

Energy-Efficient Task Offloading in Edge Computing: A Survey of Deep Reinforcement Learning Approaches

Eleftheria Papageorgiou¹, Theodoros Theodoropoulos^{1,2}[0000–0002–4618–4891],
and Konstantinos Tserpes²[0000–0001–5183–1443]

¹ Harokopio University of Athens, Kallithea, Greece
{csi23310,ttheod}@hua.gr

² National Technical University of Athens, Zografou, Greece
tserpes@mail.ntua.gr

Abstract. In edge computing, task offloading involves transferring computational tasks from the “far-edge”, which includes end-user devices or less powerful edge devices, to the “near-edge”, comprising more capable edge servers, or to the ‘core’ cloud infrastructure. This practice optimizes performance, reduces latency, and enhances overall efficiency. Energy efficiency in particular has recently become a high-priority criterion for task offloading. A prominent technique for making offloading decisions in edge computing environments is Deep Reinforcement Learning (DRL), known for its ability to adapt to complex environments and excel in multi-objective optimization tasks in terms of decision quality and speed. This paper explores the details of DRL approaches, providing an overview of recent research developments in this field. To simplify the literature analysis, we classify DRL approaches for energy-efficient task offloading between two “computing continua”: the far/near-edge continuum, and the (far-)edge-cloud continuum.

Keywords: Green Computing · Efficient Computing · Deep Reinforcement Learning · Edge Computing · Edge-Cloud · Energy Efficiency

1 Introduction

At a conceptual level, Cloud computing concentrates the execution of the larger part of applications into virtual computing “silos” in which resources are connected with nearly zero communication latency. This trend is now revised in favor of the incorporation of devices into the resource pool that may present an increased latency to the existing silos but reduced latency with regards to the data sources and end users. The tradeoff is: reduced latency across application components Vs reduced latency between application Edges (data sources, end users) and application components.

The growing dependency on data has shifted this “Edge Computing” paradigm to further extremes, with the incorporation of heterogeneous, resource-constrained

devices into the resource pool, especially in IoT-based applications. The IoT-Edge-Cloud continuum [21] refers to a spectrum of computing resources ranging from centralized Cloud servers to decentralized Edge devices, with varying degrees of proximity to end-users and data sources. As IoT-Edge-Cloud Computing gradually becomes the dominant execution paradigm, energy efficiency becomes a key challenge to overcome in the complete range of the IoT-Edge-Cloud continuum.

The potentially resource-constrained nature of devices, especially in an IoT-Edge-Cloud context, underscores the imperative to judiciously harness the available Edge resources for optimal performance [30]. Given the inherent limitations in processing and storage capacities of those devices, it becomes crucial to design and implement efficient algorithms and resource management strategies that make the most out of the constrained environment. Additionally, the challenge is exacerbated by the fact that Edge devices often operate with restricted energy autonomy, necessitating a delicate balance between computational efficiency and energy conservation. Addressing these challenges is pivotal to unlocking the full potential of Edge computing in meeting the stringent requirements of next-generation applications while ensuring sustainability in the face of energy constraints.

Task offloading [16] is one of the distinguishing operations in IoT-Edge-Cloud computing, enabling the system to select the appropriate resource across the continuum where to deploy an appropriate application component instant. The objective of this operation is to optimize resource utilization and meet QoS/QoE requirements. Task offloading algorithms typically make optimization decisions on factors like latency, bandwidth, and resource utilization. This process enhances responsiveness, and improves overall system efficiency by processing data closer to the point of generation or consumption.

The relevant scientific literature, refers to two primary paradigms for energy-efficient task offloading in resource-constrained Edge devices. The first paradigm focuses on offloading tasks from one Edge device to more suitable Edge resources, such as Edge devices, and servers. We refer to this practice as "far/near-edge continuum offloading". The second paradigm involves offloading tasks from an Edge device to a broader range of computing resources across the computing continuum, which includes both Edge resources and centralized Cloud data centers. We refer to this practice as "far-edge to cloud continuum offloading" or simply "edge-cloud continuum offloading".

