eecjournal

International Journal of Electrical, Electronics and Computers
Vol-8, Issue-6 | Nov-Dec, 2023
Available: https://aipublications.com/ijeec/
Peer-Reviewed Journal

Scalable Decentralized Multi-Agent Federated Reinforcement Learning: Challenges and Advances

Praveen Kumar Myakala[1], Srikanth Kamatala[2]

[1] Independent Researchers, Dallas, Texas, USA
Email: Praveen.K.Myakala@gmail.com

[2] Independent Researchers, Dallas, Texas, USA
Email : Kamatala.Srikanth@gmail.com

Received: 24 Oct 2023; Accepted: 30 Nov 2023; Date of Publication: 07 Dec 2023

***Abstract*—** *The increasing prevalence of decentralized multiagent systems has spurred interest in Federated Reinforcement Learning (FRL) as a privacy-preserving framework for collaborative learning. However, scaling FRL to multi-agent settings introduces significant challenges, particularly in communication efficiency, decentralized aggregation, and handling nonstationary environments. This survey explores recent advancements in Scalable Decentralized Multi-Agent Federated Reinforcement Learning (MA-FRL), with a focus on communication efficient strategies and decentralized aggregation techniques. We review key approaches such as selective agent communication, local model updates, and gradient compression, analyzing their impact on scalability, convergence, and performance trade-offs. Additionally, we highlight comparative insights into different methods, their limitations, and real-world applicability in decentralized systems such as autonomous vehicles and smart grids. By identifying open challenges, including robustness against adversarial attacks and adaptive communication mechanisms, we outline promising directions for advancing decentralized MAFRL.*

Keywords*— *Federated Reinforcement Learning, Multi-Agent Systems, Decentralized Learning, Scalability, Communication Efficiency, Aggregation Techniques, Non-Stationarity

I. INTRODUCTION

The increasing prevalence of decentralized multi-agent systems has spurred interest in Federated Reinforcement Learning (FRL) as a privacy-preserving framework for collaborative learning. FRL enables agents to learn policies without sharing raw data, making it particularly beneficial in privacy-sensitive applications such as autonomous systems, smart grids, and the Internet of Things (IoT) [1], [2]. However, scaling FRL to multi-agent environments presents significant challenges, particularly in communication efficiency, decentralized aggregation, and handling non-stationary dynamics.

Frequent model updates across many distributed agents increase bandwidth consumption, posing a major scalability challenge [3]. As multiple agents learn and adapt simultaneously, the underlying environment dynamics constantly change, making policy convergence more difficult [4]. Variations in computational resources, network connectivity, and local datasets introduce inconsistencies in the learning process, affecting overall performance [5]. Additionally, decentralized architecture must efficiently aggregate updates in a peer-to-peer fashion, as opposed to traditional centralized federated learning setups [6].

This paper presents a focused survey on Scalable Decentralized Multi-Agent Federated Reinforcement Learning (MA-FRL), emphasizing communication-efficient strategies and decentralized aggregation techniques. We analyze key approaches, including local model updates, selective agent communication, gradient compression, and decentralized aggregation (e.g., peer-to-peer and blockchain-based strategies), examining their impact on scalability, convergence, and robustness.

Our key contributions are as follows: (1) a systematic review of recent advancements in scalable MA-FRL with a focus on communication-efficient learning and decentralized aggregation, (2) a comparative analysis highlighting tradeoffs between different techniques in terms of scalability, convergence speed, and robustness, and (3) a discussion of open challenges and promising directions for future research, including adaptive communication mechanisms and robustness against adversarial agents.

II. SCALABILITY CHALLENGES IN DECENTRALIZED MA-FRL

Scaling Multi-Agent Federated Reinforcement Learning (MA-FRL) presents several challenges that hinder efficient learning in decentralized settings. These challenges stem from the need to balance communication efficiency, policy convergence, and computational constraints while ensuring robustness in dynamic and heterogeneous environments. Among these challenges, communication overhead and nonstationarity are particularly crucial, as they directly impact the scalability and stability of decentralized MA-FRL.

