Proceedings of the 2023 International Conference on Machine Learning and Automation
DOI: 10.54254/2755-2721/40/20230641

Research advanced in the integration of federated learning and reinforcement learning

Chengan Li

Nanjing University of Aeronautics and Astronautics, Nanjing, 211106, China

lchnan@nuaa.edu.cn

Abstract. Reinforcement learning (RL) and federated learning (FL) are two important machine learning paradigms. Reinforcement learning is concerned with enabling intelligence to learn optimal policies when interacting with an environment, while federated learning is concerned with collaboratively training models on distributed equipment while preserving data privacy. In recent years, the fusion and complementarity of reinforcement learning, and federated learning have attracted increasing research interest, providing new directions for the development of the machine learning community. Focusing on the integration of reinforcement learning and federated learning, this paper introduces in detail the latest technological developments in the integration of reinforcement learning and federated learning, and discusses the main challenges, existing methods and future directions of this intersection. Specifically, based on the introduction of classical reinforcement learning and federated learning. In addition, this document introduces cutting-edge results on the integration of reinforcement learning and joint learning and discusses the problems and future directions of the integration.

Keywords: Machine Learning, Federated Learning, Reinforcement Learning, Fusion.

1. Introduction

The speedy development of hardware equipment and the continuous accumulation of multimedia data provide an important foundation for large-scale model learning, which has made continuous breakthroughs in machine learning and deep learning technology in recent years. As popular research methods in the machine learning community, reinforcement learning, and federated learning have attracted a large amount of research interest and have become two integral components in the broad field.

Reinforcement learning, a method of learning optimal strategies by interacting with the environment, has been remarkably successful in many fields such as gaming, robotics, and finance. The core idea is that intelligences continuously try and learn by interacting with their environments to achieve the goal of maximizing cumulative rewards. Successful examples of reinforcement learning include AlphaGo, which beat the world Go champion through deep reinforcement learning. In addition, reinforcement learning has been used in autonomous driving, energy management, and many other fields, demonstrating its potential for a wide range of applications. Unlike reinforcement learning, which focuses on intelligent decision making. Federated learning was proposed as a result of the growing concern for data privacy and security in the modern world. In traditional centralized learning, all data is sent to a central server for training. However, this can expose sensitive information such as personal

DOI: 10.54254/2755-2721/40/20230641

health records, financial transactions, etc. Federated learning protects data privacy by training models on local devices and sharing only model parameters. Federated learning has been implemented in many real-world applications, such as smartphone keyboard prediction, medical data analysis, etc. It solves the problem of data silos and facilitates cross-organizational and cross-domain collaboration while ensuring data security and compliance.

The integration of federated learning and reinforcement learning is an emerging research direction. By applying reinforcement learning algorithms in distributed environments and sharing strategies by means of federated learning, broader and more efficient learning can be achieved. For example, in intelligent transportation systems, through federated reinforcement learning, individual vehicles can share their learned strategies to work better together and improve transportation efficiency. In healthcare, different hospitals can share reinforcement learning models to better diagnose and treat diseases while protecting patient privacy. However, the integration of federated learning and reinforcement learning also face lots of challenges. Issues such as algorithmic efficiency, data privacy, and communication constraints need to be thoroughly researched and resolved. In addition, there are many complex ethical and legal issues involved in this area, such as data ownership, liability of intelligences, etc.

Overall, the convergence of federated learning and reinforcement learning opens a promising new field that offers new solutions to many practical problems. By combining the strengths of both, broader and more efficient learning can be achieved while protecting data privacy. Research at this intersection will not only help advance science and technology, but may also have far-reaching social, economic and cultural implications. This document is intended to be a comprehensive synopsis of this intersectional domain, to promote further research and exploration, and to pave the way for future technological innovations and societal development. This study summarizes the characteristics of FL, but challenges remain in the various solutions made for the integration of RL and FL. Based on a general definition, this paper examines pertinent research on federated learning, identifies the frontiers of the literature addressing the obstacles between the two technologies, and contributes to a deeper comprehension of this convergence for potential future optimization. In Section 2, the author go discuss FL and RL in addition to the introduction. In addition, we summaries, in particular, circumstantial disclosure of information in Federated Leaning and current methods of privacy preservation in Federated Leaning. In Section 3, I identify four important challenges to the convergence of these two techniques. Section 4 gives some cutting-edge results and a discussion around them, describes some promising directions for federated learning and reinforcement learning, and provides guidance for future developments.

