



(19) 대한민국특허청(KR)
(12) 등록특허공보(B1)

(45) 공고일자 2023년03월08일
(11) 등록번호 10-2508071
(24) 등록일자 2023년03월06일

- (51) 국제특허분류(Int. Cl.)
H04B 7/06 (2017.01) G06N 3/08 (2023.01)
H04B 17/336 (2015.01) H04B 7/0452 (2017.01)
H04W 24/08 (2009.01)
- (52) CPC특허분류
H04B 7/0617 (2013.01)
G06N 3/08 (2023.01)
- (21) 출원번호 10-2022-0101909
(22) 출원일자 2022년08월16일
심사청구일자 2022년08월16일
- (56) 선행기술조사문헌
KR102388198 B1*
KR1020220069830 A
KR102154481 B1
*는 심사관에 의하여 인용된 문헌
- (73) 특허권자
세종대학교산학협력단
서울특별시 광진구 능동로 209 (군자동, 세종대학교)
- (72) 발명자
김문석
서울특별시 광진구 능동로 209(군자동)
- (74) 대리인
민영준

전체 청구항 수 : 총 10 항

심사관 : 구영희

(54) 발명의 명칭 강화 학습 기반의 빔포밍 훈련 방법 및 장치

(57) 요약

무선랜 시스템에서, 강화 학습 기반으로 빔포밍 훈련을 수행하는 방법과 장치가 개시된다. 개시된 강화 학습 기반의 빔포밍 훈련 방법은 단말로부터 SISO 피드백 정보를 수신하는 단계; 및 미리 학습된 강화 학습 모델을 이용하여, 액션 프레임의 전송에 이용되는 현재 송신 섹터 집합을 결정하는 단계를 포함하며, 상기 강화 학습 모델의 상태 집합은, 상기 SISO 피드백 정보 및 이전 MIMO 단계에서 액션 프레임의 전송에 이용된 이전 송신 섹터 집합을 포함하며, 상기 강화 학습 모델의 액션 집합은, 현재 MIMO 단계에 대한 상기 단말의 참여 정보 및 상기 현재 송신 섹터 집합에 대한 업데이트 정보를 포함한다.

대표도 - 도2



(52) CPC특허분류

H04B 17/336 (2015.01)
H04B 7/0452 (2013.01)
H04B 7/0682 (2013.01)
H04B 7/0695 (2013.01)
H04W 24/08 (2013.01)

이 발명을 지원한 국가연구개발사업

과제고유번호	1711152732
과제번호	2021-0-01816-002
부처명	과학기술정보통신부
과제관리(전문)기관명	정보통신기획평가원
연구사업명	정보통신방송혁신인재양성
연구과제명	메타버스 자율트윈 핵심기술 연구
기 여 율	1/2
과제수행기관명	세종대학교산학협력단
연구기간	2022.01.01 ~ 2022.12.31

이 발명을 지원한 국가연구개발사업

과제고유번호	1415181734
과제번호	00154678
부처명	산업통상자원부
과제관리(전문)기관명	한국산업기술평가관리원
연구사업명	시장선도를 위한 한국 주도형 K-Sensor 기술개발(R&D)
연구과제명	네트워크 기반 센서를 위한 커넥티드 지능센서 플랫폼 기술 개발
기 여 율	1/2
과제수행기관명	세종대학교산학협력단
연구기간	2022.04.01 ~ 2022.12.31

명세서

청구범위

청구항 1

단말로부터 SISO 피드백 정보를 수신하는 단계; 및

미리 학습된 강화 학습 모델을 이용하여, 액션 프레임의 전송에 이용되는 현재 송신 섹터 집합을 결정하는 단계를 포함하며,

상기 강화 학습 모델의 상태 집합은,

상기 SISO 피드백 정보 및 이전 MIMO 단계에서 액션 프레임의 전송에 이용된 이전 송신 섹터 집합을 포함하며,

상기 강화 학습 모델의 액션 집합은,

현재 MIMO 단계에 대한 상기 단말의 참여 정보 및 상기 현재 송신 섹터 집합에 대한 업데이트 정보를 포함하며,

상기 현재 및 이전 송신 섹터 집합은

위상 배열 안테나 각각에 대해 할당된 송신 섹터 중에서, 상기 액션 프레임의 전송에 이용되는 송신 섹터 조합을 포함하며,

상기 업데이트 정보는

BF 셋업 하위 단계 및 BF 선택 하위 단계를 위한 업데이트 정보이며,

상기 이전 송신 섹터 집합에, 새로운 송신 섹터 조합을 추가하는 제1정보;

상기 이전 송신 섹터 집합에 포함된 송신 섹터 조합 중에서, 적어도 하나의 송신 섹터 조합을 제거하는 제2정보;

상기 이전 송신 섹터 집합에 포함된 송신 섹터 조합 중에서, 적어도 하나를 새로운 송신 섹터 조합으로 변경하는 제3정보; 및

상기 이전 송신 섹터 집합을 상기 현재 송신 섹터 집합으로 이용하는 제4정보 중 적어도 하나를 포함하는

강화 학습 기반의 빔포밍 훈련 방법.

청구항 2

제 1항에 있어서,

상기 참여 정보는

상기 SISO 피드백 정보에 포함된 SNR값에 따라 결정되는 정보인

강화 학습 기반의 빔포밍 훈련 방법.

