Privacy-Preserving AI: Leveraging Federated Reinforcement Learning in Distributed Systems

Shiva Mehta
Centre for Research Impact & Outcome,
Chitkara University Institute of Engineering and Technology
Chitkara University
Punjab, India
shiva.2387@chitkara.edu.in

Aseem Aneja
Chitkara Centre for Research and Development,
Chitkara University,
Himachal Pradesh,174103, India
aseem.aneja.orp@chitkara.edu.in

2024 International Conference on Electrical Electronics and Computing Technologies (ICEECT) | 979-8-3503-7809-2/24/$31.00 ©2024 IEEE | DOI: 10.1109/ICEECT61758.2024.10739054

***Abstract*— Organizing computing distributed resources in conjunction with machine learning algorithms in the age of big data is critical. This paper presents the FRL approach that combines Federated Learning and Reinforcement Learning to address problems in scale, speed, and data privacy. The overall framework of the Federated Reinforcement Learning (FRL) model bestows the use of distributed computing at different agents to train a standard RL model. This framework guarantees data privacy utilizing differential privacy and secure multi-party computation (SMPC). Quantitative analysis proves the proposed FRL framework retrieves a better result than a standard RL strategy implemented in the central controller. In the 100th episode of the study, the authors' FRL framework had better results than the centralized RL approach in autonomous navigation, resource management and game-playing. Thus, the FRL framework gained a total benefit of 95 cumulative. 6, 98. 4, and 98. 3, and the rewards obtained for the centralized RL approach were 90. 5, 92. 0, and 93. 8, respectively. Also, there was lower communication overhead about the generic FRL framework, where the assessment was recorded at 33. At the same time, 100 parameter exchanges were performed; the decentralized RL method was measured at 6 units, while the centralized RL method was measured at 41. 3 units. The joint plot analysis established that the traditional FRL framework training strategy offers quicker convergence and heightened model robustness, as the collaborative training approach provides. The heatmap visualization confirms the enhancement of the capability to control the communication overhead.**

***Keywords*— *data privacy, machine learning, fed avg approach, security, federated learning, privacy preservation*.**

I. Introduction

Using distributed computing resources and sophisticated machine learning algorithms is a subject of interest in the era of big data and AI. Such activities require fast data processing and analysis simultaneously, and the demand for confidentiality and security has led to the emergence of practical approaches[1]. Two major paradigms that have contributed significantly to this development are Federated Learning (FL), a decentralized machine learning procedure in which numerous devices or nodes collaborate to build a standard model. At the same time, each retains its data[2]. This method effectively solves privacy problems because the data does not go anywhere out of a particular space; it just exchanges updates concerning the model. FL is most relevant when data is distributed across several sites and sources like mobiles, edge servers or corporate databases. It can be used effectively in conjunction with other datasets for data analytics but simultaneously achieves user anonymity and data security [3].

On the other hand, a class of machine learning known as Reinforcement Learning (RL) involves making choices in a dynamically changing environment. Learning reinforcement methods allows agents to develop the correct strategies for interacting with the world by perceiving the consequences as a reward or punishment [4]. This methodology has also been implemented in several fields, including robots, self-driving cars, game AI, and resource allocation. RL is a stunning tool for solving practical questions because of its versatility and capability to handle complex decision-making procedures [5]. These studies have two paradigms: decentralized reinforcement learning and federated reinforcement learning, which combine the decentralized reinforcement learning paradigm with the classical federated learning paradigm and offer a feasible approach to solve the problem of effective and private learning using distributed computing resources. Federated Reinforcement Learning (FRL) is the extension of federated learning and reinforcement learning that enables multiple independent learners to train an RL model [6]. Each agent learns from its context of operations, shares the experience it has gained, and changes to the global model during its operation. The ability of students to work in teams is the key to achieving better and more developed reinforcement learning models; besides, sensitive information may be shared safely. FRL solves many vital challenges in today's artificial intelligence applications using the prospects of distributed computational resources [7]. It enhances scalability because computation is distributed among numerous agents and decreases the agents' tendency to breach data privacy via localized data storage; it also improves the reinforcement learning models' generalization due to the collaborative training process [8]. FRL can revolutionise several domains, including financial cities, since it allows for correct and privately considerate decision-making models. This paper develops a fresh new framework called Federated Reinforcement Learning (FRL) that enables a coordinated use of various remote computing resources for controlling significant challenges in today's AI usage, including data privacy, extensiveness, and computation speed [9]. It is a compound of FL and RL in which several autonomous agents can cooperatively train RL but never share their local data. This research strives to enhance performance in complex environments that handle large amounts of data by providing a comprehensive architectural design [10]. It also discusses the issues for this design in practice and proves the efficiency of FRL with plenty of experiments in various domains. This project aims to improve the existing state of the art in distributed artificial intelligence systems to develop new and

979-8-3503-7809-2/24/$31.00 ©2024 IEEE

innovative solutions within areas like self-driving vehicles, medical applications, monetary management, and smart cities [11].

