

QML-IDS em Open RAN sob Restrições NISQ

Diego Abreu¹, Murilo Silva¹, Fiterlinge Sousa¹, Fernando Farias¹

¹Rede Nacional de Ensino e Pesquisa – RNP

{diego.abreu,murilo.silva,firterlinge.sousa,fernando.farias}@rnp.br

Abstract. *The disaggregation of Open RAN increases the attack surface of 5G/6G networks, creating the need for efficient intrusion detection mechanisms. This work proposes and evaluates an intrusion detection pipeline based on classical and quantum machine learning using the public Netslab-5G-ORAN-IDD dataset, which contains 1.7 million records distributed across six traffic classes. Five classical classifiers and three quantum approaches are compared: QKSVM, VQC, and QIK-SVM. QKSVM achieved $93.9\% \pm 1.2\%$ binary accuracy and outperformed SVM-RBF in all 10 evaluated train/test pairs, with an average gain of 0.85 ± 0.56 pp. Under heuristic NISQ noise scenarios inspired by IBM Quantum backends, binary accuracy degradation ranged from 1.6 to 3.3 pp, whereas VQC failed to converge and performed below the majority-class baseline. The main contribution of this work is a repeated comparative evaluation of quantum kernel-based methods for intrusion detection in O-RAN, including their sensitivity to NISQ noise and the limitations associated with their deployment in the near-RT RIC.*

Resumo. *A desagregação do Open RAN amplia a superfície de ataque das redes 5G/6G, tornando necessária a adoção de mecanismos eficientes de detecção de intrusão. Este trabalho propõe e avalia um pipeline de detecção baseado em aprendizado de máquina clássico e quântico, utilizando o conjunto público Netslab-5G-ORAN-IDD, com 1,7 milhão de registros e seis classes de tráfego. São comparados cinco classificadores clássicos e três abordagens quânticas: QKSVM, VQC e QIK-SVM. O QKSVM alcançou $93,9\% \pm 1,2\%$ de acurácia binária e superou o SVM-RBF nos 10 pares treino/teste avaliados, com ganho médio de $0,85 \pm 0,56$ pp. Em cenários heurísticos de ruído NISQ inspirados em backends IBM Quantum, a degradação binária foi de 1,6 a 3,3 pp, enquanto o VQC não convergiu e ficou abaixo da baseline majoritária. A principal contribuição é uma avaliação comparativa e repetida de métodos baseados em quantum kernels para IDS em O-RAN, incluindo sua sensibilidade a ruído NISQ e as limitações para uso no near-RT RIC.*

1. Introdução

A arquitetura *Open RAN* (O-RAN) substitui soluções monolíticas por componentes desagregados, interfaces abertas e controle baseado em software [O-RAN Alliance 2021]. Essa flexibilidade favorece interoperabilidade e inovação, mas também amplia a superfície de ataque: interfaces como E2, O1 e A1, componentes virtualizados e *xApps* no *near-RT RIC* passam a constituir novos vetores de intrusão.

A detecção de intrusão nesse ambiente deve lidar com tráfego heterogêneo, ataques dinâmicos e ciclos de controle do *near-RT RIC* entre 10 ms e 1 s. Nesse contexto, *Quantum Machine Learning* (QML) é investigado como alternativa para explorar

representações de alta dimensão. *Quantum kernels*, como os construídos com *ZZFeature-Map*, mapeiam dados clássicos para espaços de Hilbert potencialmente úteis à separação de classes [Havlíček et al. 2019]. Entretanto, dispositivos atuais operam no regime NISQ e estão sujeitos a decoerência, erros de portas e de leitura [Preskill 2018], além do custo de preparação, execução e medição dos circuitos.

Este trabalho contribui com: (i) um *pipeline* de IDS aplicado ao *Netslab-5G-ORAN-IDD* [Samarakoon et al. 2022]; (ii) a comparação de cinco modelos clássicos com QKSVM, VQC e QIK-SVM; (iii) uma avaliação repetida, com média e desvio-padrão, do desempenho dos modelos; e (iv) um estudo de sensibilidade do QKSVM a quatro perturbações de ruído NISQ aplicadas heurísticamente ao kernel, acompanhado de uma discussão sobre as limitações de implantação no *near-RT RIC*.

