qRL: Reinforcement Learning Routing for Quantum Entanglement Networks

1st Diego Abreu
Federal University of Pará (UFPA)

2nd Antonio Abelém
Federal University of Pará (UFPA)

Abstract—**Quantum Internet aims to enable quantum communication between any two points, offering applications such as quantum key distribution (QKD), distributed quantum computing, and entanglement networks. However, the current quantum technology presents challenges with low entanglement (EPR pairs) generation rates, limited quantum memory capacity, and decoherence rates that often lead to unusable EPR pairs due to low fidelity. This presents a significant challenge for tasks such as routing. In this paper, we tackle this challenge by introducing qRL, a quantum-aware routing protocol that utilizes reinforcement learning to optimize quantum routing decisions. We show that qRL consistently outperforms traditional methods by maintaining higher fidelity routes and request success rates in different network configuration scenarios.**

Index Terms—**Quantum Network, Routing, Reinforcement Learning**

I. INTRODUCTION

Quantum Internet is a rapidly growing research area that aims to transform information transmission and processing, making it more secure [1]. By enabling quantum communication between multiple network nodes, the Quantum Internet offers promising applications, such as Quantum Key Distribution (QKD), distributed quantum computing [2], and the creation of entanglement networks [3]. These applications have the potential to transform communication security and information processing capacity, playing a key role in the next generation of communication and computing technologies [4].

However, the practical realization of the Quantum Internet faces significant challenges. The currently available quantum technology has limitations related to the rates of generation of entanglements (EPR pairs - Einstein-Podolsky-Rosen pairs), the limited capacity of qubits (quantum bits) available at each node, and quantum decoherence that often results in low-quality EPR pairs and qubits, due to the loss of coherence of quantum states, fidelity. These limitations directly impact the effectiveness of operations carried out on quantum networks, particularly in information routing [5].

At the network layer, the routing of quantum information in quantum entanglement networks represents a major challenge [6]–[8]. Unlike classical (non-quantum) networks, quantum information encoded in the qubit cannot be copied and forwarded (due to the quantum no-cloning theorem) and depends on specific quantum communication operations to be transmitted. Thus, quantum information must travel complex routes while preserving fidelity, consuming as few resources as possible. Efficient and reliable information routing is an essential component for the effective functioning of any communication network, including quantum networks.

To address these challenges, this paper presents the qRL protocol, a quantum routing protocol that utilizes reinforcement learning to optimize routing decisions in quantum entanglement networks. It will be demonstrated that qRL consistently outperforms traditional routing methods, maintaining high-fidelity routes in different network configuration scenarios. This approach represents a significant advancement in the search for effective solutions for building the Quantum Internet. In this context, the main contributions consist of:

- Development of qRL, a reinforcement learning approach for routing in quantum entanglement networks.
- Analysis of qRL in terms of route quality and request success rate, with four network configurations.
- Comparison of qRL with other routing protocols for quantum entanglement networks available in the state of the art.

The remainder of this paper is structured as follows: In Section II, we establish the theoretical background by exploring routing in quantum entanglement networks. Section III provides a comprehensive overview of related works in the field, setting the context for our proposed solution. In Section IV, we delve into the details of our proposal, qRL. Section V focuses on a case study, presenting details about the dataset, the proposed evaluation methodology, and providing context for the experiments conducted. Following this, Section VI presents the results obtained from our experiments. Finally, Section VII serves as the conclusion, and outlines potential directions for future research.

II. ROUTING IN QUANTUM ENTANGLEMENT NETWORKS

Quantum communication through entanglement has a variety of applications, including quantum teleportation [1], cryptography [9], sensing [10], and blind quantum computing [11]. These applications rely on the generation and distribution of EPR pairs via quantum channels, capable of transmitting quantum states such as qubits over telecommunications fibers or free space. EPR pairs are considered the primary resource of a quantum network, generated typically through heralding entanglement generation protocols [12]. To generate an EPR pair between two nodes, the emitting node entangles a memory qubit with a transmission qubit and sends it through an existing quantum channel to the receiving node, where it's entangled with the receiver's memory qubit, creating an EPR

pair between the nodes. This operation's outcome is announced over a classical channel, such as the Internet, indicating the creation of the EPR pair. This process has a probability of failure, introducing challenges to the reliability of quantum communications.

