# Learning-based Route Selection in Noisy Quantum Communication Networks

Vini Chaudhary, Kai Li, Kaushik Chowdhury
*Institute for the Wireless Internet of Things at Northeastern University,* Boston, MA, USA
{vi.chaudhary, ka.li}@northeastern.edu, krc@ece.neu.edu

***Abstract*—Finding a path with the least overall noise from quantum memories, fibers, and gate operations in a quantum network involves the challenge of acquiring knowledge of these noises, their sources, and the time of occurrences. In this paper, we propose a reinforcement learning-based route selection approach that uses a multi-arm bandit algorithm to find the least noisy path from a transmitter (Tx) to a receiver (Rx), without considering any information on qubit decoherence due to probabilistic noises inherent in quantum memories and imperfect gate operations. It only uses network deployment knowledge to find a set of feasible paths from Tx to Rx. We provide a key finding from a network design perspective which says that performing entanglement swapping on nodes within a path in a non-synchronized and parallel manner not always reduces the decoherence experienced in achieving the end-to-end entanglement on that path. Further, we design and open-source a new simulator for simulating probabilistic noises encountered during entanglement distribution between Tx-Rx on a path, which has supporting callable functions for connecting the unknown network environment required for interaction with the multi-arm bandit agent. The simulation results demonstrate that our proposed route selection approach provides a path up to $\sim 33\%$ better fidelity (less noise) compared to conventional, distance-based route selection approach for the considered quantum network.**

***Index Terms*—Quantum communication network, Route selection, Entanglement, Fidelity, Multi-arm Bandit**

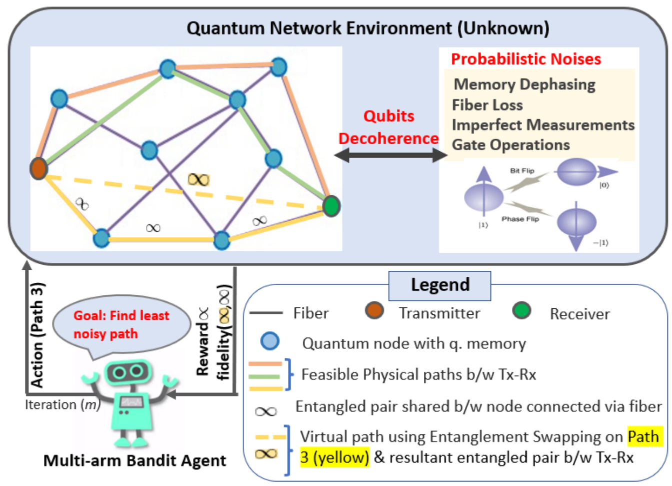



Fig. 1: An overview of our proposed approach for route selection in noisy quantum network environment is shown (unknown probabilistic noises, known deployment) using MAB algorithm. It learns the least noisy path from a set of feasible paths over multiple iterations. Paths $\{1, 2,$ and $3\}$ are shown in {orange, green, and yellow} colors in the network.

## I. Introduction

A quantum communication network is needed for realizing applications like distributed quantum computing [1], secure communication [2], and clock synchronization [3]. It enables the distribution of entangled quantum bits (qubits) among remote quantum nodes for transmission of data qubits through synergistic use of quantum and classical links [4]. However, the qubits are vulnerable to decoherence (i.e., degradation of their quantum properties) due to the noise experienced during their interaction with the quantum channel and surrounding environment. For instance, photonic qubits suffer loss in fiber quantum channel during transmission, which scales exponentially with the transmission distance. To suppress this quantum noise, quantum repeaters may be deployed for supporting long-distance entanglements using entangled pairs shared between the nodes at smaller distances via the entanglement swapping process [5]. This approach reduces the fiber-based degradation of the qubits at the cost of introducing harmful effects of noise within quantum memories and errors due to imperfect quantum gate operations within the nodes [6]. This necessitates the consideration of inevitable loss in qubits' quality in designing route selection algorithms for transmission of data qubits from transmitter (Tx) to receiver (Rx) in large-scale quantum networks.