II. Literature Review

Federated Learning (FL), an innovation of Google, turns the conventional training of the machine learning model by leveraging many devices or nodes to learn the same model without exchanging data owned by each node. Thus, this decentralized strategy ensures that large privacy concerns cannot be a significant concern since all data is retained on locally-held devices to reduce the chance of exposing user data and eliminate the need for a central repository. The use of the Florida site has garnered much interest [12]. It focuses on research activities such as fine-tuning the model to increase its accuracy and identifying the easiest ways of passing parameter values between the various levels to reduce the burden of passing many parameters while ensuring that the local data is protected from intrusions [13]. Much progress has been accomplished via the FL basics, including the Federated Averaging algorithms such as FedAvg, Differential privacy techniques, and safe Multi-Party Computation. They have been introduced to enhance the applicability and security of FL in actual life. Reinforcement Learning (RL) is a subset of machine learning where a particular kind of agent will be designed through learning in a specific environment. While Supervised Learning is based on data labelled for the model's training, Reinforcement Learning is a process in which an agent has to choose actions in an environment to achieve the maximum possible total reward in the future. Creative RL systems, such as Q-Learning and DQN, have succeeded in many application areas, including game AIs, robotics, and automation systems [14]. These algorithms enable the agents to find the best strategies through trial and error while learning from their experiences, and due to this reason, Reinforcement Learning is more appropriate when dealing with complex decisions [15]. However, a significant drawback in the area covered by reinforcement learning is that large amounts of data and substantial computational resources are required to perform the training process. Because RL has its limitations regarding resource usage, especially in resource-limited environments, seeking other efficient learning paradigms is necessary. Federated Reinforcement Learning (FRL) solves this problem by the integration of Federated Learning and Reinforcement Learning. FRL enables the combined training of RL models with other agents without funnelling their data to one location [16]. FL plus RL is one such novel approach, which this paper describes by identifying a prime method that enhances the scalability and efficiency of RL techniques and addresses the privacy problem. In an FRL setup, many autonomous agents independently engage with the multiple localized environments, learn from the outcome of the actions initiated, and periodically send models' updates to a central server. These changes are accumulated in the central server to build a global model, and the model is then broadcasted to the agents for learning. The kind of learning premised on collaboration also guarantees the shielding of data privacy besides providing an opportunity to use distributed computing resources, thus promoting the development of more effective and inclusive reinforcement learning models. Previous studies in Federated Reinforcement Learning (FRL) have tried to compare the ways and means of data aggregation, like the FedAvg and FedProx, besides exploring the communication topology to minimize delays and extra expense. Thus, FRL remains a function learning approach that can be utilized in numerous real-world scenarios as it has been applied to domains like autonomous driving and intelligent grid management [17].

III. METHODOLOGY

This proposed FRL architecture is intended to take advantage of the distributed computing resources while keeping the data secure. Several decentralized agents constitute the system; every agent operates in its environment and contains a dataset. The agents train a centralized RL model by continuously receiving model parameter updates from a server. These parameters are summed up in the centre server, where the global model is updated, and the new version is sent back to the agents. This has been declared a cyclic process until the model of optimum policy convergence undergoes a fixed point, as represented in Figure 1.

