面向物联网网络入侵检测的量子机器学习落地

Domenico Ciuonzo、Idio Guarino、Alfredo Nascita、Antonio Pescap`e、Vincenzo Spadar i

arXiv:2611.32314v1 [cs.NI] 24 Nov 2026

摘要——**物联网（IoT）的迅速发展增加了网络系统遭受网络威胁的风险，使网络入侵检测系统（NIDSs）成为保护异构环境的关键。深度学习（DL）在流量分析方面表现出色，而量子机器学习（QML）近期也作为一种学习复杂流量模式的替代方法受到关注。我们评估了量子机器学习在异构物联网场景中用于入侵检测的实际适用性。我们提出一种量子—经典混合网络入侵检测系统，并在三个近期的物联网安全数据集上对其进行评估，**评估条件贴近实际运行环境。**除检测准确率外，我们还从**五个运行层面**将量子机器学习与经典深度学习基线进行比较：(*i*) 数据稀缺情况下的性能，(*ii*) 跨域评估中的迁移能力，(*iii*) 时间约束韧性，(*iv*) 在多个含噪后端下对量子非理想性的鲁棒性，以及 (*v*) 通过校准分析评估预测可信度。二分类和多分类滥用检测实验表明，在理想和含噪条件下，深度学习通常优于量子机器学习。不过，在数据稀缺和领域偏移场景中，量子机器学习仍具竞争力，并且在时间约束下能够超越深度学习基线。**

索引词—**深度学习、网络入侵检测系统、量子机器学习、可信人工智能**。

引言

物联网（IoT）的广泛采用推动了远程监控、智能基础设施和安全关键应用等各种服务的发展 [1]。包括可穿戴设备和智能家居系统在内的互联设备数量不断增加，显著扩大了网络流量的规模和异质性。[1] 然而，物联网环境高度分布式且具有异质性的特点，也扩大了恶意行为者可利用的攻击面。针对物联网系统的网络攻击可能造成严重后果，包括隐私泄露、经济损失、服务中断以及对人身安全的威胁 [2]。因此，大量研究致力于开发能够及时、可靠地检测网络威胁的解决方案。其中，网络入侵检测系统（NIDSs）旨在区分合法网络活动与恶意流量，并识别对手实施的具体攻击类型。然而，传统方法往往难以应对现代网络流量日益增加的复杂性、规模和加密程度。机器学习（ML），以及近年来的深度学习（DL）技术，已成为现代网络入侵检测系统的核心，尤其是在物联网场景中，互联设备数量的激增大幅扩大了攻击面。量子计算（QC）代表着信息科学中一个充满前景的前沿领域。面对物联网流量数据量的持续增长，量子计算有望成为一种强有力的应对手段，能够在大数据处理方面提供理论加速，并获得更准确的计算结果。量子计算范式与机器学习的交叉融合催生了量子机器学习（QML）。量子机器学习将量子力学原理与人工智能驱动的方法相结合 [3, 4]，为捕捉复杂数据中的非线性现象提供了新途径，也为物联网环境中的普适安全带来了潜在解决方案。尽管该领域的关注度不断上升，实用量子优势仍是一个悬而未决的问题 [4]，实证性能提升大多局限于小规模的合成数据集或过时数据集评估。固有的量子硬件噪声会影响电路测量，进而影响量子机器学习输出 [5, 6]。基于量子机器学习的网络入侵检测系统还面临训练数据稀缺、领域可迁移性挑战和时序约束。最后，与分类决策相关的置信度分数是否能够准确反映实际预测性能，仍不明确 [7]。

在本研究中，我们探讨早期网络入侵检测面临的以下运营挑战（*OCs*）：

- **数据效率（OC1）：** 在数据稀缺的情况下，网络入侵检测系统如何执行二分类和多分类误用检测？
- **跨域泛化（OC2）：** 网络入侵检测系统能否迁移到不同于其原始领域的其他领域？
- **时间约束韧性（OC3）：** 在严格时间约束下，网络入侵检测系统的检测鲁棒性如何？
- **对噪声和采样误差的鲁棒性（OC4）：** 物理硬件噪声和有限测量次数会如何影响基于量子机器学习的网络入侵检测系统的性能？
- **可信度评估（OC5）：** 网络入侵检测系统的预测有多可靠，校准效果如何？

为应对这些挑战，我们首先讨论用于网络入侵检测的量子机器学习，并说明本研究相对于既有工作的贡献。然后，我们介绍所提出的量子-经典混合网络入侵检测系统和实验方法。最后，我们通过五项开放性挑战评估该框架，以确定量子机器学习在物联网网络入侵检测中的实际应用机遇与局限。所有结果的复现所需数据和代码可在 https://github.com/viciEnzoh/qml-nids 获取。

量子增强网络入侵检测的兴起

在物联网环境中，检测网络流量中的攻击至关重要，因为恶意活动可能导致数据泄漏、服务中断和经济损失。传统机器学习和深度学习方法在解决

[1] https://www.ericsson.com/en/reports-and-papers/mobility-report/reports/november-2025