2. Trabalhos Relacionados

[Groen et al. 2024] apresentaram um dos primeiros conjuntos de KPIs O-RAN com tráfego malicioso real, incluindo DDoS e DoS sob diferentes padrões de mobilidade, enquanto [Samarakoon et al. 2022] introduziram o 5G-NIDD com fluxos gerados em infraestrutura 5G real. Em QML, [Liu et al. 2021] mostraram que *quantum kernels* podem ser competitivos em dados com estruturas específicas, e [Cerezo et al. 2021] discutiram limitações de VQCs, incluindo *barren plateaus*. Diferentemente desses estudos, avaliamos QML diretamente em um problema de IDS associado a O-RAN, com comparação pareada entre QKSVM e SVM-RBF, análise do VQC e estudo de sensibilidade a múltiplos canais de ruído.

3. Dataset e Metodologia

3.1. Dataset e pré-processamento

O *Netslab-5G-ORAN-IDD* contém 1.723.817 registros e 26 atributos derivados de fluxos TCP/UDP/ICMP. A Tabela 1 resume as seis classes.

Tabela 1. Distribuição das classes no *Netslab-5G-ORAN-IDD*.

Categoria	Registros	%	Tipos de ataque
DoS	632.507	36,7%	SYN flood, HTTP flood
DDoS	420.282	24,4%	UDP, TCP ACK, ICMP, Slowloris
Web	288.304	16,7%	XSS, SQL injection, dir. bruteforce
Probe	183.293	10,6%	TCP/UDP portscan, OS fingerprinting
Benigno	170.865	9,9%	HTTP, SSH, DNS, NTP
Bruteforce	28.566	1,7%	SSH, FTP

Foram removidos `uid`, `src_ip`, `dst_ip` e `history`, além de `attack_type` e `traffic_type`, estes últimos para evitar vazamento em relação à variável-alvo `attack_category`. Variáveis de bytes receberam transformação $\log(1 + x)$ e foram criadas as características `byte_ratio`, `pkt_ratio`, `bytes_per_pkt_src` e `bytes_per_pkt_dst`. Variáveis categóricas foram convertidas por *label encoding*.

Um *Random Forest* auxiliar (`n_estimators=100`, `max_depth=10`, semente 42) selecionou 15 características para os modelos clássicos. Para os experimentos quânticos, uma segunda seleção reteve seis atributos, um por qubit: `dst_port`, `conn_state_enc`,

`http_status_error`, `src_port`, `duration` e `log_dst_bytes`. Esses atributos foram reescalados para $[0, \pi]$. Os modelos clássicos utilizaram 120.000 amostras balanceadas (20.000 por classe), enquanto os experimentos QML utilizaram até 1.500 amostras (250 por classe), devido ao custo do kernel quântico, que cresce como $\mathcal{O}(N^2 2^n)$ na simulação empregada.

3.2. Modelos Quânticos

ZZFeatureMap e QKSVM: O QKSVM utiliza um kernel quântico para medir similaridade entre amostras após sua codificação em estados quânticos. Neste trabalho, empregamos uma codificação com $n = 6$ qubits e $R = 2$ repetições. A operação de codificação é definida por:

$$U_\phi(x) = \prod_{r=1}^R \left[\prod_{q=1}^n (H_q \cdot R_z^{(q)}(2x_q)) \cdot \prod_{(q,k) \in S} \text{CNOT}_{q,k} \cdot R_z^{(k)}(2(\pi - x_q)(\pi - x_k)) \cdot \text{CNOT}_{q,k} \right] \quad (1)$$

A primeira parte do circuito aplica portas de Hadamard e rotações R_z dependentes dos atributos de entrada. Em seguida, as portas CNOT e rotações condicionadas introduzem correlações entre pares de qubits. Assim, o *ZZFeatureMap* não codifica apenas atributos individuais, mas também interações não lineares entre eles. Essa propriedade é relevante para IDS, pois muitos ataques são caracterizados por combinações de métricas, e não por um único atributo isolado. O kernel quântico mede a fidelidade entre os estados codificados:

$$K(x_a, x_b) = |\langle \phi(x_a) | \phi(x_b) \rangle|^2, \quad \mathbf{K} = |\Phi^\dagger \Phi|^2, \quad (2)$$

onde $\Phi \in \mathbb{C}^{N \times 2^n}$ contém os estados quânticos das N amostras. A matriz $\mathbf{K}$ é então usada por um SVM clássico, substituindo kernels tradicionais como RBF.

VQC: O *Variational Quantum Classifier* combina uma etapa de codificação com um circuito variacional treinável. Utilizamos o *ZZFeatureMap* com uma repetição, seguido de $L = 2$ camadas de rotações R_y e R_z e portas CNOT em anel. A saída do modelo é dada por $p(1|x) = (1 - \langle Z_0 \rangle)/2$, e os parâmetros são otimizados via COBYLA. Diferentemente do QKSVM, que depende de uma matriz de kernel, o VQC tenta aprender diretamente os parâmetros do circuito. Isso reduz a dependência de uma matriz quadrática em N , mas introduz desafios de otimização, especialmente em circuitos ruidosos ou com gradientes pequenos.

