Critical Synthesis of "Swarm Learning: A Survey of Concepts, Applications, and Trends"

Samir Nait Sadi, Arthur Macdonald

Index Terms—Swarm Learning, Blockchain, Distributed Machine Learning, Federated Learning, Privacy Preservation, Edge Computing

1 Introduction

Swarm Learning (SL) represents a new approach to distributed machine learning that addresses critical limitations of traditional centralized and federated learning paradigms, as presented in the survey by Elham Shammar, Xiaohui Cui, and Mohammed A. A. Al-qaness [1]. As the Internet of Things (IoT) continues to expand and data privacy regulations become increasingly rigid, the need for decentralized, privacy-preserving machine learning frameworks has become predominant. Unlike Federated Learning (FL), which relies on a central server for parameter aggregation, SL leverages blockchain technology to enable peer-to-peer model training and secure parameter sharing across edge nodes [1], [2]. In this synthesis, we summarize the state-of-the-art in SL technology as reported by the authors, examining its architecture, real-world applications across multiple sectors, and challenges that define current and future research directions.

2 Context and Related Work

The survey positions SL within the broader landscape of distributed learning approaches by comparing it to FL, Decentralized FL (DFL), and Swarm Intelligence (SI) [4], [5]. While traditional FL relies on a central server, the authors note that SL and DFL move towards a peer-to-peer structure. They also make a key distinction between SL and SI; while they share biological inspiration, SI focuses on optimization algorithms (such as Ant Colony Optimization or Particle Swarm Optimization) rather than training machine learning models.

The authors highlight several fundamental structural differences between these paradigms. Architecturally, while FL relies on a star topology with a central bottleneck, SL utilizes a peer-to-peer mesh secured by blockchain. This integration provides SL with high fault tolerance and a rigorous security framework (via immutable logs and access control) that standard DFL typically lacks [5]. Furthermore, regarding coordination, SL employs dynamic leader election algorithms rather than a fixed server (FL) or simple gossip protocols (DFL), ensuring that control remains decentralized yet secure.

Université de Bordeaux, Master 2, 2025–2026.
Synthesis of the article "Swarm Learning: A Survey of Concepts, Applications, and Trends" by E. Shammar, X. Cui, and M. A. A. Al-qaness [1].

3 Contribution

We identify the survey's core contribution as the consolidation and structuring of a rapidly growing and fragmented body of work on Swarm Learning. It provides a conceptual framework for SL, compares it with related paradigms, and synthesizes applications, challenges, and future directions.

3.1 Architecture, Components, and Features

The article describes SL as a distributed learning technology that relies on local model training and secure parameter sharing via a blockchain [1]. It is important to note that, according to the authors, the blockchain is not merely an optional feature but a defining component of SL. While simpler decentralized aggregation methods exist (often categorized as standard DFL), the survey argues that blockchain is essential in the SL environment to replace the central coordinator with smart contracts. This provides an immutable ledger for auditability and trust, which the authors claim is necessary for secure enterprise-grade collaboration, distinguishing SL from simple peer-to-peer parameter exchange.

The authors explain that this method guarantees data confidentiality and improves security through a comprehensive set of distinctive features. SL encompasses five core characteristics: **Privacy Preservation** (data remains at each node, minimizing breach risks), **Decentralization** (eliminates central authority and single points of failure), **Continuous Learning** (models adapt to new data and evolving conditions), **Data Diversity and Volume** (handles heterogeneous datasets from multiple sources), and **Collaborative Learning** (nodes share insights without exposing raw data). These features position SL as a strong framework for privacy-sensitive domains.

According to the survey [1], Swarm Learning is designed to eliminate the central server entirely, unlike Federated Learning which often still relies on a central aggregator. In SL, data and parameters remain at the edge. The architecture

is built upon two primary layers: an **application layer** (containing the ML platform, blockchain, and SL Library) and an **infrastructure layer** (hardware and data sources) [2], [3].

