International Journal For Advanced Research In Science & Technology
A peer reviewed international journal www.ijarst.in
ISSN: 2457-0362

HYBRID QUANTUM MACHINE TEACHING APPROACHES FOR NEXT-GENERATION PHISHING URL DETECTION

MAHESH S
RESEARCH SCHOLAR, SUNRISE UNIVERSITY, ALWAR RAJASTHAN

DR. PRASADU PEDDI
PROFESSOR, SUNRISE UNIVERSITY, ALWAR RAJASTHAN

ABSTRACT

Phishing remains one of the most widespread and persistent cybersecurity threats, exploiting deceptive Uniform Resource Locators (URLs) to lure users into revealing sensitive information. While machine learning and deep learning models have achieved substantial success in phishing detection, they often face limitations in scalability, interpretability, and adaptability to evolving attacks. This research introduces a theoretical framework for Hybrid Quantum Machine Teaching (HQMT) approaches in phishing URL detection. Unlike conventional machine learning, where models are trained purely on classical data, quantum-enhanced learning leverages the principles of superposition, entanglement, and parallel computation to accelerate pattern recognition in complex feature spaces. By combining quantum computing paradigms with teacher–student learning architectures, the proposed model seeks to enhance detection accuracy, resilience against zero-day attacks, and efficiency in handling large-scale datasets. This paper presents the theoretical underpinnings of HQMT, discusses its integration into phishing detection pipelines, and outlines potential performance gains over purely classical systems.

Keywords: Phishing Detection; Quantum Machine Learning; Machine Teaching; Hybrid Systems; Network Security

I. INTRODUCTION

Phishing is one of the most persistent and damaging threats in modern cybersecurity. As an attack vector, it exploits human trust and technological loopholes by tricking users into clicking malicious URLs that masquerade as legitimate websites. Through these deceptive links, attackers obtain sensitive data such as login credentials, banking details, or personal identification information. Over the past two decades, phishing has evolved in sophistication, ranging from simple email scams to complex, multi-stage attacks involving artificial intelligence and large-scale automation. The speed at which attackers adapt to detection strategies, coupled with the ease of deploying phishing campaigns, makes it an especially challenging threat to contain. Traditional methods, such as blacklist-based filters or signature matching, quickly become obsolete in the face of zero-day phishing URLs, underscoring the need for more adaptive and intelligent detection systems.

International Journal For Advanced Research In Science & Technology
A peer reviewed international journal www.ijarst.in
ISSN: 2457-0362

In response to these challenges, machine learning (ML) and deep learning (DL) have been widely adopted for phishing URL detection. By extracting lexical, host-based, and behavioral features from URLs, ML models such as Random Forests, Support Vector Machines, and Gradient Boosting have been used to classify phishing and legitimate links. Deep learning models, including Convolutional Neural Networks (CNNs) and Recurrent Neural Networks (RNNs), have further improved detection by automatically capturing sequential and contextual dependencies in URL structures. These advancements have yielded higher detection rates and reduced reliance on manual feature engineering. However, existing ML and DL approaches face critical limitations. First, they often require vast amounts of labeled data to achieve robust generalization. Second, they can be computationally intensive, limiting their scalability for real-time applications in enterprise networks or large-scale cloud systems. Third, adversarial phishing URLs designed to fool classifiers remain a significant challenge. These shortcomings call for an exploration of alternative paradigms that can deliver higher efficiency, adaptability, and resilience.

Quantum computing has emerged as a transformative technology with the potential to reshape numerous fields, including cybersecurity. Unlike classical computing, which relies on binary bits, quantum systems use qubits capable of existing in superposition states and leveraging entanglement. This unique property allows quantum computers to process and represent information in exponentially larger feature spaces. In the context of phishing detection, quantum machine learning (QML) algorithms can encode complex patterns in high-dimensional Hilbert spaces, offering improved discrimination between phishing and legitimate URLs. Early studies on Quantum Support Vector Machines (QSVMs) and Quantum Neural Networks (QNNs) suggest that QML can achieve superior accuracy on structured and semi-structured data compared to classical models. Nonetheless, practical implementation is hindered by limitations of current quantum hardware, including noise, limited qubit counts, and high error rates.

