(19) 대한민국특허청(KR)
(12) 공개특허공보(A)

(11) 공개번호 10-2025-0094964
(43) 공개일자 2025년06월26일

(51) 국제특허분류(Int. Cl.)
H04L 9/08 (2006.01) *G06N 3/092* (2023.01)
(52) CPC특허분류
H04L 9/0852 (2023.08)
G06N 3/092 (2023.01)
(21) 출원번호 10-2023-0185526
(22) 출원일자 2023년12월19일
심사청구일자 2023년12월19일

(71) 출원인
한국기술교육대학교 산학협력단
충청남도 천안시 동남구 병천면 충절로 1600 (한국기술교육대학교내)
(72) 발명자
한연희
대전광역시 유성구 가정로 306-6, 도룡SKVIEW아파트 101-1201
석영준
충청남도 천안시 동남구 북면 충절로 1456, 중앙아파트 105동 1204호
최요한
충청남도 천안시 동남구 병천면 가전7길 1-30, 105호
(74) 대리인
김견수

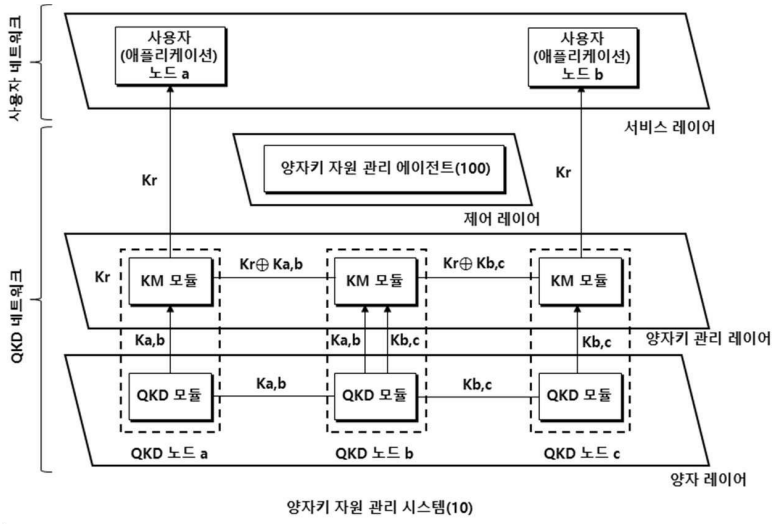
전체 청구항 수 : 총 10 항

(54) 발명의 명칭 **심층 강화학습 기반 최적화 기법을 통한 양자키 자원 관리 시스템 및 그 방법**

(57) 요 약

본 발명은 심층 강화학습 기반 최적화 기법을 통한 양자키 지원 관리 시스템 및 그 방법에 관한 것으로, 사용자 네트워크의 현재 상태에 대한 세션정보에 따라 양자키 분배 네트워크를 구성하는 개별 노드의 타입, 양자키를 요구하는 세션의 수, 양자키의 라이프 타임, 각 노드 간 연결정보, 각 노드 간 연결타입을 포함하는 양자키 분배 네트워크의 상태데이터를 강화학습하여 다음 상태에 대한 간접키의 생성량을 결정하여 양자키가 할당되지 못하는 세션을 최소화함으로써 안정적인 통신 상태를 유지할 수 있도록 하는 시스템 및 그 방법에 관한 것이다.

대 표 도 - 도1



양자키 자원 관리 시스템(10)

(52) CPC특허분류
H04L 9/0819 (2013.01)
H04L 9/0891 (2013.01)

이 발명을 지원한 국가연구개발사업

과제고유번호	1345347164
과제번호	2018R1A6A1A03025526223128209934020 2
부처명	교육부
과제관리(전문)기관명	한국연구재단
연구사업명	중점연구소지원사업
연구과제명	상호작용 가상현실 기반 몰입형 교육 훈련 플랫폼
기 여 율	1/1
과제수행기관명	한국기술교육대학교
연구기간	2018.09.01 ~ 2026.02.28

명 세 서

청구범위

청구항 1

복수의 사용자 노드를 포함하는 사용자 네트워크; 및

상기 복수의 사용자 노드에 대응하는 복수의 양자키 분배노드와 상기 복수의 양자키 분배노드를 제어하기 위한 양자키 자원 관리 에이전트를 포함하는 양자키 분배 네트워크;를 포함하며,

상기 양자키 자원 관리 에이전트는, 심층 강화학습을 통해서 적어도 하나 이상의 양자키 분배노드에 대한 간접키 생성량을 결정함으로써 직접키와 간접키를 포함하는 양자키의 관리 자원을 최적화하는 것을 특징으로 하는 심층 강화학습 기반 최적화 기법을 통한 양자키 자원 관리 시스템.

청구항 2

청구항 1에 있어서,

상기 양자키 자원 관리 에이전트는,

타임스텝별로 상기 양자키를 할당받지 못하는 세션의 수를 최소화하도록 상기 적어도 하나 이상의 양자키 분배노드에 대한 간접키의 생성량을 결정하는 것을 특징으로 하는 심층 강화학습 기반 최적화 기법을 통한 양자키 자원 관리 시스템.

청구항 3

청구항 1에 있어서,

상기 양자키 자원 관리 에이전트는,

상기 사용자 네트워크로부터 사용자 노드 간 세션정보를 수신하는 세션정보 수신부;

상기 세션정보가 수신되면 상기 복수의 양자키 분배노드로부터 각각 저장 및 관리되는 직접키와 간접키의 수를 포함하는 양자키 분배 네트워크의 상태정보를 수집하는 상태정보 수집부;

상기 수신한 세션정보와 상기 수집한 상태정보를 기반으로 상기 양자키 분배 네트워크에 대한 상태데이터를 생성하는 상태데이터 생성부; 및

상기 생성한 상태데이터를 심층 강화학습모델에 적용하여 적어도 하나 이상의 양자키 분배노드에 대한 간접키 생성량을 결정하는 간접키 생성량 결정부;를 포함하며,

상기 상태데이터는, 상기 복수의 양자키 분배노드, 각 양자키 분배노드에서 상기 직접키를 저장 관리하는 직접키 풀 및 상기 간접키를 저장 관리하는 간접키 풀에 따른 노드의 타입, 상기 직접키 혹은 간접키에 관련된 세션의 수, 상기 직접키 및 간접키에 대한 라이프 타임 또는 이들의 조합을 포함하는 노드 특징, 상기 노드 간 연결정보, 상기 노드 간 연결타입 또는 이들의 조합을 포함하는 것을 특징으로 하는 심층 강화학습 기반 최적화 기법을 통한 양자키 자원 관리 시스템.

청구항 4

청구항 1에 있어서,

상기 양자키 자원 관리 에이전트는,

상기 간접키 생성량을 결정한 양자키 분배노드와 상기 간접키 생성량에 따라 생성할 간접키를 공유할 또 다른 양자키 분배노드 간의 최단경로를 산출하는 최단경로 산출부; 및

상기 산출한 최단경로 상에 위치하는 상호 인접한 양자키 분배노드의 직접키를 상기 간접키 생성량을 결정한 양자키 분배노드로 제공하여 간접키를 생성하고, 상기 간접키 생성에 소모된 상기 직접키를 삭제하도록 제어하는 양자키 분배노드 제어부;를 더 포함하는 것을 특징으로 하는 심층 강화학습 기반 최적화 기법을 통한 양자키 자

원 관리 시스템.

청구항 5

청구항 1에 있어서,

상기 양자키 자원 관리 에이전트는,

상기 간접키 생성량에 따라 각 양자키 분배노드에서 간접키를 모두 생성하면, 양자키를 할당받지 못한 세션의 수에 대한 마이너스 값을 보상으로 산출하고, 양자키를 할당받지 못한 세션의 수가 0이면 0의 값을 보상으로 산출하는 보상 산출부; 및

상기 산출한 보상을 심층 강화학습모델에 역전파(back propagation)시켜, 누적보상이 최대가 되는 액션을 추출하여 상기 간접키 생성량을 결정하도록 상기 심층 강화학습모델을 강화학습하는 심층 강화학습부;를 더 포함하는 것을 특징으로 하는 심층 강화학습 기반 최적화 기법을 통한 양자키 자원 관리 시스템.

청구항 6

청구항 1에 있어서,

상기 양자키 자원 관리 에이전트는,

세션정보에 따라 사용자 노드로부터 해당 사용자 노드에 해당하는 양자키 분배노드에 대한 양자키 요청이 각각 있으면, 상기 양자키 요청을 받은 각 양자키 분배노드가 인접하면 직접키를 할당하도록 하고, 인접하지 않으면 간접키를 할당하여 암호화된 데이터를 송수신할 수 있도록 하고,

상기 할당한 직접키 혹은 간접키를 삭제하도록 제어하는 것을 더 포함하는 것을 특징으로 하는 심층 강화학습 기반 최적화 기법을 통한 양자키 자원 관리 시스템.

청구항 7

양자키 자원 관리 에이전트에서, 사용자 네트워크로부터 사용자 노드 간 세션정보를 수신하는 세션정보 수신 단계;

상기 양자키 자원 관리 에이전트에서, 상기 세션정보가 수신되면 복수의 양자키 분배노드로부터 각각 저장 및 관리되는 직접키와 간접키의 수를 포함하는 양자키 분배 네트워크의 상태정보를 수집하는 상태정보 수집 단계;

상기 양자키 자원 관리 에이전트에서, 상기 수신한 세션정보와 상기 수집한 상태정보를 기반으로 상기 양자키 분배 네트워크에 대한 상태데이터를 생성하는 상태데이터 생성 단계; 및

상기 양자키 자원 관리 에이전트에서, 상기 생성한 상태데이터를 심층 강화학습모델에 적용하여 적어도 하나 이상의 양자키 분배노드에 대한 간접키 생성량을 결정하는 간접키 생성량 결정 단계;를 포함하며,

상기 심층 강화학습모델에 대한 심층 강화학습을 통해서 직접키와 간접키를 포함하는 양자키를 할당받지 못하는 세션의 수를 최소화하도록 적어도 하나 이상의 양자키 분배노드에 대한 간접키 생성량을 결정함으로써, 상기 양자키의 관리 자원을 최적화하는 것을 특징으로 하는 심층 강화학습 기반 최적화 기법을 통한 양자키 자원 관리 방법.

청구항 8

청구항 7에 있어서,

상기 양자키 자원 관리 방법은,

상기 양자키 자원 관리 에이전트에서, 상기 간접키 생성량을 결정한 양자키 분배노드와 상기 간접키 생성량에 따라 생성할 간접키를 공유할 또 다른 양자키 분배노드 간의 최단경로를 산출하는 최단경로 산출 단계;

상기 양자키 자원 관리 에이전트에서, 상기 산출한 최단경로 상에 위치하는 상호 인접한 양자키 분배노드의 직접키를 상기 간접키 생성량을 결정한 양자키 분배노드로 제공하여 간접키를 생성하고, 상기 간접키 생성에 소모된 상기 직접키를 삭제하도록 제어하는 양자키 분배노드 제어 단계;

상기 양자키 자원 관리 에이전트에서, 상기 간접키 생성량에 따라 각 양자키 분배노드에서 간접키를 모두 생성

하면, 양자키를 할당받지 못한 세션의 수에 대한 마이너스 값을 보상으로 산출하고, 양자키를 할당받지 못한 세션의 수가 0이면 0의 값을 보상으로 산출하는 보상 산출 단계; 및

상기 양자키 자원 관리 에이전트에서, 상기 산출한 보상을 상기 심층 강화학습모델에 역전파(back propagation)시켜, 누적보상이 최대가 되는 액션을 추출하여 상기 간접키 생성량을 결정하도록 상기 심층 강화학습모델을 강화학습하는 심층 강화학습 단계;를 더 포함하는 것을 특징으로 하는 심층 강화학습 기반 최적화 기법을 통한 양자키 자원 관리 방법.

청구항 9

청구항 8에 있어서,

상기 양자키 분배노드 제어 단계는,

상기 세션정보에 따라 사용자 노드로부터 해당 사용자 노드에 해당하는 양자키 분배노드에 대한 양자키 요청이 각각 있으면, 상기 양자키 요청을 받은 각 양자키 분배노드가 인접하면 직접키를 할당하도록 하고, 인접하지 않으면 간접키를 각각 할당하여 암호화된 데이터를 송수신할 수 있도록 하고,

상기 할당한 직접키 혹은 간접키를 삭제하도록 제어하는 것을 더 포함하는 것을 특징으로 하는 심층 강화학습 기반 최적화 기법을 통한 양자키 자원 관리 방법.

