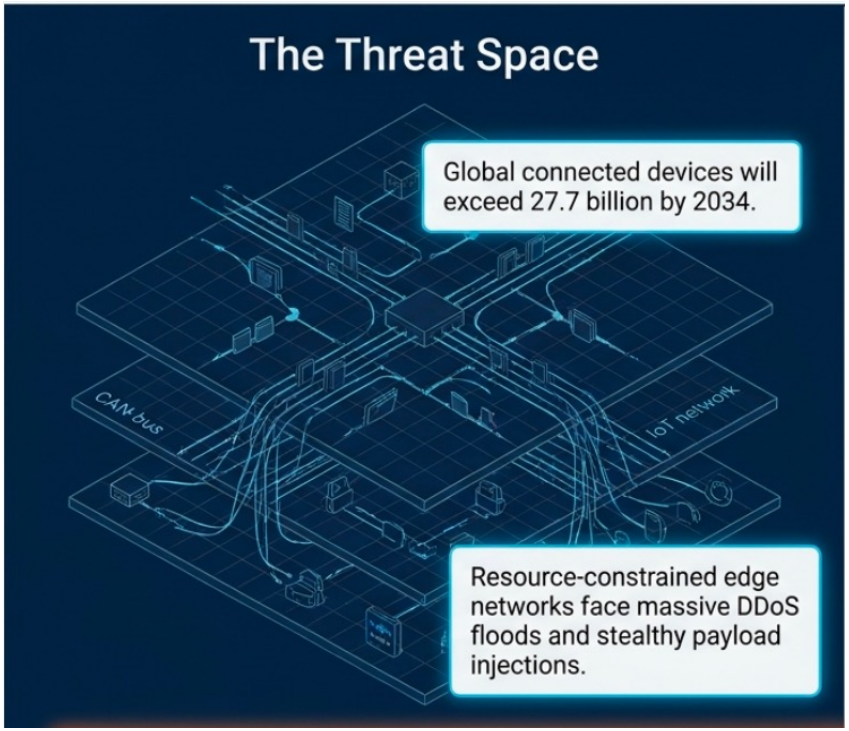
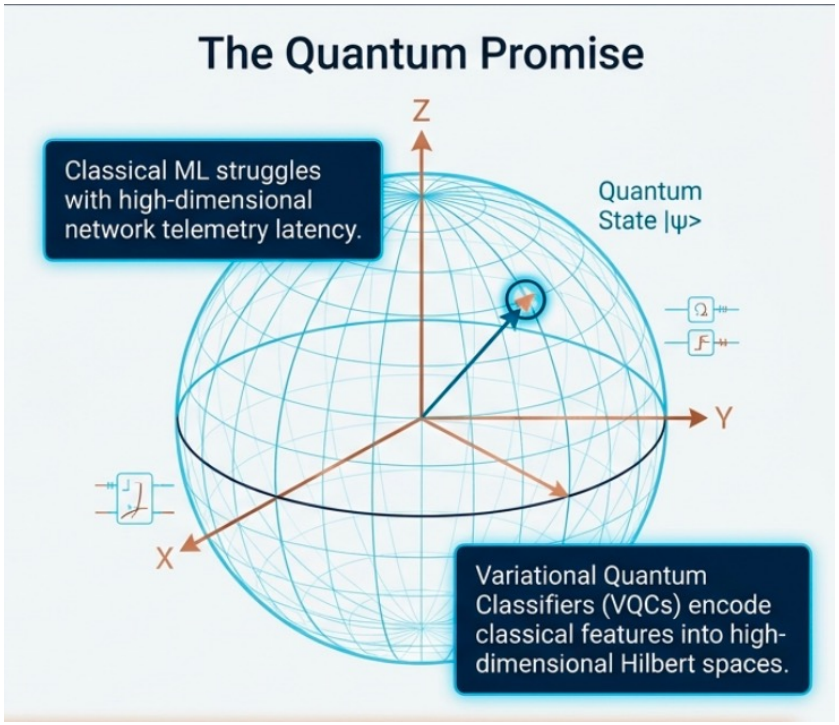
Mitigating Cyber Threat: A class-Imbalance Aware Quantum Machine Learning Framework

