



(19) 대한민국특허청(KR)
(12) 공개특허공보(A)

(11) 공개번호 10-2024-0121486
(43) 공개일자 2024년08월09일

(51) 국제특허분류(Int. Cl.)

H04W 24/02 (2009.01) G06N 20/00 (2019.01)
G06N 5/01 (2023.01) H04W 24/08 (2009.01)
H04W 28/16 (2019.01) H04W 4/40 (2018.01)
H04W 48/18 (2009.01) H04W 72/04 (2009.01)

(52) CPC특허분류

H04W 24/02 (2013.01)
G06N 20/00 (2021.08)

(21) 출원번호 10-2023-0014181

(22) 출원일자 2023년02월02일

심사청구일자 2023년02월02일

(71) 출원인

인하대학교 산학협력단

인천광역시 미추홀구 인하로 100(용현동, 인하대학교)

(72) 발명자

장경희

서울특별시 마포구 백범로 205, 101동 1110호(신공덕동, 펜트하우스)

주위건

인천광역시 미추홀구 한나루로463번길 99-4(용현동, E컬리지빌)

이슬람아자를

인천광역시 미추홀구 재념이길 118, 지하(용현동)

(74) 대리인

양성보

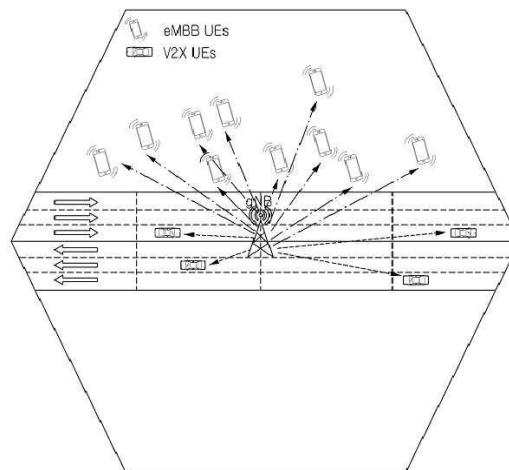
전체 청구항 수 : 총 8 항

(54) 발명의 명칭 V2X 및 eMBB 서비스를 위한 실시간 RL 기반 5G 네트워크 슬라이싱 설계

(57) 요약

V2X 및 eMBB 서비스를 위한 실시간 RL 기반 5G 네트워크 슬라이싱 설계 및 트래픽 모델 시스템 및 그 동작 방법이 제시된다. 본 발명에서 제안하는 V2X 및 eMBB 서비스를 위한 실시간 RL 기반 5G 네트워크 슬라이싱 설계 및 트래픽 모델 시스템은 무선 리소스 슬라이싱을 위한 네트워크 모델을 설계하는 네트워크 모델링부, V2X(Vehicle-to-Everything) 트래픽 모델 및 eMBB(enhanced Mobile Broadband) 트래픽 모델을 설계하는 트래픽 모델링부 및 V2X 트래픽 모델 및 eMBB 트래픽 모델에 대한 네트워크 리소스 할당을 최적화 하기 위한 강화 학습을 수행하는 강화 학습부를 포함한다.

대표도 - 도1



(52) CPC특허분류

G06N 5/01 (2023.01)
H04W 24/08 (2013.01)
H04W 28/16 (2019.01)
H04W 4/40 (2020.05)
H04W 48/18 (2013.01)
H04W 72/0446 (2023.01)
H04W 72/0453 (2023.01)

이 발명을 지원한 국가연구개발사업

과제고유번호	1711193423
과제번호	2022-0-00521-002
부처명	과학기술정보통신부
과제관리(전문)기관명	정보통신기획평가원
연구사업명	차세대자율주행차량통신기술개발
연구과제명	레벨4 이상 자율주행 서비스 지원을 위한 5G-NR-V2X 통신 기술개발
기 여 율	1/1
과제수행기관명	대보정보통신(주)
연구기간	2023.01.01 ~ 2023.12.31

명세서

청구범위

청구항 1

무선 리소스 슬라이싱을 위한 네트워크 모델을 설계하는 네트워크 모델링부;

V2X(Vehicle-to-Everything) 트래픽 모델 및 eMBB(enhanced Mobile Broadband) 트래픽 모델을 설계하는 트래픽 모델링부; 및

V2X 트래픽 모델 및 eMBB 트래픽 모델에 대한 네트워크 리소스 할당을 최적화 하기 위한 강화 학습을 수행하는 강화 학습부

를 포함하는 네트워크 슬라이싱 설계 및 트래픽 시스템.

청구항 2

제1항에 있어서,

상기 트래픽 모델링부는,

링크 유형, 차량이 전송한 메시지의 수, 다운링크의 스펙트럼 효율, 사용자의 각 전송 시간 간격(Transmission Time Interval; TTI) 및 평균 값을 계산하는 데 사용되는 시간 윈도우를 정의하는 TTI의 수를 이용하여 각 전송 시간 간격에 필요한 PRB(Physical Resource Block)의 평균 수를 나타내고, V2X 트래픽 모델을 설계하고,

다운링크에서 m번째 사용자가 필요한 비트 전송률을 얻기 위해 TTI에서 필요한 PRB 수를 이용하여 다운링크에 필요한 PRB의 평균 수를 나타내고, eMBB 트래픽 모델을 설계하는

네트워크 슬라이싱 설계 및 트래픽 시스템.

청구항 3

제1항에 있어서,

상기 강화 학습부는,

네트워크 슬라이싱을 위한 중단 확률이 최대 허용 한계보다 낮도록 보장하면서 다운링크 리소스 활용도를 최대 화하도록 각 슬라이스에 대한 슬라이싱 속도를 결정하기 위해 슬라이스 컨트롤러를 통해 일부 중간 슬라이싱 비율은 RL 알고리즘에 의해 결정하고, RL 알고리즘의 결과를 입력으로 취한 다음 슬라이싱 비율을 조정하여 최종 슬라이싱 비율을 얻는 휴리스틱 알고리즘을 이용하며, 상기 RL 알고리즘은 장기적으로 가능한 모든 동작의 탐색 및 이용을 실현하기 위한 소프트맥스 결정에 기반한 Q-러닝을 수행하고, 보상을 정의하는

네트워크 슬라이싱 설계 및 트래픽 시스템.

청구항 4

제3항에 있어서,

상기 강화 학습부는,

보상을 정의하기 위해 해당 슬라이싱 비율에 따라 선택된 동작에 대해, 해당 동작에 할당된 총 리소스에 사용된 리소스의 비율로 정의된 슬라이스에 대한 보상을 계산하는 데 정규화된 리소스 활용도 함수를 보상 함수로서 이용하고,

각 슬라이스의 보상을 최대화하기 위해 Q-러닝은 미리 정해진 시간에 이산 시간 동안 네트워크 모델과 상호 작용하고 선택된 동작에 대한 보상을 추정하여, Q-러닝의 결과에 따른 보상을 업데이트하는

네트워크 슬라이싱 설계 및 트래픽 시스템.

청구항 5

네트워크 모델링부를 통해 무선 리소스 슬라이싱을 위한 네트워크 모델을 설계하는 단계;

트래픽 모델링부를 통해 V2X 트래픽 모델 및 eMBB 트래픽 모델을 설계하는 단계; 및

강화 학습부를 통해 V2X 트래픽 모델 및 eMBB 트래픽 모델에 대한 네트워크 리소스 할당을 최적화 하기 위한 강화 학습을 수행하는 단계

를 포함하는 네트워크 슬라이싱 설계 및 트래픽 시스템의 동작 방법.

청구항 6

제5항에 있어서,

상기 트래픽 모델링부를 통해 V2X 트래픽 모델 및 eMBB 트래픽 모델을 설계하는 단계는,

링크 유형, 차량이 전송한 메시지의 수, 다운링크의 스펙트럼 효율, 사용자의 각 전송 시간 간격(Transmission Time Interval; TTI) 및 평균 값을 계산하는 데 사용되는 시간 윈도우를 정의하는 TTI의 수를 이용하여 각 전송 시간 간격에 필요한 PRB(Physical Resource Block)의 평균 수를 나타내고, V2X 트래픽 모델을 설계하고,

다운링크에서 m번째 사용자가 필요한 비트 전송률을 얻기 위해 TTI에서 필요한 PRB 수를 이용하여 다운링크에 필요한 PRB의 평균 수를 나타내고, eMBB 트래픽 모델을 설계하는

네트워크 슬라이싱 설계 및 트래픽 시스템의 동작 방법.