청구항 10

청구항 7에 있어서,

상기 상태데이터는,

상기 복수의 양자키 분배노드, 각 양자키 분배노드에서 상기 직접키를 저장 관리하는 직접키 풀 및 상기 간접키를 저장 관리하는 간접키 풀에 따른 노드의 타입, 상기 직접키 혹은 간접키에 관련된 세션의 수, 상기 직접키 및 간접키에 대한 라이프 타임 또는 이들의 조합을 포함하는 노드 특징, 상기 노드 간 연결정보, 상기 노드 간 연결타입 또는 이들의 조합을 포함하는 것을 특징으로 하는 심층 강화학습 기반 최적화 기법을 통한 양자키 자원 관리 방법.

발명의 설명

기 술 분 야

[0001] 본 발명은 심층 강화학습 기반 최적화 기법을 통한 양자키 지원 관리 시스템 및 그 방법에 관한 것으로, 더욱 상세하게는 사용자 네트워크의 현재 상태에 대한 세션정보에 따라 양자키 분배 네트워크를 구성하는 개별 노드의 타입, 양자키를 요구하는 세션의 수, 양자키의 라이프 타임, 각 노드 간 연결정보, 각 노드 간 연결타입을 포함하는 양자키 분배 네트워크의 상태데이터를 강화학습하여 다음 상태에 대한 간접키의 생성량을 결정하여 양자키가 할당되지 못하는 세션을 최소화함으로써 안정적인 통신 상태를 유지할 수 있도록 하는 시스템 및 그 방법에 관한 것이다.

배 경 기 술

[0002] 네트워크 통신의 암호체계는 높은 계산 복잡도로 안정성이 보장되고 있으나, 고도의 계산 성능을 가지는 양자컴퓨터의 출현과 발전은 현재 네트워크 통신의 암호체계에 대한 위협으로 대두되고 있다.

[0003] 이에 따라 안전한 통신을 위해 양자의 불확정성 원리에 기반한 양자 보안 통신에 대한 대중의 관심이 집중되고 있다.

[0004] 따라서 양자키 자원 관리의 중요성은 양자 보안 통신의 성공적인 구현을 위한 필수적이다.

[0005] 기본적으로 양자키는 직접키(direct key)와 간접키(indirect key)로 구성된다. 상기 직접키는 인접하는 노드 간에 사용되는 양자키를 말하며, 간접키는 인접하지 않는 노드(원거리 노드) 간에 사용되는 양자키를 말한다.

[0006] 보통 통신은 서로 인접하지 않은 노드 간에 수행되는 경우의 수가 훨씬 많고 양자키 분배에서 간접키의 생성, 분배 및 처리가 매우 중요하다.

[0007] 이러한 양자키는 각 노드에서 무분별하게 생성, 분배 및 처리되면 이와 관련한 자원이 부족하여 실제 특정 노드

간의 통신시 양자키가 할당되지 않아 블록(block)되는 경우가 발생될 수 있다.

[0008] 이에 따라 양자키 생성을 최적화하여 양자키가 할당되지 않는 세션의 발생을 최소화하는 것이 매우 중요하다.

[0009] 한편, 심층 강화학습은 환경의 상태정보(관찰정보)를 입력으로 하여 출력한 행동(action)의 결과로 보상(reword)을 받아 고도화되는 인공지능 학습의 한 종류이다. 상기 심층 강화학습은 어떤 환경에서 받는 보상을 최대화하는 정책(policy)을 찾는 것을 목표로 한다.

[0010] 즉, 심층 강화학습은 미지의 환경과 상호작용하고 시행착오를 겪어가며 매 상황에서의 선택에 대한 가치를 학습하여 다양한 분야의 문제를 해결하는데 사용될 수 있다.

[0011] 이에 따라 본 발명에서는 양자키 분배(QKD, quantum key distribution) 네트워크의 양자키 생성을 최적화하기 위해 심층강화학습을 적용하여 통신(보안통신)을 안정적으로 유지할 수 있도록 하는 방안을 제안하고자 한다. 이때, 본 발명에서는 양장키에서 간접키의 생성 최적화를 주목적으로 한다.

[0012] 또한 본 발명은 양자키 분배 네트워크의 양자키 분배노드에서 세션에 따른 양자키 분배 네트워크의 환경을 고려하여 각 양자키 분배노드의 간접키의 생성을 최적으로 제어할 수 있도록 하는 방안을 제안하고자 한다.

[0013] 또한 본 발명은 양자키 분배 네트워크를 구성하는 개별 노드의 타입, 양자키를 요구하는 세션의 수, 모든 양자키의 수, 모든 양자키의 라이프 타임, 각 노드 간 연결정보, 연결특징 또는 이들의 조합을 포함하는 양자키 분배 네트워크의 상태데이터를 사용자 네트워크의 현재 세션상태에 따라 생성하고 상기 상태데이터를 심층 강화학습모델에 입력하여 누적 행동가치가 높은 액션(간접키 생성량)을 추출하고, 상기 액션에 따른 보상을 제공하여 강화학습함으로써 실제 양자키 분배 네트워크에서 양자키의 할당을 받지 못하는 세션을 최소화하도록 각 노드에서의 간접키 생성을 최적으로 제어할 수 있도록 하는 방안을 제안하고자 한다.

[0014] 다음으로 본 발명의 기술분야에 존재하는 선행기술에 대하여 간단하게 설명하고, 이어서 본 발명이 상기 선행기술에 비해서 차별적으로 이루고자 하는 기술적 사항에 대해서 기술하고자 한다.

[0015] 먼저 한국공개특허 제2022-0080344호(2022.06.14.)는 양자키 분배에 기반하는 보안강화 전송방법 및 장치에 있어서, 하위 레이어의 네트워크 종단에서 발생할 수 있는 보안 상 허점을 방지하기 위해서, 서비스 레이어의 호출 맥락에 기반하여 양자키와 서비스 호출을 매핑시키고, 이러한 양자키를 사용하여 서비스 데이터를 암호화/복호화함으로써 전송 네트워크의 보안을 향상시키는 양자키 분배에 기반하는 보안강화 전송방법 및 장치에 관한 것이다.

[0016] 그러나 한국공개특허 제2022-0080344호는 사용자 네트워크(서비스 레이어)를 구성하는 애플리케이션 노드에 양자키 분배 네트워크에 생성한 양자키를 제공하여 상기 양자키로 데이터를 암호화하여 전송하고, 수신측에서 복호화할 수 있도록 하는 일반적인 양자키 분배 네트워크에 대한 기술을 기재하고 있을 뿐이다.

[0017] 반면에 본 발명은 양자키 분배 네트워크에서 심층 강화학습을 기반으로 양자키 중 간접키의 수량을 최적으로 결정하여 생성 및 분배함으로써 양자키 자원을 효율적으로 관리할 수 있도록 하는 것으로, 한국공개특허 제2022-0080344호는 이러한 본 발명의 기술적 특징을 기재하거나 시사 혹은 그 어떠한 암시도 없다.

[0018] 또한 한국등록특허 제2287186호(2021.08.02.)는 양자키 할당 우선순위를 고려하는 양자 통신 시스템, 장치 및 방법에 관한 것으로, 복수의 어플리케이션으로부터 양자키 할당 요청이 있으면, 양자키 할당 우선순위 정보를 고려하여 복수의 어플리케이션에 대하여 양자키를 할당하는 양자키 할당 우선순위를 고려하는 양자 통신 시스템, 장치 및 방법에 관한 것이다.

[0019] 즉, 한국등록특허 제2287186호는 양자키를 할당하는 방법에 대한 것만 기재하고 있을 뿐, 양자키의 생성에 대해서는 전혀 기재하고 있지 않다.

[0020] 반면에 본 발명은 심층 강화학습을 기반으로 인접하지 않은 노드에 대한 간접키를 생성하도록 제어할 때, 양자키 할당을 받지 않은 세션이 최소화되도록 상기 간접키의 생성량을 결정하여 상기 양자키를 생성할 수 있도록 제어하는 것으로 한국등록특허 제2287186호는 이러한 본 발명의 기술적 특징을 기재하거나 시사 혹은 그 어떠한 암시도 없다.

발명의 내용

해결하려는 과제

[0021] 본 발명은 상기와 같은 문제점을 해결하기 위해 창안된 것으로서, QKD 네트워크의 상태데이터를 기반으로 액션을 추출하기 위한 심층 강화학습모델을 통해 간접키의 수량을 결정하여 생성 및 공유할 수 있도록 함으로써 양자키를 할당받지 못하는 세션이 발생되는 것을 최소화하는 심층 강화학습 기반 최적화 기법을 통한 양자키 자원 관리 시스템 및 그 방법을 제공하는 것을 그 목적으로 한다.

[0022] 또한 본 발명은 양자키 분배 네트워크를 구성하는 개별 노드의 수, 노드 간 연결정보, 양자키를 요구하는 세션의 수, 모든 양자키의 수, 모든 양자키의 생명주기, 각 노드 간 연결타입을 포함하는 QKD 네트워크의 상태데이터를 강화학습모델에 입력하여 누적 행동가치가 높은 액션(간접키 생성량)을 추출하고 상기 액션에 따른 보상을 제공하여 강화학습함으로써 상기 QKD 네트워크에서 양자키의 할당을 받지 못하는 세션을 최소화하도록 각 노드에서의 간접키 생성을 최적으로 제어할 수 있도록 하는 시스템 및 그 방법을 제안하고자 한다.

과제의 해결 수단

[0023] 본 발명의 일 실시예에 따른 심층 강화학습 기반 최적화 기법을 통한 양자키 자원 관리 시스템은, 복수의 사용자 노드를 포함하는 사용자 네트워크 및 상기 복수의 사용자 노드에 대응하는 복수의 양자키 분배노드와 상기 복수의 양자키 분배노드를 제어하기 위한 양자키 자원 관리 에이전트를 포함하는 양자키 분배 네트워크를 포함하며, 상기 양자키 자원 관리 에이전트는, 심층 강화학습을 통해서 적어도 하나 이상의 양자키 분배노드에 대한 간접키 생성량을 결정함으로써 직접키와 간접키를 포함하는 양자키의 관리 자원을 최적화하는 것을 특징으로 한다.

[0024] 또한 상기 양자키 자원 관리 에이전트는, 타임스텝별로 상기 양자키를 할당받지 못하는 세션의 수를 최소화하도록 상기 적어도 하나 이상의 양자키 분배노드에 대한 간접키의 생성량을 결정하는 것을 특징으로 한다.

[0025] 또한 상기 양자키 자원 관리 에이전트는, 상기 사용자 네트워크로부터 사용자 노드 간 세션정보를 수신하는 세션정보 수신부, 상기 세션정보가 수신되면 상기 복수의 양자키 분배노드로부터 각각 저장 및 관리되는 직접키와 간접키의 수를 포함하는 양자키 분배 네트워크의 상태정보를 수집하는 상태정보 수집부, 상기 수신한 세션정보와 상기 수집한 상태정보를 기반으로 상기 양자키 분배 네트워크에 대한 상태데이터를 생성하는 상태데이터 생성부 및 상기 생성한 상태데이터를 상기 심층 강화학습모델에 적용하여 적어도 하나 이상의 양자키 분배노드에 대한 간접키 생성량을 결정하는 간접키 생성량 결정부를 포함하며, 상기 상태데이터는, 상기 복수의 양자키 분배노드, 각 양자키 분배노드에서 상기 직접키를 저장 관리하는 직접키 풀 및 상기 간접키를 저장 관리하는 간접키 풀에 따른 노드의 타입, 상기 직접키 혹은 간접키에 관련된 세션의 수, 상기 직접키 및 간접키에 대한 라이프 타임 또는 이들의 조합을 포함하는 노드 특징, 상기 노드 간 연결정보, 상기 노드 간 연결타입 또는 이들의 조합을 포함하는 것을 특징으로 한다.

