IJRAR.ORG E-ISSN: 2348-1269, P-ISSN: 2349-5138

INTERNATIONAL JOURNAL OF RESEARCH AND ANALYTICAL REVIEWS (IJRAR) | IJRAR.ORG

An International Open Access, Peer-reviewed, Refereed Journal

Task Scheduling Methods in Edge-Cloud Environments for Emerging Applications

D Mamatha Rani[1*], Supreethi K P[2], Bipin Bihari Jayasingh[3]

[1]Department of Computer Science,TSWRDCW, Nizamabad,Telangana 503003,India

[2]Department of Computer science and Engineering,Jawaharlal Nehru Technological University, Hyderabad, Telangana-500085, India,

[3]Professor and Head of DepartmentCVR College of Engineering/IT Department Hyderabad, Telangana-501510, India

ABSTRACT : Unprecedented growth of cloud-assisted use cases has led to compelling Cloud Service Providers (CSPs) to optimize resource usage in presence of Service Level Agreements (SLAs). Ubiquitous adoption of technological innovation such as Internet of Things (IoT) has led to emergence of fog and edge computing phenomena which leverages latency. In presence of IoT applications, scheduling of tasks is challenging for many reasons such as network hierarchy, heterogeneity of resources, mobility of devices, resource constrained devices and stochastic behavior of nodes. Traditional cloud scheduling algorithms are not sufficient to harness the power of the dynamic computing environment made up of cloud, fog and edge resources. To overcome this problem, different scheduling algorithms came into existence. Reinforcement learning is one such technique used with machine learning approach. However, the research in this paper has revealed that solutions on top of RL suffer from issues such as lack of quick adaptability and generalizability. In this paper, we investigate on state of the art methods including RL based approaches. It throws light on different techniques, their merits and demerits. It also shows the research gaps that are to be considered for further research. This research assumes significance as the optimization of cloud-fog-edge resources through efficient and intelligent scheduling has benefits for consumers and service providers. In other words, such optimization brings about equilibrium between consumer satisfaction and business growth of cloud service providers.

Keywords: Cloud Computing, Fog Computing, Edge Computing, Task Scheduling, Internet of Things (IoT), Reinforcement Learning, Machine Learning, Artificial Intelligence

1. INTRODUCTION

Technological innovations such as Artificial Intelligence (AI) and Internet of Things (IoT) have paved way for solving many real world problems. With the emerging use cases based on such technologies, diversified applications with underlying tasks with varied deadline requirements are witnessed. IoT applications produce large volumes of data and demand large scale infrastructure to deal with such data. There are many real world applications that need deployment of tasks in cloud computing environments. Certain IoT workflow applications need edge-cloud environments to satisfy latency needs. With cloud computing and edge computing combination, there is increased challenge in task scheduling. As explored in [1] and [2] task scheduling plays crucial role in resource optimization and also ensuring efficiency of cloud data centers.

Many task scheduling algorithms that focus on cloud computing environment came into existence. For instance, [1], [3], [5], [6] and [9] have task scheduling methods that focused on cloud computing and they do not support edge-computing environments. Beloglazov and Buyya [1] proposed a method for improving resource utilization in cloud through VM migration and consolidation. They found that VM live migration has potential to exploit idle nodes in cloud data centers to optimize resource utilization and reduce energy consumption.Bui *et al.* [3] proposed a task scheduling method based in Gaussian process regression method instead of heuristics approach. It exploits Gaussian process regression method to estimate resource consumption for the tasks to be scheduled.In [5] DRL based online offloading method is proposed based on deep neural networks. It is a scalable solution since it is learning based approach.Basu *et al.* [6] proposed VM live migration based RL algorithm known as Megh. It is designed to reduce cost and energy consumption. It takes dynamic workloads that are uncertain and learns with RL towards optimizing energy and cost parameters.Mao *et al.* [9] employed DDQN for efficient resource management.There are certain task scheduling methods that considered edge-computing as well as explored in [2]. These methods consider edge-cloud environment where scheduling may take place into both cloud resources and edge-resources based on the deadlines of tasks. Since scheduling has plethora of benefits to both service consumers and service providers, it is indispensable to have thorough investigation on existing methods and come up with research gaps. Towards this end, in this paper, we made review of existing methods for task scheduling. Our contributions in this paper are as follows.