청구항 7

제5항에 있어서,

상기 강화 학습부를 통해 V2X 트래픽 모델 및 eMBB 트래픽 모델에 대한 네트워크 리소스 할당을 최적화 하기 위한 강화 학습을 수행하는 단계는,

네트워크 슬라이싱을 위한 중단 확률이 최대 허용 한계보다 낮도록 보장하면서 다운링크 리소스 활용도를 최대화하도록 각 슬라이스에 대한 슬라이싱 속도를 결정하기 위해 슬라이스 컨트롤러를 통해 일부 중간 슬라이싱 비율은 RL 알고리즘에 의해 결정하고, RL 알고리즘의 결과를 입력으로 취한 다음 슬라이싱 비율을 조정하여 최종 슬라이싱 비율을 얻는 휴리스틱 알고리즘을 이용하며, 상기 RL 알고리즘은 장기적으로 가능한 모든 동작의 탐색 및 이용을 실현하기 위한 소프트맥스 결정에 기반한 Q-러닝을 수행하고, 보상을 정의하는

네트워크 슬라이싱 설계 및 트래픽 시스템의 동작 방법.

청구항 8

제7항에 있어서,

상기 강화 학습부를 통해 V2X 트래픽 모델 및 eMBB 트래픽 모델에 대한 네트워크 리소스 할당을 최적화 하기 위한 강화 학습을 수행하는 단계는,

보상을 정의하기 위해 해당 슬라이싱 비율에 따라 선택된 동작에 대해, 해당 동작에 할당된 총 리소스에 사용된 리소스의 비율로 정의된 슬라이스에 대한 보상을 계산하는 데 정규화된 리소스 활용도 함수를 보상 함수로서 이용하고,

각 슬라이스의 보상을 최대화하기 위해 Q-러닝은 미리 정해진 시간에 이산 시간 동안 네트워크 모델과 상호 작용하고 선택된 동작에 대한 보상을 추정하여, Q-러닝의 결과에 따른 보상을 업데이트하는

네트워크 슬라이싱 설계 및 트래픽 시스템의 동작 방법.

발명의 설명

기술 분야

본 발명은 V2X 및 eMBB 서비스를 위한 실시간 RL 기반 5G 네트워크 슬라이싱 설계 및 트래픽 모델 및 그 동작 방법에 관한 것이다.

[0001]

배경 기술

- [0002] 5G는 무선 기술의 업그레이드일 뿐만 아니라 네트워크 서비스 아키텍처의 변화이기도 하다. 종래기술에 따르면, 5G의 세 가지 핵심 애플리케이션은 eMBB(enhanced Mobile Broadband), mMTC(massive Machine Type Communications), uRLLC(ultra Reliable Low Latency Communications)이다. 5G는 4G와 5G의 주요 차이점인 대규모 연결과 극도로 낮은 대기 시간을 지원한다. 대규모 단말기 사용 시나리오는 네트워크 속도와 대기 시간에 대한 요구사항은 진화하고 있으며, 5G는 이를 통합 네트워크 아키텍처로 통합한다. 방대한 네트워크 요구사항에 대처하기 위해, 5G 네트워크 슬라이싱이 핵심 기술로 도입되었다.
- [0003] 네트워크 슬라이싱은 물리적 네트워크를 다중 논리 네트워크로 분할하여 성능 및 유연성 요구사항이 서로 다른 다양한 애플리케이션을 지원하도록 설계되었다. 강화 학습(Reinforcement Learning; RL)은 네트워크와의 상호 작용을 통해 경험에서 지식을 추출하고, 리소스 활용률을 극대화하도록 슬라이스에 할당된 리소스를 동적으로 조정하는 동시에 서비스 품질(QoS)을 보장한다.
- [0004] 종래기술들은 네트워크 슬라이스를 설계하고 제어하기 위한 솔루션을 제안했다. 네트워크 슬라이싱을 기반으로 하는 5G 시스템을 위한 논리적 아키텍처 및 서로 다른 액세스 네트워크 간의 이동성을 관리하는 방식이 종래기술[1]에서 제안된다. 차례로, EPC(Evolved Packet Core)를 개선하기 위한 도구로서 기능 분해 및 네트워크 슬라이싱이 종래기술[2]에서 제안된다. 그것들은 네트워크 슬라이싱과 기능 분해를 기반으로 모바일 코어 네트워크의 실현 가능성을 조사한다. 종래기술[3]은 southbound API를 통해 제어부 및 데이터부를 분리하고 RAN 인프라 엔티티 간에 서로 다른 수준의 조정을 가능하게 함으로써 RAN 제어 애플리케이션을 위한 제어부를 지원하는 데 중점을 둔 유연하고 프로그래밍 가능한 소프트웨어 정의(Software-Defined; SD)-RAN 플랫폼을 제시한다. 종래기술[4]은 각 네트워크 슬라이스에 대한 네트워크 슬라이스 트래픽 분석 및 예측, 네트워크 슬라이스 요청에 대한 승인 제어 결정에 초점을 맞추고, 측정 편차를 기반으로 예측 부하의 적응 보정을 제안한다. 그리고 종래기술[5]에서는 가상 무선 네트워크의 공동 승인 제어를 위한 저복잡도 휴리스틱 알고리즘 및 슬라이싱을 제안한다. 종래기술[6]에서는 서비스 수요 및 사용 가능한 리소스에 기반한 네트워크 슬라이싱 조정 모델을 소개하고, 5G 네트워크에 대한 교차 슬라이스 승인 제어 및 리소스 할당을 관리하기 위한 최선의 정책을 개발하고 결정하기 위한 마르코프 의사 결정 프로세스를 위한 프레임워크를 제안한다. 종래기술[7]에서는 강화 학습을 기반으로 하는 특정 슬라이싱 체계와 휴리스틱 최적화 솔루션을 제안한다.

선행기술문헌

특허문헌

- [0005] (특허문헌 0001) [1] Service Requirements for the 5G System; Stage 1 (Release 15), document 3GPP TS 22.261 v16.0.0, Jun. 2018.
- (특허문헌 0002) [2] Management of Network Slicing in Mobile Networks; Concepts, Use Cases and Requirements (Release 15), document 3GPP TS 28.530 v0.3.0, Aug. 2018.
- (특허문헌 0003) [3] System Architecture for the 5G System; Stage 2 (Release 15), document 3GPP TS 23.501 v1.0.0, Sep. 2018.
- (특허문헌 0004) [4] "White Paper: 5G Radio Access Capabilities and Technologies," Ericsson, White Paper Uen 284 23-3204 Rev C, 2016. Accessed: Apr. 5, 2019. [Online]. Available: <https://www.ericsson.com/assets/local/publications/white-papers/wp-5g.pdf>
- (특허문헌 0005) [5] Evolved Universal Terrestrial Radio Access (E-UTRA) and Evolved Universal Terrestrial Radio Access Network (E-UTRAN); Overall Description; Stage 2 (Release 15), document 3GPP TS 36.300 v15.0.0, Dec. 2017.
- (특허문헌 0006) [6] NR; Study on Vehicle-to-Everything (Release 16), document 3GPP TR 38.885 v1.0.0, Nov. 2018.
- (특허문헌 0007) [7] Albonda H D R, Perez-Romero J. An efficient RAN slicing strategy for a heterogeneous network with eMBB and V2X services[J]. IEEE access, 2019, 7: 44771-44782.

발명의 내용

해결하려는 과제

- [0006] 본 발명이 이루고자 하는 기술적 과제는 5G에 대한 두 가지 일반적인 서비스, 즉 eMBB와 셀룰러 V2X(Vehicle-to-Everything)를 제공하는 RAN 슬라이싱 문제에 대하여 네트워크 모델 특성 수집을 기반으로 온라인 네트워크 특성의 실시간 추정을 적용할 것인지, 그리고 무선 리소스를 다른 슬라이스에 할당하기 위해 강화 학습을 기반으로 한 적절한 RAN 슬라이싱 체계를 적용할 것인지 여부에 따라 두 가지 다른 체계인 실시간 RL 기반 5G 네트워크 슬라이싱 설계 및 트래픽 모델 분포를 이용한 V2X 및 eMBB 서비스 방법 및 시스템을 제공하는데 있다.