Sharifah Hafizah Syed Ariffin, Phd
Faculty of Electrical Engineering
Universiti Teknologi Malaysia

Introduction





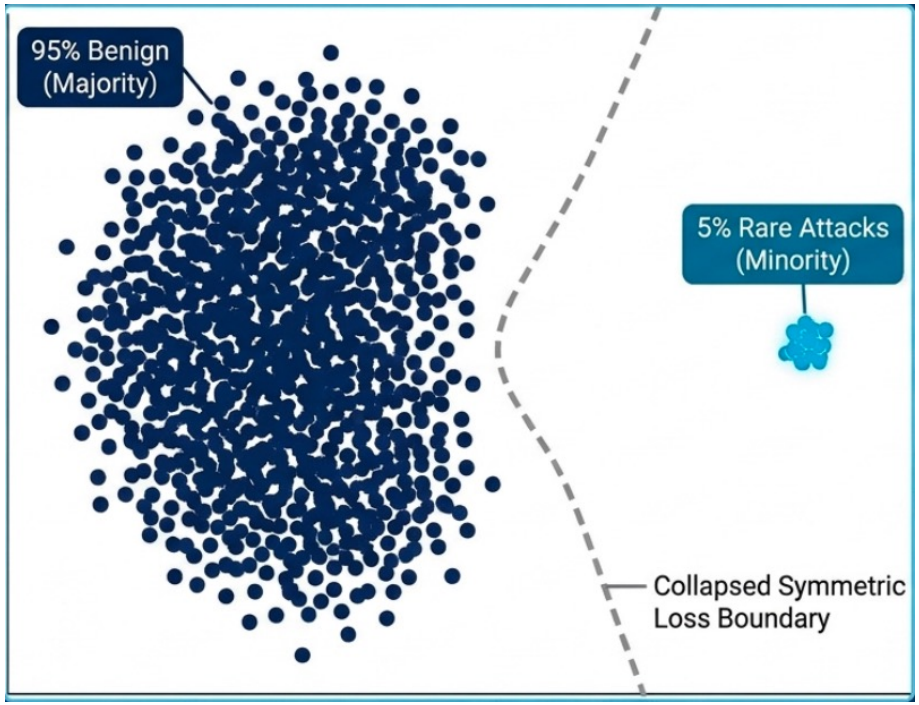
Vulnerability: real-world operational traffic is is inherently , severely imbalanced. Destructive attacks comprise only a microscopic fraction of overall traffic

Background

The Current situation	The Necessary Shift
Relies exclusively on symmetric loss function (Cross-entropy, MSE)	Imbalance need to be integrated directly into the training objective
Imbalance is addressed entirely upstream via dataset sampling (Under sampling , SMOTE)	Reshaping the loss landscape forces the quantum optimizer to be aware of rare events
Minority F1 scores for critical attacks, decrease 60%-70%	A robust decision boundary that able to resist majority-class gravitational pull.

In operational IoT and IoV networks, traffic distributions are imbalanced. Benign background telemetry or volumetric floods dominate total flow counts, while rare, highly destructive attacks comprise a small fraction of overall traffic.

Background



Symmetric loss function (MSE, Cross-entropy) treat every sample equally

The overwhelming number of easy, majority-class samples mathematically dominates the gradient updates.

The VQC model underfits the minority class, collapsing the decision boundary towards the majority and the network not able to spot the rare anomalies