1. We have made review of task scheduling papers on cloud and edge-cloud environments considering different optimization parameters.
2. We have made summary of findings in terms of different aspects considered in the existing methods and optimization parameters.
3. The research gaps are provided that could trigger further investigations in the area of task scheduling in cloud and edge-cloud environments.

The remainder of the paper is structured as follows. Section 2 reviews on state of the art of recent task scheduling methods that focused on cloud and cloud-edge environments. Section 3 presents summary of the research carried out in the existing methods in terms of number of parameters used for optimization. Section 4 concludes our work and gives scope for possible future work.

2. STATE OF THE ART OF RECENT TASK SCHEDULING METHODS

This section reviews recent literature on different techniques or methods proposed for task scheduling in cloud and edge-cloud or fog-cloud environments. It throws light into different aspects of scheduling techniques, their merits and demerits.

2.1 VM Migration and VM Consolidation

VM plays vital role in cloud infrastructure for resource provisioning. Beloglazov and Buyya [1] proposed a method for improving resource utilization in cloud through VM migration and consolidation. They found that VM live migration has potential to exploit idle nodes in cloud data centers to optimize resource utilization and reduce energy consumption. They considered dynamic environment and presence of heterogeneous cores for their task scheduling study. Their method is based on heuristics approach. It considers SLA negotiations and algorithmis designed to support optimizations such as energy efficiency and SLAs. Their algorithm monitors VMs and their resource usage. By considering VM consolidation and VM live migration, their method is aimed at reducing energy consumption and adherence to SLAs. This method lacks adaptive QoS and support for dynamic workloads.

Basu *et al*. [6] proposed VM live migration based RL algorithm known as Megh. It is designed to reduce cost and energy consumption. It takes dynamic workloads that are uncertain and learns with RL towards optimizing energy and cost parameters. It is based on Q-learning approach considering adaptive QoS and stochastic workloads. The algorithm has provision for dimensionality reduction to project state-action space dynamically. Thus it became a real-time and scalable technique for efficient task scheduling. It has initial policy in hand to start with and while learning, it improves the policy for task scheduling. It has a policy iteration approach and gets rid of curse of dimensionality. It closely observes transitions from one state to another state in the state-action space for optimal decision making.

2.2 Task Scheduling for Fog-Cloud Environment

Fog-cloud environment consists of fog computing and also cloud computing combination for rendering faster services to real time applications. Particularly IoT use cases need such environment. Pham and Huh [2] proposed a task scheduling method based on heuristics approach for such environment. It is designed to work for heterogeneous cores in fog-cloud. They considered optimizations such as energy efficiency and cost reduction by scheduling tasks in edge-cloud environment. Their algorithm is based on heuristics towards reducing cost and energy consumption. It is based on graph representation. Towards this, their method exploits task graph and processor graph. Given the two graphs representing tasks and resources, their method finds appropriate resource allocation for given tasks. It has provision for determining task priority and then choose most suitable node for execution of task.

2.3 Predictive Optimization through Gaussian Process Regression

The work in [1] and [2] is based on heuristics approach that has limitations due to lack of learning based phenomena. Bui *et al*. [3] proposed a task scheduling method based in Gaussian process regression method instead of heuristics approach. It exploits Gaussian process regression method to estimate resource consumption for the tasks to be scheduled. Then for each task, they compute the need of physical servers in the cloud through a technique known as convex optimization. They considered acceptable level of QoS for the tasks and designed algorithm in such a way that minimal resources could achieve task execution. It has mechanisms to turn off idle servers and has provision for VM migration appropriately towards saving energy. They used workload collected from Google traces in order to have their empirical study. They also considered SLA violations along with energy efficiency. Their workloads are stochastic and dynamic in nature in presence of dynamic environment and heterogeneous cores.

