Adaptive and Quantum-Resilient Intrusion Detection for Wireless Sensor Networks and IoT Environments

Mathan Kumar Mounagurusamy
Department of Computing Technologies, School of Computing, SRM Institute of Science and Technology, Kattankulathur, Chennai, Tamilnadu, India
drmathan16@gmail.com (corresponding author)

A. Anil Kumar Reddy
Department of CSE, Samskruti College of Engineering and Technology, Medchal Telangana, India
anilkumarreddyaddula0@gmail.com

C. M. Velu
Department of CSE (AIDS), Saveetha Engineering College, Thandalam, Chennai, Tamil Nadu, India
cmvelu41@gmail.com

Gera Vijaya Nirmala
Department of ECE, CVR College of Engineering, Hyderabad. Telangana, India
g.vijayanirmala@cvr.ac.in

D. Arivazhagan
AMET Business School, Academy of Maritime Education and Training Deemed to be University, Chennai, Tamilnadu, India
prof.arivazhagan@ametuniv.ac.in

Myasar Mundher Adnan
Medical Instrumentation Techniques Engineering Department, College of Engineering and Technologies, Al-Mustaqbal University, Hillah, Babil, Iraq
maiserlove06@gmail.com

Rahmaan K.
Department of Artificial Intelligence and Data Science, Mahendra Engineering College, Mallasamudram, Namakkal, Tamil Nadu, India
ksrahmaan2204@gmail.com

T. Prabhakaran
Department of CSE, JB Institute of Engineering and Technology, Hyderabad, Telangana, India
prabaakar.t@gmail.com

Received: 5 February 2025 | Revised: 26 February 2025 and 12 March 2025 | Accepted: 17 March 2025

 DOI: https://doi.org/10.48084/etasr.10464

ABSTRACT

Integrating Wireless Sensor Networks (WSNs) with the Internet of Things (IoT) has transformative potential for data acquisition, processing, and decision-making across dynamic connected environments. Ensuring the security and integrity of these systems is paramount, especially in the face of increasingly

sophisticated cyber threats. This study introduces a novel security framework that combines Quantum Key Distribution (QKD) with an adaptive Deep Reinforcement Learning (DRL)-based Intrusion Detection System (IDS), specifically designed to address the unique challenges of the WSN-IoT ecosystem. The key innovation lies in integrating QKD not only for encryption but also as a dynamic quantum-secure layer that continuously adapts to security requirements based on real-time threats and communication patterns. Unlike previous approaches that focus primarily on routing and resource allocation, the proposed framework employs DRL with Proximal Policy Optimization (PPO) to refine intrusion detection by adapting its policies based on evolving attack signatures and threat types. This dual-layer QKD-DRL approach enhances intrusion detection accuracy and establishes a self-optimizing, quantum-secure communication protocol. Tested using the CICIDS2017 dataset, the proposed model achieved a 99.75% detection rate, outperforming traditional Random Forest (97.12%) and Deep Neural Network (96.88%) models. This improvement underscores the efficacy of combining quantum cryptographic techniques with DRL-based adaptive learning, providing a robust, real-time defense mechanism for IoT-driven environments in applications such as smart cities, healthcare, and industrial IoT systems. Thus, the proposed QKD-DRL framework sets a new standard for scalable, secure communication and threat mitigation in the IoT ecosystem.

Keywords-quantum key distribution; deep reinforcement learning; wireless sensor networks; Internet of Things; intrusion detection; cyber security; deep Q-network

I. INTRODUCTION

Within the rapidly evolving field of Wireless Sensor Networks (WSNs) and Internet of Things (IoT) ecosystems, the convergence of these advanced technologies is poised to revolutionize methods of data acquisition, processing, and decision-making frameworks [1-2]. The paramount significance in addressing integrated systems' security and integrity lies in the necessity for innovative development to counter the diverse and evolving cyber threats in today's landscape [3]. This study introduces an innovative security framework designed for the IoT and WSN contexts, based on the synergistic combination of Quantum Key Distribution (QKD) and Deep Reinforcement Learning (DRL) [4].

