그래프 어텐션 네트워크 기반 심층강화학습을 통한 QKD 최적화 연구

석영준[1], 최요한[1], 최호빈[1], 한연희[1†], 임현교[2], 이찬균[2]
[1]한국기술교육대학교 컴퓨터공학과 미래융합공학전공
[2]한국과학기술정보연구원
[1]{dsb04163, yowief, chb3350, yhhan}@koreatech.ac.kr,
[2]{hk.lim, chankyunlee}@kisti.re.kr

Research on Quantum Key Distribution Optimization through Graph Attention Network-based Deep Reinforcement Learning

Yeong-Jun Seok[1], Yohan Choi[1], Ho-Bin Choi[1], Youn-Hee Han[1†], Hyun-Kyo Lim[2], Chankyun Lee[2]
[1]Future Convergence Engineering, Dept. of Computer Science and Engineering, KOREATECH
[2]Korea Institute of Science and Technology Information

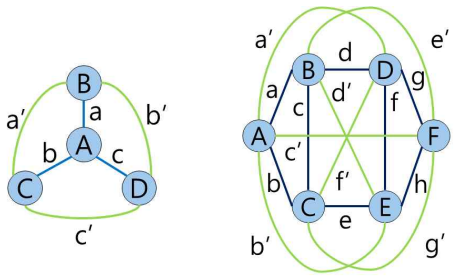
요 약

네트워크 통신의 암호체계의 안전성은 양자 컴퓨터의 등장으로 위협받고 있다. 따라서 안전한 통신 암호체계로 양자의 특성을 활용한 양자키 통신 기술이 주목받고 있다. 양자키 통신의 QKD (Qunatum Key Distribution) 기술은 직접 양자키(Direct Key)를 자원으로 간접 양자키(Indirect Key)를 생성하여 원거리 통신을 지원하기 위해 할당한다. 따라서 직접 양자키 자원의 효율적 활용을 위해 QKD의 간접키 생성을 최적화 할 필요가 있다. 본 논문은 그래프 어텐션 네트워크(Graph Attention Network, GAT) 기반 심층강화학습(Deep Reinforcement Learning, DRL)을 활용할 때 적합한 QKD 최적화 기법을 제안한다. 제안한 기법은 양자키의 모든 정보를 그래프 형태로 입력받고, DRL의 에이전트(Agent)가 행동(Action)으로 직접키의 생성 개수를 결정해 QKD의 최적화를 통한 통신 블록(Blocking) 현상을 줄이도록 한다. 또한 그리디 알고리즘(Greedy Algorithm)과의 비교 평가를 통해 QKD 최적화에서 GAT 기반 DRL 적용의 가능성을 보여준다.

Ⅰ. 서론

네트워크 통신에서 암호체계의 안전성은 네트워크 보안에서 반드시 고려되는 사항이다. 하지만 높은 계산 성능을 가진 양자 컴퓨터는 높은 계산 복잡도를 통해 안전성을 보장하는 현재 암호체계에 위협이 된다.[1-2] 따라서, 양자의 불확정성의 원리에 기반 한 양자키 통신 기술이 새로운 새로운 보안 체계로 주목받고 있다. 현재 유럽전기통신표준협회(ETSI) 및 한국정보통신기술협회(TTA)를 포함한 다양한 국내외 표준화 그룹에서 양자암호 통신망 표준화 작업을 진행 중이다. 표준화된 양자암호 통신망 구조는 양자키 분배(Qunatum Key Distribution (QKD))를 위한 네트워크 단위의 키 분배 메커니즘 및 양자키 관리 시스템이 구축되어 있다. QKD는 양자키를 직접키(Direct Key)및 간접키(Indirect Key)로 구분하고, 양자키의 생성, 할당, 삭제 등을 관리한다. 특히 간접키는 직접키를 자원으로 생성되며, 원거리 통신을 위해 사용된다. 양자암호 통신에서 QKD는 통신에 직접키 및 간접키를 할당한다. 따라서, 원활한 통신을 위해 QKD의 양자키의 생성량은 최적화될 필요가 있다.