2. Related Work

2.1. Federated Learning

In the past, smartphones, wearables, smart cars, and other smart endpoints were still just devices generating massive amounts of data in a distributed network, and they were not being utilized. As the computational and storage power of devices in a distributed network grows, it is possible to take advantage of the powerful local resources on each device. The more data there is, the better the trained model will be. Machine learning has grown from its development to the present day, and more and more people are becoming concerned about the importance of privacy. Machine Learning from its development to date, more and more people are becoming concerned about the importance of privacy. The data belongs to the user, we can neither use it without making an application, but equally need to safeguard the privacy of the data. How to securely and efficiently achieve cooperation with sensitive data is a popular discussion topic in recent years, especially in some very sensitive industries, such as banks, confidential units. In this context, federated learning has emerged, It enables model training over distributed nodes without requiring data to be centralized in one place, thus enabling value mining of data while protecting privacy. The central concept of federated learning is to share weights and updates for model training among nodes instead of raw data through distributed computing. This mechanism not only helps to protect individual privacy, but also makes full use of the local data characteristics of each node to provide a richer perspective for global model training.

DOI: 10.54254/2755-2721/40/20230641

Federated learning is a distribution learning schema that allows more than one contributor to train a model together without sharing the original database. However, even in the context of federal learning, there may be a risk of indirect information leakage. Indirect information leakage, on the other hand, refers to inferring properties of the original data by analyzing model parameters, gradients, or other information related to the learning process. We usually categorize them into model parameter leakage, gradient leakage, membership inference attack, and model inversion attack. Model parameter leakage means that by analyzing the trained model parameters, the attacker can rely on these to infer some information about the original data. Gradient leakage means that an attacker takes advantage of the fact that gradient information is often shared between clients to analyze this gradient information to imply the initial information. In the specialized domain of membership inference attacks, the adversary zealously seeks to determine whether a specific data fragment holds the status of an 'insider' in the training dataset of the model under scrutiny, particularly one designed for classification tasks [1]. Within the framework of model inference attacks, the infiltrator harnesses the trained algorithm to reverse-engineer and glean insights into the foundational data from which the model was originally sculpted [2].

Existing privacy-preserving methods play a key role in federated learning to mitigate the risk of indirect information leakage. Among them, differential privacy prevents the inference of individual data by analyzing model outputs by adding noise to the data [3]; homomorphic encryption allows computation over ciphertexts to ensure data confidentiality during model training [4]; and secure multiparty computation enables multiple participants to co-compute a function without having to share their respective inputs, thus enabling model training without disclosing the original data [5]; In addition, limiting model complexity through techniques such as model pruning and regularization reduces the risk of membership inference attacks by reducing overfitting to specific training samples [6]. Together, these approaches form a powerful toolset that helps achieve the effectiveness of federated learning while preserving privacy [7]. In addition, the implementation of federated learning faces many challenges, such as high cost of communication, hardware heterogeneity, software heterogeneity, and privacy and security.

