Mesopotamian journal of Big Data
Vol. **(2022)**, 2022, **pp**. 44–50
DOI: https://doi.org/10.58496/MJBD/2022/006 ISSN: 2958-6453
https://mesopotamian.press/journals/index.php/BigData

Research Article

A Survey on Distributed Reinforcement Learning

Maroning Useng[1,*], Suleiman Avdulrahman[2]

[1] *Department of Data Science and Analytics, Fatoni University, Pattani, Thailand*

[2] *Center for Atmospheric Research Nigeria , ICT university , Abuja , Nigeria*

ARTICLE INFO

Article History

Received 12 Sep 2022
Accepted 16 Oct 2022
Published 23 Nov 2022

Keywords

Big Data

Distributed Computing

Reinforcement Learning

DRL

ABSTRACT

In many settings, reinforcement learning (RL) has proven to be an effective tool for tackling difficult decision-making challenges. Traditional RL algorithms, on the other hand, frequently hit walls when confronted with issues of a sufficiently great scale or complexity. Distributed reinforcement learning (DRL) is a new area of study that hopes to circumvent these restrictions by dividing the learning workload among several computers. In this work, we offer a thorough overview of DRL, discussing its history, difficulties, applications, evaluation, scalability, and outstanding issues. We classify DRL approaches and frameworks and examine their similarities and differences. We also highlight the difficulties and restrictions of using DRL in real-world circumstances and examine its practical applicability in a variety of fields. We also describe current trends and future directions for evaluating DRL algorithms, and we analyse the performance of DRL algorithms on benchmark tasks. We also go through various options for distributed computing in DRL, as well as other methods for increasing the scalability and efficiency of DRL algorithms. We conclude by outlining key concerns and obstacles in DRL study, and by making suggestions for moving the subject forward. Overall, the goal of this survey is to give readers a picture of where things stand in terms of DRL study and application at the moment.

1. INTRODUCTION

In robotics, games, and finance, just to name a few, reinforcement learning (RL)[1] has proven to be a very effective method for handling difficult decision-making problems. Traditional RL algorithms, on the other hand, frequently hit walls when confronted with issues of a sufficiently great scale or complexity. To overcome these constraints, researchers have begun looking at distributed reinforcement learning (DRL)[2], which uses several agents or machines to carry out the learning process. DRL's ability to help scale RL algorithms and find answers to previously intractable issues has garnered a lot of interest in recent years.

This study aims to provide a thorough overview of DRL by discussing its history, difficulties, applications, assessment, scalability, and outstanding issues. Researchers and practitioners in the field of RL can benefit from the survey results by learning more about the present state of the art in DRL research and gaining insight into possible directions for future study. The study of distributed reinforcement learning (DRL) has become increasingly popular in recent years. Traditional reinforcement learning algorithms have limitations in terms of scalability and complexity, which is why DRL is being studied so intensively. DRL's scalability and speed of learning are both enhanced by the fact that the learning process is distributed among several agents or processors.

Robotics, video games, banking, healthcare, and transportation are just a few of the many real-world uses of DRL[3]. DRL has been implemented in a wide variety of contexts, including robotics, finance, and autonomous vehicles. These examples show how DRL can be used to boost performance and security in a variety of settings. Further, DRL can shed light on the cognitive processes of living things. Researchers can learn more about how biological organisms learn and possibly create new treatments for illnesses that influence learning and decision-making by analysing the behaviour of DRL algorithms.

DRL's capacity to address the limitations of classic RL algorithms, its many practical applications, and its promise to shed light on the ways in which biological organisms learn and make decisions all contribute to its significance and motivate further research. By expanding DRL, we can create better learning algorithms that can take on tough challenges in a wide range of fields.

*Corresponding author. Email: u.maroning@ftu.ac.th

The paper will proceed as described below. In Section 2, we introduce RL and discuss the strengths and weaknesses of the more common RL algorithms. The difficulties of DRL are discussed in Section 3. In Section 3, we give a taxonomy of DRL approaches and frameworks, and in Section 4, we compare and contrast several DRL methods. We highlight the difficulties and restrictions of using DRL in real-world contexts and analyse their implications in Section 5. In Section 6, we explore current trends and future directions for evaluating DRL algorithms, and we evaluate the performance of DRL algorithms on benchmark tasks. Section 7 explains how distributed computing in DRL can be used to make DRL algorithms more scalable and efficient. Finally, in Section 8, we detail the most pressing problems plaguing DRL studies and make suggestions for moving the field forward. This survey intends to contribute to the development of DRL by identifying relevant research topics and outstanding problems, and it provides a thorough overview of the present state of the art in DRL research and its applications.