**•Challenges and Motivations:** The idea of considering the degradation of the qubits in route selection algorithms raises several unique research challenges. Firstly, noises from quantum channels (i.e., fiber, memory, gate) are characterized by stochastic processes that vary with time and the surrounding environment [7], [8]. Secondly, these probabilistic noises are not confined to a single source, but are present in quantum memories, fibers, and operations within all the nodes constituting a transmission path. Thus, the joint modeling of the probabilistic noises from different components of a quantum network at different time instants during the end-to-end distribution of the entangled qubits on a path with multiple repeater nodes is extremely challenging and remains an open problem [9]. Lastly and most importantly, incorporating the qubits' decoherence in the network algorithm design requires prior knowledge of these probabilistic noises from all the sources. In a large-scale quantum network, this is infeasible as the unknown surrounding environment too contributes to these noises and is not always fully controllable. Most related works

[8], [10]–[13] (discussed later in Sec. II) simplify the task of finding a path by making this strong but otherwise limiting assumption. *These challenges motivate us to design a route selection approach that considers the qubits' decoherence without requiring complete knowledge of the probabilistic noises, their sources, and the time of occurrences in a quantum communication network.*

•**Proposed Solution:** To answer these challenges, we propose a learning-based approach for route selection in noisy quantum communication networks. It does not require complete prior information on the probabilistic noises inherent in the quantum memories and gate operations of the nodes during end-to-end entanglement distribution on a path in a quantum repeater network. Specifically, we assume known network deployment, i.e., the fiber connections among nodes, their lengths, and attenuation factors, to compute feasible paths from a Tx to an Rx node. We formulate the problem of selecting a high-quality (the least noisy) path from the list of feasible paths between the Tx-Rx as a multi-arm bandit (MAB) problem [14], which performs intelligent decision-making from given multiple choices. It uses an agent to interact with the quantum network (called environment) to learn the effect of probabilistic quantum noises on the fidelity of qubits on the feasible paths and find the path with the least noise, as shown in Fig. 1. We map each feasible path to an arm of a bandit and let the agent explores and exploits them over multiple iterations based on the rewards returned by them in the past iterations, with the ultimate goal of converging to a path with the highest reward. We design the reward function based on the fidelity of the entangled qubits shared between the Tx-Rx using a path. The fidelity captures the effect of noise present in the path by computing the similarity between the entangled qubits used to initiate the end-to-end entanglement distribution process over the path and the resultant end-to-end entangled qubits shared between the Tx-Rx of that path.

*To the best of our knowledge, there is no prior reinforcement learning (multi-arm bandit)-based route selection in noisy quantum networks, which enables learning of the least noisy path with minimal knowledge about the network settings.*

•**Contributions:** Our contributions are as follows:

1) We propose a route selection approach, for the transmission of data qubits from a Tx to an Rx in a noisy quantum network without any prerequisite knowledge of probabilistic noises from different quantum memories and operations. We use the MAB algorithm to find a Tx-Rx path with the least effect of these noises on the quality of entangled qubits distributed over it.
2) From a network design perspective, we found that performing entanglement swapping in a non-synchronized manner at nodes (i.e., as soon as the nodes receive two qubits, one each from the entangled pairs shared with their respective downstream and upstream links) does not necessarily minimize the decoherence experienced by the qubits, especially when a path has lesser nodes. On an average, it is better to perform swaps at the nodes in parallel and synchronize them.
3) We design a custom simulator that can simulate the decoherence of qubits due to noises in quantum memories, fibers, and imperfect gate operations in all the nodes at any time during the end-to-end entanglement distribution between a Tx-Rx in a quantum network. It is a valuable tool to create network environments unknown to a learning agent and can be accessed at [15].
4) We study the exploration-exploitation trade-off involved in the learning process. Our simulation results show that the quality (fidelity) of end-to-end entanglement achieved over a route selected by our proposed approach is $\sim 33\%$ better than that selected using state-of-the-art distance-based route selection.

## II. Related Work on Route Selection

In quantum communication networks, long-distance transmission of data qubits requires routing of entangled qubits between the end nodes. The authors in [10] propose the shortest path (distance-based) entanglement routing protocol for a square grid network topology. However, the classical distance-based routing metric often fails to fully capture the peculiarities of quantum mechanical phenomena that impact entanglement over a path, motivating us to explore quantum physics-informed routing metrics. The authors in [11] adapt the classical Dijkstra algorithm for a quantum repeater network by considering link cost as the inverse of Bell pairs per second. Likewise, the work in [8] exhaustively finds all the paths between a Tx-Rx and designs a custom end-to-end entanglement rate metric to select the best one. However, these approaches do not consider the effects of imperfect gate operations and quantum memory decoherence on qubits' fidelity. In the recent work [12], the authors propose a fidelity-based rate maximizing route selection method assuming the noise models of quantum channels (memory and fiber) are known in advance. Different from the network science approach, the work in [13] uses machine learning for allocating quantum channels to maximize the entanglement routing rate in a quantum network without addressing the impact of probabilistic noises inherent in the channels on qubits' degradation.