Entanglement Swapping (ES) extends entanglement across the network by using intermediate nodes as repeaters, generating direct EPR pairs between distant nodes through existing EPR connections, and creating an end-to-end entanglement (E2E). This process, essential for long-distance quantum communication, involves entangling qubits at repeater nodes, effectively linking distant nodes without direct quantum interaction. In a nested ES process, the E2E rate is given by $E2E\lambda = \lambda \cdot q^{\log N}$, where λ is the link level EPR generation rate, q is the link swap success probability and N is the number of repeaters in the route. The ES is probabilistic, and is also prone to noise and quantum decoherence, leading to state fidelity degradation of the final EPR pair. In contrast, quantum purification, or distillation, produces higher fidelity EPR states using multiple lower fidelity pairs. The resulting fidelity F after one purification round is given by: $F(f_1, f_2) = \frac{f_1 \cdot f_2}{f_1 \cdot f_2 + (1-f_1) \cdot (1-f_2)}$, where $f1$ and $f2$ are the original EPRs fidelity's.

Successful route establishment in a quantum network requires nodes to have available qubits for participating in entanglement swapping and quantum teleportation. Additionally, the route must have at least one EPR pair available across all channels, with the creation of EPR pairs requiring a qubit from each involved node. These requirements highlight the unique challenges and considerations in routing within quantum entanglement networks, emphasizing the network's reliance on the foundational quantum mechanical principles to achieve communication objectives. For a single request, the goal of routing can be to find an optimal path or multiple paths. The ideal definition of the optimal path is the path that has the largest number of end-to-end entangled states in a time slot, the highest fidelity, and the lowest consumption of entangled resources.

However, in dynamic and complex networks, this ideal is hard to achieve due to the variability in link capacities, physical lengths, qualities, and the fluctuating availability of quantum memory and EPR pairs. In such environments, identifying the optimal path becomes a significantly more challenging task, requiring sophisticated routing algorithms that can adapt to the changing conditions of the quantum network. These algorithms must consider a multitude of factors, including the probabilistic nature of quantum entanglement, decoherence rates, and the necessity for purification processes, to ensure the highest possible fidelity and throughput while minimizing resource consumption. The complexity of these networks requires a flexible and dynamic approach to routing, where the definition of optimal is constantly evolving in response to the network's state. Thus, in this work, we are tackling this task by proposing a novel routing framework that is capable of dynamically adapting to the changing landscape of quantum networks, aiming to optimize the path selection process in real-time based on current network conditions and the specific requirements of quantum communication tasks.

III. RELATED WORK

Various research studies investigate the development and current trends in quantum networks from different perspectives [13]–[16]. This section presents related works that address routing in quantum entanglement networks. The study by Van Meter et al. (2013) [17] is among the first to discuss quantum routing, proposing qDijkstra, an adaptation of the classic Dijkstra algorithm for quantum networks. In qDijkstra, routes are selected based on the number of hops between nodes, considering requirements such as available qubits at each node on the route and EPR pairs in each channel. In our work, qDijkstra will be compared to our proposal, qRL, which considers other metrics to choose the best route.

In this context, various metrics have been used to guide routing decisions. Several studies consider channel fidelity as an important metric in routing strategy [18]–[20]. The availability of EPR pairs is also used as a route quality metric [8], [21], [22]. Regarding the availability of memory qubits at each node, various perspectives have been adopted. For instance, Patil et al. (2021) [23] considers an infinite availability of memory qubits in their routing protocol. Other works highlight the importance of memory as a critical factor to consider [24]. Moreover, Miguel-Ramiro et al. (2023) [25] propose a routing strategy that seeks to reduce memory qubit consumption, taking into account decoherence effects. Conversely, our proposal considers fidelity, the availability of EPR pairs, and memory availability in choosing the best route.

