



(19) 대한민국특허청(KR)
(12) 공개특허공보(A)

(11) 공개번호 10-2024-0071265
(43) 공개일자 2024년05월22일

(51) 국제특허분류(Int. Cl.)
H04W 24/02 (2009.01) G06N 20/00 (2019.01)
H04B 7/145 (2006.01) H04B 7/185 (2006.01)
H04B 7/195 (2006.01) H04W 84/06 (2019.01)
(52) CPC특허분류
H04W 24/02 (2013.01)
G06N 20/00 (2021.08)
(21) 출원번호 10-2023-0029490
(22) 출원일자 2023년03월06일
심사청구일자 2023년03월06일
(30) 우선권주장
1020220152733 2022년11월15일 대한민국(KR)

(71) 출원인
경희대학교 산학협력단
경기도 용인시 기흥구 덕영대로 1732 (서천동, 경희대학교 국제캠퍼스내)
(72) 발명자
홍충선
경기도 용인시 수지구 상현로 30-10, 4813동 101호(상현동, 광고상떼빌파크뷰)
하산 셰이크 살만
경기도 용인시 기흥구 덕영대로 1732, 경희대학교 전자정보대학 352호
(뒷면에 계속)
(74) 대리인
두호특허법인

전체 청구항 수 : 총 20 항

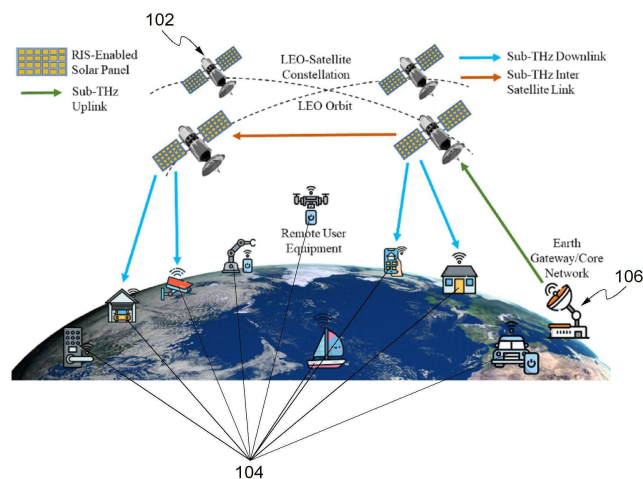
(54) 발명의 명칭 저궤도 위성 기반의 통신 방법 및 이를 수행하기 위한 컴퓨팅 장치

(57) 요약

저궤도 위성의 통신 방법 및 이를 수행하기 위한 컴퓨팅 장치가 개시된다. 개시되는 일 실시예에 따른 통신 방법은, RIS(Reconfigurable Intelligent Surface) 패넬을 구비하는 복수 개의 위성, 복수 개의 원격 사용자 단말, 및 하나 이상의 기지국 장치를 포함하는 통신 시스템의 통신 방법이고, 하나 이상의 프로세서들, 및 상기 하나 이상의 프로세서들에 의해 실행되는 하나 이상의 프로그램들을 저장하는 메모리를 구비한 컴퓨팅 장치에서 수행되는 통신 방법으로서, 위성에서 송신하는 신호를 원격 사용자 단말이 수신하는 경우, 원격 사용자 단말의 수신 신호를 산출하는 단계, 수신 신호에 대해 해당 원격 사용자 단말의 SNR(Signal to Noise Ratio)을 이용하여 원격 사용자 단말의 데이터 속도를 산출하는 단계, 및 원격 사용자 단말의 데이터 속도를 최대화 하도록 기 설정된 파라미터들을 최적화 하는 단계를 포함한다.

대표도

100



(52) CPC특허분류

H04B 7/145 (2013.01)
H04B 7/18532 (2013.01)
H04B 7/195 (2013.01)
H04W 84/06 (2019.01)

박유민

경기도 수원시 영통구 인계로 284, 401호 (매탄동, 매탄엠플타워)

(72) 발명자

박성배

서울특별시 송파구 오금로 396, 1402호(가락동, 에스케이허브파크)

이 발명을 지원한 국가연구개발사업

과제고유번호	1711194179
과제번호	207816
부처명	과학기술정보통신부
과제관리(전문)기관명	한국연구재단
연구사업명	집단연구지원
연구과제명	Meta Federated Learning 기반 Satellite-Air-Ground 통합 네트워킹 시스템 핵심 구

조 개발

기 여 율	1/3
과제수행기관명	경희대학교(국제캠퍼스)
연구기간	2023.03.01 ~ 2024.02.29

이 발명을 지원한 국가연구개발사업

과제고유번호	1711193491
과제번호	2019-0-01287-005
부처명	과학기술정보통신부
과제관리(전문)기관명	정보통신기획평가원
연구사업명	SW컴퓨팅산업원천기술개발
연구과제명	(스타랩) 분산 엣지를 위한 진화형 딥러닝 모델생성 플랫폼_2단계
기 여 율	1/3
과제수행기관명	경희대학교(국제캠퍼스)
연구기간	2023.01.01 ~ 2023.12.31

이 발명을 지원한 국가연구개발사업

과제고유번호	1711193622
과제번호	2021-0-02068-003
부처명	과학기술정보통신부
과제관리(전문)기관명	정보통신기획평가원
연구사업명	정보통신방송혁신인재양성
연구과제명	인공지능 혁신 허브 연구 개발
기 여 율	1/3
과제수행기관명	고려대학교산학협력단
연구기간	2023.01.01 ~ 2023.12.31

명세서

청구범위

청구항 1

RIS(Reconfigurable Intelligent Surface) 패널을 구비하는 복수 개의 위성, 복수 개의 원격 사용자 단말, 및 하나 이상의 기지국 장치를 포함하는 통신 시스템의 통신 방법이고,

하나 이상의 프로세서들, 및 상기 하나 이상의 프로세서들에 의해 실행되는 하나 이상의 프로그램들을 저장하는 메모리를 구비한 컴퓨팅 장치에서 수행되는 통신 방법으로서,

위성에서 송신하는 신호를 원격 사용자 단말이 수신하는 경우, 원격 사용자 단말의 수신 신호를 산출하는 단계;

상기 수신 신호에 대해 해당 원격 사용자 단말의 SNR(Signal to Noise Ratio)을 이용하여 상기 원격 사용자 단말의 데이터 속도를 산출하는 단계; 및

상기 원격 사용자 단말의 데이터 속도를 최대화 하도록 기 설정된 파라미터들을 최적화 하는 단계를 포함하는, 통신 방법.

청구항 2

청구항 1에 있어서,

상기 수신 신호(y_u)는 하기 수학식을 통해 산출되는, 통신 방법.

(수학식)

$$y_u = \sqrt{\frac{p_u}{L^{\text{tot}}}} \mathbf{g}_u^T \prod_{r=1, \dots, R_s} \Phi_r \mathbf{H}_r \mathbf{x} + z_u$$

L^{tot} : 신호가 원격 사용자 단말에 수신되기까지의 총 손실

p_u : 기지국 장치의 송신 전력

$\mathbf{g}_u$: 마지막 RIS 패널에서 원격 사용자 단말까지의 채널

$\mathbf{x}$: 기지국 장치의 전송 벡터

R_s : RIS 패널의 전체 개수

z_u : 원격 사용자 단말에서 수신된 AWGN(Additive White Gaussian Noise)

r : RIS의 인덱스

$\mathbf{H}_r$: 기지국 장치와 RIS 패널 r 사이의 채널 행렬

Φ_r : RIS 패널 r 의 위상 편이

청구항 3

청구항 2에 있어서,

상기 원격 사용자 단말의 시간 슬롯 t 에서 SNR(Signal to Noise Ratio)($\Gamma_{u,t}$)은 하기 수학식으로 산출하는,

통신 방법.

(수학식)

$$\Gamma_{u,t} = \frac{p_{s,u} \left| v_{b,u}^s \mathbf{g}_u^T(t) \prod_{r=1,\dots,R_s} \Phi_r \mathbf{H}_r(t) \mathbf{w}_u \right|^2}{N_o L^{\text{tot}}(f, d)}$$

$p_{s,u}$: 위성에서 원격 사용자 단말로 신호를 송신하기 위한 전력

$v_{b,u}^s$: 원격 사용자 단말에 대한 연관 행렬의 요소

N_o : 노이즈 스펙트럼 밀도

$\mathbf{w}_u$: 기지국 장치의 빔포밍 벡터

청구항 4

청구항 3에 있어서,

상기 원격 사용자 단말의 데이터 속도($R_{u,t}$)는 하기 수학식으로 산출하는, 통신 방법.

(수학식)

$$R_{u,t} = B_u \log_2(1 + \Gamma_{u,t})$$

B_u : 원격 사용자 단말에서 사용 가능한 총 대역폭

청구항 5

청구항 1에 있어서,

기 설정된 파라미터들을 최적화 하는 단계는,

하기 수학식의 목적 함수를 만족하도록 위성에서 원격 사용자 단말로 신호를 송신하기 위한 전력(p), 위성의 RIS 패널과 원격 사용자 단말 간의 연관 행렬($\mathbf{V}$), 위성의 RIS 패널의 위상 편이(Φ), 및 위성의 양각(α)을 최적화시키는, 통신 방법.

(수학식)

$$\max_{p, \mathbf{V}, \Phi, \alpha} \sum_{u=1}^U \sum_{t=1}^T R_{u,t}$$

$R_{u,t}$: 시간 슬롯 t 에서 원격 사용자 단말 u 의 데이터 속도

T : 시간 슬롯의 전체 개수

U : 원격 사용자 단말의 전체 개수

청구항 6

청구항 5에 있어서,

기 설정된 파라미터들을 최적화 하는 단계는,

상기 RIS 패널과 원격 사용자 단말 간의 연관 행렬($\mathbf{V}$), RIS 패널의 위상 편이(Φ), 및 위성의 양각(α)을 초기화 한 후 기 설정된 제1 알고리즘을 이용하여 상기 전력(p)을 최적화 하는 제1 최적화 단계; 및

상기 최적화된 전력(p)에 기초하여 기 설정된 제2 알고리즘을 통해 상기 RIS 패널과 원격 사용자 단말 간의 연관 행렬($\mathbf{V}$), RIS 패널의 위상 편이(Φ), 및 위성의 양각(α)을 최적화 하는 제2 최적화 단계를 포함하는, 통신 방법.

청구항 7

청구항 6에 있어서,

상기 제1 알고리즘은, WOA(Whale Optimization Algorithm)이고,

상기 제1 최적화 단계는,

하기 수학식에 의한 적합도 함수를 통해 상기 전력(p)의 최적의 값을 산출하는, 통신 방법.

(수학식)

$$\text{Fitness}(\Phi) = - \sum_{u=1}^U \sum_{t=1}^T \left\{ R_{u,t}(\Phi) + \mu F_{u,t}(f_{u,t}(\Phi)) f_{u,t}^2(\Phi) \right\}$$

$f_{u,t}(\Phi)$: 부등식 함수(inequality function)

μ : 기 설정된 상수

$F_{u,t}(f_{u,t}(\Phi))$: 인덱스 함수(index function)

상기 인덱스 함수는 상기 부등식 함수가 0이면 0이 되고, 상기 부등식 함수가 0미만이면 1의 값을 가짐.

청구항 8

청구항 6에 있어서,

상기 제2 알고리즘은, 멀티 에이전트 강화 학습(Multi Agent Reinforcement Learning) 모델을 이용한 알고리즘이고,

아래 수학식의 목적 함수를 만족하도록 상기 RIS 패널과 원격 사용자 단말 간의 연관 행렬($\mathbf{V}$), RIS 패널의 위상 편이(Φ), 및 위성의 양각(α)을 최적화 하는, 통신 방법.