A. Communication Overhead and Bandwidth Constraints

A key scalability bottleneck in decentralized MA-FRL is the high communication overhead resulting from frequent model updates across multiple agents. In traditional Federated Learning (FL), a central server aggregates local models, reducing direct inter-agent communication. However, decentralized MA-FRL lacks a central coordinator, requiring agents to exchange updates through peer-to-peer or network-based aggregation [1], [3]. As the number of agents increases, the volume of exchanged gradients, policies, and rewards grows significantly, leading to high bandwidth consumption and increased latency. Communication-efficient strategies such as gradient compression, asynchronous updates, and selective model aggregation are essential for mitigating these issues. Furthermore, heterogeneity in computational resources and network connectivity can exacerbate communication inefficiencies, as some agents may experience delays in transmitting updates, affecting overall synchronization.

B. Non-Stationarity in Multi-Agent Interactions

Scalability in decentralized MA-FRL is further complicated by the non-stationarity of the learning environment. Since agents update their policies independently, the underlying environment dynamics shift continuously, making it difficult for policies to converge [4]. This challenge is exacerbated in large-scale systems where agents interact asynchronously and must adapt to evolving behaviors of other agents. Nonstationarity also directly affects decentralized aggregation, as policy updates may become outdated or misaligned when integrated with others. Addressing non-stationarity requires approaches such as opponent modeling, meta-learning, and stabilized policy updates, which improve adaptability and convergence in dynamic settings.

C. Heterogeneity in Computational and Data Resources

Decentralized MA-FRL agents often operate under varying computational capacities, network conditions, and local datasets. Unlike centralized FL, where model updates can be synchronized, decentralized learning must account for heterogeneous agent capabilities [5]. Some agents may have limited processing power or intermittent connectivity, resulting in delayed or incomplete updates. Additionally, heterogeneous local data sets can lead to biased model updates, reducing the overall generalizability of learned policies. Strategies such as federated averaging with adaptive weighting, edge-computing-assisted federated learning, and hierarchical aggregation frameworks can help mitigate these disparities by optimizing model synchronization based on individual agent constraints.

D. Decentralized Aggregation and Model Synchronization

The absence of a central server in MA-FRL necessitates efficient decentralized aggregation techniques to ensure scalable learning. Traditional FL methods rely on server-based aggregation (e.g., FedAvg [3]), which does not directly translate to fully decentralized settings. Peer-to-peer aggregation, graph-based federated learning, and blockchain-assisted consensus mechanisms have been explored to address this limitation [6]. However, these methods introduce trade-offs in terms of scalability, security, and communication costs. Since decentralized aggregation must account for varying agent participation rates and potential network failures, it is necessary to develop adaptive aggregation strategies that balance efficiency with robustness. A key research direction is analyzing the impact of different aggregation techniques on convergence and stability in decentralized MA-FRL environments.

E. Scalability vs. Privacy Trade-Offs

Ensuring scalability while maintaining privacy is a fundamental challenge in decentralized MA-FRL because many privacy-preserving techniques, such as differential privacy and secure multi-party computation (SMPC), inherently introduce additional computational and communication overhead [2]. While these methods help protect agent data from adversarial threats, they can significantly increase latency and resource consumption, making them difficult to scale to large networks. Techniques like local differential privacy (LDP) and secure aggregation with reduced computational complexity offer

ISSN: 2456-2319
https://dx.doi.org/10.22161/eec.86.2

promising solutions but require further exploration to optimize the trade-off between privacy and scalability. Additionally, integrating privacy-preserving mechanisms with adaptive communication protocols could help improve the feasibility of secure decentralized MA-FRL systems.

F. Summary and Research Directions

Addressing scalability in decentralized MA-FRL requires a multi-faceted approach that optimizes communication, aggregation, and adaptation in dynamic multi-agent environments. Future research should explore adaptive communication protocols, hierarchically federated learning frameworks, and efficient decentralized consensus mechanisms to enhance scalability while maintaining robustness and privacy. Additionally, deeper investigations into the interplay between communication overhead, heterogeneity, and non-stationarity will be crucial in developing scalable and resilient MA-FRL frameworks. The next section reviews existing solutions that tackle these challenges.

III. EXISTING APPROACHES

Addressing the scalability challenges in decentralized Multiagent Federated Reinforcement Learning (MA-FRL) requires efficient communication strategies and robust aggregation mechanisms. Existing research has proposed various techniques to reduce communication overhead, mitigate nonstationarity, and enhance decentralized policy synchronization. This section reviews key approaches categorized into communication-efficient methods and decentralized aggregation strategies, explicitly linking each to the scalability challenges discussed in Section II.