과제의 해결 수단

- [0007] 일 측면에 있어서, 본 발명에서 제안하는 V2X 및 eMBB 서비스를 위한 실시간 RL 기반 5G 네트워크 슬라이싱 설계 및 트래픽 모델 시스템은 무선 리소스 슬라이싱을 위한 네트워크 모델을 설계하는 네트워크 모델링부, V2X(Vehicle-to-Everything) 트래픽 모델 및 eMBB(enhanced Mobile Broadband) 트래픽 모델을 설계하는 트래픽 모델링부 및 V2X 트래픽 모델 및 eMBB 트래픽 모델에 대한 네트워크 리소스 할당을 최적화 하기 위한 강화 학습을 수행하는 강화 학습부를 포함한다.
- [0008] 상기 트래픽 모델링부는 링크 유형, 차량이 전송한 메시지의 수, 다운링크의 스펙트럼 효율, 사용자의 각 전송 시간 간격(Transmission Time Interval; TTI) 및 평균 값을 계산하는 데 사용되는 시간 윈도우를 정의하는 TTI의 수를 이용하여 각 전송 시간 간격에 필요한 PRB(Physical Resource Block)의 평균 수를 나타내고, V2X 트래픽 모델을 설계하고, 다운링크에서 m번째 사용자가 필요한 비트 전송률을 얻기 위해 TTI에서 필요한 PRB 수를 이용하여 다운링크에 필요한 PRB의 평균 수를 나타내고, eMBB 트래픽 모델을 설계한다.
- [0009] 상기 강화 학습부는 네트워크 슬라이싱을 위한 중단 확률이 최대 허용 한계보다 낮도록 보장하면서 다운링크 리소스 활용도를 최대화하도록 각 슬라이스에 대한 슬라이싱 속도를 결정하기 위해 슬라이스 컨트롤러를 통해 일부 중간 슬라이싱 비율은 RL 알고리즘에 의해 결정하고, RL 알고리즘의 결과를 입력으로 취한 다음 슬라이싱 비율을 조정하여 최종 슬라이싱 비율을 얻는 휴리스틱 알고리즘을 이용하며, 상기 RL 알고리즘은 장기적으로 가능한 모든 동작의 탐색 및 이용을 실현하기 위한 소프트맥스 결정에 기반한 Q-러닝을 수행하고, 보상을 정의한다.
- [0010] 상기 강화 학습부는 보상을 정의하기 위해 해당 슬라이싱 비율에 따라 선택된 동작에 대해, 해당 동작에 할당된 총 리소스에 사용된 리소스의 비율로 정의된 슬라이스에 대한 보상을 계산하는 데 정규화된 리소스 활용도 함수를 보상 함수로서 이용하고, 각 슬라이스의 보상을 최대화하기 위해 Q-러닝은 미리 정해진 시간에 이산 시간 동안 네트워크 모델과 상호 작용하고 선택된 동작에 대한 보상을 추정하여, Q-러닝의 결과에 따른 보상을 업데이트한다.
- [0011] 또 다른 일 측면에 있어서, 본 발명에서 제안하는 V2X 및 eMBB 서비스를 위한 실시간 RL 기반 5G 네트워크 슬라이싱 설계 및 트래픽 모델 시스템의 동작 방법은 네트워크 모델링부를 통해 무선 리소스 슬라이싱을 위한 네트워크 모델을 설계하는 단계, 트래픽 모델링부를 통해 V2X 트래픽 모델 및 eMBB 트래픽 모델을 설계하는 단계 및 강화 학습부를 통해 V2X 트래픽 모델 및 eMBB 트래픽 모델에 대한 네트워크 리소스 할당을 최적화 하기 위한 강화 학습을 수행하는 단계를 포함한다.

발명의 효과

- [0012] 본 발명의 실시예들에 따르면 5G에 대한 두 가지 일반적인 서비스, 즉 eMBB와 셀룰러 V2X(Vehicle-to-Everything)를 제공하는 RAN 슬라이싱 문제에 대하여 네트워크 모델 특성 수집을 기반으로 온라인 네트워크 특성의 실시간 추정을 적용할 것인지, 그리고 무선 리소스를 다른 슬라이스에 할당하기 위해 강화 학습을 기반으로 한 적절한 RAN 슬라이싱 체계를 적용할 것인지 여부에 따라 두 가지 다른 체계인 실시간 RL 기반 5G 네트워크 슬라이싱 설계 및 트래픽 모델 분포를 이용한 V2X 및 eMBB 서비스 방법 및 시스템을 제공할 수 있다. 이를 통해 유틸리티 요구사항과 트래픽 부하의 동적 변화를 고려하여 리소스 활용 효율성을 극대화할 수 있다.

도면의 간단한 설명

- [0013] 도 1은 본 발명의 일 실시예에 따른 셀룰러 네트워크 시스템 모델을 나타내는 도면이다.
- 도 2는 본 발명의 일 실시예에 따른 V2X 및 eMBB 서비스를 위한 실시간 RL 기반 5G 네트워크 슬라이싱 설계 및

트래픽 시스템의 구성을 나타내는 도면이다.

도 3은 본 발명의 일 실시예에 따른 V2X 및 eMBB 서비스를 위한 실시간 RL 기반 5G 네트워크 슬라이싱 설계 및 트래픽 시스템의 동작 방법을 설명하기 위한 흐름도이다.

도 4는 본 발명의 일 실시예에 따른 RL 기반 RAN 슬라이싱 기법을 설명하기 위한 도면이다.

도 5는 본 발명의 일 실시예에 따른 서로 다른 에피소드들에 대한 Q-값을 나타내는 그래프이다.

도 6은 본 발명의 일 실시예에 따른 시뮬레이션 라운드에 따른 PRB 활용을 나타내는 그래프이다.

발명을 실시하기 위한 구체적인 내용

- [0014] 5G 모바일 시스템은 여러 서비스와 애플리케이션을 전송하므로 5G 네트워크 설계의 주요 과제는 단일 물리적 네트워크 인프라 내에서 서비스 품질 요구사항이 다른 다양한 유형의 사용자와 애플리케이션을 지원하는 방법이다. 이러한 과제를 해결하기 위한 방법으로 RAN(Radio Access Network) 슬라이싱이 제안된다. 본 발명은 5G에 대한 두 가지 일반적인 서비스, 즉 eMBB와 셀룰러 V2X(Vehicle-to-Everything)를 제공하는 RAN 슬라이싱 문제를 개시하고, 네트워크 모델 특성 수집을 기반으로 두 가지 다른 체계를 제안한다. 즉, 온라인 네트워크 특성의 실시간 추정을 적용할 것인지, 그리고 무선 리소스를 다른 슬라이스에 할당하기 위해 강화 학습을 기반으로 한 적절한 RAN 슬라이싱 체계를 적용할 것인지 여부이다. 이는 유틸리티 요구사항과 트래픽 부하의 동적 변화를 고려하여 리소스 활용 효율성을 극대화한다. 이하, 본 발명의 실시 예를 첨부된 도면을 참조하여 상세하게 설명한다.
- [0016] 도 1은 본 발명의 일 실시예에 따른 셀룰러 네트워크 시스템 모델을 나타내는 도면이다.
- [0017] 본 발명의 실시예에 따른 시나리오는 셀룰러 차세대 RAN 액세스 네트워크(Next Generation Radio Access Network; NG-RAN)가 단일 셀로 구성된 gNB(gNodeB)를 갖는 것이다. eMBB 셀룰러 사용자 그룹은 gNB 주위에 랜덤으로 분포하며, $m = 1, \dots, M$ 으로 번호가 매겨진다. 그리고 도 1에서와 같이 여러 독립적인 차량 흐름이 직선 고속도로를 따라 이동한다. 본 발명의 실시예에서 고속도로는 6개 차선으로 구성되어 있으며, 좌측에서 우측으로 3개 차선이 있고, 나머지 3개 차선의 주행 방향은 우측에서 좌측으로 설정되어 있다. 고속도로의 길이에 따라 더 작은 구역으로 나뉘기 때문에 고속도로 구간은 클러스터로 나뉜다. 각 차량은 동일한 클러스터 내의 다른 차량의 UE(User Equipment)와 통신할 수 있는 UE를 갖는 것으로 가정한다. 클러스터의 번호는 $j = 1, \dots, C$ 이고, j 번째 클러스터의 차량 번호는 $i = 1, \dots, V(j)$ 이다.
- [0018] 고속도로를 달리는 차량은 도착 속도가 λ_a 인 포아송 프로세스에 따라 셀로 들어가는 것으로 간주된다. eMBB UE의 수는 생성률이 λ_m 인 포아송 프로세스에 의해 생성된다.
- [0020] 도 2는 본 발명의 일 실시예에 따른 V2X 및 eMBB 서비스를 위한 실시간 RL 기반 5G 네트워크 슬라이싱 설계 및 트래픽 시스템의 구성을 나타내는 도면이다.
- [0021] 본 실시예에 따른 네트워크 슬라이싱 설계 및 트래픽 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 V2X 및 eMBB 서비스를 위한 실시간 RL 기반 5G 네트워크 슬라이싱 설계 및 트래픽 모델링 루틴(242)을 포함할 수 있다. 프로세서(210)는 네트워크 모델링부(211), 트래픽 모델링부(212) 및 강화 학습부(213)를 포함할 수 있다. 다른 실시예들에서 네트워크 슬라이싱 설계 및 트래픽 시스템(200)은 도 2의 구성요소들보다 더 많은 구성요소들을 포함할 수도 있다. 그러나, 대부분의 종래기술적 구성요소들을 명확하게 도시할 필요성은 없다. 예를 들어, 네트워크 슬라이싱 설계 및 트래픽 시스템(200)은 디스플레이나 트랜시버(transceiver)와 같은 다른 구성요소들을 포함할 수도 있다.
- [0022] 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 V2X 및 eMBB 서비스를 위한 실시간 RL 기반 5G 네트워크 슬라이싱 설계 및 트래픽 모델링 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.