(수학식)

$$\max_{\mathbf{V}, \Phi, \alpha} \sum_{u=1}^U \sum_{t=1}^T R_{u,t}$$

청구항 9

청구항 8에 있어서,

상기 제2 최적화 단계는,

하기 수학식에 따른 상기 복수 개 위성의 전역 상태(s') 및 각 위성의 로컬 상태(s'_s)를 상기 멀티 에이전트 강화 학습 모델로 입력하는 단계를 포함하는, 통신 방법.

(수학식)

$$s' = \{\{d_1\}, \{V_1\}, \{\Phi_1\}, \{p_1^*\}, \{i_1, \omega_1, \chi_1\}\}, \dots, \{d_s\}, \{V_s\}, \{\Phi_s\}, \{p_s^*\}, \{i_s, \omega_s, \chi_s\}\}$$

$$s'_s = \{\{d_s\}, \{V_s\}, \{\Phi_s\}, \{p_s^*\}, \{i_s\}, \{\omega_s\}, \{\chi_s\}\}$$

d_s : 위성 s 의 위치

V_s : 위성 s 의 RIS 패널과 사용자 단말 간의 연관 행렬

Φ_s : 위성 s 의 RIS 패널의 위상 편이

p_s^* : 위성 s 의 최적화된 전력

i_s : 위성 s 의 궤도면과 적도 사이의 교차각

ω_s : 위성 s 의 춘분점과 궤도면과 적도면의 교차점 사이의 각도

χ_s : 위성 s 의 위성의 방향과 궤도면과 적도면의 교차점이 이루는 각도

청구항 10

청구항 9에 있어서,

상기 제2 최적화 단계는,

상기 멀티 에이전트 강화 학습 모델에서 상기 각 위성의 로컬 상태에 기초하여 상기 RIS 패널과 원격 사용자 단말 간의 연관, 상기 RIS 패널의 위상 편이, 및 상기 위상의 양각에 대한 행동을 산출하는 단계; 및

상기 멀티 에이전트 강화 학습 모델에서 상기 산출된 행동에 대한 글로벌 보상을 상기 복수 개의 위성으로 부여하는 단계를 포함하는, 통신 방법.

청구항 11

청구항 10에 있어서,

상기 제2 최적화 단계는,

하기 수학식에 의한 상기 글로벌 보상($R_{s',t}$)이 최대화 되도록 상기 멀티 에이전트 강화 학습 모델을 학습하는 단계를 더 포함하는, 통신 방법.

(수학식)

$$R_{s',t} = R_t(s'_t, a_t) = \begin{cases} \sum_{u=1}^U \sum_{t=1}^T R_{u,t}, & \text{if all the associated RUEs is served by networks,} \\ -\Gamma^{\text{pen}}, & \text{otherwise impose penalty to perform better in learning.} \end{cases}$$

s'_t : 시간 슬롯 t 에서 위성의 로컬 상태

a_t : 시간 슬롯 t 에서 위성의 행동

RUE : 원격 사용자 단말

$-\Gamma^{\text{pen}}$: 페널티 값(penalty value)

청구항 12

하나 이상의 프로세서들; 메모리; 및 하나 이상의 프로그램들을 포함하고,

상기 하나 이상의 프로그램들은 상기 메모리에 저장되고, 상기 하나 이상의 프로세서들에 의해 실행되도록 구성되며,

RIS(Reconfigurable Intelligent Surface) 패널을 구비하는 복수 개의 위성, 복수 개의 원격 사용자 단말, 및 하나 이상의 기지국 장치를 포함하는 통신 시스템의 통신 방법을 수행하기 위한 컴퓨팅 장치로서,

상기 하나 이상의 프로그램들은,

위성에서 송신하는 신호를 원격 사용자 단말이 수신하는 경우, 원격 사용자 단말의 수신 신호를 산출하기 위한 명령;

상기 수신 신호에 대해 해당 원격 사용자 단말의 SNR(Signal to Noise Ratio)을 이용하여 상기 원격 사용자 단말의 데이터 속도를 산출하기 위한 명령; 및

상기 원격 사용자 단말의 데이터 속도를 최대화 하도록 기 설정된 파라미터들을 최적화 하기 위한 명령을 포함하는, 컴퓨팅 장치.

청구항 13

청구항 12에 있어서,

기 설정된 파라미터들을 최적화 하기 위한 명령은,

하기 수학식의 목적 함수를 만족하도록 위성에서 원격 사용자 단말로 신호를 송신하기 위한 전력(p), 위성의 RIS 패널과 원격 사용자 단말 간의 연관 행렬($\mathbf{V}$), 위성의 RIS 패널의 위상 편이(Φ), 및 위성의 양각(α)을 최적화시키는, 컴퓨팅 장치.

(수학식)

$$\max_{\mathbf{p}, \mathbf{V}, \Phi, \alpha} \sum_{u=1}^U \sum_{t=1}^T R_{u,t}$$

$R_{u,t}$: 시간 슬롯 t에서 원격 사용자 단말 u의 데이터 속도

T : 시간 슬롯의 전체 개수

U : 원격 사용자 단말의 전체 개수

청구항 14

청구항 13에 있어서,

기 설정된 파라미터들을 최적화 하기 위한 명령은,

상기 RIS 패널과 원격 사용자 단말 간의 연관 행렬($\mathbf{V}$), RIS 패널의 위상 편이(Φ), 및 위성의 양각(α)을 초기화 한 후 기 설정된 제1 알고리즘을 이용하여 상기 전력(p)을 최적화 하는 제1 최적화 명령; 및

상기 최적화된 전력(p)에 기초하여 기 설정된 제2 알고리즘을 통해 상기 RIS 패널과 원격 사용자 단말 간의 연관 행렬($\mathbf{V}$), RIS 패널의 위상 편이(Φ), 및 위성의 양각(α)을 최적화 하는 제2 최적화 명령을 포함하는, 컴퓨팅 장치.

청구항 15

청구항 14에 있어서,

상기 제1 알고리즘은, WOA(Whale Optimization Algorithm)이고,

상기 제1 최적화 명령은,

하기 수학적식에 의한 적합도 함수를 통해 상기 전력(p)의 최적의 값을 산출하는, 컴퓨팅 장치.

(수학적식)

$$\text{Fitness}(\Phi) = - \sum_{u=1}^U \sum_{t=1}^T \left\{ R_{u,t}(\Phi) + \mu F_{u,t}(f_{u,t}(\Phi)) f_{u,t}^2(\Phi) \right\}$$

$f_{u,t}(\Phi)$: 부등식 함수(inequality function)

μ : 기 설정된 상수

$F_{u,t}(f_{u,t}(\Phi))$: 인덱스 함수(index function)

상기 인덱스 함수는 상기 부등식 함수가 0이면 0이 되고, 상기 부등식 함수가 0미만이면 1의 값을 가짐.

청구항 16

청구항 14에 있어서,

상기 제2 알고리즘은, 멀티 에이전트 강화 학습(Multi Agent Reinforcement Learning) 모델을 이용한 알고리즘이고,

아래 수학적식의 목적 함수를 만족하도록 상기 RIS 패널과 원격 사용자 단말 간의 연관 행렬($\mathbf{V}$), RIS 패널의 위상 편이(Φ), 및 위성의 양각(α)을 최적화 하는, 컴퓨팅 장치.

(수학적식)

$$\max_{\mathbf{V}, \Phi, \alpha} \sum_{u=1}^U \sum_{t=1}^T R_{u,t}$$

청구항 17

청구항 16에 있어서,

상기 제2 최적화 명령은,

하기 수학적식에 따른 상기 복수 개 위성의 전역 상태(s') 및 각 위성의 로컬 상태(s's)를 상기 멀티 에이전트 강화 학습 모델로 입력하기 위한 명령을 포함하는, 컴퓨팅 장치.

(수학식)

$$s' = \{\{d_1\}, \{V_1\}, \{\Phi_1\}, \{p_1^*\}, \{i_1, \omega_1, \chi_1\}\}, \dots, \{d_s\}, \{V_s\}, \{\Phi_s\}, \{p_s^*\}, \{i_s, \omega_s, \chi_s\}\}$$

$$s'_s = \{\{d_s\}, \{V_s\}, \{\Phi_s\}, \{p_s^*\}, \{i_s\}, \{\omega_s\}, \{\chi_s\}\}$$

d_s : 위성 s의 위치

V_s : 위성 s의 RIS 패널과 사용자 단말 간의 연관 행렬

Φ_s : 위성 s의 RIS 패널의 위상 편이

p_s^* : 위성 s의 최적화된 전력

i_s : 위성 s의 궤도면과 적도 사이의 교차각

ω_s : 위성 s의 춘분점과 궤도면과 적도면의 교차점 사이의 각도

χ_s : 위성 s의 위성의 방향과 궤도면과 적도면의 교차점이 이루는 각도

청구항 18

청구항 17에 있어서,

상기 제2 최적화 명령은,

상기 멀티 에이전트 강화 학습 모델에서 상기 각 위성의 로컬 상태에 기초하여 상기 RIS 패널과 원격 사용자 단말 간의 연관, 상기 RIS 패널의 위상 편이, 및 상기 위상의 양각에 대한 행동을 산출하기 위한 명령; 및

상기 멀티 에이전트 강화 학습 모델에서 상기 산출된 행동에 대한 글로벌 보상을 상기 복수 개의 위성으로 부여하기 위한 명령을 포함하는, 컴퓨팅 장치.

청구항 19

청구항 18에 있어서,

상기 제2 최적화 명령은,

하기 수학식에 의한 상기 글로벌 보상($R_{s',t}$)이 최대화 되도록 상기 멀티 에이전트 강화 학습 모델을 학습하기 위한 명령을 더 포함하는, 컴퓨팅 장치.

(수학식)

$$R_{s',t} = R_t(s'_t, a_t) = \begin{cases} \sum_{u=1}^U \sum_{t=1}^T R_{u,t}, & \text{if all the associated RUEs is served by networks,} \\ -\Gamma^{\text{pen}}, & \text{otherwise impose penalty to perform better in learning.} \end{cases}$$

s'_t : 시간 슬롯 t에서 위성의 로컬 상태

a_t : 시간 슬롯 t에서 위성의 행동

RUE : 원격 사용자 단말

$-\Gamma^{\text{pen}}$: 페널티 값(penalty value)

청구항 20

RIS(Reconfigurable Intelligent Surface) 패널을 구비하는 복수 개의 위성, 복수 개의 원격 사용자 단말, 및 하나 이상의 기지국 장치를 포함하는 통신 시스템의 통신 방법을 수행하기 위한 명령어들을 포함하고, 비일시적 컴퓨터 판독 가능한 저장 매체(non-transitory computer readable storage medium)에 저장된 컴퓨터 프로그램으로서,

상기 명령어들은 하나 이상의 프로세서들을 갖는 컴퓨팅 장치에 의해 실행될 때, 상기 컴퓨팅 장치로 하여금,

위성에서 송신하는 신호를 원격 사용자 단말이 수신하는 경우, 원격 사용자 단말의 수신 신호를 산출하는 단계;

상기 수신 신호에 대해 해당 원격 사용자 단말의 SNR(Signal to Noise Ratio)을 이용하여 상기 원격 사용자 단말의 데이터 속도를 산출하는 단계; 및

상기 원격 사용자 단말의 데이터 속도를 최대화 하도록 기 설정된 파라미터들을 최적화 하는 단계를 수행하도록 하는, 컴퓨터 프로그램.

발명의 설명**기술 분야**

[0001] 본 발명의 실시예는 저궤도 위성 기반의 통신 기술과 관련된다.