QIK-SVM: O QIK-SVM é uma abordagem *quantum-inspired*, isto é, inspirada em mecanismos de codificação quântica, mas executada inteiramente em hardware clássico. Utilizamos uma codificação inspirada em Fourier, $\varphi(u) = [\cos u, \sin u, \cos 2u, \sin 2u]$, seguida de um kernel RBF no espaço transformado. Essa abordagem tem custo de inferência $\mathcal{O}(1)$ por transformação de atributo, tornando-se atraente para cenários com restrição de latência, ainda que não reproduza completamente a geometria induzida pelo *ZZFeatureMap*.

3.3. Modelos de Ruído NISQ

Avaliamos quatro canais de ruído aplicados diretamente sobre a matriz de kernel. Essa escolha permite analisar como imperfeições típicas de dispositivos NISQ afetam a simila-

ridade entre os estados quânticos codificados e, por consequência, a fronteira de decisão aprendida pelo SVM. Em vez de simular novamente todo o circuito para cada condição de ruído, os efeitos são incorporados sobre os valores de kernel, o que torna a análise mais simples e computacionalmente viável.

Depolarizante: $K_{\text{dep}}(i, j) = \alpha^d \cdot K(i, j) + (1 - \alpha^d)/2^n$, com $\alpha = 1 - 4p/3$. Esse canal representa uma perda geral de informação quântica, aproximando o estado de uma distribuição mais uniforme. Na matriz de kernel, isso reduz o contraste entre pares de amostras semelhantes e diferentes, tornando a separação entre classes menos evidente.

Dephasing (T2): $K_{\text{deph}}(i, j) = e^{-p \cdot d} \cdot K(i, j)$, $i \neq j$ — destrói as coerências de fase entre qubits. Esse canal é particularmente relevante para o *ZZFeatureMap*, pois parte da separabilidade entre as classes é codificada em relações de fase. Assim, quando o *dephasing* aumenta, o kernel perde informação discriminativa importante, o que pode degradar diretamente o desempenho do QKSVM.

Amortecimento de amplitude (T1): $K_{\text{amp}}(i, j) = f^d \cdot K(i, j) + (1 - f^d)/2^n$, com $f = 1 - \gamma/2$. Esse ruído modela a tendência de relaxação dos qubits para o estado fundamental. Em comparação ao *dephasing*, seu impacto tende a ser mais brando neste estudo, pois ele não ataca diretamente as coerências de fase exploradas pelo mapa de características, embora ainda possa reduzir a expressividade da representação quântica.

Shot noise: $K_{\text{shot}}(i, j) \sim \mathcal{N}(K(i, j), K(i, j)(1 - K(i, j))/N_{\text{shots}})$. Esse ruído representa a variabilidade estatística causada por um número finito de medições. Em hardware quântico real, as probabilidades não são observadas diretamente, mas estimadas a partir de múltiplas execuções do circuito. Portanto, quanto menor o número de *shots*, maior a incerteza na estimativa dos elementos da matriz de kernel.

Os parâmetros dos modelos de ruído foram calibrados com dados públicos de três *backends* IBM Quantum: *ibm_brisbane* 2024, *ibm_kyoto* 2023 e *ibm_nairobi* 2022. Também incluímos um cenário *Noisy extreme*, com níveis de ruído mais elevados, para avaliar o comportamento do QKSVM em condições adversas.

4. Resultados

4.1. Modelos clássicos

A Tabela 2 apresenta os resultados com 120.000 amostras e divisão 80/20. O *Random Forest* obteve o melhor F1-macro (94,99%), enquanto a *Decision Tree* apresentou desempenho próximo (94,65%) com o menor tempo de treino e predição em lote. Esses tempos não representam latência por fluxo.

Tabela 2. Desempenho dos classificadores clássicos.

Modelo	Acc.	F1-macro	Acc. bin.	F1 bin.	Tempo
Decision Tree	94,58%	94,65%	96,67%	98,02%	0,4s
Random Forest	94,92%	94,99%	96,97%	98,21%	5,7s
Gradient Boosting	94,67%	94,73%	96,54%	97,95%	116s
Logistic Regression	80,68%	80,69%	93,85%	96,43%	2,1s
KNN ($k = 5$)	93,58%	93,66%	96,20%	97,74%	1,2s

4.2. QML versus modelos clássicos

O QKSVM e os modelos clássicos de comparação foram avaliados em 10 reamostragens completas (sementes 42–51), com 1.500 amostras e divisão 75/25; o VQC foi repetido cinco vezes devido ao maior custo de otimização. A tarefa binária é desbalanceada no conjunto de teste (83,3% de ataques), razão pela qual também reportamos F1-macro.

