

(19) 대한민국특허청(KR)
(12) 공개특허공보(A)(11) 공개번호 10-2024-0117251
(43) 공개일자 2024년08월01일

(51) 국제특허분류(Int. Cl.)
G08G 1/081 (2006.01) G06N 20/00 (2019.01)
G08G 1/01 (2006.01) H04W 4/40 (2018.01)
(52) CPC특허분류
G08G 1/081 (2013.01)
G06N 20/00 (2021.08)
(21) 출원번호 10-2023-0009393
(22) 출원일자 2023년01월25일
심사청구일자 2023년01월25일

(71) 출원인
인하대학교 산학협력단
인천광역시 미추홀구 인하로 100(용현동, 인하대학교)
(72) 발명자
장경희
서울특별시 마포구 백범로 205, 101동 1110호(신공덕동, 펜트하우스)
사지드
인천광역시 미추홀구 인하로134번길 24-35(용현동)
(74) 대리인
양성보

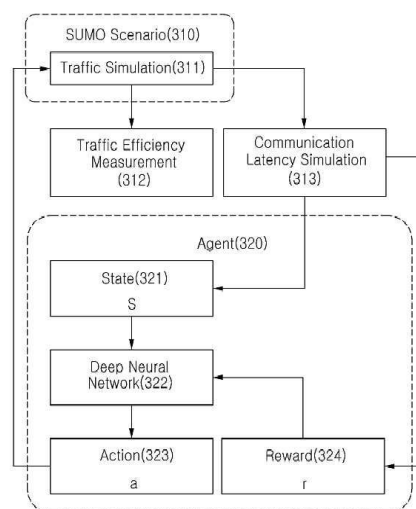
전체 청구항 수 : 총 8 항

(54) 발명의 명칭 5G-NR-V2I 데이터를 이용한 DRL 기반 교차로 교통관리 방법 및 시스템

(57) 요약

5G-NR-V2I 데이터를 이용한 DRL 기반 교차로 교통관리 방법 및 시스템이 제시된다. 본 발명에서 제안하는 5G-NR-V2I 데이터를 이용한 DRL 기반 교차로 교통관리 시스템은 강화형 기계 학습에서 학습되고 평가되는 교통 신호 제어를 최적화 하기 위한 SUMO(Simulation of Urban Mobility) 시뮬레이터 및 상기 최적화된 교통 신호 제어에 따른 교통 정보를 이용하여 교차로에서의 차량의 대기시간을 단축하기 위해 복수의 상태를 가지는 가능한 상태 공간(S), 복수의 행동을 갖는 가능한 행동 공간(A), 하나의 상태에서 다음 상태로의 전이 확률 함수 공간(P) 및 누적 기대 보상(R)을 기반으로 강화형 기계 학습(Reinforcement Learning; RL)을 수행하는 에이전트의 모델을 포함한다.

대표도 - 도3



(52) CPC특허분류

G08G 1/0145 (2013.01)

H04W 4/40 (2020.05)

이 발명을 지원한 국가연구개발사업

과제고유번호	1711175434
과제번호	2021K1A3A1A61003236
부처명	과학기술정보통신부
과제관리(전문)기관명	한국연구재단
연구사업명	국제화기반조성사업
연구과제명	[Ezbaro] 5G-NR V2X 기술을 활용한 완전 자율주행 차량 교통 흐름 최적화 및 안전성
향상 연구	
기 여 율	1/1
과제수행기관명	인하대학교
연구기간	2022.10.29 ~ 2023.10.28

명세서

청구범위

청구항 1

강화형 기계 학습에서 학습되고 평가되는 교통 신호 제어를 최적화 하기 위한 SUMO(Simulation of Urban MObility) 시뮬레이터; 및

상기 최적화된 교통 신호 제어에 따른 교통 정보를 이용하여 교차로에서의 차량의 대기시간을 단축하기 위해 복수의 상태를 가지는 가능한 상태 공간(S), 복수의 행동을 갖는 가능한 행동 공간(A), 하나의 상태에서 다음 상태로의 전이 확률 함수 공간(P) 및 누적 기대 보상(R)을 기반으로 강화형 기계 학습(Reinforcement Learning; RL)을 수행하는 에이전트의 모델

을 포함하는 교차로 교통관리 시스템.

청구항 2

제1항에 있어서,

상기 에이전트의 모델은,

URLLC(Ultra-Reliable Low Latency Communication)을 유지하기 위해, 패킷 오류를 수정하고, 패킷 오류를 줄이고, 패킷 수신 비율을 유지하기 위한 재전송에 따른 지연 시간 및 전송 대기 시간에 기초하여 각 차량의 대기 시간을 정의하고, 차선의 접근 차량 수와 차선의 대기열 크기를 정의하여 상기 상태 공간(S)을 나타내는

교차로 교통관리 시스템.

청구항 3

제1항에 있어서,

상기 에이전트의 모델은,

상기 에이전트의 모델이 신호등 시스템에 대하여 일정 시간 동안 충돌하지 않는 녹색 단계 집합을 활성화하는 선택 가능한 상기 행동 공간(A)을 나타내는

교차로 교통관리 시스템.

청구항 4

제1항에 있어서,

상기 에이전트의 모델은,

상기 행동 공간(A)을 선택한 후 환경으로부터 받은 피드백을 나타내는 누적 기대 보상(R)을 사용하여 상기 행동 공간(A)의 결과를 개선하며, 총 대기 시간의 음수 값, 운전 편의성을 제공하기 위한 가중치 및 패널티에 따라 상기 누적 기대 보상(R)을 나타내는

교차로 교통관리 시스템.

청구항 5

SUMO(Simulation of Urban MObility) 시뮬레이터를 통해 강화형 기계 학습에서 학습되고 평가되는 교통 신호 제어를 최적화 하는 단계; 및

에이전트의 모델을 통해 상기 최적화된 교통 신호 제어에 따른 교통 정보를 이용하여 교차로에서의 차량의 대기 시간을 단축하기 위해 복수의 상태를 가지는 가능한 상태 공간(S), 복수의 행동을 갖는 가능한 행동 공간(A), 하나의 상태에서 다음 상태로의 전이 확률 함수 공간(P) 및 누적 기대 보상(R)을 기반으로 강화형 기계 학습(Reinforcement Learning; RL)을 수행하는 단계

를 포함하는 교차로 교통관리 방법.

청구항 6

제5항에 있어서,

상기 복수의 상태를 가지는 가능한 상태 공간(S), 복수의 행동을 갖는 가능한 행동 공간(A), 하나의 상태에서 다음 상태로의 전이 확률 함수 공간(P) 및 누적 기대 보상(R)을 기반으로 강화형 기계 학습을 수행하는 단계는, URLLC(Ultra-Reliable Low Latency Communication)을 유지하기 위해, 패킷 오류를 수정하고, 패킷 오류를 줄이고, 패킷 수신 비율을 유지하기 위한 재전송에 따른 지연 시간 및 전송 대기 시간에 기초하여 각 차량의 대기 시간을 정의하고, 차선의 접근 차량 수와 차선의 대기열 크기를 정의하여 상기 상태 공간(S)을 나타내는

교차로 교통관리 방법.

청구항 7

제5항에 있어서,

상기 복수의 상태를 가지는 가능한 상태 공간(S), 복수의 행동을 갖는 가능한 행동 공간(A), 하나의 상태에서 다음 상태로의 전이 확률 함수 공간(P) 및 누적 기대 보상(R)을 기반으로 강화형 기계 학습을 수행하는 단계는, 상기 에이전트의 모델이 신호등 시스템에 대하여 일정 시간 동안 충돌하지 않는 녹색 단계 집합을 활성화하는 선택 가능한 상기 행동 공간(A)을 나타내는

교차로 교통관리 방법.

