define contemporary distributed computing infrastructures. Consequently, there is a growing scholarly consensus that adaptive learning-based scheduling mechanisms are required to enable real time decision making under uncertainty.

This study develops and theorizes a novel reinforcement learning driven queuing and optimization framework for dynamic task scheduling across end edge cloud systems. Drawing upon foundational work in reinforcement learning theory, evolutionary computation, and queuing based performance modeling, the research constructs an integrated conceptual architecture that fuses deep Q learning with optimal queuing dynamics and swarm intelligence. In particular, the scheduling logic is informed by the theoretical

insights of Kanikanti et al. (2025), who demonstrate that deep Q learning combined with optimal queuing theory can substantially improve dynamic task allocation in cloud computing environments. Their work provides a critical empirical and theoretical anchor for this study, which extends the paradigm into a broader end edge cloud orchestration context.

The article systematically analyzes how reinforcement learning agents can learn optimal task placement policies by observing system states such as queue length, processing capacity, latency, and energy consumption. These agents are embedded within a hybrid optimization framework that incorporates swarm-based search processes, genetic adaptation, and market-oriented scheduling strategies. The proposed model is not presented as a mathematical algorithmic artifact but as a deeply theorized scheduling ecosystem grounded in interdisciplinary research traditions. By integrating queuing theory with deep reinforcement learning, the framework accounts for both the temporal structure of task arrivals and the strategic adaptation of scheduling policies over time.

The results of this study, interpreted through extensive comparative reasoning against the existing literature, indicate that learning driven queuing-based schedulers can outperform static and metaheuristic approaches in terms of adaptability, delay minimization, and system robustness. Furthermore, the discussion situates these findings within broader debates about the future of intelligent computing infrastructures, arguing that end edge cloud systems represent not merely a technological shift but a paradigmatic transformation in how computational resources are conceptualized and governed. The article concludes by identifying key theoretical and practical implications for next generation scheduling systems and outlining a future research agenda centered on autonomous, self-optimizing distributed intelligence.

Keywords

Deep reinforcement learning, cloud task scheduling, end edge cloud computing, queuing theory, swarm intelligence, adaptive resource management.

Introduction

The evolution of distributed computing over the last three decades has been driven by an escalating demand for scalable, flexible, and cost-efficient computational resources. What began as centralized high-performance computing clusters has gradually transformed into a globally interconnected fabric of cloud data centers, edge nodes, and fog based infrastructures that collectively support applications ranging from scientific simulations to real time Internet of Things services. Within this evolving ecosystem, task scheduling has remained a foundational problem, as the manner in which computational jobs are assigned to available resources fundamentally shapes system performance, energy efficiency, and user experience. Despite decades of research, the problem of optimal task scheduling remains unresolved, largely because modern computing environments are characterized by extreme heterogeneity, stochastic workloads, and dynamically changing resource availability (Verma and Kaushal, 2017).

Historically, task scheduling in distributed systems has been addressed through heuristic and metaheuristic techniques that seek to minimize metrics such as makespan, delay, or energy consumption. Early approaches such as genetic algorithms, ant colony optimization, and particle swarm optimization were inspired by biological and social processes and demonstrated remarkable effectiveness in static or moderately dynamic environments (Dorigo et al., 1996; Eberhart and Kennedy, 1995; Holland, 1975). These methods provided a powerful toolkit for exploring large search spaces and identifying near optimal solutions to complex combinatorial problems. In cloud computing, such techniques have been widely applied to workflow scheduling, virtual machine placement, and resource provisioning, with numerous studies reporting improvements in throughput and cost efficiency (Raju et al., 2013; Khalili and Babamir, 2015; Huang et al., 2020).

However, the limitations of purely heuristic based scheduling have become increasingly evident as computing environments have grown more dynamic and decentralized. In end edge cloud architectures, tasks may arrive unpredictably, network conditions may fluctuate, and resources may join or leave the system without notice. Under such conditions, a scheduling strategy that is optimized for a particular workload distribution may rapidly become suboptimal when the environment changes (Ren et al., 2019). This challenge has motivated a shift toward learning based scheduling paradigms, in which the scheduler continuously adapts its policy based on observed system feedback.

Reinforcement learning has emerged as a particularly promising framework for adaptive task scheduling because it formalizes the problem as a sequential decision-making process under uncertainty. In reinforcement learning, an agent interacts with an environment by selecting actions, observing the resulting state transitions and rewards, and gradually learning a policy that maximizes cumulative long-term reward (Sutton and Barto, 2018). This paradigm is well suited to scheduling because the scheduler must repeatedly decide where to place incoming tasks while considering both immediate performance and future system states. Foundational work on Q learning by Watkins and Dayan (1992) and subsequent theoretical refinements have provided a rigorous basis for applying reinforcement learning to a wide range of optimization problems, including those in networking and distributed computing.

Within the domain of task scheduling, early reinforcement learning approaches focused on relatively simple environments such as wireless sensor networks, where Q learning was used to balance energy consumption and task delay (Wei et al., 2017; Khan et al., 2017). These studies demonstrated that learning based schedulers could outperform static heuristics by dynamically adapting to changes in network conditions. Subsequent research extended these ideas to more complex settings by integrating reinforcement learning with machine learning classifiers and shared value functions, thereby improving convergence speed and stability (Wei et al., 2019). Despite these advances, most reinforcement learning based scheduling systems remained limited in scale and scope, often focusing on specific subsystems rather than holistic cloud or edge cloud environments.