청구항 3

삭제

청구항 4

삭제

청구항 5

제 1항에 있어서,

상기 현재 송신 섹터 집합을 결정하는 단계는

상기 현재 MIMO 단계에 참여하고, 상기 이전 MIMO 단계에서 상기 액션 프레임 수신하지 못한 단말 중에서, 상기 SISO 피드백 정보의 SNR값이 임계값 이상인 단말의 개수가 최대인 송신 섹터 조합을 상기 이전 송신 섹터 집합에 추가하여, 상기 현재 송신 섹터 집합을 결정하는

강화 학습 기반의 빔포밍 훈련 방법.

청구항 6

제 5항에 있어서,

상기 현재 송신 섹터 집합을 결정하는 단계는

상기 현재 MIMO 단계에 참여하고, 상기 이전 MIMO 단계에서 상기 액션 프레임 수신하지 못한 단말 중에서, 상기 임계값 이상인 단말이 존재하지 않는 경우, 최대 SNR값에 대응되는 단말이 포함된 송신 섹터 조합을 상기 이전 송신 섹터 집합에 추가하여, 상기 현재 송신 섹터 집합을 결정하는

강화 학습 기반의 빔포밍 훈련 방법.

청구항 7

제 1항에 있어서,

상기 현재 송신 섹터 집합을 결정하는 단계는

상기 이전 송신 섹터 집합에 대해 상기 액션 프레임이 전송된 상황에서, 상기 액션 프레임을 수신하는 단말의 개수가 최소인 송신 섹터 조합을, 상기 이전 송신 섹터 집합에서 제거하는

강화 학습 기반의 빔포밍 훈련 방법.

청구항 8

제 1항에 있어서,

상기 현재 송신 섹터 집합을 결정하는 단계는

상기 이전 송신 섹터 집합에 대해 상기 액션 프레임이 전송된 상황에서, 상기 액션 프레임을 수신하는 단말의 개수가 최소인 송신 섹터 조합을 결정하고,

상기 현재 MIMO 단계에 참여하고, 상기 이전 MIMO 단계에서 상기 액션 프레임을 수신하지 못한 단말 중에서, 상기 SISO 피드백 정보의 SNR값이 임계값 이상인 단말의 개수가 최대인 송신 섹터 조합으로, 상기 결정된 송신 섹터 집합을 변경하는

강화 학습 기반의 빔포밍 훈련 방법.

청구항 9

제 1항에 있어서,

상기 현재 송신 섹터 집합의 송신 섹터 조합에 포함된 단말 중에서, 상기 액션 프레임의 수신이 가능한 단말을 확인하여, BF 훈련 하위 단계를 위한 송신 섹터 조합을 결정하는 단계

를 더 포함하는 강화 학습 기반의 빔포밍 훈련 방법.

청구항 10

제 1항에 있어서,

상기 강화 학습 모델의 가치 함수와 정책 함수에 대한 딥러닝 네트워크에 대해 업데이트를 수행하는 단계
를 더 포함하는 강화 학습 기반의 빔포밍 훈련 방법.

청구항 11

제 10항에 있어서,

상기 딥러닝 네트워크에 대해 업데이트를 수행하는 단계는

상기 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 가중
치를 업데이트하는 단계

를 포함하는 강화 학습 기반의 빔포밍 훈련 방법.

청구항 12

단말로부터 SISO 피드백 정보를 수신하고, 액션 프레임의 전송하는 위상 배열 안테나;

메모리;

상기 메모리와 전기적으로 연결되는 적어도 하나의 프로세서를 포함하며,

상기 프로세서는

미리 학습된 강화 학습 모델을 이용하여, 상기 액션 프레임의 전송에 이용되는 현재 송신 섹터 집합을
결정하며,

상기 강화 학습 모델의 상태 집합은

상기 SISO 피드백 정보 및 이전 MIMO 단계에서 액션 프레임의 전송에 이용된 이전 송신 섹터 집합을 포함하며,

상기 강화 학습 모델의 액션 집합은

현재 MIMO 단계에 대한 상기 단말의 참여 정보 및 상기 현재 송신 섹터 집합에 대한 업데이트 정보를 포함하며,

상기 현재 및 이전 송신 섹터 집합은

위상 배열 안테나 각각에 대해 할당된 송신 섹터 중에서, 상기 액션 프레임의 전송에 이용되는 송신 섹터 조합
을 포함하며,

상기 업데이트 정보는

BF 셋업 하위 단계 및 BF 선택 하위 단계를 위한 업데이트 정보이며,

상기 이전 송신 섹터 집합에, 새로운 송신 섹터 조합을 추가하는 제1정보;

상기 이전 송신 섹터 집합에 포함된 송신 섹터 조합 중에서, 적어도 하나의 송신 섹터 조합을 제거하는
제2정보;

상기 이전 송신 섹터 집합에 포함된 송신 섹터 조합 중에서, 적어도 하나를 새로운 송신 섹터 조합으로 변경하
는 제3정보; 및

상기 이전 송신 섹터 집합을 상기 현재 송신 섹터 집합으로 이용하는 제4정보 중 적어도 하나를 포함하는 액세스 포인트.

발명의 설명

기술 분야

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

배경 기술

[0003] IEEE 802.11 표준에 따르면, MU-MIMO 빔포밍 훈련은 SISO 단계와 MIMO 단계로 수행된다.

