



(19) 대한민국특허청(KR)  
(12) 등록특허공보(B1)

(45) 공고일자 2022년08월18일  
(11) 등록번호 10-2433577  
(24) 등록일자 2022년08월12일

(51) 국제특허분류(Int. Cl.)  
H04W 72/04 (2009.01) G06N 20/00 (2019.01)  
H04W 4/40 (2018.01) H04W 72/08 (2009.01)  
(52) CPC특허분류  
H04W 72/0493 (2013.01)  
G06N 20/00 (2021.08)  
(21) 출원번호 10-2020-0178041  
(22) 출원일자 2020년12월18일  
심사청구일자 2020년12월18일  
(65) 공개번호 10-2022-0087717  
(43) 공개일자 2022년06월27일  
(56) 선행기술조사문헌  
US20190075603 A1\*  
(뒷면에 계속)

(73) 특허권자  
연세대학교 산학협력단  
서울특별시 서대문구 연세로 50 (신촌동, 연세대학교)  
(72) 발명자  
김성륜  
서울특별시 용산구 이촌로 303, 32동 1304호 (이촌동, 현대아파트)  
오승은  
서울특별시 서대문구 성산로22길 24-18, 405호 (창천동)  
이지훈  
서울특별시 서초구 잠원로 213-10, 2동 303호 (잠원동, 한강아파트)  
(74) 대리인  
특허법인(유한)아이시스

전체 청구항 수 : 총 16 항

심사관 : 정남호

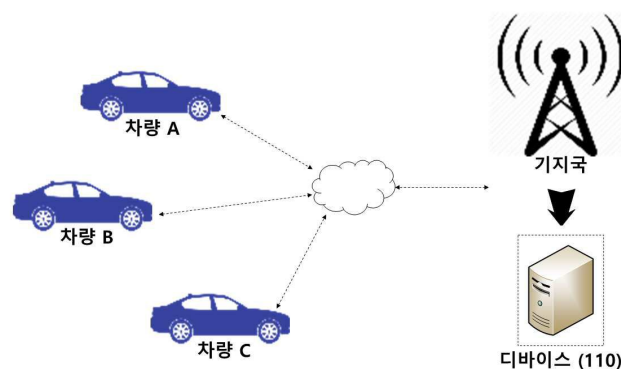
(54) 발명의 명칭 차량 네트워크의 자원 재할당 제어 방법 및 장치

(57) 요약

개시된 기술은 차량 네트워크의 자원 재할당 제어 방법 및 장치에 관한 것으로, 디바이스가 차량 네트워크를 통해 복수의 차량들에 대한 위치정보, 할당된 리소스 정보 및 리워드를 수신하는 단계; 상기 디바이스가 상기 위치정보 및 리워드를 토대로 상기 차량의 자원 재할당에 대한 제 1 확률(Probability)을 계산하고, 강화학습 모델에 상기 리소스 정보 및 리워드를 입력하여 상기 자원 재할당에 대한 제 2 확률을 계산하는 단계; 및 상기 디바이스가 상기 제 1 확률 및 상기 제 2 확률에 대한 계산 결과를 토대로 상기 차량 네트워크의 정책(Policy)을 결정하는 단계;를 포함한다.

대표도 - 도1

100



(52) CPC특허분류

*H04W 4/40* (2020.05)  
*H04W 72/0406* (2022.01)  
*H04W 72/048* (2013.01)  
*H04W 72/082* (2013.01)

(56) 선행기술조사문헌

H. Ye and G. Y. Li, "Deep reinforcement learning based distributed resource allocation for V2V broadcasting", 2018 IWCMC, IEEE, pp.440-445(2018.06.)\*  
 Y. H. Xu et al., "Deep deterministic policy gradient (DDPG)-based resource allocation scheme for NOMA vehicular communications", IEEE Access, vol.8, pp.18797-18807(2020.01.22.)\*  
 KR1020190093204 A  
 Z. Zhang, Z. Pan and M. J. Kochenderfer, "Weighted Double Q-learning", IJCAI, pp. 3455-3461(2017.08.) 1부.\*  
 \*는 심사관에 의하여 인용된 문헌

이 발명을 지원한 국가연구개발사업

과제고유번호	1711103168
과제번호	2016-0-00208-005
부처명	과학기술정보통신부
과제관리(전문)기관명	정보통신기획평가원(한국연구재단부설)
연구사업명	정보통신방송연구개발사업
연구과제명	(창조씨앗-2단계) 차세대 5G V2X 서비스 실현을 위한 정밀 측위탐색 연계 고효율 다중안테나 정보전송 및 네트워크 기술 연구
기 여 율	1/1
과제수행기관명	연세대학교 산학협력단
연구기간	2020.01.01 ~ 2020.12.31

---

## 명세서

### 청구범위

#### 청구항 1

디바이스가 차량 네트워크를 통해 제 1 차량을 포함하는 복수의 차량들에 대한 위치정보, 상기 복수의 차량들에 할당된 리소스 정보 및 리워드를 수신하는 단계;

상기 디바이스가 제 1 강화학습 모델에 제 1 상태 정보로서 상기 위치정보 및 제 1 리워드 정보로서 상기 리워드를 입력하여 상기 복수의 차량들에 대한 자원 재할당 여부를 나타내는 제 1 확률(Probability)을 계산하고, 제 2 강화학습 모델에 제 2 상태 정보로서 상기 리소스 정보 및 제 2 리워드 정보로서 상기 리워드를 입력하여 상기 복수의 차량들에 대한 자원 재할당 여부를 나타내는 제 2 확률을 계산하는 단계; 및

상기 디바이스가 상기 제 1 확률 및 상기 제 2 확률에 대한 계산 결과를 토대로 상기 차량 네트워크의 정책(Policy)을 결정하는 단계;를 포함하되,

상기 제 2 강화학습 모델은 제 2 액터 네트워크 및 제 2 크리틱 네트워크를 포함하고,

상기 제 2 크리틱 네트워크는 상기 제 1 차량 및 상기 디바이스에 각각 저장되고,

상기 디바이스가 상기 복수의 차량들 중 상기 제 1 차량으로부터 소정 거리 내에 위치한 차량들에 의한 간섭 여부에 따라 상기 제 1 차량에 저장된 상기 제 2 크리틱 네트워크를 이용할지 여부를 결정하는 단계를 포함하는 차량 네트워크의 자원 재할당 제어 방법.

#### 청구항 2

삭제

#### 청구항 3

제 1 항에 있어서,

상기 디바이스가 상기 제 1 차량에 저장된 상기 제 2 크리틱 네트워크를 이용하는 경우, 상기 디바이스는 상기 제 1 차량의 제 2 크리틱 네트워크로부터 출력된 값(Value)을 수신하여 동작하는 차량 네트워크의 자원 재할당 제어 방법.

#### 청구항 4

삭제

#### 청구항 5

제 1 항에 있어서,

상기 복수의 차량들 및 상기 디바이스는 C-V2X(Cellular Vehicle-to-Everything) 기반 통신을 이용하고,

상기 복수의 차량들은 상기 C-V2X의 업링크를 통해 상기 위치정보, 상기 리소스 정보 및 상기 리워드를 상기 디바이스에 네트워크 파라미터로 전송하고,

상기 디바이스는 상기 C-V2X의 다운링크를 통해 상기 복수의 차량들에게 상기 정책을 전송하는 차량 네트워크의 자원 재할당 제어 방법.