The recent emergence of deep reinforcement learning has further expanded the potential of learning based scheduling. By combining neural networks with Q learning, deep reinforcement learning enables agents to approximate value functions in high dimensional state spaces, making it feasible to handle the complexity of large-scale cloud environments (Sutton and Barto, 2018). This paradigm has been successfully applied to delay oriented scheduling in satellite ground integrated networks and Internet of Things systems, demonstrating significant reductions in latency and congestion (Zhou et al., 2021). These developments suggest that deep reinforcement learning can provide the adaptive intelligence required to manage the intricate dynamics of end edge cloud systems.

A particularly influential contribution in this context is the work of Kanikanti et al. (2025), who propose a deep Q learning driven dynamic optimal task scheduling framework based on optimal queuing theory. Their study represents a critical synthesis of two traditionally separate strands of research: reinforcement learning and queuing based performance modeling. By embedding a deep Q learning agent within a queuing theoretic framework, Kanikanti et al. (2025) demonstrate that the scheduler can explicitly account for queue lengths, service rates, and waiting times while learning optimal task allocation policies. This approach not only improves delay and throughput but also enhances system stability under bursty workloads. The significance of this work lies in its recognition that learning based scheduling must be grounded in rigorous performance models if it is to be both effective and interpretable.

Despite these advances, a substantial gap remains in the literature regarding the integration of deep reinforcement learning and queuing theory within end edge cloud architectures. Most existing studies focus either on centralized cloud data centers or on isolated edge or fog nodes, without fully addressing the hierarchical and distributed nature of modern computing ecosystems (Guevara and da Fonseca, 2021; Ren et al., 2020). End edge cloud computing, which encompasses cloud data centers, edge servers, and end devices within a unified orchestration framework, introduces new layers of complexity related to mobility, heterogeneity, and cross domain coordination (Duan et al., 2023). In such environments, tasks may migrate across layers based on latency constraints, energy availability, and economic considerations, rendering traditional scheduling models insufficient.

International Journal of Advance Scientific Research
(ISSN – 2750-1396)
VOLUME 06 ISSUE 01 Pages: 150-167
OCLC – 1368736135

Moreover, while deep reinforcement learning offers powerful tools for adaptive decision making, it is not without limitations. Issues such as slow convergence, instability, and sensitivity to reward design have been widely documented in the reinforcement learning literature (Kaelbling et al., 1996). In highly dynamic environments, a purely learning based scheduler may struggle to keep pace with rapid changes unless it is supported by complementary optimization mechanisms. This observation has led researchers to explore hybrid approaches that combine reinforcement learning with evolutionary and swarm-based algorithms, thereby leveraging both long term learning and short-term search capabilities (Song et al., 2020; Meshkati and Safi Esfahani, 2019).

Against this backdrop, the present study seeks to develop a comprehensive theoretical and methodological framework for deep reinforcement learning driven queuing-based task scheduling in end edge cloud environments. Building on the foundational insights of Kanikanti et al. (2025), the research extends their model to a broader and more heterogeneous computing landscape by incorporating swarm intelligence, genetic adaptation, and market-oriented scheduling principles. The central argument advanced in this article is that optimal task scheduling in contemporary distributed systems cannot be achieved through any single paradigm in isolation but requires an integrative approach that synthesizes learning, queuing, and evolutionary optimization.

The literature on swarm intelligence provides a valuable theoretical foundation for this synthesis. Particle swarm optimization, ant colony optimization, and cat swarm optimization are all inspired by collective behavior in natural systems and have been shown to effectively explore large and complex search spaces (Eberhart and Kennedy, 1995; Dorigo et al., 1996; Gabi et al., 2019). In the context of cloud scheduling, these methods have been used to minimize makespan, balance load, and reduce energy consumption by dynamically adjusting task assignments based on global and local information (Zhou et al., 2020; Sharma and Garg, 2020). However, swarm-based algorithms typically operate on a population of candidate solutions rather than learning a persistent policy, which limits their ability to adapt over time to evolving workloads.

Reinforcement learning, by contrast, is inherently policy oriented and thus well suited to environments in which decisions must be made sequentially and adaptively. By integrating swarm-based search into a reinforcement learning framework, it becomes possible to combine the exploratory power of population-based optimization with the exploitative strength of learned policies. This hybridization has been explored in various forms, such as nested particle swarm optimization for workflow scheduling (Song et al., 2020) and energy aware scheduling using particle swarm and artificial bee colony algorithms (Meshkati and Safi Esfahani, 2019). These studies suggest that hybrid approaches can achieve superior performance by mitigating the weaknesses of individual methods.

International Journal of Advance Scientific Research
(ISSN – 2750-1396)
VOLUME 06 ISSUE 01 Pages: 150-167
OCLC – 1368736135

Another important strand of the literature concerns market oriented and hierarchical scheduling strategies, which model cloud resource allocation as an economic system in which users and providers interact through pricing and bidding mechanisms (Wu et al., 2013). Such approaches are particularly relevant in end edge cloud environments, where resources are distributed across multiple administrative domains and economic incentives play a crucial role in shaping system behavior. By incorporating market-oriented principles into a reinforcement learning based scheduler, it becomes possible to align individual scheduling decisions with broader system level objectives such as cost efficiency and fairness.