[0004] SISO 단계에서 액세스 포인트는 송신 섹터별로 짧은 섹터 스윕 프레임(short sector sweep frame)을 전송하고, 단말로부터 SISO 피드백 정보를 수신한다. SISO 피드백 정보는 송신 섹터별로 전송된 짧은 섹터 스윕 프레임에 대해서, 단말이 측정한 SNR값을 포함한다.

[0005] 그리고 MIMO 단계는 세부적으로 BF 셋업(setup) 하위 단계, BF 선택(selection) 하위 단계, BF 훈련(training) 하위 단계 및 BF 피드백(feedback) 하위 단계로 이루어진다. 각 하위 단계에서 액세스 포인트는 빔포밍 훈련을 위한 액션 프레임을 단말로 전송한다.

[0006] 관련 선행문헌으로 대한민국 등록특허 제10-2288198호, 제10-2154481호, 제10-2153923호, 대한민국 공개특허 제2022-0036494호가 있다.

발명의 내용

해결하려는 과제

[0008] 본 발명은, 빔포밍 훈련을 위한 액션 프레임이 효율적으로 전송될 수 있도록, 송신 섹터를 위상 배열 안테나에 할당하는 빔포밍 훈련 방법을 제공하기 위한 것이다.

과제의 해결 수단

[0010] 상기한 목적을 달성하기 위한 본 발명의 일 실시예에 따르면, 단말로부터 SISO 피드백 정보를 수신하는 단계; 및 미리 학습된 강화 학습 모델을 이용하여, 액션 프레임의 전송에 이용되는 현재 송신 섹터 집합을 결정하는 단계를 포함하며, 상기 강화 학습 모델의 상태 집합은, 상기 SISO 피드백 정보 및 이전 MIMO 단계에서 액션 프레임의 전송에 이용된 이전 송신 섹터 집합을 포함하며, 상기 강화 학습 모델의 액션 집합은, 현재 MIMO 단계에 대한 상기 단말의 참여 정보 및 상기 현재 송신 섹터 집합에 대한 업데이트 정보를 포함하는 강화 학습 기반의 빔포밍 훈련 방법이 제공된다.

[0011] 또한 상기한 목적을 달성하기 위한 본 발명의 다른 실시예에 따르면, 단말로부터 SISO 피드백 정보를 수신하고, 액션 프레임을 전송하는 위상 배열 안테나; 메모리; 상기 메모리와 전기적으로 연결되는 적어도 하나의 프로세서를 포함하며, 상기 프로세서는 미리 학습된 강화 학습 모델을 이용하여, 상기 액션 프레임의 전송에 이용되는 현재 송신 섹터 집합을 결정하며, 상기 강화 학습 모델의 상태 집합은 상기 SISO 피드백 정보 및 이전 MIMO 단계에서 액션 프레임의 전송에 이용된 이전 송신 섹터 집합을 포함하며, 상기 강화 학습 모델의 액션 집합은 현재 MIMO 단계에 대한 상기 단말의 참여 정보 및 상기 현재 송신 섹터 집합에 대한 업데이트 정보를 포함하는 액세스 포인트가 제공된다.

발명의 효과

[0013] 본 발명의 일 실시예에 따르면, 강화 학습 모델을 이용함으로써, 빔포밍 훈련을 위한 액션 프레임의 불필요한 전송이 감소하고, 빔포밍 훈련 시간이 줄어들 수 있다.

[0014] 또한 본 발명의 일 실시예에 따르면, 가능한 많은 단말들이 빔포밍 훈련에 참여할 수 있다.

도면의 간단한 설명

- [0016] 도 1은 일반적인 빔포밍 훈련 방법을 설명하기 위한 도면이다.
- 도 2는 본 발명의 일실시예에 따른 무선랜 시스템에서 빔포밍 훈련 방법을 설명하기 위한 도면이다.
- 도 3은 본 발명의 일실시예에 따른 무선랜 시스템에서 빔포밍 훈련을 위한 SNR 예측 방법을 설명하기 위한 도면이다.
- 도 4는 본 발명의 일실시예에 따른 인공 신경망을 설명하기 위한 도면이다.