To address these challenges, researchers have begun investigating hybrid quantum–classical approaches. These systems exploit the computational strengths of quantum circuits for tasks such as feature encoding or optimization, while relying on classical algorithms for stable and interpretable decision-making. Such integration provides a feasible near-term path for applying quantum advancements to real-world problems, including cybersecurity. Within this hybrid paradigm, an emerging concept known as Hybrid Quantum Machine Teaching (HQMT) presents a promising direction. Machine teaching is a framework in which a “teacher” intelligently selects and presents the most informative examples to a “student” model, thereby reducing the sample complexity and accelerating learning. When combined with quantum computation, machine teaching can optimize the training process, guiding quantum-enhanced learners to detect phishing URLs more effectively and with fewer resources.

International Journal For Advanced Research In Science & Technology
A peer reviewed international journal www.ijarst.in
ISSN: 2457-0362

The HQMT framework for phishing detection envisions a synergistic architecture where quantum-enhanced teacher models leverage optimization algorithms, such as quantum annealing, to identify the most discriminative phishing samples. These curated datasets are then used to train student models—either classical learners or hybrid quantum–classical classifiers—thereby reducing redundancy and improving adaptability. By embedding phishing URLs into quantum feature spaces through encoding strategies like amplitude embedding or variational quantum circuits, the framework enables the capture of subtle relationships between URL tokens, domains, subdomains, and query strings that classical approaches often overlook. The ensemble decision-making layer then integrates predictions from both quantum and classical learners to ensure robustness against noise, adversarial manipulation, and hardware imperfections.

The significance of HQMT lies in its potential to deliver next-generation phishing detection systems that overcome the limitations of current models. First, it enhances efficiency by requiring fewer labeled samples, addressing the data scarcity problem often faced in cybersecurity. Second, it improves adaptability to zero-day phishing attacks by leveraging teacher–student feedback loops that enable rapid retraining with minimal data. Third, its scalability benefits from quantum parallelism, allowing large-scale URL datasets to be processed faster than classical-only systems. Finally, the hybrid integration with classical ensemble methods ensures interpretability and practical deployment, bridging the gap between cutting-edge quantum research and real-world cybersecurity applications.

As phishing attacks grow increasingly sophisticated, conventional detection strategies will continue to struggle in keeping pace. The integration of quantum computing principles, hybrid architectures, and machine teaching frameworks represents a paradigm shift in phishing detection research. By harnessing the computational power of quantum systems alongside the structured guidance of teacher–student models, HQMT approaches offer a theoretical foundation for building resilient, adaptive, and future-proof detection mechanisms. While current quantum hardware presents limitations, the rapid progress in this field makes HQMT a viable and forward-looking strategy for the cybersecurity community. This research contributes to the theoretical development of HQMT for phishing detection, positioning it as a critical pathway for advancing next-generation defenses in the ongoing battle against cyber threats.

II. TRADITIONAL AND MACHINE LEARNING APPROACHES

Phishing detection initially relied on traditional rule-based methods, such as blacklists, heuristics, and signature-based filters. Blacklists store known phishing URLs or domains and block access when a match is detected. Although simple and widely implemented in browsers and email clients, blacklist-based methods are reactive and fail against zero-day phishing attacks, which exploit newly created domains not yet recorded. Similarly, heuristics rely on manually crafted rules,

International Journal For Advanced Research In Science & Technology
A peer reviewed international journal www.ijarst.in
ISSN: 2457-0362

such as URL length, the presence of suspicious characters, or mismatched domain names. While these rules offer some defense, they lack adaptability to evolving phishing strategies and often result in high false positives. These limitations motivated the adoption of machine learning (ML) approaches, which can learn patterns from large datasets and adapt to new attack vectors without explicit rules.

