INTRUSION DETECTION USING QUANTUM GENERATIVE ADVERSARIAL NETWORKS: A FEDERATED APPROACH WITH NOISY SIMULATORS

Franco Cirillo[1*], Christian Esposito[1]

[1]Department of Computer Science, University of Salerno, Fisciano (SA), Italy
**fracirillo@unisa.it*

Keywords: QUANTUM MACHINE LEARNING (QML), QUANTUM GENERATIVE ADVERSARIAL NETWORK (QGAN), ANOMALY DETECTION, FEDERATED LEARNING (FL), QUANTUM NOISY SIMULATORS, NOISY INTERMEDIATE-SCALE QUANTUM (NISQ).

Abstract

In today's increasingly complex digital environments, Intrusion Detection Systems (IDS) play a crucial role in ensuring network security. Traditional machine learning approaches struggle with challenges such as handling high-dimensional data and maintaining performance on imbalanced datasets. Generative Adversarial Networks (GANs) offer a viable alternative by enhancing data generation, but their conventional implementations are computationally intensive and strive with capturing intricate data distributions. Quantum GANs (QGANs) leverage quantum computing to address these limitations, while a distributed approach enhances load balancing. Tested on the NSL-KDD dataset, the proposed model effectively learns the distribution of benign data using a federated hybrid QGAN architecture, which integrates quantum generators with classical discriminators. Additionally, the model has been evaluated on a quantum noisy simulator to assess performance variations under noise conditions.

1 Introduction

Intrusion Detection Systems (IDS) are critical for protecting digital infrastructures by identifying network malicious activities, unauthorized access, and policy violations [1]. However, as cyberattacks become increasingly sophisticated, traditional detection methods struggle to keep pace. Anomaly detection is a powerful approach to intrusion detection, as it relies on establishing a baseline of normal behavior rather than predefined attack signatures [2]. While effective in identifying novel attack vectors, it faces challenges such as high false positive rates and computational complexity.

Machine Learning (ML) has significantly advanced intrusion detection. However, ML's reliance on high-quality datasets and substantial computational resources presents scalability issues [3]. A key breakthrough in ML is the application of Generative Adversarial Networks (GANs) [4], which enhance intrusion detection by generating synthetic data, addressing class imbalances, and detecting subtle anomalies. Nonetheless, classical ML models, including GANs, struggle with the complexity of high-dimensional data [5].

Quantum Machine Learning (QML) offers a transformative solution to these limitations [6]. By leveraging quantum computing, QML can improve scalability and performance beyond classical ML. Specifically, Quantum
Quantum Generative Adversarial Networks (QGANs) extend classical GANs by incorporating quantum computing principles, allowing for more efficient modeling of complex data distributions [7]. A hybrid QGAN consists of two components. The first one is a Quantum Generator (G), a parameterized quantum circuit (PQC) that generates quantum states encoding data samples drawn from a prior distribution. The generator iteratively updates its parameters to approximate the target data distribution. The second component is a Classical Discriminator (D), a neural network that assesses the similarity between real and generated samples, assigning a probability score to distinguish between them.

However, real-world quantum hardware in the Noisy Intermediate-Scale Quantum (NISQ) era presents significant challenges, including decoherence, qubit errors, and gate imperfections. As a result, QML models tested under ideal conditions may not accurately reflect their real-world performance. Additionally, testing on actual quantum hardware is costly. A practical alternative is the use of noisy simulators, such as those developed by IBM, which replicate the behavior of real quantum devices under noisy conditions. These simulators enable researchers to evaluate and refine QML models without the high costs associated with physical quantum computers.

This paper presents the following key contributions:

- Development of a federated QGAN-based model specifically designed for intrusion detection.
- Experimental validation across multiple configurations on the NSL-KDD dataset, achieving an accuracy of 0.9125 and an F1-score of 0.9034, while reducing time complexity compared to the non-federated version.

- Evaluation on a noisy simulator and comparison with the ideal (noiseless) version.

To the best of our knowledge, this is the first study that evaluates a federated QGAN for intrusion detection under noisy conditions.

2 Related works

Recent advancements in machine learning, particularly GANs, have enhanced Network Intrusion Detection Systems by generating synthetic data to improve training. Studies such as [8] and [9] demonstrated how GANs enhance IDS accuracy and address class imbalance, outperforming traditional oversampling methods.