The convergence of these diverse research traditions raises profound theoretical questions about the nature of intelligence in distributed computing systems. From one perspective, the scheduler can be viewed as a rational agent that seeks to maximize a predefined reward function, as formalized in reinforcement learning theory (Sutton and Barto, 2018). From another perspective, the scheduler can be seen as an emergent phenomenon arising from the interactions of multiple optimization processes operating at different scales, as suggested by swarm intelligence and market-oriented models (Dorigo et al., 1996; Wu et al., 2013). Reconciling these perspectives requires a conceptual framework that acknowledges both the intentionality of learning agents and the self organizing dynamics of complex systems.

This article contributes to this intellectual endeavor by proposing a deeply integrated model of task scheduling that situates deep reinforcement learning within a broader ecosystem of queuing, swarm, and economic mechanisms. The theoretical foundation of the model is grounded in the recognition that queues are not merely passive buffers but active components of system dynamics that mediate the flow of tasks and resources (Kanikanti et al., 2025). By explicitly modeling queues within the learning process, the scheduler gains a more nuanced understanding of system congestion and delay, enabling more informed and proactive decision making.

In addition to its theoretical contributions, this study addresses a significant practical gap in the existing literature. While numerous algorithms have been proposed for cloud and edge scheduling, few have been articulated as comprehensive frameworks that can be systematically adapted to different application domains and infrastructure configurations. The proposed model is designed to be conceptually modular, allowing practitioners to tailor the reinforcement learning, queuing, and swarm components to the specific characteristics of their systems. This flexibility is particularly important in end edge cloud environments, where no two deployments are identical and where rigid one size fits all solutions are unlikely to succeed (Duan et al., 2023).

The remainder of this article is organized as follows. The methodology section provides a detailed and text-based exposition of the

International Journal of Advance Scientific Research
(ISSN – 2750-1396)
VOLUME 06 ISSUE 01 Pages: 150-167
OCLC – 1368736135

proposed scheduling framework, including its underlying assumptions, design rationale, and limitations. The results section offers a descriptive and interpretive analysis of the expected performance of the framework based on comparative reasoning with existing studies. The discussion section situates these findings within broader scholarly debates and explores their implications for the future of intelligent distributed computing. The conclusion synthesizes the key insights and outlines directions for future research. Throughout the article, every major claim is grounded in the existing literature, with particular emphasis on the integrative contributions of Kanikanti et al. (2025) to the field of learning based task scheduling.

Methodology

The methodological foundation of this study is rooted in the conceptual integration of deep reinforcement learning, queuing theory, and evolutionary swarm intelligence within an end edge cloud computing context. Rather than presenting a purely algorithmic or mathematical model, the methodology is articulated as a coherent and theoretically grounded framework that reflects the complex socio technical nature of modern distributed computing environments. This approach is consistent with the growing recognition that scheduling in cloud and edge systems is not merely a computational problem but a systemic phenomenon shaped by interactions between workloads, resources, networks, and adaptive decision makers (Ren et al., 2019; Duan et al., 2023).

At the core of the proposed framework lies the reinforcement learning agent, which functions as the primary decision-making entity responsible for assigning incoming tasks to available computational resources. This agent is modeled according to the principles of deep Q learning, in which a neural network approximates the Q value function that maps state action pairs to expected cumulative rewards (Sutton and Barto, 2018). The choice of deep Q learning is motivated by the high dimensionality and continuous variability of end edge cloud environments, which render traditional tabular Q learning impractical (Watkins and Dayan, 1992).

The state representation used by the agent encompasses a rich set of system variables, including but not limited to queue lengths at different nodes, estimated service rates, network latency, energy consumption, and current workload distribution. By incorporating these variables into the state, the agent is able to perceive not only the immediate availability of resources but also the temporal dynamics of congestion and processing, as emphasized in queuing-based scheduling models (Kanikanti et al., 2025). This explicit representation of queues distinguishes the framework from many existing reinforcement learning schedulers, which often treat system state in a more abstract or aggregated manner (Wei et al., 2017).

The action space of the agent consists of possible task placement decisions, such as assigning a task

International Journal of Advance Scientific Research
(ISSN – 2750-1396)
VOLUME 06 ISSUE 01 Pages: 150-167
OCLC – 1368736135

to a particular cloud data center, edge server, or fog node. In an end edge cloud environment, these options may span multiple layers of the infrastructure hierarchy, reflecting the distributed and heterogeneous nature of available resources (Guevara and da Fonseca, 2021). The reward function is designed to capture multiple system objectives, including task completion time, energy efficiency, and load balancing. In line with the multi objective scheduling literature, the reward is conceptualized as a weighted combination of these factors, allowing system designers to prioritize different performance criteria as needed (Verma and Kaushal, 2017; Meshkati and Safi Esfahani, 2019).

A critical methodological innovation of this framework is the integration of optimal queuing theory into the reinforcement learning process. Building on the insights of Kanikanti et al. (2025), the framework treats queues as first-class components of the environment rather than as incidental byproducts of scheduling decisions. This means that the agent not only observes queue lengths but also implicitly learns the relationship between its actions and the evolution of queues over time. In effect, the agent internalizes a dynamic model of system congestion, enabling it to anticipate the downstream consequences of task placement decisions. This anticipatory capability is essential for avoiding phenomena such as queue buildup and cascading delays, which can severely degrade system performance in high load scenarios (Shafi et al., 2020).

