



(19) 대한민국특허청(KR)  
(12) 등록특허공보(B1)

(45) 공고일자 2024년03월19일  
(11) 등록번호 10-2648565  
(24) 등록일자 2024년03월13일

(51) 국제특허분류(Int. Cl.)  
H04B 7/06 (2017.01) G06N 3/092 (2023.01)  
H04W 72/04 (2009.01) H04W 84/12 (2009.01)

(52) CPC특허분류  
H04B 7/06952 (2023.05)  
G06N 3/092 (2023.01)  
(21) 출원번호 10-2023-0123754  
(22) 출원일자 2023년09월18일  
심사청구일자 2023년09월18일

(56) 선행기술조사문헌  
"IEEE 802.11ay의 비대칭 링크 빔포밍 트레이닝을  
위한 Q-learning 기반의 STS 조절 기법", Journal  
of KIIT. Vol. 18, No. 11 pp. 73-79,  
(2020.11.30)\*  
\*는 심사관에 의하여 인용된 문헌

(73) 특허권자  
세종대학교산학협력단  
서울특별시 광진구 능동로 209 (군자동, 세종대학  
교)  
(72) 발명자  
김문석  
서울특별시 광진구 능동로 209(군자동)  
(74) 대리인  
수안특허법인, 민영준

전체 청구항 수 : 총 10 항

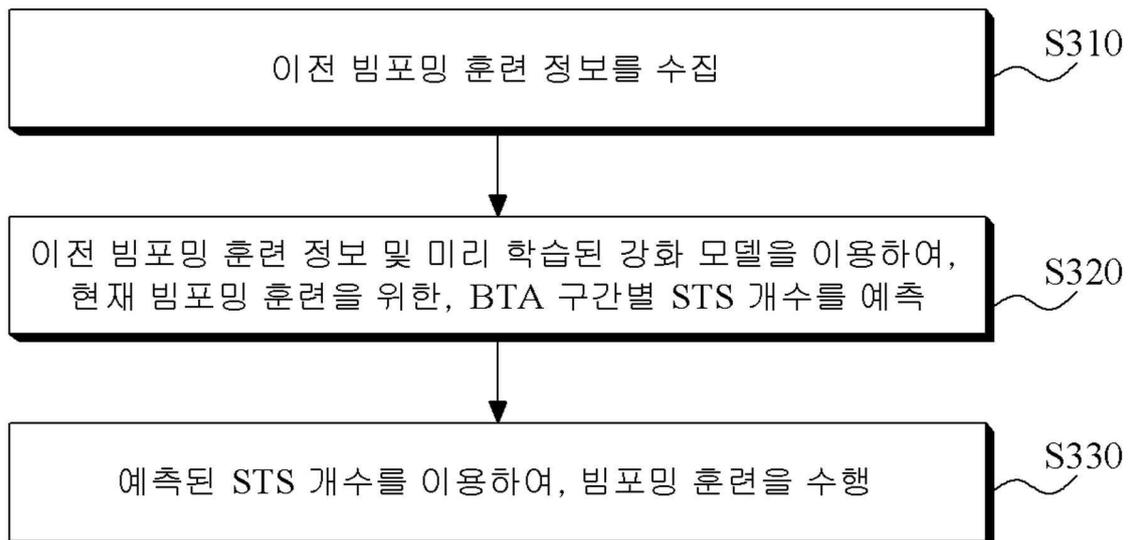
심사관 : 김수남

(54) 발명의 명칭 비대칭 링크를 위한 강화 학습 기반의 빔포밍 훈련 방법 및 장치

(57) 요약

무선랜 시스템에서 수행되는, 비대칭 링크를 위한 강화 학습 기반의 빔포밍 훈련 방법이 개시된다. 개시된 비대칭 링크를 위한 강화 학습 기반의 빔포밍 훈련 방법은 이전 빔포밍 훈련 정보를 수집하는 단계; 및 상기 이전 빔포밍 훈련 정보 및 미리 학습된 강화 학습 모델을 이용하여, 현재 빔포밍 훈련을 위한, BTA(Beamforming Training Allocation) 구간별 STS(Space-Time slot) 개수를 예측하는 단계; 및 상기 예측된 STS 개수를 이용하여, 빔포밍 훈련을 수행하는 단계를 포함하며, 상기 강화 학습 모델의 상태 정보는 상기 이전 빔포밍 훈련 정보에 따라 결정되며, 상기 강화 학습 모델의 액션 정보는 상기 BTA 구간별 STS 개수를 포함한다.

대표도 - 도3



(52) CPC특허분류

- H04B 7/0617* (2013.01)
- H04W 72/0446* (2023.01)
- H04W 72/046* (2013.01)
- H04W 84/12* (2013.01)

이 발명을 지원한 국가연구개발사업

과제고유번호	1711193909
과제번호	2021-0-01816-003
부처명	과학기술정보통신부
과제관리(전문)기관명	정보통신기획평가원
연구사업명	정보통신방송혁신인재양성
연구과제명	메타버스 자율트윈 핵심기술 연구
기여율	1/2
과제수행기관명	세종대학교 산학협력단
연구기간	2023.01.01 ~ 2023.12.31

이 발명을 지원한 국가연구개발사업

과제고유번호	1711186006
과제번호	2022R1G1A1012117
부처명	과학기술정보통신부
과제관리(전문)기관명	한국연구재단
연구사업명	개인기초연구(과기정통부)
연구과제명	딥러닝 기반 차세대 밀리미터파 무선랜 통신 및 센싱 기술 연구
기여율	1/2
과제수행기관명	세종대학교 산학협력단
연구기간	2023.03.01 ~ 2024.02.29

---

**명세서**

**청구범위**

**청구항 1**

비대칭 링크를 위한 강화 학습 기반의 빔포밍 훈련 방법에 있어서,

이전 빔포밍 훈련 정보를 수집하는 단계; 및

상기 이전 빔포밍 훈련 정보 및 미리 학습된 강화 학습 모델을 이용하여, 현재 빔포밍 훈련을 위한, BTA(Beamforming Training Allocation) 구간별 STS(Space-Time slot) 개수를 예측하는 단계; 및

상기 예측된 STS 개수를 이용하여, 빔포밍 훈련을 수행하는 단계를 포함하며,

상기 강화 학습 모델의 상태 정보는

상기 이전 빔포밍 훈련 정보에 따라 결정되며,

상기 강화 학습 모델의 액션 정보는

상기 BTA 구간별 STS 개수를 포함하며,

상기 이전 빔포밍 훈련 정보는

상기 현재 빔포밍 훈련 이전에 수행된 빔포밍 훈련 중 미리 설정된 횟수만큼의 빔포밍 훈련에 대한 정보를 포함하며, 상기 BTA 구간별로, SSW(Sector Sweep) 프레임 전송에 성공한 단말 정보와 충돌이 발생한 STS 개수 정보를 포함하며,