[0026] 또한 상기 양자키 자원 관리 에이전트는, 상기 간접키 생성량을 결정한 양자키 분배노드와 상기 간접키 생성량에 따라 생성할 간접키를 공유할 또 다른 양자키 분배노드 간의 최단경로를 산출하는 최단경로 산출부 및 상기 산출한 최단경로 상에 위치하는 상호 인접한 양자키 분배노드의 직접키를 상기 간접키 생성량을 결정한 양자키 분배노드로 제공하여 간접키를 생성하고, 상기 간접키 생성에 소모된 상기 직접키를 삭제하도록 제어하는 양자키 분배노드 제어부를 더 포함하는 것을 특징으로 한다.

[0027] 또한 상기 양자키 자원 관리 에이전트는, 상기 간접키 생성량에 따라 각 양자키 분배노드에서 간접키를 모두 생성하면, 양자키를 할당받지 못한 세션의 수에 대한 마이너스 값을 보상으로 산출하고, 양자키를 할당받지 못한 세션의 수가 0이면 0의 값을 보상으로 산출하는 보상 산출부 및 상기 산출한 보상을 상기 심층 강화학습모델에 역전파(back propagation)시켜, 누적보상이 최대가 되는 액션을 추출하여 상기 간접키 생성량을 결정하도록 상기 심층 강화학습모델을 강화학습하는 심층 강화학습부를 더 포함하는 것을 특징으로 한다.

[0028] 또한 상기 양자키 자원 관리 에이전트는, 상기 세션정보에 따라 사용자 노드로부터 해당 사용자 노드에 해당하는 양자키 분배노드에 대한 양자키 요청이 각각 있으면, 상기 양자키 요청을 받은 각 양자키 분배노드가 인접하면 직접키를 할당하도록 하고, 인접하지 않으면 간접키를 할당하여 암호화된 데이터를 송수신할 수 있도록 하고, 상기 할당한 직접키 혹은 간접키를 삭제하도록 제어하는 것을 더 포함하는 것을 특징으로 한다.

[0029] 아울러 본 발명의 일 실시예에 따른 심층 강화학습 기반 최적화 기법을 통한 양자키 자원 관리 방법은, 양자키 자원 관리 에이전트에서, 사용자 네트워크로부터 사용자 노드 간 세션정보를 수신하는 세션정보 수신 단계, 상기 양자키 자원 관리 에이전트에서, 상기 세션정보가 수신되면 상기 복수의 양자키 분배노드로부터 각각 저장 및 관리되는 직접키와 간접키의 수를 포함하는 양자키 분배 네트워크의 상태정보를 수집하는 상태정보 수집 단계, 상기 양자키 자원 관리 에이전트에서, 상기 수신한 세션정보와 상기 수집한 상태정보를 기반으로 상기 양자

키 분배 네트워크에 대한 상태데이터를 생성하는 상태데이터 생성 단계 및 상기 양자키 자원 관리 에이전트에서, 상기 생성한 상태데이터를 상기 심층 강화학습모델에 적용하여 적어도 하나 이상의 양자키 분배 노드에 대한 간접키 생성량을 결정하는 간접키 생성량 결정 단계를 포함하며, 상기 심층 강화학습모델에 대한 심층 강화학습을 통해서 직접키와 간접키를 포함하는 양자키를 할당받지 못하는 세션의 수를 최소화하도록 적어도 하나 이상의 양자키 분배노드에 대한 간접키 생성량을 결정함으로써, 상기 양자키의 관리 자원을 최적화하는 것을 특징으로 한다.

[0030] 또한 상기 양자키 자원 관리 방법은, 상기 양자키 자원 관리 에이전트에서, 상기 간접키 생성량을 결정한 양자키 분배노드와 상기 간접키 생성량에 따라 생성할 간접키를 공유할 또 다른 양자키 분배노드 간의 최단경로를 산출하는 최단경로 산출 단계 및 상기 양자키 자원 관리 에이전트에서, 상기 산출한 최단경로 상에 위치하는 상호 인접한 양자키 분배노드의 직접키를 상기 간접키 생성량을 결정한 양자키 분배노드로 제공하여 간접키를 생성하고, 상기 간접키 생성에 소모된 상기 직접키를 삭제하도록 제어하는 양자키 분배노드 제어 단계, 상기 양자키 자원 관리 에이전트에서, 상기 간접키 생성량에 따라 각 양자키 분배노드에서 간접키를 모두 생성하면, 양자키를 할당받지 못한 세션의 수에 대한 마이너스 값을 보상으로 산출하고, 양자키를 할당받지 못한 세션의 수가 0이면 0의 값을 보상으로 산출하는 보상 산출 단계 및 상기 양자키 자원 관리 에이전트에서, 상기 산출한 보상을 상기 심층 강화학습모델에 역전파(back propagation)시켜, 누적보상이 최대가 되는 액션을 추출하여 상기 간접키 생성량을 결정하도록 상기 심층 강화학습모델을 강화학습하는 심층 강화학습 단계를 더 포함하는 것을 특징으로 한다.

[0031] 또한 상기 양자키 분배노드 제어 단계는, 상기 세션정보에 따라 사용자 노드로부터 해당 사용자 노드에 해당하는 양자키 분배노드에 대한 양자키 요청이 각각 있으면, 상기 양자키 요청을 받은 각 양자키 분배노드가 인접하면 직접키를 할당하도록 하고, 인접하지 않으면 간접키를 각각 할당하여 암호화된 데이터를 송수신할 수 있도록 하고, 상기 할당한 직접키 혹은 간접키를 삭제하도록 제어하는 것을 더 포함하는 것을 특징으로 한다.

발명의 효과

[0032] 이상에서와 같이 본 발명의 심층 강화학습 기반 최적화 기법을 통한 양자키 자원 관리 시스템 및 그 방법은 양자키 분배 네트워크에서 인접하지 않은 노드사이에서 이용하기 위한 간접키를 생성할 때, 사용자 네트워크의 세션정보에 따른 양자키 분배 네트워크의 상태데이터를 생성하고, 상기 생성한 상태데이터에 따라 최적 수량의 간접키를 생성함으로써 양자키를 할당받지 못하는 세션을 최소화하여 안정적인 통신 상태를 유지할 수 있도록 하는 효과가 있다.

도면의 간단한 설명

[0033] 도 1은 본 발명의 일 실시예에 따른 심층 강화학습 기반 최적화 기법을 통한 양자키 자원 관리 시스템 및 그 방법을 설명하기 위해 나타낸 개념도이다.

도 2는 본 발명의 일 실시예에 따른 양자키 분배 네트워크를 설명하기 위해 나타낸 도면이다.

도 3은 본 발명의 일 실시예에 따른 사용자 네트워크를 설명하기 위해 나타낸 도면이다.

도 4는 본 발명의 일 실시예에 따른 간접키를 공유하는 방법을 설명하기 위해 나타낸 도면이다.

도 5는 본 발명의 일 실시예에 따른 양자키 분배 네트워크의 구성을 상세히 나타낸 도면이다.

도 6은 본 발명의 일 실시예에 따른 상태데이터 중 노드 특징을 나타낸 도면이다.

도 7은 본 발명의 일 실시예에 따른 상태데이터의 노드 특징을 도시한 도면이다.

도 8은 본 발명의 일 실시예에 따른 상태데이터의 노드 간 연결정보를 도시한 도면이다.

도 9는 본 발명의 일 실시예에 따른 상태데이터의 노드 간 연결특징을 도시한 도면이다.

도 10은 본 발명의 일 실시예에 따른 심층 강화학습모델을 통해 간접키의 생성량을 결정하여 간접키를 생성하는 방법을 설명하기 위해 나타낸 도면이다.

도 11은 본 발명의 일 실시예에 따른 양자키 관리 제어 에이전트의 구성을 나타낸 블록도이다.

도 12는 본 발명의 일 실시예에 따른 심층 강화학습 기반 최적화 기법을 통한 양자키 자원을 관리하는 절차를 나타낸 흐름도이다.

발명을 실시하기 위한 구체적인 내용

[0034] 이하, 첨부한 도면을 참조하여 본 발명의 심층 강화학습 기반 최적화 기법을 통한 양자키 자원 관리 시스템 및 그 방법에 대한 바람직한 실시예를 상세히 설명한다. 각 도면에 제시된 동일한 참조부호는 동일한 부재를 나타낸다. 또한 본 발명의 실시예들에 대해서 특정한 구조적 내지 기능적 설명들은 단지 본 발명에 따른 실시예를 설명하기 위한 목적으로 예시된 것으로, 다르게 정의되지 않는 한, 기술적이거나 과학적인 용어를 포함해서 여기서 사용되는 모든 용어들은 본 발명이 속하는 기술분야에서 통상의 지식을 가진 자에 의해 일반적으로 이해되는 것과 동일한 의미를 가지고 있다. 일반적으로 사용되는 사전에 정의되어 있는 것과 같은 용어들은 관련 기술의 문맥상 가지는 의미와 일치하는 의미를 가지는 것으로 해석되어야 하며, 본 명세서에서 명백하게 정의하지 않는 한, 이상적이거나 과도하게 형식적인 의미로 해석되지 않는 것이 바람직하다.

[0035] 도 1은 본 발명의 일 실시예에 따른 심층 강화학습 기반 최적화 기법을 통한 양자키 자원 관리 시스템 및 그 방법을 설명하기 위해 나타낸 개념도이다.

[0036] 도 1에 도시한 바와 같이, 본 발명의 일 실시예에 따른 심층 강화학습 기반 최적화 기법을 통한 양자키 자원 관리 시스템(10)(이하, 양자키 자원 관리 시스템이라 칭함)은 양자키 분배 네트워크(QKD(quantum key distribution) 네트워크)와 사용자 네트워크를 포함하여 구성되며, 심층 강화학습을 기반으로 양자키의 생성을 최적화하여 양자키 자원을 최적으로 관리하는 기능을 수행한다.

[0037] 상기 사용자 네트워크는 인터넷, 이더넷과 같은 데이터 전송채널을 통해 양자키로 암호화된 데이터를 송수신하는 사용자 노드(혹은 애플리케이션 노드)로 구성된 서비스 레이어를 포함하여 구성된다. 여기서 사용자 노드는 사용자 단말(스마트폰) 혹은 사용자 단말에 설치된 애플리케이션을 지칭할 수 있다.

[0038] 상기 양자키 분배 네트워크는 양자키 분배모듈(QKD 모듈)로 구성된 양자 레이어, 양자키 관리(KM, key management)모듈로 구성된 양자키 관리 레이어 및 양자키 분배 네트워크를 제어하기 위한 심층 강화학습 기반의 양자키 자원 관리 에이전트(100)(이하 양자키 관리 제어 에이전트라 칭함)로 구성된 제어 레이어를 포함하여 구성된다. 각 상기 네트워크는 도 2 내지 도 4를 참조하여 설명하도록 한다.

[0039] 상기 양자키 분배 모듈과 상기 양자키 관리모듈은 양자키 분배노드를 구성하며, 상기 양자키 분배노드는 사용자 노드에 각각 대응하여 구성된다.

[0040] 이때, 사용자 노드는 특정 사용자 노드와의 세션이 연결되어 데이터를 전송하고자 할 때, 사용자 노드에 해당하는 양자키 분배노드의 양자키 관리모듈에 양자키를 요청하여 할당(제공)받으며, 해당 양자키로 데이터를 암호화하여 세션이 연결된 다른 사용자 노드로 전송하여 다른 사용자 노드에서는 공유된 양자키로 상기 수신한 데이터를 복호화하여 사용한다.

[0041] 상기 양자키는 직접키와 간접키로 구성되며, 상기 직접키는 인접한 양자키 분배노드 간에 공유(분배)되며, 간접키는 인접하지 않은 양자키 분배 모드 간에 공유된다.

[0042] 이때, 사용자 노드는 완전연결형(혹은 메시형)으로 연결되도록 구성되나, 양자키 분배노드는 인접한 양자키 분배노드와 연결되도록 구성된다. 여기서 간접키는 직접키를 이용하여 중간의 양자키 분배노드에서 중계되어 공유될 수 있다.