To enhance the adaptability and robustness of the learning process, the framework incorporates swarm intelligence mechanisms as a complementary optimization layer. Specifically, a population of candidate scheduling policies is maintained and periodically evolved using particle swarm and genetic inspired operators. This population-based approach allows the system to explore a diverse set of policies in parallel, reducing the risk of premature convergence to suboptimal solutions (Eberhart and Kennedy, 1995; Holland, 1975). The best performing policies in the population are used to initialize or guide the deep Q learning agent, thereby providing a form of informed exploration that accelerates learning and improves stability (Song et al., 2020).

The rationale for integrating swarm intelligence with reinforcement learning is grounded in the recognition that learning based methods and evolutionary methods have complementary strengths and weaknesses. Reinforcement learning excels at exploiting experience to refine a policy over time, but it may struggle in the early stages when little data is available or when the reward landscape is highly nonlinear (Kaelbling et al., 1996). Swarm and genetic algorithms, by contrast, are well suited to global exploration but lack the capacity to accumulate long term experience in the form of a persistent policy (Zhou et al., 2020). By combining these approaches, the framework seeks to achieve both broad exploration and deep exploitation, thereby enhancing overall scheduling performance.

International Journal of Advance Scientific Research
(ISSN – 2750-1396)
VOLUME 06 ISSUE 01 Pages: 150-167
OCLC – 1368736135

Another methodological dimension of the framework is its incorporation of market-oriented scheduling principles. In end edge cloud environments, resources are often owned and operated by different stakeholders, each with their own cost structures and performance incentives (Wu et al., 2013). The framework models this reality by associating economic attributes such as pricing and bidding with each resource, and by including these attributes in the state and reward representations observed by the reinforcement learning agent. This allows the agent to make scheduling decisions that are not only technically efficient but also economically rational, aligning with the broader objectives of cloud service providers and users.

The methodological design also accounts for the hierarchical structure of end edge cloud systems. Tasks may be processed at the edge for low latency or offloaded to the cloud for greater computational power, depending on their requirements and current system conditions (Ren et al., 2019). The framework therefore supports multi-level scheduling, in which local agents at the edge make preliminary decisions that are coordinated with higher level agents in the cloud. This hierarchical reinforcement learning paradigm reflects the distributed nature of control in end edge cloud environments and has been shown to improve scalability and responsiveness in similar contexts (Zhou et al., 2021).

Despite its comprehensive design, the framework is subject to several limitations that must be acknowledged. One major challenge is the complexity of training deep reinforcement learning agents in highly dynamic environments. The convergence of Q learning can be slow and unstable when the state and action spaces are large and when the reward function is multi objective (Sutton and Barto, 2018). Although the integration of swarm intelligence is intended to mitigate these issues, it also introduces additional computational overhead and design complexity (Meshkati and Safi Esfahani, 2019).

Another limitation relates to the accuracy of the state representation. While the framework assumes that variables such as queue lengths and service rates can be observed or estimated, in practice these measurements may be noisy or delayed, particularly in geographically distributed systems (Pinciroli et al., 2020). Such uncertainty can degrade the quality of the learned policy and lead to suboptimal scheduling decisions. Addressing this challenge may require the incorporation of predictive analytics and real time clustering techniques, as suggested by Zhang and He (2019).

Furthermore, the economic and organizational aspects of end edge cloud computing introduce constraints that are difficult to model explicitly. Market oriented scheduling assumes that pricing and bidding mechanisms are transparent and stable, but in reality these mechanisms may be subject to strategic manipulation and regulatory constraints (Wu et al., 2013). As a result, the economic component of the framework should be viewed as an idealized abstraction rather than a fully realistic model.

International Journal of Advance Scientific Research
(ISSN – 2750-1396)
VOLUME 06 ISSUE 01 Pages: 150-167
OCLC – 1368736135

In summary, the methodology proposed in this study represents a theoretically grounded and integrative approach to task scheduling in end edge cloud environments. By combining deep Q learning, optimal queuing theory, swarm intelligence, and market-oriented principles, the framework seeks to address the multifaceted challenges of modern distributed computing. Its design is informed by and builds upon a rich body of prior research, including the seminal contributions of Kanikanti et al. (2025), and provides a foundation for the descriptive and interpretive analysis presented in the subsequent sections.

Results

The results of the proposed deep reinforcement learning driven queuing and swarm intelligence framework are best understood through a comparative and interpretive analysis grounded in the existing literature on cloud and edge task scheduling. Because the framework is conceptual and theoretically articulated rather than empirically simulated in this study, the results are presented as a synthesis of expected performance outcomes based on the established behavior of its constituent components and their documented impacts in related research. This approach is consistent with methodological traditions in systems theory and computational intelligence, where theoretical integration is often evaluated through its coherence with and extension of prior empirical findings (Verma and Kaushal, 2017).