청구항 8

제5항에 있어서,

상기 복수의 상태를 가지는 가능한 상태 공간(S), 복수의 행동을 갖는 가능한 행동 공간(A), 하나의 상태에서 다음 상태로의 전이 확률 함수 공간(P) 및 누적 기대 보상(R)을 기반으로 강화형 기계 학습을 수행하는 단계는, 상기 행동 공간(A)을 선택한 후 환경으로부터 받은 피드백을 나타내는 누적 기대 보상(R)을 사용하여 상기 행동 공간(A)의 결과를 개선하며, 총 대기 시간의 음수 값, 운전 편의성을 제공하기 위한 가중치 및 패널티에 따라 상기 누적 기대 보상(R)을 나타내는

교차로 교통관리 방법.

발명의 설명

기술 분야

[0001] 본 발명은 5G-NR-V2I 데이터를 이용한 DRL 기반 교차로 교통관리 방법 및 시스템에 관한 것이다.

배경 기술

[0002] 최근 몇 년간 인구 증가에 따른 교통의 증가는 심각한 교통 혼잡 문제를 야기하고 있다. 한 연구에 따르면 상용차의 성장은 3억 3천 5백만대, 승용차의 수는 거의 9억 5천만대에 달한다. 그 외에도, 정체의 일반적인 원인은 갑작스러운 병목 현상, 불충분한 인프라 또는 교차로의 비효율적인 교통 통제 시스템이다.

[0003] 차로에서, 일반적인 신호등 제어 시스템은 실시간 교통 정보에 동적으로 적응하지 못하기 때문에 많은 양의 교통에 그다지 효과적이지 않은 것으로 보인다. 반면, 최첨단 심층 강화 학습(Deep Reinforcement Learning; DRL) 기술은 실시간 교통 정보를 기반으로 교통 효율성을 향상시킬 수 있는 더 큰 잠재력을 보여준다. 종래 기술에서 DRL 알고리즘은 교통 상황의 입력을 취하고 교차로를 관리하는 방법을 학습한다. 이들 중 일부는 유도 루프 감지거나 고해상도 비디오 데이터와 같은 지하 센서로부터 입력 정보를 가져와 신호등 근처의 차량 정보를 얻는다. 그러나 이러한 관찰 방법은 정확성과 낮은 대기 시간을 달성하는 데 있어 프로세스가 무겁고 비효율적이다.

[0004] 짧은 대기 시간으로 실시간 데이터 수집 및 전송을 달성하기 위해 차량 대 인프라 또는 V2I(Vehicle-to-

Infrastructure)를 통한 차량 및 도로변 인프라 간 통신이 5G-NR-V2X 패러다임에서 가능하다. 기존 센싱 감지기와 비교하여 V2X(Vehicle-to-Everything)는 교통제어를 위한 광범위한 정보를 달성하기 위해 데이터를 더 빠르고 정확하게 수집할 수 있으며 더 나은 DRL 의사 결정으로 이어진다. 그러나 V2I 통신을 DRL 알고리즘과 통합하는 방법과 이 통신 패러다임이 DRL 성능에 어떤 영향을 미칠지는 아직 알려지지 않았다.

선행기술문헌

비특허문헌

- [0005] (비특허문헌 0001) [1] Ault, J., & Sharon, G. (2021, June). "Reinforcement Learning Benchmarks for Traffic Signal Control." In Thirty-fifth Conference on Neural Information Processing Systems Datasets and Benchmarks Track (Round 1).
- (비특허문헌 0002) [2] Anwar, W., Franchi, N., & Fettweis, G. (2019, September). "Physical layer evaluation of V2X communications technologies: 5G NR-V2X, LTE-V2X, IEEE 802.11 bd, and IEEE 802.11 p." In 2019 IEEE 90th Vehicular Technology Conference (VTC2019-Fall) (pp. 1-7). IEEE.
- (비특허문헌 0003) [3] Nadeem, F., Shirvanimoghaddam, M., Li, Y., & Vucetic, B. (2021). "Nonorthogonal HARQ for URLLC: Design and analysis." IEEE Internet of Things Journal, 8(24), 17596-17610.
- (비특허문헌 0004) [4] Krajzewicz, D. (2010). "Traffic simulation with SUMO-simulation of urban mobility." In Fundamentals of traffic simulation (pp. 269-293). Springer, New York, NY.

발명의 내용

해결하려는 과제

- [0006] 본 발명이 이루고자 하는 기술적 과제는 V2I-TSC(Vehicle-to-Infrastructure based Traffic Signal Control)를 제안하여 DRL(Deep Reinforcement Learning)과 V2I(Vehicle-to-Infrastructure) 통신에 의해 축적된 교통 정보를 이용하여 교차로에서 차량의 대기시간을 단축하여 교통효율을 높이는 방법 및 시스템을 제공하는데 있다. 또한, 단일 에이전트 RL(Reinforcement Learning) 프레임워크를 사용하여 SUMO(Simulation of Urban Mobility) 시뮬레이터를 통해 훈련되고 평가되는 교통 신호 제어를 최적화하고자 한다.

과제의 해결 수단

- [0007] 일 측면에 있어서, 본 발명에서 제안하는 5G-NR-V2I 데이터를 이용한 DRL 기반 교차로 교통관리 시스템은 강화형 기계 학습에서 학습되고 평가되는 교통 신호 제어를 최적화 하기 위한 SUMO(Simulation of Urban Mobility) 시뮬레이터 및 상기 최적화된 교통 신호 제어에 따른 교통 정보를 이용하여 교차로에서의 차량의 대기시간을 단축하기 위해 복수의 상태를 가지는 가능한 상태 공간(S), 복수의 행동을 갖는 가능한 행동 공간(A), 하나의 상태에서 다음 상태로의 전이 확률 함수 공간(P) 및 누적 기대 보상(R)을 기반으로 강화형 기계 학습(Reinforcement Learning; RL)을 수행하는 에이전트의 모델을 포함한다.
- [0008] 상기 에이전트의 모델은 URLLC(Ultra-Reliable Low Latency Communication)을 유지하기 위해, 패킷 오류를 수정하고, 패킷 오류를 줄이고, 패킷 수신 비율을 유지하기 위한 재전송에 따른 지연 시간 및 전송 대기 시간에 기초하여 각 차량의 대기 시간을 정의하고, 차선의 접근 차량 수와 차선의 대기열 크기를 정의하여 상기 상태 공간(S)을 나타낸다.
- [0009] 상기 에이전트의 모델은 상기 에이전트의 모델이 신호등 시스템에 대하여 일정 시간 동안 충돌하지 않는 녹색 단계 집합을 활성화하는 선택 가능한 상기 행동 공간(A)을 나타낸다.
- [0010] 상기 에이전트의 모델은 상기 행동 공간(A)을 선택한 후 환경으로부터 받은 피드백을 나타내는 누적 기대 보상(R)을 사용하여 상기 행동 공간(A)의 결과를 개선하며, 총 대기 시간의 음수 값, 운전 편의성을 제공하기 위한 가중치 및 패널티에 따라 상기 누적 기대 보상(R)을 나타낸다.
- [0011] 또 다른 일 측면에 있어서, 본 발명에서 제안하는 5G-NR-V2I 데이터를 이용한 DRL 기반 교차로 교통관리 방법은 SUMO(Simulation of Urban Mobility) 시뮬레이터를 통해 강화형 기계 학습에서 학습되고 평가되는 교통 신호 제

어를 최적화 하는 단계 및 에이전트의 모델을 통해 상기 최적화된 교통 신호 제어에 따른 교통 정보를 이용하여 교차로에서의 차량의 대기시간을 단축하기 위해 복수의 상태를 가지는 가능한 상태 공간(S), 복수의 행동을 갖는 가능한 행동 공간(A), 하나의 상태에서 다음 상태로의 전이 확률 함수 공간(P) 및 누적 기대 보상(R)을 기반으로 강화형 기계 학습(Reinforcement Learning; RL)을 수행하는 단계를 포함한다.