배경 기술

[0003] 최근, IoT(Internet of Things) 기술의 발달로 IoT 장치의 수가 급증하고 있다. 이러한 IoT 장치들의 네트워크를 활용하면 교통, 보건, 산업, 에너지, 해양 등 다양한 분야에서 효율성을 높일 수 있으며 사람들의 삶의 질을 향상시킬 수 있다. 그러나, 세계의 특정 지역에서는 여전히 인터넷 접근이 제한되고 있으며, 이를 해결하기 위해 저궤도(Low Earth Orbit: LEO) 위성 네트워크가 개발되고 있다.

[0004] 그러나, 불균일한 지리적 분포와 동적인 지상 트래픽 수요 특성으로 인해 원격 사용자 장비(Remote User Equipment: RUE)를 위한 위성 통신 리소스가 제한되어 있으며, 특히 저궤도 위성의 신호 전송 거리와 커버리지 범위는 서브테라헤르츠(THz) 주파수 대역에서 높은 전파 감쇠, 분자 흡수, 및 공간 손실로 인해 제한된다는 문제점이 있다. 따라서, 저궤도 위성의 커버리를 최대화하고 데이터 전송률을 최대화하기 위한 방안이 요구된다.

선행기술문헌**특허문헌**

[0006] (특허문헌 0001) 한국등록특허공보 제10-2393301호(2022.05.02)

발명의 내용**해결하려는 과제**

[0007] 본 발명의 실시예는 저궤도 위성의 원격 사용자 단말에 대한 데이터 속도를 최대화 할 수 있는 저궤도 위성 기반의 통신 방법 및 이를 수행하기 위한 컴퓨팅 장치를 제공하기 위한 것이다.

과제의 해결 수단

[0009] 개시되는 일 실시예에 따른 통신 방법은, RIS(Reconfigurable Intelligent Surface) 패널을 구비하는 복수 개

의 위성, 복수 개의 원격 사용자 단말, 및 하나 이상의 기지국 장치를 포함하는 통신 시스템의 통신 방법이고, 하나 이상의 프로세서들, 및 상기 하나 이상의 프로세서들에 의해 실행되는 하나 이상의 프로그램들을 저장하는 메모리를 구비한 컴퓨팅 장치에서 수행되는 통신 방법으로서, 위성에서 송신하는 신호를 원격 사용자 단말이 수신하는 경우, 원격 사용자 단말의 수신 신호를 산출하는 단계; 상기 수신 신호에 대해 해당 원격 사용자 단말의 SNR(Signal to Noise Ratio)을 이용하여 상기 원격 사용자 단말의 데이터 속도를 산출하는 단계; 및 상기 원격 사용자 단말의 데이터 속도를 최대화 하도록 기 설정된 파라미터들을 최적화 하는 단계를 포함한다.

상기 수신 신호(y_u)는 하기 수학식을 통해 산출될 수 있다.

(수학식)

$$y_u = \sqrt{\frac{p_u}{L^{\text{tot}}}} \mathbf{g}_u^T \prod_{r=1, \dots, R_s} \Phi_r \mathbf{H}_r \mathbf{x} + z_u$$

L^{tot} : 신호가 원격 사용자 단말에 수신되기까지의 총 손실

p_u : 기지국 장치의 송신 전력

$\mathbf{g}_u$: 마지막 RIS 패널에서 원격 사용자 단말까지의 채널

$\mathbf{x}$: 기지국 장치의 전송 벡터

R_s : RIS 패널의 전체 개수

z_u : 원격 사용자 단말에서 수신된 AWGN(Additive White Gaussian Noise)

r : RIS의 인덱스

$\mathbf{H}_r$: 기지국 장치와 RIS 패널 r 사이의 채널 행렬

Φ_r : RIS 패널 r 의 위상 편이

상기 원격 사용자 단말의 시간 슬롯 t 에서 SNR(Signal to Noise Ratio)($\Gamma_{u,t}$)은 하기 수학식으로 산출할 수 있다.

(수학식)

$$\Gamma_{u,t} = \frac{p_{s,u} \left| v_{b,u}^s \mathbf{g}_u^T(t) \prod_{r=1, \dots, R_s} \Phi_r \mathbf{H}_r(t) \mathbf{w}_u \right|^2}{N_o L^{\text{tot}}(f, d)}$$

$p_{s,u}$: 위성에서 원격 사용자 단말로 신호를 송신하기 위한 전력

$v_{b,u}^s$: 원격 사용자 단말에 대한 연관 행렬의 요소

N_o : 노이즈 스펙트럼 밀도

$\mathbf{w}_u$: 기지국 장치의 빔포밍 벡터

상기 원격 사용자 단말의 데이터 속도($R_{u,t}$)는 하기 수학식으로 산출할 수 있다.

(수학식)

$$R_{u,t} = B_u \log_2(1 + \Gamma_{u,t})$$

[0032] B_u : 원격 사용자 단말에서 사용 가능한 총 대역폭

[0033] 기 설정된 파라미터들을 최적화 하는 단계는, 하기 수학식의 목적 함수를 만족하도록 위성에서 원격 사용자 단말로 신호를 송신하기 위한 전력(p), 위성의 RIS 패널과 원격 사용자 단말 간의 연관 행렬($\mathbf{V}$), 위성의 RIS 패널의 위상 편이(Φ), 및 위성의 양각(α)을 최적화시킬 수 있다.

[0034] (수학식)

$$\max_{p, \mathbf{V}, \Phi, \alpha} \sum_{u=1}^U \sum_{t=1}^T R_{u,t}$$

[0035]

[0036] $R_{u,t}$: 시간 슬롯 t 에서 원격 사용자 단말 u 의 데이터 속도

[0037] T : 시간 슬롯의 전체 개수

[0038] U : 원격 사용자 단말의 전체 개수

[0039] 기 설정된 파라미터들을 최적화 하는 단계는, 상기 RIS 패널과 원격 사용자 단말 간의 연관 행렬($\mathbf{V}$), RIS 패널의 위상 편이(Φ), 및 위성의 양각(α)을 초기화 한 후 기 설정된 제1 알고리즘을 이용하여 상기 전력(p)을 최적화 하는 제1 최적화 단계; 및 상기 최적화된 전력(p)에 기초하여 기 설정된 제2 알고리즘을 통해 상기 RIS 패널과 원격 사용자 단말 간의 연관 행렬($\mathbf{V}$), RIS 패널의 위상 편이(Φ), 및 위성의 양각(α)을 최적화 하는 제2 최적화 단계를 포함할 수 있다.

[0040] 상기 제1 알고리즘은, WOA(Whale Optimization Algorithm)이고, 상기 제1 최적화 단계는, 하기 수학식에 의한 적합도 함수를 통해 상기 전력(p)의 최적의 값을 산출할 수 있다.

[0041] (수학식)

$$\text{Fitness}(\Phi) = - \sum_{u=1}^U \sum_{t=1}^T \left\{ R_{u,t}(\Phi) + \mu F_{u,t}(f_{u,t}(\Phi)) f_{u,t}^2(\Phi) \right\}$$

[0042]

[0043] $f_{u,t}(\Phi)$: 부등식 함수(inequality function)

[0044] μ : 기 설정된 상수

[0045] $F_{u,t}(f_{u,t}(\Phi))$: 인덱스 함수(index function)

[0046] 상기 인덱스 함수는 상기 부등식 함수가 0이면 0이 되고, 상기 부등식 함수가 0미만이면 1의 값을 가짐.

[0047] 상기 제2 알고리즘은, 멀티 에이전트 강화 학습(Multi Agent Reinforcement Learning) 모델을 이용한 알고리즘이고, 아래 수학식의 목적 함수를 만족하도록 상기 RIS 패널과 원격 사용자 단말 간의 연관 행렬($\mathbf{V}$), RIS 패널의 위상 편이(Φ), 및 위성의 양각(α)을 최적화 할 수 있다.

[0048] (수학식)

$$\max_{\mathbf{V}, \Phi, \alpha} \sum_{u=1}^U \sum_{t=1}^T R_{u,t}$$

[0049]

[0050] 상기 제2 최적화 단계는, 하기 수학식에 따른 상기 복수 개 위성의 전역 상태(s') 및 각 위성의 로컬 상태(s'_s)를 상기 멀티 에이전트 강화 학습 모델로 입력하는 단계를 포함할 수 있다.

[0051] (수학식)

[0052] $s' = \{\{d_1\}, \{V_1\}, \{\Phi_1\}, \{p_1^*\}, \{i_1, \omega_1, \chi_1\}\}, \dots, \{d_s\}, \{V_s\}, \{\Phi_s\}, \{p_s^*\}, \{i_s, \omega_s, \chi_s\}\}$

[0053] $s'_s = \{\{d_s\}, \{V_s\}, \{\Phi_s\}, \{p_s^*\}, \{i_s\}, \{\omega_s\}, \{\chi_s\}\}$

[0054] d_s : 위성 s의 위치

[0055] V_s : 위성 s의 RIS 패널과 사용자 단말 간의 연관 행렬

[0056] Φ_s : 위성 s의 RIS 패널의 위상 편이

[0057] p_s^* : 위성 s의 최적화된 전력

[0058] i_s : 위성 s의 궤도면과 적도 사이의 교차각

[0059] ω_s : 위성 s의 춘분점과 궤도면과 적도면의 교차점 사이의 각도

[0060] χ_s : 위성 s의 위성의 방향과 궤도면과 적도면의 교차점이 이루는 각도

[0061] 상기 제2 최적화 단계는, 상기 멀티 에이전트 강화 학습 모델에서 상기 각 위성의 로컬 상태에 기초하여 상기 RIS 패널과 원격 사용자 단말 간의 연관, 상기 RIS 패널의 위상 편이, 및 상기 위성의 양각에 대한 행동을 산출하는 단계; 및 상기 멀티 에이전트 강화 학습 모델에서 상기 산출된 행동에 대한 글로벌 보상을 상기 복수 개의 위성으로 부여하는 단계를 포함할 수 있다.

[0062] 상기 제2 최적화 단계는, 하기 수학식에 의한 상기 글로벌 보상($R_{s',t}$)이 최대화 되도록 상기 멀티 에이전트 강화 학습 모델을 학습하는 단계를 더 포함할 수 있다.

[0063] (수학식)

[0064]
$$R_{s',t} = R_t(s'_t, a_t) = \begin{cases} \sum_{u=1}^U \sum_{t=1}^T R_{u,t}, & \text{if all the associated RUEs is served by networks,} \\ -\Gamma^{\text{pen}}, & \text{otherwise impose penalty to perform better in learning.} \end{cases}$$

[0065] s'_t : 시간 슬롯 t에서 위성의 로컬 상태

[0066] a_t : 시간 슬롯 t에서 위성의 행동

[0067] RUE : 원격 사용자 단말

[0068] $-\Gamma^{\text{pen}}$: 페널티 값(penalty value)

[0069] 개시되는 일 실시예에 따른 컴퓨팅 장치는, 하나 이상의 프로세서들; 메모리; 및 하나 이상의 프로그램들을 포함하고, 상기 하나 이상의 프로그램들은 상기 메모리에 저장되고, 상기 하나 이상의 프로세서들에 의해 실행되도록 구성되며, RIS(Reconfigurable Intelligent Surface) 패널을 구비하는 복수 개의 위성, 복수 개의 원격 사용자 단말, 및 하나 이상의 기지국 장치를 포함하는 통신 시스템의 통신 방법을 수행하기 위한 컴퓨팅 장치로서, 상기 하나 이상의 프로그램들은, 위성에서 송신하는 신호를 원격 사용자 단말이 수신하는 경우, 원격 사용자 단말의 수신 신호를 산출하기 위한 명령; 상기 수신 신호에 대해 해당 원격 사용자 단말의 SNR(Signal to Noise Ratio)을 이용하여 상기 원격 사용자 단말의 데이터 속도를 산출하기 위한 명령; 및 상기 원격 사용자 단말의 데이터 속도를 최대화 하도록 기 설정된 파라미터들을 최적화 하기 위한 명령을 포함한다.