2.4 Deep Reinforcement Learning

It is the learning based approach that has potential to exploit RL method which is based on continuous evaluation of action and reward mechanism. It is found suitable and better than heuristic approaches. Huang *et al*. [5] and Mao *et al*. [19] followed DRL approach for improving task scheduling performance in cloud computing environment. In [5] DRL based online offloading method is proposed based on deep neural networks. It is a scalable solution since it is learning based approach. In [19] DeepRM is the framework proposed for task scheduling considering efficient resource management. Both the methods are based on DQN approach rather than heuristics. Both methods considered optimization parameters such as energy and cost. In other words, they are designed to reduce energy consumption and also cost incurred for task execution in cloud environments. They support stochastic workloads and adaptive QoS. However, they do not support edge-cloud environments and do not optimize SLA and response time parameters.

Rjoub *et al*. [11] proposed a methodology based DQN and RL for efficient task scheduling in cloud environment. In fact, they proposed four approaches based on RL and DQN to support large-scale workloads. Their approaches focused on reduction of task waiting time and improve resource optimization. The four approaches are based on RL, DQN, RNN-LSTIM and DRL-LSTM respectively. They considered optimization of response time parameter in task scheduling. Due to faster convergence through learning based approaches, they could achieve improved response time. They made experiments with stochastic workloads in Google Cloud platform and found that their method is efficient when compared with many existing techniques. It has support for heterogonous cores and also adaptive QoS. DRL became very popular for learning based approaches. It is widely used in different domains including mobile 5G networks. Xiong *et al*. [20] has employed DRL for optimization of 5G networks towards better allocation of resources and optimization of services.

2.5 Replication and Speculative Execution

Task scheduling in cloud environments, it is important to consider meeting deadline of tasks. Therefore, deadline-critical jobs are common in cloud computing scenarios. Xu *et al*. [7] proposed a DNN based approach for task scheduling. Their approach is named as LASER which is based on replication and speculative execution. It follows two important strategies known as speculative-resume and speculative-resume along with quantitative investigation. LASER is based on deep learning algorithm that is executed on top of MapReduce framework. It is evaluated with stochastic workloads and observed that LASER is able to support adaptive QoS. This system is able to optimize two parameters such as SLA and cost. When a job is given with deadline, their algorithm exploits a metric to compute execution time probability to reach deadline and quantify the possibility of meeting deadline. Based on this, LASTER makes decisions on migration strategy. The metric used by the algorithm plays crucial role in decision making.

2.6 DDQN Based Approaches

Deep Q-Learning networks play an important role in processing requests in a network. Zhang *et al*. [8] focused on energy efficient approach in task scheduling. It has considered edge computing and edge devices for scheduling of tasks. They considered a challenging problem associated with energy reduction in edge-cloud environment. They followed a hybrid approach considering DVFS technology and Q-learning model for energy optimization in scheduling. Further they improved their model with DDQN approach which is an optimized model for edge scheduling. It considers DVFS, production of Q-value and target network dynamics to train learning parameters. Its deep learning model exploits ReLU function to realize optimization instead of using sigmoid function to get rid of gradient vanishing. Finally, they developed a learning algorithm to train parameters and improve energy efficiency. It could save energy up to 2.4% when compared with existing models.

Li and Hu [10] also used deep RL for task scheduling in cloud computing environment. They proposed a method known as DeepJS based on RL framework. It is modelled after a problem known as bin packing problem. It has provision for computing fitness function automatically in order to reduce response time and maximize throughput of job execution dynamics. Their empirical study is based on the trace-driven approach to demonstrate faster convergence. It continuously monitors state space and action space and designs agent scheduling process for improving scheduling performance. It has training algorithm to reap benefits of RL model. Fitness computation involved in DeepJS has its important role in the overall scheduling decision making. It makes the algorithm energy efficient ensuring that cloud infrastructure is optimized for reduction of energy consumption.