Machine Learning methods have also been applied in the field of quantum entanglement networks. Le et al. 2022 [26] propose a routing protocol based on reinforcement learning. However, while Le et al. (2022) proposal focuses on allocating as many requests as possible simultaneously, our proposal, qRL, aims to identify the most efficient route for each individual entanglement request. Furthermore, in the study by Le et al. (2022), the Dijkstra algorithm is used to select routes, primarily based on the number of hops between nodes, without considering route quality in terms of fidelity. Thus, the proposed model only takes into account the qubit capacity of each node and the availability of EPR pairs, without incorporating channel fidelity or the impact of decoherence. This approach results in a model that, while useful, does not fully reflect the complexities of a real quantum network, where channel fidelity and decoherence are critical factors.

IV. QRL: REINFROCMENT LEARNING ROUTING FOR QUANTUM ENTANGLEMENT NETWORKS

qRL leverages an adaptation of the Q-Learning algorithm [27], a reinforcement learning technique [28], to optimize routing in quantum networks. This method allows the protocol to make informed decisions based on accumulated experience, without the need for a pre-trained model. Q-Learning's suitability for dynamic environments, such as quantum networks, makes it an ideal choice for qRL, which comprises the following components:

- **Agent**: The core of the reinforcement learning system, responsible for making decisions based on data received from the environment. In qRL, the agent uses information about the current state of the network to select the most appropriate action for routing optimization.
- **Environment**: Represents the external system with which the agent interacts, in this case, the quantum network itself, including all nodes, communication channels, available EPR pairs, and channel fidelity. The environment updates the agent with state changes and rewards following each action taken.
- **State (S)**: The current configuration of the network, including channel fidelity, the number of available EPR pairs, and qubit availability at each node.
- **Actions (A)**: Actions in qRL include choosing different routes and allocating EPR pairs to specific links.
- **Reward (R)**: Feedback for learning, based on the effectiveness of the taken action, such as route fidelity improvement or efficient use of network resources (fewer qubits and EPR pairs used).
- **Policy (π)**: The strategy the agent follows to choose the next action based on the current state.

The goal of qRL is to train the agent to select communication routes that optimize qubit transmission, considering metrics such as fidelity, EPR pair availability, and qubit availability at network nodes. The functioning of qRL operates as follows:

- Agent Initialization: The agent is initialized with parameters such as learning rates (α), discount factor (γ), and the probability of selecting random actions (ε). It also receives information about the network topology, represented by nodes and channels.
- State Definition: The agent constructs a representation of the current state of the network, considering the availability of EPR pairs and qubits in channels and nodes. This representation is used as input for learning.
- Action Selection: Based on the current state, the agent selects a routing action. It uses an ε-greedy policy, meaning that with probability (ε), a random action is chosen, and with a probability of 1 - (ε), the action that maximizes the Q value is chosen.
- Q-Value Update: After action selection and route execution, the agent receives a reward based on the quality of the chosen route. The agent's Q values are updated using a learning process, where the targets are the expected rewards. The Q values are updated using the formula:

$$Q(s,a) = Q(s,a) + \alpha \left[R(s,a) + \gamma \max_{a'} Q(s',a') - Q(s,a) \right] \tag{1}$$

Where:
- $Q(s,a)$ is the estimated Q value for the state-action pair.
- α is the learning rate.
- $R(s,a)$ is the received reward.
- γ is the discount factor.

- Qubit Routing: The selected route is used to efficiently route a qubit through the network. The availability of EPR pairs and qubits in channels and nodes is considered. Depending on conditions, the creation of new EPR pairs and the allocation of qubits on the route may be required.
- Reward Calculation: The reward is calculated based on the fidelity of the qubit transmission and the success of the routing. The higher the fidelity and success, the greater the reward.