*To model the effect of decoherence, most of the above works unrealistically assume a priori knowledge of: 1) accurate information on network deployment, 2) noises from the quantum channels, and 3) their time of occurrences during entanglement routing, without considering the fact that the uncontrollable surrounding environment may alter these noises. In our proposed route selection approach, we relax these assumptions by using a learning-based approach.*

## III. Problem Formulation

In this section, we describe considered quantum network model and formulation of our route selection problem.

### A. Network Model

We model the quantum communication network, shown in Fig. 1, by an undirected and connected graph $G\left(V, E\right)$, where $V = \{n\}_{n=1}^{N}$ denotes the set of quantum nodes and

$E = \{i - j; \forall i, j \in V\}$ represents the set of quantum links physically connecting two nodes in the network. The connected property of the graph enables the creation of quantum links between any two distant nodes with no direct physical link. Such links are called virtual quantum links. We assume classical information can be communicated between any two nodes via the conventional Internet. Each node $v$ is equipped with quantum processors, the number of quantum memories equal the number of physical links connecting the node, and necessary hardware to perform quantum communication-related operations and storage of qubits. The physical quantum links are made of optical fibers which support the transmission of photonic qubits.

In a quantum network, a data qubit is teleported from a Tx node $t$ to an Rx node $r$ using a pair of entangled qubits $|\Psi_{t,r}\rangle$ shared between them [4]. An entangled pair $|\Psi\rangle$ is provided to every pair of directly connected nodes by an entangled pair generator (EPG) located midway between them [8]. For creating a virtual quantum link by distributing entangled qubits ($|\Psi_{t,r}\rangle$) between end-nodes $\{t, r\}$ with no direct link, an entanglement swapping process [5] is used, which swaps the entangled pairs $|\Psi\rangle$ shared between directly connected nodes $\{i, j, \cdots, k\}$ (acting as quantum repeaters) constituting a path from the Tx to Rx given by $p_{t,r} = \{t - i - j - \cdots - k - r\}$. Further, we assume a network controller that sends control signals to trigger the generation of $|\Psi\rangle$ by the EPGs and swapping at the repeater nodes. However, the entangled qubits suffer decoherence due to inevitable probabilistic noises inherent in quantum channels (fibers, memories), imperfect measurement, and gate operations associated with the swapping process. This degrades the quality of teleported data qubits. Due to this reason, selecting a path with the least noise is considered extremely important for high-quality quantum communication.

### B. Design Objective

The decoherence of qubits varies over time and path, which depends on the noisy quantum channels present within the path and surrounding environment conditions. For instance, a single photonic qubit is so low in energy that it can be lost in a complex light background/surrounding. Thus, the overall effect of probabilistic noises on the qubits differs for different paths. Having complete knowledge of these noises, their source channels, and the time of occurrence in a network is highly improbable. To this end, we propose a challenging problem of finding the best quality (least noisy) path among the existing paths from a Tx to an Rx in a noisy quantum network with known deployment settings (fiber length and attenuation), but without any knowledge of probabilistic noises from memories and imperfect operations within the nodes. For a set of $K$ paths between the Tx node $t$ and Rx node $r$, given by $\mathcal{P}_{t,r} = \{p_{t,r}^{(k)}\}_{k=1}^{K}$, the least noisy route selection problem is mathematically expressed as:

$$p_{t,r}^{*} = \underset{p_{t,r}^{(\cdot)} \in \mathcal{P}_{t,r}}{arg\ max}\ \ f\left(p_{t,r}^{(\cdot)}\right), \tag{1}$$

where the function $f\left(p_{t,r}^{(\cdot)}\right)$ represents the quality of path $p_{t,r}^{(\cdot)}$ and $p_{t,r}^{*}$ denotes the best quality (least noisy) path, i.e., one with the highest fidelity $(F)$ of the distributed entangled qubits $|\Psi_{t,r}\rangle$, given by $F = \langle\Psi_{t,r}|\,\rho\,|\Psi_{t,r}\rangle$ with $\rho = |\Psi\rangle\langle\Psi|$.