3. MORE RECENT WORKS

This section reviews the works in the last two years to reflect on latest state of the art. Almutairi and Aldossary [21] proposed a novel method for IoT tasks to offload in edge-cloud ecosystem. It is designed to serve latency-sensitive applications in a better way. It has fuzzy logic based approach for inferring knowledge towards decision making in presence of resource utilization and dynamic resource utilization. Ding et al. [22] considered edge-cloud environment to investigate on stateful data stream applications. They proposed a method to judge state migration overhead and make partitioning decisions based on the dynamically changing network bandwidth availability. Murad et al. [23] proposed an improved version of min-min task scheduling method to deal with scientific workflows in cloud computing. It could reduce the minimum completion time besides optimizing resource utilization. Bulej et al. [24] did their research on the management of latency in edge-cloud ecosystem towards better performance in task scheduling in presence of dynamic workloads. It is designed to explore upper bound of response time and optimize the performance further. Almutairi and Aldossary [25]proposed an edge-cloud system architecture to investigate on modeling methodology on task offloading. It has offloading latency models along with variousoffloading schemes. Their simulations are made using EdgeCloudSim. They intend to improve it in future with fuzzy logic.

Zhang and Shi [26] explored on workflow scheduling in edge-cloud environment. They analyzed different possibilities in workflow scheduling in such ecosystem. They opined that workflow applications need novel approaches in scheduling process. Zhao et al. [27] focused on task scheduling along with security to prevent intrusions in edge computing environments. They considered low-rate intrusions and focused on preventing them along with task scheduling. It is a Q-learning based approach designed to meet runtime requirements based on learning process. Zhang et al. [28] proposed a time-sensitive algorithm that dynamically caters to the needs of deadline-aware tasks in edge-cloud environments. It considers job size and server capability in given dynamic and hierarchical scenario. It is multi-objective task considering execution time, cost and reduction of SLAs. Lakhan et al. [29] proposed a task scheduling approach for IoT tasks considering a hybrid mechanism consisting of task scheduling and task offloading. Singh and Bhushan [30] proposed a method for task scheduling based Cuckoo Search Optimization (CSO). It has integrated local search strategy. From these recent works it is found that they targeted IoT kind of workflows in edge-cloud environments. There is Q-Learning used in one of the papers. However, deep reinforcement learning is not found in the latest works.

3. SUMMARY OF RELATED WORKS

From the review of literature, there are many closely related works found. This section provides the summary of those works. In [1], [2],[12], [13], [14], [15] and [19] several methods based on heuristics are proposed for scheduling of tasks in edge-cloud environments. Each method is designed to optimize different parameters. Some works are meant for cloud while some other methods are designed to deal with edge-cloud environments. It is a known thing that certain heuristics are meant for generic use cases and they fail to deal with environments with dynamic changes. However, a learning-based method has potential to adapt to the runtime situations and tune different parameters accordingly. Such models are found in [3], [4], [5], [6], [7], [8], [9], [16], [17] and [18]. These models are based on ML and DL techniques that are designed to optimize resource management. Most widely used approaches are DRL and DNN methods that are learning-based phenomena. The