#### 청구항 6

제 1 항에 있어서,

상기 복수의 차량들은 특정 구간을 통과할 때 차체에 각각 탑재된 센서의 감지 반경 이내에서 수집되는 센싱 결과를 상기 위치정보로 상기 디바이스에 전송하는 차량 네트워크의 자원 재할당 제어 방법.

## 청구항 7

제 1 항에 있어서,

상기 디바이스는 지상의 기지국 또는 도로 상의 인프라에 설치되는 RSU(Road Side Unit)인 것을 특징으로 하는 차량 네트워크의 자원 재할당 제어 방법.

## 청구항 8

제 1 항에 있어서, 상기 결정된 정책은,

상기 복수의 차량들에 대한 현재 리소스 블록의 재할당 확률이 높으면 상기 현재 리소스 블록을 유지하고, 상기 복수의 차량들에 대한 현재 리소스 블록의 재할당 확률이 낮으면 복수의 리소스 블록 후보군들 중 랜덤하게 결정된 하나를 상기 복수의 차량들에 대한 새로운 리소스 블록으로 결정하는 차량 네트워크의 자원 재할당 제어 방법.

## 청구항 9

제 8 항에 있어서,

상기 디바이스는 센싱 윈도우 내 1000ms 이전까지 측정된 RSSI 값을 리소스 블록 별로 평균을 계산하고 이 중 하위 20%의 값을 상기 복수의 리소스 블록 후보군들로 결정하는 차량 네트워크의 자원 재할당 제어 방법.

## 청구항 10

제 1 항에 있어서,

상기 디바이스는 상기 제 1 확률에 대한 계산 결과 및 상기 제 2 확률에 대한 계산 결과의 가중 평균을 계산하여 상기 정책을 결정하는 차량 네트워크의 자원 재할당 제어 방법.

## 청구항 11

제 1 차량을 포함하는 복수의 차량들에 대한 위치정보, 상기 복수의 차량들에 할당된 리소스 정보 및 리워드를 수신하는 안테나;

상기 복수의 차량들에 대한 정책을 결정하기 위한 복수의 강화학습 모델들을 저장하는 메모리; 및

상기 복수의 강화학습 모델들 중 제 1 강화학습 모델에 제 1 상태 정보로서 상기 위치정보 및 제 1 리워드 정보로서 상기 리워드를 입력하여 상기 복수의 차량들에 대한 자원 재할당 여부를 나타내는 제 1 확률을 계산하고, 상기 복수의 강화학습 모델들 중 제 2 강화학습 모델에 제 2 상태 정보로서 상기 리소스 정보 및 제 2 리워드 정보로서 상기 리워드를 입력하여 상기 복수의 차량들에 대한 자원 재할당 여부를 나타내는 제 2 확률을 계산하고, 상기 제 1 확률 및 상기 제 2 확률에 대한 계산 결과를 토대로 상기 정책을 결정하는 프로세서;를 포함하되,

상기 제 2 강화학습 모델은 제 2 액터 네트워크 및 제 2 크리틱 네트워크를 포함하고,

상기 제 2 크리틱 네트워크는 상기 제 1 차량 및 상기 메모리에 각각 저장되고,

상기 프로세서는 상기 복수의 차량들 중 상기 제 1 차량으로부터 소정 거리 내에 위치한 차량들에 의한 간섭 여부에 따라 상기 제 1 차량에 저장된 상기 제 2 크리틱 네트워크를 이용할지 여부를 결정하는 차량 네트워크의 자원 재할당 제어 장치.

## 청구항 12

삭제

## 청구항 13

제 11 항에 있어서,

상기 프로세서가 상기 제 1 차량에 저장된 상기 제 2 크리틱 네트워크를 이용하는 경우, 상기 프로세서는 상기 제 1 차량의 제 2 크리틱 네트워크로부터 출력된 값(Value)을 수신하여 동작하는 차량 네트워크의 자원 재할당

제어 장치.

#### 청구항 14

삭제

#### 청구항 15

제 11 항에 있어서,

상기 프로세서는 상기 결정된 정책에 따라 상기 복수의 차량들에 대한 현재 리소스 블록의 재할당 확률이 높으면 상기 현재 리소스 블록을 유지하고 상기 복수의 차량들에 대한 현재 리소스 블록의 재할당 확률이 낮으면 복수의 리소스 블록 후보군들 중 랜덤하게 결정된 하나를 상기 복수의 차량들에 대한 새로운 리소스 블록으로 결정하는 차량 네트워크의 자원 재할당 제어 장치.

#### 청구항 16

제 15 항에 있어서,

상기 프로세서는 센싱 윈도우 내 1000ms 이전까지 측정된 RSSI 값을 리소스 블록 별로 평균을 계산하고 이 중 하위 20%의 값을 상기 복수의 리소스 블록 후보군들로 결정하는 차량 네트워크의 자원 재할당 제어 장치.

#### 청구항 17

제 1 항에 있어서,

상기 제 1 강화학습 모델은 DQN(deep Q network)를 포함하는 차량 네트워크의 자원 재할당 제어 방법.

#### 청구항 18

제 1 항에 있어서,

상기 제 1 강화학습 모델은 제 1 액터 네트워크 및 제 1 크리틱 네트워크를 포함하는 차량 네트워크의 자원 재할당 제어 방법.

#### 청구항 19

제 11 항에 있어서,

상기 제 1 강화학습 모델은 DQN(deep Q network)를 포함하는 차량 네트워크의 자원 재할당 제어 장치.

#### 청구항 20

제 11 항에 있어서,

상기 제 1 강화학습 모델은 제 1 액터 네트워크 및 제 1 크리틱 네트워크를 포함하는 차량 네트워크의 자원 재할당 제어 장치.

### 발명의 설명

### 기술 분야

[0001] 개시된 기술은 차량 네트워크의 자원 재할당을 제어하는 방법 및 장치에 관한 것이다.

### 배경 기술

[0002] 종래 서브 6GHz 대역을 활용하는 차량 네트워크의 대표적인 표준으로는 C-V2X, 그리고 IEEE 802.11p가 있다. C-V2X 표준은 자율적인 스케줄링 기반의 C-V2X 모드 4와 네트워크에서 스케줄링을 제어하는 C-V2X 모드 3를 포함하고 있다.

[0003] C-V2X 모드 4는 각 차량들이 일정 수의 패킷을 전송하게 되면 센싱 정보를 바탕으로 하여 기존에 사용하던 리소스 블록(Resource block)을 그대로 사용할 것인지 아니면 다른 리소스 블록으로 갱신하여 통신할 것인지를 결정

한다. C-V2X 모드 3는 기지국이 센싱한 정보를 바탕으로 리소스 블록을 각 차량들에게 직접 할당하는 기법이다.

[0004] 한편, 차량의 자율주행 레벨이 증가하고 차량 간 어플리케이션이 다양화됨에 따라 더 높은 데이터 전송 속도를 충족할 수 있는 MAC 연구의 필요성이 증가하고 있다. 이에 따라 Abobe-6GHz, 즉 밀리미터파를 이용한 방향성 전송에 대한 연구가 집중적으로 이루어지고 있다.