발명을 실시하기 위한 구체적인 내용

- [0017] 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.
- [0018] 이하에서, 본 발명에 따른 실시예들을 첨부된 도면을 참조하여 상세하게 설명한다.
- [0020] 도 1은 무선랜 시스템에서의 일반적인 빔포밍 훈련 방법을 설명하기 위한 도면으로서, 도 1에는 송신 섹터가 4개(TS1 내지 TS4)이고, 액세스 포인트의 위상 배열 안테나가 2개이며, 제1위상 배열 안테나(110)에는 제1 및 제2송신 섹터(TS1, TS2)가 할당되고, 제2위상 배열 안테나(120)에는 제3 및 제4송신 섹터(TS3, TS4)가 할당된 예시가 도시된다.
- [0021] 액세스 포인트는 SISO 단계에서 각 송신 섹터 별로 짧은 섹터 프레임의 전송한다. 제1위상 배열 안테나(110)는 제1 및 제2송신 섹터(TS1, TS2)로 짧은 섹터 프레임의 전송하고, 제2위상 배열 안테나(120)는 제3 및 제4송신 섹터(TS3, TS4)로 짧은 섹터 프레임의 전송한다. 액세스 포인트에 접속된 단말은 짧은 섹터 프레임의 수신하여 SNR을 측정하고, 측정된 SNR값을 포함하는 SISO 피드백 정보를 액세스 포인트로 전송한다. 즉 하나의 무선 단말은, 제1 내지 제4송신 섹터(TS1 내지 TS4)에 대해 측정된 4개의 SNR값을 포함하는 SISO 피드백 정보를 액세스 포인트로 전송한다. 예컨대 무선 단말은 제1송신 섹터(TS1)로 전송되는 짧은 섹터 스윙 프레임의 수신하여 제1송신 섹터(TS1)에 대한 SNR값을 생성하며, 제2송신 섹터(TS2)로 전송되는 짧은 섹터 스윙 프레임의 수신하여 제2송신 섹터(TS2)에 대한 SNR값을 생성한다.
- [0022] 액세스 포인트는 측정된 SNR값을 이용하여 MIMO 단계에서 빔포밍 훈련을 수행한다. 액세스 포인트는 BF 셋업 하위 단계, BF 선택 하위 단계, BF 훈련 하위 단계 및 BF 피드백 하위 단계 각각에서 액션 프레임을 전송하여, 빔포밍 훈련을 수행한다.
- [0023] 액세스 포인트는, MIMO 단계에서 액세스 포인트에 접속한 단말이 모두 수신할 수 있도록 액션 프레임을 전송하며, 이를 위해 복수의 송신 섹터의 일부 송신 섹터를 포함하는 송신 섹터 조합 중 하나를 선택하여, 액션 프레임을 전송할 수 있다. 송신 섹터 조합은, 위상 배열 안테나 각각에 할당된 송신 섹터 중에서 하나씩 선택되어 결정될 수 있다. 도 1과 같은 예시에서, 송신 섹터 조합은 (TS1, TS3), (TS1, TS4), (TS2, TS3) (TS2, TS4)와 같이, 4개로 결정될 수 있으며, 액세스 포인트는 이러한 송신 섹터 조합 중에 하나를 선택하여, 선택된 송신 섹터 조합에 포함된 복수의 송신 섹터로 액션 프레임을 전송할 수 있다.
- [0024] 모든 송신 섹터 조합으로 액션 프레임을 전송하는 것은 빔포밍 훈련 시간의 증가를 초래하며 비효율적이므로, 모든 단말이 액션 프레임을 수신하면서도 액션 프레임의 전송 횟수가 최소가 될 수 있는 효율적인 송신 섹터 조합을 결정하는 것이 필요하다. 이에 본 발명은 강화 학습을 이용하여 최적의 송신 섹터 조합을 결정하고, 이를 통해 빔포밍 훈련을 수행하는 방법을 제안한다.
- [0025] 본 발명의 일실시예에 따른 빔포밍 훈련 방법은 무선랜 시스템의 액세스 포인트에서 수행된다.
- [0027] 도 2는 본 발명의 일실시예에 따른 무선랜 시스템에서 빔포밍 훈련 방법을 설명하기 위한 도면이며, 도 3은 본 발명의 일실시예에 따른 강화 학습 모델을 도시하는 도면이다.
- [0028] 도 2를 참조하면 본 발명의 일실시예에 따른 액세스 포인트는 단말로부터 SISO 피드백 정보를 수신(S210)하고, 미리 학습된 강화 학습 모델을 이용하여, 액션 프레임의 전송에 이용되는 현재 송신 섹터 집합을 결정(S220)한다. 여기서, 송신 섹터 집합이란, 위상 배열 안테나 각각에 대해 할당된 송신 섹터 중에서, 액션 프레임의 전송

에 이용되는 송신 섹터 조합을 포함하는 집합을 의미한다.

[0029] 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 환경의 상태(state)를 관찰(observation)하며, 관찰된 상태에서 보상(reward)이 최대가 되는 행동(action)을 수행한다. 여기서 행동은 현재 송신 섹터 집합을 결정하는 것에 대응되며, 액터와 크리틱은, TD(Temporal Difference) 에러를 통해 업데이트될 수 있다.

[0030] 강화 학습 모델의 상태 집합은, SISO 피드백 정보 및 이전 MIMO 단계에서 액션 프레임의 전송에 이용된 이전 송신 섹터 집합을 포함하며, 강화 학습 모델의 액션 집합은, 현재 MIMO 단계에 대한 단말의 참여 정보 및 현재 송신 섹터 집합에 대한 업데이트 정보를 포함한다. 여기서, 이전 송신 섹터 집합은 단계 S220의 현재 송신 섹터 집합 결정 전에, 액션 프레임의 전송에 이용된 송신 섹터 집합을 의미한다. 그리고 참여 정보는 SISO 피드백 정보에 포함된 SNR값에 따라 결정되는 정보로서, 미리 설정된 임계값 이상의 SNR값을 전송한 단말이 MIMO 단계에 참여한다.

[0031] 이하에서는 MIMO 단계를 구성하는, BF 셋업 하위 단계, BF 선택 하위 단계, BF 훈련 하위 단계 및 BF 피드백 하위 단계 별로, 현재 송신 섹터 집합을 결정하는 방법을 자세히 설명하기로 한다.