Reference	**Dynamic**	**Stochastic Workload**	**Decentralized**	**Edge Cloud**	**Adaptive QoS**	**Heterogeneous**
[1]	Yes	No	No	No	No	Yes
[2]	No	No	No	Yes	No	Yes
[3]	Yes	Yes	No	No	No	Yes
[5], [9]	Yes	Yes	No	No	Yes	Yes
[6]	Yes	Yes	No	No	Yes	Yes
[7]	Yes	Yes	No	No	No	No
[8], [10]	Yes	Yes	No	No	Yes	Yes
[11]	Yes	Yes	No	No	Yes	Yes
[21]	Yes	No	No	No	Yes	Yes
[22]	Yes	No	No	Yes	Yes	No
[23]	Yes	No	No	No	Yes	Yes
[24]	Yes	No	No	No	Yes	Yes
[25]	Yes	No	No	No	Yes	Yes
[26]	Yes	Yes	No	No	Yes	Yes
[28]	Yes	No	No	No	No	No
[29]	Yes	No	No	No	Yes	Yes
[30]	Yes	No	Yes	No	Yes	Yes

primary objective of many of these works is improving energy efficiency of cloud infrastructure. Towards this end, they predict optimization possibilities and perform accordingly.

Bui *et al*. [3] proposed an optimization framework for cloud with predictive approach. They could predict dynamics of resource utilization for scheduling by employed a method named Gaussian process regression. The prediction result helped them to minimize number of servers to be used to process the requests leading to reduction of energy usage. Their method is, however, based on heuristics and is not suitable for dynamic workloads and edge-cloud environments. Cheng *et al*. [4] explored DRL based approach towards task scheduling and resource provisioning in cloud. They further optimized Q-learning method to reduce task rejection rate and improve energy efficiency. Basu *et al*. [6] focused on the problem of live migration of VMs based on RL based Q-learning process. Their methodology improve live migration and heuristics based existing approaches. Towards this end, their method exploits Megh and RL based model to have continuous adaptation to the runtime situations towards leveraging energy efficiency. Xu *et al*. [7] defined a DNN approach named LASER to support deadline critical jobs with replication and speculative execution. Their implementation of the framework is designed for Hadoop framework. Zhang *et al*. [8] defined a DDQN method towards energy efficiency in edge computing. It is based on Q-learning process and also dynamic voltage frequency scaling (DVFS) method that has potential to reduce energy usage. As Q-learning is not able to recognize continuous system states, they extended it to have double deep Q-learning.

Similar to the work of [4], Mao *et al*. [9] employed DDQN for efficient resource management. This kind of work is also found in Li *et al*. [10]. Both have employed DRL technique towards job scheduling over diversified resources. However, these learning based .methods are not able to withstand in stochastic environments. Mao *et al*. [19] and Rjoub *et al*. [11] investigated on DRL based approach for task scheduling in edge-cloud. However, they considered only response time in their research. Its drawback is that they could not exploit asynchronous method for optimization of their methods towards robustness and adaptability. There is need for improving it by considering dynamic optimization of parameters in presence of stochastic workloads.

Table 1: Merits and demerits of existing scheduling methods

As presented in Table 1, we summarize our findings leading to important research gaps. The summary is made in terms of different parameters such as dynamic environment, presence of stochastic workload, decentralized environment, usage of edge cloud, consideration for adaptive QoS and presence of heterogeneous cores for task scheduling. The advantages of the research in [1] include consideration of dynamic environment and heterogeneous cores. However, it does not consider adaptive QoS, edge cloud, decentralized environment and presence of stochastic workloads. The work

in [2] considered edge cloud and also heterogeneous cores for their task scheduling research. However, it does not support adaptive QoS, dynamic and decentralized environment, edge cloud and stochastic workloads. The merits of [3] include the consideration of dynamic environment, stochastic workloads and heterogeneous cores. But it lacks in adaptive QoS, support for edge cloud and decentralized environments. The research in [5] and [9] has similar findings. Their method has provision for considering dynamic environment, heterogeneous cores, adaptive QoS and stochastic workloads. But is not designed for edge cloud and decentralized environment. In [6], there is consideration of dynamic environment, stochastic workloads, adaptive QoS and heterogeneous cores. But does not support decentralized and edge-cloud environment. The work in [7] supports dynamic
environment and stochastic workloads. However, it has limitations to deal with heterogeneous cores, adaptive QoS, edge cloud and decentralized environments.
There is similarity in the task scheduling methods proposed in [8] and [10]. Their methods are dynamic in nature supporting adaptive QoS and stochastic workloads besides dealing with heterogeneous cores. However, they do not support decentralized and edge cloud environment. The scheduling research in [11] supports dynamic environment along with stochastic workloads. They also deal with heterogeneous cores and adaptive QoS.