상기 빔포밍 훈련 방법은

상기 이전 빔포밍 훈련 각각으로부터, 상기 현재 빔포밍 훈련 과정의 상기 BTA 구간 각각에서 타겟 단말이 상기 SSW 프레임을 전송할 확률값을 계산하는 단계;

상기 이전 빔포밍 훈련 각각에 할당된 가중치를 이용해, 상기 확률값에 대한 가중합을, 상기 BTA 구간 각각에 대해 계산하는 단계; 및

상기 가중합 중 최대 가중합에 대응되는 BTA 구간에서, 상기 타겟 단말이 상기 SSW 프레임을 전송함을 나타내는 상기 상태 정보를 생성하는 단계를 더 포함하며,

상기 타겟 단말은, 상기 이전 빔포밍 훈련 과정에서 상기 SSW 프레임 전송을 적어도 한번 성공한 단말인

비대칭 링크를 위한 강화 학습 기반의 빔포밍 훈련 방법.

**청구항 2**

삭제

**청구항 3**

삭제

**청구항 4**

제 1항에 있어서,

상기 확률값을 계산하는 단계는

제1이전 빔포밍 훈련 과정에서, 상기 타겟 단말이 타겟 BTA 구간에서 상기 SSW 프레임 전송을 성공한 경우에, 상기 타겟 단말이 상기 타겟 BTA 구간에서 상기 SSW 프레임을 전송할 확률값으로 1을 할당하고,

상기 제1이전 빔포밍 훈련 과정에서, 상기 타겟 BTA 구간에 충돌이 발생한 STS가 존재하고, 상기 타겟 단말이

타겟 BTA 구간에서 상기 SSW 프레임 전송을 미성공하되, 제2이전 빔포밍 훈련 과정에서 상기 타겟 단말이 상기 타겟 BTA 구간에서 상기 SSW 프레임 전송을 성공한 경우에, 상기 타겟 단말이 상기 타겟 BTA 구간에서 상기 SSW 프레임을 전송할 확률값으로 1을 할당하며,

상기 제2이전 빔포밍 훈련은, 상기 제1이전 빔포밍 훈련보다 이전에 수행된 훈련인  
비대칭 링크를 위한 강화 학습 기반의 빔포밍 훈련 방법.

#### 청구항 5

제 1항에 있어서,  
상기 가중치는  
상기 현재 빔포밍 훈련과, 상기 이전 빔포밍 훈련 사이의 시간 간격에 반비례하는  
비대칭 링크를 위한 강화 학습 기반의 빔포밍 훈련 방법.

#### 청구항 6

제 1항에 있어서,  
상기 현재 빔포밍 훈련 과정에서, 상기 BTA 구간 별로 충돌이 발생한 STS 개수 정보와, SSW 프레임 전송이 성공한 STS 개수 정보를 수집하는 단계; 및  
상기 수집된 STS 개수 정보를 이용하여, 상기 액션 정보에 대한 보상값을 생성하는 단계를 더 포함하는 비대칭 링크를 위한 강화 학습 기반의 빔포밍 훈련 방법.

#### 청구항 7

제 6항에 있어서,  
상기 액션 정보에 대한 보상값은  
상기 BTA 구간별로 계산된 BTA 보상값의 총합에 대응되며,  
상기 BTA 보상값은  
상기 BTA 구간의 총 STS 개수 및 상기 수집된 STS 개수 정보에 따라 결정되는  
비대칭 링크를 위한 강화 학습 기반의 빔포밍 훈련 방법.

#### 청구항 8

제 1항에 있어서,  
강화 학습 모델의 가치 함수와 정책 함수에 대한 딥러닝 네트워크를 이용하여, 상기 가치 함수 및 정책 함수의 출력값을 획득하는 단계; 및  
상기 출력 값을 손실 함수에 적용하여, 상기 손실 함수의 손실값이 최소가 되도록 상기 딥러닝 네트워크의 가중치를 업데이트하는 단계를 더 포함하는 비대칭 링크를 위한 강화 학습 기반의 빔포밍 훈련 방법.

#### 청구항 9

현재 빔포밍 훈련 이전에 수행된, 미리 설정된 횡수만큼의 이전 빔포밍 훈련에 대한 정보를 수집하는 단계;

상기 이전 빔포밍 훈련에 대한 정보를 이용하여, 상기 현재 빔포밍 훈련 과정의 BTA 구간 각각에서 SSW 프레임 을 전송할 단말의 개수를 나타내는 상태 정보를 생성하는 단계;

상기 상태 정보 및 미리 학습된 강화 학습 모델을 이용하여, 다음 빔포밍 훈련을 위한, BTA 구간별 STS개수 를 예측하는 단계; 및

상기 예측된 STS 개수를 이용하여, 빔포밍 훈련을 수행하는 단계를 포함하며,

상기 상태 정보를 생성하는 단계는

상기 이전 빔포밍 훈련 각각으로부터, 상기 현재 빔포밍 훈련 과정의 상기 BTA 구간 각각에서 타겟 단말이 상기 SSW 프레임을 전송할 확률값을 계산하는 단계;

상기 이전 빔포밍 훈련 각각에 할당된 가중치를 이용해, 상기 확률값에 대한 가중합을, 상기 BTA 구간 각각에 대해 계산하는 단계; 및

상기 가중합 중 최대 가중합에 대응되는 BTA 구간에서, 상기 타겟 단말이 상기 SSW 프레임을 전송함을 나타내는 상기 상태 정보를 생성하는 단계를 포함하며,

상기 타겟 단말은, 상기 이전 빔포밍 훈련 과정에서 상기 SSW 프레임 전송을 적어도 한번 성공한 단말인

비대칭 링크를 위한 강화 학습 기반의 빔포밍 훈련 방법.

#### 청구항 10

삭제

#### 청구항 11

제 9항에 있어서,

상기 확률값을 계산하는 단계는

제1이전 빔포밍 훈련 과정에서, 상기 타겟 단말이 타겟 BTA 구간에서 상기 SSW 프레임 전송을 성공한 경우에, 상기 타겟 단말이 상기 타겟 BTA 구간에서 상기 SSW 프레임을 전송할 확률값으로 1을 할당하고,

상기 제1이전 빔포밍 훈련 과정에서, 상기 타겟 BTA 구간에 충돌이 발생한 STS가 존재하고, 상기 타겟 단말이 타겟 BTA 구간에서 상기 SSW 프레임 전송을 미성공하되, 제2이전 빔포밍 훈련 과정에서 상기 타겟 단말이 상기 타겟 BTA 구간에서 상기 SSW 프레임 전송을 성공한 경우에, 상기 타겟 단말이 상기 타겟 BTA 구간에서 상기 SSW 프레임을 전송할 확률값으로 1을 할당하며,

상기 제2이전 빔포밍 훈련은, 상기 제1이전 빔포밍 훈련보다 이전에 수행된 훈련인

비대칭 링크를 위한 강화 학습 기반의 빔포밍 훈련 방법.