In [5], a security framework for the IoT and WSN environments was presented, based on the idea that QKD and DRL work well together. This dual strategy establishes a quantum-safe protocol for data transactions within the WSN-IoT ecosystem, protecting it against cyber threats [6-7]. The incorporation of DRL improves the security of the framework by enabling it to learn and adapt to the emergence of new threats [8]. This research represents a significant advancement with potential applications spanning various domains, including smart cities, healthcare, and industrial IoT-connected systems, where security and reliability are of paramount importance. Based on this foundation, the combination of QKD and DRL algorithms is expected to improve in the coming years to create strong and smart security solutions that work in IoT-driven environments [9].

This study presents an intrusion detection system that works in the WSN-IoT framework. It uses a DRL algorithm implemented through a Deep Q-Network (DQN) and the Proximal Policy Optimization (PPO) technique to make policies more effective. The proposed system introduces an innovative QKD framework explicitly tailored for the WSN-IoT ecosystem. This initiative seeks to establish a foundation for data exchanges resistant to quantum threats using advanced cryptographic protocols. Integrating DQN with PPO presents a compelling approach for enhancing decision-making processes in reinforcement learning frameworks. In the QKD-DRL framework, the physical layer handles quantum key transmission, while the network layer incorporates adaptive key management and routing. Application-layer compatibility is ensured through API-based interaction with existing IoT security protocols, such as TLS and IPSec, allowing easy adoption in real-world deployments. This framework demonstrates enhanced classification accuracy and a dynamic threat response capability that surpasses conventional detection models.

Recent decades have seen a significant emergence of innovative and foundational approaches to privacy-preserving communication [10, 11]. Much research has been conducted on the use of FL-based security in IoT, with a comprehensive examination of FL-enabled schemes designed to combat emerging cyber threats within extensive IoT networks presented in [12].

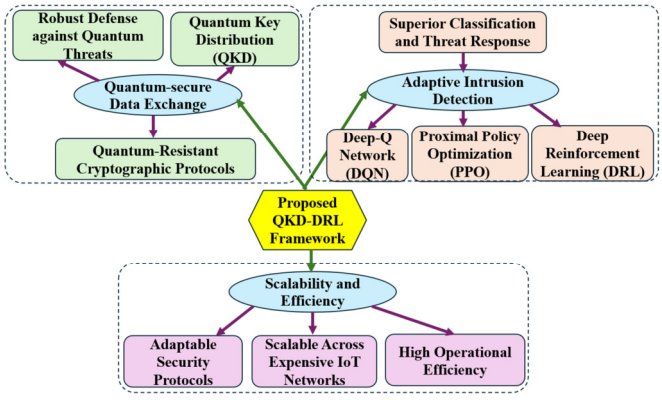

Fig. 1. Workflow of the proposed system.

II. MATERIALS AND METHODS

The proposed framework combines QKD with DRL to enhance security in IoT environments. The QKD module provides a secure way to send and receive data using special encryption methods that are resistant to quantum computing and can handle current and future threats. Figure 1 illustrates the framework of the proposed system. The system comprises

three primary modules, which work in unison to address attacks. The proposed framework aims to provide an end-to-end, adaptive security solution that leverages QKD alongside DRL.

A. Quantum-Secure Data Exchange Module (QSDEM)

This module is tasked with ensuring the security of data exchanges within an IoT ecosystem by implementing QKD. Communication channels incorporating quantum-resistant encryption techniques must be established to safeguard against potential threats posed by quantum computers, both present and future. This study selected and implemented suitable QKD protocols, including BB84 and E91, to facilitate a secure key exchange process. The BB84 protocol is based on the analogy with Alice randomly selecting a key bit sequence and choosing either a rectilinear or diagonal basis for each bit [13]. She encodes each bit into a quantum state based on the selected basis and transmits the qubits to Bob. Privacy amplification refers to transforming a lengthy key into a more concise version, eliminating the partial information accessible to an eavesdropper. The sifted-key rate R_s is defined as:

$$R_s = R \times \big(1 - 2H(e)\big) \quad (1)$$

where R is the raw key rate and $H(e)$ is the binary entropy function given by:

$$H(e) = -e\log_2(e) - (1-e)\log_2(1-e) \quad (2)$$

In the experimental setup, Alice and Bob independently select a random measurement setting, such as an angle for their polarizers, and subsequently measure the polarization state of their respective photons, meticulously documenting the results obtained. The results obtained are analyzed concerning the entanglement characteristics. Alice and Bob utilize the correlated outcomes to establish a shared secret key.

