Intrusion Detection System based on Quantum Generative Adversarial Network and Blockchain-based Federated Learning

Franco Cirillo
University of Salerno
Fisciano (SA), Italy
fracirillo@unisa.it

Christian Esposito
University of Salerno
Fisciano (SA), Italy
esposito@unisa.it

***Abstract*—Quantum Generative Adversarial Networks (QGANs) represent a promising solution for anomaly detection in network traffic, with the overarching goal of tackling the increasing complexity and sophistication of cyber threats in Next-Generation Internet infrastructures. Given the inherent noise and instability of current quantum devices, as well as the massive workload of current network flow traces, developing robust and scalable Quantum Machine Learning (QML) models that maintain performance under noisy conditions is a critical challenge. A solution focuses on designing and evaluating QGAN architectures for federated learning (F-QGAN) to address scalability and robustness. This paper presents a solution to integrate blockchain into this federated approach, enabling asynchronous implementation and improving coordination among local learners who cannot exchange weights with the centralized aggregator synchronously.**

***Index Terms*—Quantum Machine Learning, Blockchain, Federated Learning**

I. Introduction

Quantum Generative Adversarial Networks (QGANs) [1] represent a promising and innovative direction that extends classical GAN architectures to the quantum domain. They often require fewer parameters and can perform effectively even when training data is limited, making them particularly suitable for cybersecurity contexts where labelled data may be scarce or imbalanced. In [2], we have presented the design and implementation of our Quantum Generative Adversarial Network (QGAN)-based intrusion detection system, whose final circuit structure, integrating both the encoding and the ansatz components, is illustrated in Figure 1. Such a solution has been validated on real network flow datasets, providing good performance compared with traditional approaches.

While these simulations provide valuable insights, executing QGANs on actual quantum hardware is essential to assess the impact of quantum noise on model accuracy fully. However, access to real quantum hardware remains limited and cost-prohibitive, posing a significant challenge for practical experimentation. This is particularly demanding for QML algorithms, which require substantial resources because of the large number of circuit evaluations required during training.

In [3], we improved our work by exploiting distributed Machine Learning, which enhances scalability, efficiency, and privacy, making it ideal for large-scale, real-time intrusion detection [10]. The integration with QGANs can improve the resiliency and efficiency of these models. By distributing the workload across multiple nodes, this approach enhances load balancing, ensuring that computational demands are evenly distributed and managed. Additionally, it can mitigate the impact of quantum noise in NISQ devices by enabling quantum resources to collaborate. Specifically, we have presented the design and implementation of our Quantum Generative Adversarial Network (QGAN)-based intrusion detection system in a federated setting (F-QGAN), as illustrated in Figure 2. We demonstrated that each quantum client can locally train a lightweight QGAN model and participate in the collaborative learning of F-QGAN without sharing raw data. We also evaluated the model's robustness under realistic quantum noise, demonstrating its resilience and performance across various noise profiles and hardware configurations.

In this paper, we present a further extension of our framework by introducing a blockchain-integrated architecture designed to enhance the coordination, trust, and auditability of the federated quantum learning process. Although federated learning offers strong privacy guarantees by keeping data locally, it still relies on a shared infrastructure for model aggregation and validation. In classical FL systems, this role is often fulfilled by a central server or semi-trusted aggregators. However, such assumptions are problematic in adversarial or large-scale environments, particularly for quantum systems, where trust boundaries are more rigid and evaluation is more expensive.

Blockchain technology addresses these limitations by providing a decentralized, tamper-resistant ledger that can manage the federated workflow without a central authority. By embedding blockchain into our F-QGAN framework, we aim to solve several open problems. First, it enables transparent recording of model updates, evaluation scores, and client reputations, which are essential in settings where validation is costly or unreliable. Second, it enables trustless collaboration among

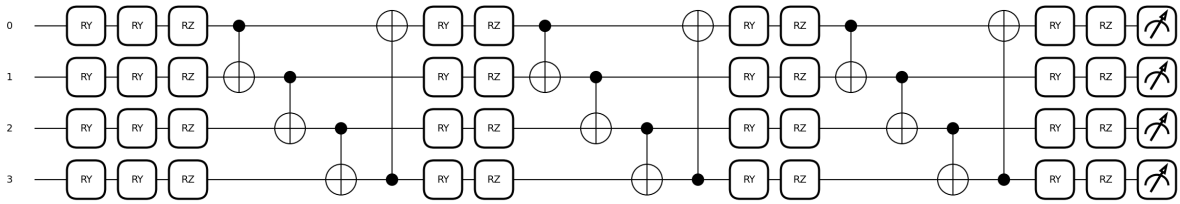

Fig. 1: The generator circuit consists of a parameterized RY encoding layer that embeds the latent vector, followed by an EfficientSU2 ansatz with depth 4.