## IV. Proposed Multi-Arm Bandit-based Solution

We use the upper confidence bound multi-arm bandit (UCB-MAB) algorithm [14] to solve our route selection problem (1). The UCB-MAB uses an agent which chooses an arm (called action) in each iteration to interact with the environment and learns the corresponding probabilistic reward, with the goal of maximizing the cumulative reward over multiple iterations. The arm with the highest mean reward at the end turns out to be the best arm. Specifically, we map the $K$ feasible paths in $\mathcal{P}_{t,r}$ connecting end nodes $\{t, r\}$ and a function of resultant fidelity of the entangled pairs distributed on these paths, respectively, to the $K$ arms and their corresponding rewards. We set the total number of iterations to $M$. In each iteration, the agent learns the effect of unknown probabilistic noises (from different quantum channels present in the chosen path) on the fidelity of the distributed entangled qubits, thus assessing the path quality. The action of choosing a path in the $m^{th}$ iteration is given by $A_m \in \mathcal{P}_{t,r}$ and the corresponding reward, $R_m$, is designed as:

$$R_m = \begin{cases} -2 \times (0.5 - F_m), & F_m \in [0, 0.25) \\ -1 \times (0.5 - F_m), & F_m \in [0.25, 0.5) \\ \ 1 \times F_m, & F_m \in [0.5, 0.6) \\ \ 2 \times F_m, & F_m \in [0.6, 0.7) \\ \ 3 \times F_m, & F_m \in [0.7, 0.8) \\ \ 4 \times F_m, & F_m \in [0.8, 0.9) \\ \ 5 \times F_m, & F_m \in [0.9, 1] \end{cases}. \tag{2}$$

Note that we penalize the reward value corresponding to the action $A_m$ severely as the resultant fidelity $F_m$ approaches 0 and vice-versa. Further, at iteration $m$, the mean reward value of the action $a \in \mathcal{P}_{t,r}$, i.e. $Q_m(a)$, is given by:

$$Q_m(a) = \frac{1}{N_m(a)} \sum_{i=1}^{m-1} R_i \mathbb{1}(A_i = a) \tag{3}$$

$$= \left(1 - \frac{1}{m-1}\right) Q_{m-1} + \frac{1}{m-1} R_{m-1}, \tag{4}$$

where $N_m(a)$ is the number of times the action $a$ has been selected prior to iteration $m$ and $\mathbb{1}$ is standard indicator function. In UCB, the action is intelligently chosen based on the function in eq. (5) that trades off the exploration of least tried actions and the exploitation of the actions with the highest mean reward value at the current iteration $m$:

$$A_m = \underset{a \in \mathcal{P}_{r,t}}{arg\ max} \left[ Q_m(a) + \gamma \sqrt{\frac{2 \times log\ m}{N_m(a)}}\ \right], \tag{5}$$

with the hyper-parameter $\gamma$ controls the level of exploration. After the last iteration $(M)$, the solution to our least noisy route selection problem in (1) is given by,

$$p_{t,r}^{*} = \underset{a \in \mathcal{P}_{t,r}}{arg\ max}\ Q_M(a). \tag{6}$$

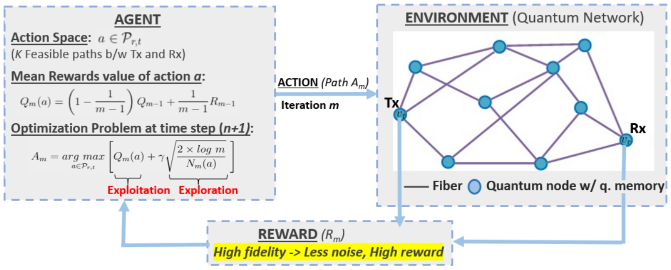



Fig. 2: Flow of our UCB-MAB based route selection approach.