[0005] 한편, C-V2X 모드 4의 경우 기존에 있는 인프라를 활용하지 않고 인프라를 활용하는 기법 대비 자원 선택의 효율이 떨어지는 문제점이 있으나 인프라를 거치면서 추가로 발생하는 지연시간(Latency)이 없다는 장점이 있다. 반면 C-V2X 모드 3의 경우 추가적인 지연시간이 발생하는 단점이 있지만 효율적인 자원 선택을 통한 데이터 전송 이득이 발생한다.

[0006] 최근 서브 6GHz 대역에서 Abobe-6GHz로 대역이 옮겨감에 따라 C-V2X 모드 3의 경우 기지국이 센싱한 정보를 바탕으로 자원을 할당해주기 때문에 모든 차량들이 방향성 안테나를 활용하는 경우 데이터 전송률이 떨어지는 문제가 발생할 수 있다. 그리고 C-V2X 모드 4의 경우에는 밀리미터파를 이용하더라도 종전의 서브 6GHz 대역과 큰 차이가 나타나지 않는다.

## 선행기술문헌

### 특허문헌

[0007] (특허문헌 0001) 한국 공개특허 제10-2020-0096096호

## 발명의 내용

### 해결하려는 과제

[0008] 개시된 기술은 차량 네트워크의 자원 재할당을 제어하는 방법 및 장치를 제공하는데 있다.

### 과제의 해결 수단

[0009] 상기의 기술적 과제를 이루기 위하여 개시된 기술의 제 1 측면은 디바이스가 차량 네트워크를 통해 복수의 차량들에 대한 위치정보, 할당된 리소스 정보 및 리워드를 수신하는 단계, 상기 디바이스가 상기 위치정보 및 리워드를 토대로 상기 차량의 자원 재할당에 대한 제 1 확률(Probability)을 계산하고, 강화학습 모델에 상기 리소스 정보 및 리워드를 입력하여 상기 자원 재할당에 대한 제 2 확률을 계산하는 단계 및 상기 디바이스가 상기 제 1 확률 및 상기 제 2 확률에 대한 계산 결과를 토대로 상기 차량 네트워크의 정책(Policy)을 결정하는 단계를 포함하는 차량 네트워크의 자원 재할당 제어 방법을 제공하는데 있다.

[0010] 상기의 기술적 과제를 이루기 위하여 개시된 기술의 제 2 측면은 복수의 차량들에 대한 위치정보, 할당된 리소스 정보 및 리워드를 수신하는 안테나, 상기 복수의 차량들에 대한 정책을 결정하기 위해서 복수의 강화학습 모델을 저장하는 메모리 및 상기 복수의 강화학습 모델 중 제 1 강화학습 모델에 상기 위치정보 및 리워드를 입력하여 상기 복수의 차량의 자원 재할당에 대한 제 1 확률을 계산하고, 상기 복수의 강화학습 모델 중 제 2 강화학습 모델에 상기 리소스 정보 및 리워드를 입력하여 상기 자원 재할당에 대한 제 2 확률을 계산하고, 상기 제 1 확률 및 상기 제 2 확률에 대한 계산 결과를 토대로 차량 네트워크의 정책을 결정하는 프로세서를 포함하는 차량 네트워크의 자원 재할당 제어 장치를 제공하는데 있다.

### 발명의 효과

[0011] 개시된 기술의 실시 예들은 다음의 장점들을 포함하는 효과를 가질 수 있다. 다만, 개시된 기술의 실시 예들이 이를 전부 포함하여야 한다는 의미는 아니므로, 개시된 기술의 권리범위는 이에 의하여 제한되는 것으로 이해되어서는 아니 될 것이다.

[0012] 개시된 기술의 일 실시예에 따르면 차량 네트워크의 자원 재할당 제어 방법 및 장치는 차량에 할당된 리소스 블록의 재할당 가능성 및 확률을 복수의 강화학습 모델을 통해 계산하여 차량 네트워크의 정책을 높은 신뢰도로 결정하는 효과가 있다.

[0013] 또한, 차량 네트워크의 통신 환경을 고려하여 높은 데이터 전송률을 유지하는 효과가 있다.

[0014] 또한, 분산 저장된 크리티컬 모델을 이용하여 데이터 전송량을 줄이는 효과가 있다.

### 도면의 간단한 설명

[0015] 도 1은 개시된 기술의 일 실시예에 따른 차량 네트워크의 자원 재할당 제어 과정을 나타낸 도면이다.

도 2는 개시된 기술의 일 실시예에 따른 차량 네트워크의 자원 재할당 제어 방법에 대한 순서도이다.

도 3은 개시된 기술의 일 실시예에 따른 차량 네트워크의 자원 재할당 제어 장치에 대한 블록도이다.

도 4는 개시된 기술의 일 실시예에 따라 디바이스의 강화학습 모델을 이용하는 것을 나타낸 도면이다.

도 5는 개시된 기술의 일 실시예에 따라 복수의 차량에 분산된 강화학습 모델을 이용하는 것을 나타낸 도면이다.

도 6은 리소스 블록 후보군을 결정하는 것을 나타낸 도면이다.

### 발명을 실시하기 위한 구체적인 내용

[0016] 본 발명은 다양한 변경을 가할 수 있고 여러 가지 실시예를 가질 수 있는 바, 특정 실시예들을 도면에 예시하고 상세한 설명에 상세하게 설명하고자 한다. 그러나, 이는 본 발명을 특정한 실시 형태에 대해 한정하려는 것이 아니며, 본 발명의 사상 및 기술 범위에 포함되는 모든 변경, 균등물 내지 대체물을 포함하는 것으로 이해되어야 한다.

[0017] 제 1, 제 2, A, B 등의 용어는 다양한 구성요소들을 설명하는데 사용될 수 있지만, 해당 구성요소들은 상기 용어들에 의해 한정되지는 않으며, 단지 하나의 구성요소를 다른 구성요소로부터 구별하는 목적으로만 사용된다. 예를 들어, 본 발명의 권리 범위를 벗어나지 않으면서 제 1 구성요소는 제 2 구성요소로 명명될 수 있고, 유사하게 제 2 구성요소도 제 1 구성요소로 명명될 수 있다. 및/또는 이라는 용어는 복수의 관련된 기재된 항목들의 조합 또는 복수의 관련된 기재된 항목들 중의 어느 항목을 포함한다.

[0018] 본 명세서에서 사용되는 용어에서 단수의 표현은 문맥상 명백하게 다르게 해석되지 않는 한 복수의 표현을 포함하는 것으로 이해되어야 한다. 그리고 "포함한다" 등의 용어는 실시된 특징, 개수, 단계, 동작, 구성요소, 부분품 또는 이들을 조합한 것이 존재함을 의미하는 것이지, 하나 또는 그 이상의 다른 특징들이나 개수, 단계, 동작, 구성요소, 부분품 또는 이들을 조합한 것들의 존재 또는 부가 가능성을 배제하지 않는 것으로 이해되어야 한다.

[0019] 도면에 대한 상세한 설명을 하기에 앞서, 본 명세서에서의 구성부들에 대한 구분은 각 구성부가 담당하는 주기능 별로 구분한 것에 불과함을 명확히 하고자 한다. 즉, 이하에서 설명할 2개 이상의 구성부가 하나의 구성부로 합쳐지거나 또는 하나의 구성부가 보다 세분화된 기능별로 2개 이상으로 분화되어 구비될 수도 있다.