网络安全任务方面已证明非常有效 [2]。然而，由于资源限制严峻，且网络流量的规模和异构性不断增加，这些方法在真实的物联网环境中仍难以部署。这些局限促使研究者探索基于量子计算的替代方案，而量子机器学习已成为网络安全领域一个颇具前景的方向 [8]。近期研究已探索将量子机器学习用于攻击检测与分类 [4, 9, 10], ，但仍有若干运营挑战尚未得到充分解决。首先，现有方法大多依赖从完整的流量聚合（例如流或双向流）中提取的统计（事后分析）特征。因此，攻击往往要等到通信结束后才能被检测到，限制了网络入侵检测系统及时响应并减轻潜在损害的能力。尽管训练数据有限是网络安全领域的常见情况，尤其是面对罕见攻击或零日攻击时，量子机器学习模型在有限训练数据下的表现仍鲜有研究。量子机器学习模型的泛化能力也尚未得到充分研究。弥补这一研究空白是实现“一套网络入侵检测系统检测所有攻击”这一范式的关键 [11, 12],；在该范式下，单一框架无需重新训练即可适用于异构网络环境。另一个关键挑战是时间约束韧性。在实际运行环境中，降低检测延迟意味着只能观察受时间限制的流量，迫使网络入侵检测系统在攻击特征完全显现之前作出判断。经典机器学习/深度学习模型已对响应速度与检测准确率之间的权衡进行了研究，但量子机器学习领域对此关注有限。除了这些运营挑战之外，当前量子硬件的局限仍制约着量子机器学习在真实网络安全环境中的部署。实际量子计算机的可用性和规模有限，限制了其部署；与此同时，当前的含噪中等规模量子（NISQ）设备也存在噪声和量子比特退相干问题，可能导致测量结果失真并降低网络入侵检测系统的有效性。近期利用模拟和真实量子硬件开展的研究 [5, 6], 对硬件引发的噪声如何影响模型可靠性仅提供了有限见解。因此，量子硬件存在噪声时，基于量子机器学习的网络入侵检测系统能否保持可信度，仍是一项尚待解决的挑战。

研究定位。 本工作扩展了我们的初步研究 [4],，该研究考察了在两个物联网数据集上、固定数据包可见性条件下，采用量子机器学习的网络入侵检测系统的早期检测。在此基础上，我们将评估范围扩大到涵盖医疗、工业和智能家居领域的三个 物联网数据集，并在不同数据稀缺性条件下开展研究。我们进一步考察此前被忽视的关键运行层面。首先，我们通过最大观测窗口评估有限流量观测时间对网络入侵检测系统性能的影响。随后，我们在多个存在噪声的量子后端上评估量子硬件噪声和有限次采样 如何影响检测性能。最后，我们通过校准分析评估网络入侵检测系统的可信度 ，以确定预测置信度是否与分类准确率一致，从而验证系统的运行可靠性。

构建量子网络入侵检测系统

以下介绍所提出的用于早期网络入侵检测系统的经典-量子混合处理流程，重点关注该架构的量子组件（图1中的*PathB*）。

经典/量子混合网络入侵检测系统处理流程。网络流量依据标准五元组划分为双向流。网络入侵检测系统的输入来自从每个双向流的首批数据包中提取的报头字段，即不检查有效载荷。网络入侵检测系统执行早期检测，从而能够在攻击完全形成之前识别恶意活动。为优化响应时间，只要有足够的资源（即量子比特数量），就可以在经典层和量子层并行处理双向流。双向流用于输入基于深度学习/量子机器学习的模型。模型处理流程的结构如下：(*i*) 经典骨干网络，用于压缩数据以满足量子输入约束；(*ii*) 量子电路，用于处理压缩后的数据；以及(*iii*) 最终的分类层，用于提供输出标签。该框架可通过以下方式扩展：(*i*) 采用针对特征映射要求定制的骨干网络压缩，以处理更高维的特征；(*ii*) 扩展分类层，以适应更广泛的攻击范围，从而支持更大规模的部署。

通过量子神经网络（QNNs）进行攻击检测与分类。量子计算依赖量子比特，通过叠加和纠缠等性质实现优化的信息编码。为此，量子机器学习通过在复杂输入处理中采用参数化量子电路，有效利用量子力学原理。在这一背景下，量子机器学习通过将经典数据投影到更高维的希尔伯特空间中，利用参数化电路对复杂的非线性进行建模 [3, 8]。量子神经网络是最常见的方法 [13, 14], ，这得益于其较低的架构复杂度和较少的参数量。如图 1 的 Quantum Circuit 模块所示，这些网络基于三个基本要素：(*i*) 特征映射、(*ii*) 拟设和 (*iii*) 测量过程。特征映射 将经典数据转换为量子表示（量子编码）。在本文中，我们考虑两种常用策略：(*i*) 振幅嵌入（AMP*e*），将归一化特征向量编码到量子态的概率振幅中，从而提供高度紧凑的数据表示；(*ii*) 角度嵌入（ANG*e*），将特征转换为旋转角度，每个字段需要一个量子比特。拟设是量子层中可训练的组件。它由连续的旋转门层组成，其角度代表权重，并在优化过程中进行调整。每一层都由单量子比特门和多量子比特门组成。量子处理的输出通过测量过程转换回经典表示。在本文中，电路生成一个大小为2^{N_q}的概率向量，其中 N_q 是量子比特的数量，每个元素表示测量得到特定计算基态的概率。这种联合表示同时捕获了单个量子比特的行为和由纠缠引起的相关性。值得注意的是，在噪声条件下，通过有限次测量进行经验概率估计（S 测量次数）影响的是整体运行时延迟，而不是电路的结构复杂度。最后，概率向量输入分类 阶段，其中 Softmax 生成类别概率，最可能的标签用于识别良性流量或特定攻击。训练遵循经典的基于梯度的优化，最小化预测的软输出与独热编码