[0023] 버스(220)는 네트워크 슬라이싱 설계 및 트래픽 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.

[0024] 네트워크 인터페이스(230)는 네트워크 슬라이싱 설계 및 트래픽 시스템(200)을 컴퓨터 네트워크에 연결하기 위한 컴퓨터 하드웨어 구성요소일 수 있다. 네트워크 인터페이스(230)는 네트워크 슬라이싱 설계 및 트래픽 시스템(200)을 무선 또는 유선 커넥션을 통해 컴퓨터 네트워크에 연결시킬 수 있다.

[0025] 데이터베이스(250)는 V2X 및 eMBB 서비스를 위한 실시간 RL 기반 5G 네트워크 슬라이싱 설계 및 트래픽 모델링을 위해 필요한 모든 정보를 저장 및 유지하는 역할을 할 수 있다. 도 2에서는 네트워크 슬라이싱 설계 및 트래픽 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.

[0026] 프로세서(210)는 기본적인 산술, 로직 및 네트워크 슬라이싱 설계 및 트래픽 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 네트워크 모델링부(211), 트래픽 모델링부(212) 및 강화 학습부(213)를 위한 프로그램 코드를 실행하도록 구성될 수 있다. 이러한 프로그램 코드는 메모리(240)와 같은 기록 장치에 저장될 수 있다.

[0027] 네트워크 모델링부(211), 트래픽 모델링부(212) 및 강화 학습부(213)는 도 3의 단계들(310~330)을 수행하기 위해 구성될 수 있다.

[0028] 네트워크 슬라이싱 설계 및 트래픽 시스템(200)은 네트워크 모델링부(211), 트래픽 모델링부(212) 및 강화 학습부(213)를 포함할 수 있다.

[0029] 본 발명의 실시예에 따른 네트워크 모델링부(211)는 무선 리소스 슬라이싱을 위한 네트워크 모델을 설계한다.

[0030] 본 발명의 실시예에 따른 트래픽 모델링부(212)는 V2X(Vehicle-to-Everything) 트래픽 모델 및 eMBB(enhanced Mobile Broadband) 트래픽 모델을 설계한다.

[0031] 본 발명의 실시예에 따른 트래픽 모델링부(212)는 링크 유형, 차량이 전송한 메시지의 수, 다운로드의 스펙트럼 효율, 사용자의 각 전송 시간 간격(Transmission Time Interval; TTI) 및 평균 값을 계산하는 데 사용되는 시간 윈도우를 정의하는 TTI의 수를 이용하여 각 전송 시간 간격에 필요한 PRB(Physical Resource Block)의 평균 수를 나타내고, V2X 트래픽 모델을 설계한다.

[0032] 본 발명의 실시예에 따른 트래픽 모델링부(212)는 다운로드에서 m번째 사용자가 필요한 비트 전송률을 얻기 위해 TTI에서 필요한 PRB 수를 이용하여 다운로드에 필요한 PRB의 평균 수를 나타내고, eMBB 트래픽 모델을 설계한다.

[0033] 본 발명의 실시예에 따른 강화 학습부(213)는 V2X 트래픽 모델 및 eMBB 트래픽 모델에 대한 네트워크 리소스 할당을 최적화 하기 위한 강화 학습을 수행한다.

[0034] 본 발명의 실시예에 따른 강화 학습부(213)는 네트워크 슬라이싱을 위한 중단 확률이 최대 허용 한계보다 낮도록 보장하면서 다운로드 리소스 활용도를 최대화하도록 각 슬라이스에 대한 슬라이싱 속도를 결정하기 위해 슬라이스 컨트롤러를 통해 일부 중간 슬라이싱 비율은 RL 알고리즘에 의해 결정하고, RL 알고리즘의 결과를 입력으로 취한 다음 슬라이싱 비율을 조정하여 최종 슬라이싱 비율을 얻는 휴리스틱 알고리즘을 이용하며, 상기 RL 알고리즘은 장기적으로 가능한 모든 동작의 탐색 및 이용을 실현하기 위한 소프트맥스 결정에 기반한 Q-러닝을 수행하고, 보상을 정의한다.

[0035] 본 발명의 실시예에 따른 강화 학습부(213)는 보상을 정의하기 위해 해당 슬라이싱 비율에 따라 선택된 동작에 대해, 해당 동작에 할당된 총 리소스에 사용된 리소스의 비율로 정의된 슬라이스에 대한 보상을 계산하는 데 정규화된 리소스 활용도 함수를 보상 함수로서 이용한다.