2. BACKGROUND

Decision-making through experience is the subject of reinforcement learning (RL), a branch of machine learning. In RL, a learner takes activities to maximise a cumulative reward signal by changing the world around it. Amazing progress has been made in a variety of fields using RL, such as gaming, robotics, and finance.

However, traditional RL algorithms[4] have difficulty dealing with situations of a sufficiently great scale or complexity. Traditional RL algorithms' computational and memory needs grow exponentially with the size of the problem space. In addition, actual outcomes can be hard to come by because the learning process is often slow and inefficient in complicated contexts. Distributed reinforcement learning (DRL) is an effort to share the load of learning across numerous agents or machines in an effort to overcome these restrictions. DRL has the potential to allow RL algorithms to scale to larger datasets and provide solutions to otherwise intractable situations.

The field of DRL[5, 6] has received a lot of attention recently, and several new methods and frameworks have been presented as a result. OpenAI Gym, a widely used RL framework, includes support for Ray-based distributed RL. Another well-liked strategy for DRL is the parameter server architecture, in which numerous agents take their cues from one master server. Actor-critic methods, in which several agents interact with the environment and learn from each other's experiences, and federated learning, in which agents learn from their local data and share the learned model with a central server, are two more approaches. To provide an overview of the state-of-the-art and indicate future research paths in DRL, numerous surveys and evaluations have been done. For instance, [7] recent review presents an in-depth summary of the difficulties and solutions in DRL, with an emphasis on the importance of communication and synchronisation in distributed learning. Another comprehensive review is [8], who classify DRL approaches and frameworks and go over their uses and drawbacks.

While these studies shed light on important questions in DRL, they do not give comprehensive coverage of the topic. The goal of this work is to present a high-level overview of DRL, covering its history, difficulties, applications, assessment, scalability, and outstanding issues. We also provide a comparison of the various DRL methods and frameworks and present a taxonomy of DRL methods and frameworks.

3. DISTRIBUTED REINFORCEMENT LEARNING

Reinforcement learning (RL) is a powerful technique for training AI agents to make decisions in complex environments. However, traditional RL can be slow, especially for complex tasks that require a lot of experience. This is where Distributed Reinforcement Learning (DRL) comes in. DRL tackles the challenge of slow training by distributing the learning process across multiple machines. Imagine a team of agents simultaneously exploring an environment, each one gathering experiences. In DRL, these experiences are then shared and used to train a central policy, which guides the actions of all the agents. To facilitate learning across a larger set of agents or computers, researchers in the field of reinforcement learning (RL) developed the concept of distributed reinforcement learning (DRL)[7, 8]. Since DRL may be scaled up to enable faster learning, it may be able to address difficult issues that regular RL algorithms cannot. In DRL, numerous agents that are capable of autonomous learning and environmental interaction must work together. A local observation is received by each agent, the agent acts in accordance with its policy, and the agent is rewarded with a signal from the environment. Agents use the rewards they earn to inform policy changes and to disseminate their knowledge to other agents. The agents will keep learning until they reach a consensus on the best course of action.

The exploration-exploitation trade-off, non-stationary environments, and the requirement for constant communication and synchronisation are only a few of the issues plaguing DRL[9]. We explore the many DRL methods and frameworks that have been offered by researchers to overcome these obstacles. One common method for DRL[10] is the parameter server architecture, in which numerous agents take their cues from one master server. Parameters are updated on a global scale when agents report their experiences and policy gradients to a central server. The agents are then prompted to adjust their policies based on the new parameters. The parameter server structure allows for asynchronous education and decreases the burden of communication between nodes.

Another method for DRL is federated learning, in which agents learn independently but then share their collectively-learned models with a master server. The server collects the agents' models and uses them to revise the master model. Since federated learning eliminates the need for agents to share data, privacy problems are mitigated, and decentralised learning is made possible. Multiple agents interact with the environment and learn from each other's experiences in what is known as an actor-critic method of DRL. An actor network, which learns the policy, and a critic network, which learns the value function, are both present in each agent. The agents communicate with one another, comparing policy and value predictions, and adapting their networks accordingly. The exploration-exploitation trade-offs can be minimised, and cooperative learning can take place, via actor-critic approaches.