QGANs [10,11] have emerged as a promising alternative, particularly in image generation, with potential advantages over classical GANs. However, their application to anomaly detection remains underexplored. In [12], QGANs were tested for particle detection, achieving 0.88 accuracy on noiseless simulators, but their generalizability was limited. Similarly, [13] introduced the SuDaI encoding technique for high-dimensional data, evaluating their approach on the NAB dataset, though without explicit performance metrics. Challenges related to quantum generator output discreteness were noted in [14], yet the study lacked detailed implementation and testing results. A hybrid quantum-classical anomaly detection approach using Variational Quantum-Classical WGANs was explored in [15], achieving an F1-score of 0.85 on a credit card fraud dataset, though its applicability to complex scenarios was limited.

Federated learning has been applied to classical GANs for model convergence [16], but research on Federated QGANs is still scarce. Early studies, such as [17], implemented a federated QGAN in a different context without reporting key performance metrics. Another work [18] introduced an unsupervised federated QGAN for optimizing wireless networks, but its reliance on quantum state transmission made it unsuitable for classification tasks and lacked accuracy evaluations.

At present, comparing the performance of Quantum Machine Learning (QML)-based approaches with classical federated learning models is not particularly insightful. For instance, the study in [19] has already achieved 98-99% accuracy on the same dataset of this work. The current focus should be on refining QML models so they can become viable once quantum hardware reaches a more advanced stage.

Despite advancements in QGAN research, their application to intrusion detection remains largely unexplored, with limited studies providing comprehensive performance evaluations. There is a need to systematically assess Federated QGANs in this domain, focusing on scalability, robustness, and real-world feasibility. This work aims to bridge this gap by evaluating their performance on high-dimensional intrusion detection datasets, analyzing key metrics such as accuracy and F1-score, and investigating the impact of quantum noise.

3 Methodology

3.1 QGAN structure

A specialized QGAN model is proposed for intrusion detection, incorporating a custom interpret function and a tailored quantum circuit. As depicted in Fig. 1, the model consists of two main components: a quantum generator implemented as a Variational Quantum Circuit (VQC) and a classical discriminator modeled as a neural network.

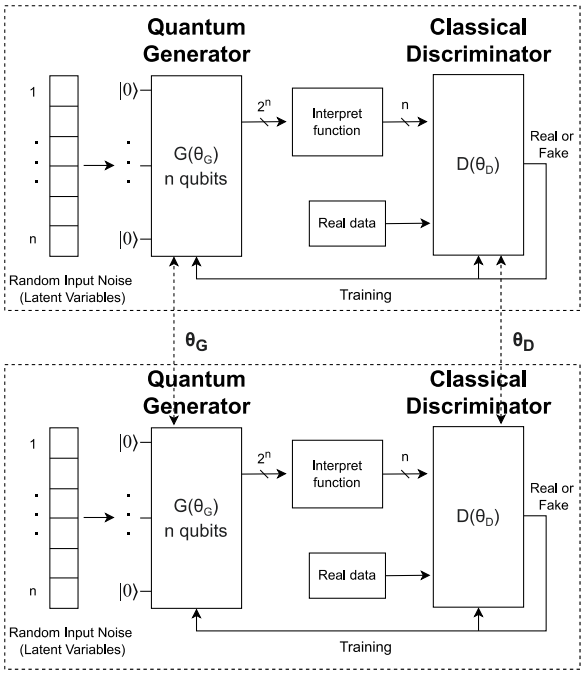

Fig. 1 Training process of the proposed federated QGAN

The training process starts with a random noise vector, which is used by the generator to produce diverse synthetic samples. The generator begins from the initial quantum state $|0\rangle$ and operates on n qubits, which is the same number of features obtained applying the Principal Component Analysis (PCA) to the NSL-KDD dataset.

Through parameterized quantum gates, controlled by trainable parameters θ_G, the generator constructs the quantum state $G(\theta_G)|0\rangle^{\otimes n}$. Since direct classical measurement of the quantum state is not feasible, an interpret function is used to map quantum outputs to a classical representation.

Meanwhile, the classical discriminator receives input data, either real or synthetic, and assigns a probability score reflecting the likelihood of an instance being an attack. The adversarial training mechanism ensures that the generator refines its output to closely resemble real data, while the discriminator enhances its ability to differentiate between authentic and generated samples.

3.2 Quantum Generator and Classical Discriminator architecture

The generator employs a variational quantum circuit, which generates synthetic data from a noise input. One of the tested

circuit configurations is depicted in Fig. 2. Each qubit in the circuit corresponds to a feature of the dataset. Initially, Hadamard gates are applied to all qubits, creating a superposition state that ensures a uniform probability distribution. A feature mapping layer, such as ZZFeatureMap, encodes latent variables into the quantum state, facilitating entanglement between qubits. The trainable component, structured using the EfficientSU2 ansatz, consists of single-qubit rotation gates and two-qubit entangling gates. Repeating these layers enhances the circuit's capacity to approximate complex distributions, with the number of repetitions acting as a key hyperparameter.