2.2. Reinforcement Learning

RL is a ML parameter that the core idea is that an Agent continuously tries and learns by interacting with the Environment in order to achieve the goal of maximizing Cumulative Reward. The roots of reinforcement learning can be found in the theory of operant conditioning in psychology and was introduced into the field of artificial intelligence in the 1950s as a computational method. The basic framework consists of states (information describing the environment), actions (operations that an intelligent body can take), rewards (criteria for evaluating how good or bad an action is), and strategies (rules defining which action to take in each state). The main algorithms for reinforcement learning are value iteration, Q-learning, policy iteration, deep Q-networks (DQN), etc. In recent years, Deep Reinforcement Learning has merged deep learning and reinforcement learning to create many new application areas. Reinforcement learning has achieved remarkable success in areas such as gaming (e.g., AlphaGo), robotics, finance, and energy management. However, despite many achievements, reinforcement learning still faces challenges such as low sample efficiency, potentially unstable training processes, and collaborative learning with multiple intelligences. As a powerful learning paradigm, reinforcement learning has demonstrated its potential and value in many areas, and future research is likely to focus on improving sample efficiency, enhancing stability and reliability, and exploring multi-intelligent body collaboration to continue to push the frontiers of AI.

2.3. Federated reinforcement learning

In today's increasingly data-driven era, machine learning has become a core technology in many fields. In particular, Federated Learning (FL) and Reinforcement Learning (RL), as two important branches of machine learning, have each made significant progress in privacy protection and automated decision making. And the fusion of these two branches, we call it Federated Reinforcement Learning (FRL). A cutting-edge machine learning strategy called federated reinforcement learning combines the advantages

DOI: 10.54254/2755-2721/40/20230641

of federated learning and reinforcement learning. It allows multiple intelligences to share policies across intelligences without sharing raw data and state information by training policies on local devices and performing policy aggregation on a central server [8]. However, merging these two domains, while full of potential, presents a number of challenges and problems.

3. Challenges

This section considers some of the challenges and issues in federated reinforcement learning that are discussed.

3.1. Data Privacy and Security

Data privacy and security have grown to be major concerns, and federated learning's main objective is to safeguard data privacy, yet reinforcement learning requires a lot of data sharing and data interaction. Although part of the ambition of federated learning is to secure data privatization, there is still the potential for sensitive information to be exposed during the process of model aggregation and communication. How to ensure privacy security while ensuring learning efficiency is an important challenge [7]. Therefore, how to conduct effective learning while preserving privacy is an important challenge that needs to be thoroughly researched and designed.

3.2. Communications overhead

Communication overhead is also an issue that cannot be ignored. Federated learning requires transferring model parameters and gradients across multiple devices, which can lead to communication bottlenecks. And in a federated reinforcement learning environment, the intelligences may be distributed in different geographical locations, involving a lot of communication and collaboration [9]. How to achieve effective multi-intelligence cooperative learning and optimize communication protocols and compression techniques to reduce the communication burden while maintaining communication efficiency is the key issue [10].

3.3. Strategy Collaboration in Heterogeneous Environments

In addition, strategy collaboration in heterogeneous environments is likewise a complex and critical challenge in the convergence of federated and reinforcement learning. In federated learning, heterogeneous environments refer to environments in which different intelligences may have different properties and dynamics. These differences may come from differences in physical environments (e.g., climatic conditions in different geographic locations), differences in task goals, inconsistencies in data distribution, and so on. Among the challenges of strategic synergy fall into three different forms:

(1) When altering the configurations of state space and action space, the level of deviation experiences a significant shift. [11]. This inconsistency complicates collaborative learning between intelligences, as they may not be able to share and understand each other's experiences and strategies directly.

(2) Non-independently and identically distributed data. In environments marked by heterogeneity, intelligences may encounter data that deviates from the norms of independent and identical distribution [12]. Such non-iid data may interfere with the federated learning process, making the training of global models difficult.

(3) Communication and synchronization. Intelligences may need to interact through different communications protocols and networks. This may lead to communication delays and synchronization problems, which may affect the collaborative updating of policies.

3.4. Model synchronization and consistency, compliance and ethical

Model synchronization involves ensuring that the model parameters of all intelligences remain consistent during training. This can be achieved through either periodic synchronization or asynchronous synchronization, where periodic synchronization updates the global model at fixed intervals, while asynchronous synchronization allows the intelligences to make a certain number of

DOI: 10.54254/2755-2721/40/20230641

updates locally before synchronizing with the global model [13]. Model consistency, on the other hand, involves ensuring that models of all intelligences maintain consistent performance and behavior during training. Maintaining consistency can be challenging as different intelligences may work with different data distributions. Methods such as the joint average (FedAvg) algorithm have been proposed to ensure global model consistency [14].