$$|\psi\rangle = \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle) \quad (3)$$

where $|00\rangle$ and $|11\rangle$ are wrangled states of the photon pairs, respectively. The correlation function $E(a,b)$for measurement settings a and b is given by:

$$E(a,b) = \langle\psi|(\sigma_a \otimes \sigma_b)|\psi\rangle \quad (4)$$

where σ_a and σ_b are the Pauli matrices for the respective measurement settings.

1) Integration of Quantum-Resistant Cryptographic Algorithms

This process ensures data integrity and confidentiality using quantum-resistant cryptographic algorithms, such as lattice-based, hash-based, and multivariate polynomial cryptography. For any lattice L with basis vectors $b_1, b_2, \ldots, b_n$, an encryption works by picking a random vector r such that

$$c = m + r \quad (5)$$

that is in the lattice L. Decryption seeks to identify the lattice point closest to c using the basis B.

2) Mitigation of Quantum Attacks

The module employs state-of-the-art cryptographic techniques to reduce the risks of quantum attacks and, thus, secure IoT networks from possible quantum adversaries. The difficulty of problems in lattice-based cryptography, like the shortest vector problem and learning with errors, provides some basic security. The problem is defined as:

$$\mathrm{As} + \mathrm{e} \equiv \mathrm{b} (\mathrm{mod} q) \quad (6)$$

where A is a known matrix, s is the secret vector, and e is the error vector. Security relies on the hardness of finding s given A and b.

Theorem 1: The QSDEM uses QKD protocols and quantum-resistant cryptographic algorithms at its core so that the confidentiality and integrity of data are maintained against any adversary with quantum computing capabilities.

The QSDEM protects any entity possessing data against any adversary, including those equipped with quantum computing capabilities, since it contains protocols for QKD and quantum-resistant cryptography. In the BB84 protocol, Alice and Bob generate and share secret keys by encoding random bit strings into quantum states. e is detected if the eavesdropping causes any disturbance to those states, as per the no-cloning theorem and Heisenberg's uncertainty principle. The possibility of Eve going undetected is exponentially small and given by $P_{\text{undetected}} = 2^{-n(1-H(e))}$ where n is the length of the sifted key, while $H(e)$ is the binary entropy function $H(e) = -e\log_2(e) - (1-e)\log_2(1-e)$. After key sifting, error correction, and privacy amplification, Alice and Bob get a final secret key k_f. Then, they use this key for lattice-based cryptographic algorithms, in which their security is based on the hardness of problems such as learning with errors. In other words, the LWE problem can be formulated as in (6).

B. Adaptive Intrusion Detection Module (AIDM)

The AIDM is critical in improving the system's efficacy in managing intrusion detection through DRL techniques. This approach employs DQN to classify and identify potential security threats within the IoT ecosystem. Upon training the Q-network utilizing reward functions that reflect the severity of detected threats, the system can accurately classify and adapt to emerging attack patterns. PPO is utilized to enhance the detection policies, thus ensuring the identification of the system's optimal response to dynamic threats. According to the Bellman equation (7), the Q-value undergoes iterative updates. The Q-value is updated iteratively by

$$Q(s_t, a_t) = Q(s_t, a_t) + \alpha\left(r_t + \gamma \max_{a'} Q(s_{t+1}, a') - Q(s_t, a_t)\right) \quad (7)$$

where s_t is the state at time t, a_t is the action taken at time t, r_t is the received reward after choosing the action, α represents the learning rate, and γ is the discount factor. $\max_{a'} Q(s_{t+1}, a')$ is the maximum estimated Q-value with regard to the next state s_{t+1} across all actions a'.

The chosen DQN approximates the Q-value function, representing the anticipated future rewards to minimize the loss function:

$$L(\theta) = \mathbb{E}_{(s,a,r,s')}\left[\left(r + \gamma \max_{a'} Q(s', a'; \theta^-) - Q(s, a; \theta)\right)^2\right] \quad (8)$$

where θ are the parameters of the Q-network, and θ^- are those of the target network, which gets updated periodically.