- [0036] 본 발명의 실시예에 따른 강화 학습부(213)는 각 슬라이스의 보상을 최대화하기 위해 Q-러닝은 미리 정해진 시간에 이산 시간 동안 네트워크 모델과 상호 작용하고 선택된 동작에 대한 보상을 추정하여, Q-러닝의 결과에 따른 보상을 업데이트한다.
- [0038] 도 3은 본 발명의 일 실시예에 따른 V2X 및 eMBB 서비스를 위한 실시간 RL 기반 5G 네트워크 슬라이싱 설계 및 트래픽 시스템의 동작 방법을 설명하기 위한 흐름도이다.
- [0039] 제안하는 V2X 및 eMBB 서비스를 위한 실시간 RL 기반 5G 네트워크 슬라이싱 설계 및 트래픽 시스템의 동작 방법은 네트워크 모델링부를 통해 무선 리소스 슬라이싱을 위한 네트워크 모델을 설계하는 단계(310), 트래픽 모델링부를 통해 V2X 트래픽 모델 및 eMBB 트래픽 모델을 설계하는 단계(320) 및 강화 학습부를 통해 V2X 트래픽 모델 및 eMBB 트래픽 모델에 대한 네트워크 리소스 할당을 최적화 하기 위한 강화 학습을 수행하는 단계(330)를 포함한다.
- [0040] 단계(310)에서, 네트워크 모델링부를 통해 무선 리소스 슬라이싱을 위한 네트워크 모델을 설계한다.
- [0041] 본 발명의 실시예에 따른 V2X와 eMBB 서비스를 함께 지원하고, eMBB에 대한 높은 데이터 속도 서비스를 위해서는 큰 대역폭을 필요로 하기 때문에 V2X 서비스는 지연 시간에 매우 민감하다. 본 발명의 실시예에 따른 네트워크는 논리적으로 두 개의 네트워크 슬라이스, 즉 V2X 서비스의 경우 RAN_slice_ID = 1, eMBB 서비스의 경우 RAN_slice_ID = 2이다.
- [0042] 전체 셀 대역폭은 대역폭 B인 리소스 블록(Resource Blocks; RB)으로 구성된다. 다운링크의 RB 수를 N_{DL} 로 나타낸다. 슬라이싱 프로세스는 두 슬라이스 사이에 DL RB를 할당해야 한다. 이를 위해, RAN_slice_ID = s에 대한 DL 리소스의 일부분으로서 $\alpha_{s,DL}$ 을 나타내고, 여기서 s = 1,2 이다. 관계식은 다음과 같다:
- [0043]
$$\sum_s \alpha_{s,DL} = 1 \quad (1)$$
- [0044] 단계(320)에서, 트래픽 모델링부를 통해 V2X 트래픽 모델 및 eMBB 트래픽 모델을 설계한다.
- [0045] 본 발명의 실시예에 따른 트래픽 모델링부는 링크 유형, 차량이 전송한 메시지의 수, 다운링크의 스펙트럼 효율, 사용자의 각 전송 시간 간격(Transmission Time Interval; TTI) 및 평균 값을 계산하는 데 사용되는 시간 윈도우를 정의하는 TTI의 수를 이용하여 각 전송 시간 간격에 필요한 PRB(Physical Resource Block)의 평균 수를 나타내고, V2X 트래픽 모델을 설계한다.
- [0046] 본 발명의 실시예에 따른 트래픽 모델링부는 다운링크에서 m번째 사용자가 필요한 비트 전송률을 얻기 위해 TTI에서 필요한 PRB 수를 이용하여 다운링크에 필요한 PRB의 평균 수를 나타내고, eMBB 트래픽 모델을 설계한다.
- [0047] 슬라이싱 전략이 없는 트래픽 모델 설계에 있어서, V2X 트래픽 모델은 다음과 같이 설계할 수 있다.
- [0048] 포아송 도착 모델에 따르면, 각 차량은 λ_{niu} (sessions/s)의 비율로 랜덤 패킷을 생성한다고 가정한다. 메시지의 크기는 S_m 이다. 셀룰러 모드에서 메시지는 DL 리소스를 사용하여 전송된다. DL에서 RAN_slice_ID = 1를 사용하는 V2X 사용자의 각 전송 시간 간격(Transmission Time Interval; TTI)에 필요한 PRB(Physical Resource Block)의 평균 수는 $\tau_{1,DL}$ 로 나타내며, 다음과 같이 쓸 수 있다:
- [0049]
$$\tau_{1,x} = \frac{\sum_{t=1}^T \sum_{j=1}^C \sum_{i=1}^{V(j)} \frac{m(j,i,t) \cdot S_m}{SP_{eff}(j,i,t)}}{T \cdot B \cdot F_d} \quad (2)$$
- [0050] 여기서 $x \in \{DL\}$ 은 링크 유형을 나타내며, m(j,i,t)은 차량이 전송한 메시지의 수이고, j번째 클러스터와 t번째 TTI의 경우 $SP_{eff,x}$ 는 다운링크의 스펙트럼 효율이며, TTI 지속 시간은 F_d 로 표시되고, T는 평균 값을 계산하는 데 사용되는 시간 윈도우를 정의하는 TTI의 수이다.
- [0051] 다음으로 슬라이싱 전략이 없는 트래픽 모델 설계에 있어서, eMBB 트래픽 모델은 다음과 같이 설계할 수 있다.
- [0052] eMBB 서비스와 관련하여, 각 eMBB 사용자는 일정한 보장 비트 전송률을 필요로 하는 세션을 생성한다. 포아송 프로세스는 세션 생성 모델로 λ_e (sessions/s)와 T_e 에 걸쳐 평균화된 세션 지속 시간의 지수 분포의 비율을 따

른다. 이러한 사용자들은 다운링크에서 eMBB 슬라이스에 할당된 PRB를 전송한다. 특정 보장 비트 전송률 R_b 를 지원하려면 $RAN_slice_ID = 2$ 인 eMBB 사용자에게 대하여 DL에 필요한 PRB의 평균 수를 $\tau_{2,DL}$ 로 나타내며 다음과 같이 쓸 수 있다:

$$\tau_{2,x} = \frac{\sum_{t=1}^T \sum_{m=1}^M \rho_x(m,t)}{T} \quad (3)$$

여기서 $x \in \{DL\}$, $\rho_x(m,t)$ 는 다운링크에서 m 번째 사용자가 필요한 비트 전송률 R_b $\rho_x(m,t) = R_b / (SP_{eff,x} * B)$ 를 얻기 위해 t 번째 TTI에서 필요한 PRB 수이다. $\tau_{2,DL}$ 은 시간 윈도우 T TTI 내에서 계산된다.

이후, RAN 슬라이싱을 위한 중단(outage) 문제를 다음과 같이 공식화한다.

목표는 슬라이스 사용자에게 대한 리소스 요구사항을 충족해야 하는 제약 조건에 따라 전체 리소스 활용도를 극대화하기 위한 최적의 슬라이스 비율 $\alpha_{s,DL}$ 을 결정하는 것이다.

DL 리소스 UDL의 총 활용도와 관련하여, 슬라이스의 합이 RAN 슬라이싱에 의해 이 슬라이스에 할당된 총 리소스, 즉 $\alpha_{s,DL} \cdot N_{DL}$ 을 초과하지 않는 제약 조건이 있다. 그렇지 않으면 중단 상황이 발생하고 활용률이 $\alpha_{s,DL} \cdot N_{DL}$ 로 제한되므로 다음과 같이 정의할 수 있다:

$$U_{DL} = \sum_s \min(\tau_{s,DL}, \alpha_{s,DL} \cdot N_{DL}^{DL}) \quad (4)$$

따라서 다운링크 최적화 문제는 중단 확률이 최대 허용 한계 P_{out} 보다 낮도록 보장하면서 DL 리소스 활용도를 최대화하는 것으로 정의된다. 이 문제는 다음과 같이 공식적으로 표현된다:

$$\max_{\alpha_{s,DL}} U_{DL} \quad (5)$$

$$\text{s.t. } \Pr[\tau_{s,DL} \geq \alpha_{s,DL} \cdot N_{DL}] < p_{out} \quad s = 1, 2 \quad (5a)$$

$$\sum_s \alpha_{s,DL} = 1 \quad (5b)$$

단계(330)에서, 강화 학습부를 통해 V2X 트래픽 모델 및 eMBB 트래픽 모델에 대한 네트워크 리소스 할당을 최적화 하기 위한 강화 학습을 수행한다.

본 발명의 실시예에 따른 강화 학습부는 네트워크 슬라이싱을 위한 중단 확률이 최대 허용 한계보다 낮도록 보장하면서 다운링크 리소스 활용도를 최대화하도록 각 슬라이스에 대한 슬라이싱 속도를 결정하기 위해 슬라이스 컨트롤러를 통해 일부 중간 슬라이싱 비율은 RL 알고리즘에 의해 결정하고, RL 알고리즘의 결과를 입력으로 취한 다음 슬라이싱 비율을 조정하여 최종 슬라이싱 비율을 얻는 휴리스틱 알고리즘을 이용하며, 상기 RL 알고리즘은 장기적으로 가능한 모든 동작의 탐색 및 이용을 실현하기 위한 소프트맥스 결정에 기반한 Q-러닝을 수행하고, 보상을 정의한다.

본 발명의 실시예에 따른 강화 학습부는 보상을 정의하기 위해 해당 슬라이싱 비율에 따라 선택된 동작에 대해, 해당 동작에 할당된 총 리소스에 사용된 리소스의 비율로 정의된 슬라이스에 대한 보상을 계산하는 데 정규화된 리소스 활용도 함수를 보상 함수로서 이용한다.

본 발명의 실시예에 따른 강화 학습부는 각 슬라이스의 보상을 최대화하기 위해 Q-러닝은 미리 정해진 시간에 이산 시간 동안 네트워크 모델과 상호 작용하고 선택된 동작에 대한 보상을 추정하여, Q-러닝의 결과에 따른 보상을 업데이트한다.

도 4는 본 발명의 일 실시예에 따른 RL 기반 RAN 슬라이싱 기법을 설명하기 위한 도면이다.

[0070] 본 발명의 실시예에 따른 RL 알고리즘을 이용한 네트워크 슬라이싱 전략에 있어서, RAN 슬라이싱은 본질적으로 네트워크 리소스 할당의 최적화 문제이며, 일련의 슬라이싱 비율 중에서 최적의 슬라이싱 비율을 선택해야 한다. 즉, 결과적인 리소스 활용 성능에 기초하여 최적의 슬라이싱 비율이 선택된다. 이 문제를 보다 실용적인 방법으로 해결하기 위해 강화 학습과 저복잡도 휴리스틱을 사용하기로 결정하고, 일반적인 접근법은 도 4에 도시되어 있다.