발명의 효과

[0071] 개시되는 실시예에 의하면, 위성의 전력, 위성의 RIS 패널과 원격 사용자 단말 간의 연관 행렬, 위성의 RIS 패널의 위상 편이, 및 위성의 양각을 최적화 함으로써, 원격 사용자 단말의 수신 신호에 따른 데이터 속도를 최대

화 할 수 있으며, 각 저궤도 위성의 커버리지를 최적화 할 수 있게 된다.

도면의 간단한 설명

- [0073] 도 1은 본 발명의 일 실시예에 따른 저궤도 위성을 이용한 통신 시스템을 개략적으로 나타낸 도면
- 도 2는 본 발명의 일 실시예에서 커버리지 영역을 제어하는데 사용되는 지구 중심에 관한 저궤도 위성의 각도를 개략적으로 나타낸 도면
- 도 3은 본 발명의 일 실시예에서 원격 사용자 단말의 데이터 속도를 최대화하기 위한 과정을 나타낸 흐름도
- 도 4는 본 발명의 일 실시예에서 멀티 에이전트 강화 학습을 위한 개략도를 나타낸 도면
- 도 5는 예시적인 실시예들에서 사용되기에 적합한 컴퓨팅 장치를 포함하는 컴퓨팅 환경을 예시하여 설명하기 위한 블록도

발명을 실시하기 위한 구체적인 내용

- [0074] 이하, 도면을 참조하여 본 발명의 구체적인 실시형태를 설명하기로 한다. 이하의 상세한 설명은 본 명세서에서 기술된 방법, 장치 및/또는 시스템에 대한 포괄적인 이해를 돕기 위해 제공된다. 그러나 이는 예시에 불과하며 본 발명은 이에 제한되지 않는다.
- [0075] 본 발명의 실시예들을 설명함에 있어서, 본 발명과 관련된 공지기술에 대한 구체적인 설명이 본 발명의 요지를 불필요하게 흐릴 수 있다고 판단되는 경우에는 그 상세한 설명을 생략하기로 한다. 그리고, 후술되는 용어들은 본 발명에서의 기능을 고려하여 정의된 용어들로서 이는 사용자, 운용자의 의도 또는 관례 등에 따라 달라질 수 있다. 그러므로 그 정의는 본 명세서 전반에 걸친 내용을 토대로 내려져야 할 것이다. 상세한 설명에서 사용되는 용어는 단지 본 발명의 실시예들을 기술하기 위한 것이며, 결코 제한적이어서는 안 된다. 명확하게 달리 사용되지 않는 한, 단수 형태의 표현은 복수 형태의 의미를 포함한다. 본 설명에서, "포함" 또는 "구비"와 같은 표현은 어떤 특성들, 숫자들, 단계들, 동작들, 요소들, 이들의 일부 또는 조합을 가리키기 위한 것이며, 기술된 것 이외에 하나 또는 그 이상의 다른 특성, 숫자, 단계, 동작, 요소, 이들의 일부 또는 조합의 존재 또는 가능성을 배제하도록 해석되어서는 안 된다.
- [0076] 이하의 설명에 있어서, 신호 또는 정보의 "전송", "통신", "송신", "수신" 기타 이와 유사한 의미의 용어는 일 구성요소에서 다른 구성요소로 신호 또는 정보가 직접 전달되는 것뿐만이 아니라 다른 구성요소를 거쳐 전달되는 것도 포함한다. 특히 신호 또는 정보를 일 구성요소로 "전송" 또는 "송신"한다는 것은 그 신호 또는 정보의 최종 목적지를 지시하는 것이고 직접적인 목적지를 의미하는 것이 아니다. 이는 신호 또는 정보의 "수신"에 있어서도 동일하다. 또한 본 명세서에 있어서, 2 이상의 데이터 또는 정보가 "관련"된다는 것은 하나의 데이터(또는 정보)를 획득하면, 그에 기초하여 다른 데이터(또는 정보)의 적어도 일부를 획득할 수 있음을 의미한다.
- [0077] 또한, 제1, 제2 등의 용어는 다양한 구성 요소들을 설명하는데 사용될 수 있지만, 상기 구성 요소들은 상기 용어들에 의해 한정되어서는 안 된다. 상기 용어들은 하나의 구성 요소를 다른 구성 요소로부터 구별하는 목적으로 사용될 수 있다. 예를 들어, 본 발명의 권리 범위를 벗어나지 않으면서 제1 구성 요소는 제2 구성 요소로 명명될 수 있고, 유사하게 제2 구성 요소도 제1 구성 요소로 명명될 수 있다.
- [0079] 도 1은 본 발명의 일 실시예에 따른 저궤도 위성을 이용한 통신 시스템을 개략적으로 나타낸 도면이다.
- [0080] 도 1을 참조하면, 저궤도 위성을 이용한 통신 시스템(100)은 위성(102), 원격 사용자 단말(104), 및 기지국 장치(106)를 포함할 수 있다.
- [0081] 위성(102)은 복수 개가 마련된다. 위성(102)은 저궤도(Low Earth Orbit: LEO) 위성일 수 있다. 각 위성(102)은 글로벌 커버리지를 제공하기 위해 최적의 위치에 배치될 수 있다. 통신 시스템(100)에서는 글로벌 커버리지를 제공하면서 기 설정된 요구 사항을 만족하는 최소 개수의 위성(102)을 포함할 수 있다.
- [0082] 각 위성(102)은 해당 커버리지 지역의 모든 원격 사용자 단말(104)과 연관(association)될 수 있다. 또한, 각 원격 사용자 단말(104)은 각 위성(102)의 서브 헤르츠 대역에서 사용 가능한 대역폭을 균등하게 나누어 사용할 수 있다.

- [0083] 위성(102)에는 RIS(Reconfigurable Intelligent Surface) 패널이 마련될 수 있다. RIS 패널은 지능형 반사 표면이라고 지칭될 수 있다. 즉, 위성(102)에는 위상 편이를 제어하고 신호 강도를 높일 수 있는 RIS 패널이 장착될 수 있다. RIS 패널은 전자기장을 변경하도록 프로그래밍하여 대상 사용자에게 신호 전력을 집중, 조정, 및 향상시킬 수 있다.
- [0084] 개시되는 실시예에서는, 위성(102)에 RIS 패널을 장착함으로써, 위성 커버리지를 최대화 할 수 있게 된다. 이는 RIS 패널의 위상 편이(phase shift)를 조절하여 구현할 수 있다. RIS 패널은 반사 위상을 변경할 수 있는 메타 표면으로 구성될 수 있다. 이때, RIS 패널에서 위상 편이를 조정하여 대상 사용자에게 빔 포밍 될 수 있다. 위성(102)에서 RIS 패널은 태양광 패널 하부에 배치될 수 있다.
- [0085] 각 RID 패널 r은 n개의 반사 요소(reflecting element)를 포함하는 행렬 $\mathbf{N}_r$ 로 구성될 수 있다. 각 반사 요소는 전파 손실을 방지하기 위해 입사 신호를 재구성할 수 있는 위상 천이기가 장착될 수 있다. 원격 사용자 단말(104)은 위성(102)을 통하여 각종 통신 서비스를 사용할 수 있다. 원격 사용자 단말(104)은 이종의 다양한 기기가 사용될 수 있다. 원격 사용자 단말(104)은 가정, 사무실, 선박, 항공기, 차량 등 다양한 장소와 장비들에 위치할 수 있으나, 각 개인이 휴대하는 장치일 수도 있다.
- [0086] 기지국 장치(106)는 위성(102)의 배치를 제어할 수 있다. 기지국 장치(106)는 위성(102)을 통한 다운 링크(downlink) 통신의 위상 편이, 위상 전송 전력, 및 빔 포밍 등을 최적화하도록 마련될 수 있다. 기지국 장치(106)는 각 위성(102)의 커버리지 최대화를 위한 스케줄링 정책, 송신 전력 할당, 위상 편이 제어, 및 위상 양각 조정을 통해 이를 구현할 수 있다. 이때, 각 위성(102)은 일종의 분산 에이전트 역할을 하고 기지국 장치(106)에 의해 중앙에서 제어될 수 있다. 이에 대한 자세한 설명은 후술하기로 한다. 이러한 최적화를 통해, 기지국 장치(106)는 위성(102)의 원격 사용자 단말(104)에 대한 평균 데이터 속도를 최대화 할 수 있다.
- [0087] 통신 시스템(100)은 위성 기반의 다중 입력 다중 출력(Multiple Input Multiple Output: MIMO) 시스템일 수 있다. 통신 시스템(100)은 원격 사용자 단말(104)에 대한 서브 테라 헤르츠(Sub-THz) 대역에서 전파 감쇠를 최소화 하도록 마련될 수 있다. 통신 시스템(100)은 지상 네트워크 인프라의 커버리지 하에 있지 않은 원격 사용자 단말(104)을 대상으로 고속 데이터 서비스를 제공할 수 있다. 또한, 도시 지역에서도 기존 지상과 네트워크가 충족할 수 없는 수요를 충족시켜 무선 서비스를 제공할 수 있다.
- [0089] 도 2는 본 발명의 일 실시예에서 커버리지 영역을 제어하는데 사용되는 지구 중심에 관한 저궤도 위성의 각도를 개략적으로 나타낸 도면이다.
- [0090] 도 2를 참조하면, 각 위성(s)의 궤도는 원형으로 가정할 수 있다. 시간 t 슬롯에서 위성(s)의 3차원 위치는 $\mathbf{d}_s(t) = [x_s(t), y_s(t), z_s(t)]^T$ 으로 나타낼 수 있다. 그리고, 특정 궤도 m에 있는 위성(s)의 다음 시간 슬롯에서의 위치는 $\mathbf{d}_s^m(t+1) = \mathbf{d}_s^m(t) + \mathbf{v}_s \Delta T$ 와 같이 나타낼 수 있다. 여기서, ΔT 는 각 시간 슬롯 간의 간격이고, $\mathbf{v}_s$ 는 위성의 속도일 수 있다.
- [0091] 한편, 적도에 대한 위성의 방향 및 위치를 좌표화하기 위해 다음의 세 가지 각도(적도 관련 위성 각도)가 사용될 수 있다. 1) 제1 각도인 경사각 i는 궤도면과 적도 사이의 교차각으로 정의될 수 있다. 90도 보다 큰 경사각은 위성의 운동이 지구의 자전 방향과 반대임을 나타낸다. 2) 제2 각도인 각도 ω 는 춘분점과 궤도면과 적도면의 교차점 사이의 각도로 정의될 수 있다. 3) 제3 각도인 각도 χ 는 위성의 방향과 궤도면과 적도면의 교차점이 이루는 각도로 정의될 수 있다.
- [0092] 지구 중심을 원점으로 하는 각 위성(s)의 시간 t 슬롯에서 데카르트 좌표($x_s(t), y_s(t), z_s(t)$)는 하기 수학적 식 1로 나타낼 수 있다.

[0093] (수학식 1)

$$x_s(t) = (h_s + R_e) [\cos \chi(t) \cos \omega(t) - \sin \chi(t) \cos i(t) \sin \omega(t)], \forall s \in \mathcal{S}_m,$$

$$y_s(t) = (h_s + R_e) [\cos \chi(t) \sin \omega(t) - \sin \chi(t) \cos i(t) \cos \omega(t)], \forall s \in \mathcal{S}_m,$$

$$z_s(t) = (h_s + R_e) [\sin \chi(t) \sin i(t)], \forall s \in \mathcal{S}_m,$$

[0094]

[0095] h_s : 위성의 고도

[0096] R_e : 지구의 반경

[0097] $\mathcal{S}_m$: 궤도면 m에 배치되는 위성의 세트

[0098] 여기서, 두 위성(s, s') 사이의 거리는 하기 수학식 2와 같이 산출할 수 있다.