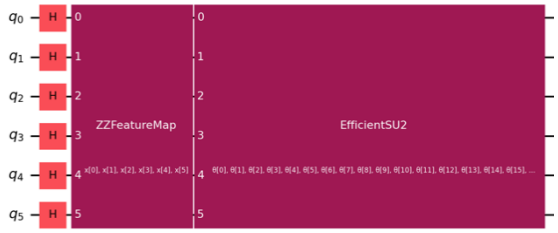

Fig. 2 Generator quantum circuit

The generator's output, represented as quantum measurements, is converted into a lower-dimensional classical vector. Each possible n-qubit measurement outcome corresponds to a probability value. The interpret function computes the marginal probability p_i for each qubit i, given by:

$$\boldsymbol{p_i} = \sum_{\boldsymbol{x}:\boldsymbol{x_i}=\mathbf{1}} \boldsymbol{P(x)}$$

where the summation includes all 2^n states where qubit i is in state 1. The resulting vector $(p_1,p_2,\ldots,p_n)$ forms a classical feature representation suitable for training.

The discriminator is implemented as a fully connected neural network with multiple hidden layers and LeakyReLU activations, ensuring stable gradient propagation. A sigmoid function in the output layer assigns a probability score indicating whether the input is real or synthetic. The network is trained using a Binary Cross-Entropy loss function, improving iteratively through adversarial feedback with the generator.

3.3 Federated Training approach

To enhance scalability and robustness, a federated training strategy is adopted, as illustrated in Fig. 1. Multiple QGAN instances are deployed across distributed nodes, each independently training on a distinct subset of the dataset. Given the sensitivity of quantum machine learning models to initialization, each quantum generator operates with a unique seed. At the end of each training epoch, the generator and discriminator weights from all nodes are aggregated separately. The next training iteration begins with the averaged weights, ensuring synchronization and collaborative learning across the distributed network.

3.4 Noisy simulation

When analyzing quantum circuits, it is crucial to consider real-world imperfections such as qubit decoherence, gate errors, and environmental noise, which affect computational reliability. The proposed QGAN configurations were first tested under noiseless conditions and then evaluated using IBM's FakeBackend to compare performance degradation and convergence time. This backend simulates quantum devices based on system snapshots, including qubit connectivity and properties. Simulations were performed with the Qiskit-Aer package using the Sampler primitive and 1024 shots. While noisy simulations yield different results, starting with noise-free simulations provides a useful baseline before assessing the impact of noise.

4 Evaluation results and analysis

To evaluate the effectiveness of the proposed Federated QGAN model, a thorough hyperparameter analysis was conducted to optimize performance, stability, and efficiency. Experiments were performed over 80 epochs, ensuring convergence of the quantum generator and discriminator, with loss stabilization observed around this point. A total of six principal components are retained, as this setting provides an optimal balance between performance and computational efficiency. A ZZFeatureMap was employed to encode noise, enhancing the generator's ability to learn complex data distributions. Different quantum ansatzes were tested, with EfficientSU2 demonstrating slightly superior performance over RealAmplitudes due to its better encoding of correlations. Increasing ansatz repetitions improved performance but significantly raised computational costs, with three repetitions providing a balanced trade-off. Learning rate adjustments revealed that an initial rate of 0.01 caused instability, while refining it to 0.001 improved stability. Further tuning led to optimal rates of 0.001 and 0.0006 for the generator and discriminator, ensuring balanced adversarial training.

A federated approach was implemented using four QGANs trained on separate data subsets, with their insights aggregated. This setup improved generalization and robustness while accelerating training. As illustrated in Fig. 3, the generator's loss decreased over time, indicating improved sample synthesis, while the discriminator's loss stabilized, reflecting a balanced adversarial process. The final model achieved an accuracy of 0.9125 and an F1-score of 0.9034.

A key advantage of the federated approach was its efficiency: while a single QGAN required 24 hours to train, the federated model achieved similar performance in around 4 hours, thanks to parallelization. This demonstrates the feasibility of federated learning in quantum machine learning applications, reducing computational overhead while maintaining high predictive accuracy.