V. SYSTEM MODEL AND EXPERIMENTS

To evaluate our proposal, the quantum network was modeled using discrete simulation through graph representation. The lattice topology was adopted for the case study, as presented in Figure 1. Formally, this network can be represented as $G = (V, E)$, where V represents the set of nodes that embody the quantum nodes. Each quantum node is capable of performing fundamental quantum communication operations, such as teleportation and entanglement swapping, enabling it to send and receive quantum information in the form of qubits. Conversely, E represents the set of edges that denote the quantum communication links between these nodes. The node capacity indicates the number of qubits stored in a node's quantum memory at any given time and can be used for teleportation tasks, entanglement generation, and swapping. The qubit decoherence rate refers to the rate at which qubits lose their quantum properties due to interaction with the environment.
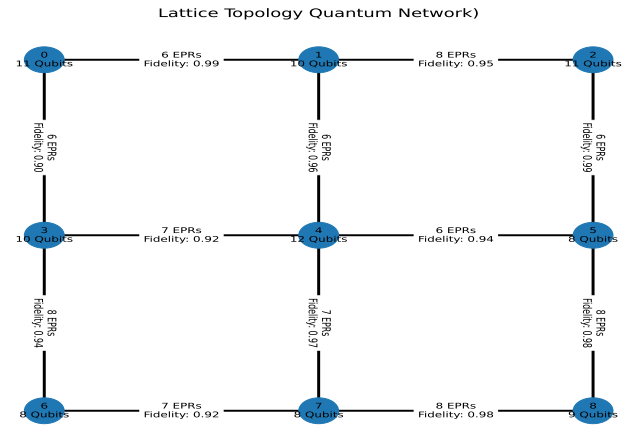

Fig. 1: Quantum Network in the Lattice Topology.

In our study, we explored four distinct network configurations, as detailed in Table I, where Mi is the initial memory available at the nodes, Fi is the initial fidelity of the created eprs, $EPRi$ is the amount of EPR initially available at each link, λ is the rate of attempts of EPR creation at each time slot at each link, $EPRp$ is the EPR creation success probability, ESp is the entanglement swap success probability, $Purp$ is the purification success probability. The configurations differ primarily in the memory size of the nodes, which is expressed as a range related to the network size (n), indicating the node's

capacity to store qubits. The other parameters, including initial fidelity, the initial quantity of EPR pairs per channel, EPR success rate, and the success rate of Entanglement Swapping (ES), have predefined initial values within specified ranges for each configuration. Specifically, Configurations 1 and 3 are designed with higher success rates for EPR generation and ES, indicative of more robust network setups, whereas Configurations 2 and 4 are characterized by lower success rates, simulating networks under more challenging conditions. To ensure a consistent basis for comparison across different network scenarios, we fixed the decoherence rate at 0.005 at each time slot. It's worth noting that time in our simulation is segmented into discrete slots, allowing for a granular examination of network dynamics over time.

TABLE I: Quantum Network Configuration Parameters.

Param	Config 1	Config 2	Config 3	Config 4
Mi	$[1.5n, 2.5n]$	$[1.5n, 2.5n]$	$[n, 1.5n]$	$[n, 1.5n]$
Fi	$[0.99, 0.95]$	$[0.85, 0.90]$	$[0.99, 0.95]$	$[0.85, 0.9]$
$EPRi$	$[7, 10]$	$[5, 7]$	$[7, 10]$	$[5, 7]$
λ	$[2, 4]$	$[1, 2]$	$[2, 4]$	$[1, 2]$
$EPRp$	$[0.95, 0.9]$	$[0.75, 0.6]$	$[0.95, 0.9]$	$[0.75, 0.6]$
ESp	$[0.95, 0.9]$	$[0.75, 0.6]$	$[0.95, 0.9]$	$[0.75, 0.6]$
$Purp$	$[0.95, 0.9]$	$[0.75, 0.6]$	$[0.95, 0.9]$	$[0.75, 0.6]$

We compare qRL with four other routing strategies. The first strategy is qDijkstra, as described by Van Meter et al. (2013) [17]. The second, named R1: fidelity, focuses on the fidelity of routes, as presented in Li et al. (2021) [18]. The third, R2: EPR, prioritizes the quantity of available EPR pairs, as discussed in Chakraborty et al. (2019) [8]. Lastly, the strategy R3: qubits considers the availability of qubits, as explored by [25].