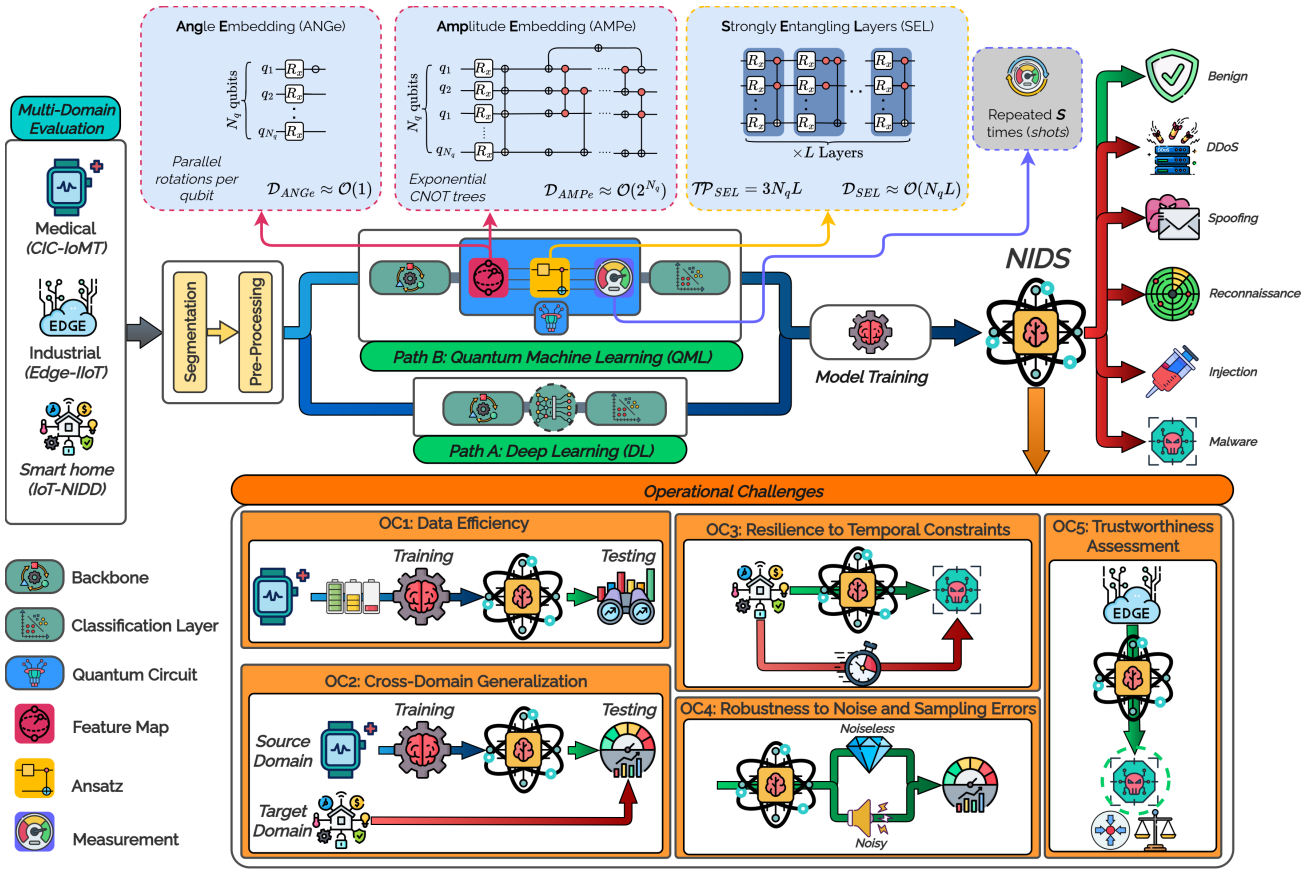


图 1. 面向多域物联网网络入侵检测系统的端到端混合流程。每个域由一个独立的数据集表示。经过流量分段和预处理后，流程分为两条路径：（A）深度学习和（B）量子机器学习；后者包含基于骨干网络的经典降维、量子变分电路和经典分类头。量子变分电路由以下部分构成：(*i*) 特征映射（AMP*e*、ANG*e*）、(*ii*) 拟设（SEL）和(*iii*) 测量过程，并以固定的测量次数（*S*）执行。该框架将网络流量映射为威胁情报，以区分良性流量与攻击流量（例如 DDoS、恶意软件），同时应对五项运营挑战（*OCs*）：（OC1）数据效率、（OC2）跨域泛化、（OC3）时间约束韧性、（OC4）对噪声和采样误差的鲁棒性，以及（OC5）可信度评估。电路复杂度以电路深度（D）和可训练参数数量（TP）来表征。

真实标签之间的损失。

实验设置

数据集。为确保全面评估，我们的分析涵盖三个不同的物联网领域：医疗、工业和智能家居。这使我们能够刻画各领域特有的威胁态势，并在实践中评估跨领域网络入侵检测系统的可移植性。因此，我们使用了三个物联网数据集，分别对应上述领域：CIC-IOMT2024（IOMT）、EDGE-IIOT（EDGE）和IOT-NIDD（NIDD）。这些数据集包含良性流量，以及分别涵盖约 604k、731k 和 201k 个双向流的 16 种、13 种和 9 种不同攻击类型，覆盖了广泛的网络威胁，例如（分布式）拒绝服务攻击、中间人攻击、注入攻击和恶意软件。对于每个双向流，我们从首批数据包中提取 4 个报头字段：分组方向（方向）、分组大小、TCP窗口大小和到达间隔时间。我们选取 10 个数据包，因为这能在早期性和检测能力之间实现适当的平衡。

