

# 양자네트워크에서 심층강화학습을 활용한 End-to-End Request 스케줄링 최적화 연구

석영준<sup>1</sup>, 최요한<sup>1</sup>, 한연희<sup>1\*</sup>, 김주봉<sup>2</sup>, 임현교<sup>2</sup>, 이찬균<sup>2</sup>

<sup>1</sup>한국기술교육대학교 컴퓨터공학과 미래융합공학전공

<sup>2</sup>한국과학기술정보연구원

<sup>1</sup>{dsb04163, yowief, yhhan}@koreatech.ac.kr,

<sup>2</sup>{jjbong, hk.lim, chankyunlee}@kisti.re.kr

## Study of End-to-End Request Scheduling Optimization Using Deep Reinforcement Learning in Quantum Networks

Yeong-Jun Seok<sup>1</sup>, Yohan Choi<sup>1</sup>, Youn-Hee Han<sup>1\*</sup>, Ju-Bong Kim<sup>2</sup>, Hyun-Kyo Lim<sup>2</sup>, Chankyun Lee<sup>2</sup>

<sup>1</sup>Future Convergence Engineering, Dept. of Computer Science and Engineering, KOREATECH

<sup>2</sup>Korea Institute of Science and Technology Information

### 요 약

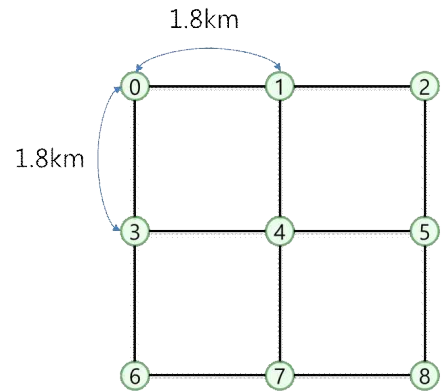
양자네트워크는 높은 보안성을 제공하지만, 네트워크의 동적인 특성과 양자 자원의 특성 때문에 효율적인 스케줄링이 필요하다. 기존의 휴리스틱 알고리즘은 Request의 우선순위를 기반으로 처리를 진행하지만, 특히 장거리 통신이 필요한 Request가 포함될 때 성능이 크게 저하되는 한계가 존재한다. 이를 극복하기 위해, 본 논문에서는 양자네트워크에서 효율적인 Request 스케줄링을 위해 심층강화학습(Deep Reinforcement Learning, DRL) 기반 알고리즘을 제안한다. 제안된 알고리즘의 성능은 NetSquid 시뮬레이션을 통해 기존 휴리스틱 알고리즘과 비교 실험을 진행했으며, 그 결과 DRL 기반 알고리즘이 더 많은 Request를 효율적으로 처리하고 지연 시간을 줄여 서비스 품질을 크게 개선하는 것을 확인할 수 있었다. 특히, 장거리 통신이 필요한 Request가 포함된 상황에서도 DRL은 기존 방법보다 더 뛰어난 성능을 보여 한계를 극복하는 데 성공했다. 본 연구는 DRL을 활용해 양자네트워크에서 Request 스케줄링을 최적화할 수 있는 가능성을 제시하며, 앞으로 다양한 네트워크 환경과 강화학습 알고리즘의 적용을 통해 성능 개선을 계속할 계획이다.

### I. 서론

양자네트워크는 고전적 네트워크에 비해 보안성과 데이터 전송의 신뢰성을 크게 강화할 수 있는 차세대 통신 기술로 주목받고 있다[1]. 양자의 특성을 활용하여 높은 수준의 보안을 제공하지만, 양자 자원의 특성과 동적 네트워크 환경에 적합한 Request 스케줄링 최적화가 필요하다. 양자네트워크에서 생성된 Request가 효율적으로 스케줄링되지 못할 경우, 서비스 지연 시간이 증가하며, 일정 시간 내에 처리되지 못한 Request는 만료되어 손실될 수 있다.

다양한 연구에서는 휴리스틱 알고리즘을 통해 이러한 Request 스케줄링 문제를 해결하고자 하였다[2]. 각 Request은 우선순위를 기반으로 하여 네트워크 상에서 양자 자원을 효율적으로 사용하도록 설계되었다. 그러나 장거리 통신이 필요한 Request가 포함될 경우, 휴리스틱 알고리즘의 성능은 크게 저하되는 한계가 존재한다. 따라서 장거리 Request가 존재하는 상황에서도 효율적으로 작동할 수 있는 새로운 접근 방식이 요구된다.