DRL algorithms include a large number of agents and machines, making it difficult to evaluate and scale them. Standard measures of DRL performance include mean reward, convergence rate, and learning stability. The number of agents, communication overhead, and available computer resources are all elements that affect the scalability of DRL algorithms. The field of DRL has a number of unanswered questions and potential future developments. Improved generalisation and transfer learning; integration of DRL with other learning paradigms including supervised and unsupervised learning; and the development of more efficient and scalable DRL algorithms all fall under this category. If these issues are resolved, DRL will be able to take on much more difficult tasks across several fields.

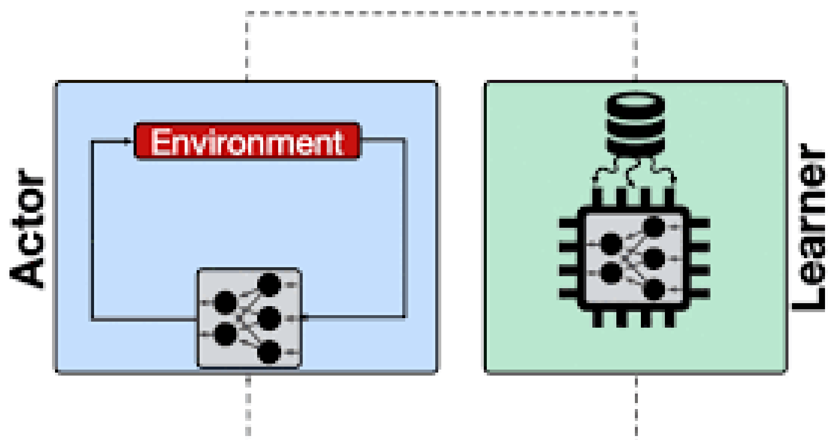

Fig. 1. Reinforcement Learning

4. APPLICATIONS OF DRL

DRL has been successfully applied to a wide range of domains, including robotics, gaming, finance, and healthcare. In this section, we discuss some of the notable applications of DRL.

- Robotics

DRL has shown promising results in robotics, where it has been used for tasks such as grasping, locomotion, and manipulation. DRL algorithms enable robots to learn complex skills from scratch, without the need for human programming. For example, DRL has been used to train a robot to play table tennis, where the robot learned to control its movements and predict the trajectory of the ball.

- Gaming

Gaming is another domain where DRL has shown remarkable results. DRL algorithms have been used to train agents to play classic games such as Atari and Go. These agents have achieved superhuman performance, outperforming even the best human players. DRL has also been used to develop new games, where the agents learn the rules and strategies of the game from scratch.

- Finance

DRL has also been applied to finance, where it has been used for tasks such as portfolio management, algorithmic trading, and risk management. DRL algorithms enable agents to learn complex trading strategies from historical data and adapt to changing market conditions. For example, DRL has been used to develop an algorithmic trading system that achieved higher returns than traditional trading algorithms.

- Healthcare

DRL has also shown potential in healthcare, where it has been used for tasks such as disease diagnosis, drug discovery, and personalized treatment. DRL algorithms enable agents to learn from large-scale medical data and provide personalized recommendations to patients. For example, DRL has been used to develop a personalized treatment plan for patients with Parkinson's disease, where the agent learned to adjust the dosage of medication based on the patient's symptoms.

5. EVALUATION AND PERFORMANCE ANALYSIS

Evaluating the performance of DRL algorithms is essential to assess their effectiveness and compare them with other approaches. In this section, we discuss some of the common evaluation metrics and performance analysis techniques used in DRL.

The following are some of the common evaluation metrics used to assess the performance of DRL algorithms:

- Reward: The reward obtained by the agent for completing a task is a common metric used in DRL. The higher the reward, the better the performance of the agent.
- Success rate: The success rate measures the percentage of times the agent successfully completes the task. It is a useful metric when the goal is to achieve a specific task.
- Exploration rate: The exploration rate measures the percentage of time the agent spends exploring new actions instead of exploiting known actions. A higher exploration rate can lead to better performance in the long run but may result in lower short-term rewards.
- Convergence rate: The convergence rate measures how quickly the agent converges to an optimal policy. A faster convergence rate is desirable as it leads to faster learning.