[0020] 그리고 이하에서 설명할 구성부 각각은 자신이 담당하는 주기능 이외에도 다른 구성부가 담당하는 기능 중 일부 또는 전부의 기능을 추가적으로 수행할 수도 있으며, 구성부 각각이 담당하는 주기능 중 일부 기능이 다른 구성부에 의해 전담되어 수행될 수도 있음은 물론이다. 따라서, 본 명세서를 통해 설명되는 각 구성부들의 존재 여부는 기능적으로 해석되어야 할 것이다.

[0021] 도 1은 개시된 기술의 일 실시예에 따른 차량 네트워크의 자원 재할당 제어 과정을 나타낸 도면이다. 도 1을 참조하면 차량 네트워크 내 노드들로 복수의 차량들과 기지국이 포함된다. 예컨대, 차량 A, 차량 B 및 차량 C는 서로 네트워크를 통해 센싱데이터를 송수신할 수 있다. 그리고 차량 뿐만 아니라 기지국에도 센싱데이터를 전송할 수 있다. 기본적인 네트워크의 구조는 셀룰러-V2X(Cellular-V2X, C-V2X)의 모드 3의 형태일 수 있다. 다만 내부적으로 처리되는 데이터 프레임워크는 C-V2X 모드 4를 기반으로 할 수 있다. 즉, C-V2X 모드 3와 모드 4가 혼용된 형태일 수 있다. 기지국에는 각 차량들에 자원을 재할당하기 위한 디바이스가 탑재된다. 디바이스(110)에는 자원 재할당 확률을 계산하는 강화학습 모델이 저장되어 있다.

[0022] 한편, 자원 재할당을 제어하는 장치는 지상의 기지국 뿐만 아니라 도로 상의 인프라에 설치되는 노변 기지국(Road Side Unit, RSU)을 이용할 수도 있다. 디바이스는 차량 네트워크를 통해 복수의 차량들에 대한 위치정보, 할당된 리소스 정보 및 리워드를 수신한다. 차량의 위치정보는 차량 내 탑재된 안테나 또는 센서의 센싱값을 의미한다. 예컨대, 복수개의 차량이 특정 구간을 통과할 때 차체에 각각 탑재된 센서의 감지 반경 이내에서 수집되는 센싱 결과를 위치정보로 이용할 수 있다. 그리고 리소스 정보는 현재 차량에 할당된 리소스 블록의 정보 및 리소스 블록의 크기 정보를 의미한다. 그리고 리워드는 차량 네트워크의 환경(Environment)에 대한 리워드를 의미한다. 즉, 복수의 차량들에 대한 위치정보, 할당된 리소스 정보 및 리워드는 차량 네트워크의 정책을 결정



하기 위한 파라미터로 이용될 수 있다.

- [0023] 한편 디바이스는 수신된 네트워크 파라미터를 이용하여 차량에 대한 자원 재할당 확률을 각각 계산할 수 있다. 여기에서 차량에 대한 자원은 통신을 위한 리소스 블록을 의미한다. 디바이스는 복수의 차량들에게 기 할당된 리소스 블록을 그대로 유지할 것인지 아니면 새로운 리소스 블록을 할당할 것인지 계산할 수 있다. 이때, 단순히 자원 재할당을 위한 확률을 강화학습 모델을 이용하여 계산하는 것이 아니라 차량의 위치정보를 토대로 자원 재할당 가능성을 추정한 결과도 함께 이용할 수 있다. 일 실시예로, 차량의 위치정보 및 리워드를 토대로 차량의 자원 재할당에 대한 제 1 확률을 계산할 수 있다. 그리고, 강화학습 모델에 리소스 정보 및 리워드를 입력하여 자원 재할당에 대한 제 2 확률을 계산할 수 있다. 그리고 제 1 확률과 제 2 확률의 계산 결과를 토대로 차량 네트워크의 정책(Policy)을 결정할 수 있다.
- [0024] 한편, 디바이스는 자원 재할당 확률을 계산하기 위한 강화학습 모델을 포함한다. 강화학습 모델은 액터 크리틱 네트워크를 이용할 수 있다. 디바이스는 차량 네트워크의 통신 상태에 따라 자신이 보유한 크리틱 네트워크를 이용할 수도 있고 차량에 저장된 크리틱 네트워크를 이용할 수도 있다. 즉, 크리틱 네트워크는 차량과 디바이스에 분산 저장될 수 있다. 예컨대, 차량 네트워크의 통신 환경이 양호하면 디바이스가 저장하는 크리틱 네트워크를 이용할 수 있고 차량 네트워크의 통신 환경이 불량하면 차량에 탑재된 크리틱 네트워크의 결과값(Value)을 수신할 수 있다. 전자의 경우는 디바이스에 집중된(Centralized) 형태의 모델일 수 있으며 후자의 경우는 차량과 디바이스에 분산된(Decentralized) 형태의 모델일 수 있다. 여기에서 차량 네트워크의 통신 환경의 양호 또는 불량을 판단하는 기준은 특정 차량의 주변에 다수의 차량이 몰려 있어서 데이터 간섭이 발생할 가능성에 따라 판단할 수 있다. 예컨대, 통신 환경의 간섭이 임계값 이상이면 복수개의 차량에 저장된 크리틱 네트워크를 이용하고 통신 환경의 간섭이 임계값 미만이면 디바이스에 저장된 크리틱 네트워크를 이용할 수 있다. 따라서 통신 환경에 따라 데이터 전송률을 조절하여 정책 결정에 따른 정확도를 유지할 수 있다.
- [0025] 한편, 상술한 바와 같이 복수의 차량 및 디바이스는 C-V2X(Cellular Vehicle-to-Everything) 기반 통신을 이용한다. 복수의 차량은 C-V2X의 업링크를 통해 위치정보, 할당된 리소스 정보 및 리워드를 디바이스에 네트워크 파라미터로 전송할 수 있다. 그리고, 디바이스는 C-V2X의 다운링크를 통해 정책을 전송할 수 있다.
- [0026] 한편, 디바이스는 자원 재할당 확률을 각각 계산한 결과에 따라 현재 리소스 블록의 재할당 확률이 높은 것으로 판단하면 현재 할당된 상태를 유지하고 재할당 확률이 낮은 것으로 판단하면 복수의 리소스 블록 후보군들 중 랜덤하게 결정된 하나를 차량에 대한 새로운 리소스 블록으로 결정할 수 있다. 전자의 경우에는 전송 속도(Data rate)가 저하되지 않을 것으로 예측한 경우이며 이 경우 현재 상태를 그대로 유지할 수 있다. 물론 다음 네트워크 파라미터가 수신되면 상술한 계산 과정을 다시 수행하게 되며 이 때 자원 재할당 확률이 낮아지면 리소스 블록을 새롭게 갱신할 수 있다.
- [0027] 한편, 디바이스는 차량의 자원 재할당 확률이 낮은 것으로 계산하면 센싱 윈도우 내 1000ms 이전까지 측정된 RSSI 값을 토대로 새로운 리소스 블록을 결정할 수 있다. 예컨대, RSSI 값을 리소스 블록 별로 평균을 계산하고 이 중 하위 20%의 값을 복수의 리소스 블록 후보군으로 결정할 수 있다. 이들 중 랜덤하게 결정된 하나의 리소스 블록이 차량에게 다시 할당하게 되는 리소스 블록으로 결정된다.
- [0028] 한편, 상술한 바와 같이 디바이스는 차량의 위치 자원 재할당에 대한 제 1 확률과 제 2 확률을 계산한 결과를 토대로 정책을 결정한다. 일 실시예로, 제 1 확률을 계산한 결과 및 제 2 확률을 계산한 결과 각각에 가중치를 곱한 값의 평균을 차량 네트워크의 정책(Policy)으로 결정할 수 있다. 이때, 강화학습 모델이 사전에 충분히 학습되지 않은 시기에는 계산 결과에 곱해지는 가중치를 조절하는 것으로 결과값의 신뢰도를 높일 수 있다. 예컨대, 학습 초기에는 제 1 확률을 계산한 결과에 대한 가중치를 높은 값으로 설정하고 학습 후기에는 제 2 확률을 계산한 결과에 대한 가중치를 높은 값으로 설정할 수 있다. 따라서 차량 네트워크의 통신 상태를 고려하여 데이터 전송량을 조절하여 네트워크 신뢰도를 높일 수 있고 가중치를 조절하여 강화학습의 불안정한 결과값을 보완할 수 있다.
- [0029] 도 2는 개시된 기술의 일 실시예에 따른 차량 네트워크의 자원 재할당 제어 방법에 대한 순서도이다. 도 2를 참조하면 차량 네트워크의 자원 재할당 제어 방법(200)은 차량 네트워크 파라미터를 수신하는 단계(210), 차량의 자원 재할당에 대한 제 1 확률 및 제 2 확률을 각각 계산하는 단계(220) 및 차량 네트워크의 정책을 결정하는 단계(230)를 포함한다. 차량 네트워크의 자원 재할당 제어 방법(200)은 디바이스를 통해 순차적으로 수행될 수 있다.
- [0030] 210 단계에서 디바이스는 차량 네트워크를 통해 복수의 차량들에 대한 네트워크 파라미터를 수신한다. 네트워크