[0071] 본 발명의 실시예에 따른 네트워크 모델(410)로서 각 슬라이스에 대한 슬라이싱 속도, $\alpha_{s,DL}$ 을 결정하기 위해 슬라이스 컨트롤러가 고려된다. 슬라이스 컨트롤러의 작동에는 두 가지 주요 부분이 있다. 첫 번째 부분에서, 일부 중간 슬라이싱 비율은 $\beta_{s,DL}$ 로 표시되는 RL 알고리즘(420)에 의해 결정된다. 두 번째 부분은 RL 알고리즘의 결과를 입력으로 취한 다음 슬라이싱 비율을 조정하여 최종 슬라이싱 비율 $\alpha_{s,DL}$ 을 얻는 휴리스틱 알고리즘(421)으로 이어진다.

[0072] 먼저, 다운링크에 대해 RL 알고리즘을 수행하여 $\beta_{s,DL}$ 을 결정한다고 가정한다. 본 발명은 두 가지 서비스 유형을 기반으로 슬라이싱 문제를 논의하므로, V2X 및 eMBB를 기반으로 20개의 슬라이싱 동작을 설정했으며, 동작 집합은 다음과 같다:

$$\begin{cases} \beta_{1,x}(k) = 0.05k \\ \beta_{2,x}(k) = 1 - 0.05k \end{cases} \quad (6)$$

[0074] 여기서, $k=1,2,\dots,20$ 이고 $x=\{DL\}$ 이다.

[0075] $\beta_{1,x}(k)$ 는 V2X 서비스를 위해 예약된 PRB의 부분이다;

[0076] $\beta_{2,x}(k)$ 는 eMBB 서비스를 위해 예약된 PRB의 부분이다;

[0077] RL의 일반적인 운영에서 최적의 솔루션을 찾기 위해 다른 동작 a_k (다른 슬라이싱 비율)를 시도한다. 선택된 동작의 결과로, RL의 프로세스는 보상 $R_{TOT,DL}(a_k)$ 을 얻을 수 있으며, 보상은 원하는 최적의 목표 측면에서 동작의 결과가 얼마나 좋은지 또는 얼마나 나쁜지를 측정한다. 보상에 기초하여, RL 알고리즘의 의사 결정을 조정하여 최고 보상으로 이어지는 동작을 점진적으로 학습할 수 있다. 동작 선택은 이용 간 트레이드 오프(trade-off between exploitation)(높은 보상 동작 시도)과 탐색(그로부터 학습하기 위한 다른 동작 시도) 사이의 균형을 유지함으로써 이루어진다.

[0078] 본 발명에서, 고려된 RL 알고리즘은 장기적으로 가능한 모든 동작의 탐색-이용을 실현할 수 있는 Softmax 결정에 기반한 Q-러닝이다. 따라서 보상을 정의할 수 있다. 아래에서, 보상 함수와 Q-러닝 알고리즘의 구체적인 운영에 대하여 설명한다.

[0079] 적절한 보상 함수는 동작의 능력을 반영해야 한다. 이를 바탕으로, 해당 슬라이싱 비율이 $\beta_{s,DL}(k)$ 인 선택된 동작 a_k 에 대해, 해당 동작에 할당된 총 리소스에 사용된 리소스의 비율로 정의된 슬라이스에 대한 보상을 계산하는 데 정규화된 리소스 활용도 함수를 사용할 수 있다.

[0080] V2X 슬라이스($s=1$)의 경우:

$$\psi_{1,DL}(a_k) = \frac{\tau_{1,DL}}{\beta_{1,DL}(k) \cdot N_{DL}} \quad (7)$$

[0082] eMBB 슬라이스($s=2$)의 경우:

$$\psi_{2,DL}(a_k) = \frac{\tau_{2,DL}}{\beta_{2,DL}(k) \cdot N_{DL}} \quad (2)$$

[0084] 이러한 방정식에 기초하여, 동작 a_k 의 결과로서의 보상 $\beta_{s,DL}(k)$ 는 다음과 같이 정의된다:

$$R_{s,DL}(a_k) = \begin{cases} e^{\psi_{s,DL}(a_k)} & \psi_{s,DL}(a_k) \leq 1 \\ e^{-\psi_{s,DL}(a_k)} & \psi_{s,DL}(a_k) > 1 \end{cases} \quad (9)$$

식(9)에서, $\psi_{s,DL}(a_k)$ 의 값이 0 내지 1 사이인 한, 보상 함수는 $\psi_{s,DL}(a_k) = 1$ 에서 최대치로 기하급수적으로 증가한다. 따라서, 이 값의 증가, 즉 더 높은 활용도로 이어지는 동작들은 또한 더 큰 보상 피드백을 받는다.

반대로, 만약 $\psi_{s,x}(a_x) > 1$ 이라면, 이 경우는 슬라이스가 중단될 것임을 나타낸다. 그러므로 보상은 줄어들 것이다.

또한, 총 보상은 고려된 모든 슬라이스에 대한 동작 효과를 고려해야 하므로, 일반적으로 슬라이스당 보상의 기하학적 평균으로 정의된다:

$$R_{TOT,DL}(a_k) = (\prod_{s=1}^2 R_{s,DL}(a_k))^{1/s} \quad (10)$$

본 발명의 실시예에 따른 슬라이스 컨트롤러의 Q-러닝 체계의 궁극적인 목표는 각 슬라이스의 예상 장기 보상을 최대화하는 조치를 찾는 것이다. 이 목표를 달성하기 위해 Q-러닝은 정해진 시간에 이산 시간 동안 네트워크 모델과 상호 작용하고 선택한 동작에 대한 보상을 추정한다. 보상에 따르면 슬라이스 컨트롤러는 동작 a_k 에 대한 경험 기록을 유지하고 동작 값 함수(Q-값이라고도 함)는 $Q_{DL}(a_k)$ 에 저장된다. 각 시간 단계에 대해 $Q_{DL}(a_k)$ 의 값은 null 할인율을 가지는 단일 상태 Q-러닝 방법에 따라 업데이트된다:

$$Q_{DL}(a_k) \leftarrow (1 - \alpha)Q_{DL}(a_k) + \alpha \cdot R_{TOT,DL}(a_k) \quad (11)$$

여기서 $\alpha \in (0,1)$ 은 학습 속도이고, $R_{TOT,DL}(a_k)$ 는 V2X 및 eMBB 슬라이스에 대해 동작 a_k 를 수행하는 데 대한 총 보상이다. 동작 a_k 가 이전에 사용된 적이 없는 경우 초기화 시 $Q_{DL}(a_k)$ 는 임의의 값으로 초기화된다.

소프트맥스 정책에 따른 선정기준은 다음과 같다.

$Q_{DL}(a_k)$ 에 기초한 동작의 선택은 소프트맥스 전략에 기초하며, 여기서 다른 동작의 선택은 확률적이다. 구체적으로, 동작 a_k 의 선택에 해당하는 확률 $P(a_k)$ 는 다음과 같이 정의된다:

$$P(a_k) = \frac{e^{Q_{DL}(a_k)/\tau}}{\sum_{j=1}^{A_x} e^{Q_{DL}(a_j)/\tau}} \quad (12)$$

여기서 τ 은 온도 매개변수라고 하는 양의 정수이며, 선택 확률을 제어한다. τ 값이 높을수록 동작 확률은 동일하다. 반대로 τ 값이 낮으면 Q 값이 다른 동작의 선택 확률에 큰 차이가 발생한다. 소프트맥스 전략은 탐색과 이용 사이의 효과적인 절충을 가능하게 한다. 즉, 높은 보상을 산출하는 높은 확률의 동작을 선택하는 동시에 특정 확률의 새로운 동작을 계속 탐색하는 것은 미래에 더 나은 결정으로 이어질 수 있다.

본 발명의 실시예에 따른 다운로드 시뮬레이션에서 코드를 실행할 때 Matlab 명령 윈도우에서 다음을 수행한다:

```
episode=0998: a_k=007, Gamma1=40.07, Gamma2=91.52, R_DL(V2X)=1.77,
R_DL(eMBB)=2.02, R_TOT_DL=1.89, beta_V2X=0.35, beta_eMBB=0.65
episode=0999: a_k=003, Gamma1=48.24, Gamma2=92.23, R_DL(V2X)=0.20,
R_DL(eMBB)=1.72, R_TOT_DL=0.59, beta_V2X=0.15, beta_eMBB=0.85
episode=1000: a_k=011, Gamma1=51.66, Gamma2=81.53, R_DL(V2X)=1.60,
R_DL(eMBB)=2.47, R_TOT_DL=1.99, beta_V2X=0.55, beta_eMBB=0.45
```

여기서 a_k 는 슬라이싱 비율에 대한 동작이다. "Gamma"는 각 슬라이스에 필요한 평균 RB 수이다. "R_DL"은 각 슬라이스에 대한 보상 계산이다. "R_TOT_DL"은 다운로드에 대한 총 보상 계산이다. "beta"는 각 슬라이스에 해

당하는 슬라이스 비율이다. 구현에는 두 가지 슬라이싱 서비스가 포함된다. 즉, 두 슬라이싱 비율의 합은 1이다. 위에서 알 수 있듯이, 네트워크 리소스는 각 슬라이스에 성공적으로 할당된다.