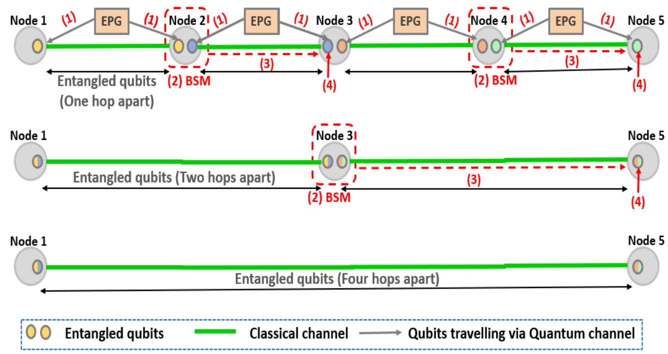



Fig. 3: Steps involved in end-to-end entanglement swapping.

The flow of our UCB-MAB based solution is shown in Fig. 2. Here, the rewards are probabilistic in nature due to the probabilistic noises that impact fidelity. We assume that the network controller sends the control signal to generate entangled pairs by the EPGs existing between the nodes one-hop apart on the chosen path in each iteration.

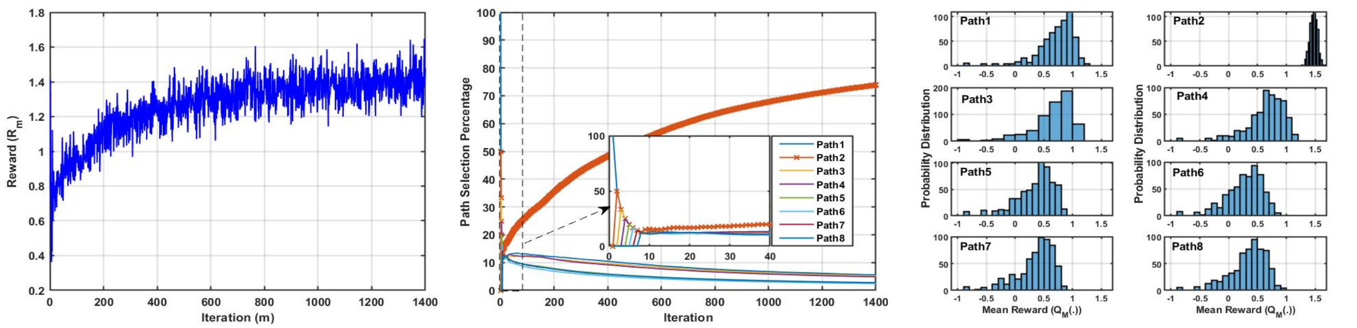
## V. Simulation Results and Discussions

In this section, we describe the quantum network environment created using our simulator and performance of our approach.

### A. *Quantum Network Simulator*

We design our custom simulator which (i) tracks time a qubit spends in memory, fiber, and gate operations during different steps (explained later) of entanglement swapping, (ii) calculates probability of occurrence of different noises, and (iii) generate them with these computed probabilities in every iteration. Thus, there is no need to create time events for each nodes, memories, and errors in different gate operation explicitly as required in other existing simulators. This requires prior knowledge of when the dephasing and other noises will occur in different qubits at different nodes in a large-scale network. Hence, our simulator provides ease of use without requiring in-depth and accurate knowledge of timings of different events. Further, it offers flexibility to do analysis of synchronized and non-synchronized swapping operations on nodes. This enables application of learning-based algorithms to interact with quantum network environment, which is not straight-forward with other existing simulators that mostly gives an average behavior of noises/errors on qubits' density matrix. Our simulator implements the effect of these errors on qubits' states at every instance. We design this simulator in MATLAB using QLib [16], which is a MATLAB package for quantum information theory calculations. This simulator can be accessed for research purposes at [15].