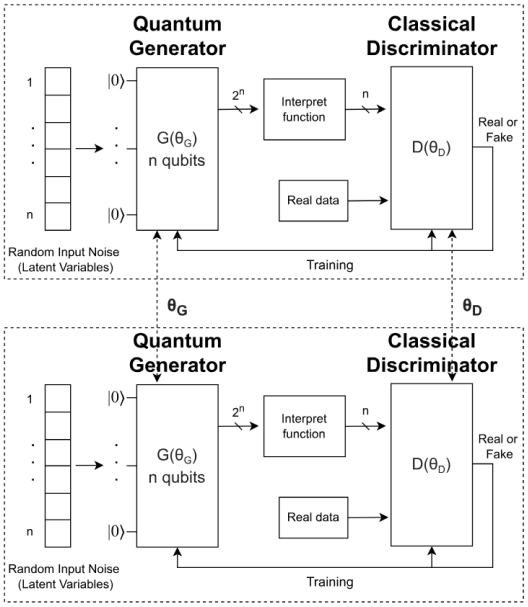

Fig. 2: Training process of the proposed F-QGAN.

nodes by leveraging smart contracts to automate classification, reward distribution, and access control, reducing the risk of manipulation or freeloading. Third, blockchain supports the integration of incentive mechanisms that reward clients based on their contributions to the quality of the global model and their efficient use of resources. Finally, by combining blockchain with off-chain storage, we achieve a scalable, auditable, and verifiable model for sharing without overloading the chain itself.

The paper is organized as follows: We begin with a state-of-the-art review that analyzes current approaches to integrating blockchain with FL and discusses their suitability for quantum environments. The core of the paper is the Architecture section, where we present our fully integrated framework, describing the structure and function of each layer: quantum machine learning, blockchain coordination, and federated workflow. We conclude by outlining directions for future work.

II. State of the Art

The integration of FL with blockchain technologies has led to the development of various hybrid architectures aimed at improving the trustworthiness, traceability, and decentralization of collaborative machine learning. Existing systems vary in their degree of devolution, consensus protocols, incentive mechanisms, and methods for preserving privacy and integrity. However, these approaches present limitations when applied to the more constrained and fragile context of Quantum FL (QFL), which inherits the complexities of both quantum computation and distributed learning.

Several FL-blockchain integration strategies adopt a partially decentralized architecture [4], [5], [6], in which selected super nodes are responsible for maintaining the blockchain ledger and validating transactions. These super nodes are typically equipped with stronger computation and communication capabilities, enabling faster consensus and reduced overhead. However, this design introduces a degree of centralization, thereby sacrificing transparency and fairness and potentially contradicting the foundational goals of fully distributed systems. In contrast, fully decentralized frameworks [7], [8] assume that all clients act symmetrically, participating both as model trainers and blockchain validators or miners. While this approach improves fairness and transparency, it often relies on the unrealistic assumption that all clients possess equivalent computational power. In practice, client heterogeneity (primarily in quantum systems) challenges this assumption and complicates the design of fair and efficient systems. Different consensus mechanisms have been explored to reconcile performance, fault tolerance, and security. Protocols such as Proof of Stake (PoS) [8], Practical Byzantine Fault Tolerance (PBFT) [6], and Raft have been used in FL-blockchain systems to ensure agreement on model updates and reputation scores. In some systems [9], randomness-based committee rotation mechanisms are used to dynamically select a subset of nodes responsible for consensus in each round. While this enhances security by limiting exposure to Sybil attacks, it assumes a low proportion of malicious nodes and may degrade under adversarial settings.