In the case study, simulations are conducted with multiple routing requests, each with a distinct origin and destination node. The protocol under analysis evaluates the best route for each request. For all simulations, the values of α, γ and ε were empirically adjusted to 0.2, 0.8, and 0.8, respectively. This method provides a comprehensive assessment of qRL's performance compared to existing routing strategies, across a variety of network scenarios.

VI. RESULTS AND DISCUSSION

This section details the outcomes of our experiments, wherein we compare qRL with four distinct routing strategies. We focus on highlighting the quality of end-to-end fidelity routes and the success rate of requests.

A. End-to-End Fidelity Route Quality Results

The experiments results for the four configurations are illustrated in Figure 2. The graph depicts the average end-to-end route fidelity across various network sizes, with the number of nodes (n) ranging from 4 to 64. For this experiment, the number of request was fixed in 1000 requests. In each configuration, the performance of qRL is compared to four other routing strategies. In this scenario, each network node has different initial values for node capacity and decoherence rate, and each network channel has varying values for fidelity,

the quantity of EPR pairs, and purification, as outlined in Table I (Section V).

(a) Config1   (b) Config2
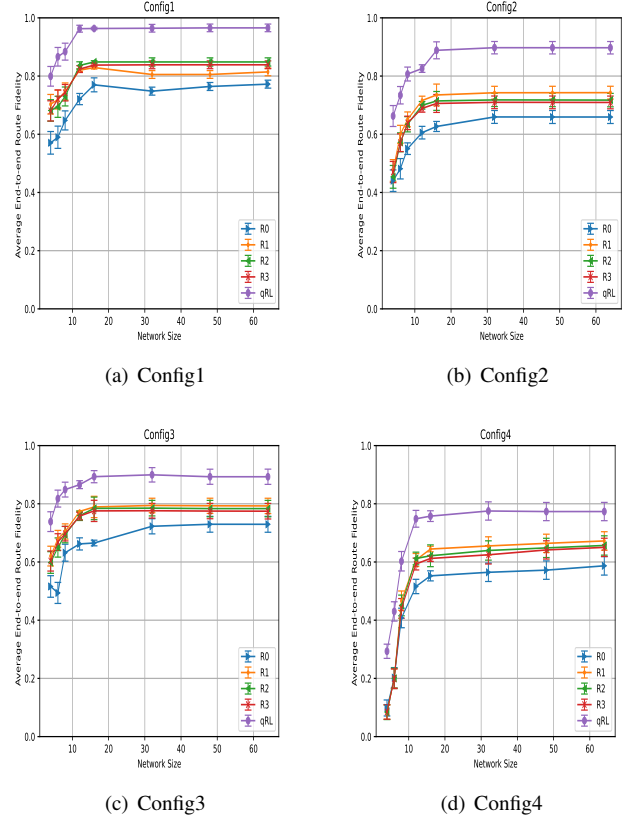
(c) Config3   (d) Config4

Fig. 2: Average End-to-End Fidelity at each quantum network configuration.

The evaluation of end-to-end fidelity across various configurations illustrates qRL's dominance, particularly against traditional strategies which yield similar results to each other. Notably, qDijkstra underperforms, while R3: qubits exhibits more effective results, only outmatched by qRL. Among the configurations, Config1 stands out with the highest fidelity, descending through Config2, Config3, to Config4, which shows the lowest. This trend reveals that average fidelity is typically subdued in smaller networks but escalates with the network's expansion, stabilizing beyond 30 nodes due to enhanced redundancy and alternative pathways that uphold fidelity despite potential isolated failures. In large-scale networks, qRL leverages the abundant data to refine and optimize routes. Conversely, small networks grapple with limited routing alternatives and resources, leading to higher susceptibility to failures and diminished route fidelity.