A. Communication-Efficient Strategies

Communication efficiency is crucial for scalable MA-FRL, as excessive message exchange can lead to network congestion and high latency. A key trend in communication-efficient MAFRL is the development of adaptive strategies, where agents dynamically adjust communication parameters based on factors such as network conditions, computational resources, and learning progress. This adaptability is crucial for optimizing performance in heterogeneous and dynamic environments.

Local Model Updates: Local model updates reduce communication overhead by allowing agents to perform multiple training iterations before synchronizing with others. This directly addresses the communication overhead challenge by decreasing the frequency of global synchronizations. Methods such as FedAvg [3] and FedProx [2] adjust local computation intensity based on agent resources. However, this approach introduces a trade-off, as excessive local updates can lead to model divergence, particularly in non-IID (non-independent and identically distributed) environments.

Gradient Compression and Quantization: Gradient compression techniques such as sparse updates, low-bit quantization, and sketching [7] aim to reduce bandwidth usage by transmitting only the most significant updates or encoding them efficiently. These methods specifically address the bandwidth constraints in large-scale multi-agent systems. While compression reduces communication costs, it may introduce approximation errors that affect convergence. Recent approaches incorporate adaptive error compensation to balance communication efficiency with learning performance.

Selective Agent Communication: Instead of broadcasting updates to all agents, selective communication strategies identify the most relevant peers for information exchange. This technique is particularly effective in addressing both communication overhead and heterogeneity, as agents can focus updates on influential neighbors while ignoring low-impact interactions. Methods such as gossip learning [8] and attention-based message dynamically determine which agents contribute the most to learning. By prioritizing high-impact interactions, these methods improve communication efficiency while maintaining policy performance.

Asynchronous Communication: Traditional FL relies on synchronous updates, where all agents must wait for each other before aggregating updates. In contrast, asynchronous communication allows agents to update their models at different times, reducing idle time and improving responsiveness. This approach is particularly useful for addressing the heterogeneity challenge, as it prevents slower agents from delaying overall learning progress. Methods such as FedAsync [9] and staleness-aware aggregation compensate for delayed updates, ensuring robustness against network delays and computational disparities.

B. Decentralized Aggregation Strategies

Decentralized MA-FRL eliminates reliance on a central server, requiring alternative aggregation mechanisms for combining agent policies. These methods address the decentralized aggregation challenge by ensuring that agents can collaboratively learn without centralized coordination.

Peer-to-Peer Aggregation: In peer-to-peer (P2P) learning, agents directly share updates with their neighbors without a central coordinator. This method enhances scalability by removing bottlenecks associated with server-based aggregation. Graph-based techniques such as D-FL [10] and Diffusion FL enable agents to propagate updates across a network topology, gradually converging to a global policy. However, P2P aggregation is vulnerable to network partitioning, requiring robust neighbor selection policies.

ISSN: 2456-2319
https://dx.doi.org/10.22161/eec.86.2

Blockchain-Assisted Federated Learning: Blockchain technology has been explored as a means of enabling decentralized and trustworthy aggregation in FL [11]. By storing model updates on a distributed ledger, blockchain based FL ensures transparent and tamper-resistant aggregation. Techniques such as smart contract-driven FL enable secure model sharing without centralized control. However, the computational and storage overhead of blockchain integration remains a challenge, particularly for large-scale deployments.

Hierarchical Federated Learning: Hierarchical FL introduces intermediate aggregation layers, where groups of agents first synchronize locally before communicating with the global network. This reduces overall communication complexity and allows for more structured policy updates [12]. Hierarchical aggregation is particularly useful in large-scale multi-agent environments where direct agent-to-agent communication may be impractical. However, designing an optimal hierarchy requires balancing local and global updates to ensure efficient learning.

Decentralized Consensus Mechanisms: Consensus-based approaches, such as federated averaging with decentralized optimization [13], allow agents to reach agreement on policy updates without a central server. These techniques specifically target the scalability challenge by enabling largescale distributed learning while reducing reliance on predefined network structures. However, ensuring stability and efficiency in fully decentralized settings remains an open research challenge.

C. Comparative Analysis of Existing Methods

The methods discussed above offer different trade-offs in terms of communication efficiency, convergence speed, and robustness. ***Table I*** provides a comparative summary of key approaches, highlighting their advantages, limitations, and scalability impact.