More advanced mechanisms have also been proposed. Proof of Reputation (PoR) [10] restricts each node to validating only one other node during consensus, substantially reducing validation latency. However, this strategy performs poorly in heterogeneous environments where reputation scores may not correlate well with actual contribution quality. Another method, Proof-of-Federalism (PoF) [11], selects the most fre-

quently adopted global model during cross-validation rounds. The node whose model is most selected over time becomes the consensus winner. While conceptually aligned with FL's objective, this model is vulnerable to manipulation in noisy or unbalanced datasets. A more recent trend is the use of multi-aggregator frameworks, in which model aggregation is not centralized but distributed among several trusted or semi-trusted participants. This helps mitigate single-point failures and reduces computational bottlenecks.

Incentive mechanisms are crucial for maintaining participation and fairness in collaborative environments. Several architectures provide rewards based on model accuracy, data quality, and resource contribution. Reward structures must be carefully designed to balance incentives, fairness, and resistance to manipulation.

Efficiency is further improved by lightweight blockchain networks, such as micro-chains, which reduce overhead by narrowing the scope of information recorded per round. These structures maintain local consensus without needing a full-chain update, significantly improving scalability. However, their security guarantees under adversarial scenarios remain an open question. As for storage, systems increasingly rely on decentralized solutions such as IPFS [12], not only to alleviate the burden of on-chain storage but also to support model auditing and versioning. Techniques like chunking and content addressing further reduce storage overhead and enhance retrieval performance.

While many of these solutions are effective in classical FL settings, applying them to QFL introduces additional challenges. Synchronizing training rounds is inherently more difficult due to variable quantum decoherence and gate noise, which make local training and validation expensive and unreliable. The cost and feasibility of quantum model validation remain critical bottlenecks, as fidelity evaluation and error-mitigation techniques are still in the early stages of maturity.

III. Architecture

The proposed system introduces a fully integrated framework for Federated QGANs tailored for intrusion detection, in which blockchain is embedded directly into a decentralized peer-to-peer topology, as depicted in Figure 3. This architecture comprises three interdependent layers: the quantum machine learning layer, the blockchain layer, and the federated learning workflow. Together, they form a cohesive and resilient design that eliminates the need for centralized coordination while maintaining security, trust, and efficiency.

A. Quantum Machine Learning Layer

At the foundation of the system lies the quantum machine learning layer. Each node in the network operates its own Quantum GAN instance, consisting of both a generator and a discriminator. These models are trained locally, allowing each node to learn from its own data without ever sharing raw samples. This decentralized learning structure not only enhances privacy but also reduces communication requirements. Quantum GANs are particularly effective in this context

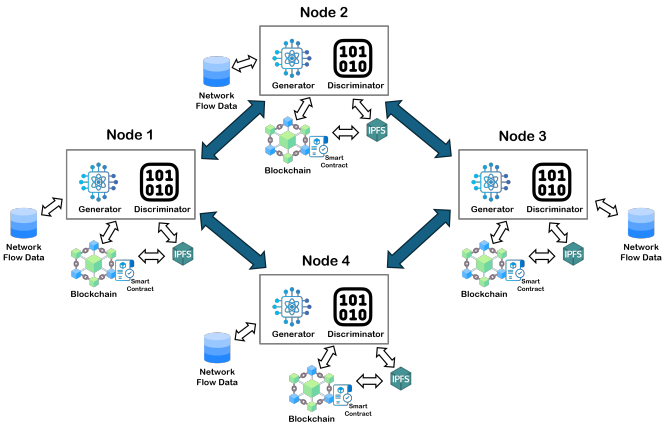

Fig. 3: Comprehensive architecture of the integration of F-QGAN with blockchain

because they naturally reduce the dimensionality of input data and require fewer trainable parameters. This compactness minimizes the data exchanged during model updates, making the approach more suitable for federated environments. Furthermore, QGANs are adept at handling imbalanced datasets, which are common in intrusion detection, by generating gate fidelity and data that strengthen local learning.

B. Blockchain Layer

To coordinate collaboration among nodes, the system integrates a blockchain layer that enables decentralized control and validation. All clients in the network are treated as equal peers, with comparable quantum computational resources, including gate fidelity and the number of qubits. This symmetry justifies the use of a fully distributed architecture, avoiding any form of centralized aggregation or orchestration.