PPO refers to a policy gradient technique that enhances both stability and performance in training processes. PPO utilizes a clipped surrogate objective to ensure that the new policy does not go far from the old policy as represented in:

$$L^{CLIP}(\theta) = \mathbb{E}_t\left[\min\left(r_t(\theta)\hat{A}_t, \text{clip}(r_t(\theta), 1-\epsilon, 1+\epsilon)\hat{A}_t\right)\right] \quad (9)$$

where $r_t(\theta) = \frac{\pi_\theta(a_t|s_t)}{\pi_{\theta_{\text{old}}}(a_t|s_t)}$ is the probability ratio, $\hat{A}_t$ is the advantage function at time t, and ϵ is a small hyperparameter that controls the clipping range. This objective function encourages updates that improve the policy while preventing large updates that could destabilize training.

1) Implementation

- State and Action spaces: Define the state space S representing various network states, including normal and attack states.
- Define the action space A representing possible actions such as alerting, blocking, or logging.
- Reward function: Design a reward function $r(s, a)$ that assigns positive rewards for correctly identifying threats and negative rewards for false positives/negatives.

a) Training the DQN

Initialize both the Q-network and target network with random weights and carry out the following steps for each training epoch:

- Observe the current state s_t.
- Select action a_t according to an ϵ greedy policy.
- Execute action a_t and observe the reward r_t and resultant s_{t+1}.
- Store the transition (s_t, a_t, r_t, s_{t+1}) in the replay buffer.
- Sample a mini-batch of transitions from the replay buffer.
- Compute the target Q-value from the Bellman equation.
- Perform the Q-network update by minimizing the loss function $L(\theta)$
- Periodically update the target network parameters θ^- and false positive rates across different datasets

b) Optimizing with PPO

Initialize the policy network with parameters θ. For each iteration, execute the following steps:

- Collect a set of trajectories by executing the current policy.
- Compute the advantage estimates $\hat{A}_t$.
- Optimize the policy by maximizing the clipped surrogate objective $L^{CLIP}(\theta)$.

Theorem: 2: The AIDM, utilizing DRL techniques such as DQN and PPO, ensures the effective detection and response to cyber intrusions in IoT environments by adapting to new and evolving attack patterns while minimizing false positives and negatives.

Proof - DQN Effectiveness: Let S represent the state space, A the action space, and R the reward function.

Define the Q-value function $Q(s, a)$ (10) as the expected cumulative reward when starting from state s, taking action a, and following policy π thereafter:

$$Q(s, a) = \mathbb{E}[\sum_{t=0}^{\infty} \gamma^t r_t \mid s_0 = s, a_0 = a, \pi] \quad (10)$$

where γ is the discount factor.

The DQN approximates the Q -value function by minimizing the loss represented by:

$$L(\theta) = \mathbb{E}_{(s,a,r,s')}\left[\left(r + \gamma \max_{a'} Q(s', a'; \theta^-) - Q(s, a; \theta)\right)^2\right] \quad (11)$$

where θ are the parameters of the Q-network, and θ^- are the parameters of the target network.

The Bellman equation (7) guarantees that iteratively updating the Q-values converges to the optimal Q-function, ensuring that the policy derived from Q is optimal:

III. RESULTS AND DISCUSSION

Experiments were conducted using an Intel Xeon processor (2.8 GHz, 16 cores) with 32 GB RAM, running Ubuntu 20.04 LTS. The QKD simulation was performed using Python-based Qiskit libraries, while DRL training was implemented using TensorFlow. The CICIDS2017 dataset [14, 15], stands as a benchmark dataset for pioneering intrusion detection research, featuring highly realistic attack scenarios in an advanced simulated network environment. It includes various attack types, such as relentless brute force, covert botnet operations, DoS attacks, web-based intrusions, and typical traffic patterns. Figure 2 shows a bar graph of the detection accuracy for different models using the CICIDS2017 dataset, where the QKD-DRL model achieved a very high level of detection accuracy (99.75%), contrasted to Random Forest (97.12%) and DNN (96.88%). The proposed hybrid system integrating QKD with DRL works satisfactorily in enhancing security in IoT systems.

