Journal of Advanced Technology Research Vol.9, No.1, pp.13-16, ISSN: 0974-6269

심층 강화학습 기반 세션 블록 최소화를 위한 양자 키 분배 네트워크의 종단 간 키 생성 방법

End-to-End Key Generation Method in Quantum Key Distribution Networks for Minimizing Session Blocks Using Deep Reinforcement Learning

석 영 준(Yeongjun Seok), 김 주 봉(Ju-Bong Kim), 한 연 희(Youn-Hee Han)

Advanced Technology Research Center
Korea University of Technology and Education
{dsb04163, rlawnqhd, yhhan}@koreatech.ac.kr

요 약

양자 컴퓨터와 양자 알고리즘의 발전으로 인해 전통적인 암호체계가 위협받고 있다. 이로 인해, 네트워크 통신의 안전한 보안을 위해 양자 키 분배(Quantum Key Distribution, QKD) 기술이 주목받고 있다. QKD 시스템은 양자 물리학의 원리를 기반으로 한 보안 기술로, 양자 컴퓨터의 위협으로부터 안전한 통신을 보장한다. 하지만 생성된 암호키는 사용에 제한이 있고, 재사용이 불가능기 때문에 효율적인 키 관리가 필수적이다. 본 논문에서는 그래프 어텐션 네트워크(Graph Attention Network, GAT)와 LSTM(Long Short-Term Memory)을 DRL과 결합하여 QKD의 키 생성 및 할당 최적화 기법을 제안합니다. 이를 통해 QKD 네트워크에서 발생하는 세션 블록 현상을 최소화하도록 한다. 또한 그리디 알고리즘보다 더 효과적으로 키 자원을 관리하고 블록 현상을 줄일 수 있다는 것을 비교 평가를 통해 증명한다.

키워드: 양자 키 분배, 심층 강화 학습, 종단 간 키 생성, 세션 블록 최소화, GAT, LSTM

Abstract

The advancement of quantum computers and quantum algorithms poses a threat to traditional cryptographic systems. Consequently, Quantum Key Distribution (QKD) technology, which ensures secure communications by leveraging the principles of quantum physics, is gaining attention for securing network communications. However, the generated cryptographic keys in QKD systems are limited in usage and cannot be reused, necessitating efficient key management. In this paper, we propose an optimization technique for QKD key generation and allocation by combining Graph Attention Networks (GAT) and Long Short-Term Memory (LSTM) with Deep Reinforcement Learning (DRL). This approach aims to minimize session block occurrences in QKD networks. Furthermore, we demonstrate through comparative evaluation that this method manages key resources more effectively and reduces blocking occurrences more efficiently than the greedy algorithm.

Key words: Quantum Key Distribution, Deep Reinforcement Learning, End-to-End Key Generation, Session Block Minimization, GAT, LSTM

1. 서 론

양자 컴퓨터와 양자알고리즘의 발전과 함께 전통적인 암호화 체계는 위협을 받고 있다[1]. 현재 널리 사용되는 암호화 방식은 수학적 복잡성에 의존하고 있지만, 양자 알고리즘은 이를 쉽게 해독할 수 있는 잠재력을 지니고 있다. 이로 인해 금융 및 다양한 산업 분야에서 기밀성과 신뢰성을 유지하기 위한 새로운 보안 기술이 절실히 필요하다. 이러한 배경에서 양자 키 분배(Quantum Key Distribution, QKD) 기술이 주목받고 있다[2].

QKD는 양자 물리학의 원리를 활용한 완벽한 보안을 제공하는 암호 키를 생성하고 분배하는 시스템이다. 양자 키 분배(QKD) 네트워크는 키의 생성, 분배 및 관리에 있어 많은 도전과제를 안고 있습니다. 특히, QKD 네트워크에서는 암호 키의 재사용이 불가능하고, 키의 수명이 제한적이기 때문에 효율적인 키 관리가 중요합니다. 이러한 특성 때문에 QKD 네트워크에서 생성된 키는 다양한 인증 서비스에 할당되어야 하며, 이를 위해 효과적인 자원 할당 방법이 요구된다. 따라서 QKD 네트워크의 성능을 최적화하고 보안을 강화하기 위해 자원 할당 문제를 해결할 수 있는 새로운 접근 방법이 필요합니다

본 연구에서는 QKD 네트워크에서 효율적인 키 관리와 자원 할당을 위해 심층 강화 학습(DRL) 기반의 새로운 방법론을 제안합니다. 그래프 어텐션 네트워크(Graph Attention Network, GAT)와 장단기 메모리(Long Short-Term Memory, LSTM)을 결합한 DRL 기법을 제안하여 QKD의 키 생성 및 할당 최적화를 목표로 한다. 이를 통해 QKD 네트워크에서 발생하는 세션 블록 현상을 최소화하고, 그리디 알고리즘보다 더 효과적으로 키 자원을 관리할 수 있음을 비교 평가를 통해 증명하고자 한다. 실험 결과, 제안된 DRL 기반 방법론이 기존의 휴리스틱 방법보다 더 나은 성능을 보이며, QKD 네트워크의 효율성을 크게 향상시킬 수 있음을 확인할 수 있다.