[0043] 예를 들어, 사용자 노드 a에서 사용자 노드 c로 세션 연결을 통한 암호화된 데이터를 전송하고자 하는 경우, 사용자 노드 a에 해당하는 양자키 분배노드 a와 사용자 노드 c에 해당하는 양자키 분배노드 c는 직접 연결되지 않고 중간의 양자키 분배노드 b를 통해 연결된 것으로 가정하자. 이때, 사용자 노드 a에 해당하는 양자키 분배노드 a의 양자키 모듈에서 간접키(Kr)를 생성하여 양자키 관리모듈로 제공하고, 상기 간접키를 양자키 분배노드 b를 경유하여 양자키 분배노드 c와 공유한다.

[0044] 상기 예에서, 양자키 분배노드 a는 양자키 분배노드 b와 공유한 직접키(Ka,b)로 간접키(Kr)를 암호화하여 양자키 분배노드 b로 전송하고, 상기 양자키 분배노드 b는 해당 직접키(Ka,b)로 복호화한 다음, 양자키 분배노드 c와 공유한 직접키(Kb,c)로 다시 암호화하여 양자키 분배노드 c로 전송한다. 이후, 양자키 분배노드 c는 상기 직접키(Kb,c)로 복호화하고 사용자 노드 c의 요청에 따라 간접키(Kr)를 사용자 노드 c로 제공함으로써 사용자 노드 a로부터 수신한 암호화된 데이터를 해당 간접키로 복호화하여 사용할 수 있도록 한다.

[0045] 이때, 양자키 분배 네트워크 내에서의 데이터(즉, 양자키) 송수신은 양자채널로 이루어지며, 사용자 노드 간의 데이터 송수신은 일반적인 이더넷이나 인터넷과 같은 데이터 채널로 이루어지며, 사용자 네트워크(즉, 사용자

노드)와 양자키 분배 네트워크(예: 양자키 분배노드의 양자키 관리모듈)간의 데이터 송수신은 상기 데이터 채널로 이루어질 수 있다.

[0046] 또한 간접키는 양자키 관리모듈을 통해 중계되어 공유될 수 있으며, 직접키는 양자키 분배모듈을 통해 직접 공유되거나, 양자키 관리모듈을 통해 공유될 수 있다.

[0047] 한편, 간접키는 인접하지 않은 양자키 분배노드 간에 생성 및 공유되기 때문에 사용자 네트워크에서 사용자 노드 간의 세션 연결에 대비하여 모든 간접키를 생성하게 되면, 사용자 노드가 무수하게 많아지면 무한대의 간접키는 기하급수적으로 생성될 수 있다.

[0048] 이러한 문제점을 해결하기 위해서 본 발명의 심층 강화학습 기반의 양자키 자원 관리 에이전트(100)(이하 양자키 관리 제어 에이전트라 칭함)는 환경과 상호작용하여 강화학습모델에 대한 강화학습을 수행하여 양자키 분배노드에서의 양자키(더욱 자세하게는 간접키)의 보유량을 최적으로 설정하여 생성 및 분배하도록 제어함으로써 양자키가 할당되지 않는 세션을 최소화하여 양자키 자원을 최적으로 운용하고 통신 상태를 안정적으로 유지할 수 있도록 한다.

[0049] 즉, 양자키 자원 관리 에이전트(100)는 사용자 네트워크의 세션정보(즉, 상태)에 따라 다음 세션정보에 대한 양자키 요구량을 예측하여 각 양자키 분배모드의 양자키(간접키) 생성량을 결정하여 생성 및 공유하도록 제어하는 것이다.

[0050] 도 2는 본 발명의 일 실시예에 따른 양자키 분배 네트워크를 설명하기 위해 나타낸 도면이다.

[0051] 도 2에 도시한 바와 같이, 본 발명의 일 실시예에 따른 양자키 분배 네트워크(G^Q)는 양자키 분배노드의 집합(V^Q)와 양자키 분배노드 간의 링크 집합(E^Q)로 정의된다. 또한 도 2의 N_i^Q는 양자키 분배노드(i)의 인접한 양자키 분배노드(이웃노드)의 집합을 말한다. 예를 들어, N_0^Q = {1, 2}로 표현된다. 즉, 양자키 분배노드 0에 인접한 양자키 분배노드는 1, 2로 표시된 양자키 분배노드인 것이다. 또 다른 예로, N_3^Q은 {1, 4, 5}로 표현된다.

[0052] 또한 양자키 분배 네트워크는 양자키 분배노드 이외에 양자키를 저장하여 관리하기 위한 키 풀(key pool)(미도시)을 포함하여 구성된다. 상기 키 풀은 직접키를 위한 직접키 풀(DKP, direct key pool)과 간접키 풀(indirect key pool)을 포함하여 구성된다.

[0053] 직접키 풀은 $DKP_i^{(i,j)}$ 및 $DKP_j^{(i,j)}$로 표현되며, 인접한 두개의 양자키 분배노드에서 각각 유지 관리된다. 즉, 직접키 풀은 인접한 두개의 양자키 분배노드에서 상호 공유된 직접키를 관리하기 위한 노드를 말한다. 상기 직접키는 인접한 두개의 양자키 관리노드에서 각각 생성된 후, 공유되며 이와 동시에 각 직접키 풀에 저장 및 관리된다.

[0054] 상기 간접키 풀은 사용자 네트워크의 사용자 노드 간 링크에 따른 간접키를 관리하기 위한 것으로 상기 사용자 노드에 해당하는 양자키 분배노드에서 각각 유지 관리된다. 상기 간접키 풀은 도 3을 참조하여 상세히 설명하도록 한다.

[0055] 도 3은 본 발명의 일 실시예에 따른 사용자 네트워크를 설명하기 위해 나타낸 도면이다.

[0056] 도 3에 도시한 바와 같이, 본 발명의 일 실시예에 따른 사용자 네트워크($\hat{G}^Q$)는 양자키 분배노드에 해당하는 사용자 노드의 집합($\hat{V}^Q$)와 사용자 노드 간의 링크(사용자 네트워크 링크) 집합($\hat{E}^Q$)으로 정의된다.

[0057] 이때, 양자키 분배 네트워크의 간접키 풀은 사용자 네트워크 링크($\hat{i}$,$\hat{j}$)에 따라 해당하는 사용자 노드의 양자키 분배노드에서 간접키를 각각 저장 및 관리하기 위한 것으로, 간접키 풀은 $IPK_i^{(\hat{i},\hat{j})}$ 및 $IPK_j^{(\hat{i},\hat{j})}$로 표현된다.

[0058] 상기 간접키는 네트워크 링크를 통해 사용자가 노드 간의 세션이 연결될 때, 각 사용자 노드에 해당하는 양자키 분배노드가 이웃하지 않는 경우에 생성되어 공유된다. 본 발명에서 간접키는 양자키 분배 네트워크의 경로에 따라 공유된다.

[0059] 상기 간접키를 공유하는 것은 도 4를 참조하여 상세히 설명하도록 한다.

[0060] 도 4는 본 발명의 일 실시예에 따른 간접키를 공유하는 방법을 설명하기 위해 나타낸 도면이다.

[0061] 도 4에 도시한 바와 같이, 본 발명의 일 실시예에 따른 양자키 자원 관리 에이전트(100)는 현재 사용자 네트워크의 현재 세션정보에 따른 양자키 분배 네트워크의 상태데이터를 생성하고, 상기 생성한 상태데이터를 심층 강화학습모델에 입력하여, 적어도 하나 이상의 양자키 분배노드의 간접키 생성량을 결정하여 생성 및 분배(공유)할 수 있도록 한다.

[0062] 상기 세션($\hat{S}^{Q}$)는 ($\hat{i}$,$\hat{j}$, st)로 정의되며, st는 포아송 분포(poisson distribution)에 따른 세션도착시간을 말한다. 세션성립에 의한 데이터전송에 따라 하나의 직접키 혹은 간접키가 소모된다. 이때, 사용자 노드 간에는 직접키와 간접키가 해당하는 각 양자키 분배노드에 의해 할당된다.

[0063] 이때, 두 개의 양자키 분배노드 간의 간접키를 공유하기 위해서 양자키 자원 관리 에이전트(100)는 다음의 [수학식 1]에 따라 직접키 풀의 직접키 수와 직접키의 임계값 수, 결정(요구)된 간접키의 수에 따른 링크의 가중치를 계산하고, 다익스트라(dijkstra) 알고리즘에 따른 가중치의 합이 최소가 되는 최단경로를 탐색하여 상기 간접키를 공유할 수 있도록 제어한다.

[0064] [수학식 1]

$$w^{(i,j)} = \begin{cases} 100 + \dfrac{1000}{\left|DKP_i^{(i,j)}\right| - N_{th}^{d} - \widehat{N_{ij}}}, & \left|DKP_i^{(i,j)}\right| > N_{th} + \widehat{N_{ij}} \\ \infty, & \left|DKP_i^{(i,j)}\right| \le N_{th} + \widehat{N_{ij}} \end{cases}$$

[0065]

[0066] 여기서, N_{th}^{d}는 링크(i,j)에서, 남아있는 직접키의 임계값 수를 말한다. 예를 들어, 직접키 풀에 대한 직접키의 임계값 수(즉, 용량)가 10이고 현재 직접키의 수가 8이면 남아 있는 직접키의 임계값 수는 2를 말한다. 또한 N_{th}는 남아있는 간접키의 임계값 수를 말하며, $\widehat{N}_{i,j}$는 결정(요구된) 간접키의 수를 말한다.

[0067] 즉, 양자키 자원 관리 에이전트(100)는 최단경로 상에 위치하는 양자키 분배노드 간의 직접키를 활용하여 특정 양자키 분배노드에서 생성한 간접키를 암복호화를 통하여 목적지 양자키 분배노드로 전달(경유)시켜 공유할 수 있도록 하는 것이다.

[0068] 예를 들어, 도 4에 도시한 것과 같이, 양자키 분배노드 0에서 간접키를 생성하여 양자키 분배노드 5로 공유하고자 할 때, 도 4에서 최단경로(p)가 {(0,1), (1,3), (3,5)}로 결정 난 경우, 최단경로 상의 양자키 분배노드 0에서 양자키 분배노드 1과 공유한 직접키로 암호화하여 양자키 분배노드 1로 전송하고, 양자키 분배노드 1은 양자키 분배노드 0과 공유한 직접키로 복호화하여 양자키 분배노드 3과 공유한 양자키로 암호화하여 양자키 분배노드 3으로 전송하며, 상기 양자키 분배노드 3은 양자키 분배노드 3과 공유한 직접키로 복호화하여 다시 양자키 분배노드 5와 공유한 직접키로 암호화하여 양자키 분배노드 5로 전송함으로써 양자키 분배노드 0과 양자키 분배노드 5 간의 간접키를 공유하도록 한다.

[0069] 이때, 간접키는 양자키 분배노드 0과 양자키 분배노드 5에서 각각 관리하는 간접키 풀에 저장되며, 각 양자키 분배노드에서 사용된 직접키는 암복호화에 각각 소비되어 삭제된다.

[0070] 도 5는 본 발명의 일 실시예에 따른 양자키 분배 네트워크의 구성을 상세히 나타낸 도면이다.

[0071] 도 5에 도시한 바와 같이, 본 발명의 일 실시예에 따른 양자키 분배 네트워크는 복수의 양자키 분배노드, 복수의 직접키 풀과 간접키 풀 및 양자키 자원 관리 에이전트(100)를 포함하여 구성된다.

[0072] 상기 양자키 분배노드는 양자키 생성모듈(QKD 모듈), 양자키 관리모듈을 포함하여 구성된다. 이때, 양자키 생성모듈은 양자 레이어에 구성되며, 양자키 관리모듈과 직접키 풀 및 간접키 풀을 양자키 관리 레이어에 구성된다. 또한 양자키 자원 관리 에이전트(100)는 제어 레이어에 구성된다.

[0073] 상기 양자키 자원 관리 에이전트(100)는 사용자 네트워크로부터 사용자 노드에 대한 세션정보를 수신하고, 상기 현재 타임스텝에서 세션정보에 따른 양자키 분배 네트워크의 상태데이터를 심층 강화학습모델에 입력하여 다음 타임스텝에 따른 적어도 하나 이상의 양자키 분배노드에 대한 간접키 생성량을 결정하고, 상기 간접키를 다른

양자키 분배노드에 공유하기 위한 최단경로에 따라 공유할 수 있도록 제어한다.

[0074] 상기 심층 강화학습모델은 타임스텝(time step)별로 누적보상이 최대가 되는 적어도 하나 이상의 양자키 분배노드에 대한 간접키 생성량(액션)을 추출하도록 심층 강화학습된 것이다.