所提出的模型。 我们比较经典深度学习架构与多种量子机器学习配置的性能。对于二分类和多分类滥用检测，我们采用 CNN+LSTM [15],作为稳健的基线，该模型由交替排列的卷积层和池化层组成，后接一个 LSTM层。量子机器学习模型采用 AMP*e* 和 ANG*e* 特征映射，并在其前添加一个初始全连接层（全连接层）以进行数据压缩。对于变分组件，我们采用强纠缠层拟设。此外，由于评估结果表明 AMP*e*是最有效的特征映射，我们提出了 AMP*e* (CNN+LSTM)，这是一个扩展模型，在 AMP*e* 阶段之前利用 CNN+LSTM骨干网络，整合基于深度学习的数据压缩。在所有混合模型中，骨干网络均作为独立的降维阶段，使经典流程与下游量子电路相互解耦。值得注意的是，CNN+LSTM 与 AMP*e* (CNN+LSTM)的深度压缩结构相同，但瓶颈层大小不同，分别为 100 和 32。因此，它们代表两种完整的模型配置，其参数在训练过程中分别独立优化。与当前先进方法中的量子比特数量一致 [6, 10],，鉴于 NISQ时代的限制，所有模拟均采用 5 个量子比特。

量子器件仿真。 这些模型通过 PennyLane 仿真设备 default.qubit 实现。初始训练在解析模式 下进行，以确立

理想（无噪声）条件下的性能。不过，为了考虑 NISQ设备固有的噪声、退相干和门误差，我们将评估扩展到含噪场景。我们使用 qiskit.aer模拟器来复现真实 IBM硬件（虚拟后端）的噪声特征。具体而言，这些含噪特征代表真实量子处理器在其生命周期某一时点的静态快照，记录了目标设备当时的物理错误率、量子比特退相干和读出不保真度。对于给定的虚拟后端，通过有限数量的测量次数（S）估算期望值，以近似电路基态的概率分布，从而考虑硬件运行约束。无噪声 基准采用的解析模式对应于 $S=\infty$。因此，与该基准的偏差共同反映了（*i*）有限采样测量误差，以及（*ii*）模拟的量子设备噪声。

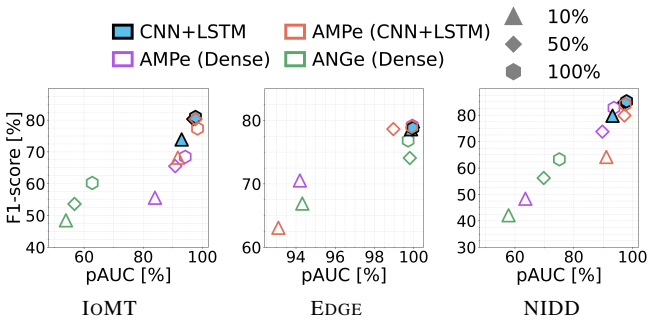


图 2. 在不同训练数据比例（$\delta \in \{10、50、100\}$%）下，DL 和 QML 模型在所有数据集上的 F1分数与部分曲线下面积（FPR 最大值为 1%）对比。数值取 5 折的平均值。

评估策略。 评估遵循成熟的分层五折交叉验证协议。对于每一折，我们保留训练数据的20%用于验证，并仅使用训练集拟合*MinMax* 归一化，以防止数据泄漏。实验分别针对二分类 和多分类误用检测场景进行，主要指标分别采用部分曲线下面积（pAUC）和（宏平均）F1分数。pAUC衡量固定假阳性率（FPR）区间内受试者工作特征（ROC）曲线下的（部分）面积（在本实验设置中，pAUC取FPR为1%时的值）。为满足现实NIDS部署中网络服务连续性的运行要求，我们优先采用较低的FPR阈值，从而限制误报。相反，（宏平均）F1分数对各类别精确率和召回率的调和平均值取平均，为良性流量和每个特定攻击类别提供平衡的性能度量。它不仅量化NIDS区分良性流量与恶意流量的能力，还衡量其区分不同攻击类型的能力。最后，在多分类场景中，我们通过预期校准误差（ECE）评估校准情况；该指标利用对最大预测类别概率划分的10个区间，量化预测置信度与实际准确率之间的偏差。ECE越高，表明校准效果越差；在区分不同流量类别时，NIDS置信度估计的可靠性也越低。

量子网络入侵检测系统的实用化：五个关键视角

OC1：数据效率。我们首先分析模型在训练数据有限的情况下的泛化能力。这种设定反映了现实世界中的物联网安全场景：新出现的威胁和零日攻击往往缺乏足够的样本来训练网络入侵检测系统。为模拟这些场景，我们系统地调整训练集比例 δ，并评估其对性能的影响。我们将 $\delta \in \{10, 50, 100\}$%，跨度设为一个数量级，同时保留足够数量的训练样本。图 2 比较了随着训练集规模增大，各量子机器学习模型的二分类性能（假阳性率为 1% 时的部分曲线下面积）和多分类误用检测性能（F1 分数）。作为对比，我们也纳入了深度学习基线（CNN+LSTM）的结果。分析揭示了三个 决定所提方案有效性的主要性能因素

架构：（*i*）特征映射的表达能力；（*ii*）数据压缩骨干网络的深度；以及（*iii*）量子处理的影响。在各种量子编码策略中，ANG*e* 效果最差，尤其是在 IOMT 和 NIDD 数据集上。例如，在前者中，其性能未超过 54% 的 F1 分数和 61% 的部分曲线下面积，凸显了它在这些环境中捕捉复杂威胁模式的能力有限。相比之下，AMP*e* 展现出更强的表征能力，在同一数据集上的 F1 分数（部分曲线下面积）最高可达 69%（95%）。这种编码还体现出显著的数据效率：仅使用 50% 的数据集进行训练，性能就与使用完整数据集训练的模型相差不超过 3%，这表明量子嵌入能够以显著更少的样本捕捉威胁特征。