3.5. Compliance and ethical considerations

Primarily, it's imperative that FL client software residing on end-user gadgets be fortified against illicit entry, cyber intrusions, and unauthorized data exfiltration either from the device itself or during data exchanges with the orchestration hub. To uphold this level of privacy, several robust methodologies are employed, such as secure data amalgamation, the principles of differential privacy, and the utilization of homomorphic encryption schemes [15]. Even though the objective of federated reinforcement learning is to acquire knowledge without exposing the raw data, there remains an imperative to perpetually safeguard information security and uphold data confidentiality. Secondly, FRL may involve multiple organizations and stakeholders, so there is a need to consider ethical and social responsibility to ensure fair and transparent use of the technology. Ethical and compliant implementation of FRL can be ensured through transparent protocols, third-party audits, and compliance assessments.

4. Cutting-edge achievements

Artificial intelligence (AI) research on the intersection of federated and reinforcement learning is growing in popularity, and its application to the problem of intelligent decision-making in dispersed environments is anticipated. Research in this area involves not only algorithmic and technological innovations, but also multiple challenges such as privacy, communication efficiency, and multi-intelligence collaboration. In recent years, researchers have achieved a series of impressive cutting-edge results that not only advance theories but also provide new possibilities for practical applications. For example, collaborative learning in heterogeneous environments provides new approaches for collaborative decision-making of distributed intelligences; privacy-preserving reinforcement learning provides safeguards for applications involving sensitive data; efficient communication protocols enable federated reinforcement learning to be realized in resource-constrained environments; and the study of multi-intelligence collaboration provides new perspectives on intelligent decision-making in large-scale and dynamic environments. Together, these cutting-edge results reveal the potential and challenges of the convergence of federated and reinforcement learning, providing rich insights and guidance for future research and applications.

Firstly, the work of Hsieh provides a fresh perspective by proposing a non-independently identically distributed (Non-IID) federated reinforcement learning approach. They call this approach FedDRL, and this approach aims to solve the problem of collaborative learning in heterogeneous environments. In traditional reinforcement learning, intelligences usually learn and make decisions in a fixed environment. However, in real-world applications, intelligences need to work in multiple heterogeneous environments that may have different data distributions, dynamics, and constraints.

In addition, privacy protection is another central issue in federated reinforcement learning. Hence the emergence of Secure Federated Learning, an approach to protect privacy and security when sharing data or models between multiple intelligences, which prevents malicious attackers from stealing or tampering with the data or models, thus ensuring the interests and trust of each intelligence. The challenge of secure federated reinforcement learning lies in designing effective encryption, authentication, and motivation methods to address issues such as sensitivity, integrity, and reliability of data or models. For example, FedCoin is a blockchain-based federated reinforcement learning method for decentralized data sharing and model updating. The method uses a smart contract-based incentive mechanism for rewarding agents that contribute data or models and punishing agents that cheat or attack. Meanwhile, Zhu and his team explored the application of differential privacy in federated reinforcement learning. They proposed a new learning algorithm which can achieve effective learning while protecting data privacy. The significance of this work is that it not only solves the privacy problem in federated learning, but also

DOI: 10.54254/2755-2721/40/20230641

ensures the efficiency and accuracy of learning [1]. This opens up the possibility of using federated reinforcement learning in the future in healthcare, finance and other applications involving sensitive data.

Overall, now have a new viewpoint on the challenge of intelligent decision making in remote environments thanks to research at the interface of federated learning and reinforcement learning. The cutting-edge results in this area demonstrate its potential and diversity, providing valuable guidance for future research and applications. As technology continues to advance, we can expect more breakthroughs and innovations in this area. Meanwhile, we extensively discussed the application of FRL in various tasks, such as intelligent transport, control optimization, etc.