Proposed Method: Margin-based Class Wise Focal Loss for Bipolar readouts

Constraint

VQCs output expectation values bounded in [+1,+1] not classsical probabilities

Mapping

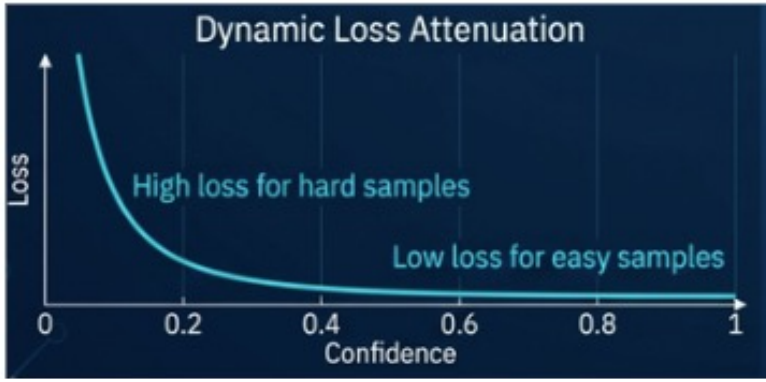
Bipolar Probabilities

$$p_t = \frac{1}{1 + e^{-\hat{y}y}}$$

Formulation

Margin-based Focal Loss

$$L = -\alpha (1 - p_t)^{\gamma} \log(p_t)$$

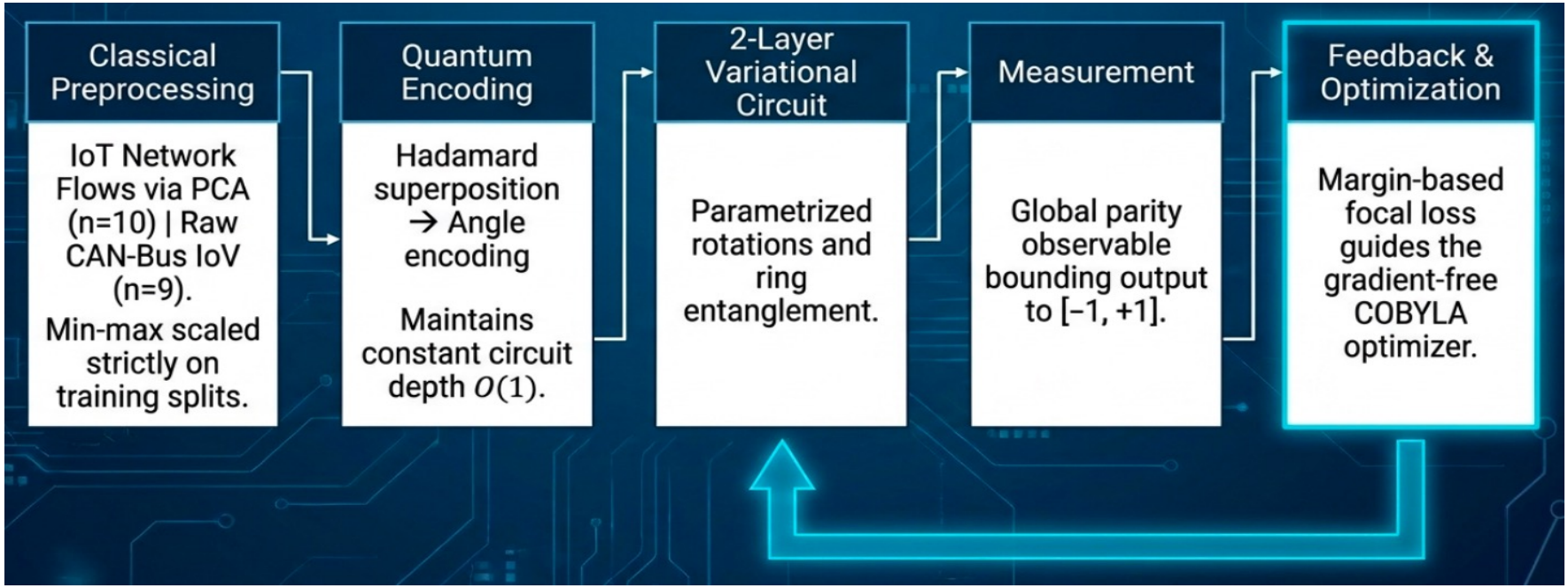


Independent Parameters:

- Weighting: Minority = 0.75, Majority = 0.25
- Focusing: Exponent = 2.0

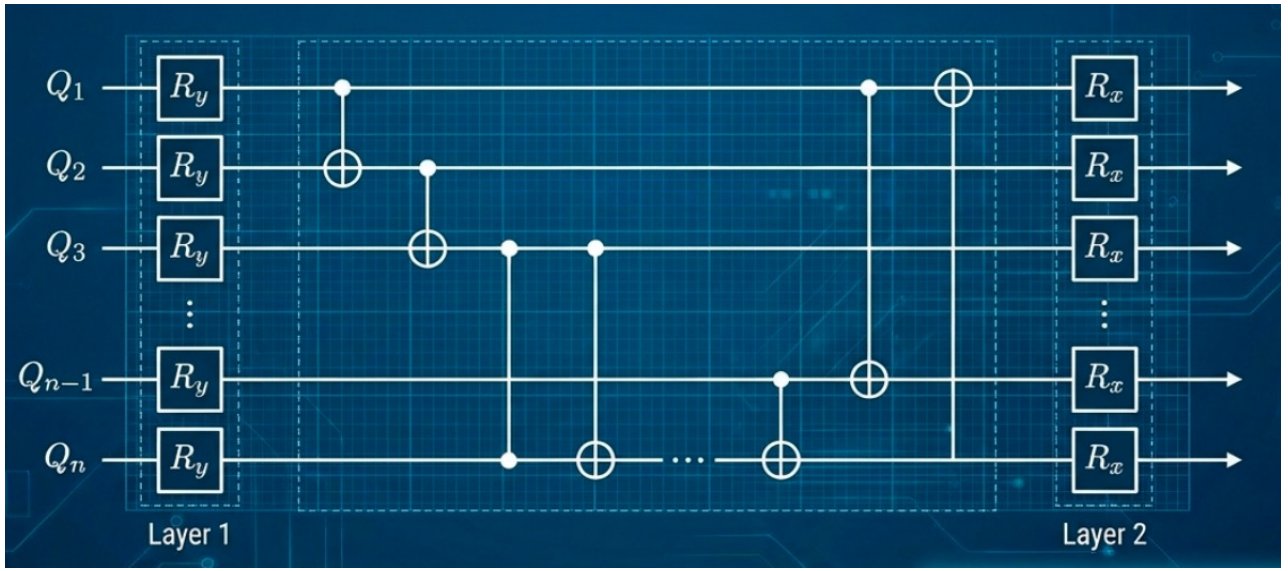
This formulation explicitly protects the rare attack vectors from gradient starvation.

Imbalance-aware quantum classification pipeline



The proposed scheme: PCA-based feature reduction in the classical preprocessing, angle encoding in the quantum encoding, a two-layer variational quantum classifier, and a sign-threshold decision rule. The pipeline is trained end-to-end with a margin-based focal loss objective.

Training Procedure



Topology

A compact two-layer variational circuit utilizing a ring entangled CNOT architecture.

Ultra-lightweight

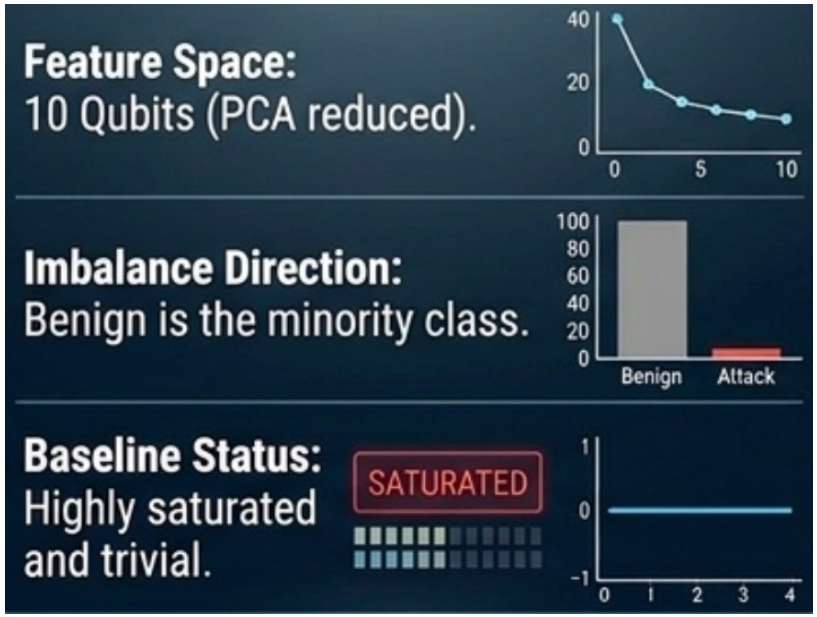
Only 2n trainable parameters (20 parameters for 10-qubit IoT, 18 parameters for 9-qubit IoV