Table I: Comparison of Communication-Efficient and Decentralized Aggregation Strategies

Method	Key Advantage	Potential Limitation	Scalability Impact
Local Updates	Reduces communication frequency	Can lead to model divergence	High
Gradient Compression	Minimizes bandwidth usage	May introduce approximation errors	Moderate
Selective Communication	Focus updates on relevant agents	Requires efficient neighbor selection	High
Asynchronous Updates	Improves responsiveness	Needs staleness-aware aggregation	High
Peer-to-Peer Aggregation	Fully decentralized	Vulnerable to network partitioning	Moderate
Blockchain FL	Ensures secure aggregation	High computational overhead	Low
Hierarchical FL	Reduce direct agent communication	Requires structured hierarchy	High
Decentralized Consensus	Scalable without central server	Needs efficient optimization protocols	High

D. Summary and Research Directions

Existing solutions address various aspects of scalability in decentralized MA-FRL, yet significant challenges remain. Communication-efficient strategies such as adaptive gradient compression and dynamic agent clustering could further optimize bandwidth usage. Additionally, improving decentralized aggregation through hybrid peer-to-peer and hierarchical models could enhance stability in large-scale deployments. Future work should also explore the integration of privacy-preserving techniques with adaptive communication strategies to balance security and efficiency. The next section provides a comparative analysis of these techniques, evaluating their effectiveness in different MA-FRL scenarios.

IV. COMPARATIVE ANALYSIS OF KEY TECHNIQUES

The scalability of decentralized Multi-Agent Federated Reinforcement Learning (MA-FRL) depends on a balance between communication efficiency, policy convergence, and robustness. This section provides a comparative analysis of the existing techniques discussed in Section III, focusing on their trade-offs in terms of scalability, communication overhead, convergence speed, and adaptability in heterogeneous environments.

A. Trade-offs in Communication Efficiency

Communication-efficient strategies are crucial for scaling MA-FRL, as they impact the frequency, size, and necessity of inter-agent message exchanges. Methods that reduce

ISSN: 2456-2319
https://dx.doi.org/10.22161/eec.86.2

communication overhead can improve scalability but may introduce delays in convergence due to limited information sharing.

Table II summarizes the trade-offs of key communication efficient approaches.

Table II: *Comparison of Communication-Efficient Strategies*

Method	**Communication Overhead**	**Convergence Speed**	**Scalability**
Local Model Updates	Low (fewer global updates)	Moderate (risk of slower convergence)	High (reduces communication needs)
Gradient Compression	Low (smaller message sizes)	Potential for Slower Convergence (approximation errors)	Moderate (depending on compression level)
Selective Communication	Low (fewer agents contacted)	High (focus on key updates)	High (minimizes unnecessary interactions)
Asynchronous Communication	Moderate (independent updates)	High (avoids waiting delays)	High (less synchronization required)

The effectiveness of these communication strategies directly influences decentralized aggregation methods, which must balance scalability with policy synchronization.

B. Effectiveness of Decentralized Aggregation Strategies

Decentralized aggregation eliminates the need for a central server but requires efficient coordination to ensure policy alignment. Aggregation methods must handle communication constraints while maintaining robustness against network failures. ***Table III*** presents a comparative summary of key aggregation techniques.

Table III: Comparison of Decentralized Aggregation Strategies

Method	**Scalability**	**Robustness to Failures**	**Computational Cost**
Peer-to-Peer Aggregation	High (distributed updates)	Moderate (robustness depends on network topology)	Low (minimal overhead)
Blockchain FL	Low (ledger overhead)	High (secure and tamper-proof)	High (computationally intensive)
Hierarchical FL	High (structured aggregation)	High (localized stability)	Moderate (complexity of hierarchy management)
Decentralized Consensus	High (distributed optimization)	Moderate (depends on consensus protocol)	Moderate (computational cost depends on the specific protocol)

Scalability in decentralized aggregation is also influenced by the level of heterogeneity among agents and the stability of learning policies in non-stationary environments.

C. Impact of Heterogeneity and Non-Stationarity

Heterogeneity in agent resources and dynamic environment shifts present significant obstacles to scalable MA-FRL. Adaptive strategies, such as dynamic model synchronization and staleness-aware aggregation, help mitigate these issues. However, non-stationarity remains a fundamental challenge, requiring techniques such as opponent modeling and meta-learning to improve learning stability in changing environments.