4.1. Intelligent Transportation System (ITS)

The proliferation of vehicles on urban roads has led to increased traffic congestion, longer travel times, and an increase in the number of road accidents [16]. Conventional traffic management systems often rely on predefined rules, which may not be sufficient to cope with the growing complexity and unpredictability of traffic scenarios [17]. Federated reinforcement learning, on the other hand, allows multiple intelligences (e.g., vehicles or traffic signals) to learn collaboratively without directly sharing data [18]. Accordingly, each intelligent can train its model locally before sending its updated model or model parameters to a centralized server for aggregate. This approach not only protects data privacy, but also lowers communication costs, enhances the system's real-time performance, and increases reliability. Federated Reinforcement Learning has several potential applications in intelligent transport systems. Firstly, it can be used for route optimization. Traditional route planning methods typically rely on a central server to collect and analyses traffic data, and then provide optimal route recommendations for each vehicle. Federated Reinforcement Learning, on the other hand, allows vehicles to collaboratively learn how to choose the best route without directly sharing position and speed data. This not only protects driver privacy, but also allows for more accurate predictions of traffic flow and congestion, leading to more optimized route suggestions. Additionally, traffic light management can be done using federated reinforcement learning. Traditional traffic signal control methods typically rely on fixed timings and predetermined strategies. Federated Reinforcement Learning, on the other hand, allows traffic signals to adaptively adjust to real-time traffic flow and the surrounding environment [19]. This means that traffic signals can dynamically adjust the timings of traffic lights according to the actual situation, such as traffic flow, pedestrian flow and weather conditions, thus improving traffic flow and safety. Federated reinforcement learning can also be used for accident prevention and energy management [20]. Through collaborative learning between vehicles, possible traffic accidents can be predicted and avoided earlier. For electric vehicles, Federated Reinforcement Learning can help optimize battery charging and discharging strategies to extend battery life and improve energy efficiency.

4.2. Optimization of control

Federated Reinforcement Learning is gaining attention for industrial control applications, providing a way for distributed systems and devices to learn and make decisions collaboratively without directly sharing sensitive or private data. In complex production lines and automation systems, individual devices and robots can learn and optimize their control strategies locally, and push the model parameters or learnt techniques to a centralized server for update and aggregation. This approach not only improves the operational efficiency and stability of the overall system, but also ensures data privacy and security. Federated Reinforcement Learning also enables cross-device, cross-factory, cross-geographical optimization, creating new opportunities. Industrial Internet of Things (IIoT) advancement has made it possible for the federated reinforcement learning, and it may become a key technology to support large-scale, distributed industrial systems [18,19,21,23].

5. Conclusion

Federated Reinforcement Learning, as an interdisciplinary amalgamation of Federated Learning and Reinforcement Learning, holds immense potential and challenges. This paper encapsulates the

DOI: 10.54254/2755-2721/40/20230641

principles of Federated Reinforcement Learning, outlines its technological challenges and corresponding solutions, and delves into the prospective directions for its future development. On the ever-evolving path ahead, researchers and practitioners in Federated Reinforcement Learning will persist in exploring novel theories, algorithms, and applications, paving the way for increased possibilities and innovations in achieving distributed intelligent decision-making and problem-solving.

References

[1] Baluta, T., Shen, S., Hitarth, S., Tople, S., & Saxena, P. (2022, November). Membership inference attacks and generalization: A causal perspective. In Proceedings of the 2022 ACM SIGSAC Conference on Computer and Communications Security (pp. 249-262).

[2] Zuobin Ying, Yun Zhang, Ximeng Liu. "Privacy-Preserving in Defending against Membership Inference Attacks", Proceedings of the 2020 Workshop on Privacy-Preserving Machine Learning in Practice, 2020

[3] Dwork, C. (2008). Differential Privacy: A Survey of Results. Theory and Applications of Models of Computation.

[4] Gentry, C. (2009). Fully Homomorphic Encryption Using Ideal Lattices. STOC '09.

[5] Yao, A. C. (1982). Protocols for Secure Computations. 23rd Annual Symposium on Foundations of Computer Science.