#### 청구항 12

제 9항에 있어서,

상기 가중치는

상기 현재 빔포밍 훈련과, 상기 이전 빔포밍 훈련 사이의 시간 간격에 반비례하는

비대칭 링크를 위한 강화 학습 기반의 빔포밍 훈련 방법.

#### 청구항 13

제 9항에 있어서,

상기 이전 빔포밍 훈련에 대한 정보를 수집하는 단계는

상기 BTA 구간별로 SSW 프레임 전송에 성공한 단말 정보와 충돌이 발생한 STS 개수 정보

를 포함하는 비대칭 링크를 위한 강화 학습 기반의 빔포밍 훈련 방법.

**청구항 14**

삭제

**발명의 설명**

**기술 분야**

[0001] 본 발명은 무선랜 시스템의 비대칭 링크를 위한 빔포밍 훈련 방법 및 장치에 관한 것으로서, 더욱 상세하게는 강화 학습 기반의 빔포밍 훈련 방법 및 장치에 관한 것이다.

**배경 기술**

[0003] 초연결(Hyper-connected) 시대에 진입함에 따라 6GHz 이하의 주파수 대역은 포화 상태이며, 데이터 수요는 나날이 기하급수적으로 증가하고 있다. 이에 넓은 대역폭을 활용하여 대량의 데이터 전송이 가능한 Millimeter-wave (mmWave)가 많은 주목을 받고 있다. 2012년 승인된 IEEE 802.11ad는 처음으로 60GHz mmWave를 사용한 Wi-Fi 기술이다. 그리고 Augmented Reality (AR)/Virtual Reality (VR), 실시간 multi-view 8K video 등과 같이 새로운 애플리케이션을 위해서 IEEE 802.11ad를 기반으로 100Gbps의 처리량 목표의 IEEE 802.11ay 표준이 제안되었다.

[0004] mmWave는 넓은 대역폭을 사용한다는 장점이 있는 반면, 주파수가 높아진 만큼 대기 중이나 전송선로에서 손실률이 많이 증가한다는 단점이 있다. 이 같은 문제 해결을 위해 mmWave 통신 시스템에서는 위상 배열 안테나 (phased-array antenna)를 사용해 원하는 방향으로 안테나 빔 (beam)을 움직여 지향성 빔(directional beam)을 생성하며, 이 같은 기술을 빔포밍(Beamforming, BF)이라고 한다. 빔포밍 사용 시 송수신 노드가 통신을 위해 사용할 적합한 지향성 빔을 찾는 것이 중요하며, 이러한 지향성 빔을 찾는 과정을 빔포밍 훈련(BF Training)이라고 한다. 빔포밍 훈련 수행시에, 송신 섹터 개념을 사용하며, 송신 섹터란 안테나 영역을 방향성 조절을 통해 여러 개로 세분화한 영역을 뜻한다.

[0005] 빔포밍 과정에서 액세스 포인트(AP)는 송신 섹터 별로 프레임을 전송하며, 단말(station)은 프레임이 어느 방향에서 전송되는지 모르기 때문에, 가장 넓은 빔 폭으로 프레임을 수신하기 위해 유사 전방향(quasi omnidirectional) 모드를 이용한다. 반대로 단말이 프레임을 전송할 때는 액세스 포인트가 유사 전방향 모드를 이용한다.

[0006] 일반적으로 단말은 액세스 포인트보다 적은 수의 안테나를 가지고 있기 때문에, 단말의 전송 파워는 액세스 포인트보다 약하다. 따라서, 액세스 포인트가 전송하는 지향성 프레임을, 단말은 유사 전방향 모드에서 수신 가능하지만, 단말이 전송하는 지향성 프레임을, 액세스 포인트는 유사 전방향 모드에서 수신하지 못하는 비대칭 링크 상황이 발생하며, IEEE 802.11ay에서는 이러한 비대칭 링크 상황을 위한 빔포밍 훈련 방법이 제안되었다.

[0007] IEEE 802.11ay에서 제안된 비대칭 링크 상황을 위한 빔포밍 훈련 방법에서는 BI(Beacon Interval)이 반복되며, BI는 액세스 포인트와 단말 사이에 컨트롤 프레임이 송수신되는 BHI(Beacon Header Interval)과 액세스 포인트와 단말 사이에 데이터가 송수신되는 DTI(Data Transmission Interval)로 구성된다. 그리고 BHI는 BTI(Beacon Transmission Interval), A-BFT(Association Beamforming Training) 및 ATI(Announcement Transmission Interval)로 구성된다.

[0008] 관련 선행문헌으로 특허 문헌인 대한민국 등록특허 제10-2508071호, 대한민국 공개특허 제2022-0149228호가 있으며, 비특허 문헌인 "Beamforming Training for Asymmetric Links in IEEE 802.11ay: Implementation and Performance Evaluation, Yena Kim, Journal of The Korea Society of Computer and Information Vol. 25 No. 11, pp. 89-95, November 2020"가 있다.

**발명의 내용**

**해결하려는 과제**

[0010] 본 발명은, 비대칭 링크를 위한 빔포밍 훈련이 효율적으로 수행되어, 비대칭 링크를 위한 빔포밍 훈련에 소요되는 시간이 감소할 수 있는, 강화 학습 기반의 빔포밍 훈련 방법 및 장치를 제공하기 위한 것이다.

**과제의 해결 수단**

[0012] 상기한 목적을 달성하기 위한 본 발명의 일 실시예에 따르면, 이전 빔포밍 훈련 정보를 수집하는 단계; 및 상기 이전 빔포밍 훈련 정보 및 미리 학습된 강화 학습 모델을 이용하여, 현재 빔포밍 훈련을 위한, BTA(Beamforming Training Allocation) 구간별 STS(Space-Time slot) 개수를 예측하는 단계; 및 상기 예측된 STS 개수를 이용하여, 빔포밍 훈련을 수행하는 단계를 포함하며, 상기 강화 학습 모델의 상태 정보는 상기 이전 빔포밍 훈련 정보에 따라 결정되며, 상기 강화 학습 모델의 액션 정보는 상기 BTA 구간별 STS 개수를 포함하는 비대칭 링크를 위한 강화 학습 기반의 빔포밍 훈련 방법이 제공된다.