One of the most significant expected outcomes of the framework is a reduction in task completion time and system wide delay. The integration of optimal queuing theory into the deep reinforcement learning process, as advocated by Kanikanti et al. (2025), ensures that the scheduler explicitly accounts for waiting times and service rates when making decisions. In traditional heuristic based schedulers, tasks may be assigned to resources that appear available but are in fact subject to long queues, leading to hidden delays that degrade overall performance (Shafi et al., 2020). By contrast, a queuing aware learning agent is expected to avoid such pitfalls by internalizing the temporal structure of system congestion.

Comparative reasoning with prior studies supports this expectation. For example, Zhang and He (2019) demonstrate that real time clustering and traffic prediction can significantly improve scheduling decisions by anticipating future load patterns. Similarly, Pinciroli et al. (2020) show that predictive analytics can enhance resource management for burstable cloud instances. The proposed framework extends these ideas by embedding predictive and queuing information directly into the learning process, thereby enabling the agent to not only react to current conditions but also anticipate future congestion. This anticipatory capability is a key driver of delay reduction in learning-based scheduling systems (Zhou et al., 2021).

Another important result pertains to system adaptability and robustness under dynamic workloads. Traditional metaheuristic algorithms

International Journal of Advance Scientific Research
(ISSN – 2750-1396)
VOLUME 06 ISSUE 01 Pages: 150-167
OCLC – 1368736135

such as particle swarm optimization and genetic algorithms have been shown to perform well in static or slowly changing environments but often struggle when workload patterns shift rapidly (Huang et al., 2020; Zhou et al., 2020). Reinforcement learning, by contrast, continuously updates its policy based on new experience, making it inherently adaptive (Sutton and Barto, 2018). By combining reinforcement learning with swarm-based exploration, the proposed framework is expected to exhibit both rapid adaptation and sustained performance across a wide range of operating conditions.

This dual capability is particularly important in end edge cloud environments, where tasks may migrate between cloud, fog, and edge layers in response to latency and energy constraints (Ren et al., 2019). Studies of hierarchical and distributed scheduling strategies indicate that multi-level coordination can significantly improve system resilience and scalability (Wu et al., 2013; Zhou et al., 2021). The proposed framework builds on these insights by enabling local and global learning agents to coordinate their decisions through shared state representations and reward structures. As a result, the system is expected to maintain stable performance even when individual nodes experience failures or overloads.

Energy efficiency is another domain in which the framework is anticipated to yield positive results. Energy aware scheduling has been a major focus of cloud computing research, with numerous studies demonstrating that intelligent task placement can reduce power consumption without sacrificing performance (Meshkati and Safi Esfahani, 2019; Khan et al., 2017). By including energy consumption as a component of the reward function, the reinforcement learning agent is incentivized to favor energy efficient resources and to balance load in a way that minimizes unnecessary power usage. The queuing-based perspective further supports this goal by preventing the overloading of specific nodes, which can lead to inefficient energy utilization due to thermal and performance constraints (Kanikanti et al., 2025).

The incorporation of market-oriented scheduling principles also leads to expected improvements in cost efficiency and fairness. Wu et al. (2013) argue that hierarchical and market-based strategies can align individual scheduling decisions with broader economic objectives, thereby enhancing the sustainability of cloud services. By allowing the reinforcement learning agent to observe and respond to pricing signals, the proposed framework enables it to trade off performance and cost in a rational manner. This capability is particularly relevant in end edge cloud systems, where resources may be owned by different providers and subject to varying pricing models (Duan et al., 2023).

Another noteworthy result is the anticipated improvement in scalability. As cloud and edge systems continue to grow in size and complexity, centralized scheduling approaches become increasingly impractical (Guevara and da Fonseca, 2021). The hierarchical reinforcement learning structure of the proposed framework distributes decision making across multiple agents, each responsible for a subset of the system. This distribution reduces computational and communication overhead while preserving global coordination through shared learning objectives. Prior studies on distributed and hierarchical control support the conclusion that such architectures can scale more effectively than monolithic schedulers (Ren et al., 2020).

The framework is also expected to exhibit improved stability and convergence properties compared to purely reinforcement learning based schedulers. One of the challenges of deep Q learning is its sensitivity to reward design and hyperparameter tuning, which can lead to oscillatory or divergent behavior (Sutton and Barto, 2018). By seeding the learning process with swarm-based policy candidates and periodically refreshing the population through evolutionary operators, the framework introduces an additional layer of stability that guides the agent toward promising regions of the policy space (Song et al., 2020). This hybrid approach is consistent with findings in the evolutionary computing literature, which suggest that combining learning and evolution can yield more reliable optimization outcomes (Sulaiman et al., 2021).

From a broader systems perspective, the results of the framework can be interpreted as a move toward self-organizing and autonomic computing infrastructures. By continuously learning from system feedback and adjusting scheduling policies accordingly, the framework embodies the principles of autonomic computing, in which systems manage themselves with minimal human intervention (Ren et al., 2019). The explicit modeling of queues, energy, and economic factors ensures that this self-management is grounded in realistic operational constraints rather than abstract optimization objectives.

In summary, the results of the proposed framework, as inferred from its theoretical underpinnings and alignment with the existing literature, suggest that deep reinforcement learning driven queuing and swarm intelligence can significantly enhance the performance, adaptability, and sustainability of end edge cloud task scheduling. These results are not presented as isolated claims but as a coherent extension of a rich body of research, including the influential work of Kanikanti et al. (2025), which provides a critical empirical and conceptual foundation for the framework.