발명의 효과

[0012] 본 발명의 실시예들에 따르면 V2I-TSC(Vehicle-to-Infrastructure based Traffic Signal Control)를 제안하여 DRL(Deep Reinforcement Learning)과 V2I(Vehicle-to-Infrastructure) 통신에 의해 축적된 교통 정보를 이용하여 교차로에서 차량의 대기시간을 단축하여 교통효율을 높일 수 있다. 또한, 단일 에이전트 RL(Reinforcement Learning) 프레임워크를 사용하여 SUMO(Simulation of Urban MObility) 시뮬레이터를 통해 훈련되고 평가되는 교통 신호 제어를 최적화할 수 있다.

도면의 간단한 설명

[0013] 도 1은 본 발명의 일 실시예에 따른 V2I-TSC 애플리케이션 시나리오의 가정을 나타내는 도면이다.

도 2는 본 발명의 일 실시예에 따른 5G-NR-V2I 데이터를 이용한 DRL 기반 교차로 교통관리 시스템의 구성을 나타내는 도면이다.

도 3은 본 발명의 일 실시예에 따른 상태, 동작, 보상 함수를 갖춘 DQN 기반 알고리즘을 설명하기 위한 도면이다.

도 4는 본 발명의 일 실시예에 따른 5G-NR-V2I 데이터를 이용한 DRL 기반 교차로 교통관리 방법을 설명하기 위한 흐름도이다.

도 5는 본 발명의 일 실시예에 따른 통신 대기 시간을 고려한 에이전트의 평균 대기 시간 및 트래픽 단계 기간과 트래픽 효율성 사이의 균형을 나타내는 그래프이다.

도 6은 본 발명의 일 실시예에 따른 평균 대기시간을 나타내는 그래프이다.

발명을 실시하기 위한 구체적인 내용

[0014] 강화 학습(Reinforcement Learning; RL) 기반 교통 관리에 대한 최근 연구는 유망한 결과를 보여주지만, 교통량이 증가하고 실시간 교통 정보가 부족하기 때문에 중요한 문제이다. RL 알고리즘과 V2X(Vehicle-to-Everything) 통신 기술의 개선은 더 나은 트래픽 효율성을 달성하기 위한 새로운 전망을 창출하고 있다. 본 발명은 5G-NR-V2X 패러다임에서 V2I(Vehicle-to-Infrastructure) 통신을 사용하여 현실적인 교통 상태를 포착하기 위한 새로운 방법인 V2I-TSC를 제안한다. 단일 에이전트 RL 프레임워크를 사용하여 SUMO(Simulation of Urban MObility) 시뮬레이터를 통해 훈련되고 평가되는 교통 신호 제어를 최적화한다. 실험 결과는 제안하는 방법이 일반적인 교통 통제 방법에 비해 교차로에서 교통 효율성을 향상시킨다는 것을 보여준다. 이하, 본 발명의 실시 예를 첨부된 도면을 참조하여 상세하게 설명한다.

[0016] 기존의 신호등 제어 시스템은 실시간 교통을 통합하지 않고 고정 프로그램을 배치하거나 시각 기반, 센서 기반 또는 V2X 기반 감지에 기반한 실시간 교통 조건을 제한된 정도로 고려하여 교통 혼잡을 줄이려고 한다. 교통 상황의 관찰은 감지 능력과 감지 반경의 가정에 의해 제한된다. 예를 들어, 일부 종래기술은 최첨단 기술을 가정하고 대기열로 정의된 정지된 차량의 수, 누적 대기 시간 등을 포함한 실시간 관측을 고려하는 반면, 다른 종래 기술에서는 덜 유익한 감지 능력을 가정하고 대기열 앞에 보이는 차량의 가시적 대기열 길이 또는 대기 시간만 고려한다. 감지 반경 또한 50미터에서 들어오는 도로 전체에 이르기까지 다양하다. 그러나 제한된 관측치는 전체적인 교통 조건을 제공하지 않는다. 동시에, 특히 도로 전체를 감지하는 경우에는 감지 오류나 현실적인 감지 조건을 고려하지 않기 때문에 관찰에 부정확한 교통 정보가 포함될 수 있다. 교통 분석에서는 대기 시간, 처리량, 지연, 대기열과 같은 다양한 측정값이 성능을 측정하는 데 사용된다.

[0018] 도 1은 본 발명의 일 실시예에 따른 V2I-TSC 애플리케이션 시나리오의 가정을 나타내는 도면이다.

[0019] 도 1은 V2I-TSC 애플리케이션 시나리오의 가정을 보여준다. 도 1은 각 접근로가 신호등 신호와 여러 개의 진입

및 출구 차선을 포함하는 4개의 접근 교차로를 보여준다. 4개의 접근로는 각각 길이가 870m, 920m, 791m, 356m, 인 동(E), 서(W), 남(S), 북(N)으로 표시되어 있다. 교통 신호등은 교차로의 교통 흐름을 조절하는 RL 에이전트에 의해 제어된다. RL 에이전트는 교차로에 위치한 RSU로부터 교통 조건을 수신한다. 이 정보는 V2I 링크를 통해 다음 교차로 영역 내에 위치한 차량에서 RSU로 전송되는 주기적 메시지에서 수집된다. 이 정보를 기반으로, RL 에이전트는 신호등으로 다시 전송되는 최적의 위상 조합과 위상 지속 시간을 결정하는 방법을 학습한다.

[0021] 도 2는 본 발명의 일 실시예에 따른 5G-NR-V2I 데이터를 이용한 DRL 기반 교차로 교통관리 시스템의 구성을 나타내는 도면이다.

[0022] 본 실시예에 따른 DRL 기반 교차로 교통관리 시스템(200)은 프로세서(210), 버스(220), 네트워크 인터페이스(230), 메모리(240) 및 데이터베이스(250)를 포함할 수 있다. 메모리(240)는 운영체제(241) 및 DRL 기반 교차로 교통관리 루틴(242)을 포함할 수 있다. 프로세서(210)는 SUMO(Simulation of Urban Mobility) 시뮬레이터(211) 및 에이전트 모델(212)을 포함할 수 있다. 다른 실시예들에서 DRL 기반 교차로 교통관리 시스템(200)은 도 2 구성요소들보다 더 많은 구성요소들을 포함할 수도 있다. 그러나, 대부분의 종래기술적 구성요소들을 명확하게 도시할 필요성은 없다. 예를 들어, DRL 기반 교차로 교통관리 시스템(200)은 디스플레이나 트랜시버(transceiver)와 같은 다른 구성요소들을 포함할 수도 있다.