[0013] 또한 상기한 목적을 달성하기 위한 본 발명의 다른 실시예에 따르면, 현재 빔포밍 훈련 이전에 수행된, 미리 설정된 횟수만큼의 이전 빔포밍 훈련에 대한 정보를 수집하는 단계; 상기 이전 빔포밍 훈련에 대한 정보를 이용하여, 상기 현재 빔포밍 훈련 과정의 BTA 구간 각각에서 SSW 프레임 전송할 단말의 개수를 나타내는 상태 정보를 생성하는 단계; 상기 상태 정보 및 미리 학습된 강화 학습 모델을 이용하여, 다음 빔포밍 훈련을 위한, BTA 구간별 STS개수를 예측하는 단계; 및 상기 예측된 STS 개수를 이용하여, 빔포밍 훈련을 수행하는 단계를 포함하는 비대칭 링크를 위한 강화 학습 기반의 빔포밍 훈련 방법이 제공된다.

[0014] 또한 상기한 목적을 달성하기 위한 본 발명의 또 다른 실시예에 따르면, 위상 배열 안테나; 메모리; 및 상기 메모리와 전기적으로 연결된 적어도 하나의 프로세서를 포함하며, 상기 프로세서는 이전 빔포밍 훈련 정보를 수집하고, 상기 이전 빔포밍 훈련 정보 및 미리 학습된 강화 학습 모델을 이용하여, 현재 빔포밍 훈련을 위한, BTA 구간별 STS 개수를 예측하며, 상기 강화 학습 모델의 상태 정보는 상기 이전 빔포밍 훈련 정보에 따라 결정되며, 상기 강화 학습 모델의 액션 정보는 상기 BTA 구간별 STS 개수를 포함하는, 비대칭 링크를 위한 강화 학습 기반의 빔포밍 훈련 장치가 제공된다.

**발명의 효과**

[0016] 본 발명의 일실시예에 따르면, BTA 구간에 할당될 STS의 개수를 예측함으로써, 비대칭 링크를 위한 빔포밍 훈련에 소요되는 시간이 감소될 수 있다.

**도면의 간단한 설명**

- [0018] 도 1은 무선랜 시스템에서 비대칭 링크를 위한 빔포밍 훈련 방법을 설명하기 위한 도면이다.
- 도 2는 송신 섹터의 일예를 도시하는 도면이다.
- 도 3은 본 발명의 일실시예에 따른 비대칭 링크를 위한 강화 학습 기반의 빔포밍 훈련 방법을 설명하기 위한 도면이다.
- 도 4는 본 발명의 일실시예에 따른 강화 학습 모델을 설명하기 위한 도면이다.
- 도 5는 본 발명의 일실시예에 따른 강화 학습 모델의 업데이트 방법을 설명하기 위한 도면이다.
- 도 6은 본 발명의 다른 실시예에 따른 비대칭 링크를 위한 강화 학습 기반의 빔포밍 훈련 방법을 설명하기 위한 도면이다.

**발명을 실시하기 위한 구체적인 내용**

[0019] 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. 각 도면을 설명하면서 유사한 참조부호를 유사한 구성요소에 대해 사용하였다.

[0020] 이하에서, 본 발명에 따른 실시예들을 첨부된 도면을 참조하여 상세하게 설명한다.

- [0022] 도 1은 무선랜 시스템에서 비대칭 링크를 위한 빔포밍 훈련 방법을 설명하기 위한 도면이며, 도 2는 송신 섹터의 일예를 도시하는 도면이다.
- [0023] 도 2에서는, 액세스 포인트가 2개의 위상 배열 안테나(210, 220)을 이용하며, 제1위상 배열 안테나(210)에는 제 1 및 제2송신 섹터(TS1, TS2)가 할당되고, 제2위상 배열 안테나(220)에는 제3 및 제4송신 섹터(TS3, TS4)가 할당된 예시가 도시된다.
- [0024] 비대칭 링크 상황에서 빔포밍 훈련을 위해 액세스 포인트는, BTI 구간에서 송신 섹터 별로 비콘(beacon) 프레임을 전송한다. 액세스 포인트는 BTI 구간에서 송신 섹터 별로 비콘 프레임을 전송할 때, 각 비콘 프레임에 TRN-R(Training-RX) 서브 필드를 추가하여 전송한다.
- [0025] 단말은 액세스 포인트가 전송한 비콘 프레임을 수신할 때마다, TRN-R 서브필드를 제외한 비콘 프레임은 유사 전 방향 패턴으로 수신하고, TRN-R 서브필드들을 수신할 때는 서로 다른 수신 섹터를 적용하여, 수신 품질이 가장 좋은 최적의 수신 섹터를 결정할 수 있다. 안테나의 쌍대성(reciprocal)을 고려했을 때, 액세스 포인트로부터의 전송에 대한 단말의 최적 수신 섹터는, 반대로 단말이 액세스 포인트로 데이터를 전송할 때의 최적 송신 섹터에 대응된다.
- [0026] 액세스 포인트는 단말 각각과의 최적 송신 및 수신 섹터를 확인하기 위해, DTI 구간에, BTA(Beamforming Training Allocation) 구간을 할당한다. DTI 구간에 할당되는 BTA 구간의 개수는, 액세스 포인트의 송신 섹터의 개수에 대응된다. 즉, 송신 섹터 별로 BTA 구간이 할당되며, 도 2와 같은 예시에서는 4개의 BTA 구간이 할당될 수 있다. 그리고 하나의 BTA 구간 내에는 하나의 단말이 SSW(Sector Sweep) 프레임을 전송할 수 있는 적어도 하나의 STS(space-time slot)와, 액세스 포인트가 ACK(Ack) 프레임을 전송할 수 있는 시간이 할당된다.
- [0027] DTI 구간의 각 BTA 구간마다 액세스 포인트는, 단말로부터 전송되는 SSW 프레임의 수신을 대기하고, BTA 구간에 할당된 STS의 개수만큼을 대기한 이후, SSW 프레임을 수신한 송신 섹터로 ACK 프레임을 전송한다. 단말이 SSW 프레임을 전송할 때, 랜덤한 타이머를 이용하며, 타이머 값이 0이 되면 SSW 프레임을 전송한다. 액세스 포인트의 ACK 프레임에는 SSW 프레임 전송에 성공한 단말의 정보가 포함된다.
- [0028] 이와 같은 빔포밍 훈련 과정에서, 만일 특정 송신 섹터에 많은 단말이 밀집되어 있는 상태에서 STS의 개수가 충분하지 않다면, 해당 송신 섹터의 BTA 구간에서 단말들이 동시에 SSW 프레임을 전송하여 SSW 프레임의 충돌이 발생할 가능성이 높아진다. 충돌을 일으킨 단말은 다시 빔포밍 훈련을 통해 SSW 프레임을 전송해야 하므로, 무선랜 시스템 전체적으로 빔포밍 훈련 시간이 증가한다. 반대로 STS의 개수가 지나치게 많은 경우에도, 빔포밍 훈련 시간이 증가하기 때문에, 적절한 STS의 개수를 BTA 구간에 할당할 필요가 있다.
- [0029] 이에 본 발명은 BTA 구간 별로 적절한 STS의 개수를 예측하고, 예측된 STS의 개수에 기반하여 빔포밍 훈련을 수행하는 방법을 제안한다.
- [0030] 본 발명의 일실시예는 적절한 STS의 개수를 예측하기 위해, 강화 학습 모델을 이용한다. 강화 학습 모델의 상태 정보는 이전 빔포밍 훈련 정보에 따라 결정되며, 강화 학습 모델의 액션 정보는 BTA 구간별 STS 개수를 포함한다. 본 발명의 일실시예는 예측된 STS 개수만큼 BTA 구간에 STS를 할당하여, 빔포밍 훈련을 수행한다.
- [0031] 본 발명의 일실시예에 따른 빔포밍 훈련 방법은 컴퓨팅 장치에서 수행될 수 있으며, 이러한 컴퓨팅 장치는 위상 배열 안테나, 메모리 및 메모리와 전기적으로 연결된 적어도 하나의 프로세서를 포함한다. 액세스 포인트는 전술된 컴퓨팅 장치의 일예이며, 프로세서는 빔포밍 훈련을 위한 일련의 과정을 수행한다.
- [0033] 도 3은 본 발명의 일실시예에 따른 비대칭 링크를 위한 강화 학습 기반의 빔포밍 훈련 방법을 설명하기 위한 도면이며, 도 4는 본 발명의 일실시예에 따른 강화 학습 모델을 설명하기 위한 도면이다.
- [0034] 도 3을 참조하면, 본 발명의 일실시예에 따른 액세스 포인트는 현재 빔 포밍 훈련을 위해 이전 빔포밍 훈련 정보를 수집(S310)한다. 여기서, 이전 빔포밍 훈련 정보는 현재 빔포밍 훈련의 이전에 수행된 빔포밍 훈련에 대한 정보에 대응되며, 가장 최근 미리 설정된 횟수만큼 수행된 빔포밍 훈련에 대한 정보를 포함한다. 그리고 이전 빔포밍 훈련 과정에서의, BTA 구간별로 SSW 프레임 전송에 성공한 단말 정보와, BTA 구간별로 충돌이 발생한 STS 개수 정보를 포함할 수 있다.
- [0035] 그리고 액세스 포인트는 이전 빔포밍 훈련 정보 및 미리 학습된 강화 학습 모델을 이용하여, 현재 빔포밍 훈련을 위한, BTA 구간별 STS 개수를 예측(S320)한다. 여기서, 강화 학습 모델의 상태 정보는 이전 빔포밍 훈련 정보에 따라 결정되며, 강화 학습 모델의 액션 정보는 BTA 구간별 STS 개수를 포함한다.