Discussion

The discussion of the proposed deep reinforcement learning driven queuing and swarm intelligence framework must be situated within the broader intellectual landscape of distributed computing, optimization theory, and artificial intelligence. At its core, the framework represents a synthesis of multiple paradigms that have historically evolved along parallel but often disconnected trajectories. By bringing together reinforcement learning, queuing theory, evolutionary computation, and market-oriented scheduling, the framework challenges conventional assumptions about how task scheduling should be conceptualized and

International Journal of Advance Scientific Research
(ISSN – 2750-1396)
VOLUME 06 ISSUE 01 Pages: 150-167
OCLC – 1368736135

implemented in end edge cloud environments (Ren et al., 2019; Duan et al., 2023).

One of the most profound theoretical implications of the framework is its reconceptualization of queues as dynamic and informative elements of the learning process rather than as passive buffers. In classical queuing theory, queues are modeled as stochastic processes that characterize system performance in terms of waiting times and service rates. While this perspective has been invaluable for capacity planning and performance analysis, it has often been divorced from the decision-making processes that generate those queues (Shafi et al., 2020). The work of Kanikanti et al. (2025) bridges this gap by embedding queuing dynamics directly into a deep Q learning scheduler, thereby transforming queues into actionable signals that guide adaptive decision making.

This integration has far reaching implications for the design of intelligent systems. By learning how its actions affect queue evolution, the scheduler effectively acquires a form of temporal foresight, enabling it to avoid myopic decisions that might yield short term gains at the expense of long term stability. This capability resonates with broader debates in reinforcement learning about the importance of modeling environment dynamics and delayed rewards (Sutton and Barto, 2018). In the context of cloud and edge computing, where congestion and overload can propagate across the network, such foresight is essential for maintaining quality of service.

The hybridization of reinforcement learning with swarm intelligence further enriches this theoretical landscape. Swarm based algorithms are often celebrated for their ability to explore complex search spaces through simple local interactions, giving rise to emergent global behavior (Dorigo et al., 1996; Eberhart and Kennedy, 1995). Reinforcement learning, by contrast, is typically framed in terms of a single agent optimizing a global reward function. By combining these paradigms, the framework blurs the distinction between centralized and decentralized intelligence, suggesting that effective scheduling may require both individual learning and collective adaptation (Song et al., 2020).

This synthesis also addresses a longstanding tension in the optimization literature between exploration and exploitation. Reinforcement learning agents must balance the need to try new actions with the desire to exploit known good strategies, a trade off that becomes increasingly difficult in large and dynamic environments (Kaelbling et al., 1996). Swarm and genetic algorithms offer a complementary mechanism for exploration by maintaining a diverse population of candidate solutions. The framework leverages this diversity to guide and stabilize the learning process, thereby enhancing convergence and reducing the risk of suboptimal equilibria (Sulaiman et al., 2021).

From a practical standpoint, the framework has significant implications for the management of end edge cloud infrastructures. The shift toward distributed and hierarchical computing

International Journal of Advance Scientific Research
(ISSN – 2750-1396)
VOLUME 06 ISSUE 01 Pages: 150-167
OCLC – 1368736135

architectures has created new challenges for resource orchestration, as tasks must be dynamically allocated across layers with different performance, energy, and cost characteristics (Ren et al., 2019; Guevara and da Fonseca, 2021). Traditional centralized schedulers are ill equipped to handle this complexity, while purely local heuristics lack the global perspective required for optimal coordination. The hierarchical reinforcement learning approach adopted in the framework provides a middle ground by enabling local decision making that is aligned with global objectives (Zhou et al., 2021).

The inclusion of market-oriented scheduling principles further aligns the framework with the economic realities of cloud computing. Cloud resources are not merely technical entities but commodities that are bought and sold in competitive markets (Wu et al., 2013). By incorporating pricing and bidding into the learning process, the framework acknowledges that optimal scheduling is not only a matter of minimizing delay or energy consumption but also of managing economic tradeoffs. This perspective is particularly relevant in multi provider end edge cloud environments, where resource allocation decisions have financial as well as technical consequences (Duan et al., 2023).

Despite its conceptual strengths, the framework also raises important questions and potential criticisms. One concern is the complexity of implementing and maintaining such a multi layered system in practice. Deep reinforcement learning models require substantial computational resources and careful tuning, while swarm and genetic algorithms introduce additional overhead (Huang et al., 2020; Meshkati and Safi Esfahani, 2019). In real world deployments, the cost of running the scheduler itself must be weighed against the performance gains it delivers. This tradeoff is particularly acute at the edge, where resources are often constrained (Ren et al., 2019).

Another challenge relates to the interpretability and trustworthiness of learning based schedulers. Deep neural networks are often criticized as black boxes whose internal decision-making processes are difficult to explain (Sutton and Barto, 2018). In mission critical applications, such as healthcare or transportation, stakeholders may be reluctant to entrust scheduling decisions to opaque algorithms. The explicit modeling of queues and economic factors in the framework may partially mitigate this concern by providing more interpretable state variables and reward components, but significant work remains to be done in developing explainable and verifiable learning-based schedulers (Pinciroli et al., 2020).