Machine learning techniques introduced a more data-driven paradigm for phishing URL detection. Instead of relying on pre-defined rules, ML models extract lexical, host-based, and content-based features from URLs to distinguish phishing from legitimate sites. Lexical features include token patterns, URL length, and the presence of special symbols such as "@" or "-". Host-based features cover attributes like domain age, IP address information, and DNS records, which often reveal anomalies in phishing domains. Content-based features examine webpage structure, embedded links, and script behavior to capture malicious intent. Using these features, models such as Support Vector Machines (SVMs), Naïve Bayes classifiers, and Random Forests demonstrated strong performance in early phishing detection research. For instance, Random Forests benefit from ensemble learning, combining multiple decision trees to reduce variance and improve robustness, while SVMs are particularly effective for high-dimensional feature spaces.

The adoption of ensemble learning methods further advanced phishing detection by combining multiple classifiers to reduce bias and variance. Techniques such as Gradient Boosting, AdaBoost, and XGBoost have been applied to phishing URL datasets, achieving higher accuracy compared to single models. These methods work by iteratively focusing on difficult-to-classify examples, thereby improving generalization and resilience against diverse phishing patterns. However, they require substantial computational resources for large-scale deployment and may still struggle with highly dynamic phishing attacks designed to bypass detection systems.

Despite their success, classical ML approaches often suffer from feature engineering bottlenecks. The effectiveness of these models depends heavily on the choice and quality of extracted features. Crafting effective feature sets requires domain expertise and does not always capture the evolving nature of phishing attacks. Moreover, attackers often manipulate lexical and structural features of URLs to mimic legitimate domains, making it challenging for static feature-based models to generalize. This vulnerability has driven researchers toward more automated feature extraction techniques, particularly through deep learning.

Deep learning (DL) models such as Convolutional Neural Networks (CNNs) and Recurrent Neural Networks (RNNs) have demonstrated significant promise in phishing URL detection. Unlike traditional ML methods, deep learning models automatically learn hierarchical feature representations directly from raw URLs or webpage content. CNNs, for instance, can capture spatial and sequential dependencies in URL strings, while RNNs are capable of modeling long-term dependencies, making

International Journal For Advanced Research In Science & Technology
A peer reviewed international journal www.ijarst.in
ISSN: 2457-0362

them effective in recognizing subtle manipulations in phishing URLs. Hybrid models, such as CNN-LSTM architectures, have further improved detection accuracy by combining local pattern recognition with long-sequence modeling. These approaches significantly reduce the dependency on manual feature engineering while improving adaptability to unseen phishing attacks.

Nevertheless, both traditional ML and deep learning methods face certain limitations. Classical models often lack scalability and struggle with zero-day attacks, while deep learning models, despite higher accuracy, demand extensive labeled datasets and computational resources. Moreover, adversarial phishing URLs crafted to exploit model weaknesses remain a critical concern. These challenges highlight the need for next-generation detection frameworks that can handle evolving attack strategies, reduce reliance on massive training data, and enhance resilience to adversarial manipulations. This has set the stage for exploring quantum-enhanced and hybrid learning approaches, which combine the strengths of classical ML with the computational advantages of quantum systems.

III. QUANTUM MACHINE LEARNING (QML)

Quantum Machine Learning (QML) is an emerging interdisciplinary field that combines the principles of quantum computing with the methodologies of machine learning (ML) to address complex problems in data analysis and pattern recognition. Unlike classical computing, which encodes information in binary bits, quantum computing leverages qubits, which can exist in a superposition of states. This property allows quantum systems to process information in exponentially larger feature spaces compared to classical systems. Furthermore, quantum entanglement and interference enable unique computational advantages, particularly in tasks involving optimization, high-dimensional feature encoding, and probabilistic modeling. These properties make QML highly attractive for cybersecurity applications, including phishing URL detection, where complex patterns and evolving attack strategies pose significant challenges for classical models.