Our simulator allows simulation of a quantum network with user-defined number of nodes placed randomly in a desired area. It establishes a link between any two nodes whose distance is smaller than a configurable value and ensures that the resultant network is a connected graph. Further, it assumes presence of an EPG located midway on each link for generation of entangled pair and distribution of one of its entangled qubit to the downstream and upstream nodes each. To share an entangled qubits pair between Tx and Rx end nodes on a path, it simulates entanglement swapping at the intermediate nodes by carrying out the following steps: (1) transmission of entangled qubits from EPG to the nodes forming that link, (2) performing Bell state measurements (BSM) [5] on the two qubits (from different pairs) at the nodes based on the decided entanglement swapping order, (3) transmission of results of qubits' measurements to the nodes on respective upstream classical links, and (4) perform gate operations on the qubits of these upstream links' nodes to achieve entanglement with that on the respective downstream links' nodes. This results in sharing entangled qubits between nodes that are more than one-hop apart, as shown in Fig. 3. For each set of nodes $\{i, j, k\}$ connected via links $i-j$ and $j-k$, the time required for operation in steps (1)-(4) while attempting swapping at node $j$ are, respectively, $t_f =$ (fiber link length)/$(2 \times 10^8)$ s, $t_{BSM} = 10$ ns, $t_c =$ (classical link length)/$(3 \times 10^8)$ s, and $t_{op} = 10$ ns [12], [17]. To reduce decoherence of qubits, swapping is done at multiple nodes in parallel order. For instance, if a path with 5 nodes is given by $\{1-2-3-4-5\}$, then firstly the swapping will be done at nodes 2 and 4 to create virtual path $\{1-3-5\}$ (level-i), followed by swapping at node 3 to generate virtual path $\{1-5\}$ (level-ii). We consider two scenarios in this parallel swapping: (a) synchronized and (b) non-synchronized operations at the nodes involved in swapping at different levels. In the former case, measurement in step (2) happens after all the nodes (here $2, 4$) receive qubits from EPGs during level-i. Likewise, step (4) happens when classical information reaches all the respective nodes during each level [18]. In the latter case, a node initiates the steps $(1)-(4)$ as soon as the two qubits from different entangled pairs arrive. Thus, the time of start of operations in steps $(1)-(4)$ are different for different nodes [18]. Further, our simulator considers depolarization noise for modeling loss in fiber, dephasing noise in memory, bit flip error during BSM, and bit-phase flip error in gate operations. It generates the instances of these noises/errors in each iteration of UCB-MAB with probabilities $p_{depol} = f$(fiber length) [12], $p_{dephas} = f$(time qubit sits in memory) [12], $p_{BSM} = 0.2$, and $p_{gate} = 0.2$.




(a) (b) (c)

Fig. 4: (a) Rewards corresponding to the path selected in each iteration, (b) percentage of selection of each path with increasing iterations, and (c) distribution of mean reward of each path in our proposed approach for least noisy route selection.

### B. Simulation Parameters

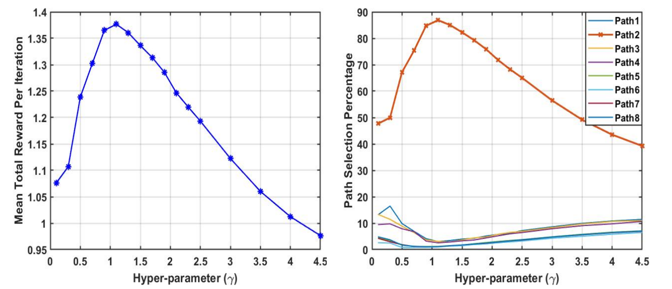
For the simulations in the following subsections, we consider a network spanning $20 \times 20$ km$^2$ area containing total $N = 15$ randomly deployed nodes. Quantum links are considered between nodes with inter-node distances $\leq 10$ Km. We set node 2 as Tx, node 15 as Rx, and considered $K = 8$ paths between them as feasible path set $\mathcal{P}_{t,r}$. These paths are a subset of all possible paths between the nodes $\{2, 5\}$, obtained by applying the knowledge of network deployment and attenuation of different fibers to infer loss in fidelity of end-to-end entangled qubits [12]. Note that we have used only fiber loss characterized by depolarization quantum channel model because noises from other channels are unknown. We consider the synchronized operations scenario in swapping process unless specified and provide the results averaged over 600 experiments, each with $M = 1400$ total iterations.

### C. Performance of Proposed Learning-based Route Selection

In Fig. 4(a), we show that our proposed approach improves the cumulative reward as iteration progresses, which highlights its convergence towards selection of the path with the highest rewards (fidelity). We further validate this in Fig. 4(b) which illustrates that the percentage of selection of path 2 increases with the iterations, while that for other paths decreases, and Fig. 4(c) which shows that the path 2 indeed has the highest mean reward $(1.45)$. Also, we observe that the value of fidelity corresponding to the least noisy (best) path using our simulated network is in the range $[0.7, 0.8]$ because we have considered decoherence due to fibers, memories, imperfect measurements, and errors in gate operations.