[0036] 강화 학습 모델은 일실시예로서 도 3의 액터-크리틱(actor-critic) 모델일 수 있다. 강화 학습 모델에서 에이전트(RL Agent)는 액세스 포인트에 대응되며, 환경(Environment)은 빔포밍 훈련에 대응된다. 그리고 에이전트는 액터와 크리틱으로 구성될 수 있다. 에이전트, 즉 액세스 포인트는 현재 빔포밍 훈련의 상태(state)를 관찰(observation)하며, 관찰된 상태에서 보상(reward)이 최대가 되는 행동(action)을 수행한다. 액터와 크리틱은, TD(Temporal Difference) 에러를 통해 업데이트될 수 있다.

[0037] 즉, 강화 학습 모델은, 이전 빔포밍 훈련 상태에서부터 현재 빔포밍 훈련을 위한, BTA 구간별 STS 개수를 예측하는 액션을 수행한다. 전송된 바와 같이, BTA 구간은 액세스 포인트의 송신 섹터에 대응되므로, 강화 학습 모델은, 액세스 포인트의 송신 섹터의 개수에 대응되는 BTA 구간별로, STS 개수를 예측하는 액션을 수행한다.

[0038] 일실시예로서, 최근  $N_{win}$  번의 이전 빔포밍 훈련에 대한 정보가 수집되는 경우, 상태 정보 S는 [수학식 1]과 같이 표현될 수 있다.

**수학식 1**

$$S = (BFT_1, \dots, BFT_i, \dots, BFT_{N_{win}})$$

[0039]

[0040] 여기서,  $BFT_1$  는 현재 빔포밍 훈련의 바로 직전에 수행된 가장 최근의 빔포밍 훈련에 대한 정보를 나타내며, i는 이전 빔포밍 훈련에 대한 인덱스를 나타낸다.

[0041] 액세스 포인트의 송신 섹터의 개수, 즉 DTI 구간에 할당된 BTA 구간의 개수가  $N_t$  일 때,  $BFT_i$  는 [수학식 2]와 같이 표현될 수 있고,  $BTA_j$  는 [수학식 3]과 같이 표현될 수 있다.

**수학식 2**

$$BFT_i = (BTA_1, \dots, BTA_j, \dots, BTA_{N_t})$$

[0042]

**수학식 3**

$$BTA_j = (D, N_c)$$

[0043]

[0044] 여기서  $BTA_j$  는 송신 섹터 j에 대한 BTA 정보를 나타내며, j는 송신 섹터 및 BTA 구간에 대한 인덱스를 나타낸다. 그리고 D는 BTA 구간에서 SSW 프레임을 성공적으로 전송한 단말에 대한 식별 정보를 나타내며,  $N_c$  는 BTA 구간에서 충돌이 발생한 STS 개수 정보를 나타낸다.

[0045] 실시예에 따라서, 액세스 포인트는 도 6에서 후술되는 바와 같이, BTA 구간 각각에 포함되는 단말의 개수를 상태 정보로 이용할 수 있다.

[0046] 액세스 포인트는 예측된 STS 개수를 이용하여, 빔포밍 훈련을 수행(S330)한다. 액세스 포인트는 현재 빔포밍 훈련 과정에서 BTA 구간 별로 예측된 STS 개수만큼의 STS를 BTA 구간 별로 할당하고, 단말로부터 SSW 프레임을 수신한다.

[0047] 현재 빔포밍 훈련이 수행된 이후, 액세스 포인트는 액션 정보에 대한 보상값을 생성한다. 액세스 포인트는 일실

시예로서, 현재 빔포밍 훈련 과정에서, BTA 구간 별로 충돌이 발생한 STS 개수 정보와, SSW 프레임 전송이 성공한 STS 개수 정보를 수집하고, 수집된 STS 개수 정보를 이용하여, 액션 정보에 대한 보상값을 생성할 수 있다.