[0033] **BF 셋업 하위 단계, BF 선택 하위 단계**

[0034] BF 셋업 하위 단계 및 BF 선택 하위 단계에서 액세스 포인트는, 전송된 강화 학습 모델을 이용하여 현재 송신 섹터 집합을 결정한다.

[0035] 빔포밍 훈련에 참여하는 단말의 개수를 N 이라고 하고, 상태 집합에 포함된 이전 송신 섹터 집합과 SISO 피드백 정보를 각각 T_N 및 F_N 으로 표현할 때, 이전 송신 섹터 집합은 [수학식 1]과 같이 표현될 수 있다.

수학식 1

[0036]
$$T_N = \{t_1, \dots, t_i, \dots, t_N\}$$

[0037] 여기서, $t_i (1 \leq i \leq N)$ 는 i 번째 액션 프레임을 전송할 때 이용된 송신 섹터 조합을 나타낸다. 빔포밍 훈련에 참여하는 단말의 개수가 N 개이기 때문에, 액션 프레임은 최대 N 회 전송되며, 송신 섹터 조합의 최대 개수는 N 개이다.

[0038] 그리고 위상 배열 안테나의 개수를 M 개라고 할 때, i 번째 액션 프레임의 전송에 이용되는 송신 섹터 조합은 [수학식 2]와 같이 표현될 수 있고, $TS_j (1 \leq j \leq M)$ 는 j 번째 위상 배열 안테나에 할당된 송신 섹터를 나타낸다.

수학식 2

[0039]
$$t_i = (TS_1, \dots, TS_j, \dots, TS_M)$$

[0040] SISO 피드백 정보는 [수학식 3]과 같이 표현될 수 있으며, 여기서 $R_{fb}^{STAk} (1 \leq k \leq N)$ 는, k 번째 단말로부터 수신된 SISO 피드백 정보를 나타낸다.

수학식 3

$$F_N = \{R_{fb}^{STA1}, \dots, R_{fb}^{STAk}, \dots, R_{fb}^{STAN}\}$$

그리고 R_{fb}^{STAk} 는 [수학식 4]와 같이 표현될 수 있으며, $r_{TSl}(1 \leq l \leq L)$ 은 송신 섹터(TS)의 총 개수를 L이라고 할 때, SISO 단계에서 1번째 송신 섹터로 전송된 짧은 섹터 스왑 프레임에 대한 SNR값을 나타낸다. 즉, 단말은 모든 송신 섹터 각각에 대한 SNR값을 포함하는 SISO 피드백 정보를 액세스 포인트로 전송한다.

수학식 4

$$R_{fb}^{STAk} = \{r_{TS1}, \dots, r_{TSL}, \dots, r_{TSL}\}$$

행동 집합에 포함된 참여 정보와 업데이트 정보를 각각 P_N , e라고 할 때, 참여 정보는 [수학식 5]와 같이 표현될 수 있고, k 번째 단말이 현재 MIMO 단계에 참여하는 경우는 p_k 는 1, 그렇지 않은 경우 p_k 는 0으로 설정될 수 있다.

수학식 5

$$P_N = \{p_1, \dots, p_k, \dots, p_N\}$$

단말이 전송한 SNR값 중에서 미리 설정된 제1임계값 이상의 SNR값이 1개 이상이라면, 해당 단말은 현재 MIMO 단계에 참여하는 것으로 결정된다.

업데이트 정보는 일실시예로서, 제1 내지 제4정보를 포함할 수 있으며, 액세스 포인트는 업데이트 정보에 따라서 현재 송신 섹터 집합을 결정한다. 먼저, 제1정보(e=1)는 이전 송신 섹터 집합에, 적어도 하나의 새로운 송신 섹터 조합을 추가하는 정보이다. 따라서 이 경우, 액션 프레임의 전송 횟수가 증가한다. 이 때, 액세스 포인트는, 이전 송신 섹터 집합에 포함된 송신 섹터 조합을 제외한 송신 섹터 조합 중에서, 액션 프레임을 추가로 전송할 적어도 하나의 송신 섹터 조합을 이전 송신 섹터 집합에 추가하여, 현재 송신 섹터 집합을 결정한다.

액세스 포인트는 이전 빔포밍 훈련의 BF 셋업 하위 단계, BF 선택 하위 단계에서 액션 프레임을 수신하지 못한 단말들이 포함된 송신 섹터들중에서, 액션 프레임을 추가로 전송할 송신 섹터 조합을 결정하여 이전 송신 섹터 집합에 추가한다. 단말의 액션 프레임의 수신 여부는, 단말이 전송하는 ACK 정보로 확인할 수 있다. 이 때, 액세스 포인트는, 액션 프레임을 수신하지 못한 단말들이 전송한 SNR값 중 미리 설정된 제2임계값 이상인 SNR값이 최대가 되도록, 액션 프레임을 추가로 전송할 송신 섹터 조합을 결정한다.

다시 말해 액세스 포인트는 현재 MIMO 단계에 참여하고, 이전 MIMO 단계에서 액션 프레임을 수신하지 못한 단말 중에서, SISO 피드백 정보의 SNR값이 제2임계값 이상인 단말의 개수가 최대인, 송신 섹터 조합을 이전 송신 섹터 집합에 추가하여, 현재 송신 섹터 집합을 결정한다.