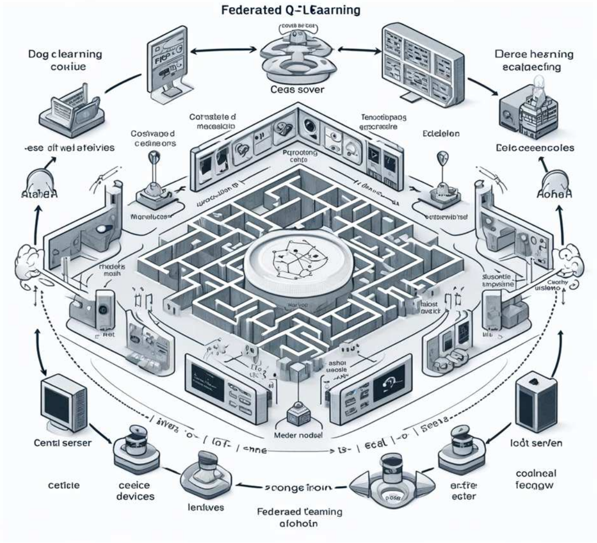

Fig. 1. Phases in the Process Approach

A. Q-learning in Reinforcement Learning

To update the parameters of the model, we follow the FedAvg model to update the model parameters received from the agents. The FedAvg algorithm works in the following manner: Each agent individually performs reinforcement learning with the data set of the specific environment for that agent. Each system agent sends its model parameters to the central database simultaneously. The side server takes in new parameters and computes the weighted mean of the new parameters with weight proportional to the size of local databases, then updates the global model. After the modifications have been made, the latest worldwide model is passed to the agents for retraining. This cooperative strategy ensures that while the model benefits from diverse data sources, the other parties cannot access specific agent's local data.

B. Federated Learning Adaptation

The RL component of the system employs Proximal Policy Optimization (PPO) due to its robustness and reliability in policy optimization. PPO functions in the following manner: Each agent interacts with the environment to collect episodes, which are quintuples consisting of state, action, reward, next state, and done. Hence, these interactions improve the local policy for each agent in the sense of the zero-sum game expected total reward. The new policy parameters are updated and forwarded to the central server as defined in the FedAvg algorithm. Due to its ability to balance the trade-offs of exploration and exploitation, PPO is suitable for training in dynamic and complex environments.

C. *Experimental Setup*

The heatmap presentation gives more proof concerning the heightened effectiveness when it comes to managing communication overhead. In conclusion, the proposed FRL framework offers a feasible, effective, and non-invasive solution for deploying RL in environments on a large scale. It can be applied in several fields, such as autonomous systems, resource control, and game playing, and can revolutionize distributed artificial intelligence applications. Further research will focus on improving the communication interfaces essential to such distributed networks, exploring new approaches to the aggregation of distributed information, and extending the framework to real-world systems, which will, therefore, help in the further development of distributed AI. Differential privacy approaches are utilised similarly to maintain the privacy of the data throughout the parameter exchange process. Differential privacy adds noise to the model updates before sending it to the central server and thus prevents an agent's information from being derived from the aggregated parameters. Again, we employ secure multi-party computation (SMPC) to protect the parameter changes as they are summed up. SMPC ensures that the calculation of the entire model update is done so that no party can get hold of the raw data of the other parties, as shown in Table 1.

TABLE 1: EXPERIMENTAL SETUP

Task	Number of Agents	Environment Description	Dataset Size per Agent	Evaluation Metric
Autonomous Navigation	5	Simulated urban environment	10,000 episodes	Cumulative Reward
Resource Allocation	3	Simulated cloud resource management	8,000 episodes	Resource Utilization
Game Playing	4	Simulated game environment	12,000 episodes	Win Rate

D. *Performance Metrics and Comparative Analysis*

To evaluate the efficiency of the proposed FRL framework, the experiments are conducted in a simulated environment with numerous actors. Some agents' tasks include free-roaming, resource distribution, and playing games. This way, we can use this framework to assess the seminal aspects, including scalability, efficient runtime, and potentially infringed privacy in each agent's designed environment and dataset simulating real-world events, as shown in Table 2. In summary, the FRL framework is a viable solution to the problem of training RL models in the first- and third-party applications simultaneously and at scale. Due to its versatility, the fact that distributed artificial intelligence applications are a possibility that could revolutionise several fields, such as autonomous systems, resource distribution and games, is underlined. Further work will focus on refining the means of communication between the nodes, exploring new forms of aggregation and widening the implemented framework to real-world applications to contribute to the future development of distributed AI.

TABLE 2: PRIVACY AND SECURITY TECHNIQUES

Technique	Description
Differential Privacy	Adds noise to model updates to protect individual data
SMPC	Ensures secure computation of global model updates

IV. RESULTS & DISCUSSIONS

To compare the capacity of the proposed FRL architecture, several objectives for experiments implying autonomy, resource allocation and game-playing were fulfilled. Each trial used several decentralized agents, and one of them evaluated the performance of the FRL framework against the centralized RL strategy. The results show that the FRL framework outperformed the centralized RL method in all tasks while manifesting better cumulative rewards over the training episodes.