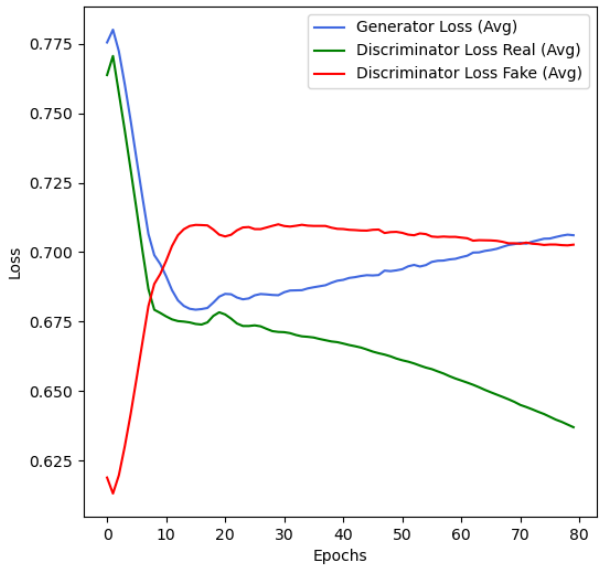

Fig. 3 Training process of the proposed federated QGAN

The best-performing configuration was further evaluated on the FakeNairobi backend, a noisy quantum simulator with seven qubits, to assess its robustness under realistic conditions. Initial results showed that the generator's loss increased without balancing with the discriminator's loss, indicating that the generator struggled against the more dominant discriminator. To address this, the discriminator's learning rate was reduced to 0.0001, achieving an optimal balance between the two models, as illustrated in Fig. 4.

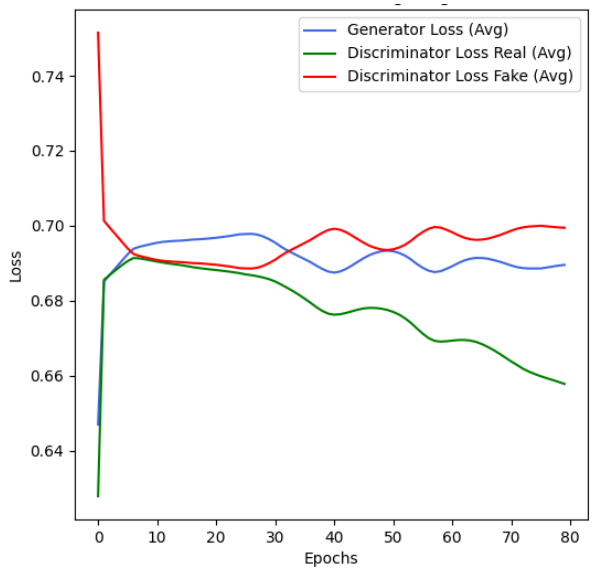

Fig. 4 Training process of the proposed federated QGAN

With this adjustment, the final model reached an accuracy of 0.8738 and an F1-score of 0.8797, lower than the noiseless version but a solid starting point for further improvements. These results highlight the importance of balancing the generator and discriminator's capabilities to achieve stable training and reasonable performance in noisy environments. The learning rate proves to be a key tunable parameter in this regard. Further experiments are necessary to refine the model, enhance its performance, and develop strategies to mitigate the effects of quantum noise.

5 Conclusion and Future Work

This study examined the effectiveness of a Federated QGAN model for intrusion detection, emphasizing performance, stability, and computational efficiency. The final model achieved predictive accuracy of 0.9125 and an F1-score of 0.9034 while significantly reducing training time through parallelization. To assess robustness in realistic conditions, the model was tested on the FakeNairobi noisy simulator, revealing challenges in generator-discriminator balance. Adjusting the discriminator's learning rate improved stability, leading to an accuracy of 0.8738 and an F1-score of 0.8797. These results highlight the importance of tuning hyperparameters, particularly the learning rate, to mitigate noise effects.

Future work will focus on refining the model to enhance performance under noise, exploring alternative quantum feature maps, ansatz structures, and aggregation strategies. Additionally, further experiments with different federated setups, data partitioning techniques, and noise-mitigation methods will help improve scalability and robustness. Integrating hybrid quantum-classical approaches or reinforcement learning could further optimize training dynamics and overall effectiveness.

6 Acknowledgements

This research was funded by the NGIsargasso project (Europe Horizon Grant No. 101092887), Open Call 4 FRQGAN4AD project. We acknowledge the use of IBM Quantum services for this work. The views expressed are those of the authors, and do not reflect the official policy or position of IBM or the IBM Quantum team.