骨干网络执行的数据压缩同样具有决定性作用。事实上，在保持 AMP*e* 特征映射不变的情况下，我们观察到，从浅层架构（全连接层）转向更深层架构（CNN+LSTM）会带来显著的性能提升，即使在极端数据稀缺（δ =10%）的情况下也是如此；在 IOMT 上，部分曲线下面积仍保持在 91% 以上。有趣的是，在 EDGE 上，性能在训练数据达到 50% 时趋于饱和，部分曲线下面积达到 $>$ 99.8%。这表明，一个规模更小且分层良好的训练集可以充当隐式正则化器，使模型能够忽略随机网络波动，并优先提取通用的威胁特征。

最后，在完整数据场景下，采用相同骨干网络架构的 CNN+LSTM 和 AMP*e* (CNN+LSTM) 之间的性能差距在 F1分数 上最多为 3%，在部分曲线下面积 上最多为 1%，凸显了共同压缩结构的主导作用。不过，由于这两个模型在训练过程中调整的参数并不相同，这一差距并不能完全归因于量子电路本身。然而，在极端数据稀缺条件下，二者出现了显著差异：当数据量为δ = 10%时，AMP*e* (CNN+LSTM) 架构的 F1分数 显著低于其经典对应模型，在 NIDD 数据集上的差距为 $\approx$ 15%。这表明，与经典层相比，量子电路需要更多样本才能达到相同水平的训练稳定性。基于这些发现，鉴于 AMP*e* (CNN+LSTM) 展现出最高的整体有效性，我们选择它与 CNN+LSTM 深度学习基线进行后续深入分析。

OC2：跨域泛化。我们开展跨域评估，以评估深度学习和量子机器学习的鲁棒性

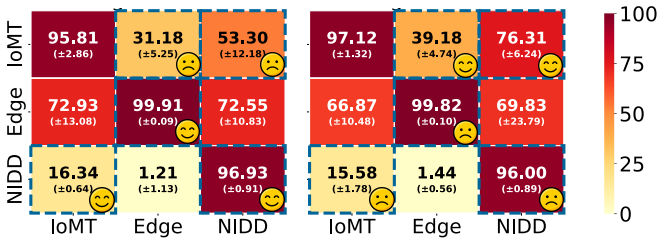


图3. 交叉评估设置中 CNN+LSTM（左）和 AMP*e* (CNN+LSTM)（右）的部分曲线下面积（假阳性率最大为 1%）。行/列分别表示源/目标数据集。数值为 $mean \pm std.dev.$ 个折的结果。虚线框标出 ☺ 在 ≥ 4/5 个折中部分曲线下面积高于 ☹ 的配置。

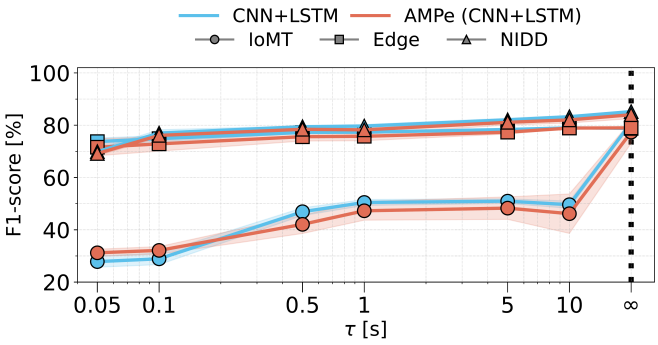


图4. 随着最大观测窗口（τ）增大，AMP*e* (CNN+LSTM) 和 CNN+LSTM 的 F1分数变化。 $\tau = \infty$：无约束场景。数值为 $avg. \pm std.$ 个折的结果。 x轴采用对数刻度。

模型在域偏移下的表现。我们的目标是确定模型能否泛化到陌生的网络环境，而无需进行站点特定再训练。这项分析对于网络入侵检测系统的实际部署尤其重要，因为在相同网络数据分布上训练和测试的模型，可能会无意中掩盖限制其可移植性的环境特定模式。在评估中，我们在源数据集上训练模型，并使用源域中的良性双向流和目标域中的恶意双向流进行评估（源→目标）。这一策略使我们能够评估在源网络上学习到的攻击模式，是否有助于检测来自其他网络的恶意流量。此外，在训练和测试中都保留源域的良性样本，有助于减少部署后的误报 [12]。图3显示了所有源→目标配对在假阳性率为1%时的部分曲线下面积，并对比AMP*e* (CNN+LSTM)和CNN+LSTM。域内基线显示在主对角线上，其中良性样本和恶意样本均来自同一数据集。正如预期，域内性能始终较高，两个模型在所有数据集上的部分曲线下面积均超过95%。这证实，当训练与测试分布一致时，这些架构十分有效。然而，对域偏移的鲁棒性因源数据集而异。使用EDGE训练的模型展现出更强的可移植性，在所有陌生环境中均保持至少67%的部分曲线下面积。相比之下，使用NIDD训练会导致性能显著下降，有时甚至是灾难性的（部分曲线下面积 < 17%）模型难以区分良性流量和恶意流量。泛化能力不足可能是由于NIDD的特定组成所致，其中约80%的双向流仅包含单个数据包。如此有限的观测窗口无法支持模型学习复杂的时间或结构模式，迫使模型过拟合于环境特定伪影，而这些伪影无法转化为可移植的攻击表征。对比AMP*e* (CNN+LSTM)和CNN+LSTM可见，两种配置的跨域泛化存在明显差异。在IOMT→NIDD中，与深度学习基线相比，AMP*e* (CNN+LSTM)的平均部分曲线下面积达到 +23%，但在EDGE→IOMT上的表现下降了6%，这表明量子增强特征映射可能能够提取更具鲁棒性的跨域表征。然而，各折之间较大的标准差使得这使得难以明确判断哪种模型更胜一筹。不过，通过逐折检查可以看出一种趋势。因此，图3还标出了在若干配置中，某个模型至少在 5 个折中的 4 个折上取得更高部分曲线下面积的情况。这些结果表明，跨域鲁棒性仍是一项关键但常被忽视的挑战。