강화학습(Reinforcement Learning)의 맥락에서, 에이전트(Agent)는 환경(Environment)과 상호작용 하여 누적 보상(Reward)을 최대화하는 정책(Policy)를 찾는다. 정책은 에이전트가 출력하는 일련의 행동(Action)들을 나타낸다. 에이전트는 상태(Observation)를 입력받아 행동을 출력하고, 환경으로부터 보상을 받는다. 심층강화학습(Deep Reinforcement Learning (DRL))[3]은 심층 신경망(Deep Neural Network (DNN))을 통해 기존 강화학습의 한계를 해결한다. 그래프 어텐션 네트워크(Graph Attention Network (GAT))[4]는 그래프의 중요도를 반영하여, 직접 분석 가능한 DNN의 일종이다.



(a) Toy (b) Butterfly

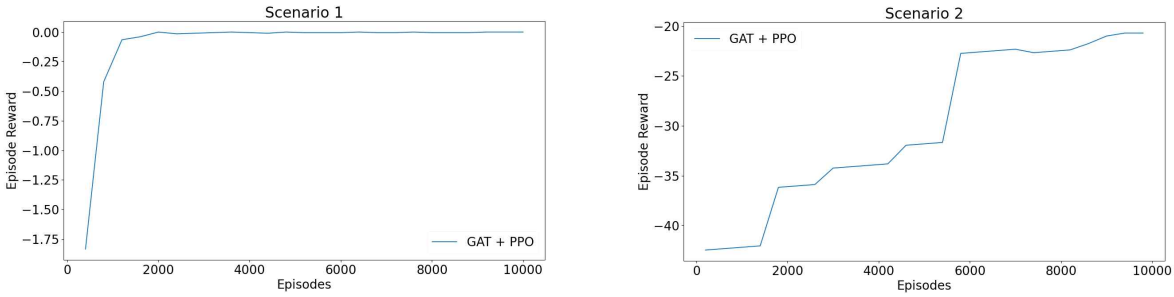
[그림 1] 본 연구에서 활용하는 네트워크 토폴로지

본 논문은 QKD 시스템의 최적화를 위해 GAT 기반의 DRL 기법을 제안한다. 표준화된 양자암호통신망 구조를 강화학습 환경으로 구현하여 베이스 라인(Base-Line)인 그리디 알고리즘(Greedy Algorithm)과 비교실험을 통해 GAT 기반 DRL의 우수성을 증명한다.

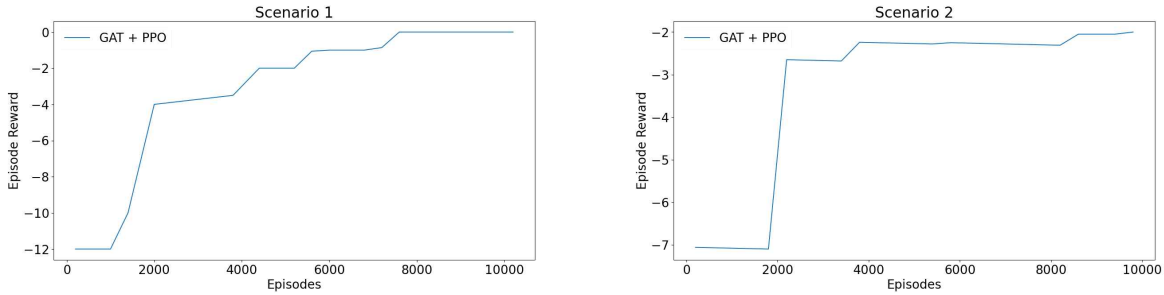
Ⅱ. 본론

본 논문에서 모든 양자키는 생성 후 키 풀(Key Pool)에 저장되고 사용된다. 모든 양자키는 생명 주기(Life Cycle)을 가지고 있고, 생명 주기가 만료되면 삭제된다. 양자키 중 간접키는 직접키를 자원으로 생성된다. 자원으로 사용되는 직접키의 종류 및 개수는 시스템의 경로 설정을 통해 결정된다.