Furthermore, the literature promotes Deep Reinforcement Learning (DRL) [1] as a prominent technique to enable energy efficient offloading. DRL is a machine learning algorithm where an agent learns to make decisions by interacting with an environment to achieve certain goals. In the context of task offloading, DRL can assist in decision-making by dynamically determining which tasks should be processed locally on Edge devices or Cloud resources to be processed [26]. Through continuous learning and using feedback from the environment, DRL algorithms optimize task allocation based on factors such as network condi-

tions, device capabilities, QoS requirements and most recently -and importantly- based on energy efficiency requirements.

Recognizing the importance of the contribution of DRL in task offloading, we performed a review on contemporary solutions that are based on the use of DRL algorithms for energy efficient task offloading across the two compute continua. The analysis follows the following structure: Section 2 briefly describes various contemporary DRL algorithms that shall be explored within the frame of this work. Section 3 showcases various DRL solutions for energy efficient task offloading within the frame of exclusive Edge resources (far/near-edge continuum). Section 4 showcases various DRL solutions for energy efficient task offloading at the edge-cloud continuum. Finally, Section 5 summarizes the merits and findings of this work, and proposes potential future research directions.

2 Background: Deep Reinforcement Learning Algorithms

2.1 Deep Q-Network (DQN)

Deep Q-Network (DQN) [20] combines Q-learning with deep neural networks to approximate Q-values, which represent the expected cumulative rewards for actions in given states. DQN trains a neural network to predict Q-values directly from raw inputs, enabling learning of complex behaviors in high-dimensional environments. It uses experience replay and target networks to stabilize training and improve efficiency, learning from past experiences to mitigate issues of correlated data and non-stationarity. By iteratively interacting with the environment and updating network parameters via gradient descent, DQN optimizes decisions to maximize expected cumulative rewards over time.

2.2 Double Deep Q-Network (DDQN)

Double Deep Q-Network (DDQN) [28] enhances DQN by addressing the over-estimation bias in traditional Q-learning. DDQN uses two neural networks—the online and target networks—to decouple action selection and value estimation. The online network selects actions, while the target network estimates their values, reducing bias and stabilizing learning. DDQN also uses experience replay and target network updates to further improve stability and sample efficiency, enabling robust learning of optimal policies in complex environments with discrete action spaces.

2.3 Deep Deterministic Policy Gradient (DDPG)

Deep Deterministic Policy Gradient (DDPG) [24] is a model-free, off-policy algorithm for continuous action spaces. It uses neural networks to approximate both the policy (actor) and value function (critic), combining policy gradient methods and Q-learning in an actor-critic framework. DDPG employs target networks and experience replay to mitigate issues like non-stationarity and correlated data,

learning effectively from past experiences. By updating policy parameters in the direction of the estimated gradient of expected cumulative rewards, DDPG generates actions that maximize long-term rewards, facilitating complex behaviors in challenging environments.

2.4 Multi-Agent Deep Reinforcement Learning (MADRL)

Multi-Agent Deep Reinforcement Learning (MADRL) [12] extends DRL to scenarios with multiple agents interacting with each other and the environment. MADRL trains neural network-based agents to learn individual policies and strategies for cooperation or competition. It addresses challenges like non-stationarity and mixed interactions among agents. MADRL typically uses decentralized training with centralized execution, where agents share information during training but maintain decentralized control during execution. This approach enables agents to develop sophisticated behaviors and strategies for achieving collective goals in dynamic environments.

Multi-Agent Deep Deterministic Policy Gradient (MADDPG) [9] extends DDPG to multi-agent environments. MADDPG uses centralized training with decentralized execution, where each agent has its own actor and critic networks. During training, critics have access to all agents' observations and actions, enhancing coordination and learning of optimal policies. Decentralized execution allows agents to operate independently based on their observations during testing. MADDPG uses experience replay and target networks to stabilize learning, improve efficiency, and handle continuous action spaces, enabling robust learning in complex multi-agent environments.

3 Task Offloading in Far/Near-Edge Compute Continuum

Task offloading in the far/near-edge compute continuum involves the delegation of computational tasks from resource-constrained Edge devices to more powerful computing resources located in nearby Edge servers or to other Edge devices. This process aims to optimize performance, reduce latency, and conserve energy by distributing computing tasks strategically across the available resources. Task offloading decisions are typically based on factors such as the computational requirements of the task, the availability of resources, network conditions, and user preferences. Offloading decisions can be made dynamically in real-time using various algorithms and heuristics, ensuring efficient utilization of Edge resources while meeting the quality-of-service requirements of applications.