[0048] 이 때 액세스 포인트는 BTA 구간별로 BTA 보상값을 계산하고, BTA 보상값의 총합을 액션 정보에 대한 보상값으로 계산할 수 있다.

[0049] 일실시예로서, [수학식 4]를 이용하여 BTA 보상값( $\Gamma_j$ )을 계산할 수 있으며, [수학식 5]를 이용하여, 액션 정보에 대한 보상값(R)을 계산할 수 있다.

**수학식 4**

[0050] 
$$\Gamma_j = \frac{N_s^j}{N_{STS}^j + N_c^j}$$

**수학식 5**

[0051] 
$$R = \sum_{j=1}^{N_t} \Gamma_j$$

[0052] 여기서,  $N_s^j$  는 송신 섹터 j에 대응되는 BTA 구간에서 SSW 프레임 전송에 성공한 단말의 개수를 나타내며,  $N_c^j$  는 송신 섹터 j에 대응되는 BTA 구간에서 충돌이 발생한 STS의 개수를 나타낸다. 그리고  $N_{STS}^j$  는 송신 섹터 j에 대응되는 BTA 구간에 할당된 STS의 개수를 나타낸다.

[0053] 본 발명의 일실시예에 따르면, BTA 구간에 할당될 STS의 개수를 예측함으로써, 비대칭 링크를 위한 빔포밍 훈련에 소요되는 시간이 감소될 수 있다.

[0055] 도 5는 본 발명의 일실시예에 따른 강화 학습 모델의 업데이트 방법을 설명하기 위한 도면이다.

[0056] 본 발명의 일실시예에 따른 액세스 포인트는, 강화 학습 모델의 가치 함수와 정책 함수 딥러닝 네트워크에 대해 업데이트를 수행한다. 즉, 액세스 포인트는 강화 학습 모델을 이용해 STS의 개수를 예측하면서, 수집된 데이터를 이용하여 강화 학습 모델에 대한 학습을 수행하여, 딥러닝 네트워크를 업데이트한다.

[0057] 본 발명의 일실시예에 따른 강화 학습 모델에서 에이전트의 액터는 정책 함수를 이용하여 행동, 즉 STS의 개수를 예측한다. 그리고 에이전트의 크리틱(비평가)는 가치 함수를 이용하여 액터의 행동, 즉 예측된 STS의 개수가 적절한지 여부를 평가한다. 정책 함수와 가치 함수는 딥러닝 네트워크로 구현될 수 있으며, 가치 함수의 평가 결과가, 정책 함수 딥러닝 네트워크의 파라미터(가중치) 업데이트에 이용된다. 그리고 가치 함수 딥러닝 네트워크는, 행동의 수행 결과, 즉 보상에 따라 파라미터가 업데이트된다.

[0058] 도 5를 참조하면, 액세스 포인트는 강화 학습 모델의 정책 함수 및 가치 함수에 대한 딥러닝 네트워크를 이용하여 정책 함수 및 가치 함수의 출력값을 획득(S510)한다.

[0059] 정책 함수( $\pi^\theta(S, A)$ )는 특정 상태(S)에서 특정 행동(A)이 결정될 확률을 계산하기 위한 함수로서, 여기서  $\theta$  는 정책 함수 딥러닝 네트워크의 가중치를 나타낸다. 정책 함수 딥러닝 네트워크는 상태 정보를 입력받아,

입력된 상태 정보에 대해서 특정 행동을 결정할 확률에 대한, 평균( $\mu(S)$ )과 표준 편차( $\sigma(S)$ )를 출력하도록 학습된 네트워크이다. 액세스 포인트는 정책 함수 딥러닝 네트워크의 출력값으로부터 정규 분포

$\mathbb{N}(A|\mu(S), \sigma(S))$  ( )를 생성하여, 특정 행동을 결정할 확률을 계산할 수 있다. 예컨대, 전송된 업데이트 정보 중 가장 높은 확률에 대응되는 정보에 따라서, 액세스 포인트는, BTA 구간별로 할당될 STS의 개수를 결정할 수 있다.

[0060] 가치 함수( $V^v(S)$ )는 입력된 상태에 대한 보상의 기대값을 계산하기 위한 함수이며, 여기서  $v$  는 가치 함수 딥러닝 네트워크의 가중치를 나타낸다. 가치 함수 딥러닝 네트워크는 상태 정보를 입력받아, 입력된 상태 정보에 대한 보상 기대값을 출력하도록 학습된 네트워크이다.

[0061] 액세스 포인트는 정책 함수 및 가치 함수 딥러닝 네트워크의 출력값을 손실 함수에 적용하여, 손실 함수의 손실값이 최소가 되도록 정책 함수 및 가치 함수 딥러닝 네트워크의 가중치를 업데이트(S520)한다.

[0062] 액세스 포인트는 일실시예로서, 시간차 학습(Temporal Difference Learning)을 이용해, 정책 함수 및 가치 함수 딥러닝 네트워크의 가중치를 업데이트할 수 있다. 시간 t 시점에서 상태를  $S_t$  라 하고 해당 시점에서의 보상을

$\Gamma_t$  라고 할 때,  $V^v(S_t)$  의 실제 값에 대한 추정치 TD 타겟(Temporal Difference target,  $y_t$ )은 [수학식 6]을 통해 계산될 수 있다.

**수학식 6**

$$y_t = \Gamma_t + \gamma \cdot V^v(S_{t+1})$$

[0063]

여기서,  $\gamma$  은 감가율(Discount Factor)을 나타낸다.

[0064]

가치 함수 딥러닝 네트워크는 [수학식 7]과 같이 계산되는 TD 에러값( $\epsilon_t$ )이 최소가 되도록 업데이트될 수 있다.

[0065]

**수학식 7**

$$\epsilon_t = (y_t - V^v(S_t))^2$$

[0066]

그리고 정책 함수 딥러닝 네트워크는 [수학식 8]와 같이 표현되는 손실 함수의 손실값이 최소가 되도록 업데이트될 수 있다.

[0067]

**수학식 8**

$$Loss = -\log(\mathbb{N}(A|\mu(S_t), \sigma(S_t))) \cdot \epsilon_t$$

[0068]

도 6은 본 발명의 다른 실시예에 따른 비대칭 링크를 위한 강화 학습 기반의 빔포밍 훈련 방법을 설명하기 위한 도면이다.