OC3：时间约束韧性。 我们进行了细粒度分析，以评估网络入侵检测系统在有限流量观测时间内提供准确检测的能力；对于在攻击下的网络设备或用户而言，这是快速进行威胁缓解和减少损害的关键因素。为模拟有时间限制的观测，我们为每个双向流强制设置最大观测窗口 τ，仅使用在该窗口内观测到的数据包进行推理。这种运行设置不同于早期性，后者将分析限制在固定的数据包数量（N_p）内，而 τ 施加的是严格的时间约束。因此，如果在 τ 到期前未能完整收集最初的 N_p个数据包，则会在该时刻仅使用可用数据包触发分类。图4展示了随着 τ 增加，CNN+LSTM和AMP*e* (CNN+LSTM)模型在所有数据集上取得的F1分数。我们还纳入了“无约束”场景（$\tau = \infty$）。不同数据集上的实验结果显示，对观测窗口 τ的敏感性呈现出明显不同的行为。在EDGE上，当施加严格的运行延迟约束（$\tau \leq 1$秒）时，AMP*e* (CNN+LSTM)与CNN+LSTM之间的性能差距约为2%，差异较小。随着 τ增加，这一差距逐渐缩小，两种架构早在 $\tau = 10$秒时就表现相近。相反，NIDD上的性能基本不受观测窗口影响。事实上，无论 τ如何，两种模型都表现出近乎相同的行为，这与我们此前对该数据集构成的观察一致：由于大多数双向流都极短，施加时间期限对可用于检测的信息影响甚微。最显著的差异出现在IOMT上。在高响应区间（$\tau \leq 100$毫秒）内，AMP*e* (CNN+LSTM)取得了高于CNN+LSTM的F1分数，表明在有限流量可见性下，量子增强特征映射可能更有效地提取威胁特征。然而，这一趋势

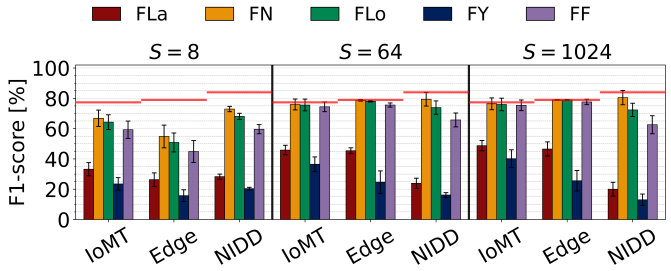


图 5。AMP 的 F1-score$e$ （CNN+LSTM）在不同后端和测量次数下的（S）。—：无噪声后端。数值为 avg。$\pm std$。在5 折上的平均值。

表 I AMP 的预期校准误差e （CNN+LSTM）。数值为5 折上的平均值。绿色（分别为红色）单元格突出显示欠自信（分别为过度自信）的模型。🗄：数据集。**S**：测量次数。：无噪声。🏁：深度学习基线。

🗄	S	Fake backend					💎	🏁
		FLa	FN	FLo	FY	FF		
IoMT	8	14.4 ±6.5	2.3 ±2.3	2.8 ±2.5	17.9 ±8.9	3.8 ±3.0		
	64	9.7 ±2.5	2.0 ±0.6	3.0 ±1.2	15.0 ±4.5	4.9 ±1.7	0.7 ±0.6	0.6 ±0.3
	1024	17.9 ±4.8	1.9 ±0.9	2.9 ±1.3	27.0 ±5.7	4.5 ±1.7		
Edge	8	13.8 ±6.7	1.3 ±0.8	2.0 ±1.4	20.3 ±10.8	3.3 ±2.3		
	64	7.8 ±2.6	0.5 ±0.1	0.8 ±0.3	9.3 ±3.7	1.6 ±0.6	0.1 ±0.0	0.2 ±0.1
	1024	10.1 ±3.8	0.5 ±0.1	0.7 ±0.2	22.8 ±14.5	1.4 ±0.5		
NIDD	8	7.7 ±3.3	3.3 ±1.0	3.6 ±1.2	10.0 ±2.5	4.1 ±1.6		
	64	13.1 ±3.9	3.0 ±0.5	4.6 ±1.1	16.1 ±4.3	7.6 ±1.5	0.2 ±0.1	0.3 ±0.2
	1024	20.6 ±3.5	2.7 ±0.6	4.2 ±1.1	22.7 ±7.9	7.4 ±1.8		