파라미터는 차량의 위치정보, 차량에 할당된 리소스 정보 및 리워드를 포함한다.

- [0031] 220 단계에서 디바이스는 네트워크 파라미터에 포함된 차량의 위치정보 및 리워드를 토대로 차량의 자원 재할당에 대한 제 1 확률을 계산한다. 그리고, 강화학습 모델에 리소스 정보 및 리워드를 입력하여 자원 재할당에 대한 제 2 확률을 계산한다. 제 1 확률을 계산하는 것은 차량의 센싱데이터를 토대로 차량이 향후 이동할 위치를 고려하여 자원 재할당 여부를 추정하는 것이며 제 2 확률을 계산하는 것은 현재 할당된 리소스 블록을 그대로 이용할 것인지 또는 다른 리소스 블록을 할당할 것인지 확률을 계산하는 것을 의미한다. 디바이스는 강화학습 모델을 이용하여 제 2 확률을 계산할 수 있다. 물론 제 1 확률의 계산에 있어서도 강화학습 모델을 이용할 수 있다.
- [0032] 230 단계에서 디바이스는 제 1 확률 및 제 2 확률 각각의 계산 결과를 토대로 차량 네트워크의 정책(Policy)을 결정한다. 정책은 제 1 확률과 제 2 확률의 계산 결과의 가중 평균을 구한 값으로 결정할 수 있다. 예컨대, 제 1 확률의 계산 결과에 특정 값의 가중치를 곱하고 제 2 확률의 계산 결과에 서로 다른 가중치를 곱하고 두 결과값의 평균을 취하는 것으로 정책을 결정할 수 있다. 가중치의 비중은 강화학습 모델의 학습 정도에 따라 서로 달라질 수 있다. 예컨대, 학습 초기에는 제 1 확률의 계산 결과에 대한 가중치를 높은 값으로 설정하고 학습 후기에는 제 2 확률의 계산 결과에 대한 가중치를 높은 값으로 설정할 수 있다.
- [0033] 도 3은 개시된 기술의 일 실시예에 따른 차량 네트워크의 자원 재할당 제어 장치에 대한 블록도이다. 도 3을 참조하면 차량 네트워크의 자원 재할당 제어 장치(300)는 안테나(310), 메모리(320), 및 프로세서(330)를 포함한다.
- [0034] 안테나(310)는 복수의 차량들에 대한 위치정보, 할당된 리소스 정보 및 리워드를 수신한다. 안테나는 차량에 탑재된 방향성 안테나와 통신할 수 있는 장치이다. 안테나(310)는 기지국에 탑재된 안테나를 이용하거나 주파수 센서를 이용할 수 있다.
- [0035] 메모리(320)는 복수의 차량들에 대한 정책을 결정하기 위해서 복수의 강화학습 모델을 저장한다. 메모리(320)는 소정의 저장공간을 갖는 하드디스크와 같은 장치로 구현된다. 메모리(320)는 강화학습 모델과 네트워크 파라미터를 저장할 수 있는 용량을 갖는다.
- [0036] 프로세서(330)는 복수의 강화학습 모델 중 제 1 강화학습 모델에 위치정보 및 리워드를 입력하여 복수의 차량의 자원 재할당에 대한 제 1 확률을 계산한다. 그리고, 복수의 강화학습 모델 중 제 2 강화학습 모델에 리소스 정보 및 리워드를 입력하여 자원 재할당에 대한 제 2 확률을 계산한다. 제 1 확률의 계산과 제 2 확률의 계산에는 각각 서로 다른 강화학습 모델을 이용할 수 있다. 예컨대, 제 1 확률을 계산하기 위해서 DQN(Deep Q-network)과 같은 모델을 이용할 수 있다. 물론 제 2 강화학습 모델과 마찬가지로 액터 크리틱 네트워크를 이용할 수도 있다. 프로세서(330)는 두 강화학습 모델을 통해 결과값이 출력되면 각각의 가중 평균을 계산하여 이를 네트워크의 정책으로 결정한다.
- [0037] 한편, 상술한 자원 재할당 제어 장치(300)는 컴퓨터와 같은 디바이스에서 실행될 수 있는 실행가능한 알고리즘을 포함하는 프로그램(또는 어플리케이션)으로 구현될 수 있다. 상기 프로그램은 일시적 또는 비일시적 판독 가능 매체(non-transitory computer readable medium)에 저장되어 제공될 수 있다.
- [0038] 비일시적 판독 가능 매체란 레지스터, 캐쉬, 메모리 등과 같이 짧은 순간 동안 데이터를 저장하는 매체가 아니라 반영구적으로 데이터를 저장하며, 기기에 의해 판독(reading)이 가능한 매체를 의미한다. 구체적으로는, 상술한 다양한 어플리케이션 또는 프로그램들은 CD, DVD, 하드 디스크, 블루레이 디스크, USB, 메모리카드, ROM(read-only memory), PROM(programmable read only memory), EPROM(Erasable PROM, EPROM) 또는 EEPROM(Electrically EPROM) 또는 플래시 메모리 등과 같은 비일시적 판독 가능 매체에 저장되어 제공될 수 있다.
- [0039] 일시적 판독 가능 매체는 스태틱 램(Static RAM, SRAM), 다이내믹 램(Dynamic RAM, DRAM), 싱크로너스 디램(Synchronous DRAM, SDRAM), 2배속 SDRAM(Double Data Rate SDRAM, DDR SDRAM), 증강형 SDRAM(Enhanced SDRAM, ESDRAM), 동기화 DRAM(SyncLink DRAM, SLDRAM) 및 직접 램버스 램(Direct Rambus RAM, DRRAM) 과 같은 다양한 RAM을 의미한다
- [0040] 도 4는 개시된 기술의 일 실시예에 따라 디바이스의 강화학습 모델을 이용하는 것을 나타낸 도면이다. 도 4를 참조하면 복수의 차량과 노변기지국으로 이루어진 차량 네트워크를 형성할 수 있으며 이러한 차량 네트워크의 데이터 프레임워크 구조는 C-V2X 모드 4 구조일 수 있다. 차량 네트워크의 통신 상태에 따라 업링크를 통해 전송되는 네트워크 파라미터가 달라질 수 있다. 도 4의 경우에는 차량 네트워크의 통신 상태가 양호한 상태이므로