The following are some of the common performance analysis techniques used in DRL:

- Learning curves: Learning curves show the performance of the agent over time as it learns from experience. They are useful for assessing the effectiveness of the algorithm and identifying areas for improvement.
- Hyperparameter tuning: DRL algorithms often have many hyperparameters that need to be tuned to achieve optimal performance. Hyperparameter tuning involves testing different combinations of hyperparameters and selecting the best performing one.
- Visualization: Visualizing the behavior of the agent can provide insights into its learning process and help identify areas for improvement. For example, visualizing the action-value function can reveal which actions are most valuable in different states.
- Ablation study: An ablation study involves testing the performance of the agent with different components removed or modified. It can help identify which components are essential for achieving optimal performance.

Evaluating the performance of DRL algorithms is crucial for assessing their effectiveness and improving their performance. By using appropriate evaluation metrics and performance analysis techniques, researchers can gain insights into the strengths and weaknesses of different DRL algorithms and identify ways to improve their performance. Number equations consecutively.

6. SCALABILITY AND EFFICIENCY OF DRL

The practical implementation of DRL algorithms relies heavily on their scalability and efficiency. In this section, we explore some of the obstacles and solutions to DRL's scalability and efficiency problems. Some difficulties in increasing DRL algorithms' throughput and scalability include the following. DRL algorithms can be computationally costly, necessitating a large amount of computing in order to train the agents. The algorithm's scalability and performance may be impaired as a result of this. In distributed DRL, communication between agents might be a bottleneck, especially if they are located in different physical locations. This might cause delays and lower productivity. In a distributed setting, it can be difficult to provide the quantity of memory, disc space, and computing power that DRL algorithms require.

The following are some methods for making DRL algorithms more efficient and scalable:

Scalability and efficiency of DRL algorithms can be greatly increased by computing them in parallel. Data parallelism, model parallelism, and pipeline parallelism are all methods that can be used to accomplish this goal. Through the use of distributed computing, the computational weight of DRL algorithms can be spread across numerous machines, allowing for the usage of larger datasets. Parameter servers, federated learning, and distributed reinforcement learning are among methods that can help with this. DRL models can be made smaller through the use of model compression techniques without suffering a noticeable drop in performance. DRL algorithms may need less storage space and RAM as a result of this. Hardware acceleration: Methods like graphics processing units (GPUs) and tensor processing units (TPUs) can greatly improve the efficiency of DRL algorithms by accelerating their computation.

Robotics, video games, the financial sector, and even healthcare have all benefited from the implementation of scalable and effective DRL algorithms. Robots have been trained to do complicated tasks like grasping and manipulation with the help of effective DRL algorithms. Algorithmic trading and portfolio management are two financial applications that have made use of scalable DRL algorithms. The practical implementation of DRL algorithms relies heavily on their scalability and efficiency. Scalability and efficiency of DRL algorithms can be enhanced through the use of relevant techniques including parallelization, distributed computing, model compression, and hardware acceleration, allowing for their implementation in practical settings.

7. CHALLENGES AND OPEN PROBLEMS

Despite the recent advances in DRL, there are still many challenges and open problems that need to be addressed. In this section, we discuss some of the most significant challenges and open problems in DRL.

1. Scalability One of the most significant challenges in DRL is scalability. While distributed DRL can help address this issue to some extent, there are still many open problems in this area. For example, how can we scale DRL algorithms to handle extremely large datasets or highly complex environments? How can we minimize communication overhead and ensure efficient use of resources?
2. Exploration Another significant challenge in DRL is exploration and exploitation. DRL algorithms often require a significant amount of exploration to learn an optimal policy, but excessive exploration can lead to high computational and time costs. How can we balance exploration and exploitation in DRL algorithms to achieve optimal performance while minimizing the computational and time costs?
3. Generalization is another important challenge in DRL. DRL algorithms often require a large number of training samples to learn an optimal policy, but the policy may not generalize well to new, unseen environments. How can we improve the generalization performance of DRL algorithms?
4. Safety is an important concern in many DRL applications, such as robotics and healthcare. How can we ensure that DRL agents behave safely in these applications? How can we design DRL algorithms that are robust to uncertainties and adversarial attacks?
5. Explainability is another important challenge in DRL. DRL algorithms can learn complex policies that are difficult to interpret, making it challenging to understand how the algorithm arrived at a particular decision. How can we design DRL algorithms that are transparent and explainable?
6. Transfer Learning is an important problem in DRL, particularly for applications where training data is limited or expensive to obtain. How can we leverage knowledge from previous tasks to improve the learning performance of DRL algorithms? How can we design DRL algorithms that can transfer knowledge between tasks efficiently? DRL has made significant progress in recent years, but there are still many challenges and open problems that need to be addressed. By addressing these challenges and open problems, researchers can further improve the scalability, efficiency, safety, and generalization performance of DRL algorithms and enable their use in real-world applications.