[0075] 상기 양자키 자원 관리 에이전트(100)는 액션에 따라 간접키를 생성하도록 제어하며, 간접키가 모두 생성된 경우, 모든 세션에 대한 양자키의 할당여부에 따른 소정의 보상을 심층 강화학습모델에 적용한다.

[0076] 즉, 간접키를 모두 생성한 경우, 양자키 자원 관리 에이전트(100)는 양자키를 할당받지 못한 세션이 존재하면 상기 양자키를 할당받지 못한 세션의 수에 대한 마이너스 값의 보상을 심층 강화학습모델에 적용하거나, 양자키를 할당받지 못한 세션의 수가 0이면 0의 보상을 심층 강화학습모델에 적용하여 고도화함으로써 행동가치가 높은 액션을 추출하도록 하는 것이다.

[0077] 간접키를 모두 생성하지 못하는 상황에서도 세션이 존재하지 않는다면, 에이전트는 양자키를 생성하지 않아도 되는 상황이므로, 이러한 상황을 학습하는 것이므로 0의 보상을 준다. 결국, 에이전트는 간접키의 생성 여부가 아닌 양자키를 할당받지 못한 세션의 수로만 보상을 산출하고, 그 보상의 범위는 [-무한대, 0]이다. 여기서 무한대는 이론적인 값이며, 실질적으로는 양자키를 할당받지 못한 세션의 수에 해당하는 값이다.

[0078] 상기 심층 강화학습은 사용자 네트워크와 양자키 분배 네트워크를 가상화한 상태에서 심층 강화학습을 수행한 후, 실제 환경에 적용되거나, 상기 실제 환경에서 심층 강화학습을 수행한 후, 실제 환경에 적용될 수 있다.

[0079] 상기 상태데이터는 노드 특징, 노드 간 연결정보, 노드 간 연결특징을 포함하는 것으로, 도 6 내지 도 8을 참조하여 상세히 설명하도록 한다.

[0080] 도 6은 본 발명의 일 실시예에 따른 상태데이터 중 노드 특징을 나타낸 도면이다. 본 발명에서 상태데이터는 심층 강화학습에 이용되거나, 혹은 각 양자키 분배노드의 간접키 생성량을 결정하기 위해 생성된다.

[0081] 본 발명은 양자키 자원을 최적으로 관리하기 위해 강화학습을 위한 상태데이터의 정의가 필요하다. 양자키 자원 관리 에이전트(100)는 사용자 네트워크의 세션정보에 따른 양자키 분배 네트워크의 상태데이터를 생성함으로써 심층 강화학습모델을 위한 입력을 생성한다.

[0082] 상기 상태데이터(St)는 타임스텝 t에서의 노드 특징(NFt), 노드 간 연결정보(EIt), 연결특징(EFt)을 포함하여 구성된다.

[0083] 상기 노드 특징은 행동(action) 적용 여부(actin required), 노드 타입(node type), 간접키와 직접키와 관련된 세션의 수(number of sessions), 간접키 및 직접키의 라이프 타임(life time of keys)을 포함하여 구성된다. 상기 노드 특징은 도 7을 참조하여 상세히 설명하도록 한다.

[0084] 노드 간 연결정보는 타임스텝 t에서 노드 간 연결을 인덱스로 표시한 것이며, 연결특징은 노드 간 타임스텝 t에서 노드 간 에지의 타입을 인덱스로 표시한 것이다. 상기 노드 간 연결정보 및 연결특징은 도 8 및 도 9를 참조하여 상세히 설명하도록 한다.

[0085] 도 7은 본 발명의 일 실시예에 따른 상태데이터의 노드 특징을 도시한 도면이다.

[0086] 도 7에 도시한 바와 같이, 본 발명의 일 실시예에 따른 노드 특징은 노드 인덱스(node index)와 특징 인덱스(feature index)의 행렬로 구성된다.

[0087] 상기 노드 인덱스는 양자키 분배 네트워크를 구성하는 양자키 분배노드, 직접키 풀, 간접키 풀을 각각 나타낸다. 상기 노드 인덱스는 각 노드에 부여된 번호 혹은 식별자로 구성될 수 있다.

[0088] 상기 특징 인덱스 0은 각 노드에 행동 적용여부를 나타낸다. 0은 행동이 적용되지 않음을 나타내며, 1은 행동이 적용됨을 나타낸다. 본 발명에서는 각 양자키 분배노드에 대한 간접키의 생성량만을 제어하므로 간접키의 생성 및 간접키에 의한 세션처리 후, 라이프 타임이 만료된 양자키의 삭제, 직접키의 생성에 대한 처리 및 직접키에 대한 세션처리가 진행된다.

[0089] 상기 특징 인덱스 1은 노드의 종류를 나타낸다. 0은 양자키 분배노드, 1은 직접키 풀, 2는 간접키 풀을 나타낸다. 특징 인덱스 1의 값은 심층 강화학습 동안 항상 동일한 값을 유지한다.

[0090] 상기 특정 인덱스 2는 간접키를 요구하는 세션의 수를 나타낸다. 또한 특정 인덱스 3 내지 10은 간접키, 직접키를 포함하는 모든 양자키의 라이프 타임을 나타낸다. 모든 양자키의 라이프타임을 나타내기 위해서 키 풀의 용량과 동일한 수의 특징 인덱스를 구성한다. 특징 인덱스 3 내지 10에 해당하는 각 열은 타임스텝 t에서 양자키

하나에 대한 남아 있는 라이프 타임을 각각 나타낸다.

[0091] 직접키 풀에 저장된 양자키의 총수(Mdk)(임계값)는 정해져 있으며, 라이프 타임이 종료되면 직접키 풀에서 제거되며 주기적으로 생성된다.

[0092] 또한 직접키는 사용자 노드 간에 세션이 연결되었을 때, 각 사용자 노드에 대응하는 양자키 분배노드가 인접한 것이면 소비되어 삭제된다.

[0093] 또한 간접키를 생성할 때, 최단경로 상에 위치하는 상호 인접한 양자키 분배노드에 대해서 생성한 직접키를 소비하여 생성하므로, 해당 직접키는 삭제된다.

[0094] 도 8은 본 발명의 일 실시예에 따른 상태데이터의 노드 간 연결정보를 도시한 도면이다.

[0095] 도 8에 도시한 바와 같이, 본 발명의 일 실시예에 따른 노드 간 연결정보는 양자키 분배 네트워크에서 각 노드에 대한 에지들의 연결정보를 인덱스로 표시한 것을 말한다.

[0096] 예를 들어, 인덱스 0 내지 5를 양자키 분배노드로 인덱싱하고, 인덱스 6 내지 13을 직접키 풀로 인덱싱하며, 인덱스 14 내지 20을 간접키 풀로 인덱싱한다고 가정하면, 양자키 분배노드 간 연결정보, 양자키 분배노드와 직접키 풀 간의 연결정보 및 양자키 분배노드와 간접키 풀 간의 연결정보를 도 8에 도시한 것과 같이 인덱스로 표현할 수 있다.

[0097] 도 9는 본 발명의 일 실시예에 따른 상태데이터의 노드 간 연결특징을 도시한 도면이다.

[0098] 도 9에 도시한 바와 같이, 본 발명의 일 실시예에 따른 노드 간 연결특징은 노드 간 연결되는 에지의 타입을 의미하는 것으로, 양자키 분배노드 간의 연결(에지)은 1, 양자키 분배노드와 직접키 풀 간의 연결은 2 및 양자키 분배노드와 간접키 풀 간의 연결은 3으로 표시된다.

[0099] 도 10은 본 발명의 일 실시예에 따른 심층 강화학습모델을 통해 간접키의 생성량을 결정하여 간접키를 생성하는 방법을 설명하기 위해 나타낸 도면이다.

[0100] 도 10에 도시한 바와 같이, 본 발명의 일 실시예에 따른 심층 강화학습모델을 통해 간접키의 생성량을 결정하기 위해서, 양자키 자원 관리 에이전트(100)는 사용자 네트워크로부터 세션연결에 따른 세션정보가 수신되면, 현재 타임스텝 t에서 양자키 분배 네트워크의 상태정보를 수집하여 상기 양자키 분배 네트워크에 대한 상태데이터를 생성한다.

[0101] 상기 세션정보는 사용자 노드 간의 세션연결에 대한 출발지 사용자 노드 및 목적지 사용자 노드의 정보(예: 식별자 혹은 IP주소 등)를 포함하여 구성될 수 있다.

[0102] 또한 양자키 분배 네트워크의 상태정보는 현재 타임스텝에서 양자키 분배노드 간에 생성되어 공유한 간접키의 수($NI_t^{(\hat{i},\hat{j})}$)와 양자키 분배노드 간에 생성되어 공유한 직접키의 수($ND_t^{(k,l)}$)를 포함한다. 여기서, $NI_t^{(\hat{i},\hat{j})}$는 사용자 노드 $\hat{i}$에 대응하는 양자키 분배노드 i에서 관리하는 $IPK_i^{(\hat{i},\hat{j})}$(간접키 풀)의 간접키의 수를 말하며, $ND_t^{(k,l)}$는 양자키 분배노드 k에서 관리하는 직접키 풀의 직접키의 수를 말한다.

[0103] 이때, 양자키 자원 관리 에이전트(100)는 각 양자키 분배노드별로 간접키의 수와 직접키의 수(상태정보)와 세션정보를 기반으로 현재 타임스텝에서의 상태데이터를 생성한다. 이때, 각 양자키 분배노드별로 직접키의 수와 간접키의 수를 합(σ)하여 상태데이터를 생성할 수 있다.

[0104] 또한 양자키 자원 관리 에이전트(100)는 심층 강화학습모델에 생성한 상태데이터를 입력하여 액션을 추출한다. 상기 액션은 현재 타임스텝에서의 세션상태에 따른 양자키 분배 네트워크의 상태에서 다음의 세션상태에 따른 양자키 분배 네트워크의 상태를 예측하여 적어도 하나 이상의 양자키 분배노드별 간접키 생성량($NGI_t^{(\hat{i},\hat{j})}$)을 포함하여 구성된다. 즉, 미리 간접키를 생성하여 안정적인 통신상태를 유지할 수 있도록 하는 것이다.

[0105] 이때, 양자키 자원 관리 에이전트(100)는 양자키 분배노드에서 현재 타임스텝 t에서 세션이 연결된 사용자 노드에 해당하는 각 양자키 분배노드가 인접한 것이면, 사용자 노드의 요청과 사용자 노드의 출발지 및 목적지에 따른 직접키를 제공하도록 제어하여 소비(즉, 직접세션(direction session)생성을 지원하며, 소비된 직접키의 수($NCD_t^{(i,j)}$)에 따라 직접키 풀에 포함된 직접키의 개수를 갱신한다.

[0106] 또한 양자키 자원 관리 에이전트(100)은 간접키 생성량이 결정되면 양자키 분배노드 간의 최단경로를 산출하여 최단경로 상에 위치하는 인접한 적어도 하나 이상의 두 양자키 분배노드에 대해 생성된 직접키를 출발지 사용자 노드에 대응하는 양자키 분배노드로 제공하여 간접키를 생성하고 최단경로를 통해 목적지 사용자 노드에 대응하는 양자키 분배노드로 전송(공유)할 수 있도록 한다.

[0107] 즉, 최단경로가 산출되면 현재 타임스텝 t에서 간접키를 생성을 지원하기 위한 직접키의 소비량($NPD_t^{(i,j)}$)이 결정되며, 각 양자키 분배노드는 간접키 생성 지원을 위해 소비된 직접키를 직접키 풀에서 삭제한다.

[0108] 또한 양자키 자원 관리 에이전트(100)는 현제 타임스텝 t에서 세션이 연결된 사용자 노드에 해당하는 각 양자키 분배노드가 인접하지 않는 것이면, 각 사용자 노드의 요청과 사용자 노드의 출발지 및 목적지에 따른 간접키를 제공하도록 각 양자키 분배노드를 제어하여 간접키를 제공(소비)하도록 제어하여 간접키를 통한 데이터 송수신을 지원(즉, 간접세션 생성지원)한다. 이때, 각 양자키 분배노드는 간접세션 생성지원을 위해 소비된 간접키의 수($NCI_t^{(\hat{i},\hat{j})}$)에 따라 간접키 풀에서 간접키를 제거하고 직접키의 개수를 갱신한다. 상기 도 10에 나타낸 $NEI_t^{(\hat{i},\hat{j})}$ 는 제거한 간접키의 수를 말하며, $NGD_t^{(i,j)}$ 는 주기에 따라 생성된 직접키의 수를 말한다.