One of the central motivations for applying QML in phishing detection is its ability to capture non-linear and high-dimensional relationships within URL structures and network features. Phishing URLs often contain subtle variations in tokens, domains, and query parameters that are difficult to detect with traditional models. By encoding URL features into quantum states, QML algorithms can map classical data into quantum Hilbert spaces, enabling more expressive representations of structural and contextual dependencies. For instance, Quantum Support Vector Machines (QSVMs) leverage quantum kernels to achieve superior classification performance compared to their classical counterparts, particularly in cases where feature relationships are too complex for conventional kernel methods.

Another promising direction in QML involves Quantum Neural Networks (QNNs) and Variational Quantum Circuits (VQCs). QNNs adapt the architecture of classical neural networks into quantum

circuits, where layers are represented by quantum gates and parameters are optimized through variational techniques. VQCs, in particular, have shown potential in learning patterns from structured and unstructured data while requiring fewer parameters compared to deep classical models. These quantum-enhanced models are not only capable of capturing intricate dependencies in phishing data but also provide resilience against adversarial manipulations, which often exploit weaknesses in deterministic decision boundaries of classical models.

QML also plays a significant role in addressing the data efficiency problem faced by classical machine learning. In phishing detection, labeled datasets are often limited, especially for zero-day phishing URLs. Quantum-enhanced models, by exploiting the high-dimensional representational capacity of quantum states, can potentially achieve better generalization with fewer samples. This property aligns well with the needs of cybersecurity systems, which require rapid adaptation to new and previously unseen threats. Early studies on quantum kernel methods suggest that QML can achieve strong performance even with limited training data, making it a compelling candidate for phishing URL detection.

Despite its promise, QML is not without challenges. Current quantum hardware suffers from noisy intermediate-scale quantum (NISQ) limitations, including decoherence, limited qubit counts, and high error rates. These constraints make it difficult to train large-scale quantum models or deploy them in real-time production environments. Furthermore, the development of efficient quantum feature encoding strategies remains an open research area. Encoding classical URL data into quantum states without losing essential information is computationally demanding and may affect scalability. Additionally, hybrid optimization—where classical optimizers are used to tune parameters in quantum circuits—faces issues such as barren plateaus, where gradients vanish and hinder learning.

To overcome these challenges, researchers are actively exploring hybrid quantum–classical frameworks, where quantum circuits are integrated with classical machine learning pipelines. In such systems, quantum modules are typically used for feature mapping or kernel estimation, while classical learners handle decision-making and optimization. This synergy leverages the strengths of both paradigms: the expressive power of quantum representations and the stability of classical models. For phishing detection, such hybrid approaches offer a practical near-term solution, enabling the deployment of quantum-inspired techniques on NISQ devices while benefiting from classical computational stability.

In summary, Quantum Machine Learning represents a transformative step in advancing phishing detection and broader cybersecurity applications. By exploiting quantum properties such as superposition, entanglement, and interference, QML can provide richer data representations, higher classification accuracy, and improved adaptability to zero-day phishing attacks. Although current hardware limitations restrict its full-scale deployment, the rapid

International Journal For Advanced Research In Science & Technology
A peer reviewed international journal www.ijarst.in
ISSN: 2457-0362

progress in quantum computing technologies, coupled with hybrid quantum–classical approaches, makes QML a promising frontier for building next-generation phishing URL detection systems. This foundation sets the stage for exploring Hybrid Quantum Machine Teaching (HQMT) frameworks, which aim to enhance QML with intelligent teacher–student paradigms for more efficient and adaptive phishing defense mechanisms.