Due to the storage limitations and performance constraints of blockchains, storing full quantum model parameters or gradients on-chain would be impractical and costly. To address this, the system decouples model storage from the blockchain by using the InterPlanetary File System (IPFS) [12]. IPFS is a decentralized, peer-to-peer file storage protocol that enables content-addressed data storage. When a node uploads its model to IPFS, the system generates a unique content identifier (CID) based on the cryptographic hash of the file. This CID serves as a tamper-evident pointer to the file's exact version and content, and is the only reference stored on the blockchain.

This separation brings several advantages. First, it significantly reduces the on-chain data footprint, thereby minimizing gas costs and blockchain congestion. Second, it enhances auditability and reproducibility, as the CID guarantees that any node retrieving the model receives the exact version used in prior rounds. Third, IPFS enables distributed load balancing and content persistence across the network, eliminating the need for a centralized file server. Nodes can cache popular models locally or retrieve them from nearby peers, improving latency and fault tolerance.

The blockchain itself functions as a transparent and immutable layer of coordination. A smart contract deployed on the chain governs the lifecycle of training rounds by managing model submissions, registering updates, and granting access to peer evaluations. After each round, nodes submit performance metrics such as grades or similarity scores, which are also recorded on-chain. These values are later aggregated to update trust levels across participants.

A generic public consensus mechanism secures the blockchain, ensuring that all nodes agree on the system's state without relying on privileged actors. To further reduce operational costs, the system minimizes the amount of data written to the chain, relying on IPFS to handle the bulk of model-related storage. The compact architecture of QGANs contributes to this efficiency by requiring fewer trainable parameters and reducing the size and frequency of updates that must be referenced or retrieved.

C. Federated QML Workflow

The federated quantum learning workflow connects the computational and blockchain layers into a unified loop. Each node not only trains a model but also validates those received from peers. At the end of each training cycle, a node retrieves models from other participants using IPFS links referenced in the blockchain. These models are then aggregated through a linear combination, where the contribution of each peer is weighted by a grade that reflects its previous performance. This grade is based on the similarity between the last local weights and the aggregated model's results.

After training for a fixed number of epochs, each node evaluates the alignment between the newly obtained weights and its previous state. Based on this comparison, it assigns updated grades to its peers, quantifying each peer's contribution to convergence or divergence. These grades are submitted to the blockchain smart contract, which aggregates them periodically and updates each node's reputation. Over time, this process creates a trust-based structure that reinforces cooperation and penalizes unhelpful or dishonest behavior. To enforce reliability, the system implements a threshold-based mechanism. Nodes whose grades fall below a warning threshold are flagged, while those that consistently underperform are eventually excluded from participation. This process allows the network to self-regulate and adapt without any external validators.

Figure 4 illustrates the convergence of the proposed QFL framework. The plot shows the evolution of their overall accuracy over rounds, demonstrating that the federated mechanism yields smoother convergence and greater stability than isolated training. This stability arises from the continuous exchange and aggregation of model updates, thereby mitigating local fluctuations and reinforcing global learning dynamics.

D. A Unified and Efficient QFL Framework

This architecture is designed to eliminate the overhead in traditional federated learning frameworks, in which training, validation, and aggregation are treated as separate steps. In the proposed system, these processes are unified by integrating blockchain and quantum learning. The result is a highly efficient and transparent framework, well-suited for secure and scalable intrusion detection in quantum-enabled environments.

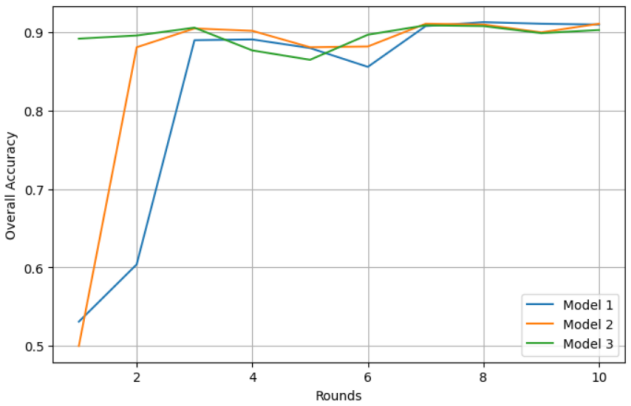

Fig. 4: Convergence plot of three models in the proposed F-QGAN framework.