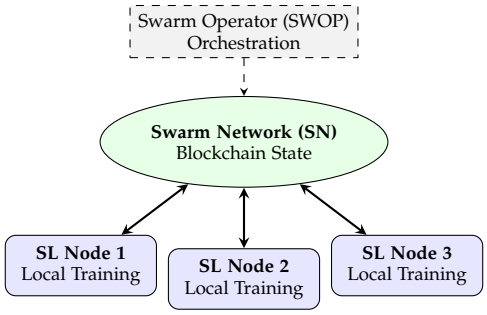

Fig. 1. Swarm Learning Architecture. The Swarm Network (SN) connects distributed SL nodes via blockchain, while the Swarm Operator (SWOP) handles orchestration.

As illustrated in Figure 1, the architecture relies on three primary node types: **Swarm Learning (SL)** nodes that train local models, the **Swarm Network (SN)** which uses blockchain to track state metadata (without sharing raw parameters), and the **Swarm Operator (SWOP)** for orchestration. Additional components, such as the SPIRE Server for identity verification and the License Server (LS), support the framework's security and management operations.

The learning process, depicted in Figure 2, is driven by a dynamic **Leader Election Algorithm (LEA)**. Instead of a permanent central server, a temporary leader is elected via blockchain consensus (e.g., Proof of Stake) for each training cycle to merge model parameters. This cyclical process of training, synchronization, election, and aggregation eliminates single points of failure [3].

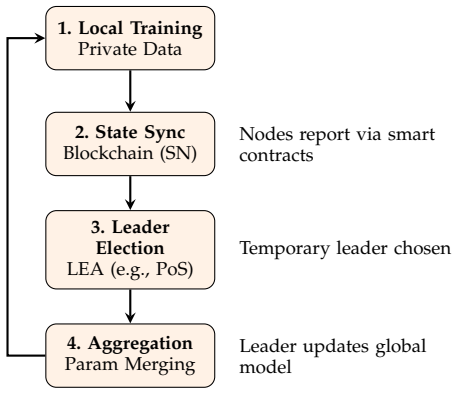

Fig. 2. The Swarm Learning Training Cycle. Iterative process of local training, blockchain synchronization, leader election, and parameter aggregation.

3.2 Applications as Evidence of Contribution

Beyond the conceptual framework, the survey provides an extensive mapping of SL applications across multiple domains, which constitutes an important part of its contribution. The authors identify **Healthcare** as a primary domain, where SL enables COVID-19 and cancer diagnosis while complying with GDPR and HIPAA regulations. In **Transportation** and **Energy**, SL facilitates traffic management and PV forecasting by leveraging distributed sensor data. Other critical applications include **Industry** (predictive maintenance, digital twins), **Finance** (P2P credit scoring, fraud detection), and **Smart Homes** (anomaly detection). Emerging fields such as the **Metaverse** and **Robotic Systems** are also highlighted, where SL supports 6G communication and swarm robot coordination [1].

3.3 Identified Challenges and Future Directions

A further contribution of the article is to synthesize the main technical and methodological obstacles that currently limit SL adoption. While SL enhances privacy, the survey categorizes critical hurdles into several key areas. **Integrity Attacks**, such as data poisoning and backdoors, pose severe security risks. **Network Availability** is threatened by Eclipse attacks (node isolation) and Denial of Service. Additionally, **Privacy Leakage** remains a concern through model inversion attacks, while **Data Heterogeneity** (non-IID data) and **Fairness** issues complicate the training of unbiased global models [1], [6], [7].

Addressing these challenges defines the roadmap for future development. The authors propose prioritizing research into **Security & Trust**, specifically suggesting the integration of Homomorphic Encryption. **Scalability** and **Node Management** are cited as critical for industrial adoption, requiring optimized communication protocols and dynamic incentive mechanisms. Finally, the survey emphasizes the need for **Resource-Efficient** ("Green") algorithms and robust handling of **Non-IID Data** to ensure the long-term viability of Swarm Learning systems.

4 Discussion

In this section, we critically evaluate the survey itself, focusing on its contributions, limitations, and methodological choices.

4.1 Assessment of the Article's Contributions

We consider a major strength of the article to be its range. It covers SL concepts, architecture, security aspects, and a wide range of real applications. By consolidating a large number of primary studies into a single structured narrative, the survey significantly lowers the entry barrier for researchers new to the topic.