IV. MACHINE TEACHING

Machine Teaching is an emerging paradigm within artificial intelligence that complements and extends the traditional framework of machine learning. While machine learning focuses on enabling algorithms to extract patterns and knowledge from data, machine teaching shifts the emphasis toward optimally designing the training process itself. In this approach, a "teacher" entity is responsible for carefully selecting, curating, or synthesizing the most informative examples that can accelerate and enhance the learning process of a "student" model. By doing so, machine teaching reduces the data and computational resources required for effective training, addressing one of the fundamental bottlenecks of machine learning in domains with scarce or expensive labeled data, such as phishing URL detection.

The concept of machine teaching is particularly valuable in adversarial environments like cybersecurity, where attackers continuously evolve their tactics. Instead of overwhelming the model with vast and often redundant training data, machine teaching enables the teacher to identify and prioritize critical phishing examples that expose the vulnerabilities of the student model. This selective exposure helps the model generalize more effectively to unseen phishing attacks, including zero-day threats, while minimizing training time and resource consumption. For phishing URL detection, the teacher can, for instance, emphasize URLs with obfuscated tokens, unusual domain hierarchies, or deceptive query parameters, which are more likely to confuse traditional classifiers.

Machine teaching differs from related concepts such as curriculum learning or active learning. In curriculum learning, training samples are presented in a structured order, usually from simple to complex, to facilitate gradual improvement. Active learning focuses on enabling the learner to query the most informative data points. In contrast, machine teaching explicitly models the teacher–student interaction to optimize the overall learning objective. The teacher does not merely provide ordered examples but strategically designs the teaching set or teaching strategy to minimize the learner's sample complexity and maximize performance. This distinction is critical for phishing detection, where the evolving threat landscape requires not just incremental learning but a proactive teaching mechanism capable of steering the learner toward robust and resilient decision boundaries.

The applications of machine teaching span multiple areas, including computer vision, natural language processing, and cybersecurity. In phishing detection, one promising approach is synthetic data

International Journal For Advanced Research In Science & Technology
A peer reviewed international journal www.ijarst.in
ISSN: 2457-0362

teaching, where the teacher generates realistic but diverse phishing URLs to supplement limited training datasets. Another approach is adversarial teaching, where the teacher intentionally introduces adversarially crafted phishing samples during training to harden the student model against real-world evasion attempts. These strategies ensure that the student not only learns to identify known phishing patterns but also develops the capacity to detect novel manipulations.

The integration of machine teaching with hybrid quantum–classical systems presents a novel opportunity to enhance phishing detection. In such a framework, the teacher could leverage quantum optimization algorithms or quantum-enhanced feature selection to identify the most discriminative phishing URLs, while the student could be a quantum, classical, or hybrid learner. This interaction would combine the strengths of machine teaching—efficient sample selection and learning guidance—with the representational power of quantum computing, leading to a more scalable and adaptive detection system. For instance, the teacher could identify URL components that contribute most to adversarial success and encode these into quantum feature spaces, ensuring that the student model learns from the most challenging and informative examples.

Machine teaching offers a paradigm shift in training strategies, moving from passive data-driven learning to an active, guided learning process. By enabling a teacher model to strategically curate training data, machine teaching reduces the dependency on massive datasets, improves adaptability to adversarial threats, and accelerates the convergence of student models. When combined with Quantum Machine Learning (QML), this paradigm becomes particularly powerful for phishing detection, where zero-day attacks and evolving adversarial strategies demand efficient, intelligent, and resilient learning systems. This synergy between machine teaching and quantum computing paves the way for the proposed Hybrid Quantum Machine Teaching (HQMT) framework, which seeks to establish next-generation defenses against phishing URL attacks.