IV. Conclusion

In this paper, we have presented the integration of blockchain technology into our federated Quantum GAN-based intrusion detection system. Building on previous work that demonstrated the feasibility and noise robustness of quantum-enhanced models in a federated setting, we proposed a fully decentralized architecture where each node participates as both a quantum learner and a blockchain validator.

Blockchain integration provides a secure, transparent coordination layer, enabling trustless collaboration among peers without relying on a central aggregator. By storing only model metadata and performance metrics on-chain, and using IPFS for off-chain model storage, the system achieves scalability and auditability with minimal overhead. Smart contracts govern training rounds, reputation updates, and incentive mechanisms, ensuring a fair and efficient learning process across heterogeneous nodes.

This framework addresses key challenges in distributed quantum learning, such as validation cost, reputation management, and secure model exchange.

Acknowledgements

This research was partially funded by the NGIsargasso project (European Horizon Grant No. 101092887), the Open Call 4 FRQGAN4AD project, and the SERICS project (PE00000014) under the NRRP MUR program funded by the EU - NGEU.

References

[1] L. Beshaj and G. Tyagi, "Quantum generative adversarial networks (qgans) using quantum kernel learning for discriminators," in *Quantum Information Science, Sensing, and Computation XVII*, vol. 13451. SPIE, 2025, pp. 66–71.

[2] F. Cirillo and C. Esposito, "Intrusion detection system based on quantum generative adversarial network," in *Proceedings of the 17th International Conference on Agents and Artificial Intelligence, ICAART 2025 - Volume 1, Porto, Portugal, February 23-25, 2025*, 2025, pp. 830–838.

[3] ——, "Intrusion detection using quantum generative adversarial networks: a federated approach with noisy simulators," in *IET Space and Communications Conference 2025*, vol. 2025. IET, 2025, pp. 31–35.

[4] C. Fang, Y. Guo, J. Ma, H. Xie, and Y. Wang, "A privacy-preserving and verifiable federated learning method based on blockchain," *Computer Communications*, vol. 186, pp. 1–11, 2022.

[5] C. Feng, B. Liu, K. Yu, S. K. Goudos, and S. Wan, "Blockchain-empowered decentralized horizontal federated learning for 5g-enabled uavs," *IEEE Transactions on Industrial Informatics*, vol. 18, no. 5, pp. 3582–3592, 2021.

[6] S. Guo, K. Zhang, B. Gong, L. Chen, Y. Ren, F. Qi, and X. Qiu, "Sandbox computing: A data privacy trusted sharing paradigm via blockchain and federated learning," *IEEE Transactions on Computers*, vol. 72, no. 3, pp. 800–810, 2022.

[7] Y. Qu, S. R. Pokhrel, S. Garg, L. Gao, and Y. Xiang, "A blockchained federated learning framework for cognitive computing in industry 4.0 networks," *IEEE Transactions on Industrial Informatics*, vol. 17, no. 4, pp. 2964–2973, 2020.

[8] A. Wahrstätter, S. Khan, and D. Svetinovic, "Openfl: A scalable and secure decentralized federated learning system on the ethereum blockchain," *Internet of Things*, vol. 26, p. 101174, 2024.

[9] Y. Qu, C. Xu, L. Gao, Y. Xiang, and S. Yu, "Fl-sec: Privacy-preserving decentralized federated learning using signsgd for the internet of artificially intelligent things," *IEEE Internet of Things Magazine*, vol. 5, no. 1, pp. 85–90, 2022.

[10] D. C. Nguyen, M. Ding, P. N. Pathirana, A. Seneviratne, and A. Y. Zomaya, "Federated learning for covid-19 detection with generative adversarial networks in edge cloud computing," *IEEE Internet of Things Journal*, vol. 9, no. 12, pp. 10 257–10 271, 2021.

[11] Y. Qu, L. Gao, Y. Xiang, S. Shen, and S. Yu, "Fedtwin: Blockchain-enabled adaptive asynchronous federated learning for digital twin networks," *IEEE Network*, vol. 36, no. 6, pp. 183–190, 2022.

[12] J. Benet, "Ipfs - content addressed, versioned, p2p file system," 2014. [Online]. Available: https://arxiv.org/abs/1407.3561