[0109] 이때, 심층 강화학습모델을 통해 결정한 간접키의 생성량은 간접키 풀에 남아 있는 간접키의 개수와 합하였을 때, 간접키 풀의 최대용량을 초과하지 못한다.

[0110] 또한 양자키 자원 관리 에이전트(100)는 간접키의 생성량을 결정하여 간접키 생성을 제어하였을 때, 간접키가 모두 생성되면, 모든 세션에 대한 양자키의 할당 여부에 따라 보상을 산출하여 심층 강화학습모델에 적용하여 심층 강화학습을 수행한다. 즉, 산출한 보상을 심층 강화학습모델에 역전파(back propagation)시켜 누적보상(합)이 최대가 되는 액션을 추출하도록 강화학습을 수행한다.

[0111] 상기 보상(r_t)은 간접키 생성량에 따라 모든 간접키가 생성된 후, 양자키를 할당받지 못한 세션의 수에 대한 마이너스 값(-(SBDt + SBIt))으로 산출되며, 상기 양자키를 할당받지 못한 세션의 수가 0이면 0의 값으로 산출된다.

[0112] 상기 SBDt는 타임스텝 t에서 직접키를 할당받지 못한 세션의 수를 말하며, 상기 SBIt는 간접키를 할당받지 못한 세션의 수를 말한다.

[0113] 즉, 심층 강화학습모델은 강화학습을 수행하는 동안 발생되는 보상의 누적합이 최대가 되도록 강화학습되므로, 양자키를 할당받지 못한 세션의 수에 (-)를 붙여 누적합이 0에 최대한 가까운 액션을 추출하도록 구성되는 것이다.

[0114] 상기에서 설명한 것과 같이, 심층 강화학습모델은 다음의 [수학식 2]와 같이 타임스텝 t에서, 간접키 풀의 최대용량(Mik)에서 간접키 풀에 남아있는 간접키의 수($NI_t^{(\hat{i},\hat{j})}$)를 차감한 결과를 초과하지 않는 범위 내에서 간접키의 생성량을 결정하여 미리 생성함으로써, 다음 상태에서 양자키를 할당받지 못하는 세션의 수를 최소화할 수 있도록 하는 것이다.

[0115] [수학식 2]

[0116] $$\textbf{decide } NGI_t^{(\hat{i},\hat{j})} \in \left\{0, \ldots, M_{ik} - NI_t^{(\hat{i},\hat{j})}\right\} \textbf{ to minimize } \lim_{T\to\infty} \frac{\sum_{t=0}^{T} SB_t}{T}$$

[0117] 여기서, SBt는 직접키와 간접키를 포함하는 양자키를 할당받지 못한 세션(블록된 세션, session block)의 수를 말한다.

[0118] 도 11은 본 발명의 일 실시예에 따른 양자키 관리 제어 에이전트의 구성을 나타낸 블록도이다.

[0119] 도 11에 도시한 바와 같이, 본 발명의 일 실시예에 따른 양자키 자원 관리 에이전트(100)는 세션정보 수신부(110), 상태정보 수집부(120), 상태데이터 생성부(130), 간접키 생성량 결정부(140), 최단경로 산출부(150), 양자키 분배노드 제어부(160), 보상 산출부(170) 및 심층 강화학습부(180)를 포함하여 구성된다.

[0120] 상기 세션정보 수신부(110)는 사용자 네트워크로부터 사용자 노드 간 세션정보를 수신한다. 상기 세션정보는 출발지, 목적지에 따른 사용자 노드의 정보를 포함할 수 있다.

[0121] 상기 상태정보 수집부(120)는 상기 세션정보가 수신되면 현재 타임스텝에서의 양자키 분배 네트워크의 상태정보를 수집한다. 상기 상태정보는 개별 양자키 분배노드에 요청하여 수집하거나, 혹은 양자키 분배노드에서 양자키의 소모, 생성에 따른 변경되면 상기 상태정보를 제공받아 수집될 수 있다.

[0122] 상기 상태정보는 각 양자키 분배노드에서 각각 관리하는 직접키 풀, 양자키 풀에 저장된 직접키 및 간접키의 수를 포함하여 구성된다. 이때, 상기 상태정보는 각 양자키 분배노드의 식별정보(ID), 각 직접키 및 간접키의 라이프타임을 더 포함하여 구성될 수 있다.

[0123] 상기 상태데이터 생성부(130)는 상기 수신한 세션정보와 상기 수집한 상태정보에 따라 양자키 분배 네트워크의 상태데이터를 생성한다.

[0124] 상기 상태데이터는 도 6 및 도 7을 참조하여 설명하였으므로 여기서는 생략한다.

[0125] 상기 간접키 생성량 결정부(140)는 생성한 상태데이터를 심층 강화학습모델에 적용하여 적어도 하나 이상의 양자키 분배노드에 대한 간접키의 생성량을 결정한다. 상기 간접키의 생성량은 상기 상태데이터를 심층 강화학습모델에 입력하여 간접키 생성량을 나타내는 액션을 추출함으로써 수행된다.

[0126] 상기 최단경로 산출부(150)는 상기 적어도 하나 이상의 양자키 분배노드에 대한 간접키 생성량을 결정하면 상기 생성될 간접키를 공유한 또 다른 양자키 분배노드에 대한 최단경로를 산출한다. 상기 최단경로 산출부(150)는 각 양자키 분배노드 간의 연결(에지)에 대한 가중치를 계산하고 다익스트라(dijkstra) 알고리즘을 활용하여 상기 가중치가 최소가 되는 최단경로를 산출한다.

[0127] 상기 양자키 분배노드 제어부(160)는 상기 산출한 최단경로에 위치하는 양자키 분배노드에 대해서 생성한 직접키를 상기 간접키 생성량을 결정한 양자키 분배노드에 제공하여 간접키를 생성하도록 제어한다.

[0128] 또한 양자키 분배노드 제어부(160)는 각 양자키 분배노드를 제어하여 간접키 생성에 의해 소비된 직접키를 삭제할 수 있도록 한다.

[0129] 또한 양자키 분배노드 제어부(160)는 각 양자키 분배노드에서, 사용자 노드로부터 양자키 요청이 있는 경우, 세션정보에 따라 사용자 노드에 대응하는 각 양자키 분배노드가 인접해 있으면 직접키를 제공하도록 하고, 인접해 있지 않으면 간접키를 제공하도록 제어한다.

[0130] 이때, 직접키와 간접키는 라이프타임이 얼마 남지 않은 것을 제공하도록 제어할 수 있다. 물론 제공되는 직접키와 간접키는 상호 공유된 것임은 자명하다.

[0131] 상기 보상 산출부(170)는 간접키 생성량에 따라 적어도 하나 이상의 양자키 분배노드에서 간접키를 생성하였을 때, 각 세션에 대한 양자키의 할당 여부에 따라 보상을 산출한다. 상기 보상을 산출하는 것은 도 10을 참조하여 설명하였으므로 여기서는 생략한다.

[0132] 상기 심층 강화학습부(180)는 산출한 보상을 심층 강화학습모델에 역전파시켜 업데이트함으로써 상기 심층 강화학습모델에 대한 강화학습을 수행한다.

[0133] 상기 심층 강화학습은 시뮬레이션 환경에서 충분히 수행되어 실제 환경에 적용될 수 있으며, 실제 환경에서도 지속적으로 수행되어 점진적으로 고도화될 수 있다.

[0134] 도 12는 본 발명의 일 실시예에 따른 심층 강화학습 기반 최적화 기법을 통한 양자키 자원을 관리하는 절차를 나타낸 흐름도이다.

[0135] 도 12에 도시한 바와 같이, 본 발명의 일 실시예에 따른 심층 강화학습 기반 최적화 기법을 통한 양자키 자원을 관리하는 절차는 우선, 양자키 자원 관리 에이전트(100)는 사용자 네트워크로부터 사용자 노드 간 세션정보를 수신하는 세션정보 수신 단계를 수행한다(S110).

[0136] 상기 세션정보는 출발지와 목적지에 따른 사용자 노드 간의 데이터 전송을 위한 세션연결을 말한다.

[0137] 다음으로 양자키 자원 관리 에이전트(100)는 세션정보가 수신(S120)되면, 양자키 분배 네트워크의 상태정보를 수집하는 상태정보 수신 단계를 수행한다(S130).

[0138] 상기 상태정보는 양자키의 변경이 있을 때마다 수집되거나, 혹은 양자키 자원 관리 에이전트(100)의 요청에 따라 각 양자키 분배노드로부터 수집될 수 있다.

[0139] 다음으로 양자키 자원 관리 에이전트(100)는 세션정보와 상태정보를 이용하여 양자키 분배 네트워크의 상태데이

터를 생성하는 상태데이터 생성 단계를 수행한다(S140).

[0140] 상기 상태데이터는 양자키 분배 네트워크를 구성하는 각 노드에 대한 노드 특징, 노드 간 연결정보 및 노드 간 연결특징을 포함함은 상술한 바와 같다.

[0141] 다음으로 양자키 자원 관리 에이전트(100)는 상태데이터를 심층 강화학습모델에 적용하여 적어도 하나 이상의 양자키 분배노드에 대한 간접키 생성량을 결정하는 간접키 생성량 결정 단계를 수행한다(S150).

[0142] 상기 심층 강화학습모델은 현재의 상태데이터, 다음의 상태데이터, 액션에 따른 보상을 심층 강화학습하여 생성된 것으로, 현재의 상태데이터로부터 누적보상이 최대인 액션(간접키 생성량)을 추출하도록 구성된다.

[0143] 다음으로 양자키 자원 관리 에이전트(100)는 간접키 생성량을 결정한 양자키 분배노드와 상기 간접키 생성량에 따라 생성한 간접키를 공유할 양자키 분배노드 간의 최단경로를 산출하는 최단경로 산출 단계를 수행한다(S160).

[0144] 상기 최단경로를 산출하는 것은 다익스트라(dijkstra) 알고리즘을 통해 수행됨은 상술한 바와 같다.

[0145] 다음으로 양자키 자원 관리 에이전트(100)는 최단경로 상에 위치하는 양자키 분배노드 간에 생성한 직접키를 간접키 생성량을 결정한 양자키 분배노드로 제공하여 간접키를 생성하도록 제어하는 양자키 분배노드 제어 단계를 수행한다(S170).

[0146] 이때, 양자키 분배노드 제어 단계는 각 양자키 분배노드를 제어하여 간접키 생성에 사용되는 직접키를 삭제하도록 제어한다.

[0147] 또한 양자키 분배노드 제어 단계는 사용자 노드의 요청에 따라 양자키 분배노드를 제어하여 직접키 혹은 간접키를 사용자 노드에 할당(제공)할 수 있도록 제어하는 것을 더 포함한다.

[0148] 다음으로 양자키 자원 관리 에이전트(100)는 간접키 생성결과에 따른 보상을 산출하는 보상 산출 단계를 수행한다(S180).

[0149] 상기 보상은 결정한 간접키를 모두 생성하면, 모든 세션에 대한 양자키의 할당여부에 따라 산출된다.

[0150] 다음으로 양자키 자원 관리 에이전트(100)는 산출한 보상을 심층 강화학습모델에 역전파시켜 업데이트(강화학습)하는 심층 강화학습 단계를 수행한다(S190).

[0151] 심층 강화학습 단계는 보상을 심층 강화학습모델에 적용하여 누적보상이 최대인 액션을 추출하도록 상기 심층 강화학습모델에 대한 강화학습을 수행하는 것이다.

[0152] 이상에서 설명한 바와 같이, 본 발명은 심층 강화학습을 통해서 현재의 상태로부터 간접키의 생성량을 결정함으로써 최적으로 양자키 자원을 관리할 수 있도록 함과 동시에 안정적인 통신을 유지할 수 있도록 하는 효과가 있다.