在 $\tau \geq 500$毫秒时反转，此时基线成为性能最佳的模型。值得注意的是，在 $\tau \in [0。5, 10]$秒时，两种模型都表现出更高的不稳定性，这一点可由更大的方差得到证明。这表明，这些时间间隔对应于一个过渡阶段，在此阶段，可用数据包可能尚未形成底层威胁的稳定指纹。最后，在系统未施加响应性约束时，性能差距保持一致，CNN+LSTM的表现优于AMPe (CNN+LSTM)，F1分数高出 $\approx$3%。

OC4：对噪声和采样误差的鲁棒性。当代量子硬件天然容易受到环境退相干和门级错误的影响，这些因素会成为网络入侵检测系统处理流程中的信号失真因素。为量化物理噪声和有限次采样对检测能力的影响，我们将理想模拟（无噪声）与五个 含噪后端进行比较：*FakeLagos* （FLa）、*FakeNairobi* （FN）、*FakeLondon*（FLo）、*FakeFractional* （FF）和*FakeYorktown* （FY）。这些后端涵盖了广泛的噪声特征，展现了含噪声中等规模量子技术从旧式平台（如 FY）到中代架构（如 FLa 和 FN）的演进。这种多样性使我们能够在不同物理条件下评估该框架，包括由不完善的量子比特测量导致的读出错误，以及有限的量子比特相干时间，即量子比特在发生环境退相干前保持其量子态的时长。重要的是，这些后端最多支持 7 个量子比特，与我们的特征映射维度相匹配，从而确保能够真实评估物理含噪声中等规模量子系统的限制。图 5 展示了 AMPe (CNN+LSTM) 在所有数据集上的 F1分数。我们通过改变 $S \in \{8、64、1024\}$ 来评估随机采样，考察测量开销与检测准确率之间的权衡，并确定实现令人满意的运行性能所需的最小采样深度。正如预期， S 是影响性能的主要因素，将其从 8 增至 1024 通常会显著提升 F1分数。一个显著的例外是 NIDD：随着 S 增加，FLa 和 FY 的 F1分数反而下降。一种可能的解释是，在高噪声条件下（例如读出错误率较高或相干时间较短），增加 S 只是进一步巩固系统性硬件偏差，反而可能降低性能。一方面，在所有后端上仅采用 8 次测量在实际中都不可行；与无噪声相比，差距达 $\geq$10%（例如 FN 在 IOMT 上），这是低测量次数下的采样方差与硬件噪声共同作用的结果。另一方面，FN 和 FLo 是抗噪性最强的架构。这可能源于它们高效的拓扑结构（H形和T形）以及现代原生门（cx/sx），仅需 64 次测量便能达到近乎最优的性能，并且在几乎所有数据集上接近基线（例如，FN 在 IOMT 和 EDGE 上的性能损失最多为 2%）。

OC5：可信度评估。为评估所提出流程在实际部署中的预测可靠性，我们进行了校准分析。尽管模型校准对经典网络入侵检测系统至关重要 [7]，，但在NISQ时代对其进行评估同样关键。事实上，硬件引发的噪声、系统性偏差和门误差会直接扭曲量子电路的输出概率分布，可能导致预测置信度失准。因此，模型可能达到较高的准确率，却在实际表现与预测置信度之间存在严重偏差。这种不匹配会表现为较高的预期校准误差（ECE），表明量子网络入侵检测系统存在校准失配。从实际运行角度看，硬件引发的校准失配可能导致严重的安全故障，例如产生误报，或未能发现真实的入侵行为。相反，经过良好校准的模型能够确保输出概率准确反映其预测的真实可能性。因此，ECE可用于评估现代量子后端的进步能否转化为可靠的实际效益，即通过有效的阈值策略减少警报过载。表I报告了AMPe (CNN+LSTM)在所有数据集上针对每一组（后端、测量次数）组合达到的ECE。作为比较，我们还报告了无噪声场景和深度学习基线的表现。由于两个模型采用相同的骨干网络架构，这项比较有双重目的：(*i*)评估经典-量子混合设计对模型校准的影响；(*ii*)评估与解析无噪声场景相比，校准在含噪的有限采样执行下如何变化。一方面，采用混合模型并不会降低校准水平，反而在EDGE和NIDD上略有提升，使两个模型在所有数据集上均表现出出色的校准效果，ECE始终低于1%。另一方面，将各后端与无噪声场景进行比较，证实了我们先前分析中观察到的趋势，表明含噪的有限采样执行可能是校准性能下降的原因。具体而言，在真实硬件环境下，ECE始终较高

于无噪声基线。此外，部分含噪后端的高度波动进一步证实了校准不稳定性，尤其是FLa和FY，其ECE的标准差较大。在评估的后端中，只有FN的ECE会随着测量次数增加而持续改善，且仅在EDGE上低于1%的ECE阈值。对可靠性图的检查（因篇幅所限未列出）显示，量子电路往往会降低预测置信度，导致置信不足。尽管这有助于在无噪声条件下维持校准，但硬件噪声会加剧这一效应，降低输出概率的质量并造成误分类。这些发现表明，未来容错量子架构所追求的无噪声运行环境能够实现校准良好的基于量子机器学习的网络入侵检测系统，而当前量子硬件则会因硬件噪声和有限测量次数不确定性而导致校准性能下降。