The parameters Mi and Fi significantly influence the network's initial state, with higher ranges of Mi and higher fidelity values (Fi) providing a robust starting point for qRL to optimize routes effectively. For instance, configurations with higher Mi and Fi values typically facilitate better performance due to the increased likelihood of high-fidelity connections between nodes from the outset. Conversely, lower

values in these parameters challenge qRL to find efficient paths through a more constrained network, impacting the overall success rate and fidelity of the routes. The $EPRi$, λ, $EPRp$, ESp, and $Purp$ parameters directly affect the dynamic capabilities of the network to sustain and enhance quantum connections over time. A higher initial amount of EPR pairs ($EPRi$) and higher success probabilities for EPR creation ($EPRp$), entanglement swaps (ESp), and purification ($Purp$) enable the network to more readily adapt to and recover from degradations in fidelity, thereby maintaining higher overall route quality. The rate of EPR creation attempts (λ) plays a critical role in replenishing and expanding the network's quantum resources, thus facilitating continuous improvement in routing efficiency and fidelity as the network evolves.

B. Requests Success Results

The presented graphs illustrated in Figure 3 show the Request Success Rate for four different configurations in our quantum network experiment, with the number of requests ranging from 10 to 3000. For this experiment, the network size was fixed in 30 nodes. Across all configurations, qRL consistently maintains the highest success rates, suggesting its superior adaptability and efficiency in handling quantum communication requests. In Config1 and Config2, qRL begins with success rates comparable to other strategies but exhibits a slower decline as the number of requests increases, indicating its robustness in maintaining network performance under load. Interestingly, in Config3 and Config4, qRL showcases a noticeable advantage over other strategies from the outset, maintaining higher success rates throughout and demonstrating its effectiveness in more demanding network conditions. In contrast, traditional routing strategies R0, R1, R2, and R3 show varying degrees of decline in success rates as the number of requests grows, with R0 performing the worst in all configurations. This decline is particularly steep in Config3 and Config4, where the network's increasing request load appears to challenge these strategies' capacity to sustain high success rates. The results also highlight the importance of network configuration in determining the performance of routing strategies. Config1, with its initially high success rates across all strategies, suggests a network setup that is more suited to handle a higher volume of requests. As the configurations progress to Config4, the more pronounced differentiation in the success rates of qRL compared to other strategies emphasizes qRL's strength in managing more complex and request-heavy quantum network environments.

VII. CONCLUSION AND FUTURE WORKS

The quantum entanglement networks still face several challenges in their practical implementation, especially concerning routing in complex networks. This work introduced qRL, a reinforcement learning-based routing protocol for quantum entanglement networks, addressing critical challenges in building the Quantum Internet. This protocol was designed to optimize routing decisions by leveraging a reinforcement