† 한연희(Youn-Hee Han, yhhan@koreatech.ac.kr): 교신저자



(a) Toy 네트워크 토폴로지에서의 GAT + PPO 훈련 중 테스트 에피소드 보상 결과



(b) Butterfly 네트워크 토폴로지에서의 GAT + PPO 훈련 중 테스트 에피소드 보상 결과

[그림 2] GAT + PPO 훈련 중 테스트 에피소드 보상 결과

(표 1) 평균 블록 횟수 비교

	Scenario 1(B)		Scenario 2(B)	
Algorithm	Toy	Butterfly	Toy	Butterfly
GAT + PPO	0.0	0.0	23.19	2.09
Greedy	3.0	13.0	30.11	4.63

양자키를 할당받지 못한 통신은 블록(Blocking, B)되며, 누락된다. 본 논문에서는 원할한 통신을 위해 시간당 B인 B_t를 감소시키기 위해 문제의 목적함수를 다음과 같이 나타낸다.

$$mininize \lim_{T\to\infty} \frac{\sum_{t=0}^{T} B_t}{T} \qquad (1)$$

에이전트는 환경으로부터 모든 키 풀의 정보를 그래프 형태로 입력받고, 간접키의 생성 개수를 행동으로 출력한다. 시간 t에서 보상은 r_t로 나타내고 B_t가 증가하면 누적 보상 R이 감소한다.

$$R = \sum_{t=0}^{T} r_t = \sum_{t=0}^{T} -B_t \qquad (2)$$

실험에서는 GAT와 심층강화학습 알고리즘인 PPO (Proximal Policy Optimization)의 결합과 그리디 알고리즘을 비교한다. 다양한 비교실험을 위해 네트워크 토폴로지 및 시나리오를 각각 2개 사용하여 총 4번의 실험을 실시한다. 실험의 결과는 그림 2로 나타낸다.

총 100회 테스트의 발생한 평균 B의 값을 표로 제시한다. 시나리오 1은 통신이 정적으로 발생하고, 특정 통신이 발생한 시간에 간접키를 만들 수 없다. 시나리오 2는 통신이 푸아송 분포(Poisson Distribution)를 따라 동적으로 발생한다. 시나리오 1에서 그리디 알고리즘은 B가 발생하고, 심층강화학습은 학습을 통해 미리 간접키를 생성해 B가 발생하지 않았다. 시나리오 2에서 심층강화학습의 B가 그리디 알고리즘보다 적다. 따라서 심층강화학습이 모든 시나리오에서 그리디 알고리즘보다 성능이 높은 것을 알 수 있다.

Ⅲ. 결론

본 논문에서는 QKD 최적화를 위한 GAT 기반 DRL을 기법을 제안한다. 실험 결과를 통해 제안하는 기법을 활용할 때 DRL은 통신의 발생 패턴을 학습할 수 있고, 그리디 알고리즘 보다 적절한 직접키의 생성 개수를 결정할 수 있음을 증명했다. 향후 연구에서는 실제 QKD가 적용된 네트워크를 환경으로 실험을 진행할 예정이다.

ACKNOWLEDGMENT

이 논문은 2023년 및 2018년도 정부(교육부)의 재원으로 한국연구재단의 지원을 받아 수행된 연구임 (NRF-2023R1A2C1003143 & NRF-2018R1A6A1A03025526).

참고문헌

[1] 권오성, 김용수, 한상욱, & 문성욱. (2014). 미래통신 보안기술: 양자암호통신 연구 현황 및 전망. Telecommunications Review, 24(3), 404-418.

[2] Arute, Frank, et al. "Quantum supremacy using a programmable superconducting processor." Nature 574.7779 (2019): 505-510.

[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.

[4] Veličković, P., Cucurull, G., Casanova, A., Romero, A., Lio, P., & Bengio, Y. (2017). Graph attention networks. arXiv preprint arXiv:1710.