2. 양자 키 분배 네트워크 모델

본 장에서는 양자 키 분배 네트워크의 시스템 모델을 설명한다. 시스템 모델은 양자 키 분배 네트워크 모델, 세션 모델로 구성된다.

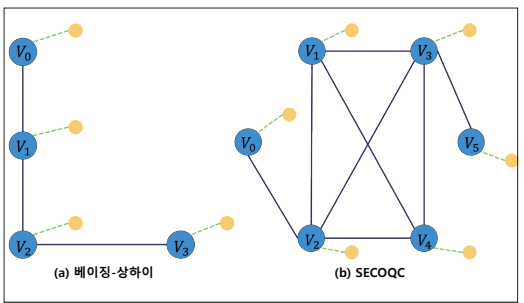

그림 1. 실제 QKD 시스템이 적용된 네트워크

2.1. 양자 키 분배 네트워크 모델

그림 1은 본 연구에서 사용한 양자 키 분배 네트워크의 토폴로지이다. 양자 키 분배 네트워크 토폴로지를 구성하는 임의의 노드 V_i는 QKD 노드를 나타낸다. QKD 노드를 연결하는 링크는 양자 링크를 나타낸다. 양자 링크로 연결된 노드들은 양자 키 풀을 보유하고 있고, 양자 키 풀에 양자 키를 저장한다. 양자 링크로 연결되지 않은 노드들은 종단 간 키 풀을 보유하고 있고, 종단 간 키 풀에 종단 간 키를 저장한다.

모든 키는 생명주기를 가지고 있고, 생명주기에 따라 생성 및 삭제된다. 양자 키는 주기적으로 생성되지만, 종단 간 키는 키 릴레이 방식으로 생성된다. 키 릴레이는 양자 링크로 연결되지 않은 노드들 간의 최단 경로를 통해 랜덤 넘버를 공유해 종단 간 키를 생성하는 방법이다. 랜덤 넘버를 공유하는 과정에서 랜덤 넘버의 유출을 막기 위해 양자 키를 사용한 암호화 방법이 사용된다. 따라서 종단 간 키는 생성될 때 다수의 양자 키를 소모하게 된다. 그림 1 (a)에서 키 릴레이 과정은 다음과 같다.

- V_0와 V_2간의 종단 간 키 생성을 위한 최단 경로 V_0-V_1-V_2 생성

- V_0, V_2와 V_2, V_2의 양자 키 x개 소모
- V_0와 V_2간의 종단 간 키 x개 생성

2.2. 세션 모델

세션은 QKD 노드와 연결된 서비스 노드 간에 발생하는 암호화를 요구하는 인증 또는 통신이다. 모든 세션은 푸아송 분포 (Poisson Distribution)에 따라 확률적으로 생성되며, 서비스 노드들과 연결된 QKD 노드들의 관계에 따라 두 가지 종류로 구분된다. QKD 노드들이 양자 링크로 연결되어 있다면 직접 세션이라 정의되며, 암호화를 위해 양자 키를 소모한다. QKD 노드들이 양자 링크로 연결되지 않았다면 간접 세션이라 정의되며 암호화를 위해 종단 간 키를 소모한다. 양자 키 또는 종단 간 키의 부족으로 암호화에 실패한다면, 이것은 세션 블록 (Session Blocks)이라 정의된다. 시간 t에서 발생한 세션 블록의 수를 SB_t로 정의한다.

3. Markov Decision Process

본 장에서는 시스템 모델에서 세션 블록을 최소화하기 위한 심층강화학습의 마르코프 결정 과정 (Markov Decision Process, MDP)을 (S,A,R,T,γ)로 정의한다. 시간 t에서 QKD 네트워크 환경에서 상태 $s_t \in S$가 에이전트에게 입력되고, 행동 $a_t \in A$가 결정된다. 에이전트는 자신의 행동에 따라 보상 r_t를 받고, 환경은 상태 전이 확률 $T(s_t,a_t,s_{t+1})=P(s_{t+1} \mid s_t,a_t)$에 따라 다음 상태 s_{t+1}로 전이된다. 그러나 에이전트는 상태 전이 확률을 알 수 없다. γ는 현재 보상이 미래의 보상보다 더 중요한지를 나타내는 수치인 할인 인자(discount factor)이다. 여기서 상태 s_t는 QKD 네트워크의 토폴로지 정보, 모든 키의 수와 생명주기 정보를 포함한다. 행동 a_t는 종단 간 키의 생성량을 나타낸다. 보상 r_t는 $-SB_t$이다. 에이전트는 각 상태에서 수행할 행동을 결정하는 정책 함수 π를 가진다. 학습을 통해 최적의 정책 π^*가 발견되면 MDP를 해결할 수 있다. 따라서 최적의 종단 간 키 공급 정책을 얻기 위해 DRL 방식을 통해 에이전트를 QKDN 환경에서 학습시킨다.