The relationship between heterogeneity, non-stationarity, and scalability is summarized in ***Table IV***.

***Table IV**: Impact of Heterogeneity and Non-Stationarity on Scalability*

Challenge	Effect on Scalability	Existing Mitigation	Open Issues
Heterogeneous Resources	Moderate (slows convergence)	Adaptive Synchronization (weight based on resources)	Unbalanced Model Contributions
Non-Stationary Agents	High (policies may diverge)	Opponent Modeling (anticipate agent shifts)	Slow Adaptation to Changes
Variable Data Distribution	High (bias in policy updates)	Meta-Learning (fast adaptation)	Reduced Policy Generalization

D. Summary and Future Research Directions

The analysis highlights the necessity of hybrid strategies that combine multiple approaches to address different scalability challenges. While local updates and gradient compression reduce communication costs, they need to be balanced with aggregation techniques that ensure convergence. Similarly, decentralized aggregation must be designed with robustness and efficiency in mind, favoring hybrid approaches such as hierarchical peer-to-peer models that integrate structured communication with flexibility.

Future research should focus on:

- Adaptive Aggregation Mechanisms: Developing methods that dynamically adjust model weights based on agent contributions, network conditions, and learning performance to optimize scalability.
- Privacy-Preserving Scalability: Investigating hybrid privacy-preserving techniques that combine differential privacy with secure aggregation protocols to minimize computational overhead.
- Resilient Learning in Dynamic Systems: Developing fast adaptation techniques for non-stationary environments, including meta-learning and dynamic reinforcement strategies for scalable multi-agent collaboration.

The next section explores open challenges and potential directions for advancing scalable decentralized MA-FRL.

V. OPEN CHALLENGES AND FUTURE DIRECTIONS

Despite significant advancements in Scalable Decentralized Multi-Agent Federated Reinforcement Learning (MA-FRL), several open challenges remain. Among these challenges, addressing non-stationarity and ensuring privacy-preserving scalability are particularly critical, as they directly impact learning stability and deployment feasibility. Addressing these challenges is essential for improving the scalability, efficiency, and robustness of decentralized MA-FRL in real-world applications. This section outlines key unresolved issues and potential future directions for research.

A. Scalability of Communication-Efficient Strategies

Communication overhead remains a major bottleneck in large-scale MA-FRL systems. While techniques such as gradient compression, selective agent communication, and asynchronous updates have been proposed, their effectiveness depends on network conditions and agent heterogeneity. A key research direction is the development of adaptive communication protocols that dynamically adjust message frequency, size, and relevance based on real-time network conditions and agent learning progress. Exploring hybrid strategies that combine local updates with decentralized gossip-based aggregation could further enhance scalability.

B. Decentralized Aggregation in Dynamic Environments

Most existing decentralized aggregation methods assume relatively stable agent participation. However, real-world applications, such as IoT networks and autonomous multiagent systems, involve dynamic environments where agents frequently join or leave the system. This introduces challenges related to fault tolerance, trust mechanisms, and adaptive aggregation. Future research should focus on designing resilient decentralized aggregation frameworks that incorporate self-healing mechanisms, agent dropout detection, and adaptive peer selection. Blockchain-based federated learning could provide tamper-resistant aggregation but requires further optimization to reduce its computational and storage overhead.

C. Handling Heterogeneity in Large-Scale MA-FRL

Heterogeneity in computational resources, network bandwidth, and local datasets significantly impacts model convergence and fairness in decentralized learning. Furthermore, heterogeneity can exacerbate non-stationarity, as agents with different learning speeds or local reward distributions may create inconsistent updates. Current approaches, such as adaptive weighting in federated averaging, are limited in handling extreme disparities among agents. Future work should explore federated reinforcement learning with hierarchical and cluster-based aggregation, where agents with similar capabilities are grouped before global synchronization. Additionally, optimizing personalized federated learning for multi-agent settings could help balance local adaptation with global model generalization.

D. Stability in Non-Stationary Multi-Agent Systems

Non-stationarity remains a critical challenge in MA-FRL, as agents continuously adapt their policies while interacting in a shared environment. Existing stabilization techniques, such as meta-learning and opponent modeling, require further refinement for decentralized settings. Research should explore adaptive policy stabilization mechanisms that leverage predictive modeling of agent behavior and dynamically adjust learning rates based on environmental changes. Another promising direction is the integration of curriculum learning, where agents progressively adapt to increasing levels of complexity, improving robustness in non-stationary environments.