[0102] 도 5는 본 발명의 일 실시예에 따른 서로 다른 에피소드들에 대한 Q-값을 나타내는 그래프이다.

[0103] 본 발명의 실시예에 따른 시뮬레이션에서 네트워크 부하는 경량이다. 네트워크에서, matlab 콘솔의 로그에 나타난 바와 같이, 차량 UE로부터의 PRB 요구는 일반적으로 eMBB UE로부터의 요구보다 낮으므로, 더 나은 슬라이싱 전략은 더 많은 PRB를 eMBB 서비스에 할당해야 한다.

[0104] 1000 에피소드 이후의 최종 Q-값 표는 동작 #5/#6/#7이 다른 동작보다 더 바람직하다는 것을 나타낸다. 동작 #5/#6/#7은 V2X 서비스의 PRB 비율이 0.3, 0.35, 0.4이고 eMBB의 PRB 비율이 0.7, 0.65, 0.6이므로 더 많은 PRB가 eMBB 서비스에 할당되었다.

[0105] 본 발명의 실시예에 따르면, 리소스 수요에 기초하여 RL에 의해 선택된 중간 슬라이싱 비율인 $\beta_{s,DL}$ 을 미세 조정하기 위해 휴리스틱 체계를 적용한다.

[0106] 이 아이디어는 실제 RB 요구사항과 슬라이싱 비율 $\beta_{s,DL}$ 을 기반으로 한다. 이 알고리즘은 한 슬라이스가 다운 링크에서 실제로 필요한 것보다 더 많은 리소스를 가지고 있는지를 추정한다. 즉 $\psi_{s,DL}(a_{DL,sel}) < 1$ 이다. 반면, 다른 슬라이스는 필요한 리소스보다 적은 리소스를 가지고 있다. 즉 $\psi_{s,DL}(a_{DL,sel}) > 1$ 이다. 이 경우 슬라이스는 다른 슬라이스로 전송할 수 있는 추가 용량 $\Delta C_{s,DL}$ 을 남긴다. 세부적으로 추가 용량은 다음과 같이 표시된다:

$$[0107] \Delta C_{s,DL} = (1 - \psi_{s,DL}(a_{DL,sel})) \cdot \omega \quad (13)$$

[0108] 여기서 ω 는 RB 소비 변동에 대한 여유를 남겨두기 위한 [0,1] 범위의 구성 매개변수이다. 이에 기초하여 슬라이스의 슬라이싱 비율은 $\Delta C_{s,DL}$ 만큼 줄어들 것이다. 즉 $a_{s,DL} = \beta_{s,DL} - \Delta C_{s,DL}$ 이다. 반대로 다른 슬라이스의 슬라이싱 비율은 $\Delta C_{s,DL}$ 만큼 증가할 것이다. 즉 $a_{s',DL} = \beta_{s',DL} + \Delta C_{s,DL}$ 이다.

[0110] 도 6은 본 발명의 일 실시예에 따른 시뮬레이션 라운드에 따른 PRB 활용을 나타내는 그래프이다.

[0111] 파란색 선은 ignore_online_nw_characteristics=False(즉, 온라인 네트워크 모델의 네트워크 특성을 매개변수 추정에 사용)일 때의 선이고 빨간색 선은 ignore_online_nw_characteristics=True(즉, 네트워크 특성 추정 없음)일 때의 선이다.

[0112] 도 6에 도시된 그래프는 다음을 나타낸다:

[0113] (1) RL 알고리즘을 적용하기 전에 매개변수 추정을 위해 온라인 네트워크 모델의 네트워크 특성을 사용하는 경우 네트워크 특성을 무시할 때보다 PRB 활용도가 더 우수하다. 이와 같이 강화 학습을 위해 적합 네트워크 특성을 사용할 때의 이점을 볼 수 있다.

[0114] (2) 더 나은 매개변수 추정을 위해 더 많은 네트워크 특성을 사용할 수 있고 더 나은 RL 학습 알고리즘을 위해 더 많은 샘플을 사용할 수 있기 때문에 시뮬레이션을 더 많이 수행할 수록 PRB 활용도가 높아진다.

[0115] (3) 그래프의 커브에 약간의 변동이 있으며 이는 시뮬레이션 시간이 충분하지 않음을 나타낸다. 이러한 시뮬레이션에서 sim_nr_drops_per_round=20은 너무 짧고, 그래프의 커브를 부드럽게 만들기 위해 sim_nr_drops_per_round는, 예를 들어 100보다 커야 하지만 이러한 경우 시뮬레이션 시간이 더 오래 걸리는 단점이 있다.

[0117] 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.

[0118] 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, 가상장치(virtual equipment), 컴퓨터 저장 매체 또는 장치에 구체화(embody)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.

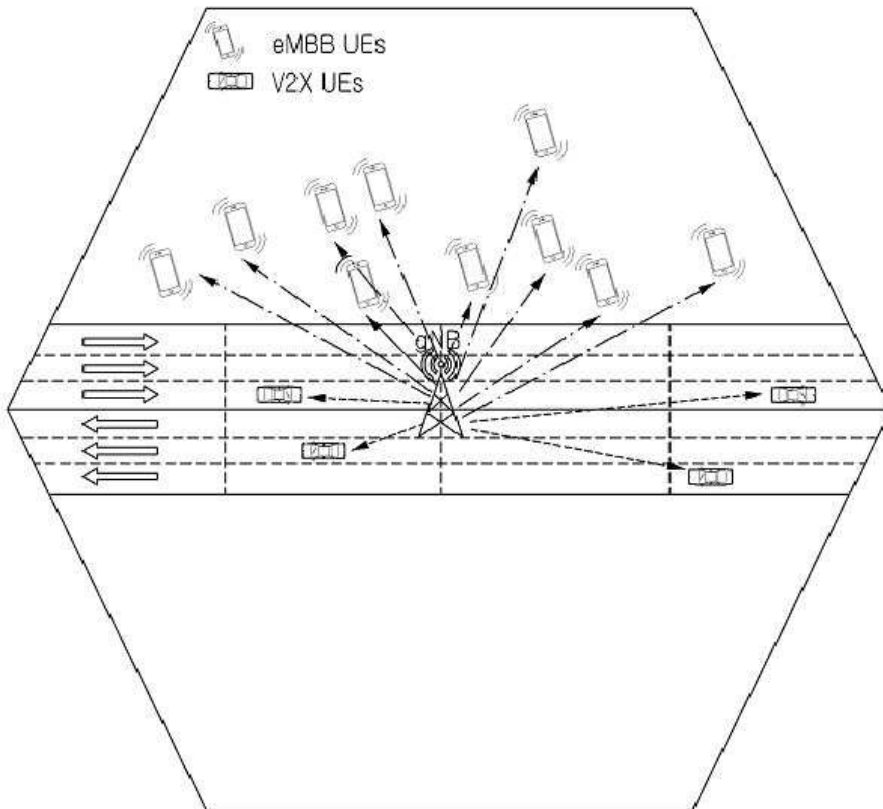
[0119] 실시예에 따른 방법은 다양한 컴퓨터 수단을 통하여 수행될 수 있는 프로그램 명령 형태로 구현되어 컴퓨터 판독 가능 매체에 기록될 수 있다. 상기 컴퓨터 판독 가능 매체는 프로그램 명령, 데이터 파일, 데이터 구조 등을 단독으로 또는 조합하여 포함할 수 있다. 상기 매체에 기록되는 프로그램 명령은 실시예를 위하여 특별히 설계되고 구성된 것들이거나 컴퓨터 소프트웨어 당업자에게 공지되어 사용 가능한 것일 수도 있다. 컴퓨터 판독 가능 기록 매체의 예에는 하드 디스크, 플로피 디스크 및 자기 테이프와 같은 자기 매체(magnetic 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.