The organization of the paper is also a positive point. The authors clearly separate conceptual foundations (architecture, components, features) from application domains and from the discussion of challenges and future directions. This makes the survey easy to navigate and allows readers to focus on the part most relevant to their interests.

Finally, the explicit comparison between SL, FL, DFL, and SI helps clarify terminology in a field where names are potentially confusing. This positioning with respect to existing paradigms is consistent with the goal of a state-of-the-art paper.

4.2 Critical Observations and Missing Aspects

Despite its comprehensive scope, we note that the survey remains largely descriptive. Most of the analysis is qualitative, and the article does not provide systematic quantitative comparisons (for example, common benchmarks or performance metrics) between SL and alternative distributed learning frameworks. Consequently, we find it difficult to assess how much of the claimed benefit of SL over FL or DFL holds across different tasks and datasets.

Another limitation we observed is that the paper devotes relatively little space to practical deployment aspects. Issues such as engineering complexity, operational costs of maintaining a blockchain infrastructure, and integration within existing industrial pipelines are mentioned but not analyzed in depth.

4.3 Link to Challenges and Future Directions

We acknowledge that the survey appropriately identifies a wide range of open challenges and research directions, especially around security, non-IID data, fairness, and resource efficiency. The structured presentation in tables (challenges and future work) gives a useful "research roadmap" for the community.

Nevertheless, we argue that the link between this roadmap and the earlier methodological analysis could be tighter. For instance, the article could have more explicitly prioritized research gaps based on the number or quality of existing works, or highlighted which domains (such as healthcare vs. transportation) are more mature and which remain underexplored.

Overall, the survey offers a broad and well-organized entry point into Swarm Learning and usefully maps its challenges and future directions, but we believe its contribution would be stronger with systematic quantitative evidence, a clearer prioritization of research gaps, and more detailed analysis of practical deployment constraints.

5 Conclusion

Swarm Learning (SL) emerges as a comprehensive decentralized machine learning framework that integrates blockchain technology with edge computing to ensure data confidentiality, model integrity, and resilience against system failures [1], [2]. While SL's decentralized architecture and blockchain integration provide substantial advantages in privacy, security, and fault tolerance [4], [5], the technology faces formidable challenges including sophisticated security threats (poisoning attacks, backdoors, inference attacks), handling strongly non-IID data distributions, ensuring fairness across heterogeneous participants, and managing the computational and communication overhead of blockchain consensus mechanisms [6], [7]. These challenges delineate SL as an evolving research field rather than a mature technology. Moving forward, we suggest that crucial research avenues include developing lightweight consensus protocols to reduce energy consumption, exploring synergies with homomorphic encryption for deeper privacy, and establishing interoperability standards to facilitate large-scale industrial deployment.

References

[1] E. Shammar, X. Cui, and M. A. A. Al-qaness, "Swarm Learning: A Survey of Concepts, Applications, and Trends," *arXiv preprint arXiv:2405.00556v2*, 2025.

[2] S. Warnat-Herresthal *et al.*, "Swarm learning for decentralized and confidential clinical machine learning," *Nature*, vol. 594, no. 7862, pp. 265–270, 2021.

[3] J. Han, Y. Ma, and Y. Han, "Demystifying swarm learning: A new paradigm of blockchain-based decentralized federated learning," *arXiv preprint arXiv:2201.05286*, 2022.

[4] E. T. Martínez Beltrán *et al.*, "Decentralized federated learning: Fundamentals, state of the art, frameworks, trends, and challenges," *IEEE Communications Surveys & Tutorials*, 2023.

[5] H. Zhang, S. Jiang, and S. Xuan, "Decentralized federated learning based on blockchain: Concepts, framework, and challenges," *Computer Communications*, vol. 216, pp. 140–150, 2024.

[6] K. Chen *et al.*, "Backdoor attacks against distributed swarm learning," *ISA Transactions*, 2023.

[7] H. A. Madni, R. M. Umer, and G. L. Foresti, "Blockchain-based swarm learning for the mitigation of gradient leakage in federated learning," *IEEE Access*, vol. 11, pp. 65491–65556, 2023.