7 References

[1] O. H. Abdulganiyu, T. Ait Tchakoucht, and Y. K. Saheed, "A systematic literature review for network intrusion detection system (ids)," International journal of information security, vol. 22, no. 5, pp. 1125–1162,2023.

[2] S. H. Rafique, A. Abdallah, N. S. Musa, and T. Murugan, "Machine learning and deep learning techniques for internet of things network anomaly detection—current research trends," Sensors, vol. 24, no. 6, p.1968, 2024.

[3] S. Muneer, U. Farooq, A. Athar, M. Ahsan Raza, T. M. Ghazal, and S. Sakib, "A critical review of artificial intelligence based approaches in intrusion detection: A comprehensive analysis," Journal of Engineering, vol. 2024, no. 1, p. 3909173, 2024.

[4] J. Gui, Z. Sun, Y. Wen, D. Tao, and J. Ye, "A review on generative adversarial networks: Algorithms, theory, and applications," IEEE transactions on knowledge and data engineering, vol. 35, no. 4, pp. 3313–3332, 2021.

[5] H. Yamasaki, N. Isogai, and M. Murao, “Advantage of quantum machine learning from general computational advantages,” 2023.
[6] M. Cerezo, G. Verdon, H.-Y. Huang, L. Cincio, and P. J. Coles, “Challenges and opportunities in quantum machine learning,” Nature Computational Science, vol. 2, no. 9, pp. 567–576, 2022.
[7] C. Zoufal, A. Lucchi, and S. Woerner, “Quantum generative adversarial networks for learning and loading random distributions,” npj Quantum Information, vol. 5, no. 1, p. 103, 2019.
[8] M. H. Shahriar, N. I. Haque, M. A. Rahman, and M. Alonso, “G-ids: Generative adversarial networks assisted intrusion detection system,” in 2020 IEEE 44th Annual Computers, Software, and Applications Conference (COMPSAC), 2020, pp. 376–385.
[9] J. Lee and K. Park, “Gan-based imbalanced data intrusion detection system,” Personal and Ubiquitous Computing, vol. 25, no. 1, pp. 121– 128, 2021.
[10] S. Lloyd and C. Weedbrook, “Quantum generative adversarial learning,”
Physical review letters, vol. 121, no. 4, p. 040502, 2018.
[11] P.-L. Dallaire-Demers and N. Killoran, “Quantum generative adversarial networks,” Physical Review A, vol. 98, no. 1, p. 012324, 2018.
[12] E. Bermot, C. Zoufal, M. Grossi, J. Schuhmacher, F. Tacchino, S. Vallecorsa, and I. Tavernelli, “Quantum Generative Adversarial Networks For Anomaly Detection In High Energy Physics,” in 2023 IEEE International Conference on Quantum Computing and Engineering (QCE), vol. 01, Sep. 2023, pp. 331–341.
[13] B. Kalfon, S. Cherkaoui, J. Laprade, O. Ahmad, and S. Wang, “Successive data injection in conditional quantum GAN applied to time series anomaly detection,” IET Quantum Communication, vol. 5, no. 3, pp. 269–281, Sep. 2024.
[14] M. A. Rahman, H. Shahriar, V. Clincy, M. F. Hossain, and M. Rahman, “A Quantum Generative Adversarial Network-based Intrusion Detection System,” in 2023 IEEE 47th Annual Computers, Software, and Applications Conference (COMPSAC). Torino, Italy: IEEE, Jun. 2023, pp. 1810–1815.
[15] D. Herr, B. Obert, and M. Rosenkranz, “Anomaly detection with variational quantum generative adversarial networks,” Quantum Science and Technology, vol. 6, no. 4, p. 045004, Oct. 2021.
[16] M. Rasouli, T. Sun, and R. Rajagopal, “Fedgan: Federated generative adversarial networks for distributed data,” arXiv preprint arXiv:2006.07228, 2020.
[17] R. B. G´omez, C. O’Meara, G. Cortiana, C. B. Mendl, and J. Bernab´e Moreno, “Towards autoqml: A cloud-based automated circuit architecture search framework,” in 2022 IEEE 19th International Conference on Software Architecture Companion (ICSA-C). IEEE, 2022, pp. 129–136.
[18] T. Jamaluddin, B. Narottama, and S. Y. Shin, “Unsupervised federated quantum gan for optimizing wireless communications,” Authorea Preprints, 2023.
[19] P. K. Sarkar, H.-H. Nguyen, and D. M. Farid, “Flgan-ids: Intrusion detection using gans with federated learning,” in International Conference on Intelligent Systems and Data Science. Springer, 2024, pp. 216–230.