업링크를 통해 A2C 알고리즘 동작에 필요한 네트워크 파라미터로 위치정보, 리소스 정보 및 리워드를 전송할 수 있다.

[0041] 도 5는 개시된 기술의 일 실시예에 따라 복수의 차량에 분산된 강화학습 모델을 이용하는 것을 나타낸 도면이다. 앞서 도 4의 경우에는 차량 네트워크의 통신 상태가 양호한 것을 가정하였으나 도 5의 경우에는 통신 상태가 고르지 못한 경우를 가정한다. 이 경우에는 크리티크 네트워크를 차량에 분산시키고 차량의 크리티크 네트워크를 통해 출력된 결과값을 네트워크 파라미터로 수신할 수 있다. 즉, 통신 상태에 따라 강화학습 모델의 구조가 디바이스에 집중된 형태이거나 디바이스와 차량에 분산된 형태일 수 있다.

[0042] 추가로 도 5에서는 각 차량이 크리티크 모델을 가지고 있기에 디바이스의 추정 기법을 사용하지 않는다는 전제 하에 위치정보이나 리워드를 주고 받지 않는 대신에 차량의 크리티크 네트워크에서 출력된 값(Value)과 리소스 정보(Rsc Info) 값을 업로드하여 통신량을 줄이는 효과를 볼 수 있다. 이와 같은 분산형 A2C 기법은 차량 별 채널 환경 등 통신 상황에 따라 업로드할 정보량을 다르게 가져갈 수 있다. 물론 일부 통신 환경이 좋은 차량들에 한해서는 추정 기법이 적용될 수 있다.

[0043] 한편, 아래 수학적 식 1과 같이 두 가지 기법을 모두 포함하는 수식을 정리할 수 있다.

[0044] [수학적 식 1]

$$p_i(0) = \left(1 - \frac{n_i - n}{N - n}\right) * 0.8, \quad \text{where}$$

$$n_i = \sum_{j \neq i} 1(d_{ij} < d_{th}), \quad N = \max(n_i), \quad n = \min(n_i)$$

[0045]

[0047] 가령 주변에 차량이 많이 위치해 간섭을 주거나 받을 확률이 높은 차량들의 리소스 블록을 더 빈번하게 바꿔줌으로써 러닝 초기에도 비교적 높은 성능을 노려볼 수 있다.

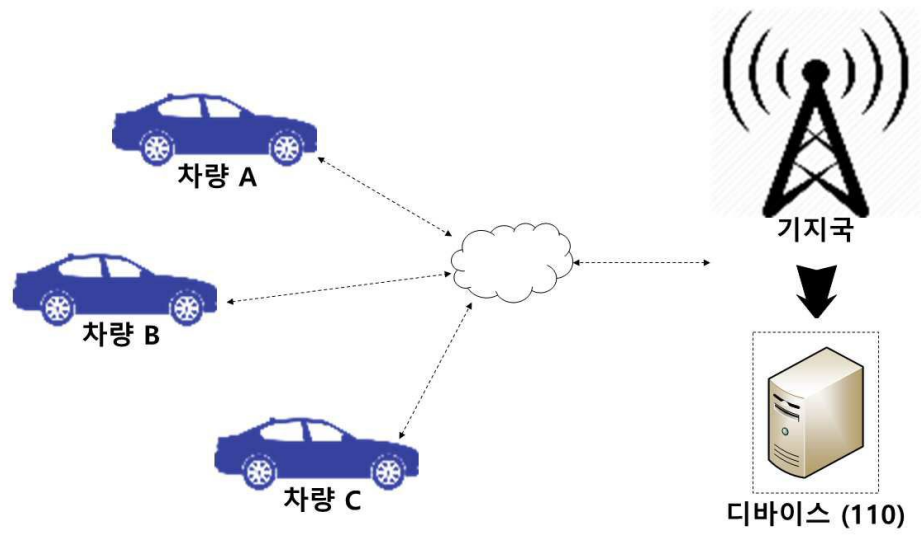
[0048] 도 6은 리소스 블록 후보군을 결정하는 것을 나타낸 도면이다. 도 6을 참조하면 차량 네트워크의 기본적인 데이터 프레임워크는 C-V2X 모드 4 프레임워크를 기반으로 한다. 각 차량들이 센싱 및 데이터 전송을 하던 중 일정 수의 패킷을 전송하게 되면 셀렉션 윈도우(Selection Window)에 도입하게 되고 자원 재할당 가능성(1-P, P=[0,0.8])에 따라 기존 통신에 활용하던 리소스 블록을 그대로 사용할지 혹은 새로운 리소스 블록 후보군들 중 랜덤하게 재할당할 것인지를 결정한다. 이 때, 센싱 윈도우(Sensing Window)로부터 1000ms 이전까지의 측정된 수신 세기(Received Signal Strength Indication, RSSI) 값을 리소스 블록 별로 평균을 내고 그 중 하위 20% 값을 새로운 리소스 블록 후보군으로 지정한다. 이들 중 랜덤하게 결정된 하나가 새로운 리소스 블록으로 결정된다.

[0049] 개시된 기술의 일 실시예에 따른 차량 네트워크의 자원 재할당 제어 방법 및 장치는 이해를 돕기 위하여 도면에 도시된 실시 예를 참고로 설명되었으나, 이는 예시적인 것에 불과하며, 당해 분야에서 통상적 지식을 가진 자라면 이로부터 다양한 변형 및 균등한 타 실시 예가 가능하다는 점을 이해할 것이다. 따라서, 개시된 기술의 진정한 기술적 보호범위는 첨부된 특허청구범위에 의해 정해져야 할 것이다.