본 논문에서는 심층강화학습(Deep Reinforcement Learning, DRL)을 기반으로 한 Request 스케줄링 알고리즘을 제안한다. DRL에서는 환경(Environment)과 상호작용하는 에이전트(Agent)가 환경의 상태 정보(Observation)를 입력받아 행동(Action)을 출력하고, 그 결과에 따라 보상(Reward)을 받으며 보상을 최대화하는 정책(Policy)을 학습한다. 다양한 상태 정보를 바탕으로 최적의 행동을 학습하는 에이전트의 특성은 복잡한 최적화 문제를 해결하는 데 적합하다[3].



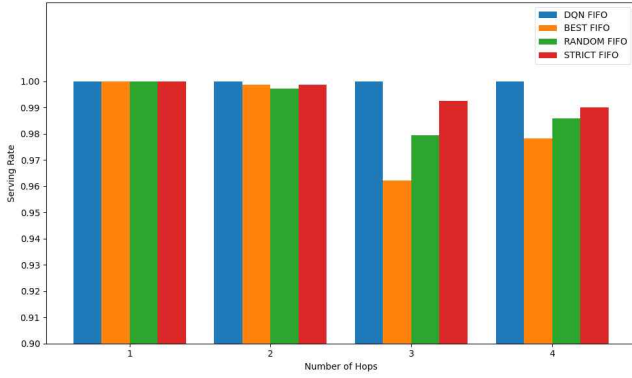
[그림 1] GRID Topology

본 연구에서는 DRL 기반의 Request 스케줄링 알고리즘을 제안하고, 이를 NetSquid 시뮬레이션을 통해 기존 휴리스틱 알고리즘과 성능 비교 실험을 수행하였다. 실험은 3x3 네트워크 구조에서 진행되었으며, 결과적으로 제안된 심층강화학습 기반 접근 방식의 우수성을 입증하였다.

### II. 본론

본 절에서는 DRL 환경의 시나리오, Observation, Action, 그리고 Reward에 대해 자세히 설명한다. NetSquid 시뮬레이션을 통해 생성된 양자 암호 통신망의 구조는 [그림 1]과 같다. 이 네트워

\* 한연희(Youn-Hee Han, [yhhan@koreatech.ac.kr](mailto:yhhan@koreatech.ac.kr)): 교신저자



[그림 2] Serving Rate in GRID Network

크에서 Request는 무작위 노드 간에 푸아송 분포(Poisson Distribution)를 따라 발생하며, Request Queue에서 대기 상태에 놓인다. Request Queue에서는 FIFO(First-In-First-Out) 방식으로 Request의 서비스 여부가 결정된다. Request는 발생한 노드 간의 양자 자원이 존재하는 양자 경로(End-to-End Entanglement)가 있을 경우, 알고리즘에 의해 서비스가 결정된다. 경로가 없거나 알고리즘이 'Skip'을 결정하면, Request는 서비스되지 않고 Request Queue에 다시 추가된다. 만약 Request Queue에 Skip된 Request만 남아있을 경우, 이들에 대해서는 한 번 더 서비스 여부를 결정한다.

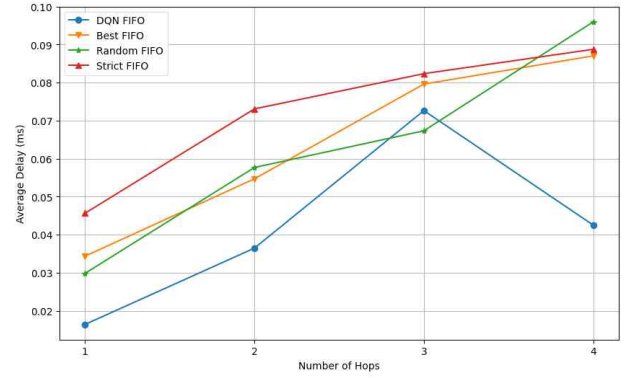
DRL 에이전트의 Observation은 양자 암호 통신망의 Topology 정보와 Request Queue 정보로 구성된다. Topology 정보는 서비스 대기 중인 Request의 경로 정보를 포함하며, 이를 통해 에이전트는 Request를 서비스할 경우 소모될 양자 자원을 파악할 수 있다. Request Queue 정보는 Queue에서 대기 중인 Request들의 정보를 포함하여, 향후 서비스해야 할 Request들을 예측할 수 있도록 한다.