[0023] 메모리(240)는 컴퓨터에서 판독 가능한 기록 매체로서, RAM(random access memory), ROM(read only memory) 및 디스크 드라이브와 같은 비소멸성 대용량 기록장치(permanent mass storage device)를 포함할 수 있다. 또한, 메모리(240)에는 운영체제(241)와 DRL 기반 교차로 교통관리 루틴(242)을 위한 프로그램 코드가 저장될 수 있다. 이러한 소프트웨어 구성요소들은 드라이브 메커니즘(drive mechanism, 미도시)을 이용하여 메모리(240)와는 별도의 컴퓨터에서 판독 가능한 기록 매체로부터 로딩될 수 있다. 이러한 별도의 컴퓨터에서 판독 가능한 기록 매체는 플로피 드라이브, 디스크, 테이프, DVD/CD-ROM 드라이브, 메모리 카드 등의 컴퓨터에서 판독 가능한 기록 매체(미도시)를 포함할 수 있다. 다른 실시예에서 소프트웨어 구성요소들은 컴퓨터에서 판독 가능한 기록 매체가 아닌 네트워크 인터페이스(230)를 통해 메모리(240)에 로딩될 수도 있다.

[0024] 버스(220)는 DRL 기반 교차로 교통관리 시스템(200)의 구성요소들간의 통신 및 데이터 전송을 가능하게 할 수 있다. 버스(220)는 고속 시리얼 버스(high-speed serial bus), 병렬 버스(parallel bus), SAN(Storage Area Network) 및/또는 다른 적절한 통신 기술을 이용하여 구성될 수 있다.

[0025] 네트워크 인터페이스(230)는 DRL 기반 교차로 교통관리 시스템(200)을 컴퓨터 네트워크에 연결하기 위한 컴퓨터 하드웨어 구성요소일 수 있다. 네트워크 인터페이스(230)는 DRL 기반 교차로 교통관리 시스템(200)을 무선 또는 유선 커넥션을 통해 컴퓨터 네트워크에 연결시킬 수 있다.

[0026] 데이터베이스(250) 5G-NR-V2I 데이터를 이용한 DRL 기반 교차로 교통관리를 위해 필요한 모든 정보를 저장 및 유지하는 역할을 할 수 있다. 도 2에서는 DRL 기반 교차로 교통관리 시스템(200)의 내부에 데이터베이스(250)를 구축하여 포함하는 것으로 도시하고 있으나, 이에 한정되는 것은 아니며 시스템 구현 방식이나 환경 등에 따라 생략될 수 있고 혹은 전체 또는 일부의 데이터베이스가 별개의 다른 시스템 상에 구축된 외부 데이터베이스로서 존재하는 것 또한 가능하다.

[0027] 프로세서(210)는 기본적인 산술, 로직 및 DRL 기반 교차로 교통관리 시스템(200)의 입출력 연산을 수행함으로써, 컴퓨터 프로그램의 명령을 처리하도록 구성될 수 있다. 명령은 메모리(240) 또는 네트워크 인터페이스(230)에 의해, 그리고 버스(220)를 통해 프로세서(210)로 제공될 수 있다. 프로세서(210)는 SUMO 시뮬레이터(211) 및 에이전트 모델(212)을 위한 프로그램 코드를 실행하도록 구성될 수 있다. 이러한 프로그램 코드는 메모리(240)와 같은 기록 장치에 저장될 수 있다.

[0028] SUMO 시뮬레이터(211) 및 에이전트 모델(212)은 도 4 단계들(410~420)을 수행하기 위해 구성될 수 있다.

[0029] DRL 기반 교차로 교통관리 시스템(200)은 SUMO 시뮬레이터(211) 및 에이전트 모델(212)을 포함할 수 있다.

[0030] 본 발명의 실시예에 따른 SUMO 시뮬레이터(211)는 강화형 기계 학습에서 학습되고 평가되는 교통 신호 제어를 최적화 하기 위한 SUMO 시뮬레이션을 수행한다.

[0031] 도 1을 참조하면 본 발명의 실시예에 따른 V2I-TSC 애플리케이션 시나리오의 가정에서, 각 접근로가 신호등 신호와 여러 개의 진입 및 출구 차선을 포함하는 4개의 접근 교차로를 보여준다. 4개의 접근로는 각각 길이가

870m, 920m, 791m, 356m, 인 동(E), 서(W), 남(S), 북(N)으로 표시되어 있다. 교통 신호등은 교차로의 교통 흐름을 조절하는 RL 에이전트에 의해 제어된다. RL 에이전트는 교차로에 위치한 RSU로부터 교통 조건을 수신한다. 이 정보는 V2I 링크를 통해 다음 교차로 영역 내에 위치한 차량에서 RSU로 전송되는 주기적 메시지로부터 수집된다. 이 정보를 기반으로, RL 에이전트는 신호등으로 다시 전송되는 최적의 위상 조합과 위상 지속 시간을 결정하는 방법을 학습한다.

[0032] 본 발명에서는 현실적인 V2I 기반 데이터 수집 방법을 고려하여 단일 RL 에이전트를 제안한다.

[0033] 본 발명의 실시예에 따른 에이전트 모델(212)은 상기 최적화된 교통 신호 제어에 따른 교통 정보를 이용하여 교차로에서의 차량의 대기시간을 단축하기 위해 복수의 상태를 가지는 가능한 상태 공간(S), 복수의 행동을 갖는 가능한 행동 공간(A), 하나의 상태에서 다음 상태로의 전이 확률 함수 공간(P) 및 누적 기대 보상(R)을 기반으로 강화형 기계 학습(Reinforcement Learning; RL)을 수행한다.

[0034] 본 발명의 실시예에 따른 에이전트 모델(212)은 URLLC(Ultra-Reliable Low Latency Communication)을 유지하기 위해, 패킷 오류를 수정하고, 패킷 오류를 줄이고, 패킷 수신 비율을 유지하기 위한 재전송에 따른 지연 시간 및 전송 대기 시간에 기초하여 각 차량의 대기 시간을 정의하고, 차선의 접근 차량 수와 차선의 대기열 크기를 정의하여 상기 상태 공간(S)을 나타낸다.

[0035] 본 발명의 실시예에 따른 에이전트 모델(212)은 상기 에이전트의 모델이 신호등 시스템에 대하여 일정 시간 동안 충돌하지 않는 녹색 단계 집합을 활성화하는 선택 가능한 상기 행동 공간(A)을 나타낸다.

[0036] 본 발명의 실시예에 따른 에이전트 모델(212)은 상기 행동 공간(A)을 선택한 후 환경으로부터 받은 피드백을 나타내는 누적 기대 보상(R)을 사용하여 상기 행동 공간(A)의 결과를 개선하며, 총 대기 시간의 음수 값, 운전 편의성을 제공하기 위한 가중치 및 패널티에 따라 상기 누적 기대 보상(R)을 나타낸다. 도 3을 참조하여 본 발명의 실시예에 따른 상태, 동작, 보상 함수를 갖춘 DQN 기반 알고리즘을 더욱 상세히 설명한다.

[0038] 도 3은 본 발명의 일 실시예에 따른 상태, 동작, 보상 함수를 갖춘 DQN 기반 알고리즘을 설명하기 위한 도면이다.

[0039] 본 발명의 실시예에 따르면, SUMO 시나리오(310)를 통해 교통 시뮬레이션(311)을 수행하여 강화형 기계 학습에서 학습되고 평가되는 교통 신호 제어를 최적화한다. 그리고, 교통 효율 측정(312) 및 통신 지연 시뮬레이션(313)을 통한 현실적인 V2I 기반 데이터 수집 방법을 고려하여 단일 RL 에이전트 모델(320)을 제안한다.