V. HYBRID QUANTUM–CLASSICAL SYSTEMS

Hybrid quantum–classical systems have emerged as a practical and powerful paradigm in the transition from traditional machine learning to full-scale quantum machine learning. Current quantum hardware, often referred to as Noisy Intermediate-Scale Quantum (NISQ) devices, is constrained by limited qubit counts, noise, and short coherence times. These limitations make it infeasible to run large quantum algorithms entirely on quantum hardware. To address this, researchers have proposed hybrid architectures that integrate quantum modules with classical computational resources, leveraging the strengths of both paradigms. In such systems, quantum computing is typically employed for tasks that benefit from quantum properties, such as high-dimensional feature mapping or kernel estimation, while classical computing handles optimization, large-scale data management, and stability. This division of labor enables scalable, near-term applications of quantum-enhanced

learning in real-world domains, including phishing URL detection.

In the context of phishing URL detection, hybrid quantum–classical systems offer unique advantages. Classical models such as Random Forests, Gradient Boosting, or Deep Neural Networks are highly effective at processing large-scale lexical and host-based features extracted from URLs. However, they often struggle with capturing non-linear relationships and subtle structural dependencies in phishing data. Quantum modules, on the other hand, can embed URL tokens, domain hierarchies, and query structures into quantum feature spaces, enabling richer data representations that are difficult to achieve with classical models alone. By combining these capabilities, hybrid systems allow the classical learner to benefit from quantum-enhanced embeddings, improving overall detection accuracy and resilience against obfuscated or adversarial phishing attacks.

One common architecture for hybrid systems is the variational hybrid framework, where a quantum circuit is parameterized and trained using classical optimizers. In this workflow, URL data is first encoded into quantum states using methods such as amplitude encoding or angle encoding. The quantum module, often implemented as a Variational Quantum Circuit (VQC), extracts expressive features through quantum gates, entanglement, and interference. The output of the quantum circuit is then passed to a classical optimizer—such as stochastic gradient descent or evolutionary algorithms—that updates circuit parameters. This iterative feedback loop between quantum and classical components facilitates efficient learning while maintaining compatibility with existing machine learning pipelines. For phishing detection, such architectures allow the system to adapt dynamically to evolving threat patterns while operating within current quantum hardware limitations.

Hybrid systems also support ensemble-style integration, where multiple learners—some quantum, some classical—contribute to the final decision through mechanisms like weighted voting or stacking. For example, a hybrid phishing detection system could consist of a quantum kernel estimator combined with classical classifiers such as Logistic Regression or Random Forest. The quantum module provides non-linear decision boundaries that enhance the robustness of phishing detection, while classical ensemble methods contribute stability and interpretability. Such multi-modal approaches significantly reduce the variance and bias present in standalone models, thereby improving generalization to unseen phishing URLs.

Another advantage of hybrid quantum–classical systems is their potential to improve training efficiency. Quantum-enhanced feature extraction reduces the need for large datasets by producing high-dimensional representations that increase separability between phishing and legitimate URLs. This is particularly important in detecting zero-day phishing attacks, where limited labeled data is available. Moreover, quantum circuits can capture correlations between URL components that are invisible to linear classical models. When combined with

International Journal For Advanced Research In Science & Technology
A peer reviewed international journal www.ijarst.in
ISSN: 2457-0362

classical post-processing, these systems achieve superior detection without incurring the prohibitive costs of training deep classical networks on massive datasets.

Despite their promise, hybrid systems face challenges that must be addressed for effective deployment. Data encoding remains a significant bottleneck, as transforming large-scale phishing datasets into quantum states is computationally expensive. Moreover, optimization in variational circuits is prone to barren plateaus, where gradients vanish and hinder convergence. Scalability is another concern, as integrating quantum components into large classical pipelines introduces communication overhead between quantum processors and classical hardware. Addressing these issues will require innovations in quantum feature encoding, noise mitigation, and hybrid optimization algorithms.

Hybrid quantum–classical systems represent a practical bridge between current computational realities and the promise of full-scale quantum learning. By combining the high-dimensional representational capacity of quantum circuits with the robustness and scalability of classical algorithms, hybrid architectures offer an effective pathway for next-generation phishing URL detection. These systems not only enhance accuracy and adaptability but also provide resilience against adversarial manipulations and zero-day threats. As quantum hardware continues to evolve, hybrid approaches will likely form the foundation of deployable, real-world cybersecurity solutions, paving the way for advanced frameworks such as Hybrid Quantum Machine Teaching (HQMT).