Reference	**Method**	**Optimization Parameters**			
		SLA Violations	**Cost**	**Response Time**	**Energy**
[1]	Heuristics	Yes	No	No	Yes
[2]	Heuristics	No	Yes	No	Yes
[3]	Gaussian Process Regression	Yes	No	No	Yes
[5], [9]	DQN	No	Yes	No	Yes
[6]	Q Learning	No	Yes	No	Yes
[7]	DNN	Yes	Yes	No	No
[8], [10]	DDQN	No	No	No	Yes
[11]	DRL (REINFORCE)	No	No	Yes	No
[21]	SJF	No	No	Yes	Yes
[22]	Cloud Computing	No	No	Yes	No
[24]	Cloud computing	No	Yes	Yes	Yes
[26]	CSA	No	Yes	No	No
[27]	Cloud computing	No	Yes	Yes	Yes
[28]	cloud computing	No	Yes	No	No
[30]	CSP	No	Yes	Yes	No

Table 2: Optimization parameters considered by existing scheduling method

However, the drawback is that those methods do not consider decentralized and edge-cloud environment. Table 1 also provides summary of most recent works [21]-[30] in terms of support for different features.

As presented in Table 2, we summarized the existing methods in terms of optimization parameters and the approach considered in the task scheduling research. The optimization parameters considered for comparative study of existing methods are SLA violations, cost, response time and energy. Research in [1] is based on heuristics approach and considers energy and SLA violations parameters. Their research lacks in the study of response time and cost of scheduling that are crucial for task scheduling. The work in [2] is also based on heuristics method but considers cost and energy parameters. It does not throw light on response time and SLA violations. In [3], their method is based on Gaussian process regression and considers two parameters such as energy and SLAs. It has no support for optimization of cost and response time. The task scheduling research in [5] and [9] is based on DQN method and supports cost and energy parameters for optimization. However, they have no optimization of SLAs and response time. In [6] Q-learning based phenomenon is used considering energy and cost dynamics for optimization. But it lacks optimization of response time and SLAs. DNN is the scheduling method used in [7] and it has support for optimization of cost and SLA parameters. It lacks support for energy and response time optimizations. The work in [8] and [10] is based on DDQN method and it supports only energy parameter for optimization. It lacks support for response time, cost and SLA optimizations. In [11] DRL method is used for task scheduling by considering response time for optimization. However, it does not support optimization of SLAs, cost and energy. Table 1 also provides summary of most recent works [21]-[30] in terms of support for different optimization parameters.

4. CONCLUSION AND FUTURE WORK

In this paper, we investigate on state of the art methods including RL based approaches. It throws light on different techniques, their merits and demerits. It also shows the research gaps that are to be considered for further research. This research assumes significance as the optimization of cloud-fog-edge resources through efficient and intelligent scheduling has benefits for consumers and service providers. In other words, such optimization brings about equilibrium between consumer satisfaction and business growth of cloud service providers. Existing methods are studied and compared in terms of different aspects such as dynamic environment, presence of stochastic workload, decentralized environment, usage of edge cloud, consideration for adaptive QoS and presence of heterogeneous cores for task scheduling. This research also considered optimization parameters and the approach considered in the task scheduling research. The optimization parameters considered for comparative study of existingmethods are SLA violations, cost, response time and energy. The research in this paper has resulted in many insights or research gaps as summarized in Table 1 and Table 2. In future we would like to consider research gaps found in this paper and propose a deep learning based framework and algorithms for efficient scheduling of IoT application tasks in Edge-Cloud computing environments.