[0040] 에이전트 모델(320)은 네 가지 항목 {S,A,R,P}를 기반으로 한다. S가 상태 $s_t \in S$ 를 가지는 가능한 상태 공간(321)이고 A가 행동 $a_t \in A$ 를 갖는 가능한 행동 공간(323)이라고 하자. P는 한 상태에서 다음 상태로의 전이 확률 함수 공간이고 R은 누적 기대 보상(324)이다. 에이전트 모델(320)은 $R(s, a)$ 에 의해 주어진 상태 테스트에서 강화형 기계 학습(322)을 통해 즉각적인 효용을 나타내는 보상뿐만 아니라 새로운 상태(s_{t+1})가 생성되는 방식으로 환경에 영향을 미치는 행동(a_t)을 선택한다. 목표는 관찰된 보상을 사용하여 R 값이 최대화될 수 있도록 정책을 조정하는 것이다. 에이전트 모델(320)은 교차로의 각 진입 차선에서 접근 중이거나 대기 중인 모든 차량으로부터 현재 상태(s_t) 정보를 획득해야 한다. 상태 정보는 차량에서 차선 시작까지의 거리, 각 차량의 속도 및 신호에서 모든 정지 차량의 대기 시간을 포함한다. 이러한 값을 계산하기 위해 차량 사용자(VUE)와 RSU 간의 V2I 링크를 통해 차량의 위치 업데이트를 주기적으로 수집하며, 통신 지연으로 인해 차량의 위치 업데이트가 지연될 경우 각 차량의 대기 시간 T_w 에 지연 시간 T_d 가 가산된다. 대기 시간은 차량의 정지 상태 지속 시간에 의해 정량화되고 위치 업데이트에 의해 차량의 정지 여부가 결정되기 때문이다. 따라서 각 차량 T_v 의 대기 시간은 다음과 같이 전송 대기 시간의 영향을 받는다.

$$T_v = T_w + T_d \quad (1)$$

[0042] N을 교차로의 L 차선에서 대기 중인 차량의 수라 하면, 궁극적인 최적화 목표는 다음과 같이 주어진 교차로에서 모든 차량의 평균 대기 시간을 최소화하는 것이다.

$$\text{minimize} : \frac{\sum_{l=1}^L \sum_{n=1}^N T_w}{N} \quad (2)$$

[0044] 본 발명의 실시예에 따르면, V2I 링크가 DRL에 미치는 영향을 분석하고 대기 시간을 줄이기 위해 상태, 동작, 보상 함수를 갖춘 DQN 기반 알고리즘을 제안한다. 제안된 방법은 또한 V2I 통신 성능을 조사하고 통신이 최적화 대상에 미치는 영향을 평가한다.

[0045] 차선에 위치한 차량의 대기 시간을 W_{ci} 로 하고 VUE에서 RSU까지의 거리를 d 라 하자. d 가 커지면 옆 방향 링크에서 패킷 오류가 발생할 수 있다[2]. 이 오류로 인해 잘못된 위치 데이터가 전송되어 시스템 성능이 저하될 수 있다. 5G-NR-V2X에서, URLLC(Ultra-reliable Low Latency Communication)을 유지하기 위해, 패킷 오류는 HARQ를 사용하여 수정된다[3]. 본 발명의 실시예에서는 오류를 줄이고 90% 패킷 수신 비율(Packet Reception Ratio; PRR)을 유지하기 위해 HARQ가 거리가 100m를 초과할 때 두 번 이상 재전송해야 한다고 가정한다. 5G-NR-V2X에서 단일 전송의 최소 지연 시간은 100바이트의 경우 0.25ms이고 1500바이트의 경우 1ms이다[2]. 제안된 시스템의 경우 패킷 크기에 100바이트가 필요한 URLLC가 고려된다. 차량이 다른 애플리케이션에 대해 V2I 링크를 동시에 사용한다고 가정하면 더 큰 패킷 크기와 더 많은 지연 시간이 예상된다. 따라서 이 작업은 각 전송에 대한 $T_L=0.50ms$ 지연 시간을 고려한다. 따라서, 전송 대기 시간을 고려한 각 차량의 대기 시간은 다음과 같이 정의된다.

$$W_{d_i} = W_{c_i} + T_L T_n \text{ where } T_n = \begin{cases} 1 & \text{if } d \leq 100 \\ 2 & \text{if } d \leq 150 \\ 3 & \text{if } d \leq 200 \\ 4 & \text{otherwise} \end{cases} \quad (3)$$

[0047] 반면에, 접근하는 차량의 수와 차선의 대기열에 있는 차량의 수는 지속 시간과 관련이 없기 때문에 전송 대기 시간의 영향을 받지 않는다. 차량의 속도가 v_c 라고 가정하면, 차선의 접근 차량 수와 차선의 대기열 크기는 다음과 같이 정의할 수 있다.

$$N_a = \sum_{i=1}^n c_i \text{ where } c_i = \begin{cases} 1 & \text{if } v_c > 0 \\ 0 & \text{otherwise} \end{cases} \quad (4)$$

$$N_q = \sum_{i=1}^n c_i \text{ where } c_i = \begin{cases} 1 & \text{if } v_c \neq 0 \\ 0 & \text{otherwise} \end{cases} \quad (5)$$

[0050] 상태 S 는 접근하는 차량의 수, 대기열의 크기, 모든 진입 차선의 대기 시간으로 정의된다.

$$S = \{S_W, S_{N_a}, S_{N_q}\} \\ \text{where } S_W = \sum W_{d_i}, S_{N_a} = \sum N_{a_i} \text{ and } S_{N_q} = \sum N_{q_i} \quad (6)$$

[0052] 본 발명의 실시예에 따른 행동 집합은 에이전트 모델(320)이 선택할 수 있는 가능한 행동을 나타낸다. 에이전트 모델(320)이 신호등 시스템이므로 일정 시간 동안 충돌하지 않는 녹색 단계 집합을 활성화하는 것이 행동이다. 형식적으로, 행동 공간은 집합 A 에 정의된다. 이 집합에는 도 1과 같이 에이전트 모델(320)이 수행할 수 있는 모든 가능한 행동이 포함되어 있다.

$$A = \{P_1, P_2, P_3, P_4\} \quad (7)$$

[0054] 본 발명의 실시예에 따른 강화 학습에서 보상은 에이전트 모델(320)이 행동을 선택한 후 환경으로부터 받은 피드백을 나타낸다. 에이전트 모델(320)은 보상을 사용하여 선택한 행동의 결과를 이해하고 향후 행동의 선택을 위해 모델을 개선한다. 본 발명에서는 피드백을 위해 총 대기 시간의 음수 값을 고려하는 식 (8)의 보상 함수를 구현한다. 대기 시간은 교통 효율성의 성능 척도이기 때문에 사용되며, 선택한 작업이 교통 효율성을 감소시키거나 증가시키는지 에이전트 모델(320)이 이해할 수 있도록 지원한다. 동시에, 운전자가 운전 편의성을 제공하기 위해서는, 운전자가 교통 상황을 자주 바꾸지 않는 것이 필요하다. 이를 인정하기 위해 가중치 S_{N_q} 와 함께 패널티 p 도 보상 함수에서 고려된다.

$$R(s_i, a_i) = R_w - S_{N_q} * p$$

$$\text{where, } R_w = -S_w \text{ and } p = \begin{cases} 0 & \text{if } a_i \neq a_{i-1} \\ 1 & \text{if } a_i = a_{i-1} \end{cases} \quad (8)$$

[0055]

[0057] 도 4는 본 발명의 일 실시예에 따른 5G-NR-V2I 데이터를 이용한 DRL 기반 교차로 교통관리 방법을 설명하기 위한 흐름도이다.