E. Privacy-Preserving and Secure Decentralized Learning

Ensuring privacy and security in decentralized MA-FRL remains an ongoing challenge, particularly when dealing with sensitive applications such as healthcare, finance, and industrial automation. However, privacy-preserving techniques often introduce communication overhead, making scalability even more difficult. While techniques like differential privacy and secure multiparty computation (SMPC) enhance data protection, they often impose additional computational costs. Future research should focus on lightweight privacy-reserving techniques, such as efficient secure aggregation and privacy-aware policy learning, which balance security and efficiency. Additionally, adversarial robustness mechanisms need further investigation to defend against poisoning attacks and model inversion threats in decentralized settings.

F. Benchmarking and Standardized Evaluation Metrics

A major limitation in current MA-FRL research is the lack of standardized benchmarking frameworks and evaluation metrics, hindering direct comparisons between different approaches. Most existing studies evaluate their methods on different testbeds, making it difficult to assess their scalability, convergence, and communication efficiency under uniform conditions. Establishing benchmark environments with standardized performance metrics would enable more systematic comparisons. Future work should also focus on developing real-world testbeds that simulate decentralized, large-scale multi-agent interactions, providing a more comprehensive evaluation of proposed methods.

G. Towards Hybrid and Adaptive Federated Reinforcement Learning

The future of MA-FRL lies in developing hybrid models that integrate multiple learning paradigms to enhance scalability and adaptability. Potential research areas include:

- Hybrid Centralized-Decentralized Learning: Combining the benefits of centralized coordination with decentralized agent autonomy. One approach could involve using a subset of agents as super-agents responsible for periodic centralized aggregation while most agents continue decentralized learning.
- Multi-Tier Federated Reinforcement Learning: Implementing hierarchical structures where agents communicate within localized sub-networks before contributing
- to global updates. For example, clustering agents based on proximity or task similarity could improve local efficiency while reducing global communication overhead.
- Adaptive Federated Reinforcement Learning: Designing self-adjusting mechanisms where agents dynamically switch between synchronous and asynchronous updates based on network congestion, task complexity, or learning phase. This would enable greater flexibility in large-scale, resource-constrained environments.

H. Summary and Future Outlook

Addressing these open challenges is essential for realizing the full potential of scalable decentralized MA-FRL. Future advancements should focus on:

- Enhancing communication efficiency through adaptive and hybrid message-passing techniques.
- Developing fault-tolerant decentralized aggregation strategies that are resilient to dynamic agent participation.
- Improving policy stability in non-stationary environments with adaptive learning rate adjustments and predictive behavior modeling.

ISSN: 2456-2319
https://dx.doi.org/10.22161/eec.86.2

- Strengthening privacy-preserving methods that minimize computational overhead while maintaining security.
- Establishing benchmarking frameworks for systematic evaluation and comparison of MA-FRL techniques.

By addressing these challenges, future research can pave the way for scalable, secure, and adaptive decentralized learning systems applicable to real-world multiagent environments.

VI. CONCLUSION

Scalable Decentralized Multi-Agent Federated Reinforcement Learning (MA-FRL) represents a promising paradigm for enabling collaborative learning in multi-agent systems without relying on a central coordinator. This survey explored key challenges, existing approaches, comparative trade-offs, and future research directions in MA-FRL, focusing on communication efficiency, decentralized aggregation, heterogeneity, non-stationarity, privacy, and benchmarking.

Our analysis identified several major challenges that hinder the scalability of decentralized MA-FRL. Communication overhead, exacerbated by frequent inter-agent synchronization, remains a primary bottleneck. While strategies such as gradient compression, selective agent communication, and asynchronous updates reduce communication costs, they introduce trade-offs related to convergence speed and stability. Decentralized aggregation further complicates scalability, requiring efficient mechanisms to aggregate local policies while ensuring robustness to network failures and agent dropouts. Additionally, heterogeneous agent capabilities and nonstationary environments impact learning stability, necessitating adaptive learning techniques for effective collaboration.

Existing approaches provide partial solutions to these challenges, but limitations persist. Communication-efficient strategies improve scalability but must be balanced with policy consistency. Decentralized aggregation methods such as peer-to-peer, blockchain-assisted, and hierarchical learning enhance autonomy but introduce computational trade-offs. Privacy-preserving mechanisms, while crucial for secure decentralized learning, often impose additional communication and computational burdens, which can further hinder scalability.