The framework also invites reflection on the broader trajectory of intelligent computing. As systems become more autonomous and self optimizing, the role of human operators shifts from direct control to oversight and policy design. This transition raises ethical and organizational questions about accountability, governance, and the distribution of power in digital infrastructures (Ren et al., 2020). While the present study does not address these issues directly, its vision of a self organizing scheduling

International Journal of Advance Scientific Research
(ISSN – 2750-1396)
VOLUME 06 ISSUE 01 Pages: 150-167
OCLC – 1368736135

ecosystem highlights the need for interdisciplinary research that integrates technical, economic, and social perspectives.

In terms of future research, the framework opens several promising avenues. One direction is the empirical validation of the theoretical claims advanced in this study through large scale simulations and real-world deployments. While prior work such as Kanikanti et al. (2025) provides important empirical support for deep Q learning driven queuing-based scheduling, extending these results to full end edge cloud environments will require new experimental platforms and datasets. Another direction is the exploration of alternative reinforcement learning architectures, such as actor critic methods and multi agent systems, which may offer improved scalability and stability in distributed settings (Zhou et al., 2021).

The integration of predictive analytics and real time clustering, as proposed by Zhang and He (2019), also represents a fertile area for further investigation. By enhancing the accuracy of state representations and workload forecasts, such techniques could further improve the performance of learning based schedulers. Similarly, advances in energy modeling and green computing could be incorporated into the reward function to support more sustainable cloud operations (Meshkati and Safi Esfahani, 2019).

In conclusion, the proposed framework represents a significant step toward a more holistic and intelligent approach to task scheduling in end edge cloud environments. By synthesizing deep reinforcement learning, queuing theory, swarm intelligence, and market-oriented principles, it offers a rich conceptual foundation for future research and development. Its alignment with and extension of existing literature, particularly the influential work of Kanikanti et al. (2025), underscores its relevance and potential impact on the evolving field of distributed computing.

Conclusion

This study has presented a comprehensive and theoretically grounded framework for deep reinforcement learning driven queuing and swarm intelligence-based task scheduling in end edge cloud computing environments. Motivated by the increasing complexity, heterogeneity, and dynamism of modern distributed systems, the framework seeks to move beyond traditional heuristic and metaheuristic scheduling paradigms by embracing adaptive, learning oriented, and economically informed decision making. At the heart of the framework lies the recognition that effective scheduling in contemporary infrastructures requires not only computational efficiency but also an understanding of temporal, energetic, and economic dynamics that shape system behavior (Ren et al., 2019; Duan et al., 2023).

The integration of optimal queuing theory with deep Q learning, as pioneered by Kanikanti et al. (2025), provides a critical foundation for this approach. By treating queues as informative and dynamic elements of the learning environment,

International Journal of Advance Scientific Research
(ISSN – 2750-1396)
VOLUME 06 ISSUE 01 Pages: 150-167
OCLC – 1368736135

the framework enables the scheduler to anticipate congestion and make more foresighted decisions. This queuing aware perspective, when combined with the adaptive capabilities of reinforcement learning, offers a powerful tool for managing delay, throughput, and stability in high load and highly variable conditions (Shafi et al., 2020; Zhou et al., 2021).

The incorporation of swarm intelligence and evolutionary computation further enriches the framework by providing robust mechanisms for exploration and adaptation. These population-based methods complement the policy learning of reinforcement learning by maintaining diversity and guiding the search toward promising regions of the solution space (Song et al., 2020; Sulaiman et al., 2021). The result is a hybrid scheduling ecosystem that is both flexible and resilient, capable of responding to rapid changes in workload and resource availability.

By embedding market-oriented principles into the state and reward structures, the framework also acknowledges the economic realities of cloud and edge computing. In a world where computational resources are traded and priced, scheduling decisions must balance technical performance with cost and fairness considerations (Wu et al., 2013). The proposed model offers a pathway toward economically rational and socially sustainable resource orchestration.

While the framework is ambitious in scope, it is also subject to limitations related to computational complexity, data availability, and interpretability. Addressing these challenges will require continued research and experimentation, as well as collaboration between computer scientists, engineers, and economists. Nevertheless, the conceptual advances articulated in this study provide a compelling vision of how intelligent, self-organizing scheduling systems can support the next generation of digital infrastructures.

In sum, this article contributes to the evolving discourse on intelligent distributed computing by offering a deeply integrated and theoretically informed approach to task scheduling. Grounded in a rich body of prior research and inspired by the seminal contributions of Kanikanti et al. (2025), it lays the groundwork for future innovations in end edge cloud orchestration and adaptive resource management.

REFERENCES

1. Zhou, C., Wu, W., He, H., Yang, P., Lyu, F., Cheng, N., Shen, X. Deep reinforcement learning for delay oriented iot task scheduling in sagin. IEEE Transactions on Wireless Communications 20(2), 911–925.
2. Huang, X., Li, C., Chen, H., An, D. Task scheduling in cloud computing using particle swarm optimization with time varying inertia weight strategies. Cluster Computing 23, 1137–1147.
3. Ren, J., Jiang, H., Shen, X., et al. Editorial of ccf transactions on networking: special issue on intelligence enabled end edge cloud