[0058]

제안하는 5G-NR-V2I 데이터를 이용한 DRL 기반 교차로 교통관리 방법은 SUMO(Simulation of Urban MObility) 시뮬레이터를 통해 강화형 기계 학습에서 학습되고 평가되는 교통 신호 제어를 최적화 하는 단계(410) 및 에이전트의 모델을 통해 상기 최적화된 교통 신호 제어에 따른 교통 정보를 이용하여 교차로에서의 차량의 대기시간을 단축하기 위해 복수의 상태를 가지는 가능한 상태 공간(S), 복수의 행동을 갖는 가능한 행동 공간(A), 하나의 상태에서 다음 상태로의 전이 확률 함수 공간(P) 및 누적 기대 보상(R)을 기반으로 강화형 기계 학습(Reinforcement Learning; RL)을 수행하는 단계(420)를 포함한다.

[0059]

단계(410)에서, SUMO 시뮬레이터를 통해 강화형 기계 학습에서 학습되고 평가되는 교통 신호 제어를 최적화 하기 위한 SUMO 시뮬레이션을 수행한다.

[0060]

도 1을 참조하면 본 발명의 실시예에 따른 V2I-TSC 애플리케이션 시나리오의 가정에서, 각 접근로가 신호등 신호와 여러 개의 진입 및 출구 차선을 포함하는 4개의 접근 교차로를 보여준다. 4개의 접근로는 각각 길이가 870m, 920m, 791m, 356m, 인 동(E), 서(W), 남(S), 북(N)으로 표시되어 있다. 교통 신호등은 교차로의 교통 흐름을 조절하는 RL 에이전트에 의해 제어된다. RL 에이전트는 교차로에 위치한 RSU로부터 교통 조건을 수신한다. 이 정보는 V2I 링크를 통해 다음 교차로 영역 내에 위치한 차량에서 RSU로 전송되는 주기적 메시지에서 수집된다. 이 정보를 기반으로, RL 에이전트는 신호등으로 다시 전송되는 최적의 위상 조합과 위상 지속 시간을 결정하는 방법을 학습한다.

[0061]

본 발명에서는 현실적인 V2I 기반 데이터 수집 방법을 고려하여 단일 RL 에이전트를 제안한다.

[0062]

단계(420)에서, 에이전트의 모델을 통해 상기 최적화된 교통 신호 제어에 따른 교통 정보를 이용하여 교차로에서의 차량의 대기시간을 단축하기 위해 복수의 상태를 가지는 가능한 상태 공간(S), 복수의 행동을 갖는 가능한 행동 공간(A), 하나의 상태에서 다음 상태로의 전이 확률 함수 공간(P) 및 누적 기대 보상(R)을 기반으로 강화형 기계 학습을 수행한다.

[0063]

본 발명의 실시예에 따른 에이전트 모델은 URLLC(Ultra-Reliable Low Latency Communication)을 유지하기 위해, 패킷 오류를 수정하고, 패킷 오류를 줄이고, 패킷 수신 비율을 유지하기 위한 재전송에 따른 지연 시간 및 전송 대기 시간에 기초하여 각 차량의 대기 시간을 정의하고, 차선의 접근 차량 수와 차선의 대기열 크기를 정의하여 상기 상태 공간(S)을 나타낸다.

[0064]

본 발명의 실시예에 따른 에이전트 모델은 상기 에이전트의 모델이 신호등 시스템에 대하여 일정 시간 동안 충돌하지 않는 녹색 단계 집합을 활성화하는 선택 가능한 상기 행동 공간(A)을 나타낸다.

[0065]

본 발명의 실시예에 따른 에이전트 모델은 상기 행동 공간(A)을 선택한 후 환경으로부터 받은 피드백을 나타내는 누적 기대 보상(R)을 사용하여 상기 행동 공간(A)의 결과를 개선하며, 총 대기 시간의 음수 값, 운전 편의성을 제공하기 위한 가중치 및 패널티에 따라 상기 누적 기대 보상(R)을 나타낸다.

[0067]

도 5는 본 발명의 일 실시예에 따른 통신 대기 시간을 고려한 에이전트의 평균 대기 시간 및 트래픽 단계 기간과 트래픽 효율성 사이의 균형을 나타내는 그래프이다.

[0068]

도 5(a)는 통신 대기 시간을 고려한 에이전트의 평균 대기 시간을 나타내는 그래프이고, 도 (b)는 트래픽 단계 기간과 트래픽 효율성 사이의 균형을 나타내는 그래프이다.

[0069]

V2I 통신의 효과를 분석하기 위해 도 5(a)에서 첫 100회 평균 대기 시간을 제시하여 통신 대기 시간이 있는 경우와 없는 경우의 시뮬레이션 결과를 비교한다. 도 5(a)는 에이전트가 상태 정보에 부정확한 대기 시간 데이터가 포함되어 있어도 평균 대기 시간을 줄이는 것을 보여준다. 대기 시간에 포함된 지연이 있더라도 상태의 대기 시간이 줄어들고 이에 비례하여 보상이 증가하므로 75회 학습 후에는 대기 시간의 차이가 미미해진다. 도 5(b)

는 DRL에 의해 최적화된 차선 대기열, 시간 손실 및 대기 시간에 대한 다양한 교통 신호 단계 지속 시간의 영향을 보여준다. 감소된 대기 시간, 즉 향상된 교통 효율성과 운전 편의성 사이에는 상충 관계가 있음이 분명하다. 단계 기간이 감소함에 따라 트래픽 효율성이 증가한다. 반면에 짧은 단계 기간은 운전의 편안함을 방해한다. 이 두 가지 상충되는 요소 사이의 균형을 맞추기 위해 단계 기간을 단축하고 누적 기대 보상을 제안한다.

[0071] 도 6은 본 발명의 일 실시예에 따른 평균 대기시간을 나타내는 그래프이다.

[0072] 도 6은 V2I-TSC가 고정 시간 제어 및 IDQN에 비해 대기 시간을 더 줄이고 트래픽 효율성을 향상시키는 것을 보여준다.

[0073] 이와 같이 본 발명은 V2I 통신과 RL 접근을 이용한 교차로 교통 효율성을 향상시킬 수 있다. 5G-NR-V2I 통신 패러다임을 사용하여 DRL 에이전트를 위한 더 나은 학습 환경을 만들기 위해 현실적인 교통 데이터를 합성하고, 그 결과 긍정적인 학습률과 최소 평균 대기 시간이 고정 시간 제어에 비해 각각 51%, IDQN에 비해 22% 감소한 것으로 나타났다.

[0074] 이상에서 설명된 장치는 하드웨어 구성요소, 소프트웨어 구성요소, 및/또는 하드웨어 구성요소 및 소프트웨어 구성요소의 조합으로 구현될 수 있다. 예를 들어, 실시예들에서 설명된 장치 및 구성요소는, 예를 들어, 프로세서, 콘트롤러, ALU(arithmetic logic unit), 디지털 신호 프로세서(digital signal processor), 마이크로컴퓨터, FPGA(field programmable gate array), PLU(programmable logic unit), 마이크로프로세서, 또는 명령(instruction)을 실행하고 응답할 수 있는 다른 어떠한 장치와 같이, 하나 이상의 범용 컴퓨터 또는 특수 목적 컴퓨터를 이용하여 구현될 수 있다. 처리 장치는 운영 체제(OS) 및 상기 운영 체제 상에서 수행되는 하나 이상의 소프트웨어 애플리케이션을 수행할 수 있다. 또한, 처리 장치는 소프트웨어의 실행에 응답하여, 데이터를 접근, 저장, 조작, 처리 및 생성할 수도 있다. 이해의 편의를 위하여, 처리 장치는 하나가 사용되는 것으로 설명된 경우도 있지만, 해당 기술분야에서 통상의 지식을 가진 자는, 처리 장치가 복수 개의 처리 요소(processing element) 및/또는 복수 유형의 처리 요소를 포함할 수 있음을 알 수 있다. 예를 들어, 처리 장치는 복수 개의 프로세서 또는 하나의 프로세서 및 하나의 콘트롤러를 포함할 수 있다. 또한, 병렬 프로세서(parallel processor)와 같은, 다른 처리 구성(processing configuration)도 가능하다.