To address these challenges, we outlined several promising research directions. Adaptive communication protocols, integrating real-time network optimization, could dynamically regulate inter-agent communication, reducing redundancy without sacrificing performance. Resilient decentralized aggregation frameworks, incorporating self-healing mechanisms and trust-aware peer selection, could improve fault tolerance in dynamic environments. Enhancing policy stability in nonstationary settings through predictive modeling, opponent-aware learning, and curriculum learning could lead to more robust multi-agent collaboration. Further research is also needed in lightweight privacy-preserving techniques to balance security with efficiency, ensuring privacy without excessive computational overhead. Finally, establishing standardized benchmarking frameworks will be critical to enabling systematic comparisons of MA-FRL approaches across different application domains.

The future of scalable MA-FRL lies in the development of hybrid and adaptive learning models, integrating techniques such as hybrid centralized-decentralized learning, multi-tier federated reinforcement learning, and adaptive federated reinforcement learning. These approaches can enable a balance between scalability, efficiency, and flexibility in large-scale distributed systems.

By addressing these challenges, future research can unlock the full potential of scalable decentralized MA-FRL, enabling robust, secure, and adaptable multi-agent learning systems. As MA-FRL continues to evolve, interdisciplinary efforts across distributed systems, reinforcement learning, and secure computing will be necessary to bridge the gap between theoretical advancements and real-world deployment. The insights presented in this survey provide a strong foundation for future research in scalable decentralized MA-FRL, guiding the development of next-generation distributed learning systems for real-world multi-agent applications.

REFERENCES

[1] 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.

[2] Tian Li, Anit Kumar Sahu, Ameet Talwalkar, and Virginia Smith. Federated learning: Challenges, methods, and future directions. *IEEE Signal Processing Magazine*, 37(3):50–60, 2020.

[3] McMahan, B., Moore, E., Ramage, D., Hampson, S., & y Arcas, B. A. (2017, April). Communication-efficient learning of deep networks from decentralized data. In *Artificial intelligence and statistics* (pp. 1273-1282). PMLR.

[4] Fu, H., Tang, H., Hao, J., Lei, Z., Chen, Y., & Fan, C. (2019). Deep multi-agent reinforcement learning with discrete-continuous hybrid action spaces. *arXiv preprint arXiv:1903.04959*.

ISSN: 2456-2319
https://dx.doi.org/10.22161/eec.86.2

[5] Zhang, C., Xie, Y., Bai, H., Yu, B., Li, W., & Gao, Y. (2021). A survey on federated learning. *Knowledge-Based Systems*, *216*, 106775.
[6] Foerster, J., Assael, I. A., De Freitas, N., & Whiteson, S. (2016). Learning to communicate with deep multi-agent reinforcement learning. *Advances in neural information processing systems*, *29*.
[7] Aji, A. F., & Heafield, K. (2017). Sparse communication for distributed gradient descent. *arXiv preprint arXiv:1704.05021*.
[8] Blot, M., Picard, D., Cord, M., & Thome, N. (2016). Gossip training for deep learning. arXiv preprint arXiv:1611.09726.
[9] Xie, C., Koyejo, S., & Gupta, I. (2019). Asynchronous federated optimization. arXiv preprint arXiv:1903.03934.
[10] Lalitha, A., Kilinc, O. C., Javidi, T., & Koushanfar, F. (2019). Peer-to-peer federated learning on graphs. arXiv preprint arXiv:1901.11173.
[11] Xu, G., Zhou, Z., Dong, J., Zhang, L., & Song, X. (2023). A blockchain-based federated learning scheme for data sharing in industrial internet of things. IEEE Internet of Things Journal.
[12] Briggs, C., Fan, Z., & Andras, P. (2020, July). Federated learning with hierarchical clustering of local updates to improve training on non-IID data. In *2020 international joint conference on neural networks (IJCNN)* (pp. 1-9). IEEE.
[13] Nedic, A., Olshevsky, A., Ozdaglar, A., & Tsitsiklis, J. N. (2008, December). Distributed subgradient methods and quantization effects. In *2008 47th IEEE conference on decision and control* (pp. 4177-4184). IEEE.

ISSN: 2456-2319
https://dx.doi.org/10.22161/eec.86.2