[0070]

도 6을 참조하면, 본 발명의 일실시예에 따른 액세스 포인트는 현재 빔포밍 훈련 이전에 수행된, 미리 설정된

[0071]

횟수만큼의 이전 빔포밍 훈련에 대한 정보를 수집(S610)한다. 이전 빔포밍 훈련에 대한 정보는 전술된 바와 같이, BTA 구간별로 SSW 프레임 전송에 성공한 단말 정보와 충돌이 발생한 STS 개수 정보를 포함할 수 있다.

[0072] 그리고 액세스 포인트는, 단계 S610에서 수집된 이전 빔포밍 훈련에 대한 정보를 이용하여, 현재 빔포밍 훈련 과정의 BTA 구간 각각에서 SSW 프레임을 전송할 단말의 개수를 나타내는 상태 정보를 생성(S620)한다. 즉, 액세스 포인트는 BTA 구간 즉, 송신 섹터에 포함된 단말의 개수를 나타내는 상태 정보를 생성한다.

[0073] 단계 S620에서 액세스 포인트는 상태 정보를 생성하기 위해, 먼저 이전 빔포밍 훈련 각각으로부터, 현재 빔포밍 훈련 과정의 BTA 구간 각각에서 타겟 단말이 SSW 프레임을 전송할 확률값을 계산한다. 즉, 타겟 단말이 SSW 프레임을 전송할 확률값은, 이전 빔포밍 훈련 각각에 대해 계산된다. 여기서, 타겟 단말은 이전 빔포밍 훈련 과정에서 SSW 프레임을 적어도 한번 성공적으로 전송한 단말을 나타낸다.

[0074] 액세스 포인트는 제1이전 빔포밍 훈련 과정에서, 타겟 단말이 타겟 BTA 구간에서 SSW 프레임 전송을 성공한 경우에, 타겟 단말이 타겟 BTA 구간에서 SSW 프레임을 전송할 확률값으로 1을 할당한다. 그리고 제1이전 빔포밍 훈련 과정에서, 타겟 BTA 구간에 충돌이 발생한 STS가 존재하고, 타겟 단말이 타겟 BTA 구간에서 상기 SSW 프레임 전송을 미성공하더라도, 제2이전 빔포밍 훈련 과정에서 타겟 단말이 타겟 BTA 구간에서 SSW 프레임 전송을 성공한 경우라면, 타겟 단말이 타겟 BTA 구간에서 SSW 프레임을 전송할 확률값으로 1을 할당한다. 여기서 제2이전 빔포밍 훈련은, 제1이전 빔포밍 훈련보다 이전에 수행된 훈련을 나타낸다. 예컨대, 제1이전 빔포밍 훈련의 인덱스가  $i$ 인 경우, 제2이전 빔포밍 훈련의 인덱스는  $i-1$ 일 수 있다.

[0075] 전술된 케이스 이외의 경우, 액세스 포인트는, 타겟 단말이 타겟 BTA 구간에서 SSW 프레임을 전송할 확률값으로 0을 할당한다.

[0076] 액세스 포인트는  $N_{win}$  번의 이전 빔포밍 훈련 각각에 대해서, 현재 빔포밍 훈련 과정의 BTA 구간 각각에서 타겟 단말이 SSW 프레임을 전송할 확률값을 계산하고, 이전 빔포밍 훈련 각각에 할당된 가중치를 이용해, 확률값에 대한 가중합을, BTA 구간 각각에 대해 계산한다. 액세스 포인트는 [수학식 9]를 이용하여, 확률값에 대한

가중합(  $p_j(a)$  )을 계산할 수 있다.

**수학식 9**

$$p_j(a) = \sum_{i=1}^{N_{win}} \gamma_i \cdot p_{ij}(a)$$

[0077]

[0078] 여기서,  $\gamma_i$  는  $i$ 번째 이전 빔포밍 훈련에 대한 가중치를 나타내며,  $p_{ij}(a)$  는  $i$ 번째 이전 빔포밍 훈련으로부터 획득된, 현재 빔포밍 훈련 과정의  $j$ 번째 BTA 구간에서 타겟 단말이 SSW 프레임을 전송할 확률값을 나타낸다.

[0079] 그리고 가중치는 현재 빔포밍 훈련과, 이전 빔포밍 훈련 사이의 시간 간격에 반비례하도록 설정될 수 있다. 즉,  $N_{win}$  번의 이전 빔포밍 훈련 중, 가장 최근의 이전 빔포밍 훈련에 할당되는 가중치가 가장 크고, 가장 처음에 수행된 이전 빔포밍 훈련에 할당되는 가중치가 가장 작을 수 있다.

[0080] 그리고 액세스 포인트는 가중합 중 최대 가중합에 대응되는 BTA 구간에서, 타겟 단말이 SSW 프레임을 전송함을 나타내는 상태 정보를 생성할 수 있다. 예컨대, 5번의 이전 빔포밍 훈련으로부터 상태 정보가 생성되며, 4개의 송신 섹터가 이용된 경우, 타겟 단말이 현재 빔포밍 훈련 과정에서 4개의 BTA 구간 각각에 포함될 확률값이 5개가 계산되며, BTA 구간 별로 5개의 확률값에 대한 가중합이 계산된다. 그리고 현재 빔포밍 훈련 과정에서, 가중합이 가장 큰 BTA 구간에서, 타겟 단말이 SSW 프레임을 전송함을 나타내는 상태 정보가 생성될 수 있다.

[0081] 액세스 포인트는 상태 정보 및 미리 학습된 강화 학습 모델을 이용하여, 현재 빔포밍 훈련을 위한, BTA 구간

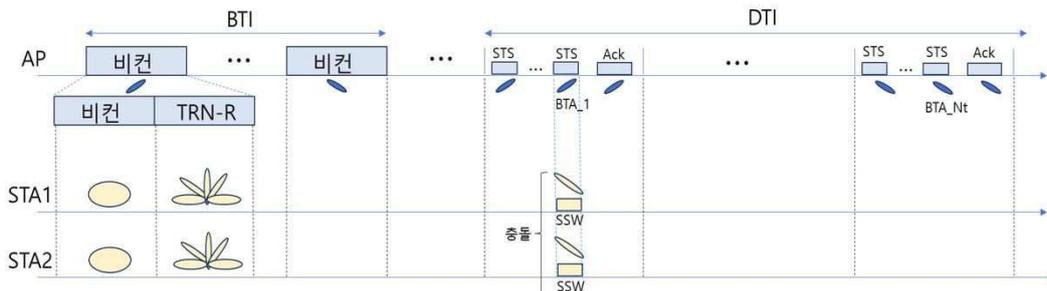
별 STS 개수를 예측(S630)하고, 예측된 STS 개수를 이용하여, 빔포밍 훈련을 수행(S640)한다.