结论

本研究对基于量子机器学习的网络入侵检测系统在具有实际运行意义的相关条件下进行早期入侵检测的实用性进行了评估，涵盖数据效率、跨域泛化、时间韧性、对噪声和采样误差的鲁棒性以及预测可信度。在若干场景中，量子机器学习在训练数据有限的情况下仍保持竞争力（**OC1**），表明其可能适用于新兴威胁和零日攻击。结果凸显了经典压缩对于实现高质量量子嵌入和电路优化的重要性，有助于缩小当前量子机器学习与深度学习之间的性能差距。与此同时，在异构网络环境中的可移植性仍是一项主要挑战（**OC2**）。跨域性能取决于训练数据集的特征，这凸显了对于经典模型和量子模型而言，学习能够超越特定环境流量模式进行泛化的攻击表征存在困难。对时间约束韧性的分析（**OC3**）表明，量子机器学习能够改善早期入侵检测，尤其是在可用流量信息有限的情况下。本研究强调，在实际部署中不能忽视当前量子硬件的局限性（**OC4**）。检测性能和模型校准容易受到硬件引发的噪声、量子比特退相干以及后端特定特征的影响。尽管近期的含噪后端在有利的运行条件下已接近无噪声行为，但某些硬件配置仍会导致校准不准确的预测，从而可能损害所生成警报的可靠性（**OC5**）。总体而言，量子机器学习有望成为未来智能网络安全系统的可行组成部分，尤其适用于异构物联网环境中的早期和自适应入侵检测。从部署角度来看，云端排队和编排开销可能阻碍在线检测，而紧凑的少量量子比特架构则有助于实现本地部署。

未来工作 将探索拥有更多量子比特的更大型量子架构、先进的特征映射、拟设、测量方案以及瓶颈匹配消融实验，以分离量子贡献。我们将把 OC2 扩展到完整的部署场景，从目标网络中获取正常流量和攻击流量，以评估误报和可移植性。此外，通过采用匹配的测量次数预算进行无噪声模拟，可以将有限测量次数采样误差与物理噪声区分开来。我们计划开展量子硬件实体部署，以评估模拟之外的端到端决策延迟和吞吐量，并通过事后校准缓解采样和噪声相关的校准偏差。最后，我们将研究容错运行机制，以了解硬件噪声变得不再重要时 OC 的表现，同时仍需考虑有限采样和网络约束。

参考文献

[1] J. Zhou 等，“基于云的物联网的安全与隐私：挑战”，IEEE通信杂志，第55卷，第1期，第26–33页，2017年。[2] N. Bugshan 等，“基于入侵检测的集成学习与零接触网络微服务”，IEEE通信杂志，第61卷，第6期，第86–92页，2023年。[3] L. Alchieri 等，“量子机器学习简介：从量子逻辑到量子深度学习”，量子机器智能，第3卷，第2期，第28页，2021年。[4] V. Spadari 等，“通过量子机器学习实现网络入侵检测：现实检验”，载于 IEEE INFOCOM研讨会，2025年，第1–6页。[5] A. Tudisco 等，“评估变分量子线路模型在金融欺诈检测中的计算优势”，IEEE Access，第12卷，第102 918–102 940页，2024年。[6] F. Cirillo 等，“在含噪量子计算机上对入侵检测量子机器学习方法进行基准测试”，量子机器智能，第8卷，第1期，第27页，2026年。[7] A. Nascita 等，“面向互联网流量分类与预测及入侵检测的可解释人工智能综述”，IEEE通信综述与教程，第27卷，第5期，第3165–3198页，2025年。[8] J. Biamonte 等，“量子机器学习”，《自然》，第549卷，第7671期，第195–202页，2017年。[9] D. Abreu 等，“QML-IDS：量子机器学习入侵检测系统”，载于 IEEE计算机与通信国际研讨会（ISCC），2024年，第1–6页。[10] I. Ahmad 等，“用于异常检测的量子机器学习：更智能、更安全的物联网网络的未来”，IEEE网络，2026年。[11] S. Ismail 等，“物联网与工业物联网中的入侵检测：使用 TON IoT、WUSTL-IIOT-2021 和 EdgeIIoTset 数据集比较轻量级机器学习技术”，IEEE Access，第13卷，第73 468–73 485页，2025年。[12] G. Apruzzese 等，“网络入侵检测系统的机器学习交叉评估”，IEEE网络与服务管理汇刊，第19卷，第4期，第5152–5169页，2022年。[13] A. Reejisha 等，“利用量子神经网络改进图像分类”，载于 IEEE计算机、电子与电气工程及其应用国际会议（IC2E3），2023年，第1–12页。[14] X. Ding 等，“QFMLNet：面向6G天地一体化网络的人工智能驱动量子通信架构”，IEEE通信杂志，第63卷，第7期，第34–40页，2025年。[15] M. Lopez-Martin 等，“用于物联网的卷积神经网络与循环神经网络网络流量分类器”，IEEE Access，第5卷，第18 042–18 050页，2017年。

Domenico Ciuonzo [SM] (domenico.ciuonzo@unina.it) 是那不勒斯费德里科二世大学的副教授。他的研究涉及数据融合、网络分析、物联网、信号处理和人工智能。

Idio Guarino [M] (idio.guarino@unibo.it) 是意大利博洛尼亚大学的助理教授。他的研究重点是互联网流量分析，致力于开发数据驱动方法，以对网络流量进行建模、分类和预测。

Alfredo Nascita (alfredo.nascita@unina.it) 是那不勒斯费德里科二世大学的助理教授。他的研究兴趣包括互联网流量分析、机器学习与深度学习，以及可解释人工智能。

Antonio Pescap`e [SM](pescape@unina.it) 是那不勒斯费德里科二世大学的正教授。他的工作重点是互联网的测量、监测与分析。

Vincenzo Spadari (vincenzo.spadari@unina.it) 是那不勒斯费德里科二世大学的博士生。他的研究重点是将量子机器学习应用于网络流量数据分析。