3.1 DQN

In Mobile Edge Computing (MEC), mobile terminals can offload tasks to nearby Edge servers for computational offloading. A DQN algorithm can be employed to determine the optimal task offloading and resource allocation strategy, ensuring proper resource allocations via MEC [15]. Here, each potential outcome

is represented as a state space, and transitions between states are considered actions. Utilizing the DQN approach reduces both energy consumption and task delay costs.

In wireless-powered MEC, multiple wireless devices connect to the same network and follow a binary offloading policy, either executing computation tasks locally or remotely on a MEC server [19]. The Deep Reinforcement learning-based Online Offloading (DROO) [14] algorithm maximizes the weighted sum computation rate in these networks, significantly reducing computational complexity by avoiding combinatorial optimization problems, especially in large networks. An adaptive procedure is proposed to automatically adjust DROO algorithm parameters, enhancing the algorithm’s learning from past experiences to improve offloading actions. This leads to reduced computation time per task and lower system energy consumption. Unlike the TADPG agent that runs on each mobile device, DROO operates a single agent for all wireless devices.

In vehicular Edge computing, deciding whether computation tasks should be executed locally or on Edge servers in nearby vehicles is crucial. A DQN-based algorithm incorporating Bayesian inference, leveraging prior distributions and statistical data, addresses this need [25]. By offloading tasks to vehicular Edge servers, the overall delay and energy costs are minimized.

3.2 DDQN

In 5G networks, MEC enables user equipment to offload tasks requiring substantial computation resources and energy to nearby MEC servers. The goal is to minimize energy consumption in a dynamic multi-user MEC system through joint optimization of computation offloading and resource allocation, considering task delay and requirements constraints. A study explores Q-learning and DDQN solutions [32]. The Q-learning agent learns optimal behaviors independently over time, whereas DDQN uses DNNs to estimate Q-learning’s action-value function, separating action selection and evaluation. Both methods effectively reduce energy consumption and achieve desirable average task delays.

3.3 DDPG

Using the DDPG paradigm, a Temporal Attentional Deterministic Policy Gradient (TADPG) approach [4] addresses computation offloading and resource allocation. Each mobile device runs a TADPG agent for dynamic partial task offloading and resource allocation decisions. The TADPG agent features a temporal feature extraction network, consisting of an attentional LSTM [23] and a 1-dimensional convolution residual block, along with a rank-based priority experience replay to prioritize significant experiences. This approach speeds up training and improves stability, reducing task completion time and energy consumption.

In MEC, IoT devices offload intensive computations to Edge servers to achieve low latency and reduced energy consumption. When multiple devices are connected, computing delays can occur. This issue can be addressed with a DDPG

approach, training a centralized policy to optimize local computing CPU frequencies and offloading policies, minimizing energy consumption and average computing delay [18].

3.4 MADRL

In the Intelligent Internet of Things (IIoT) paradigm, each smart device is an agent, known as a machine-type agent (MTA). In factories, MTAs communicate, generating large datasets that increase communication costs. A MADRL algorithm optimizes offloading policies [5], deciding whether to offload computations based on channel conditions and past data to minimize delay and interference, enhancing channel access success and reducing costs and energy consumption.

As IIoT evolves, multi-access MEC offloads Cloud services to Edge clients, often leading to high energy consumption. In dynamic network environments, intelligent task offloading strategies are challenging due to limited battery capacity [3]. The multi-agent deep deterministic policy gradient (MA2DDPG) computation offloading framework addresses this, using centralized training and distributed execution for optimal resource allocation and minimizing energy consumption and transmission latency [11]. This framework features in-layer optimization for resource allocation and outer-layer modules for task scheduling, energy transmission, and channel interference management.

4 Task Offloading in the Edge-Cloud Compute Continuum

Task offloading in the IoT-Edge-Cloud continuum [6] involves the strategic allocation of computing tasks across the spectrum of computing resources. This approach aims to optimize performance, reduce latency, and conserve energy by dynamically distributing tasks based on factors such as computational requirements, network conditions, and user preferences. Tasks can be offloaded from Edge devices to nearby Edge servers for low-latency processing or to the Cloud for high-complexity computations. Offloading decisions are made dynamically in real-time, considering the trade-offs between latency, bandwidth, and resource availability, to ensure efficient utilization of both Edge and Cloud resources while meeting application requirements.