References

[1] A. Beloglazov and R. Buyya, “Optimal online deterministic algorithms and adaptive heuristics for energy and performance efficient dynamic consolidation of virtual machines in cloud data centers,” Concurrency and Computation: Practice and Experience, vol. 24, no. 13, pp. 1397–1420, 2012.

[2] Xuan-Qui Pham and Eui-Nam Huh, “Towards task scheduling in a cloud-fog computing system,” in 2016 18th Asia-Pacific Network Operations and Management Symposium, Oct 2016, pp. 1–4.

[3] D.-M. Bui, Y. Yoon, E.-N. Huh, S. Jun, and S. Lee, “Energy efficiency for cloud computing system based on predictive optimization,” Journal of Parallel and Distributed Computing, vol. 102, pp. 103–114, 2017.

[4] M. Cheng, J. Li, and S. Nazarian, “DRL-cloud: Deep reinforcement learning-based resource provisioning and task scheduling for cloud service providers,” in Proceedings of the 23rd Asia and South Pacific Design Automation Conference. IEEE Press, 2018, pp. 129–134.

[5] L. Huang, S. Bi, and Y. J. Zhang, “Deep reinforcement learning for online computation offloading in wireless powered mobile-edge computing networks,” IEEE Transactions on Mobile Computing, pp. 1–1, 2019.

[6] D. Basu, X. Wang, Y. Hong, H. Chen, and S. Bressan, “Learn-asyou-go with megh: Efficient live migration of virtual machines,” IEEE Transactions on Parallel and Distributed Systems, vol. 30, no. 8, pp. 1786–1801, 2019.

[7] M. Xu, S. Alamro, T. Lan, and S. Subramaniam, “Laser: A deep learning approach for speculative execution and replication of deadline-critical jobs in cloud,” in Proceedings of the 26th International Conference on Computer Communication and Networks (ICCCN). IEEE, 2017, pp. 1–8.

[8] Q. Zhang, M. Lin, L. T. Yang, Z. Chen, S. U. Khan, and P. Li, “A double deep Q-learning model for energy-efficient edge scheduling,” IEEE Transactions on Services Computing, 2018, (preprint).

[9] H. Mao, M. Alizadeh, I. Menache, and S. Kandula, “Resource management with deep reinforcement learning,” in Proceedings of the 15th ACM Workshop on Hot Topics in Networks. ACM, 2016, pp. 50–56.

[10] F. Li and B. Hu, “Deepjs: Job scheduling based on deep reinforcement learning in cloud data center,” in Proceedings of the 2019 4th International Conference on Big Data and Computing, 2019, pp. 48–53.

[11] G. Rjoub, J. Bentahar, O. A. Wahab, and A. S. Bataineh, “Deep and reinforcement learning for automated task scheduling in largescale cloud computing systems.”

[12] O. Skarlat, M. Nardelli, S. Schulte, M. Borkowski, and P. Leitner, “Optimized IoT service placement in the fog,” Service Oriented Computing and Applications, vol. 11, no. 4, pp. 427–443, 2017.

[13] X.-Q. Pham, N. D. Man, N. D. T. Tri, N. Q. Thai, and E.-N. Huh, “A cost-and performance-effective approach for task scheduling based on collaboration between cloud and fog computing,” International Journal of Distributed Sensor Networks, vol. 13, no. 11, pp. 1–16, 2017.

[14] A. Brogi and S. Forti, “QoS-aware deployment of IoT applications through the fog,” IEEE Internet of Things Journal, vol. 4, no. 5, pp. 1185–1192, 2017.

[15] T. Choudhari, M. Moh, and T.-S. Moh, “Prioritized task scheduling in fog computing,” in Proceedings of the ACMSE Conference, ser. ACMSE ’18. New York, NY, USA: ACM, 2018, pp. 22:1–22:8.

[16] H. Li, K. Ota, and M. Dong, "Learning IoT in edge: Deep learning for the Internet of Things with edge computing," IEEE Network, vol. 32, no. 1, pp. 96–101, 2018.