만일 제2임계값 이상의 SNR값이 존재하지 않는다면, 액세스 포인트는 최대 SNR값을 전송한 단말이 포함된 송신 섹터가 포함되도록, 액션 프레임을 추가로 전송할 송신 섹터 조합을 결정한다. 즉, 액세스 포인트는 현재 MIMO 단계에 참여하고, 이전 MIMO 단계에서 액션 프레임을 수신하지 못한 단말 중에서, 제2임계값 이상인 단말이 존재하지 않는 경우, 최대 SNR값에 대응되는 단말이 포함된 송신 섹터 조합을 이전 송신 섹터 집합에 추가하여, 현재 송신 섹터 집합을 결정한다.

제2정보(e=2)는 이전 송신 섹터 집합에 포함된 송신 섹터 조합 중에서, 적어도 하나의 송신 섹터 조합을 제거하

는 정보이다. 따라서 이 경우 액션 프레임의 전송 횟수가 감소한다. 액세스 포인트는, 이전 송신 섹터 집합에 포함된 송신 섹터 조합 각각으로 액션 프레임이 전송됐을 때, 액션 프레임을 수신한 단말의 개수가 최소인 송신 섹터 조합을 이전 송신 섹터 집합에서 제거할 수 있다.

[0052] 제3정보(e=3)는 이전 송신 섹터 집합에 포함된 송신 섹터 조합 중에서, 적어도 하나를 새로운 송신 섹터 조합으로 변경하는 정보이다. 따라서, 이 경우 이전 송신 섹터 집합에 포함된 송신 섹터 조합의 개수와, 현재 송신 섹터 집합에 포함된 송신 섹터 조합의 개수는 동일하다. 액세스 포인트는 업데이트 정보가 제2정보인 경우의 제거될 송신 섹터 조합을 결정하는 방법을 이용해, 송신 섹터 조합을 제거하고, 업데이트 정보가 제1정보인 경우의 추가될 송신 섹터 조합을 결정하는 방법을 이용해, 송신 섹터 조합을 추가한다.

[0053] 제4정보(e=4)는 이전 송신 섹터 집합을 현재 송신 섹터 집합으로 이용하는 정보이다. 즉, 이 경우 액세스 포인트는 이전 송신 섹터 집합을 현재 송신 섹터 집합으로 결정하고, 액션 프레임을 전송한다.

[0054] 액세스 포인트는 전송된 바와 같이 보상이 최대가 될 수 있도록 업데이트 정보 중 하나를 결정하고, 결정된 업데이트 정보에 따라서 현재 송신 섹터 집합을 결정할 수 있다. 그리고 이 때, 보상(Γ)은 [수학식 6]과 같이 결정될 수 있다.

수학식 6

$$\Gamma = \frac{N_S + (N - N_t)}{N}$$

[0055]

[0056] 여기서, N_S 는 BF 셋업 하위 단계에서 적어도 하나의 액션 프레임을 수신한 단말의 개수를 나타내며, N_t 는 BF 셋업 하위 단계에서의 액션 프레임의 전송 횟수를 나타낸다.

[0058] BF 훈련 하위 단계

[0059] BF 훈련 하위 단계에서 액세스 포인트는 현재 송신 섹터 집합의 송신 섹터 조합에 포함된 단말 중에서, 액션 프레임의 수신이 가능한 단말을 확인하여, 송신 섹터 조합을 결정한다.

[0060] 액세스 포인트는 이전 송신 섹터 집합의 송신 섹터 조합에 포함된 단말은 액션 프레임의 수신이 가능한 것으로 판단한다. 그리고 현재 송신 섹터 집합의 송신 섹터 조합 중 새로 추가된 송신 섹터 조합에 포함된 단말에 대해, 새로 추가된 송신 섹터 조합에 포함된 단말에서 전송된 SNR값 중 적어도 하나가 미리 설정된 제2임계값인 경우 액션 프레임의 수신이 가능한 것으로 판단한다.

[0061] 액세스 포인트는 액션 프레임의 수신이 가능한 단말 중에서, BF 훈련 하위 단계에서의 빔포밍 훈련이 필요한 단말이 포함된 송신 섹터 조합이 현재 송신 섹터 집합에 포함되도록, 현재 송신 섹터 집합을 갱신할 수 있다.

[0063] BF 피드백 하위 단계

[0064] BF 피드백 하위 단계에서 액세스 포인트는, 결정된 현재 송신 섹터 집합과 무관하게, 최대 SNR값을 전송한 단말이 포함된 송신 섹터로 액션 프레임을 전송한다.

[0066] 도 4는 본 발명의 일실시예에 따른 강화 학습 모델의 업데이트 방법을 설명하기 위한 도면이다.

[0067] 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 현재 송신 섹터 집합을 결정하면서, 수집된 데이터를 이용하여 강화 학습 모델에 대한 학습을 수행하여, 딥러닝 네트워크를 업데이트한다.

[0068] 본 발명의 일실시예에 따른 강화 학습 모델에서 에이전트의 액터는 정책 함수를 이용하여 행동, 즉 현재 송신 섹터 집합의 결정을 수행한다. 그리고 에이전트의 크리틱(비평가)는 가치 함수를 이용하여 액터의 행동, 즉 현재 송신 섹터 집합의 결정이 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.