[0099] (수학식 2)

$$d_{s,s'} = \sqrt{(R_e + h_s)^2 + (R_e + h_{s'})^2 - 2(R_e + h_s)(R_e + h_{s'}) \cos \delta}$$

[0101] δ : 지구 중심에서 볼 때 하나의 궤도에 있는 두 위성 사이의 각도

[0102] 두 위성(s, s') 사이의 각도 δ 는 $\delta = |\alpha_s - \alpha_{s'}|$ 로 나타낼 수 있고, α 는 각 위성의 양각(elevation angle)이다.

[0103] 그리고, 기지국 장치(b)와 위성(s) 사이의 시간 t 슬롯에서 거리($d_{b,s}$)는 하기 수학식 3으로 나타내고, 위성(s)과 원격 사용자 단말(u) 사이의 시간 t 슬롯에서 거리($d_{s,u}$)는 하기 수학식 4로 나타낼 수 있다.

[0104] (수학식 3)

$$d_{b,s} = \sqrt{R_e^2 \sin^2 \alpha_b + h_s^2 + 2h_s R_e - R_e \sin \alpha_b}$$

[0106] (수학식 4)

$$d_{s,u} = \sqrt{R_e^2 \sin^2 \alpha_u + h_s^2 + 2h_s R_e - R_e \sin \alpha_u}$$

[0108] 또한, 기지국 장치(b)와 위성(s) 사이의 시간 t 슬롯에서 양각은 하기 수학식 5로 나타내고, 위성(s)과 원격 사용자 단말(u) 사이의 시간 t 슬롯에서 양각은 하기 수학식 6으로 나타낼 수 있다.

[0109] (수학식 5)

$$\alpha_{b,s}(t) = \arccos \left(\frac{R_e^2 + d_{b,s}(t) - (R_e + h)^2}{2R_e d_{b,s}(t)} \right) - \frac{\pi}{2}$$

[0110]

[0111] (수학식 6)

$$\alpha_{s,u}(t) = \arccos \left(\frac{R_e^2 + d_{s,u}(t) - (R_e + h)^2}{2R_e d_{s,u}(t)} \right) - \frac{\pi}{2}$$

[0112]

[0113] 여기서, 가시선(Line-of-Sight: LoS) 전파 거리와 기지국 장치(b) 및 원격 사용자 단말(u)의 최소 양각은 각 시간 슬롯 t에서 위성(s)가 제공하는 커버리지 영역을 결정할 수 있다. 이때, 최소 양각이 커지면 위성의 커버리지는 작아지게 된다.

[0115] 통신 시스템(100)에서 간접 통신 링크는 위성-기지국 장치 간 링크, 위성-위성 간 링크, 및 위성-원격 사용자 단말 간의 링크가 있다. 여기서, 각 링크는 동일한 기능을 공유하므로 동일한 채널 모델(channel model)을 사용할 수 있다. 그리고, 서브 테라 헤르츠 대역에서 신호 전파에는 확산 손실(spreading loss), 지구 근처의 분자 흡수 손실(molecular absorption loss), 비와 구름에 의한 손실, 및 전리층 플라즈마 손실 등에 의해 영향을 받

게 된다.

[0116] 확산 손실은 등방성 수신기에서 감지할 수 있는 등방성 송신기에서 방출된 전력의 비율을 의미할 수 있다. 분자 흡수 손실은 진동하는 분자 내에서 운동 에너지로 변환된 전자기 에너지의 비율을 나타낼 수 있다. 비와 구름에 의한 손실은 신호가 대기를 통해 이동할 때 비와 구름에 의해 산란되고 흡수되는데 따른 손실을 의미할 수 있다. 전리층 플라즈마 손실은 신호가 전리층의 플라즈마를 통과할 때 발생하는 손실을 의미할 수 있다.

[0117] 여기서, 주파수 f 의 신호가 거리 d 만큼 전파될 때 발생하는 총 손실은 하기 수학적 식 7로 나타낼 수 있다.

[0118] (수학적 식 7)

$$L^{\text{tot}}(f, d) = \frac{(4\pi df)^2 L^{\text{abs}}(f, d) L^{\text{rain}} L^{\text{cloud}} L^{\text{atten}}}{c^2 G_b G_u d_x d_y A^\alpha \left| \sum_{n_s=1}^{N_r} \frac{\sqrt{F_{n_s}^{\text{tot}}}}{d_{b,s} d_{s,s'} d_{s,u}} \right|^\alpha}$$

[0119]

[0120] L^{abs} : 분자 흡수 손실

[0121] L^{rain} : 비에 의한 손실

[0122] L^{cloud} : 구름에 의한 손실

[0123] L^{atten} : 감쇠 계수(신호의 전파 경로에 따른 단위 거리당 손실)

[0124] c : 빛의 속도

[0125] G_b : 기지국 장치의 안테나 이득

[0126] G_u : 원격 사용자 단말의 안테나 이득

[0127] d_x, d_y : 위성의 단위 셀의 x방향과 y방향의 크기

[0128] A : 위성의 커버리지

[0129] α : 위성의 양각

[0130] $d_{b,s}$: 위성 and 기지국 장치 간의 거리

[0131] $d_{s,s'}$: 위성 and 다른 위성 간의 거리

[0132] $d_{s,u}$: 위성 and 원격 사용자 단말 간의 거리

[0133] N_r : RIS 패널의 반사 요소의 개수

[0134] n_s : RIS 패널의 s 번째 반사 요소

[0135] $F_{n_s}^{\text{tot}}$: 수신 신호 전력에서 고려되는 정규화된 총 방사 패턴

[0137] 한편, 기지국 장치(106)에는 전용 구역 내의 복수 개의 원격 사용자 단말(102)과 통신을 수행하는 복수 개(K개)의 안테나가 장착될 수 있다. 기지국 장치(106)는 서브 테라헤르츠 통신의 커버리지 범위를 향상시키기 위해 여러 RIS 패널을 원격 사용자 단말(102)로 지원하여 K개의 안테나로부터 데이터 스트림을 동시에 전송할 수 있다. RIS 패널을 활성화 하면, 각 데이터 스트림은 하나의 원격 사용자 단말(102)로 빔 형성되게 된다.

[0138] 여기서, $\mathbf{H}_1 \in \mathbb{C}^{(N_1 \times K)}$ 을 기지국 장치와 첫 번째 위성의 RIS 사이의 채널 행렬로 정의하고, $\mathbf{H}_{r+1} \in \mathbb{C}^{(N_{r+1} \times N_r)}$ 을 RIS r 과 RIS $r+1$ 사이의 채널 행렬로 정의하면, 원격 사용자 단말(u)에서 수신된 신호(y_u)는 아래 수학적 식 8로 나타낼 수 있다. 이때, 기지국 장치(b)에서 전송된 신호는 원격 사용자 단말(u)에 도

달하기 전에 RIS 홉 R_s ($R_s \leq S$, S 는 위성의 전체 개수)을 거쳐 이동한다고 가정할 수 있다. 즉, R_s 는 신호가 원격 사용자 단말에 도달하기 전에 거치는 RIS 패널을 의미할 수 있다.

(수학식 8)

$$y_u = \sqrt{\frac{p_u}{L^{\text{tot}}}} \mathbf{g}_u^T \prod_{r=1, \dots, R_s} \Phi_r \mathbf{H}_r \mathbf{x} + z_u$$

L^{tot} : 신호가 원격 사용자 단말에 수신되기까지의 총 손실

p_u : 기지국 장치의 송신 전력

$\mathbf{g}_u$: 마지막 RIS에서 원격 사용자 단말까지의 채널

$\mathbf{x}$: 기지국 장치의 전송 벡터

z_u : 원격 사용자 단말에서 수신된 AWGN(Additive White Gaussian Noise)

$\mathbf{H}_r$: 기지국 장치와 RIS r 사이의 채널 행렬

Φ_r : RIS r 의 위상 편이

여기서, r 번째 RIS 패널의 위상 편이(Φ_r)은 아래의 수학식 9로 나타낼 수 있다.

(수학식 9)

$$\Phi_r \triangleq \text{diag}[A_{r1}e^{j\theta_{r1}}, A_{r2}e^{j\theta_{r2}}, \dots, A_{rn_r}e^{j\theta_{rn_r}}]$$

A_{rn} : r 번째 RIS 패널에서 n 번째 요소의 진폭

θ_{rn} : r 번째 RIS 패널에서 n 번째 요소의 반사 계수

또한, R 개의 모든 RIS 패널과 U 개의 모든 원격 사용자 단말에 대한 연관 행렬(association matrix)는 $\mathbf{V} \in \mathbb{C}^{(U \times R)}$ 로 나타낼 수 있다. 각 원격 사용자 단말(u)에 대한 연관 행렬 $\mathbf{V}$ 의 요소는 아래 수학식 10으로 나타낼 수 있다.

(수학식 10)

$$v_{b,u}^s = \begin{cases} 1, & \text{if GBS } b \text{ transmits to RUE } u \text{ via } s^{\text{th}} \text{ satellite} \\ 0, & \text{otherwise.} \end{cases}$$

GBS : 기지국 장치

RUE : 원격 사용자 단말

또한, 전송 벡터($\mathbf{x}$)는 $\mathbf{x} \triangleq \sum_{u=1}^U \mathbf{w}_u s_u$ 으로 나타낼 수 있다. 여기서, $\mathbf{w}_u \in \mathbb{C}^{(K \times 1)}$ (K 는 기지국 장치의 안테나 개수)는 빔포밍 벡터이고, $s_u \in \mathcal{CN}(0, 1)$ 는 각 원격 사용자 단말(u)의 가우시안 신호일 수 있다.

수학식 8로 주어지는 수신 신호에 대해 각 시간 슬롯 t 에서 원격 사용자 단말(u)의 순간 SNR(Signal to Noise Ratio)은 하기 수학식 11로 나타낼 수 있다.

[0160] (수학식 11)

$$\Gamma_{u,t} = \frac{p_{s,u} \left| v_{b,u}^s \mathbf{g}_u^T(t) \prod_{r=1,\dots,R_s} \Phi_r \mathbf{H}_r(t) \mathbf{w}_u \right|^2}{N_o L^{\text{tot}}(f, d)}$$

[0161]

[0162] $p_{s,u}$: 위성에서 원격 사용자 단말로 신호를 송신하기 위한 전력

[0163] N_o : 노이즈 스펙트럼 밀도

[0164] 그리고, 시간 슬롯 t 에서 원격 사용자 단말(u)의 실행 가능한 데이터 속도($R_{u,t}$)는 하기 수학식 12로 나타낼 수 있다.

[0165] (수학식 12)

$$R_{u,t} = B_u \log_2(1 + \Gamma_{u,t})$$

[0166]

[0167] B_u : 원격 사용자 단말에서 사용 가능한 총 대역폭

[0168] 따라서, RIS 패널과 SNR 레벨을 증가시키면 데이터 속도를 향상시킬 수 있게 된다.

[0170] 통신 시스템(100)에서는 기 설정된 제한 조건에서 RIS 패널과 원격 사용자 단말 간의 연관, 위상 편이, 위성 전력, 및 양각을 최적화 하여 원격 사용자 단말(102)의 데이터 속도를 최대화 할 수 있다. 이에 따른, 통신 시스템(100)의 목적 함수는 아래 수학식 13과 같이 나타낼 수 있다.