Further, from the magnified portion in Fig. 4(b), we observe that the best quality path 2 is unanimously obtained from iteration 30 onward and prior to that there is no specific trend in the percentage of selection of different paths. Thus, our proposed approach can give the final decision of least noisy path quickly, i.e., it does not require 1400 iterations. We chose this number to analyse the long-term behavior of our UCB-MAB-based solution.

Next, we compare the performance of our proposed approach with classical distance-based route selection approach for the simulated network in Table I. We observe that the path selected by the distance-based algorithm has the shortest length and low fidelity (quality), while that selected by our proposed approach has $\sim 33\%$ better fidelity. However, this gain is achieved at the cost of increase in computation time due to trial and error involved in the learning algorithm, as opposed to the calculation of distances of all the paths between each Tx and Rx pair in the distance-based algorithm.




(a) (b)

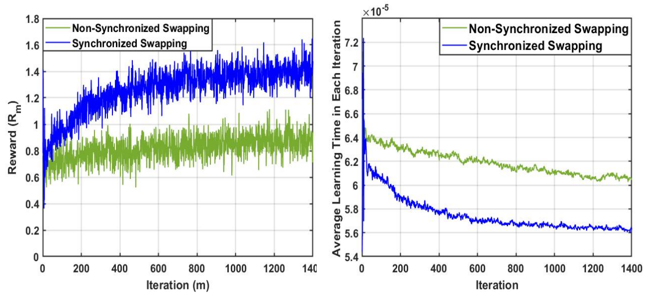
Fig. 5: Variation of (a) mean total reward per iteration $(\overline{R})$ and (b) percentage of selection of paths with different values of hyper-parameter $\gamma$.

### D. Selection of Hyper-parameter

To set the value of hyper-parameter $\gamma$, we study how the mean total reward per iteration $(\overline{R} = \frac{1}{M}\sum_{m=1}^{M} R_m)$ and mean reward of each path (arm) after final iteration $(Q_M(\cdot))$ change on varying the value of $\gamma$ in range [0.1,4.5]. From Fig. 5 (a), we observe that mean total reward first increases with increase in the level of exploration up to $\gamma = 2$ and thereafter decreases. This is because each path returns a distinct reward from a limited range of possible values and these paths are unlikely to generate rewards in the same range to keep the $\overline{R}$ same for different levels of exploration. Further, Fig. 5(b) depicts a peak in percentage of selection of the best quality path$-2$, which indicates the optimal value of $\gamma$ equals $1$ for balancing the exploration-exploitation trade-off in route selection. On further increasing $\gamma$ $(> 1)$, the selection percentage of best quality path keeps decreasing till the point

| Approaches | Selected Path | Path length (Km) | Average Reward ($Q_M(\cdot)$) | Fidelity |
|---|---|---|---|---|
| Proposed | 2 | 19.094 | 1.4729 | $\sim[0.7, 0.8]$ |
| Distance-based | 1 | 18.006 | 0.7011 | $\sim[0.6, 0.7]$ |

TABLE I: Comparison with state-of-the-art technique.




(a) (b)

Fig. 6: (a) Rewards returned by the network and (b) learning time per iteration for the synchronized & non-synchronized entanglement swapping scenarios.

the selection algorithm behaves as random search with each path having equal percentage of selection.

### *E. Network Design: Synchronized vs. Non-synchronized Entanglement Swapping*

In Figs. 6(a)-(b), we compare the two scenarios of synchronized and non-synchronized swapping at multiple nodes of a path in each level, explained in Sec. V-A. We observe that, on an average, the non-synchronized one takes more time to complete end-to-end entanglement distribution (equivalent to learning time per iteration), thereby yielding lower rewards (hence, fidelity). To further analyse, we take $5$ different paths containing total $4, 5, 6, 7,$ and $8$ nodes, respectively, and observe that the end-to-end entanglement distribution time is less in synchronized case for the paths with lesser nodes ($4, 5, 6$ nodes' path). To understand the reason, take an example of a $5$ nodes' path $\{1-2-3-4-5\}$. In synchronized case, two parallel swaps at nodes $2, 4$ can take place at level-i to form virtual path $\{1-3-5\}$, followed by one swap at node $3$. While in the non-synchronized one, if the one qubit each from two entangled pairs from its upstream link $(3-4)$ and downstream link $(2-3)$ arrive at node $3$ earlier than that at nodes $2, 4$, then at level-i swap will take place at node $3$ resulting in virtual path $\{1-2-4-5\}$ and other nodes will wait for next level. In level-ii, swap will take place at either node $2$ or $4$ based on qubits arrival, followed by final swap at level-iii. Thus, the non-synchronized case may end up taking more levels to complete swapping (hence, more time and less resultant fidelity).