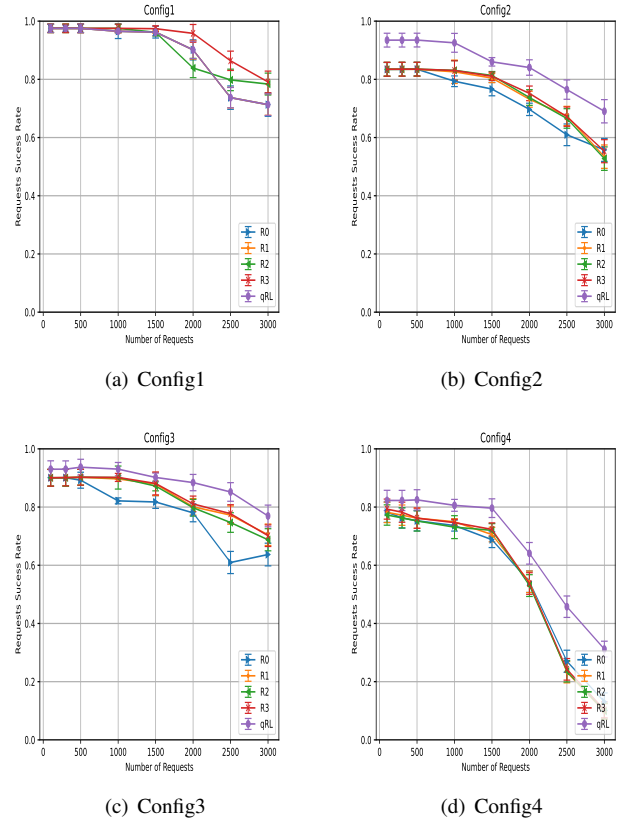(a) Config1   (b) Config2

(c) Config3   (d) Config4

Fig. 3: Average request success rate at each quantum network configuration.

learning approach and demonstrated to be an effective solution, consistently outperforming traditional routing methods. Through tests in four distinct configuration scenarios, qRL proved its versatility and efficiency, adapting and learning effectively in diverse network environments while maintaining high-fidelity routes and high request success rates in different network configurations.

For future works, an area of interest is the integration of learning mechanisms that can quickly adapt to changes in network topology and variable conditions of qubits and channels. Additionally, exploring the application of qRL in large-scale network scenarios and its interaction with classical communication protocols may provide insights for the practical implementation of the Quantum Internet. Another promising direction involves investigating strategies to minimize resource consumption, such as bandwidth and energy, while maintaining route fidelity, contributing to the sustainability and efficiency of future quantum networks.

REFERENCES

[1] W. Kozlowski, S. Wehner, R. Van Meter, B. Rijsman, A. Cacciapuoti, M. Caleffi, and S. Nagayama, "Rfc 9340 architectural principles for a quantum internet," *Architecture*, vol. 4, p. 4, 2023.
[2] D. Cuomo, M. Caleffi, and A. S. Cacciapuoti, "Towards a distributed quantum computing ecosystem," *IET Quantum Communication*, vol. 1, no. 1, pp. 3–8, 2020.

[3] J. Nötzel and S. DiAdamo, "Entanglement-enhanced communication networks," in *2020 IEEE International Conference on Quantum Computing and Engineering (QCE)*, 2020.

[4] A. Abelém, G. Vardoyan, and D. Towsley, "Quantum internet: The future of internetworking," in *Minicursos do XXXVIII Simpósio Brasileiro de Redes de Computadores e Sistemas Distribuídos*. SBC, dec 2020, pp. 48–90.

[5] L. Gyongyosi, S. Imre, and H. V. Nguyen, "A survey on quantum channel capacities," *IEEE Communications Surveys & Tutorials*, vol. 20, no. 2, pp. 1149–1205, 2018.

[6] L. Gyongyosi and S. Imre, "Routing space exploration for scalable routing in the quantum internet," *Scientific reports*, vol. 10, no. 1, p. 11874, 2020.

[7] M. Pant, H. Krovi, D. Towsley, L. Tassiulas, L. Jiang, P. Basu, D. Englund, and S. Guha, "Routing entanglement in the quantum internet," *npj Quantum Information*, 2019.

[8] K. Chakraborty, F. Rozpedek, A. Dahlberg, and S. Wehner, "Distributed routing in a quantum internet," *arXiv preprint arXiv:1907.11630*, 2019.

[9] V. J. Geddada and P. Lakshmi, "Distance based security using quantum entanglement: a survey," in *2022 13th International Conference on Computing Communication and Networking Technologies (ICCCNT)*. IEEE, 2022, pp. 1–4.

[10] C. L. Degen, F. Reinhard, and P. Cappellaro, "Quantum sensing," *Reviews of modern physics*, vol. 89, no. 3, p. 035002, 2017.

[11] A. Broadbent, J. Fitzsimons, and E. Kashefi, "Universal blind quantum computation," in *2009 50th annual IEEE symposium on foundations of computer science*, 2009.