[0171] (수학식 13)

$$\max_{\mathbf{p}, \mathbf{V}, \Phi, \alpha} \sum_{u=1}^U \sum_{t=1}^T R_{u,t}$$

[0172]

[0173] $\mathbf{p}$: 위성에서 원격 사용자 단말로 신호를 송신하기 위한 전력

[0174] $\mathbf{V}$: 위성의 RIS 패널과 원격 사용자 단말 간의 연관 행렬

[0175] Φ : 위성의 RIS 패널의 위상 편이

[0176] α : 위성의 양각

[0177] 도 3은 본 발명의 일 실시예에서 원격 사용자 단말의 데이터 속도를 최대화하기 위한 과정을 나타낸 흐름도이다. 도 3을 참조하면, 원격 사용자 단말의 데이터 속도를 최대화하는 문제는 크게 2가지로 나눌 수 있다. 먼저, 연관 행렬, 위상 편이, 및 양각을 초기화 하여 위성의 전력을 최적화한 후(S 101), 최적화 된 위성의 전력에 기초하여 연관 행렬, 위상 편이, 및 양각을 최적화 할 수 있다(S 103). 이때, 원격 사용자 단말의 데이터 속도가 최대가 되는지를 확인하여(S 105), 데이터 속도가 최대화 되었으면 위성의 전력, 연관 행렬, 위상 편이, 및 양각의 최적화 값을 출력할 수 있다(S 107). 만약, 데이터 속도가 최대화 되지 않았으면 위의 과정을 다시 반복할 수 있다.

[0178] 구체적으로, 위성의 송신 전력을 최적화 하는 문제는 WOA(Whale Optimization Algorithm)을 이용하여 해결할 수 있다. RIS 패널과 원격 사용자 단말 간의 연관 행렬($\mathbf{V}$), RIS 패널의 위상 편이(Φ), 및 위성의 양각(α)을 초기화하면, 위성의 송신 전력을 최적화 하는 문제는 아래 수학식 14와 같이 나타낼 수 있다.

[0179] (수학식 14)

$$\max_p \sum_{u=1}^U \sum_{t=1}^T R_{u,t}(p)$$

[0180]

[0181] 여기서, 고래 최적화 알고리즘(WOA)은 혹등 고래의 사냥 특성을 모사한 메타-휴리스틱 알고리즘이다. 혹등 고래는 먹이를 발견한 후 먹이 아래로 깊이 잠수하고 나선형의 거품을 생성하는데, 이러한 거품은 먹이를 더 작은 범위로 포위하는 역할을 한다.

[0182] 고래 최적화 알고리즘에서 개별 고래(agent)는 서로 다른 제어 변수들을 나타내며, 고래와 먹이 간 거리를 목적 함수의 비용을 의미한다. 이때, 개별 고래의 시간 별 위치는 1) 먹이 포위(Encircling Prey), 2) 버블-넷 공격(Bubble-Net Attacking), 및 3) 먹이 탐색(Search for Prey)의 과정을 통해 산출된다. 고래 최적화 알고리즘(WOA)은 기 공지된 알고리즘이므로 이에 대한 자세한 설명은 생략하기로 한다.

[0183] 개시되는 실시예에서, 위성의 전력(p)이 고래의 위치가 될 수 있다. 최적의 고래 위치(즉, 위성의 전력 값)은 하기 수학식 15에 의한 적합도 함수를 통해 산출할 수 있다.

[0184] (수학식 15)

$$\text{Fitness}(\Phi) = - \sum_{u=1}^U \sum_{t=1}^T \left\{ R_{u,t}(\Phi) + \mu F_{u,t}(f_{u,t}(\Phi)) f_{u,t}^2(\Phi) \right\}$$

[0185]

[0186] $f_{u,t}(\Phi)$: 부등식 함수(inequality function)

[0187] μ : 기 설정된 상수

[0188] $F_{u,t}(f_{u,t}(\Phi))$: 인덱스 함수(index function)

[0189] 여기서, 인덱스 함수 $F_{u,t}(f_{u,t}(\Phi))$ 는 부등식 함수 $f_{u,t}(\Phi)$ 가 0 이상이면 0이 되고, 부등식 함수 $f_{u,t}(\Phi)$ 가 0미만이면 1의 값을 가질 수 있다. 즉, 인덱스 함수와 부등식 함수의 관계는 아래 수학식 16으로 나타낼 수 있다.

[0190] (수학식 16)

$$F_{u,t}(f_{u,t}(\Phi)) = 0 \text{ if } f_{u,t}(\Phi) \geq 0$$

[0191]

$$F_{u,t}(f_{u,t}(\Phi)) = 1 \text{ if } f_{u,t}(\Phi) < 0$$

[0192]

[0193] 수학식 15에서 적합도 함수의 우변에 음의 부호를 두는 것은 최대화 문제를 최소화 문제로 변환하기 위한 것이다.

[0194] 이와 같이, 위성의 전력(p)이 최적화 되면, 통신 시스템(100)은 멀티 에이전트 강화 학습(Multi Agent Reinforcement Learning)을 통해 RIS 패널과 원격 사용자 단말 간의 연관 행렬($\mathbf{V}$), RIS 패널의 위상 편이(Φ), 및 위성의 양각(α)을 최적화 할 수 있다. 이는 아래의 수학식 17과 같이 나타낼 수 있다.

[0195] (수학식 17)

$$\max_{\mathbf{V}, \Phi, \alpha} \sum_{u=1}^U \sum_{t=1}^T R_{u,t}$$

[0196]

[0197] 도 4는 본 발명의 일 실시예에서 멀티 에이전트 강화 학습을 위한 개략도를 나타낸 것이다. 도 4를 참조하면, RIS-LEO 네트워크 환경은 저궤도 위성, 사용되는 전력, RIS의 입력 및 출력 신호, 위성의 고도 및 각도의 모든

정보를 포함하는 전체 네트워크 환경을 포함할 수 있다. 이때, 네트워크 내 정보는 각 저궤도 위성(에이전트)을 공동으로 학습하는데 사용될 수 있다. 각 에이전트는 네트워크 환경으로부터 상태에 대한 관찰을 받을 수 있다. 여기서, 각 위성은 에이전트(agent)의 역할을 하게 된다. 멀티 에이전트 강화 학습 모델은 기지국 장치에 설치될 수 있으나, 이에 한정되는 것은 아니며 그 이외의 서버 컴퓨팅 장치에 설치될 수도 있다.

[0198] 멀티 에이전트 강화 학습 모델은 정책 네트워크 θ_s 와 가치 네트워크 σ_s 를 가질 수 있다. 여기서, 정책 네트워크(policy network) 및 가치 네트워크(value network)는 멀티 에이전트 강화 학습을 위한 인공 신경망일 수 있다. 가치 네트워크는 각 에이전트의 로컬 비평가 역할을 할 수 있다. 가치 네트워크는 각 에이전트(위성)가 새 정책을 학습하면 그 정확성에 대한 즉각적인 피드백을 제공할 수 있다. 정책 네트워크는 주요 정책의 학습 단위 역할을 할 수 있다. 정책은 각 에이전트에서 의사 결정(즉, 행동)을 내리는데 사용될 수 있다.

[0199] 각 에이전트가 분산적으로 취한 행동은 중앙 집중식 비평가에서 평가에 사용될 수 있다. 중앙 집중식 비평가는 모든 에이전트 간의 협력을 보장하며, 모든 에이전트의 전역 손실을 계산하고 시스템의 성능이 전역적으로 수렴 되도록 할 수 있다. 분산된 각 에이전트의 결과를 분석하기 위해 중앙 집중식 비평가(Centralized Critic)에게 현재의 보상(reward)을 보낼 수 있으며, 각 보상 및 상태는 경험 풀(Experience Pool)에 저장될 수 있다.

[0200] 정책 네트워크는 액터(actor)로 지칭되고, 가치 네트워크는 비평가(critic)로 지칭될 수 있다. 정책 네트워크(즉, 액터)는 로컬 정보(즉, 해당 에이전트의 정보)만 보고 이를 기반으로 행동(action)을 생성하나, 가치 네트워크(즉, 비평가)는 자신을 더 잘 최적화하기 위해 다른 에이전트(즉, 다른 위성들)의 전체 정보를 볼 수 있다.

[0201] 여기서, 전역 상태(global state)를 s' 라 하고, 에이전트 s (즉, 위성 s)의 로컬 상태(local state)를 s'_s 라 하면, 이는 다음의 수학적 식 18과 같이 나타낼 수 있다. 각 에이전트의 로컬 상태들(즉, 전역 상태)는 강화 학습 모델의 입력으로 사용된다.

[0202] (수학적 식 18)

$$s' = \{\{d_1\}, \{V_1\}, \{\Phi_1\}, \{p_1^*\}, \{i_1, \omega_1, \chi_1\}\}, \dots, \{d_s\}, \{V_s\}, \{\Phi_s\}, \{p_s^*\}, \{i_s, \omega_s, \chi_s\}\}$$

$$s'_s = \{\{d_s\}, \{V_s\}, \{\Phi_s\}, \{p_s^*\}, \{i_s\}, \{\omega_s\}, \{\chi_s\}\}$$

[0205] d_s : 위성 s 의 위치

[0206] V_s : 위성 s 의 RIS 패널과 사용자 단말 간의 연관 행렬

[0207] Φ_s : 위성 s 의 RIS 패널의 위상 편이

[0208] p_s^* : 위성 s 의 최적화된 전력

[0209] i_s : 위성 s 의 궤도면과 적도 사이의 교차각

[0210] ω_s : 위성 s 의 춘분점과 궤도면과 적도면의 교차점 사이의 각도

[0211] χ_s : 위성 s 의 위성의 방향과 궤도면과 적도면의 교차점이 이루는 각도

[0212] 에이전트 s 는 담당 지역에서 최적의 전력 $\{p_s^*\}$ 을 얻은 후, 로컬 상태 s'_s 를 액터(actor)(정책 네트워크)에게 전달할 수 있다. 현재의 로컬 상태 s'_s 에서 액터는 RIS 패널과 원격 사용자 단말 간의 연관, RIS 패널의 위상 편이, 및 위상의 양각을 최적화하는 방법을 계산할 수 있다. 이때, 에이전트의 액터가 수행한 행동(a_s)은 하기 수학적 식 19로 나타낼 수 있다. 즉, 에이전트(위성)은 RIS 패널과 원격 사용자 단말 간의 연관, RIS 패널의 위상 편이, 및 위상의 양각 중 하나 이상을 변경하는 행동을 할 수 있다.

[0213] (수학적 식 19)

$$a_{s'} = \{\{\alpha_{s'}\}, \{V_{s'}\}, \{\phi_{s'}\}\}$$

[0215] 행동 확률(action probability)을 생성한 후, 액터는 행동 분포를 샘플링하고 로컬 환경과 상호 작용하여 다음

상태($\tilde{s}'$)로 이동하기 전에 보상(R_s)을 얻을 수 있다. 그리고, 경험 풀(experience pool) D에는 경험 ($s'_s, a_s, R_s, \tilde{s}'_s$)이 저장된다. 이때, 수학적 17을 최대화 하기 위한 데이터 처리량과 QoS 공정성을 즉각적인 보상으로 사용할 수 있다. 또한, 에이전트들의 협력을 강화하기 위해 수학적 20으로 표현되는 글로벌 보상을 받을 수 있다.

(수학적 20)

$$R_{s',t} = R_t(s'_t, a_t) = \begin{cases} \sum_{u=1}^U \sum_{t=1}^T R_{u,t}, & \text{if all the associated RUEs is served by networks,} \\ -\Gamma^{\text{pen}}, & \text{otherwise impose penalty to perform better in learning.} \end{cases}$$

$-\Gamma^{\text{pen}}$: 페널티 값(penalty value)

여기서, 글로벌 보상 $R_{s',t}$ 은 시간 슬롯 t에서 에이전트 s'의 즉각적인 보상이다. 모든 에이전트(모든 위성)들은 글로벌 보상을 공유할 수 있다. 강화 학습 모델은 글로벌 보상이 최대화 되도록 학습될 수 있다.