VI. THEORETICAL FRAMEWORK OF HQMT

The proposed Hybrid Quantum Machine Teaching (HQMT) framework represents a conceptual advancement in the intersection of quantum machine learning, machine teaching, and hybrid quantum–classical architectures. Unlike conventional phishing detection approaches that rely solely on machine learning to passively learn from large datasets, HQMT introduces a proactive and adaptive paradigm in which a teacher module strategically guides a quantum–classical student system. This teacher–student dynamic aims to reduce sample complexity, improve resilience against adversarial phishing attempts, and enable faster adaptation to zero-day attacks. By combining the expressive representational power of quantum learning with the guidance of machine teaching, HQMT offers a theoretical blueprint for next-generation phishing URL detection.

At the core of HQMT lies the teacher–student interaction loop, in which the teacher is tasked with curating or generating the most informative phishing and legitimate URL examples. These examples are selected not to maximize dataset size, but to maximize learning efficiency. The student model, implemented as a hybrid quantum–classical system, receives these tailored samples to optimize its detection capabilities. For instance, phishing URLs with obfuscated subdomains or deceptive query strings may be prioritized by the teacher, as they often exploit vulnerabilities in existing models.

International Journal For Advanced Research In Science & Technology
A peer reviewed international journal www.ijarst.in
ISSN: 2457-0362

This process ensures that the student model is consistently challenged and exposed to critical adversarial examples, enabling stronger generalization to novel phishing techniques.

The quantum component of the HQMT framework serves as a high-dimensional feature mapper, encoding URL tokens into quantum states that capture intricate structural and semantic dependencies. Variational Quantum Circuits (VQCs) or Quantum Kernel Estimators are particularly well-suited for this role, as they exploit superposition and entanglement to uncover correlations that classical learners may overlook. Meanwhile, the classical component handles large-scale optimization, ensemble integration, and decision aggregation, ensuring stability and scalability. This division of computational labor enables HQMT to overcome the limitations of current Noisy Intermediate-Scale Quantum (NISQ) hardware while still harnessing its advantages.

The teaching mechanism in HQMT can adopt several strategies. In curriculum-style teaching, the teacher introduces phishing samples in a structured progression, beginning with simpler manipulations (e.g., misspelled domains) before advancing to more complex obfuscations (e.g., multi-level redirects). In adversarial teaching, the teacher generates or selects adversarial phishing URLs specifically designed to challenge the student's weaknesses, thereby enhancing robustness against real-world attacks. In synthetic data teaching, the teacher leverages generative models to produce realistic but diverse phishing examples that supplement limited real-world datasets. These strategies are not mutually exclusive and may be dynamically combined, depending on the detection environment.

The HQMT framework also integrates a feedback mechanism between the student and teacher. When the student misclassifies a phishing URL, the teacher analyzes the misclassification and adjusts its teaching set accordingly. This creates a closed-loop system of guided learning, where the teacher continuously adapts to the evolving performance of the student. Such adaptivity is particularly important in phishing detection, as attackers frequently update their methods to evade detection. By embedding adaptability into the teaching–learning cycle, HQMT ensures that the system evolves alongside emerging threats, rather than lagging behind them.

From a theoretical standpoint, HQMT provides a pathway toward sample-efficient, adversary-resilient, and generalizable phishing detection systems. It reduces reliance on massive labeled datasets, which are often unavailable for zero-day phishing attacks, and shifts the emphasis toward strategic knowledge transfer. Moreover, by embedding quantum feature representations within the teaching loop, HQMT enhances the expressive power of the learning process, enabling the detection of subtle, high-dimensional phishing patterns that classical models cannot easily capture.