Tabela 3. Comparativo QML e clássico (média $\pm$ desvio-padrão).

Categoria	Modelo	Tarefa	Acc.	F1-macro
Quântico	QKSVM	Bin.	$93,9 \pm 1,2\%$	$87,1 \pm 3,1\%$
	VQC	Bin.	$67,1 \pm 4,4\%$	$57,4 \pm 7,1\%$
	QIK-SVM	Multi.	$84,5 \pm 1,5\%$	$84,4 \pm 1,6\%$
Clássico	SVM-RBF	Bin.	$93,0 \pm 1,0\%$	$84,9 \pm 2,7\%$
	RF	Multi.	$89,2 \pm 1,0\%$	$89,2 \pm 1,0\%$
	LR	Bin.	$92,4 \pm 1,3\%$	$84,5 \pm 3,0\%$
Baseline	Majoritária	Bin.	$83,3 \pm 0,1\%$	$45,4 \pm 0,0\%$
	Majoritária	Multi.	$16,5 \pm 0,0\%$	$4,7 \pm 0,0\%$

O principal resultado é a comparação pareada entre QKSVM e SVM-RBF: embora as médias sejam próximas, o QKSVM superou o SVM-RBF nas 10 de 10 repetições, com ganho de $0,85 \pm 0,56$ pp em acurácia. Em F1-macro, os valores foram 87,1% e 84,9%, respectivamente. O resultado indica uma vantagem pequena, porém consistente, do kernel quântico sobre o RBF no mesmo regime de dados; não implica superioridade sobre todos os modelos clássicos.

O VQC apresentou $67,1\% \pm 4,4\%$ de acurácia e não convergiu em nenhuma das cinco repetições, ficando abaixo da baseline majoritária. Esse comportamento reforça a dificuldade de otimização de circuitos variacionais [Cerezo et al. 2021]. Já o QIK-SVM alcançou $84,5\% \pm 1,5\%$ na tarefa multiclasse, abaixo do RF ($89,2\% \pm 1,0\%$), mas muito acima da baseline majoritária.

Também avaliamos o regime de poucas amostras. Com $N = 30$, o QKSVM obteve $93,2\% \pm 0,7\%$ contra $89,7\% \pm 4,3\%$ do SVM-RBF. A diferença cai para menos de 0,3 pp em $N = 80$, e o SVM-RBF supera ligeiramente o QKSVM em $N = 120$. Assim, os resultados sugerem maior estabilidade do QKSVM em amostras muito pequenas, mas não sustentam uma vantagem geral de acurácia nesse regime.

5. Impacto do Ruído NISQ

Para a análise de ruído, utilizamos uma baseline única sem perturbação: $93,08\% \pm 1,10\%$ na tarefa binária e $82,33\% \pm 4,48\%$ na multiclasse, com $R = 8$ reamostragens de 600 amostras (100 por classe, divisão 75/25). A Tabela 4 apresenta os cenários heurísticos inspirados em *backends* IBM.

Na tarefa binária, a degradação varia de 1,58 a 3,25 pp; na multiclasse, de 2,08 a 6,33 pp. Como os desvios-padrão são da mesma ordem de alguns desses deltas, o ordenamento entre cenários deve ser interpretado qualitativamente. Entre os canais avaliados, o *dephasing* (T2) foi o mais prejudicial: em $p = 0,05$, a acurácia binária caiu para $87,75\% \pm 1,53\%$, redução de 5,33 pp em relação à baseline. Esse efeito é coerente com o *ZZFeatureMap*, cuja representação depende de relações de fase entre qubits.

Tabela 4. Sensibilidade do QKSVM a cenários heurísticos de ruído NISQ.