누적 할인 보상(cumulative discounted reward)은 할인을 γ (0과 1 사이의 값)을 사용하여 수학적 21을 통해 산출할 수 있다.

(수학적 21)

$$R(t) = \sum_{k=0}^{\infty} \gamma^k R_{s',t}(t+k+1)$$

멀티 에이전트들의 통합 정책(combined policy)을 $\pi = \{\pi_s | s \in \mathcal{S}\}$ 라고 하면, 상태-가치 함수(state-value fuction)(V^π)는 수학적 22로 나타낼 수 있다.

(수학적 22)

$$V^\pi(s'(t)) = \mathbb{E}_{a(t), s'(t+1), \dots} [R(t) | s'(t)]$$

여기서, 최적화의 문제는 초기 상태 s'(0)에 대해 예상되는 누적 할인 보상을 최대화하기 위한 최선의 방법을 결정하는 것으로 적용될 수 있으며, 따라서 멀티 에이전트 강화 학습을 통한 수학적 17의 최적화 문제는 하기 수학적 23으로 나타낼 수 있다. 즉, 수학적 17의 최적화 문제는 시간 경과에 따른 예상 누적 보상을 최대화하는 최적의 정책을 탐색하는 것으로 표현될 수 있다.

(수학적 23)

$$\begin{aligned} \pi^* &= \arg \max_{\pi} \mathbb{E}_{s'(0) \sim \rho_s(s'(0))} [V^\pi(s'(0))] \\ &= \arg \max_{\pi} \kappa(\pi). \end{aligned}$$

한편, 주어진 π^{old} 를 가지고 $\kappa(\pi)$ 를 최대화 하는 것은 하기 수학적 24를 최대화 하는 것과 같다.

(수학적 24)

$$\mathbb{E}_{\pi} [A^{\pi^{\text{old}}}(s'(t), a(t))]$$

$A^\pi(s'(t), a(t))$: 이점 함수(advantage function)

여기서, 이점 함수(advantage function)는 $A^\pi(s'(t), a(t)) = \{Q^\pi(s'(t), a(t)) - V^\pi(s'(t))\}$ 와 같다.

그리고, $Q^\pi(s'(t), a(t))$ 는 상태-행동 가치 함수(state-action value function)로 이는 $Q^\pi(s'(t), a(t)) = \mathbb{E}_{a(t+1), s'(t+1), \dots}[R(t)|(s'(t), a(t))]$ 으로 나타낼 수 있다. 또한, 클립 함수(clip function)를 사용하여 $\mathbb{E}_\pi[A^{\pi^{\text{old}}}(s'(t), a(t))]$ 를 근사화하면 수학적 23의 최적화 문제를 아래 수학적 25와 같이 변환할 수 있다.

(수학적 25)

$$\max_{\theta=\{\theta_s|s \in \mathcal{S}\}} \mathbb{E}_{\pi^{\text{old}}} \left\{ \min[r'(\theta_s)A^{\pi^{\text{old}}}(s'(t), a(t)), \text{clip}(r'(\theta_s), \epsilon)A^{\pi^{\text{old}}}(s'(t), a(t))] \right\}$$

$r'(\theta)$: 결합 확률 비율(joint probability ratio)

여기서, 결합 확률 비율은 $r'(\theta) = \frac{\pi_s(a|s'; \theta)}{\pi_s^{\text{old}}(a|s'; \theta^{\text{old}})}$ 으로 나타낼 수 있다. 결합 확률 비율의 값은 정책 업데이트를 작은 범위로 유지하는 클립 함수를 사용하여 $[1 - \epsilon_1, 1 + \epsilon_1]$ 의 범위 내에서 유지될 수 있다.

에이전트 s 를 제외한 모든 에이전트들의 결합 행동으로 $\mathbf{a}^{-s}$ 을 사용하고, 에이전트 s 를 제외한 모든 에이전트들의 통합 정책으로 π^{-s} 을 사용할 수 있다. 정책들은 독립적이기 때문에 $\pi(\mathbf{a}|\boldsymbol{\tau}) = \prod_{s=1}^S \pi^s(a^s|\boldsymbol{\tau}^s)$ 라고 믿는 것이 안전하다. 여기서, $\boldsymbol{\tau}$ 는 행동 관찰 히스토리(action-observation history)이다. 이러한 인수분해(factorization)로부터 다음 수학적 26과 같은 목표를 생성할 수 있다.

(수학적 26)

$$\max_{\theta^1, \dots, \theta^S} \mathbb{E}_{\mathbf{a} \sim \pi_{\text{old}}} \left\{ \min \left[\left(\prod_{s=1}^S r'_s \right) A^\pi, \text{clip} \left(\left(\prod_{s=1}^S r'_s \right), 1 - \epsilon, 1 + \epsilon \right) A^\pi \right] \right\}$$

θ^s : 에이전트 s 의 정책 파라미터(policy parameter)

여기서, $r_s = \frac{\pi^s(a^s|\boldsymbol{\tau}^s; \theta^s)}{\pi_{\text{old}}^s(a^s|\boldsymbol{\tau}^s; \theta_{\text{old}}^s)}$ 이다. 이점 함수 A^π 는 정의되어 있지만, 각 에이전트의 개별 기여도를 결정하기는 어렵다. 신용 할당을 용이하게 하기 위해, 결합 이점 함수(joint advantage function)은 에이전트의 특정 로컬 에이전트로 다음과 같이 분해될 수 있다. $A^\pi(s, \mathbf{a}) = \sum_{i=1}^N c^s \cdot A^s(s, (a^s, \mathbf{a}^{-s}))$

여기서, $A^s(s, (a^s, \mathbf{a}^{-s})) = Q^\pi(s, (a^s, \mathbf{a}^{-s})) - \mathbb{E}[Q^\pi(s, (\tilde{a}^s, \mathbf{a}^{-s}))]$ 는 반대 사실 이점(counterfactual advantage)을 나타내고, c^s 는 음수가 아닌 가중치를 나타낸다.

A^π 의 음이 아닌 분해로 인해 글로벌 최적과 로컬 최적 사이에 단조로운 연결이 있으며, 이는 $\arg \max_{\theta^1, \dots, \theta^S}$ 에서 $\arg \max_{\theta^1}, \dots, \arg \max_{\theta^S}$ 로의 변환을 의미한다. 수학적 26의 최적화는 각 에이전트의 개별 목표를 최대화하도록 다음 수학적 27과 같이 변경할 수 있다.

(수학적 27)

$$L(\theta^s) = \mathbb{E}_{\mathbf{a} \sim \pi_{\text{old}}} \left\{ \min \left[\left(\prod_{j \neq s} r^j \right) A^s, \text{clip} \left(\left(\prod_{j \neq s} r^j \right) r^s, 1 - \epsilon, 1 + \epsilon \right) A^s \right] \right\}$$

그러나, 각 에이전트가 실행 전반에 걸쳐 완전히 독립적인 경우에는 과도한 분산(variance)의 가능성을 높이므로, 에이전트의 수에 따라 분산이 기하급수적으로 증가할 수 있다. 이에 이중 클리핑(double clipping) 접근 방

식을 통해 수학식 27을 다음의 수학식 28과 같이 수정할 수 있다.

(수학식 28)

$$L(\theta^s) = \mathbb{E}_{\mathbf{a} \sim \pi_{\text{old}}} \{ \min [g(\mathbf{r}^{-s})r^s A^s, \text{clip}(g(\mathbf{r}^{-s})r^s, 1 - \epsilon_1, 1 + \epsilon_1)A^s] \}$$

여기서, $g(\mathbf{r}^{-s}) = \text{clip}(\prod_{j \neq s} r^j, 1 - \epsilon_2, 1 + \epsilon_2)$, $\epsilon_2 < \epsilon_1$ 이다. 수학식 27에서 $\prod_{j \neq s} r^j$ 의 존재는 $\prod_{j \neq s} r^j$ 의 가중치로 인해 에이전트의 목표에 영향을 미친다. $\prod_{j \neq s} r^j$ 에 대한 클리핑 결과, 에이전트 s에 대한 다른 에이전트의 업데이트 효과는 $[1 - \epsilon_2, 1 + \epsilon_2]$ 로 제한되어 다른 에이전트에 의해 생성된 변형(variation)을 제한할 수 있다.

액터 네트워크(actor network)의 매개변수가 수정되면 샘플링된 행동(action)이 비평가 네트워크(critic network)로 전달되고, 비평가 네트워크는 행동의 추정된 가치를 추정하게 된다. 예측된 값을 기반으로 정책 기울기 접근 방식을 사용하여 액터를 변경할 수 있다. 보존된 샘플을 사용하여 훈련 단계에서 이점 함수 $A^{\pi^{\text{old}}}$ 의 기대치를 추정할 수 있다. 즉, 정책은 기울기를 사용하여 다음의 수학식 29와 같이 업데이트 될 수 있다.

(수학식 29)

$$\Delta \theta_s = \nabla_{\theta_s} \hat{\mathbb{E}} \left[\min [g(\mathbf{r}^{-s})r^s A^s, \text{clip}(g(\mathbf{r}^{-s})r^s, 1 - \epsilon_1, 1 + \epsilon_1)A^s] \right]$$

여기서, A^s 는 일반화 이점 추정(Generalized Advantage Estimation: GAE)이며, 이는 다음의 수학식 30으로 정의할 수 있다.

(수학식 30)

$$A^s(s(t), a(t)) = \hat{Q}(s'(t), a(t)) - V_{\tau_s^{\text{old}}}(s'(t))$$

여기서, $\hat{Q}_s(s'(t), a(t))$ 는 k-스텝 부트스트래핑 선형 조합(k-step bootstrapping linear combination)을 사용하여 다음 수학식 31과 같이 추정할 수 있다.

(수학식 31)

$$\hat{Q}(s'(t), a(t)) = \sum_{k=t}^{\infty} (\Xi v)^{k-t} \delta(t) + Q_{\sigma_s^{\text{old}}}(s(t), a(t))$$

여기서, $\delta(t)$ 는 시간 차(temporal difference)이며 이는 수학식 32로 정의할 수 있다.

(수학식 32)

$$\delta(t) = r(t) + \Xi Q_{\sigma_s}(s'(t+1), a(t+1)) - Q_{\sigma_s^{\text{old}}}(s'(t), a(t))$$

각 위성 s는 동일한 결합된 행동-상태 가치 함수를 추정하기 때문에, 가치 네트워크의 손실 함수는 다음 수학식 33과 같이 나타낼 수 있으며, 이러한 손실 함수를 최소화 하도록 가치 네트워크가 학습될 수 있다.

(수학식 33)

$$L^{\text{value}}(t, \sigma_s) = (\hat{Q}_s(s'(t), a(t)) - Q_{\sigma_s^{\text{old}}}(s'(t), a(t)))^2$$

이를 위해, 경사 하강 방식이 사용될 수 있으며, 각 위성(에이전트)의 기울기 값은 아래 수학식 34를 통해 산출

할 수 있다.

[0267] (수학식 34)

$$\Delta\sigma_s = \nabla_{\sigma_s} \hat{\mathbb{E}} \left[(\hat{Q}_s(s'(t), a(t)) - Q_{\sigma_s^{\text{old}}}(s'(t), a(t)))^2 \right]$$

[0269] 여기서, 매개변수 θ 와 σ 는 각각 수학식 29와 수학식 34에 따라 정책 네트워크와 가치 네트워크의 손실 함수가 수렴할 때까지 변경될 수 있다.

[0270] 개시되는 실시예에 의하면, 위성의 전력, 위성의 RIS 패널과 원격 사용자 단말 간의 연관 행렬, 위성의 RIS 패널의 위상 편이, 및 위성의 양각을 최적화 함으로써, 원격 사용자 단말의 수신 신호에 따른 데이터 속도를 최대화 할 수 있으며, 각 저궤도 위성의 커버리지를 최적화 할 수 있게 된다.