4.1 DQN

Vehicle Ad Hoc Networks (VANETs) enable vehicles to communicate within smart city infrastructure. With the increasing number of vehicles, risks arise when offloading tasks through mobile devices like roadside base units. Blockchain and smart contracts ensure system safety, while a DRL algorithm [31] using extended DQN addresses the offloading problem by finding the optimal strategy for each vehicle, considering data amount, MEC processing power, throughput, and

bandwidth. Blockchain-based access control efficiently identifies and prevents unauthorized offloading, ensuring security and privacy.

A DRL-driven optimization scheme for resource allocation in IoT-Edge-Cloud environments includes intelligent content caching and request aggregation mechanisms. A DQN solution addresses asymmetrical control and multi-layer heterogeneous resource allocation problems by optimizing content distribution [10]. The DQN algorithm enhances content caching and task scheduling based on past user requests and available resources, improving bandwidth and resource use, reducing energy consumption, and enhancing environmental sustainability.

The Internet of Vehicles (IoV) connects vehicles to their environment, providing efficient, low-latency transmission services [22]. Dynamic vehicle environments impact communication, causing latency-sensitive services to face limitations. A cooperative IoT-Edge-Cloud content-delivery scheme uses a DQN approach [7] to optimize computing, caching, and communication resources, minimizing network latency and improving content distribution and QoS in asymmetrical IoV environments. Furthermore, Vehicle Edge networks, consisting of Cloud servers, roadside units, and moving vehicles, constitute an optimization problem in computational offloading and resource allocation. A distributed DQN learning algorithm [17] with parallel DNNs is capable of solving this problem, minimizing system costs in terms of time and energy.

4.2 DDQN

In [29], DDQN is proposed as an offloading algorithm for multi-user computing offloading in Cloud-assisted mobile Edge computing environments [13]. Multiple devices offload tasks simultaneously, evenly distributing wireless bandwidth. The optimal strategy found by the algorithm allows MEC and Cloud computing to collaborate, reducing delay and energy consumption effectively.

4.3 DDPG

5G beyond involves leveraging the capabilities of end devices, Edge servers, and the Cloud for computation-intensive and delay-sensitive applications. Thus, the decision-making process for computation offloading is quite complex. In [8], a DDPG algorithm is proposed to minimize system energy consumption. The DDPG algorithm solves joint computation offloading and resource allocation problems using an Actor-Critic framework, where the Actor generates actions and the Critic guides the Actor. Action refinement in DRL supports joint offloading and resource allocation, minimizing system energy consumption.

4.4 MADRL

The GreenKube framework [27] employs the MADRL concept to address latency issues and enhances computational infrastructure using AI methodologies like Deep Learning, Deep Reinforcement Learning, and Graph Neural Networks. Intelligent, Monitoring, and Forecasting Agents interact with the Kubernetes³

³ <https://kubernetes.io/>

platform to predict QoS metrics, guide strategies for Autoscaling, Self-Healing, and Task Offloading. These strategies ensure efficient use of computing power, minimize costs and energy consumption, and reduce latency for sensitive tasks. The framework aims to improve container orchestration by meeting QoS requirements and minimizing energy consumption using AI methodologies.

5 Conclusion

In summary, this paper investigates multiple solutions for optimizing task offloading in the Edge-Cloud compute continuum. We distinguish the works to those that promote edge to edge resource task offloading (Far/Near-Edge compute continuum) and those that promote edge to cloud resource offloading (Edge-Cloud compute continuum). The solutions proposed are based on DRL algorithms, like DQN, DDQN, DDPG, and MADRL, for optimizing task offloading while considering the system’s energy efficiency. A related work taxonomy based on the findings of this work is showcased in Table 1. As technology advances, the integration of innovative DRL algorithms with the Edge computing environment is essential so that the full potential of next-generation applications can be experienced by the users and simultaneously to tackle sustainability issues in the face of energy limitations.