[0153] 상기에서는 본 발명에 따른 바람직한 실시예를 위주로 상술하였으나, 본 발명의 기술적 사상은 이에 한정되는 것은 아니며 본 발명의 각 구성요소는 동일한 목적 및 효과의 달성을 위하여 본 발명의 기술적 범위 내에서 변경 또는 수정될 수 있을 것이다.

[0154] 또한, 이상에서는 본 발명의 바람직한 실시예에 대하여 도시하고 설명하였지만, 본 발명은 상술한 특정의 실시예에 한정되지 아니하며, 청구범위에서 청구하는 본 발명의 요지를 벗어남이 없이 당해 발명이 속하는 기술분야에서 통상의 지식을 가진 자에 의해 다양한 변형 실시가 가능한 것은 물론이고, 이러한 변형 실시들은 본 발명의 기술적 사상이나 전망으로부터 개별적으로 이해되어서는 안 될 것이다.

부호의 설명

[0155] 10: 심층 강화학습 기반 최적화 기법을 통한 양자키 자원 관리 시스템

100: 양자키 자원 관리 에이전트 110: 세션정보 수신부

120: 상태정보 수집부 130: 상태데이터 생성부

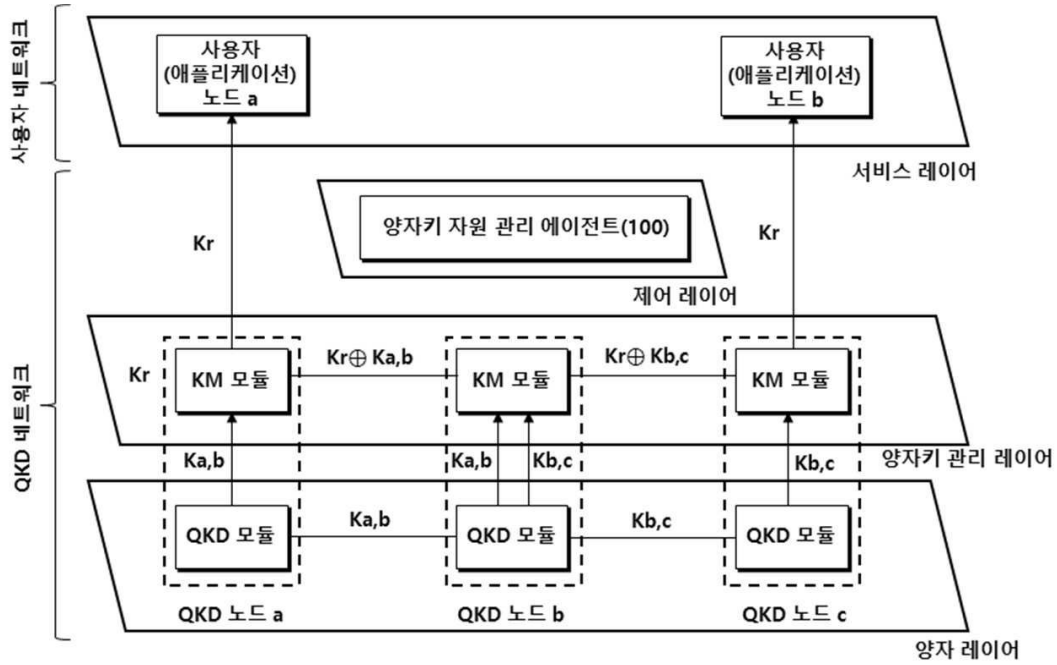
140: 간접키 생성량 결정부 150: 최단경로 산출부

160: 양자키 분배노드 제어부 170: 보상 산출부

180: 심층 강화학습부

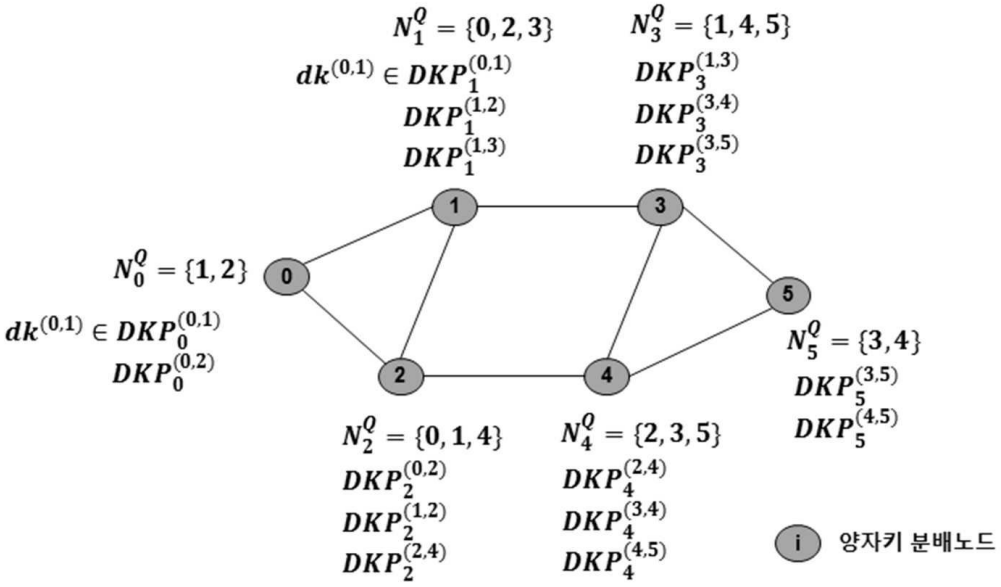
도면

도면1

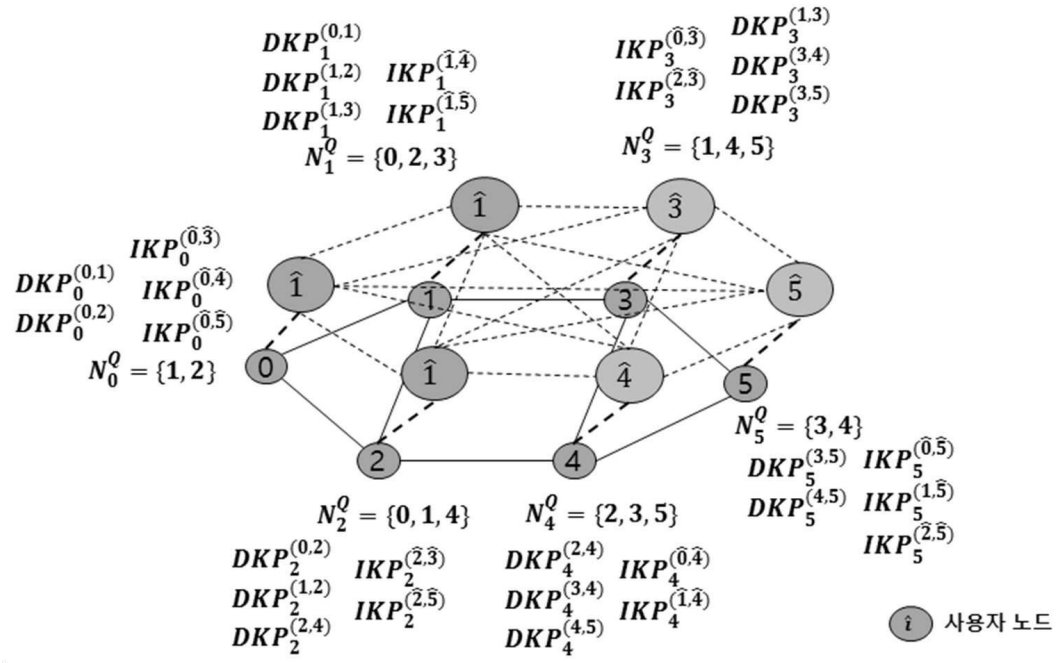


양자키 자원 관리 시스템(10)

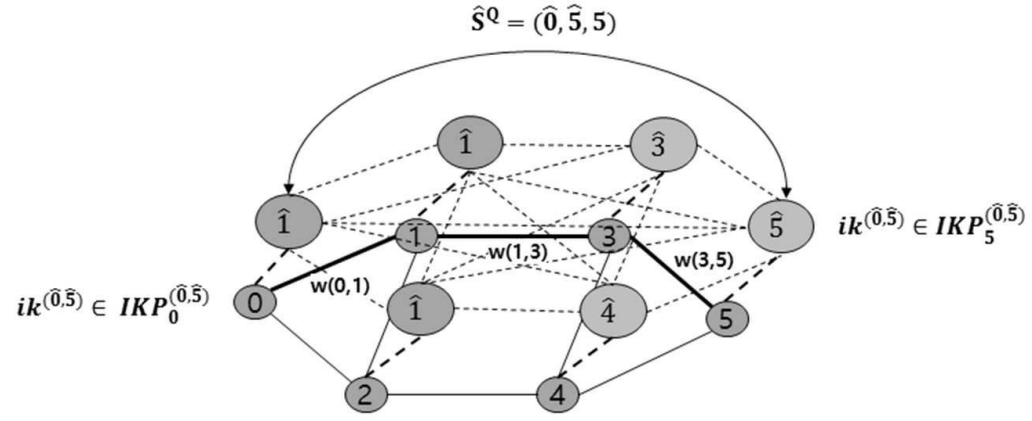
도면2



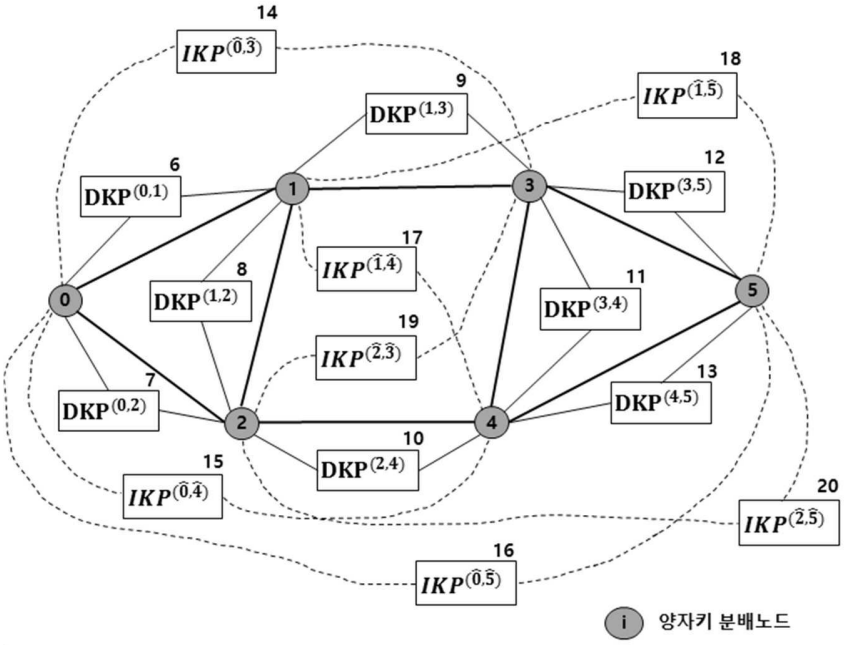
도면3


$DKP_1^{(0,1)}$
$DKP_1^{(1,2)}$ $IKP_1^{(\hat{1},\hat{4})}$
$DKP_1^{(1,3)}$ $IKP_1^{(\hat{1},\hat{5})}$
$N_1^Q = \{0, 2, 3\}$
$IKP_3^{(\hat{0},\hat{3})}$ $DKP_3^{(1,3)}$
$IKP_3^{(\hat{2},\hat{3})}$ $DKP_3^{(3,4)}$
$DKP_3^{(3,5)}$
$N_3^Q = \{1, 4, 5\}$
$IKP_0^{(\hat{0},\hat{3})}$
$DKP_0^{(0,1)}$ $IKP_0^{(\hat{0},\hat{4})}$
$DKP_0^{(0,2)}$ $IKP_0^{(\hat{0},\hat{5})}$
$N_0^Q = \{1, 2\}$
$N_5^Q = \{3, 4\}$
$DKP_5^{(3,5)}$ $IKP_5^{(\hat{0},\hat{5})}$
$DKP_5^{(4,5)}$ $IKP_5^{(\hat{1},\hat{5})}$
$IKP_5^{(\hat{2},\hat{5})}$
$N_2^Q = \{0, 1, 4\}$
$DKP_2^{(0,2)}$ $IKP_2^{(\hat{2},\hat{3})}$
$DKP_2^{(1,2)}$ $IKP_2^{(\hat{2},\hat{5})}$
$DKP_2^{(2,4)}$
$N_4^Q = \{2, 3, 5\}$
$DKP_4^{(2,4)}$ $IKP_4^{(\hat{0},\hat{4})}$
$DKP_4^{(3,4)}$ $IKP_4^{(\hat{1},\hat{4})}$
$DKP_4^{(4,5)}$
$\hat{i}$ 사용자 노드