The theoretical framework of HQMT unites three complementary paradigms: quantum machine learning for expressive feature extraction, machine teaching for guided knowledge transfer, and hybrid architectures for practical scalability.

Together, these components form a cohesive system capable of addressing the limitations of traditional phishing detection models. Although currently constrained by hardware and algorithmic challenges, HQMT provides a forward-looking theoretical foundation for next-generation phishing defense mechanisms, capable of adapting intelligently and efficiently in an adversarial cyber landscape.

VII. CONCLUSION

This paper presented a theoretical exploration of Hybrid Quantum Machine Teaching (HQMT) for phishing URL detection. By combining quantum-enhanced feature encoding with teacher–student architectures and hybrid ensemble layers, the proposed framework addresses key limitations of current phishing detection systems. While practical deployment depends on the advancement of quantum hardware, this research positions HQMT as a forward-looking strategy for building resilient, adaptive, and next-generation defenses against phishing attacks.

REFERENCES

1. Adebowale, M.A., Lwin, K.T., Sanchez, E., Hossain, M.A.: Intelligent web-phishingdetection and protection scheme using integrated features of images, frames andtext. Expert Syst. Appl. 115, 300–313 (2019)

2. Al-Ahmadi, S., Alotaibi, A., Alsaleh, O.: PDGAN: phishing detection with gener-ative adversarial networks. IEEE Access 10, 42459–42468 (2022)

3. Al-Alyan, A., Al-Ahmadi, S.: Robust URL phishing detection based on deep learn-ing. KSII Trans. Internet Inf. Syst. (TIIS) 14(7), 2752–2768 (2020)

4. Al-Haija, Q.A., Al Badawi, A.: URL-based phishing websites detection via machinelearning. In: 2021 International Conference on Data Analytics for Business andIndustry (ICDABI), pp. 644–649. IEEE (2021)

5. AlEroud, A., Karabatis, G.: Bypassing detection of URL-based phishing attacksusing generative adversarial deep neural networks. In: Proceedings of the SixthInternational Workshop on Security and Privacy Analytics, pp. 53–60 (2020)

6. U. Z. Asif et al.7. Aljofey, A., Jiang, Q., Qu, Q., Huang, M., Niyigena, J.P.: An effective phishingdetection model based on character level convolutional neural network from URL.Electronics 9(9), 1514 (2020)

7. Alshingiti, Z., Alaqel, R., Al-Muhtadi, J., Haq, Q.E.U., Saleem, K., Faheem, M.H.:A deep learning-based phishing detection system using CNN, LSTM, and LSTM-CNN. Electronics 12(1), 232 (2023)

8. Aung, E.S., Yamana, H.: URL-based phishing detection using the entropy of non-alphanumeric characters. In: Proceedings of the 21st International Conferenceon Information Integration and Web-

based Applications & Services, pp. 385–392(2019)

9. Bozkir, A.S., Dalgic, F.C., Aydos, M.: GramBeddings: a new neural network forURL based identification of phishing web pages through n-gram embeddings. Comput. Secur. 124, 102964 (2023)

10. Butnaru, A., Mylonas, A., Pitropakis, N.: Towards lightweight URL-based phishingdetection. Future Internet 13(6), 154 (2021)

11. Chai, Y., Zhou, Y., Li, W., Jiang, Y.: An explainable multi-modal hierarchicalattention model for developing phishing threat intelligence. IEEE Trans. Dependable Secure Comput. 19(2), 790–803 (2021)

12. Common crawl. https://commoncrawl.org/16. Curlie. https://curlie.org/17. Dutta, A.K.: Detecting phishing websites using machine learning technique. PLoSONE 16(10), e0258361 (2021)

13. 19. Feng, T., Yue, C.: Visualizing and interpreting RNN models in URL-based phishingdetection. In: Proceedings of the 25th ACM Symposium on Access Control Modelsand Technologies, pp. 13–24