Table 1. Related Work Taxonomy

Related Work	Algorithm	Domain
[15]	DQN	far/near-Edge
[14]	DQN	far/near-Edge
[25]	DQN	far/near-Edge
[32]	DDQN	far/near-Edge
[4]	DDPG	far/near-Edge
[18]	DDPG	far/near-Edge
[5]	MADRL	far/near-Edge
[3]	MADRL	far/near-Edge
[31]	DQN	Edge-Cloud
[10]	DQN	Edge-Cloud
[7]	DQN	Edge-Cloud
[17]	DQN	Edge-Cloud
[2]	DDQN	Edge-Cloud
[29]	DDQN	Edge-Cloud
[8]	DDPG	Edge-Cloud
[27]	MADRL	Edge-Cloud

Acknowledgments. This project has received funding from the Horizon Europe Framework Programme under Grant agreement No 101135775 (PANDORA). This paper reflects only the authors’ view and the Commission is not responsible for any use that may be made of the information it contains.

References

1. Arulkumaran, K., Deisenroth, M.P., Brundage, M., Bharath, A.A.: Deep reinforcement learning: A brief survey. *IEEE Signal Processing Magazine* **34**(6), 26–38 (2017)
2. C. Nguyen, D., N. Pathirana, P., Ding, M., Seneviratne, A.: Secure computation offloading in blockchain based iot networks with deep reinforcement learning. *IEEE Transactions on Network Science and Engineering* **8**(4), 3192–3208 (2021)
3. Cao, B., Zhang, L., Li, Y., Feng, D., Cao, W.: Intelligent offloading in multi-access edge computing: A state-of-the-art review and framework. *IEEE Communications Magazine* **57**, 56–62 (03 2019)
4. Chen, J., Xing, H., Xiao, Z., Xu, L., Tao, T.: A drl agent for jointly optimizing computation offloading and resource allocation in mec. *IEEE Internet of Things Journal* **PP**, 1–1 (05 2021)
5. Chen, X., Jiao, L., Li, W., Fu, X.: Efficient multi-user computation offloading for mobile-edge cloud computing. *IEEE/ACM Transactions on Networking* **24** (10 2015)
6. Chourlias, A., Theodoropoulos, T., Violos, J., Leivadreas, A., Tserpes, K., Zalachoris, C.K.: A brief review of population-based methods for task offloading in cloud-to-edge continuum. In: 2023 IEEE 12th International Conference on Cloud Networking (CloudNet). pp. 448–453. IEEE (2023)
7. Cui, T., Yang, R., Fang, C., Yu, S.: Deep reinforcement learning-based resource allocation for content distribution in iot-edge-cloud computing environments. *Symmetry* **15**(1) (2023)
8. Dai, Y., Zhang, K., Maharjan, S., Zhang, Y.: Edge intelligence for energy-efficient computation offloading and resource allocation in 5g beyond. *IEEE Transactions on Vehicular Technology* **69**(10), 12175–12186 (2020)
9. Du, J., Kong, Z., Sun, A., Kang, J., Niyato, D., Chu, X., Yu, F.R.: Maddpg-based joint service placement and task offloading in mec empowered air-ground integrated networks. *IEEE Internet of Things Journal* **11**(6), 10600–10615 (2024)
10. Fang, C., Zhang, T., Huang, J., Xu, H., Hu, Z., Yang, Y., Wang, Z., Zhou, Z., Luo, X.: A drl-driven intelligent optimization strategy for resource allocation in cloud-edge-end cooperation environments. *Symmetry* **14**(10) (2022)
11. Gong, Y., Yao, H., Wang, J., Wu, D., Zhang, N., Richard Yu, F.: Decentralized edge intelligence-driven network resource orchestration mechanism. *IEEE Network* **37**(2), 270–276 (2023)
12. Gronauer, S., Diepold, K.: Multi-agent deep reinforcement learning: a survey. *Artificial Intelligence Review* pp. 1–49 (2022)
13. Hortelano, D., de Miguel, I., Barroso, R.J.D., Aguado, J.C., Merayo, N., Ruiz, L., Asensio, A., Masip-Bruin, X., Fernández, P., Lorenzo, R.M., Abril, E.J.: A comprehensive survey on reinforcement-learning-based computation offloading techniques in edge computing systems. *Journal of Network and Computer Applications* **216**, 103669 (2023)
14. Huang, L., Bi, S., Zhang, Y.J.: Deep reinforcement learning for online computation offloading in wireless powered mobile-edge computing networks. *IEEE Transactions on Mobile Computing* **PP**, 1–1 (07 2019)
15. Huang, L., Feng, X., Zhang, C., Qian, L., Wu, Y.: Deep reinforcement learning-based joint task offloading and bandwidth allocation for multi-user mobile edge computing. *Digital Communications and Networks* **5** (10 2018)