DRL 에이전트의 Action은 경로가 존재하는 Request의 서비스 여부를 결정하는 것이다. Action은 0 또는 1로 표현되며, 0일 경우 Skip을, 1일 경우 해당 Request의 서비스를 의미한다. 한 번 Skip된 Request는 Request Queue에서 대기하며, 이후 경로가 존재할 경우 Action은 항상 1이 되어 서비스가 이루어진다. 따라서 한 번 Skip된 Request는 Request Queue에서 한 번 순환된 후 경로의 존재 여부에 따라 반드시 서비스될 수 있다.

DRL 에이전트의 Reward( $R$ )은 Request Queue에 경로가 존재하지 않는 Request만 남았을 때, 서비스된 Request에 대해 다음과 같은 식으로 정의된다:

$$R = \frac{1}{P} \sum_{i=0}^{P-1} \frac{f_i}{d_i} \quad (1)$$

수식 (1)에서  $P$ 는 서비스된 Request의 수,  $f_i$ 는  $i$ 번째로 서비스된 Request가 사용한 경로의 충실도(Fidelity), 그리고  $d_i$ 는  $i$ 번째로 서비스된 Request가 Queue에서 대기한 시간(Delay)이다. 이러한 설계를 통해 에이전트는 Request Queue에 대기 중인 Request를 최대한 많이 서비스하며, Delay를 줄이는 방향으로 동작하도록 학습된다. DRL 에이전트는 다수의 학습 과정을 통해 Reward를 최대화하는 방식으로 동작을 최적화한다. 이 과정에서 에이전트는 가능한 많은 Request를 서비스하고 Delay를 최소화하는 방향으로 행동한다.



[그림 3] Average Delay in GRID Network

[그림 2]는 기존의 휴리스틱 알고리즘과 DRL 기반 알고리즘(DQN, Deep Q-Network)의 성능을 비교한 실험 결과를 보여준다. 실험은 총 100회 진행되었으며, 각 실험의 결과는 평균값으로 제시된다. X축은 서비스된 Request의 경로 길이를 나타내는 'Number of Hops'를, Y축은 발생한 Request가 서비스될 확률을 나타내는 'Serving Rate'를 나타낸다. 기존의 휴리스틱 알고리즘과 비교했을 때, DRL 기반 알고리즘은 모든 Request에 대해 가장 높은 서비스 성공률을 보였다.

[그림 3]은 100회 실험에서 서비스된 Request의 경로 길이에 따른 평균 Delay를 보여준다. DRL은 대부분의 경로 길이에서 비교적 낮은 평균 Delay를 기록하며, 서비스 품질을 향상시킨다.

### III. 결론

본 논문에서는 양자네트워크에서 DRL을 활용한 Request 스케줄링 알고리즘을 제안한다. 제안하는 DRL 기반 알고리즘은 기존의 휴리스틱 알고리즘에 비해 더 많은 Request를 효율적으로 처리할 수 있으며, 서비스된 Request의 Delay를 줄여 서비스 품질을 향상시킨다. 특히, 장거리 Request가 존재하는 환경에서 DRL 알고리즘은 보다 높은 서비스 품질을 제공한다. 향후 연구에서는 다양한 네트워크 환경에서 DRL 알고리즘의 적용 가능성을 검토하고, 다양한 강화학습 알고리즘과의 성능 비교를 통해 제안 방법의 효율성을 더욱 개선할 계획이다.

### ACKNOWLEDGMENT

이 논문은 2023년도 정부(교육부)의 재원으로 한국연구재단의 지원을 받아 수행된 기초연구사업임 (No. NRF-2023R1A2C1003143).

### 참고문헌

- [1] 권오성, 김용수, 한상욱, 문성욱. 미래통신 보안기술: 양자암호통신 연구 현황 및 전망. Telecommunications Review, 24(3), 404-418. 2014.
- [2] C. Cicconetti, M. Conti and A. Passarella, "Request Scheduling in Quantum Networks," in IEEE Transactions on Quantum Engineering, vol. 2, pp. 2-17, 2021
- [3] Arulkumaran, K., Deisenroth, M. P., Brundage, M., & Bharath, A. A. (2017). Deep reinforcement learning: A brief survey. IEEE Signal Processing Magazine, 34(6), 26-38.