[17] M. Cheng, J. Li, and S. Nazarian, "DRL-cloud: Deep reinforcement learning-based resource provisioning and task scheduling for cloud service providers," in Proceedings of the 23rd Asia and South Pacific Design Automation Conference. IEEE Press, 2018, pp. 129–134.

[18] Q. Zhang, M. Lin, L. T. Yang, Z. Chen, and P. Li, "Energyefficient scheduling for real-time systems based on deep qlearning model," IEEE Transactions on Sustainable Computing, vol. 4, no. 1, pp. 132–141, 2017.

[19] H. Mao, M. Alizadeh, I. Menache, and S. Kandula, "Resource management with deep reinforcement learning," in Proceedings of the 15th ACM Workshop on Hot Topics in Networks, 2016, pp. 50–56.

[20] Z. Xiong, Y. Zhang, D. Niyato, R. Deng, P. Wang, and L.-C. Wang, "Deep reinforcement learning for mobile 5g and beyond: Fundamentals, applications, and challenges," IEEE Vehicular Technology Magazine, vol. 14, no. 2, pp. 44–52, 2019.

[21] JaberAlmutairi;Mohammad Aldossary; (2021). A novel approach for IoT tasks offloading in edge-cloud environments . Journal of Cloud Computing, p1-19.

[22] ShaoshuaiDing;LeiYang;JiannongCao;WeiCai;MingkuiTan;Zhenyu Wang; (2021). Partitioning Stateful Data Stream Applications in Dynamic Edge Cloud Environments . IEEE Transactions on Services Computing, p1-14.

[23] SALLAR SALAM MURAD1 , ROZIN BADEEL, NASHAT SALIH ABDULKARIM ALSANDI3 , RAFI. (2022). OPTIMIZED MIN-MIN TASK SCHEDULING ALGORITHM FOR SCIENTIFIC WORKFLOWS IN A CLOUD ENVIRONMENT. Journal of Theoretical and Applied Information Technology. 100(2), pp.480-506.

[24] Bulej, LubomÃ¬r; BureÅ¡, TomÃ¡Å¡; Filandr, Adam; HnÄ› tynka, Petr; HnÄ› tynkovÃ¡, Iveta; PacovskÃ½, Jan; Sandor, Gabor; Gerostathopoulos, Ilias (2021). Managing latency in edgeâ€ " cloud environment. Journal of Systems and Software, 172, p1-15.

[25] Jaber Almutairi1 and Mohammad Aldossary. (2021). Investigating and Modelling of Task Offloading Latency in Edge-Cloud Environment. Computers, Materials & Continua. 68(3), pp.1-18.

[26] RuiZhang;Wenyu Shi; (2021). Research on Workflow Task Scheduling Strategy in Edge Computer Environment . Journal of Physics: Conference Series, p1-6.

[27] Xu Zhao;GuangqiuHuang;LingGao;MaozhenLi;Quanli Gao; (2021). Low load DIDS task scheduling based on Q-learning in edge computing environment . Journal of Network and Computer Applications, p1-12.

[28] Yu Zhang 1,2 , Bing Tang, Jincheng Luo and Jiaming Zhang. (2022). Deadline-Aware Dynamic Task Scheduling in Edge–Cloud Collaborative Computing. MDPI. 11(.), pp.1-24.

[29] Abdullah Lakhan 1,2 ,Mazin Abed Mohammed 3 , Karrar Hameed Abdulkareem 4,5 , Mu. (2022). Delay Optimal Schemes for Internet of Things Applications in Heterogeneous Edge Cloud Computing Networks. MDPI. 22(.), pp.1-30.

[30] Mandeep Singh1 , Shashi Bhushan. (2022). CS Optimized Task Scheduling for Cloud Data Management. International Journal of Engineering Trends and Technology. 70(6), pp.114-121.