Barren Plateau Resistant

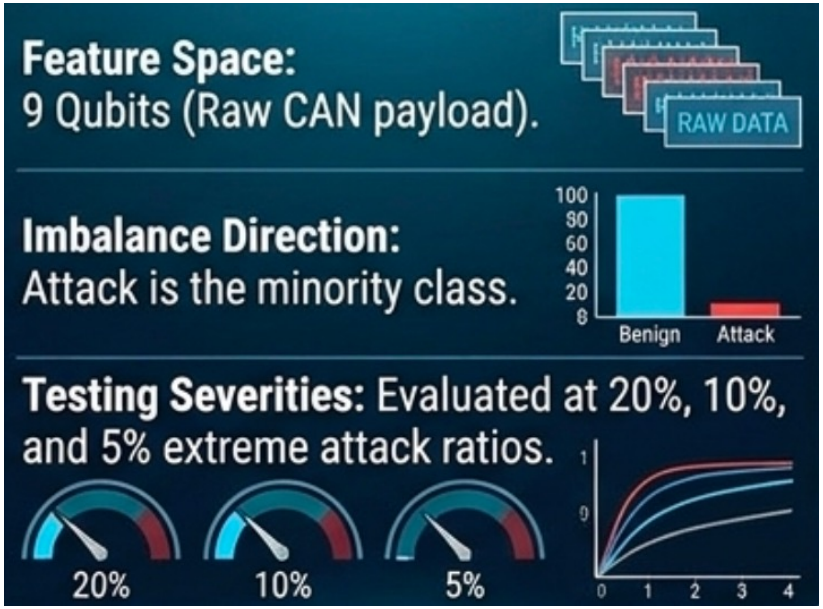
Shallow circuit depth deliberately constraint the landscape, avoiding explonential flat gradient regions

Experimental Dataset

CICIOT2023 - Network Flow



CICIOV2024 – CAN-Bus



All models trained over multiple random seed to rigorously test gradience-free optimize variance

CICIOT2023: Training Objective Comparison

Objective	F1	Bal. Acc.	MCC	AUC-PR
MSE	0.9913 ± 0.0007	0.9913 ± 0.0008	0.9587 ± 0.0033	0.9982 ± 0.0001
Cross-entropy	0.9915 ± 0.0006	0.9912 ± 0.0009	0.9596 ± 0.0029	0.9982 ± 0.0001
Focal (ours)	0.9895 ± 0.0006	0.9871 ± 0.0048	0.9502 ± 0.0037	0.9980 ± 0.0001