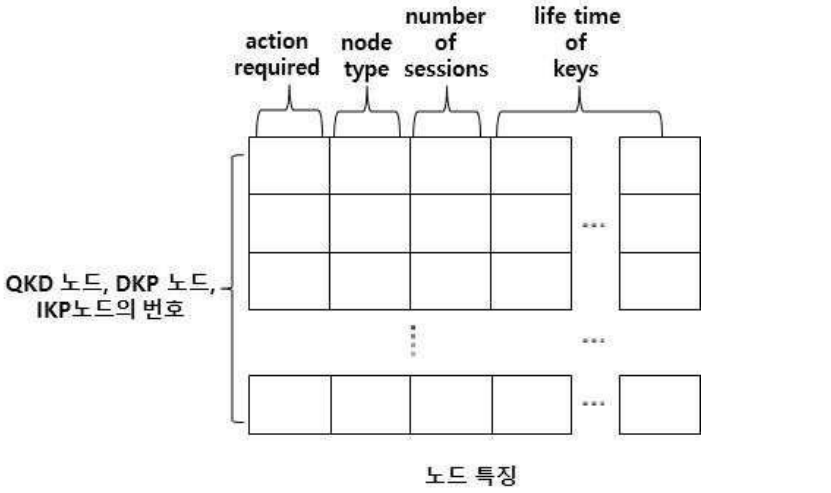
도면4


$\hat{S}^Q = (\hat{0}, \hat{5}, 5)$
$ik^{(\hat{0},\hat{5})} \in IKP_0^{(\hat{0},\hat{5})}$
$ik^{(\hat{0},\hat{5})} \in IKP_5^{(\hat{0},\hat{5})}$
w(0,1)
w(1,3)
w(3,5)

도면5


14
$IKP^{(\hat{0},\hat{3})}$
18
$IKP^{(\hat{1},\hat{5})}$
9
$DKP^{(1,3)}$
6
$DKP^{(0,1)}$
12
$DKP^{(3,5)}$
17
$IKP^{(\hat{1},\hat{4})}$
8
$DKP^{(1,2)}$
11
$DKP^{(3,4)}$
19
$IKP^{(\hat{2},\hat{3})}$
7
$DKP^{(0,2)}$
13
$DKP^{(4,5)}$
10
$DKP^{(2,4)}$
15
$IKP^{(\hat{0},\hat{4})}$
20
$IKP^{(\hat{2},\hat{5})}$
16
$IKP^{(\hat{0},\hat{5})}$
0
1
2
3
4
5
i
양자키 분배노드

도면6


action required
node type
number of sessions
life time of keys
QKD 노드, DKP 노드, IKP노드의 번호

노드 특징

도면7

노드 인덱스

		0	1	2	3	4	...	10
QKD 노드	0	0	0	-1	-1	-1	...	-1
	1	0	0	-1	-1	-1	...	-1
	2	0	0	-1	-1	-1	...	-1
	3	0	0	-1	-1	-1	...	-1
	4	0	0	-1	-1	-1	...	-1
	5	0	0	-1	-1	-1	...	-1
DPK 노드	6	0	1	2	3	3	...	0
	7	0	1	0	3	3	...	0
	8	0	1	1	3	2	...	0
	9	0	1	2	3	3	...	0
	10	0	1	0	3	3	...	0
	11	0	1	2	3	3	...	0
	12	0	1	1	3	3	...	0
	13	0	1	0	3	2	...	0
IPK 노드	14	1	2	2	3	2	...	0
	15	1	2	3	3	3	...	0
	. .	. .	. .	. .	. .	. .	. .	. .
	20	0	2	1	1	1	...	1

특징 인덱스

노트 특징

도면8

2 X E0 = 16 E0 : QKD 노드 간 연결

0 1,0,2,1,2,1,3,2,4,3,4,3,5,4,5
1 0,2,0,2,1,3,1,4,2,4,3,5,3,5,4

2 X E1 = 32 E1 : QKD 노드와 DKP 노드 간 연결

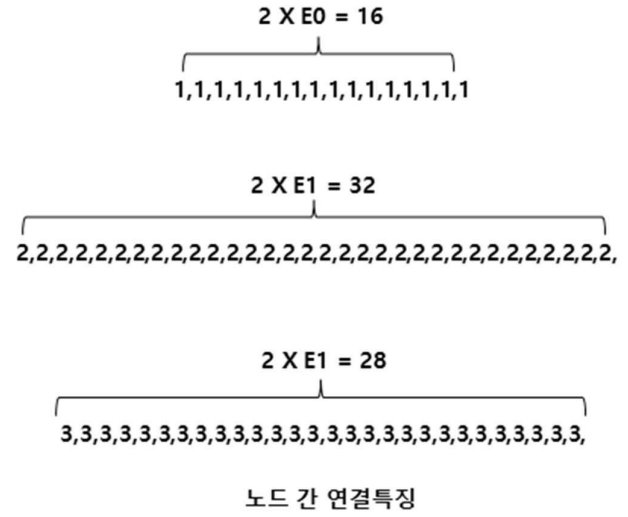
0 6,1,6,0,7,2,7,1,8,2,8,1,9,3,9,2,10,4,10,3,11,4,11,3,12,5,12,4,13,5,13,
6 0,6,1,7,0,7,2,8,1,8,2,9,1,9,3,10,2,10,4,11,3,11,4,12,3,12,5,13,4,13,5

2 X E1 = 28 E1 : QKD 노드와 IPK 노드 간 연결

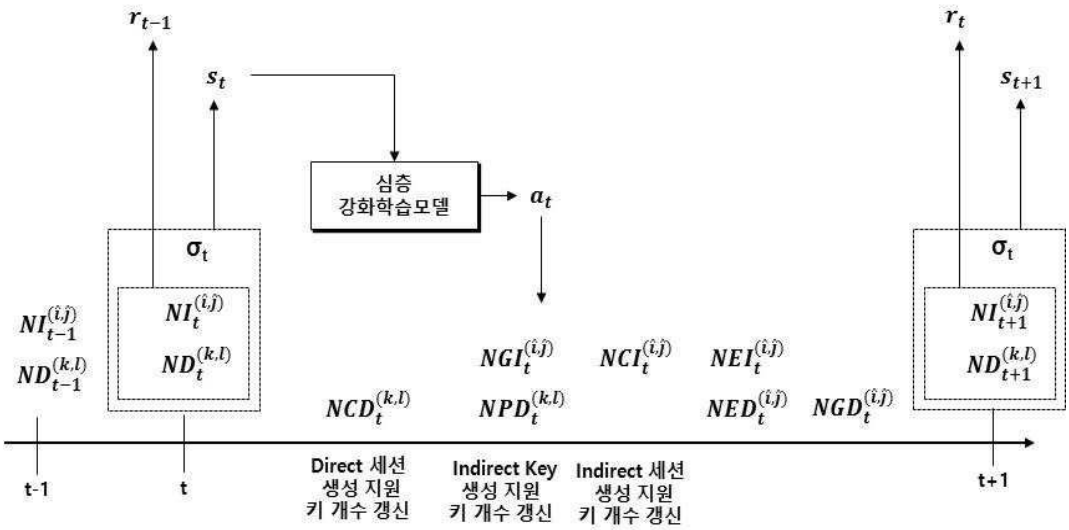
0, 14,3,14,0,15,4,15,0,16,5,16,1,17,4,17,1,18,5,18,2,19,3,19,2,20,5,20
14 0,14,3,15,0,15,4,16,0,16,5,17,1,17,4,18,1,18,5,19,2,19,3,20,2,20,5

노드 간 연결정보

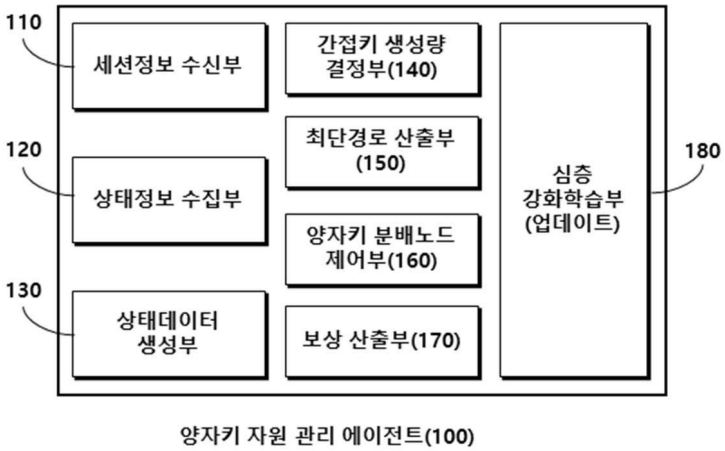
도면9



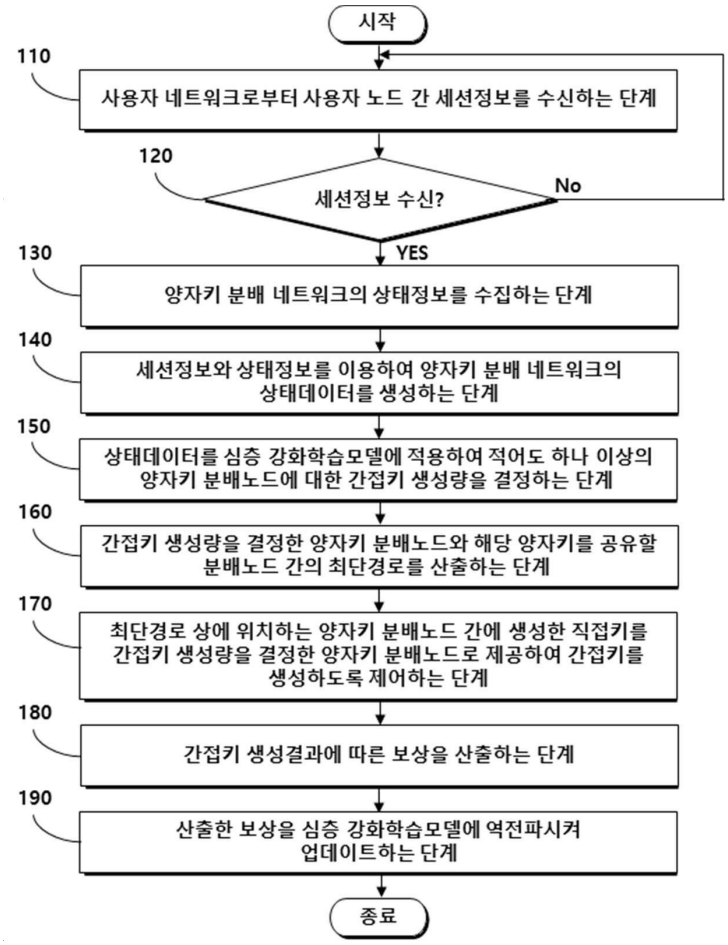
도면10



도면11



도면12


시작
110
사용자 네트워크로부터 사용자 노드 간 세션정보를 수신하는 단계
120
세션정보 수신?
No
YES
130
양자키 분배 네트워크의 상태정보를 수집하는 단계
140
세션정보와 상태정보를 이용하여 양자키 분배 네트워크의 상태데이터를 생성하는 단계
150
상태데이터를 심층 강화학습모델에 적용하여 적어도 하나 이상의 양자키 분배노드에 대한 간접키 생성량을 결정하는 단계
160
간접키 생성량을 결정한 양자키 분배노드와 해당 양자키를 공유할 분배노드 간의 최단경로를 산출하는 단계
170
최단경로 상에 위치하는 양자키 분배노드 간에 생성한 직접키를 간접키 생성량을 결정한 양자키 분배노드로 제공하여 간접키를 생성하도록 제어하는 단계
180
간접키 생성결과에 따른 보상을 산출하는 단계
190
산출한 보상을 심층 강화학습모델에 역전파시켜 업데이트하는 단계
종료