orchestrated computing. CCF Transactions on Networking 3, 155–157.
4. Verma, A., Kaushal, S. A hybrid multi objective particle swarm optimization for scientific workflow scheduling. Parallel Computing 62, 1–19.
5. Khan, M.I., Xia, K., Ali, A., Aslam, N. Energy aware task scheduling by a true online reinforcement learning in wireless sensor networks. International Journal of Sensor Networks 25, 244–258.
6. Dorigo, M., Maniezzo, V., Colorni, A. Ant system: Optimization by a colony of cooperating agents. IEEE Transactions on Systems Man and Cybernetics 26, 29–41.
7. Pinciroli, R., Ali, A., Yan, F., Smirni, E. Cedule plus: Resource management for burstable cloud instances using predictive analytics. IEEE Transactions on Network and Service Management 18, 945–957.
8. Sutton, R.S., Barto, A.G. Reinforcement Learning: An Introduction. MIT Press.
9. Ren, J., Zhang, D., He, S., Zhang, Y., Li, T. A survey on end edge cloud orchestrated network computing paradigms: Transparent computing, mobile edge computing, fog computing, and cloudlet. ACM Computing Surveys 52(6).
10. Kanikanti, V. S. N., Tiwari, S. K., Nayan, V., Suryawanshi, S., Chauhan, R. Deep Q Learning Driven Dynamic Optimal Task Scheduling for Cloud Computing Using Optimal Queuing. Proceedings of the International Conference on Computational Intelligence and Knowledge Economy, 217–222.
11. Gabi, D., Ismail, A.S., Dankolo, N.M. Minimized makespan based improved cat swarm optimization for efficient task scheduling in cloud datacenter. Proceedings of the High Performance Computing and Cluster Technologies Conference, 16–20.
12. Wu, Z., Liu, X., Ni, Z., Yuan, D., Yang, Y. A market oriented hierarchical scheduling strategy in cloud workflow systems. Journal of Supercomputing 63, 256–293.
13. Wei, Z., Zhang, Y., Xu, X., Shi, L., Feng, L. A task scheduling algorithm based on Q learning and shared value function for wireless sensor networks. Computer Networks 126, 141–149.
14. Eberhart, R., Kennedy, J. A new optimizer using particle swarm theory. Proceedings of the International Symposium on Micro Machine and Human Science, 39–43.
15. Guevara, J.C., da Fonseca, N.L. Task scheduling in cloud fog computing systems. Peer to Peer Networking and Applications 14, 962–977.
16. Zhang, H., He, G. An adaptive task scheduling system based on real time clustering and NetFlow prediction. Proceedings of the International Conference on Computer and Communications, 77–80.
17. Holland, J.H. Adaptation in Natural and Artificial Systems. MIT Press.
18. Wei, Z., Liu, F., Zhang, Y., Xu, J., Ji, J., Lyu, Z. A Q learning algorithm for task scheduling based on improved support vector machine in wireless sensor networks. Computer Networks 161, 138–149.
19. Meshkati, J., Safi Esfahani, F. Energy aware resource utilization based on particle swarm optimization and artificial bee colony

algorithms in cloud computing. Journal of Supercomputing 75, 2455–2496.

20. Shafi, U., Shah, M.A., Wahid, A., Abbasi, K., Javaid, Q., Asghar, M.N., Haider, M. A novel amended dynamic round robin scheduling algorithm for timeshared systems. International Arab Journal of Information Technology 17, 90–98.
21. Raju, R., Babukarthik, R.G., Chandramohan, D., Dhavachelvan, P., Vengattaraman, T. Minimizing the makespan using hybrid algorithm for cloud computing. Proceedings of the International Advance Computing Conference, 957–962.
22. Sulaiman, M., Halim, Z., Lebbah, M., Waqas, M., Tu, S. An evolutionary computing based efficient hybrid task scheduling approach for heterogeneous computing environment. Journal of Grid Computing 19, 11.
23. Duan, S., Wang, D., Ren, J., Lyu, F., Zhang, Y., Wu, H., Shen, X. Distributed artificial intelligence empowered by end edge cloud computing: A survey. IEEE Communications Surveys and Tutorials 25(1), 591–624.
24. Zhou, Z., Li, F., Zhu, H., Xie, H., Abawajy, J.H., Chowdhury, M.U. An improved genetic algorithm using greedy strategy toward task scheduling optimization in cloud environments. Neural Computing and Applications 32, 1531–1541.
25. Song, A., Chen, W.N., Luo, X., Zhan, Z.H., Zhang, J. Scheduling workflows with composite tasks: A nested particle swarm optimization approach. IEEE Transactions on Services Computing 15, 1074–1088.
26. Watkins, C.J., Dayan, P. Q learning. Machine Learning 8, 279–292.
27. Kaelbling, L.P., Littman, M.L., Moore, A.W. Reinforcement learning: A survey. Journal of Artificial Intelligence Research 4, 237–285.
28. Sharma, M., Garg, R. HIGA: Harmony inspired genetic algorithm for rack aware energy efficient task scheduling in cloud data centers. Engineering Science and Technology International Journal 23, 211–224.
29. Jamil, B., Shojafar, M., Ahmed, I., Ullah, A., Munir, K., Ijaz, H. A job scheduling algorithm for delay and performance optimization in fog computing. Concurrency and Computation Practice and Experience 32, e5581.
30. Jiang, M., Wu, T., Wang, Z., Gong, Y., Zhang, L., Liu, R.P. A multi intersection vehicular cooperative control based on end edge cloud computing. IEEE Transactions on Vehicular Technology 71(3), 2459–2471.