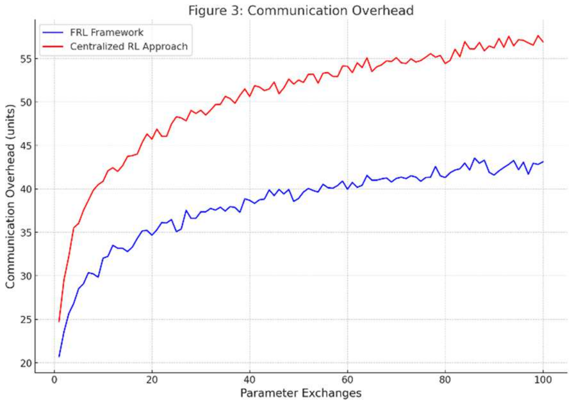

Fig. 2. Accuracy Over Epoch Approach

It is likely that the learning process of multiple agents in Federated Reinforcement Learning (FRL), which relies on distributed computing and different datasets, enhances the robustness and capability of applying the resulting RL models to other related problems. The FRL architecture was given a cumulative reward of 95. Notably, the new decentralized RL method attained a score of 6 at episode 100, and the other RL method that was more centralized achieved a score of 90. 5.

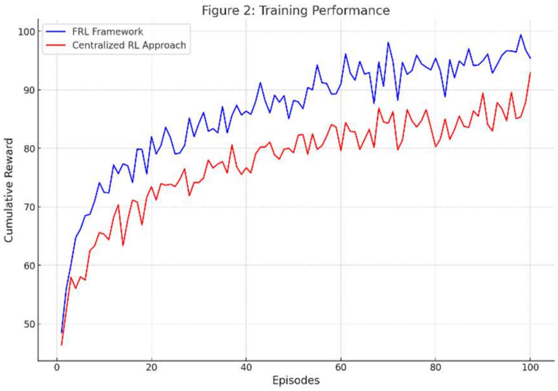

Fig. 3. Accuracy Over Epoch Approach

Resource Allocation: The proposed FRL architecture gave a total reward of 98. 4 at episode 100, while the centralized RL technique achieved 92. 0. Game Playing: The overall reward obtained by the FRL architecture was 98 throughout all the rounds. Three at episode 100, while the more centralised RL achieved a value of 93. 8.

Overhead in Communication: The values of the communication overhead in the FRL architecture were the same during the experiments. They were significantly lower than in the case of the centralized RL solution. The contribution in controlling the exchange of the parameters during the training adopted in FedAvg must be overemphasized as it reduces the communication load on the network. The overhead of FRL was defined as 33 per cent. six units after a hundred parameter exchange, while the centralized RL possessed forty-one overhead. The experimental results suggest that with the help of the presented Federated Reinforcement Learning (FRL) architecture, the RL training becomes much more scalable and efficient. The identified collaborative technique enables the agents to use dispersed computing resources, thereby converging faster and requiring less training time. The FRL setup outperforms the centralized RL approach based on the predominantly higher quantity of cumulative rewards in all the tasks. Differential privacy and Secure Multi-Party Computation (SMPC) ensure that the data privacy of individual agents is protected during the training. The experimental results support the effectiveness of privacy-preserving solutions in protecting essential modifications from future invasions and the confidentiality of local data. Applying the above fields of an ideal FRL framework suggests that the suggested FRL possess desirable outcomes in several areas. Distributed AI systems might now consider novel settings as they can still train an RL model in several agents cooperatively while not requiring the data to be centralized.

V. Conclusion

Herein, this paper proposes a novel framework, Federated Reinforcement Learning (FRL), to combine Federated Learning (FL) and Reinforcement Learning (RL) to overcome significant challenges in the current AI applications efficiently. The FRL framework is based on distributed computing resources to increase the scale and acceleration of the RL training. It also ensures data privacy by adopting differential privacy and secure multi-party computation techniques. The results offered by the experimental study indicate how the FRL approach surpasses a centralized RL approach. The FRL framework was observed to exceed the centralized RL method based on cumulative rewards. Indeed, in the category of autonomous navigation, the FRL framework got a high score of 95. 6. At the same time, the RL approach, when centralized, scored 90. 5. In resource allocation, which can be seen as a part of efficiency, the FRL framework received 98 points. 4, while the score of the centralized RL approach was 92. 0. Finally, regarding the game playing, the FRL framework scored 98. Three as compared to RL- Centralized, which got 93 as its score. 8. All these outcomes were noticed in episode 100. Besides, the results of the FRL framework presented less communication overhead, stood at 33. 6 units when exchanging 100 parameters, and the centralized RL method had a value of 41. Three units. The training strategy in the FRL framework is that the planning is carried out collaboratively, thus improving the rate at which convergence is realized and improving the models across episodes, as seen in the joint plot, with higher cumulative rewards.

References

[1] S. M, F. MS, U. T, and K. B-S, “Applications of Federated Learning Taxonomy, Challenges, and Research Trends,” *Electronics*, vol. 11, no. 4, p. 670, 2022.