[12] A. Dahlberg, M. Skrzypczyk, T. Coopmans, L. Wubben, F. Rozpedek, M. Pompili, A. Stolk, P. Pawełczak, R. Knegjens, J. de Oliveira Filho *et al.*, "A link layer protocol for quantum networks," in *Proceedings of the ACM special interest group on data communication*, 2019, pp. 159–173.

[13] J. Illiano, M. Caleffi, A. Manzalini, and A. S. Cacciapuoti, "Quantum internet protocol stack: A comprehensive survey," *Computer Networks*, vol. 213, p. 109092, 2022.

[14] S. R. Hasan, M. Z. Chowdhury, M. Saiam, and Y. M. Jang, "Quantum communication systems: Vision, protocols, applications, and challenges," *arXiv preprint arXiv:2212.13333*, 2022.

[15] A. Kumar, R. Krishnamurthi, G. Sharma, S. Jain, P. Srikanth, K. Sharma, and N. Aneja, "Revolutionizing modern networks: Advances in ai, machine learning, and blockchain for quantum satellites and uav-based communication," *arXiv preprint arXiv:2303.11753*, 2023.

[16] M. Caleffi, M. Amoretti, D. Ferrari, D. Cuomo, J. Illiano, A. Manzalini, and A. S. Cacciapuoti, "Distributed quantum computing: a survey," *arXiv preprint arXiv:2212.10609*, 2022.

[17] R. Van Meter, T. Satoh, T. D. Ladd, W. J. Munro, and K. Nemoto, "Path selection for quantum repeater networks," *Networking Science*, vol. 3, no. 1, pp. 82–95, 2013.

[18] C. Li, T. Li, Y.-X. Liu, and P. Cappellaro, "Effective routing design for remote entanglement generation on quantum networks," *npj Quantum Information*, vol. 7, no. 10, 2021.

[19] L. Gyongyosi and S. Imre, "Decentralized base-graph routing for the quantum internet," *Physical Review A*, vol. 98, no. 2, p. 022310, 2018.

[20] A. Pirker and W. Dür, "A quantum network stack and protocols for reliable entanglement-based networks," *New Journal of Physics*, vol. 21, no. 3, p. 033003, 2019.

[21] A. Patil, M. Pant, D. Englund, D. Towsley, and S. Guha, "Entanglement generation in a quantum network at distance-independent rate," *npj Quantum Information*, 2022.

[22] A. Fischer and D. Towsley, "Distributing graph states across quantum networks," in *2021 IEEE International Conference on Quantum Computing and Engineering (QCE)*, 2021.

[23] A. Patil, J. I. Jacobson, E. Van Milligen, D. Towsley, and S. Guha, "Distance-independent entanglement generation in a quantum network using space-time multiplexed greenberger–horne–zeilinger (ghz) measurements," in *2021 IEEE International Conference on Quantum Computing and Engineering (QCE)*. IEEE, 2021, pp. 334–345.

[24] P. Nain, G. Vardoyan, S. Guha, and D. Towsley, "On the analysis of a multipartite entanglement distribution switch," *Proceedings of the ACM on Measurement and Analysis of Computing Systems*, vol. 4, no. 2, pp. 1–39, 2020.

[25] J. Miguel-Ramiro, A. Pirker, and W. Dür, "Optimized quantum networks," *Quantum*, vol. 7, p. 919, 2023.

[26] L. Le and T. N. Nguyen, "Dqra: Deep quantum routing agent for entanglement routing in quantum networks," *IEEE Transactions on Quantum Engineering*, vol. 3, pp. 1–12, 2022.

[27] A. Kumar, A. Zhou, G. Tucker, and S. Levine, "Conservative q-learning for offline reinforcement learning," *Advances in Neural Information Processing Systems*, 2020.

[28] W. Qiang and Z. Zhongli, "Reinforcement learning model, algorithms and its application," in *2011 International Conference on Mechatronic Science, Electric Engineering and Computer (MEC)*. IEEE, 2011, pp. 1143–1146.