## VI. Conclusion

In this paper, we propose a learning-based solution for the route selection problem in noisy quantum communication networks. Our solution does not require any prior knowledge of probabilistic noises from quantum memories and gate operations that inherently affect the quality of qubit. It leverages UCB multi-arm bandit to select the least noisy (the best quality) path with minimal knowledge of network deployment topology. Further, we provide insights for performing synchronized entanglement swapping at different nodes on a path over the non-synchronized one, so as to obtain a high fidelity (better quality) end-to-end entanglement, especially on paths with lesser nodes. We created a simulator that provides flexibility to play with different parameters involved in end-to-end entanglement distribution process and their order to provide network design level insights. Finally, we showed through simulations that our proposed approach selects the path with higher fidelity compared to the distance-based approach.

## References


[1] M. Caleffi, A. S. Cacciapuoti, and G. Bianchi, "Quantum internet: From communication to distributed computing!" in *Proceedings of the 5th ACM International Conference on Nanoscale Computing and Communication*, 2018, pp. 1–4.

[2] N. Gisin, G. Ribordy, W. Tittel, and H. Zbinden, "Quantum cryptography," *Reviews of modern physics*, vol. 74, no. 1, p. 145, 2002.

[3] P. Komar, E. M. Kessler, M. Bishof, L. Jiang, A. S. Sørensen, J. Ye, and M. D. Lukin, "A quantum network of clocks," *Nature Physics*, vol. 10, no. 8, pp. 582–587, 2014.

[4] D. Bouwmeester, J.-W. Pan, K. Mattle, M. Eibl, H. Weinfurter, and A. Zeilinger, "Experimental quantum teleportation," *Nature*, vol. 390, no. 6660, pp. 575–579, 1997.

[5] R. Van Meter and J. Touch, "Designing quantum repeater networks," *IEEE Communications Magazine*, vol. 51, no. 8, pp. 64–71, 2013.

[6] H.-J. Briegel, W. Dür, J. I. Cirac, and P. Zoller, "Quantum repeaters: the role of imperfect local operations in quantum communication," *Physical Review Letters*, vol. 81, no. 26, p. 5932, 1998.

[7] J. E. Martinez, P. Fuentes, P. M. Crespo, and J. Garcia-Frias, "Quantum outage probability for time-varying quantum channels," *Physical Review A*, vol. 105, no. 1, p. 012432, 2022.

[8] M. Caleffi, "Optimal routing for quantum networks," *IEEE Access*, vol. 5, pp. 22 299–22 312, 2017.

[9] A. S. Cacciapuoti, M. Caleffi, F. Tafuri, F. S. Cataliotti, S. Gherardini, and G. Bianchi, "Quantum internet: networking challenges in distributed quantum computing," *IEEE Network*, vol. 34, no. 1, pp. 137–143, 2019.

[10] 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*, vol. 5, no. 1, pp. 1–9, 2019.

[11] 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.

[12] K. Li, V. Chaudhary, S. G.Sanchez, and K. Chowdhury, "Q-FiRM: Fidelity-based Rate Maximizing Routes for Quantum Networks," in *IEEE Consumer Communications & Networking Conference (CCNC)*. IEEE, 2023.

[13] 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.

[14] T. Lattimore and C. Szepesvári, *Bandit algorithms*. Cambridge University Press, 2020.

[15] "Simulator for creating quantum network environment for Multi-arm bandit algorithm," https://genesys-lab.org/source-codes.

[16] S. Machnes, "Qlib-a matlab package for quantum information theory calculations with applications," *arXiv preprint arXiv:0708.0478*, 2007.

[17] C. Cicconetti, M. Conti, and A. Passarella, "Request scheduling in quantum networks," *IEEE Transactions on Quantum Engineering*, vol. 2, pp. 2–17, 2021.

[18] W. Kozlowski, A. Dahlberg, and S. Wehner, "Designing a quantum network protocol," in *Proceedings of the 16th International Conference on emerging Networking EXperiments and Technologies*, 2020, pp. 1–16.