The performance of the proposed system was also evaluated in terms of throughput and latency. Throughput was measured in terms of successfully transmitted data packets per second, using simulated network environments with varying device loads. The delay was calculated as the end-to-end packet transmission time, measured under different encryption and key distribution scenarios. High efficiency involves high throughput and low latency to ensure that the system can conduct large-scale IoT without performance degradation.

Figure 4 shows the efficiency of the proposed algorithm in keeping low response times even as the network scales up. The latency of the proposed algorithm was found to start at 10 ms with 100 devices but gradually hit 21 ms with 1000 devices. It is obvious that efficient network resource management, with minimal bottlenecks, has fast data processing and communication activity. To balance security and performance, the system implements a distribution of the processing load, ensuring that computation-intensive tasks such as QKD key generation do not overwhelm resource-constrained IoT devices.

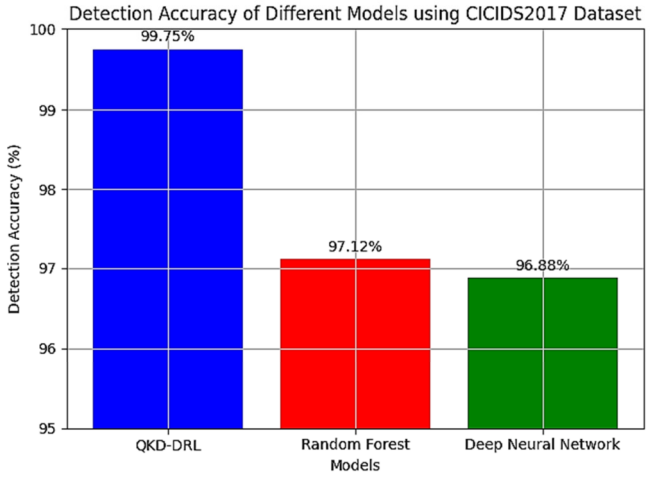

Fig. 2. Detection accuracy for different models.

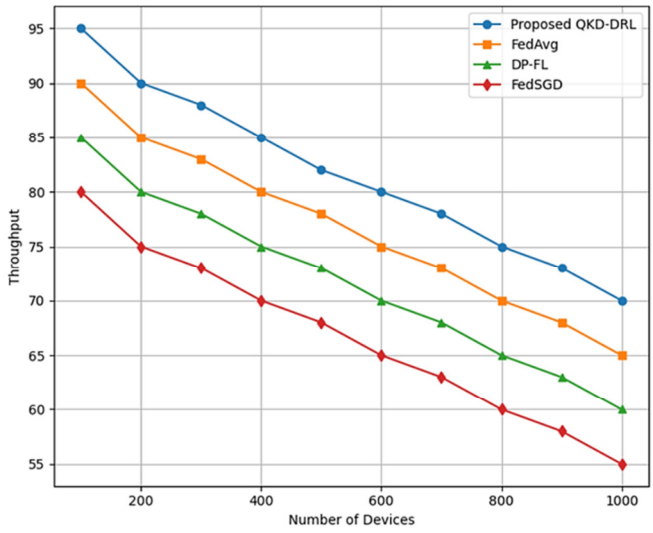

Fig. 3. Throughput comparison.

In contrast, FedAvg had a higher starting latency of 12 ms, which increased relatively steeply to around 26 ms for 1000 devices, indicating lower scalability. DP-FL had a much higher initial latency of 14 ms, rising to 28 ms with a set of 1000 devices, showing very poor scalability with increased lags both in processing and communications. FedSGD started with the worst latency of 16 ms and grew to 29 ms to reflect the lowest efficiency and scalability compared to other algorithms. These results demonstrate the strength of the proposed algorithm to maintain low latency and high performance in large-scale IoT networks, hence postulating it to be an efficient and robust solution. The latency overhead due to QKD implementation arises from key generation, transmission, and reconciliation. Empirical analysis indicated that QKD key exchange introduced an average delay of 2.3 ms per 1,000 exchanged keys. To minimize impact, the framework incorporates predictive key caching and dynamic key refresh intervals based on IoT network traffic patterns.