[6] Shokri, R., et al. (2017). Membership Inference Attacks Against Machine Learning Models. 2017 IEEE Symposium on Security and Privacy (SP).

[7] Bonawitz, K., et al. (2019). Towards Federated Learning at Scale: System Design. Proceedings of the 2nd SysML Conference.

[8] Huang, X., Leng, S., Maharjan, S., & Zhang, Y. (2021). Multi-agent deep reinforcement learning for computation offloading and interference coordination in small cell networks. IEEE Transactions on Vehicular Technology, 70(9), 9282-9293.

[9] Chang, M., Kaushik, S., Weinberg, S. M., Griffiths, T., & Levine, S. (2020, November). Decentralized reinforcement learning: Global decision-making via local economic transactions. In International Conference on Machine Learning (pp. 1437-1447). PMLR.

[10] Sattler, F., Wiedemann, S., Müller, K. R., & Samek, W. (2019). Robust and communication-efficient federated learning from non-iid data. IEEE transactions on neural networks and learning systems, 31(9), 3400-3413.

[11] Zhao, X., Wang, L., Zhou, Y., Pan, B., Wang, R., Wang, L., & Yan, X. (2022). Energy management strategies for fuel cell hybrid electric vehicles: Classification, comparison, and outlook. Energy Conversion and Management, 270, 116179.

[12] Durrant, A., Markovic, M., Matthews, D., May, D., Enright, J., & Leontidis, G. (2022). The role of cross-silo federated learning in facilitating data sharing in the agri-food sector. Computers and Electronics in Agriculture, 193, 106648.

[13] Li, M., et al. (2014). Scaling Distributed Machine Learning with the Parameter Server. OSDI.

[14] McMahan, H. B., et al. (2016). Communication-Efficient Learning of Deep Networks from Decentralized Data. arXiv preprint arXiv:1602.05629.

[15] Truong, N., Sun, K., Wang, S., Guitton, F., & Guo, Y. (2021). Privacy preservation in federated learning: An insightful survey from the GDPR perspective. Computers & Security, 110, 102402.

[16] Elallid, B. B., Benamar, N., Hafid, A. S., Rachidi, T., & Mrani, N. (2022). A comprehensive survey on the application of deep and reinforcement learning approaches in autonomous driving. Journal of King Saud University-Computer and Information Sciences, 34(9), 7366-7390.

[17] lallid, B. B., Bagaa, M., Benamar, N., & Mrani, N. (2023, June). A reinforcement learning based approach for controlling autonomous vehicles in complex scenarios. In 2023 International Wireless Communications and Mobile Computing (IWCMC) (pp. 1358-1364). IEEE.

DOI: 10.54254/2755-2721/40/20230641

[18] Yang, Q., Liu, Y., Chen, T., & Tong, Y. (2019). Federated machine learning: Concept and applications. ACM Transactions on Intelligent Systems and Technology (TIST), 10(2), 1-19.
[19] Wiering, M. A. (2000). Multi-agent reinforcement learning for traffic light control. In Machine Learning: Proceedings of the Seventeenth International Conference (ICML'2000) (pp. 1151-1158).
[20] Li, L., Lv, Y., & Wang, F. Y. (2016). Traffic signal timing via deep reinforcement learning. IEEE/CAA Journal of Automatica Sinica, 3(3), 247-254.
[21] Riedmiller, M., Hafner, R., Lampe, T., Neunert, M., Degrave, J., Wiele, T., ... & Springenberg, J. T. (2018, July). Learning by playing solving sparse reward tasks from scratch. In International conference on machine learning (pp. 4344-4353). PMLR.
[22] Wan, J., Tang, S., Shu, Z., Li, D., Wang, S., Imran, M., & Vasilakos, A. V. (2016). Software-defined industrial internet of things in the context of industry 4.0. IEEE Sensors Journal, 16(20), 7373-7380.
[23] Li, T., Sahu, A. K., Talwalkar, A., & Smith, V. (2020). Federated learning: Challenges, methods, and future directions. IEEE signal processing magazine, 37(3), 50-60.