[0083] 앞서 설명한 기술적 내용들은 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 실시예들을 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. 하드웨어 장치는 실시예들의 동작을 수행하기 위해 하나 이상의 소프트웨어 모듈로서 작동하도록 구성될 수 있으며, 그 역도 마찬가지이다.

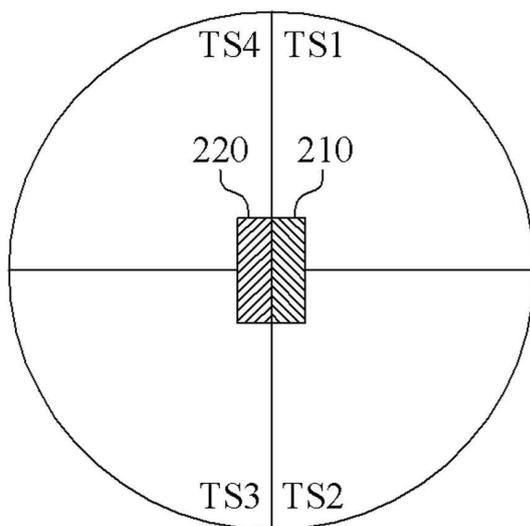
[0085] 이상과 같이 본 발명에서는 구체적인 구성 요소 등과 같은 특정 사항들과 한정된 실시예 및 도면에 의해 설명되었으나 이는 본 발명의 보다 전반적인 이해를 돕기 위해서 제공된 것일 뿐, 본 발명은 상기의 실시예에 한정되는 것은 아니며, 본 발명이 속하는 분야에서 통상적인 지식을 가진 자라면 이러한 기재로부터 다양한 수정 및 변형이 가능하다. 따라서, 본 발명의 사상은 설명된 실시예에 국한되어 정해져서는 아니되며, 후술하는 특허청구범위뿐 아니라 이 특허청구범위와 균등하거나 등가적 변형이 있는 모든 것들은 본 발명 사상의 범주에 속한다고 할 것이다.

도면

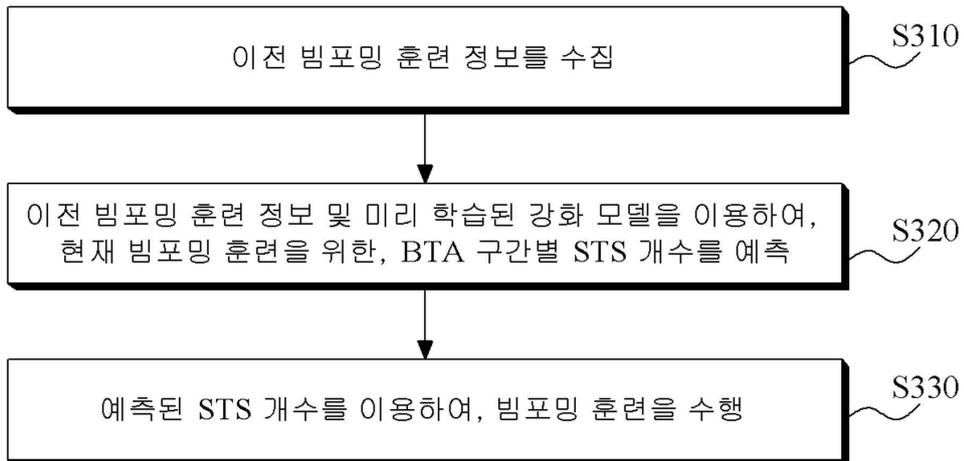
도면1



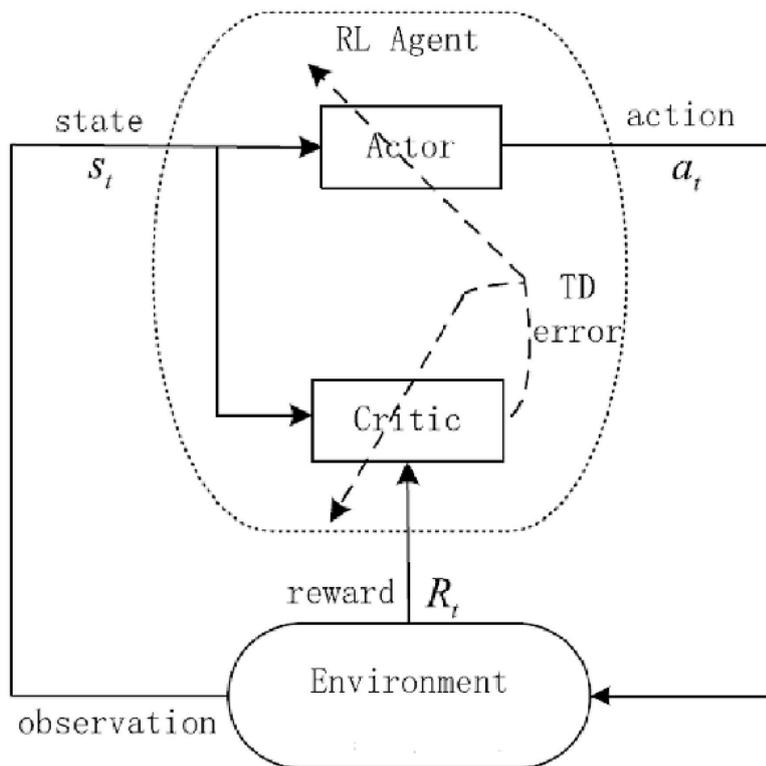
도면2



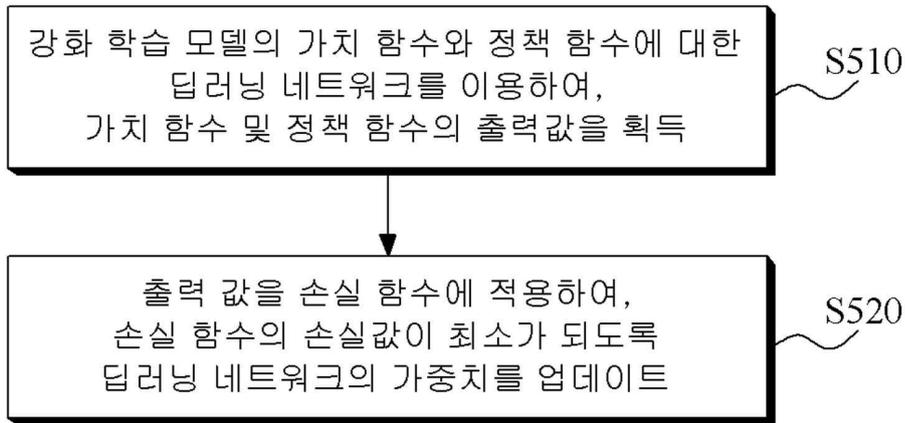
도면3



도면4



도면5



도면6