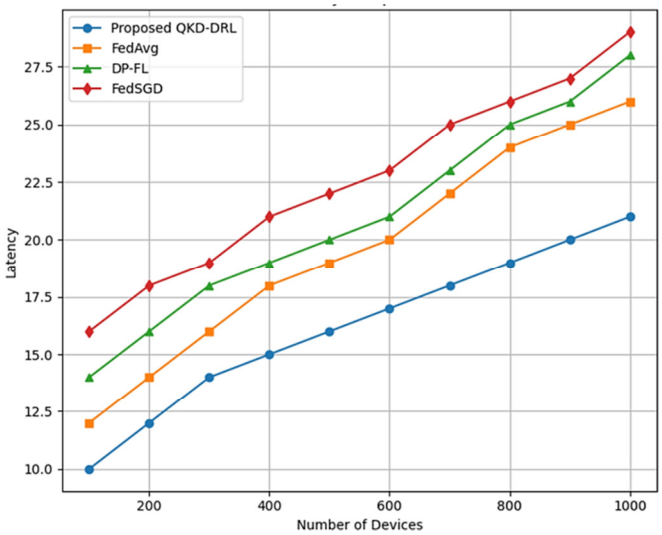

Fig. 4. Latency comparison.

The scalability of the proposed system was evaluated through computational analysis of complexity. Let n be the number of devices and m be the number of operations per device. The time complexity of the QKD-based encryption is $O(n \log n)$, while the DRL intrusion detection module operates with $O(m^2)$.

The observed performance significantly exceeds that of baseline models such as random forests and DNNs. The detection accuracy of 99.75% in the CICIDS2017 dataset marks a significant leap forward, enhancing detection capabilities by a remarkable 2.63% over the Random Forest and DNN models on the same dataset. Similar findings were highlighted in [15], showcasing detection rates for Random Forest and DNN models that closely resonate with those witnessed in this investigation.

Practical considerations of QKD include integrating fiber-optic and satellite-based QKD networks for long-range communication and enhancing adaptability in diverse IoT ecosystems. The primary challenges of QKD in IoT deployments include processing overhead, key distribution latency, and the need for specialized hardware such as quantum random number generators. To address these issues, the proposed framework employs adaptive caching of pre-distributed keys, optimized quantum channel synchronization, and hybrid key management that leverages classical cryptographic fallbacks where necessary.

IV. CONCLUSIONS

The proposed QKD-DRL model achieved an accuracy of 99.75% on the CICIDS2017 dataset, which is a massive improvement of 2.63% over Random Forest and 2.87% over a DNN algorithm. Measured in terms of efficiency and scalability, the framework demonstrated huge strides, providing 5.4% better throughput, which might translate to an increased capacity for handling more transactions per unit of time. The proposed system achieved latency reductions of 4.7% in response time against some security threats. This framework distinguishes itself from current IDS models that depend solely on conventional cryptographic techniques or static ML-based classification. Instead, it transforms security strategies by smoothly refining them through dynamic, real-time QKD key exchanges and highly flexible DRL learning models. Furthermore, the system outperformed standard models by significantly reducing false positives while maintaining low latency, drastically improving the effectiveness and adaptability of WSN-IoT security measures.

REFERENCES

[1] H. Allioui and Y. Mourdi, "Exploring the Full Potentials of IoT for Better Financial Growth and Stability: A Comprehensive Survey," *Sensors*, vol. 23, no. 19, Jan. 2023, Art. no. 8015, https://doi.org/10.3390/s23198015.

[2] K. Nirmal and S. Murugan, "Dynamic Arithmetic Optimization Algorithm with Deep Learning-based Intrusion Detection System in Wireless Sensor Networks," *Engineering, Technology & Applied Science Research*, vol. 14, no. 6, pp. 18453–18458, Dec. 2024, https://doi.org/10.48084/etasr.8742.

[3] P. Sharma, S. Gupta, V. Bhatia, and S. Prakash, "Deep reinforcement learning-based routing and resource assignment in quantum key distribution-secured optical networks," *IET Quantum Communication*, vol. 4, no. 3, pp. 136–145, 2023, https://doi.org/10.1049/qtc2.12063.