[0069] 도 4를 참조하면, 액세스 포인트는 정책 함수 및 가치 함수에 대한 딥러닝 네트워크를 이용하여 정책 함수 및 가치 함수의 출력값을 획득(S410)한다.

[0070] 정책 함수($\pi^{\theta}(S, A)$)는 특정 상태(S)에서 특정 행동(A)이 결정될 확률을 계산하기 위한 함수로서, 여기서 θ 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($N(A|\mu(S), \sigma(S))$)를 생성하여, 특정 행동을 결정할 확률을 계산할 수 있다. 예컨대, 전송된 업데이트 정보 중 가장 높은 확률에 대응되는 정보에 따라서, 액세스 포인트는, 현재 송신 섹터 집합을 결정할 수 있다.

[0071] 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.

[0072] 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(S420)한다.

[0073] 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

Γ_t 라고 할 때, $V^v(S_t)$ 의 실제 값에 대한 추정치 TD 타겟(Temporal Difference target, y_t)은 [수학식 7]을 통해 계산될 수 있다.

수학식 7

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

[0074] 여기서, γ 은 감가율(Discount Factor)을 나타낸다.

[0076] 가치 함수 딥러닝 네트워크는 [수학식 8]과 같이 계산되는 TD 에러값(ϵ_t)이 최소가 되도록 업데이트될 수 있다.

수학식 8

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

[0078] 그리고 정책 함수 딥러닝 네트워크는 [수학식 9]와 같이 표현되는 손실 함수의 손실값이 최소가 되도록 업데이트될 수 있다.

수학식 9

$$Loss = -\log(N(A|\mu(S_t), \sigma(S_t))) \cdot \varepsilon_t$$

[0079]

[0081]

도 5는 본 발명의 일실시예에 따른 강화 학습 기반의 빔포밍 훈련 장치를 설명하기 위한 도면으로서, 도 5에서 는 액세스 포인트가 일실시예로서 도시된다.

[0082]

도 5를 참조하면, 본 발명의 일실시예에 따른 액세스 포인트는 위상 배열 안테나(510), 메모리(520) 및 적어도 하나의 프로세서(530)를 포함한다.

[0083]

위상 배열 안테나(510)는 단말로부터 SISO 피드백 정보를 수신하고, 액션 프레임의 전송한다.

[0084]

프로세서(530)는 메모리(520)와 전기적으로 연결되며, 미리 학습된 강화 학습 모델을 이용하여, 액션 프레임의 전송에 이용되는 현재 송신 섹터 집합을 결정한다. 강화 학습 모델은 전송된 강화 학습 모델로서, 강화 학습 모델의 상태 집합은 SISO 피드백 정보 및 이전 MIMO 단계에서 액션 프레임의 전송에 이용된 이전 송신 섹터 집합을 포함하며, 강화 학습 모델의 액션 집합은 현재 MIMO 단계에 대한 상기 단말의 참여 정보 및 상기 현재 송신 섹터 집합에 대한 업데이트 정보를 포함한다.

[0086]

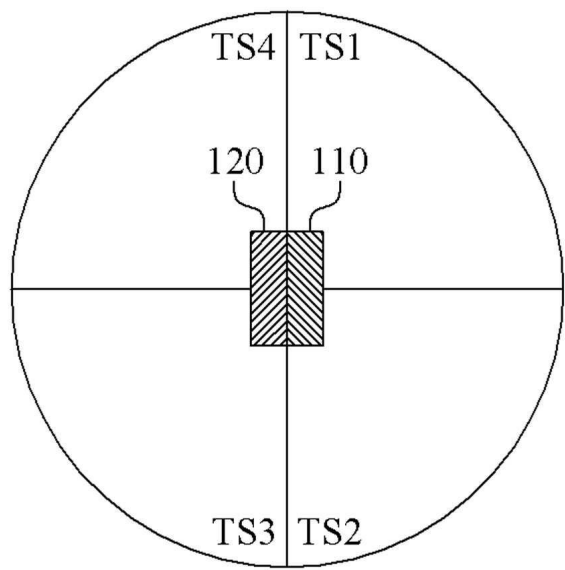
앞서 설명한 기술적 내용들은 다양한 컴퓨터 수단을 통하여 수행될 수 있는 프로그램 명령 형태로 구현되어 컴퓨터 판독 가능 매체에 기록될 수 있다. 상기 컴퓨터 판독 가능 매체는 프로그램 명령, 데이터 파일, 데이터 구조 등을 단독으로 또는 조합하여 포함할 수 있다. 상기 매체에 기록되는 프로그램 명령은 실시예들을 위하여 특별히 설계되고 구성된 것들이거나 컴퓨터 소프트웨어 당업자에게 공지되어 사용 가능한 것일 수도 있다. 컴퓨터 판독 가능 기록 매체의 예에는 하드 디스크, 플로피 디스크 및 자기 테이프와 같은 자기 매체(magnetic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

[0088]