8. CONCLUSION

Finally, distributed reinforcement learning (DRL) stands out as a potentially game-changing field that could completely alter the way in which difficult problems that need for extensive thought are approached. This survey provides a thorough introduction to DRL, including the field's basic concepts, various approaches, and real-world implications. As we delve deeper into DRL's potential, we also shine a light on the challenges and open questions that now plague this field. Significant difficulties exist that researchers and practitioners must overcome, such as scalability, exploration and exploitation trade-offs, generalization capabilities, safety issues, explainability, and transfer learning. Despite these limitations, DRL has shown encouraging outcomes in numerous domains, including robotics, gaming, finance, and healthcare. If we work to increase our understanding of DRL and tackle the outstanding challenges, we can unlock the technology's full potential. It is clear that future breakthroughs in AI and other fields will be enabled by ongoing research and innovation in DRL. Persevering and working together to overcome obstacles is the key to realizing DRL's transformative potential for addressing difficult, real-world situations.

Funding

The author's paper explicitly states that no funding was received from any institution or sponsor.

Conflicts of Interest

The authors declare that there is no conflict of interests regarding the publication of this paper.

Acknowledgment

The authors would like to express their gratitude to the Department of Data Science and Analytics, Fatoni University for their moral support. Please accept my sincere gratitude for the useful recommendations and constructive remarks provided by the anonymous reviewers.

References

[1] G. Weiß, "Distributed reinforcement learning," in *The Biology and technology of intelligent autonomous agents*, 1995, pp. 415-428: Springer.

[2] E. Liang *et al.*, "RLlib: Abstractions for distributed reinforcement learning," in *International Conference on Machine Learning*, 2018, pp. 3053-3062: PMLR.

[3] A. H. Ali, "A survey on vertical and horizontal scaling platforms for big data analytics," *International Journal of Integrated Engineering*, vol. 11, no. 6, pp. 138-150, 2019.

[4] A. H. Ali and M. Z. Abdullah, "Recent trends in distributed online stream processing platform for big data: Survey," in *2018 1st Annual International Conference on Information and Sciences (AiCIS)*, 2018, pp. 140-145: IEEE.

[5] A. H. Ali and M. Z. Abdullah, "A novel approach for big data classification based on hybrid parallel dimensionality reduction using spark cluster," *Computer Science*, vol. 20, no. 4, 2019.

[6] A. H. Ali and M. Z. Abdullah, "An efficient model for data classification based on SVM grid parameter optimization and PSO feature weight selection," *International Journal of Integrated Engineering*, vol. 12, no. 1, pp. 1-12, 2020.

[7] M. Littman and J. Boyan, "A distributed reinforcement learning scheme for network routing," in *Proceedings of the international workshop on applications of neural networks to telecommunications*, 2013, pp. 55-61: Psychology Press.

[8] S. Kapturowski, G. Ostrovski, J. Quan, R. Munos, and W. Dabney, "Recurrent experience replay in distributed reinforcement learning," in *International conference on learning representations*, 2019.

[9] M. W. Hoffman *et al.*, "Acme: A research framework for distributed reinforcement learning," *arXiv preprint arXiv:2006.00979*, 2020.

[10] J. Hu, H. Zhang, L. Song, R. Schober, and H. V. Poor, "Cooperative internet of UAVs: Distributed trajectory design by multi-agent deep reinforcement learning," *IEEE Transactions on Communications*, vol. 68, no. 11, pp. 6807-6821, 2020.