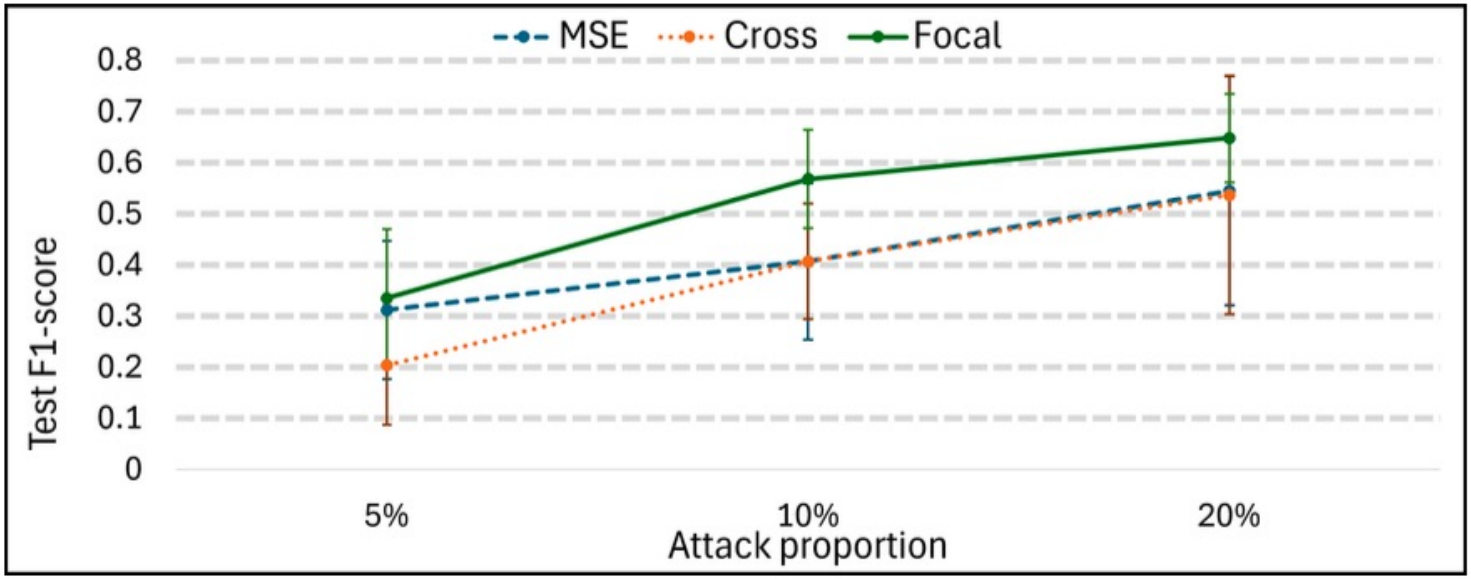
Using a 10-qubit variational quantum circuit, 10 PCA components, and a test partition of 2,000 records. All training objectives achieving an F1-score above 0.99 and a balanced accuracy above 0.988. Differences among the objectives are smaller than the seed-to-seed standard deviation for each objective.

CICIOV2024: Training Objective Comparison

Objective	F1	Bal. Acc.	MCC	AUC-PR
MSE	0.545 ± 0.224	0.712 ± 0.104	0.498 ± 0.209	0.543 ± 0.208
Cross-entropy	0.537 ± 0.233	0.722 ± 0.128	0.475 ± 0.225	0.533 ± 0.260
Focal (ours)	$\mathbf{0.648 \pm 0.086}$	$\mathbf{0.794 \pm 0.041}$	$\mathbf{0.563 \pm 0.118}$	$\mathbf{0.667 \pm 0.154}$

Using 9 qubits with raw CAN features, the proposed class-wise focal loss achieves the highest F1-score (0.648 ± 0.086), balanced accuracy, Matthews Correlation Coefficient (MCC), and area under the precision–recall curve (AUC-PR).

CICIOV2024: Test F1 Vs Attack Proportion



Focal loss consistently achieves highest F1-score as well as the best performance on imbalance robust metric.

Structured Imbalance Awareness

	Standard method		Proposed method
✗	Relies on dataset manipulation upstream	✓	Imbalance natively integrated into the objective via bipolar mapping
✗	Highly vulnerable to decision boundary collapse in real traffic	✓	Protects minority class boundaries from majority class gravity
✗	Volatile and unpredictable gradient-free VQC training	✓	Actively stabilizes the COBYLA optimizer fore reproducible training

Conclusion

Multiclass cyber treat extension: Extending the class-wise focal loss formulation to 10-class attack suites to support fine grained incident response.

Current research collaboration between Universiti Teknologi Malaysia and Bahrain Polytechnic

ASEAN smart Mobility testbed integration: deploying the trained QML-IDS model on edge gateways within connected autonomous vehicle (CAV) testbeds managed jointly by UTM and ASEAN industry/agencies/university partners

shafizah@utm.my

THANK YOU