4. 실험 및 결과

표 1. 다양한 푸아송 분포 기대 값에서 각 환경의 100회 평균 에피소드 SB

	베이징-상하이			SECOQC		
Algorithm	λ=0.8	λ=1.0	λ=1.2	λ=0.8	λ=1.0	λ=1.2
PPO	14.53	29.91	45.48	5.4	20.58	44.87
Greedy	21.76	32.01	55.13	15.3	29.1	55.81

학습에 사용된 심층강화학습 알고리즘은 Graph Attention Model (GAT) [3]과 Long-Short Term Model (LSTM) [4]을 활용한 Proximal Policy Objective (PPO) [5] 알고리즘이다. 그림 2와 그림 3은 PPO 알고리즘의 학습 결과를 나타낸다. 푸아송 분포 기대 값 λ가 1인 환경에서 5회 반복된 학습의 평균 에피소드 보상 및 95% 신뢰 구간을 나타낸다. 반복 학습에서 평균 에피소드 보상이 증가하며, 심층강화학습 기법이 에피소드 SB를 감소시키는 것을 볼 수 있다. 표 1은 심층강화학습 기법의 성능과 Heuristic한 그리디 알고리즘의 성능을 비교 실험한다. 그리디 알고리즘은 시간 t에서 발생한 세션의 수와 같은 수의 종단 간 키를 생성하도록 설계했다.

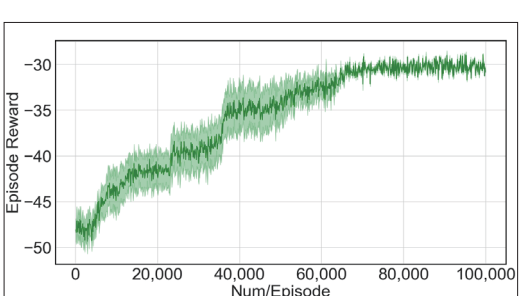

그림 2. 베이징-상하이 네트워크 토폴로지에서 훈련 중 PPO의 에피소드 보상 그래프

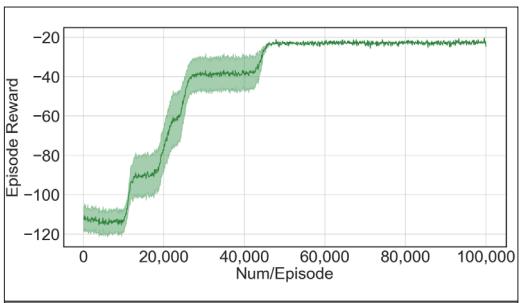

그림 3. SECOQC 네트워크 토폴로지에서 훈련 중 PPO의 에피소드 보상 그래프

5. 결론

본 논문에서는 QKD 네트워크에서 심층강화학습 기반의 효율적인 종단 간 키 생성 방법을 제안한다. 학습 결과와 비교 실험 결과를 통해 제안하는 기법이 세션의 발생 패턴을 학습할 수 있고, 다양한 세션 발생 패턴에 대응하여 그리디 알고리즘보다 효율적인 종단 간 키 생성을 할 수 있음을 증명한다. 추후 실제 양자 네트워크 시뮬레이션에서 더욱 발전한 심층강화학습 기법을 적용할 예정이다.

참고 문헌

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

[2] A. I. Nurhadi and N. R. Syambas, "Quantum Key Distribution (QKD) Protocols: A Survey," 2018 4th International Conference on Wireless and Telematics (ICWT), Nusa Dua, Bali, Indonesia, 2018, pp. 1-5

[3] Veličković, P., Cucurull, G., Casanova, A., Romero, A., Lio, P., & Bengio, Y. (2017). Graph attention networks.

[4] S. Hochreiter and J. Schmidhuber, "Long Short-Term Memory," in Neural Computation, vol. 9, no. 8, pp. 1735-1780, 15 Nov. 1997.

[5] J.Schulman,F.Wolski,P.Dhariwal,etal,Proxi malpolicyopti- mization algorithms, CoRR abs/1707.0 (2017).