[0120] 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.

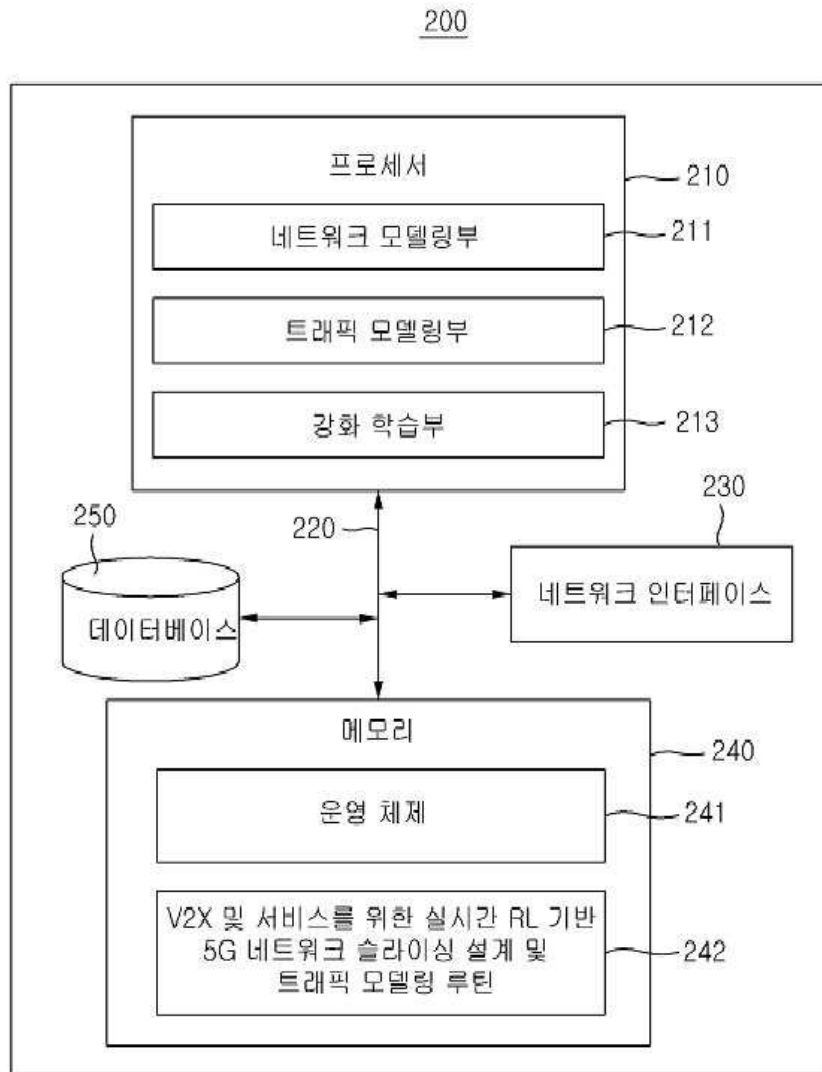
[0121] 그러므로, 다른 구현들, 다른 실시예들 및 특허청구범위와 균등한 것들도 후술하는 특허청구범위의 범위에 속한다.

도면

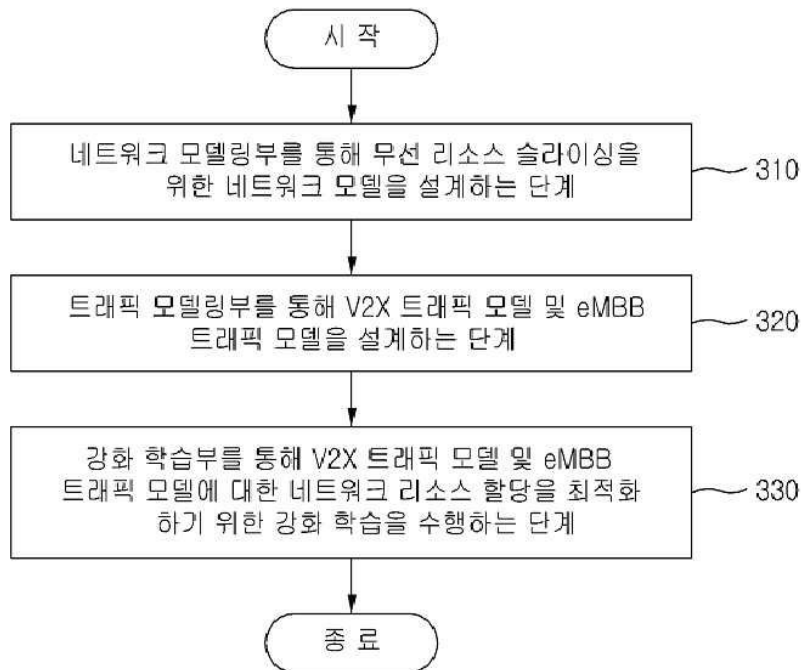
도면1



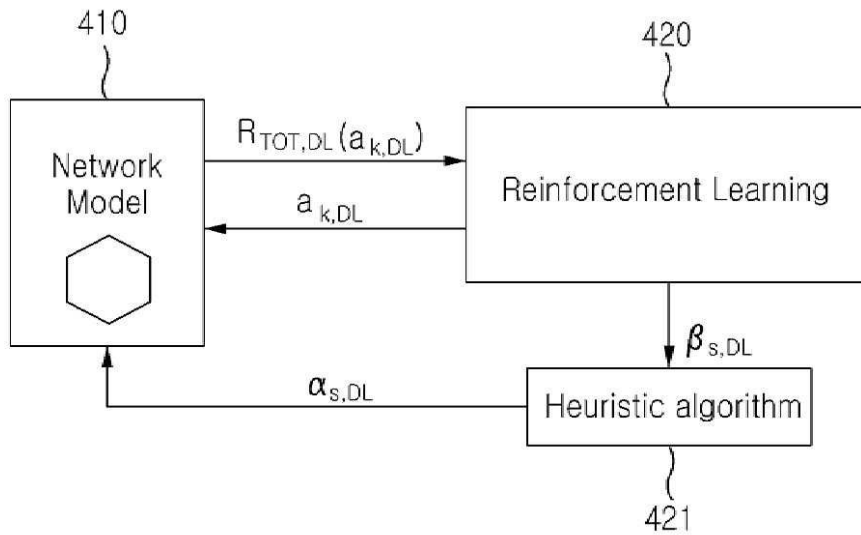
도면2



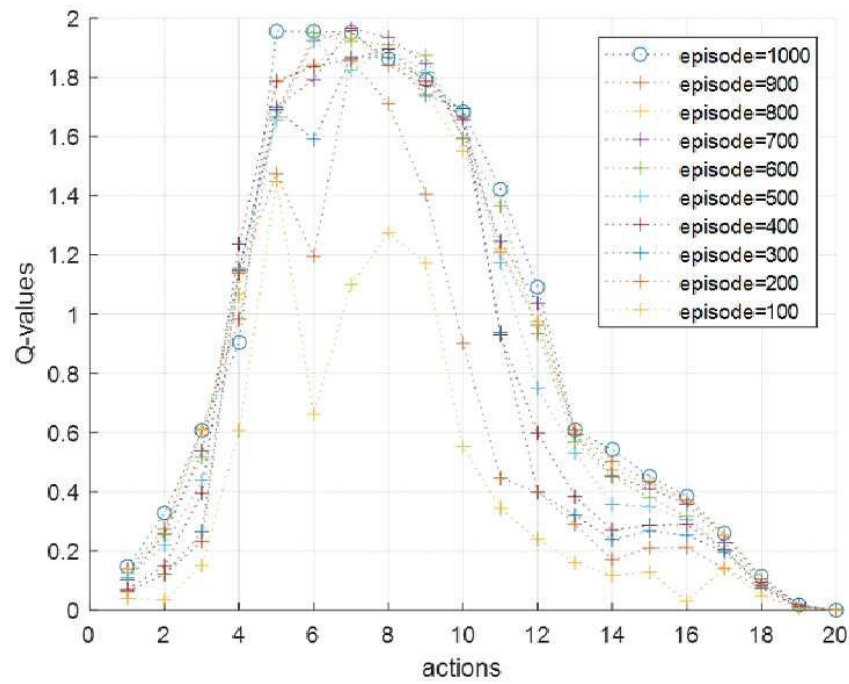
도면3



도면4



도면5



도면6