도면

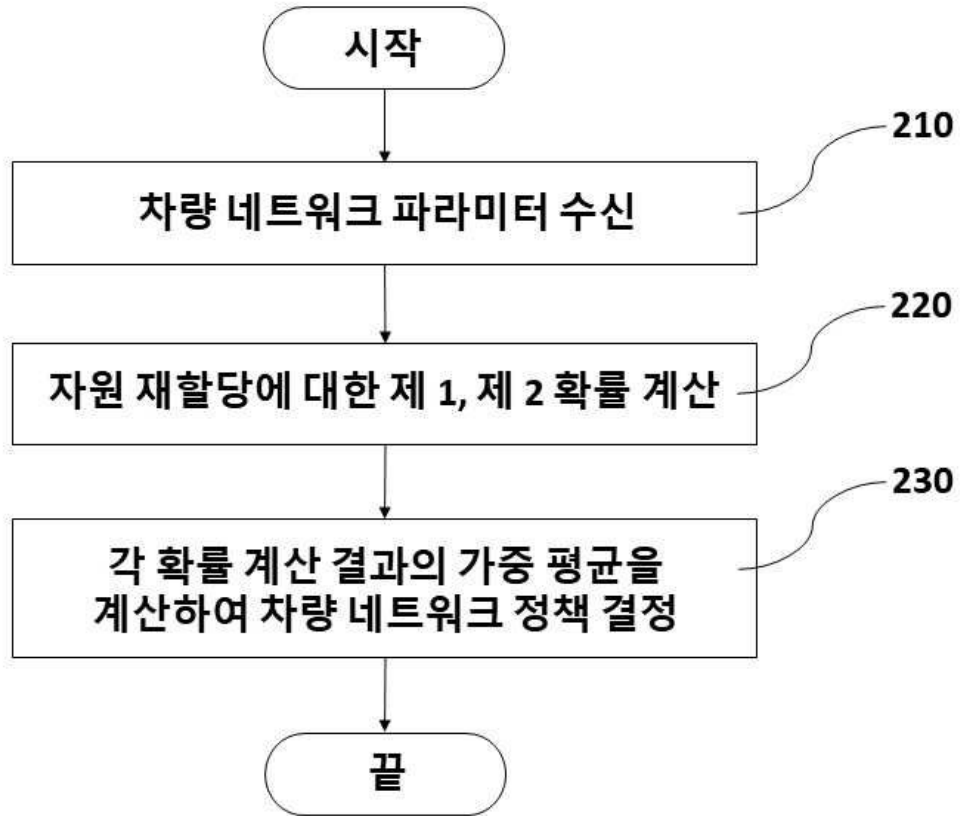
도면1

100

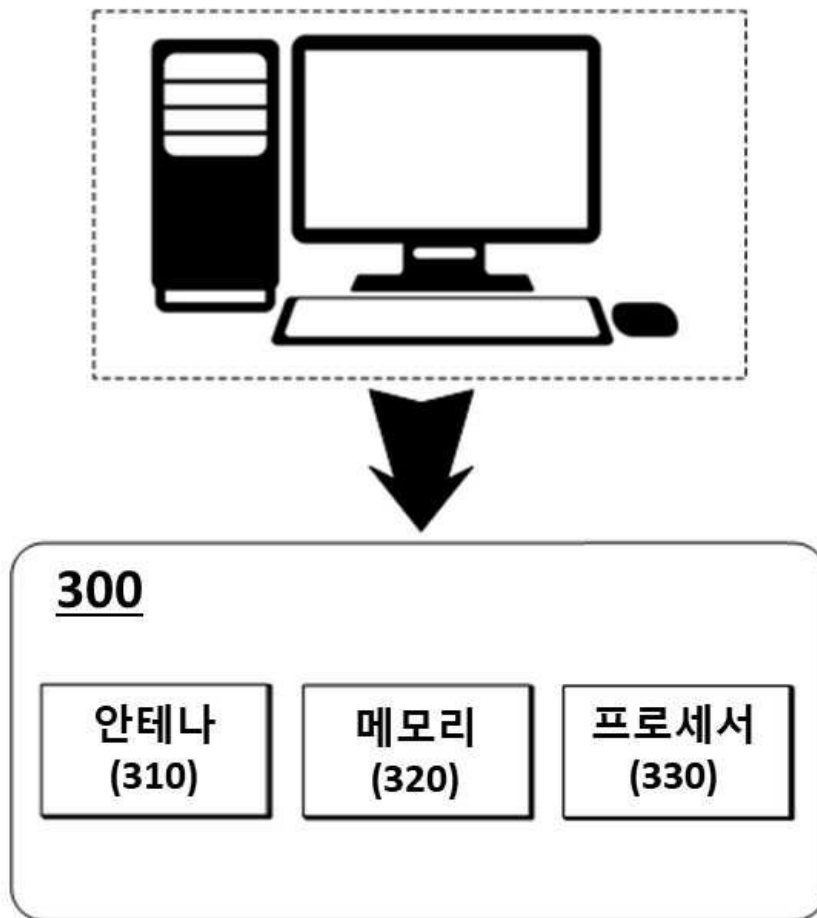


도면2

**200**

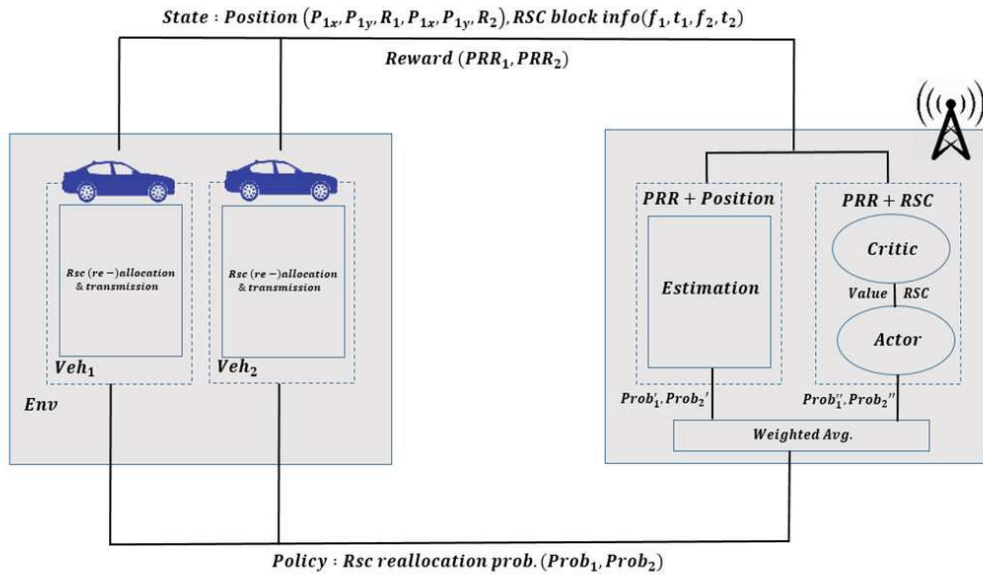


도면3



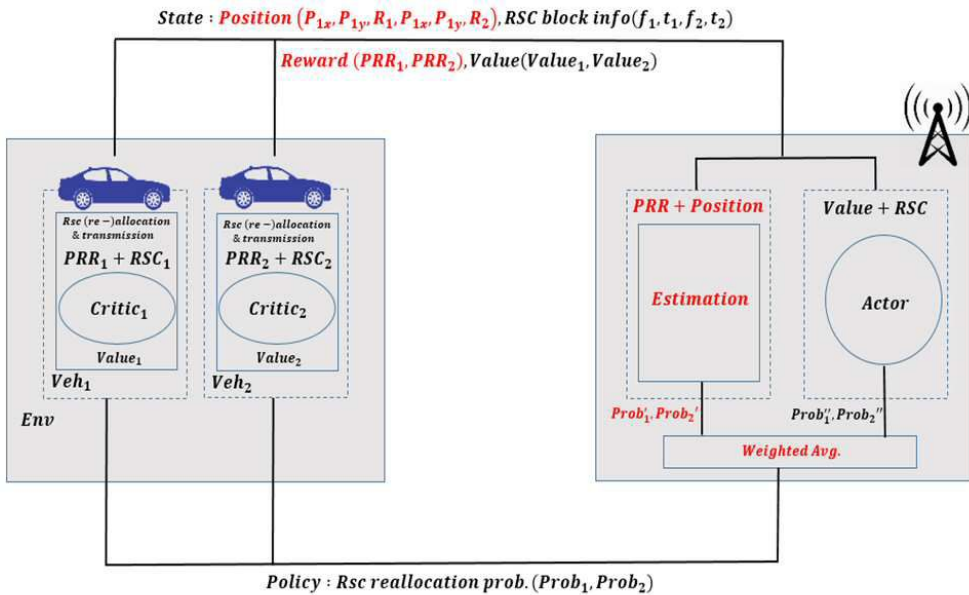
도면4

400



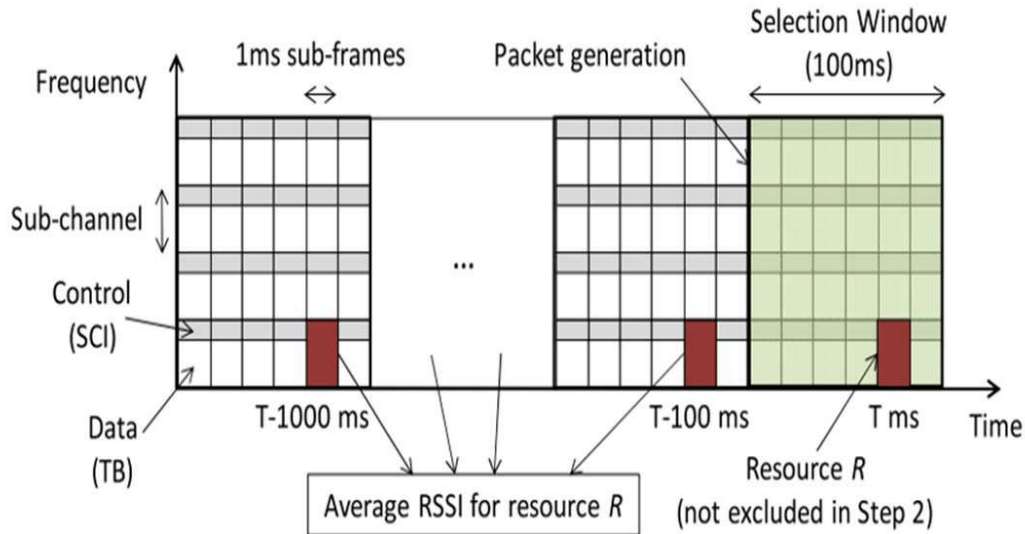
도면5

500



도면6

600



【심사관 직권보정사항】

【직권보정 1】

【보정항목】 청구범위

【보정세부항목】 청구항 11

【변경전】

제 1 차량을 포함하는 복수의 차량들에 대한 위치정보, 상기 복수의 차량들에 할당된 리소스 정보 및 리워드를 수신하는 안테나;

상기 복수의 차량들에 대한 정책을 결정하기 위한 복수의 강화학습 모델들을 저장하는 메모리; 및

상기 복수의 강화학습 모델들 중 제 1 강화학습 모델에 제 1 상태 정보로서 상기 위치정보 및 제 1 리워드 정보로서 상기 리워드를 입력하여 상기 복수의 차량들에 대한 자원 재할당 여부를 나타내는 제 1 확률을 계산하고, 상기 복수의 강화학습 모델들 중 제 2 강화학습 모델에 제 2 상태 정보로서 상기 리소스 정보 및 제 2 리워드 정보로서 상기 리워드를 입력하여 상기 복수의 차량들에 대한 자원 재할당 여부를 나타내는 제 2 확률을 계산하고, 상기 제 1 확률 및 상기 제 2 확률에 대한 계산 결과를 토대로 상기 정책을 결정하는 프로세서;를 포함하되,

상기 제 2 강화학습 모델은 제 2 액터 네트워크 및 제 2 크리틱 네트워크를 포함하고,

상기 제 2 크리틱 네트워크는 상기 제 1 차량 및 상기 디바이스에 각각 저장되고,

상기 프로세서는 상기 복수의 차량들 중 상기 제 1 차량으로부터 소정 거리 내에 위치한 차량들에 의한 간섭 여부에 따라 상기 제 1 차량에 저장된 상기 제 2 크리틱 네트워크를 이용할지 여부를 결정하는 차량 네트워크의 자원 재할당 제어 장치.

【변경후】