Cenário	Acc. bin.	Δ bin.	Acc. multi.	Δ multi.
Sem ruído	93,08 $\pm$ 1,10%	—	82,33 $\pm$ 4,48%	—
ibm_brisbane (2024)	91,50 $\pm$ 1,96%	−1,58 pp	80,25 $\pm$ 3,89%	−2,08 pp
ibm_kyoto (2023)	90,92 $\pm$ 2,31%	−2,17 pp	78,75 $\pm$ 3,22%	−3,58 pp
ibm_nairobi (2022)	90,50 $\pm$ 1,72%	−2,58 pp	78,33 $\pm$ 3,48%	−4,00 pp
Noisy extreme	89,83 $\pm$ 1,14%	−3,25 pp	76,00 $\pm$ 3,56%	−6,33 pp

5.1. Limitações para o near-RT RIC

Os experimentos não medem latência ponta a ponta nem integração como *xApp*; portanto, não estabelecem a viabilidade do QKSVM no *near-RT RIC*. Em hardware real, devem ser considerados preparação de estados, número de *shots*, filas, comunicação e leitura. Além disso, o *Netslab-5G-ORAN-IDD* contém fluxos TCP/UDP/ICMP agregados por conexão, e não KPIs nativos de E2, O1 ou A1. Assim, os resultados demonstram capacidade de classificação e sensibilidade a ruído, mas não desempenho operacional em uma implantação O-RAN em tempo quase real.

6. Conclusão

Este trabalho apresentou e avaliou um *pipeline* de IDS para O-RAN combinando aprendizado de máquina clássico, QML e análise de sensibilidade a ruído NISQ. Entre os modelos clássicos, o *Random Forest* obteve o melhor F1-macro (94,99%). Na comparação pareada em amostras reduzidas, o QKSVM alcançou 93,9% $\pm$ 1,2% de acurácia e superou o SVM-RBF nas 10 repetições, com ganho médio de 0,85 $\pm$ 0,56 pp. O VQC, por outro lado, apresentou baixa acurácia e falha de convergência.

Sob as perturbações heurísticas NISQ, o QKSVM apresentou queda de 1,58 a 3,25 pp na tarefa binária, e o *dephasing* foi o canal isolado mais crítico. Esses resultados fornecem evidência de que *quantum kernels* podem ser competitivos com kernels clássicos nesse cenário, mas não demonstram vantagem quântica geral nem viabilidade direta no *near-RT RIC*. Como trabalhos futuros, destacam-se a execução em hardware quântico real, o uso de KPIs nativos de interfaces O-RAN e a medição de latência ponta a ponta em uma implementação de *xApp*.

Agradecimentos

Este trabalho contou com o apoio do Ministério da Ciência, Tecnologia e Inovação (MCTI), por meio dos recursos da Lei nº 8.248, de 23 de outubro de 1991, conforme orientação da Secretaria de Empreendedorismo e Inovação do MCTI, no âmbito do Programa Open-RAN@Brasil.

Artefatos e Reprodutibilidade

Os códigos, scripts, configurações experimentais e demais artefatos utilizados neste trabalho estão disponíveis publicamente no repositório: https://github.com/DiegoAbreuSWB/qml_ids_wqusec_extended. O repositório reúne os recursos necessários para apoiar a reprodução dos experimentos e a verificação dos resultados apresentados neste artigo.

Declaração de Uso de Inteligência Artificial

Ferramentas de inteligência artificial generativa foram utilizadas como apoio durante o desenvolvimento deste trabalho. O Claude Code foi empregado como ferramenta auxiliar no desenvolvimento e revisão de código. O ChatRNP e o ChatGPT foram utilizados principalmente para revisão textual, clareza, gramática e aprimoramento da escrita. Todas as sugestões geradas por essas ferramentas foram revisadas criticamente pelos autores. Os autores são integralmente responsáveis pelo conteúdo científico, metodologia, resultados, interpretações e conclusões apresentados neste trabalho.

Referências

- Cerezo, M., Arrasmith, A., Babbush, R., et al. (2021). Variational quantum algorithms. *Nature Reviews Physics*, 3:625–644.
- Groen, P. et al. (2024). Open RAN dataset for intrusion detection with mobility patterns. *Computer Networks*, 242:110567.
- Havlíček, V., Córcoles, A. D., Temme, K., Harrow, A. W., Kandala, A., Chow, J. M., and Gambetta, J. M. (2019). Supervised learning with quantum-enhanced feature spaces. *Nature*, 567:209–212.
- Liu, Y., Arunachalam, S., and Temme, K. (2021). A rigorous and robust quantum speed-up in supervised machine learning. *Nature Physics*, 17:1013–1017.
- O-RAN Alliance (2021). O-RAN Architecture Description v07.00. Technical report, O-RAN Alliance.
- Preskill, J. (2018). Quantum computing in the NISQ era and beyond. *Quantum*, 2:79.
- Samarakoon, S. et al. (2022). 5G-NIDD: A comprehensive network intrusion detection dataset generated over 5G wireless network.