16. Islam, A., Debnath, A., Ghose, M., Chakraborty, S.: A survey on task offloading in multi-access edge computing. *Journal of Systems Architecture* **118**, 102225 (2021)
17. Khayyat, M., Elgendy, I.A., Muthanna, A., Alshahrani, A.S., Alharbi, S., Koucheryavy, A.: Advanced deep learning-based computational offloading for multilevel vehicular edge-cloud computing networks. *IEEE Access* **8**, 137052–137062 (2020)
18. Liu, R., Liu, X., Wang, S., Yin, C.: Deep deterministic policy gradient based computation offloading in wireless-powered mec networks. In: 2020 IEEE Globecom Workshops (GC Wkshps). pp. 1–6 (2020)
19. Mao, Y., You, C., Zhang, J., Huang, K., Letaief, K.: A survey on mobile edge computing: The communication perspective. *IEEE Communications Surveys & Tutorials* **PP**, 1–1 (08 2017)
20. Roderick, M., MacGlashan, J., Tellex, S.: Implementing the deep q-network. arXiv preprint arXiv:1711.07478 (2017)
21. Rosendo, D., Costan, A., Valduriez, P., Antoniu, G.: Distributed intelligence on the edge-to-cloud continuum: A systematic literature review. *Journal of Parallel and Distributed Computing* **166**, 71–94 (2022)
22. Singh, P.K., Nandi, S.K., Nandi, S.: A tutorial survey on vehicular communication state of the art, and future research directions. *Vehicular Communications* **18**, 100164 (2019)
23. Staudemeyer, R.C., Morris, E.R.: Understanding lstm—a tutorial into long short-term memory recurrent neural networks. arXiv preprint arXiv:1909.09586 (2019)
24. Tan, H.: Reinforcement learning with deep deterministic policy gradient. In: 2021 International Conference on Artificial Intelligence, Big Data and Algorithms (CAIBDA). pp. 82–85. IEEE (2021)
25. Tang, D., Zhang, X., Li, M., Tao, X.: Adaptive inference reinforcement learning for task offloading in vehicular edge computing systems. In: 2020 IEEE International Conference on Communications Workshops (ICC Workshops). pp. 1–6 (2020)
26. Tang, M., Wong, V.W.: Deep reinforcement learning for task offloading in mobile edge computing systems. *IEEE Transactions on Mobile Computing* **21**(6), 1985–1997 (2020)
27. Theodoropoulos, T., Makris, A., Korontanis, I., Tserpes, K.: Greenkube: Towards greener container orchestration using artificial intelligence. In: 2023 IEEE International Conference on Service-Oriented System Engineering (SOSE). pp. 135–139 (2023)
28. Van Hasselt, H., Guez, A., Silver, D.: Deep reinforcement learning with double q-learning. In: Proceedings of the AAAI conference on artificial intelligence. vol. 30 (2016)
29. Wang, Y., Ge, H., Feng, A., Li, W., Liu, L., Jiang, H.: Computation offloading strategy based on deep reinforcement learning in cloud-assisted mobile edge computing. In: 2020 IEEE 5th International Conference on Cloud Computing and Big Data Analytics (ICCCBDA). pp. 108–113 (2020)
30. Zhang, J., Hu, X., Ning, Z., Ngai, E.C.H., Zhou, L., Wei, J., Cheng, J., Hu, B.: Energy-latency tradeoff for energy-aware offloading in mobile edge computing networks. *IEEE Internet of Things Journal* **5**(4), 2633–2645 (2017)
31. Zheng, X., Li, M., Chen, Y., Guo, J., Alam, M., Hu, W.: Blockchain-based secure computation offloading in vehicular networks. *IEEE Transactions on Intelligent Transportation Systems* **22**(7), 4073–4087 (2021)
32. Zhou, H., Jiang, K., Liu, X., Li, X., Leung, V.C.M.: Deep reinforcement learning for energy-efficient computation offloading in mobile-edge computing. *IEEE Internet of Things Journal* **9**(2), 1517–1530 (2022)