[0075] 소프트웨어는 컴퓨터 프로그램(computer program), 코드(code), 명령(instruction), 또는 이들 중 하나 이상의 조합을 포함할 수 있으며, 원하는 대로 동작하도록 처리 장치를 구성하거나 독립적으로 또는 결합적으로(collectively) 처리 장치를 명령할 수 있다. 소프트웨어 및/또는 데이터는, 처리 장치에 의하여 해석되거나 처리 장치에 명령 또는 데이터를 제공하기 위하여, 어떤 유형의 기계, 구성요소(component), 물리적 장치, 가상 장치(virtual equipment), 컴퓨터 저장 매체 또는 장치에 구체화(embodiment)될 수 있다. 소프트웨어는 네트워크로 연결된 컴퓨터 시스템 상에 분산되어서, 분산된 방법으로 저장되거나 실행될 수도 있다. 소프트웨어 및 데이터는 하나 이상의 컴퓨터 판독 가능 기록 매체에 저장될 수 있다.

[0076] 실시예에 따른 방법은 다양한 컴퓨터 수단을 통하여 수행될 수 있는 프로그램 명령 형태로 구현되어 컴퓨터 판독 가능 매체에 기록될 수 있다. 상기 컴퓨터 판독 가능 매체는 프로그램 명령, 데이터 파일, 데이터 구조 등을 단독으로 또는 조합하여 포함할 수 있다. 상기 매체에 기록되는 프로그램 명령은 실시예를 위하여 특별히 설계되고 구성된 것들이거나 컴퓨터 소프트웨어 당업자에게 공지되어 사용 가능한 것일 수도 있다. 컴퓨터 판독 가능 기록 매체의 예에는 하드 디스크, 플로피 디스크 및 자기 테이프와 같은 자기 매체(magnetic media), CD-ROM, DVD와 같은 광기록 매체(optical media), 플롭티컬 디스크(floptical disk)와 같은 자기-광 매체(magneto-optical media), 및 롬(ROM), 램(RAM), 플래시 메모리 등과 같은 프로그램 명령을 저장하고 수행하도록 특별히 구성된 하드웨어 장치가 포함된다. 프로그램 명령의 예에는 컴파일러에 의해 만들어지는 것과 같은 기계어 코드뿐만 아니라 인터프리터 등을 사용해서 컴퓨터에 의해서 실행될 수 있는 고급 언어 코드를 포함한다.

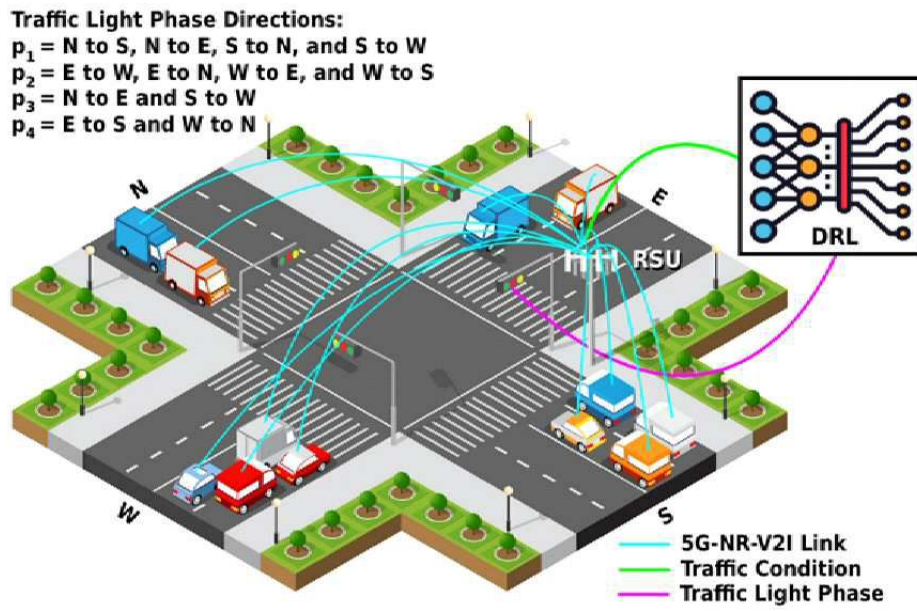
[0077] 이상과 같이 실시예들이 비록 한정된 실시예와 도면에 의해 설명되었으나, 해당 기술분야에서 통상의 지식을 가진 자라면 상기의 기재로부터 다양한 수정 및 변형이 가능하다. 예를 들어, 설명된 기술들이 설명된 방법과 다른 순서로 수행되거나, 및/또는 설명된 시스템, 구조, 장치, 회로 등의 구성요소들이 설명된 방법과 다른 형태로 결합 또는 조합되거나, 다른 구성요소 또는 균등물에 의하여 대치되거나 치환되더라도 적절한 결과가 달성될 수 있다.

[0078] 그러므로, 다른 구현들, 다른 실시예들 및 특허청구범위와 균등한 것들도 후술하는 특허청구범위의 범위에 속한

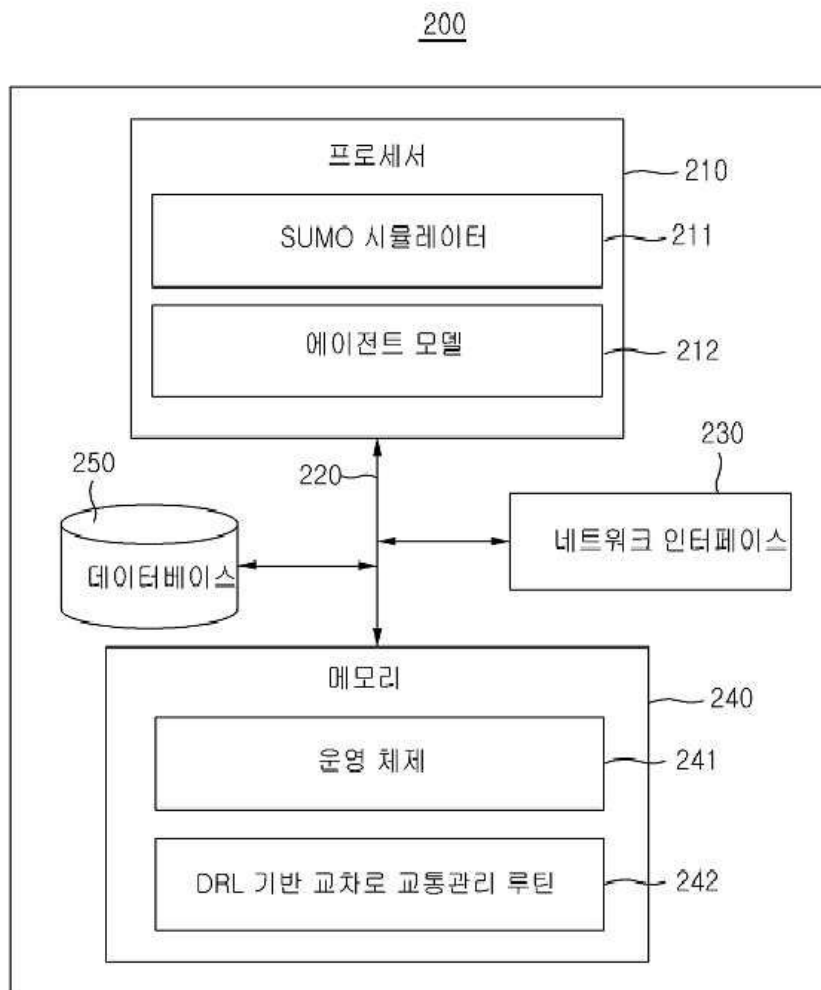
다.