제 1 차량을 포함하는 복수의 차량들에 대한 위치정보, 상기 복수의 차량들에 할당된 리소스 정보 및 리워드를 수신하는 안테나;

상기 복수의 차량들에 대한 정책을 결정하기 위한 복수의 강화학습 모델들을 저장하는 메모리; 및



상기 복수의 강화학습 모델들 중 제 1 강화학습 모델에 제 1 상태 정보로서 상기 위치정보 및 제 1 리워드 정보로서 상기 리워드를 입력하여 상기 복수의 차량들에 대한 자원 재할당 여부를 나타내는 제 1 확률을 계산하고, 상기 복수의 강화학습 모델들 중 제 2 강화학습 모델에 제 2 상태 정보로서 상기 리소스 정보 및 제 2 리워드 정보로서 상기 리워드를 입력하여 상기 복수의 차량들에 대한 자원 재할당 여부를 나타내는 제 2 확률을 계산하고, 상기 제 1 확률 및 상기 제 2 확률에 대한 계산 결과를 토대로 상기 정책을 결정하는 프로세서;를 포함하되,

상기 제 2 강화학습 모델은 제 2 액터 네트워크 및 제 2 크리틱 네트워크를 포함하고,

상기 제 2 크리틱 네트워크는 상기 제 1 차량 및 상기 메모리에 각각 저장되고,

상기 프로세서는 상기 복수의 차량들 중 상기 제 1 차량으로부터 소정 거리 내에 위치한 차량들에 의한 간섭 여부에 따라 상기 제 1 차량에 저장된 상기 제 2 크리틱 네트워크를 이용할지 여부를 결정하는 차량 네트워크의 자원 재할당 제어 장치.

## 【직권보정 2】

【보정항목】 청구범위

【보정세부항목】 청구항 13

## 【변경전】

제 11 항에 있어서,

상기 프로세서가 상기 제 1 차량에 저장된 상기 제 2 크리틱 네트워크를 이용하는 경우, 상기 프로세서는 상기 제 1 차량의 제 2 크리틱 네트워크로부터 출력된 값(Value)을 수신하여 동작하는 차량 네트워크의 자원 재할당 제어 장치.

## 【변경후】

제 11 항에 있어서,

상기 프로세서가 상기 제 1 차량에 저장된 상기 제 2 크리틱 네트워크를 이용하는 경우, 상기 프로세서는 상기 제 1 차량의 제 2 크리틱 네트워크로부터 출력된 값(Value)을 수신하여 동작하는 차량 네트워크의 자원 재할당 제어 장치.