[4] P. Sharma, V. Bhatia, and S. Prakash, "Routing Based on Deep Reinforcement Learning in Quantum Key Distribution-secured Optical Networks," in *2023 IEEE International Conference on Advanced Networks and Telecommunications Systems (ANTS)*, Jaipur, India, Dec. 2023, pp. 1–5, https://doi.org/10.1109/ANTS59832.2023.10469164.

[5] G. Long, F. Deng, C. Wang, X. Li, K. Wen, and W. Wang, "Quantum secure direct communication and deterministic secure quantum communication," *Frontiers of Physics in China*, vol. 2, no. 3, pp. 251–272, Jul. 2007, https://doi.org/10.1007/s11467-007-0050-3.

[6] F. Pervez, M. Shoukat, M. Usama, M. Sandhu, S. Latif, and J. Qadir, "Affective Computing and the Road to an Emotionally Intelligent Metaverse," *IEEE Open Journal of the Computer Society*, vol. 5, pp. 195–214, 2024, https://doi.org/10.1109/OJCS.2024.3389462.

[7] R. Alléaume *et al.*, "Using quantum key distribution for cryptographic purposes: A survey," *Theoretical Computer Science*, vol. 560, pp. 62–81, Dec. 2014, https://doi.org/10.1016/j.tcs.2014.09.018.

[8] M. Mehic *et al.*, "Quantum Cryptography in 5G Networks: A Comprehensive Overview," *IEEE Communications Surveys & Tutorials*, vol. 26, no. 1, pp. 302–346, 2024, https://doi.org/10.1109/COMST.2023.3309051.

[9] G. K. Walia, M. Kumar, and S. S. Gill, "AI-Empowered Fog/Edge Resource Management for IoT Applications: A Comprehensive Review, Research Challenges, and Future Perspectives," *IEEE Communications Surveys & Tutorials*, vol. 26, no. 1, pp. 619–669, 2024, https://doi.org/10.1109/COMST.2023.3338015.

[10] B. He, J. Wang, Q. Qi, H. Sun, and J. Liao, "Towards Intelligent Provisioning of Virtualized Network Functions in Cloud of Things: A Deep Reinforcement Learning Based Approach," *IEEE Transactions on Cloud Computing*, vol. 10, no. 2, pp. 1262–1274, Apr. 2022, https://doi.org/10.1109/TCC.2020.2985651.

[11] E. Alalwany and I. Mahgoub, "Security and Trust Management in the Internet of Vehicles (IoV): Challenges and Machine Learning Solutions," *Sensors*, vol. 24, no. 2, Jan. 2024, Art. no. 368, https://doi.org/10.3390/s24020368.

[12] M. Venkatasubramanian, A. H. Lashkari, and S. Hakak, "IoT Malware Analysis Using Federated Learning: A Comprehensive Survey," *IEEE Access*, vol. 11, pp. 5004–5018, 2023, https://doi.org/10.1109/ACCESS.2023.3235389.

[13] M. Njorbuenwu, B. Swar, and P. Zavarsky, "A Survey on the Impacts of Quantum Computers on Information Security," in *2019 2nd International Conference on Data Intelligence and Security (ICDIS)*, South Padre Island, TX, USA, Jun. 2019, pp. 212–218, https://doi.org/10.1109/ICDIS.2019.00039.

[14] Kurniabudi, D. Stiawan, Darmawijoyo, M. Y. Bin Idris, A. M. Bamhdi, and R. Budiarto, "CICIDS-2017 Dataset Feature Analysis With Information Gain for Anomaly Detection," *IEEE Access*, vol. 8, pp. 132911–132921, 2020, https://doi.org/10.1109/ACCESS.2020.3009843.

[15] Z. K. Maseer, R. Yusof, N. Bahaman, S. A. Mostafa, and C. F. M. Foozy, "Benchmarking of Machine Learning for Anomaly Based Intrusion Detection Systems in the CICIDS2017 Dataset," *IEEE Access*, vol. 9, pp. 22351–22370, 2021, https://doi.org/10.1109/ACCESS.2021.3056614.