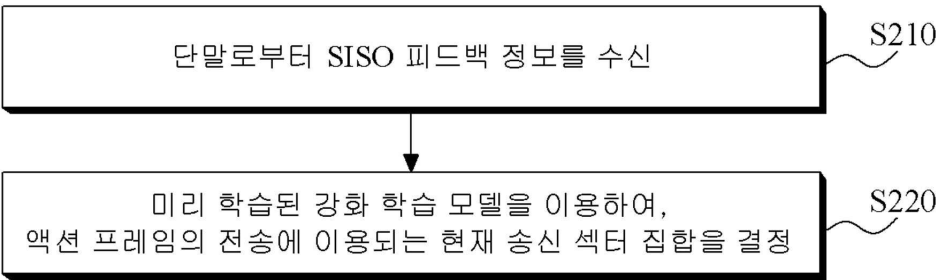
이상과 같이 본 발명에서는 구체적인 구성 요소 등과 같은 특정 사항들과 한정된 실시예 및 도면에 의해 설명되었으나 이는 본 발명의 보다 전반적인 이해를 돕기 위해서 제공된 것일 뿐, 본 발명은 상기의 실시예에 한정되는 것은 아니며, 본 발명이 속하는 분야에서 통상적인 지식을 가진 자라면 이러한 기재로부터 다양한 수정 및 변형이 가능하다. 따라서, 본 발명의 사상은 설명된 실시예에 국한되어 정해져서는 아니되며, 후술하는 특허청구범위뿐 아니라 이 특허청구범위와 균등하거나 등가적 변형이 있는 모든 것들은 본 발명 사상의 범주에 속한다고 할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

도면

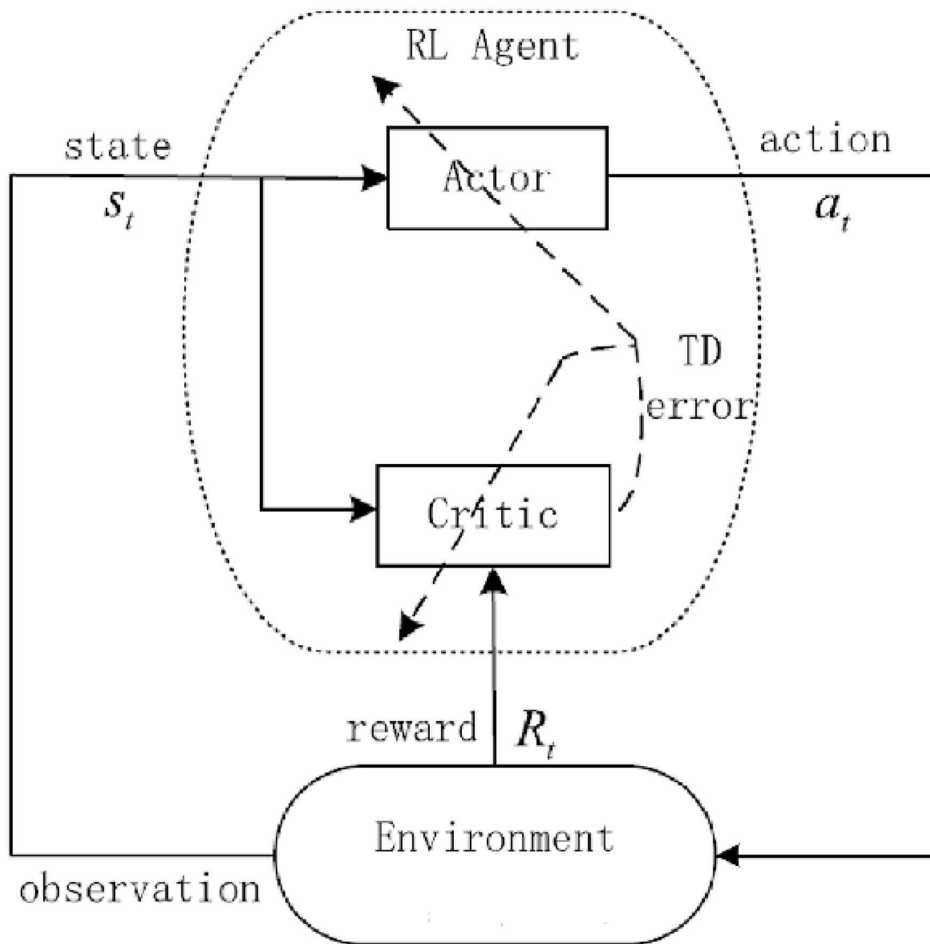
도면1



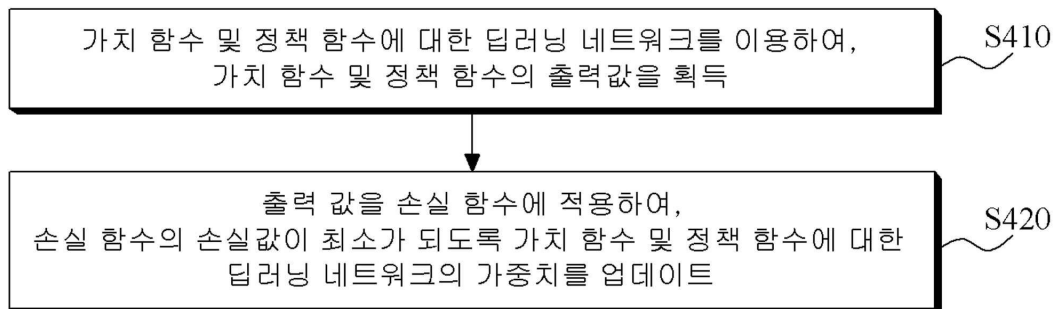
도면2



도면3



도면4



도면5