도면

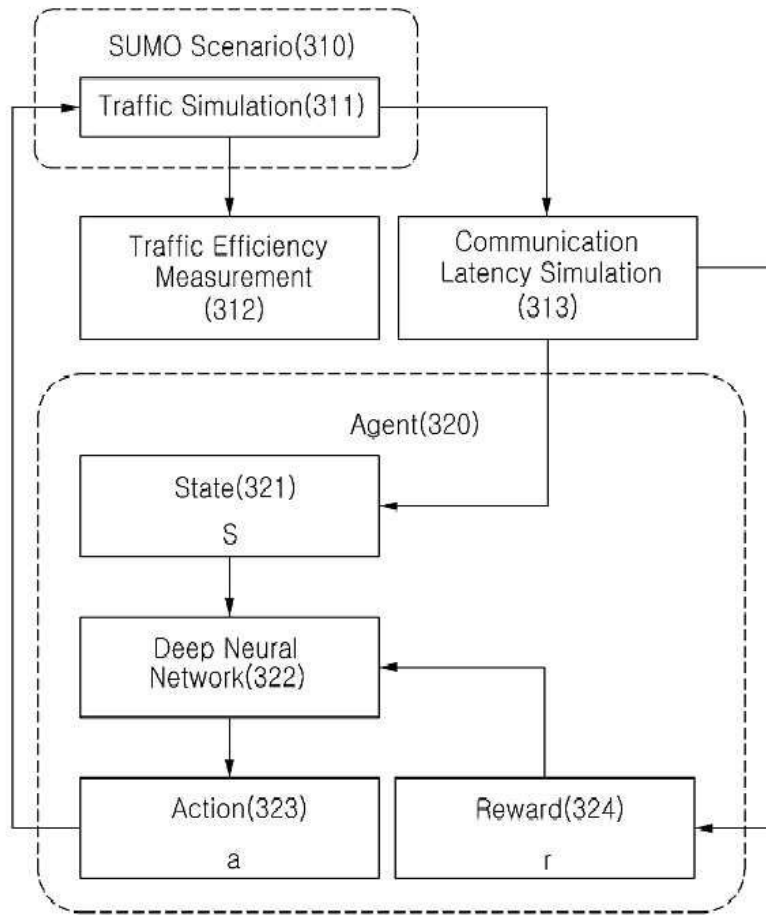
도면1



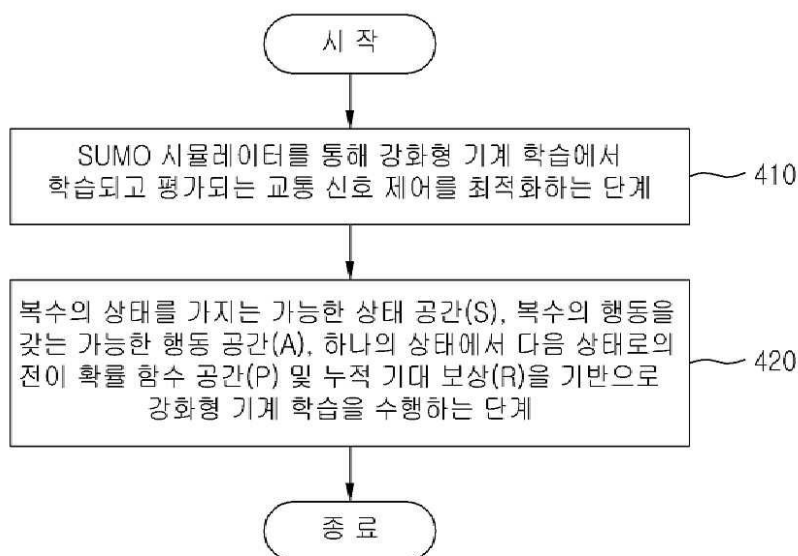
도면2



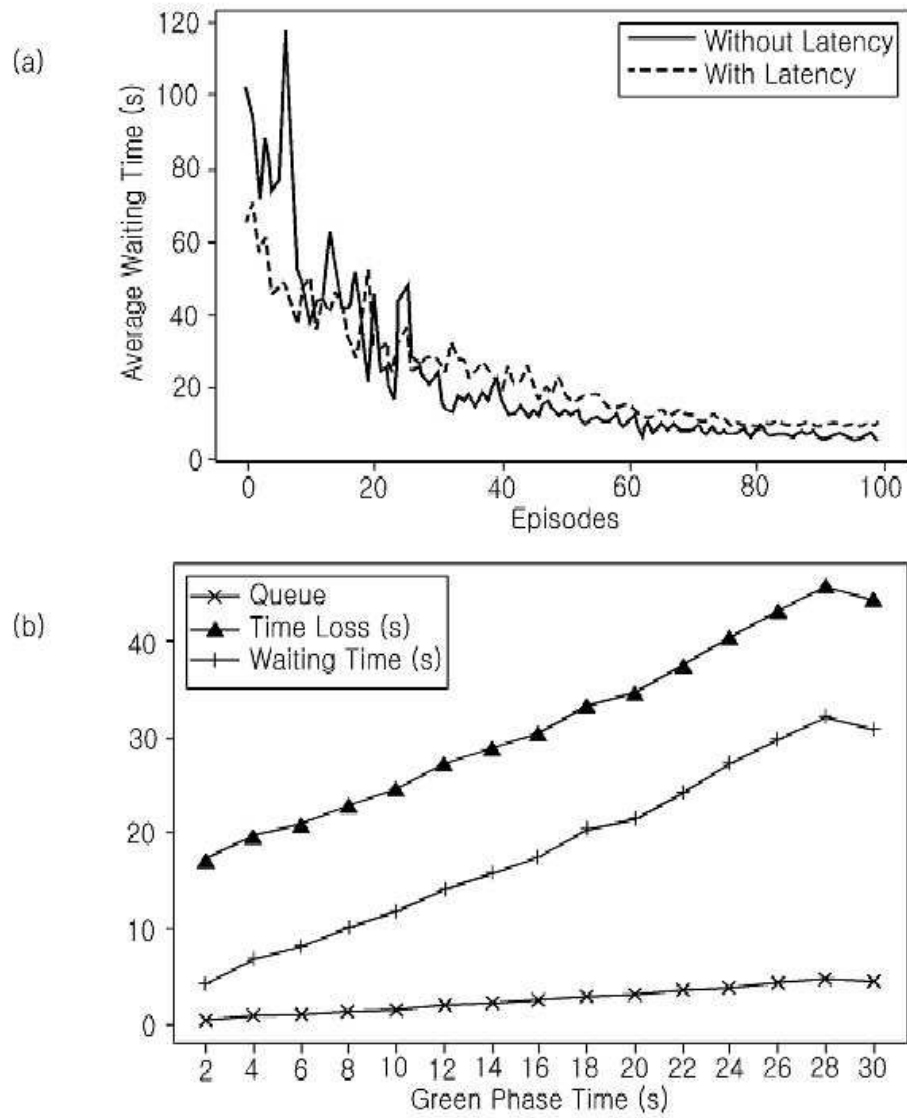
도면3



도면4



도면5



도면6