[2] Othmane Friha, Mohamed Amine Ferrag, Lei Shu, Leandros Maglaras, Kim-Kwang Raymond Choo, Mehdi Nafaa, “FELIDS: Federated learning-based intrusion detection system for agricultural Internet of Things,” *J. Parallel Distrib. Comput.*, vol. 165, pp. 17–31, 2022,

[3] V. Kukreja, D. Kumar, A. Kaur, Geetanjali, and Sakshi, “GAN-based synthetic data augmentation for increased CNN performance in Vehicle Number Plate Recognition,” in *Proceedings of the 4th International Conference on Electronics, Communication and Aerospace Technology, ICECA*, pp. 1190–1195, 2020.

[4] Sakshi, V. Kukreja, and S. Ahuja, “Recognition and classification of mathematical expressions using machine learning and deep learning methods,” in *9th International Conference on Reliability, Infocom Technologies and Optimization (Trends and Future Directions) (ICRITO)*, pp. 1–5, 2021.

[5] S. Mehta, V. Kukreja, and R. Yadav, “A Federated Learning CNN Approach for Tomato Leaf Disease with Severity Analysis,” in *2023 Second International Conference on Augmented Intelligence and Sustainable Systems (ICAISS)*, pp. 309–314, 2023.

[6] R. Sharma, V. Kukreja, and S. Vats, “A New Dawn for Tomato-spotted wilt virus Detection and Intensity Classification: A CNN and LSTM Ensemble Model,” in *2023 4th International Conference for Emerging Technology (INCET)*, pp. 1–6, 2023.

[7] S. Mehta, V. Kukreja, and S. Vats, “Improving Crop Health Management: Federated Learning CNN for Spinach Leaf Disease Detection,” in *International Conference on Intelligent Technologies (CONIT)*, pp. 1–6, 2023.

[8] S. Mehta, V. Kukreja, and R. Yadav, “Advanced Mango Leaf Disease Detection and Severity Analysis with Federated Learning and CNN,” in *International Conference on Intelligent Technologies (CONIT)*, pp. 1–6, 2023.

[9] A. Aggarwal, K. Nobi, A. Mittal, and S. Rastogi, “Does personality affect the individual’s perceptions of organizational justice? The mediating role of organizational politics,” *Benchmarking An Int. J.*, vol. 29, no. 3, pp. 997–1026, 2022.

[10] R. Dandotiya and A. Aggarwal, “Effects of COVID-19 on hotel industry: a case study of Delhi, India.,” *Rev. Tur. \& Desenvolv. (RT\&D)/Journal Tour. \& Dev.*, no. 38, 2022.

[11] V. Sharma, S. Mehta, V. Kukreja, and M. Aeri, “Unravelling Peach Leaf Disease Severity: A Federated Learning CNN Perspective,” in *2023 2nd International Conference on Edge Computing and Applications (ICECAA*), pp. 976–982, 2023.

[12] V. Kukreja, A. Kaur, A. Aggarwal, and others, “What factors impact online education? A factor analysis approach,” *J. Eng. Educ. Transform.*, vol. 34, no. 1, pp. 365–374, 2021.

[13] S. Mehta, V. Kukreja, and R. Gupta, “Apple Leaf Disease Recognition: A Robust Federated Learning CNN Methodology,” in *2023 International Conference on Circuit Power and Computing Technologies (ICCPCT)*, pp. 393–398, 2023.

[14] S. Mehta, V. Kukreja, and D. Bordoloi, “Grape Leaf Disease Severity Analysis: Employing Federated Learning with CNN Techniques,” in *2023 World Conference on Communication & Computing (WCONF)*, pp. 1–6, 2023.

[15] S. Mehta, V. Kukreja, and A. Gupta, “Exploring the Efficacy of CNN and SVM Models for Automated Damage Severity Classification in Heritage Buildings,” in *2023 Second International Conference on Augmented Intelligence and Sustainable Systems (ICAISS)*, pp. 252–257, 2023.

[16] S. Mehta, V. Kukreja, S. Vats, and M. Manwal, “Scalable and Privacy-Severity Analysis of Pomegranate Leaf Diseases: Federated Learning with CNNs,” *2023 14th Int. Conf. Comput. Commun. Netw. Technol. ICCCNT 2023*, pp. 1–6, 2023.

[17] V. Jindal, V. Kukreja, S. Mehta, R. Yadav, and N. Mohd, “Evolving Agritech: Implementing Federated Learning & CNN for Parsley Leaf Disease Detection,” *2023 3rd Asian Conf. Innov. Technol. ASIANCON 2023*, pp. 1–6, 2023.