[0272] 도 5는 예시적인 실시예들에서 사용되기에 적합한 컴퓨팅 장치를 포함하는 컴퓨팅 환경(10)을 예시하여 설명하기 위한 블록도이다. 도시된 실시예에서, 각 컴포넌트들은 이하에 기술된 것 이외에 상이한 기능 및 능력을 가질 수 있고, 이하에 기술된 것 이외에도 추가적인 컴포넌트를 포함할 수 있다.

[0273] 도시된 컴퓨팅 환경(10)은 컴퓨팅 장치(12)를 포함한다. 일 실시예에서, 컴퓨팅 장치(12)는 위성(102)일 수 있다. 또한, 컴퓨팅 장치(12)는 원격 사용자 단말(104)일 수 있다. 또한, 컴퓨팅 장치(12)는 기지국 장치(106)일 수 있다. 또한, 컴퓨팅 장치(12)는 멀티 에이전트 강화 학습 모델이 설치된 서버 컴퓨팅 장치일 수 있다.

[0274] 컴퓨팅 장치(12)는 적어도 하나의 프로세서(14), 컴퓨터 판독 가능 저장 매체(16) 및 통신 버스(18)를 포함한다. 프로세서(14)는 컴퓨팅 장치(12)로 하여금 앞서 언급된 예시적인 실시예에 따라 동작하도록 할 수 있다. 예컨대, 프로세서(14)는 컴퓨터 판독 가능 저장 매체(16)에 저장된 하나 이상의 프로그램들을 실행할 수 있다. 상기 하나 이상의 프로그램들은 하나 이상의 컴퓨터 실행 가능 명령어를 포함할 수 있으며, 상기 컴퓨터 실행 가능 명령어는 프로세서(14)에 의해 실행되는 경우 컴퓨팅 장치(12)로 하여금 예시적인 실시예에 따른 동작들을 수행하도록 구성될 수 있다.

[0275] 컴퓨터 판독 가능 저장 매체(16)는 컴퓨터 실행 가능 명령어 내지 프로그램 코드, 프로그램 데이터 및/또는 다른 적합한 형태의 정보를 저장하도록 구성된다. 컴퓨터 판독 가능 저장 매체(16)에 저장된 프로그램(20)은 프로세서(14)에 의해 실행 가능한 명령어의 집합을 포함한다. 일 실시예에서, 컴퓨터 판독 가능 저장 매체(16)는 메모리(랜덤 액세스 메모리와 같은 휘발성 메모리, 비휘발성 메모리, 또는 이들의 적절한 조합), 하나 이상의 자기 디스크 저장 디바이스들, 광학 디스크 저장 디바이스들, 플래시 메모리 디바이스들, 그 밖에 컴퓨팅 장치(12)에 의해 액세스되고 원하는 정보를 저장할 수 있는 다른 형태의 저장 매체, 또는 이들의 적합한 조합일 수 있다.

[0276] 통신 버스(18)는 프로세서(14), 컴퓨터 판독 가능 저장 매체(16)를 포함하여 컴퓨팅 장치(12)의 다른 다양한 컴포넌트들을 상호 연결한다.

[0277] 컴퓨팅 장치(12)는 또한 하나 이상의 입출력 장치(24)를 위한 인터페이스를 제공하는 하나 이상의 입출력 인터페이스(22) 및 하나 이상의 네트워크 통신 인터페이스(26)를 포함할 수 있다. 입출력 인터페이스(22) 및 네트워크 통신 인터페이스(26)는 통신 버스(18)에 연결된다. 입출력 장치(24)는 입출력 인터페이스(22)를 통해 컴퓨팅 장치(12)의 다른 컴포넌트들에 연결될 수 있다. 예시적인 입출력 장치(24)는 포인팅 장치(마우스 또는 트랙패드 등), 키보드, 터치 입력 장치(터치패드 또는 터치스크린 등), 음성 또는 소리 입력 장치, 다양한 종류의 센서 장치 및/또는 촬영 장치와 같은 입력 장치, 및/또는 디스플레이 장치, 프린터, 스피커 및/또는 네트워크 카드와 같은 출력 장치를 포함할 수 있다. 예시적인 입출력 장치(24)는 컴퓨팅 장치(12)를 구성하는 일 컴포넌트로서 컴퓨팅 장치(12)의 내부에 포함될 수도 있고, 컴퓨팅 장치(12)와는 구별되는 별개의 장치로 컴퓨팅 장치(12)와 연결될 수도 있다.

[0279] 이상에서 본 발명의 대표적인 실시예들을 상세하게 설명하였으나, 본 발명이 속하는 기술분야에서 통상의 지식을 가진 자는 상술한 실시예에 대하여 본 발명의 범주에서 벗어나지 않는 한도 내에서 다양한 변형이 가능함을 이해할 것이다. 그러므로 본 발명의 권리범위는 설명된 실시예에 국한되어 정해져서는 안 되며, 후술하는 특허

청구범위뿐만 아니라 이 특허청구범위와 균등한 것들에 의해 정해져야 한다.

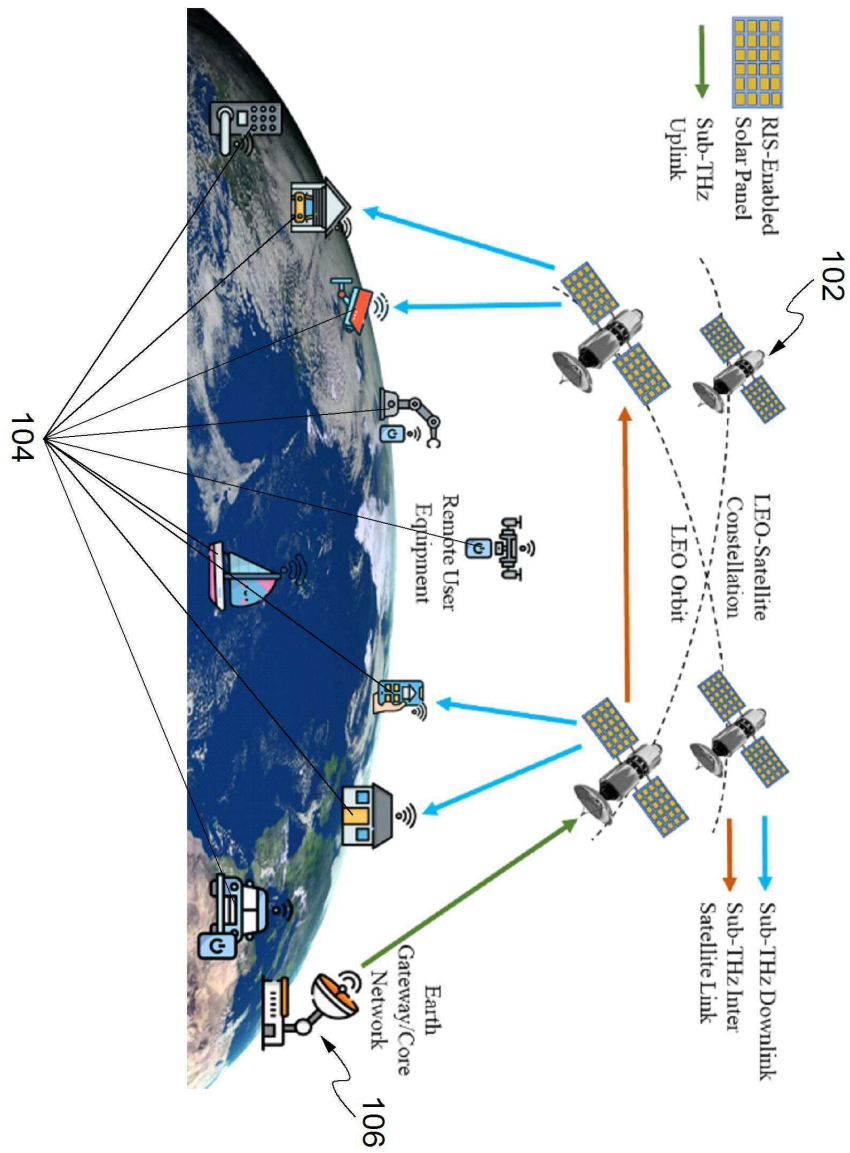
부호의 설명

[0281]

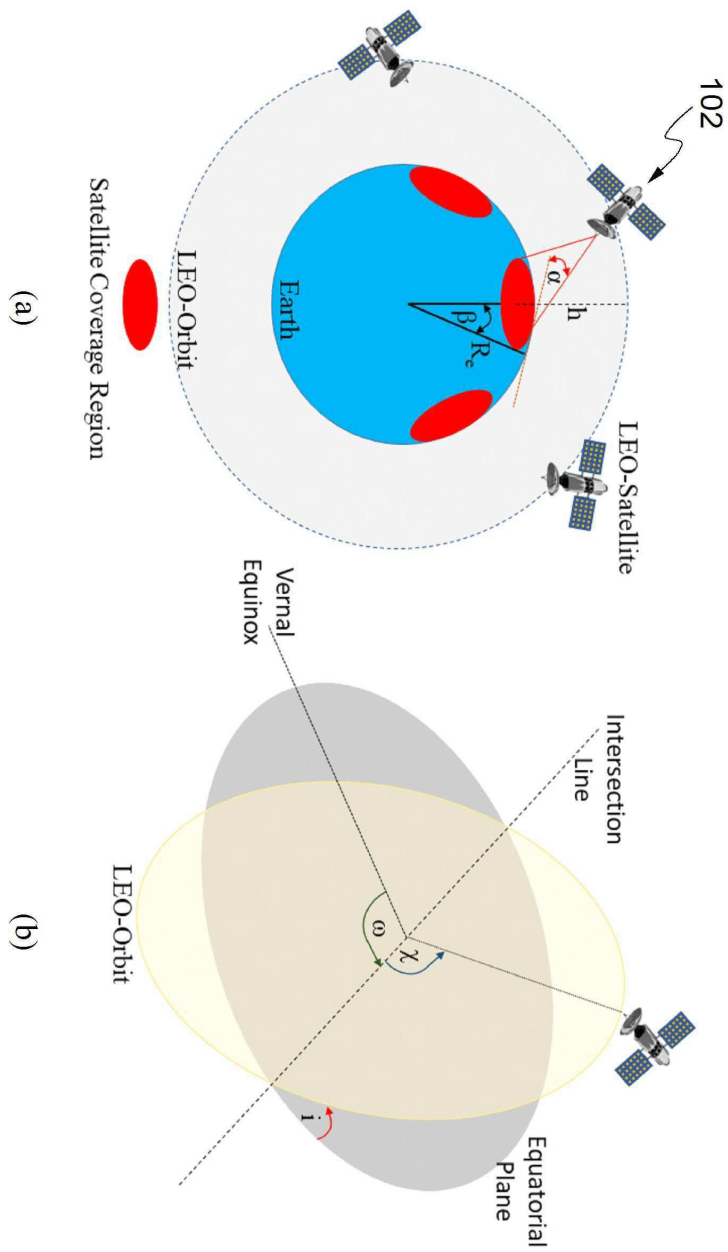
- 10 : 컴퓨팅 환경
- 12 : 컴퓨팅 장치
- 14 : 프로세서
- 16 : 컴퓨터 판독 가능 저장 매체
- 18 : 통신 버스
- 20 : 프로그램
- 22 : 입출력 인터페이스
- 24 : 입출력 장치
- 26 : 네트워크 통신 인터페이스
- 100 : 통신 시스템
- 102 : 위성
- 104 : 원격 사용자 단말
- 106 : 기지국 장치

100

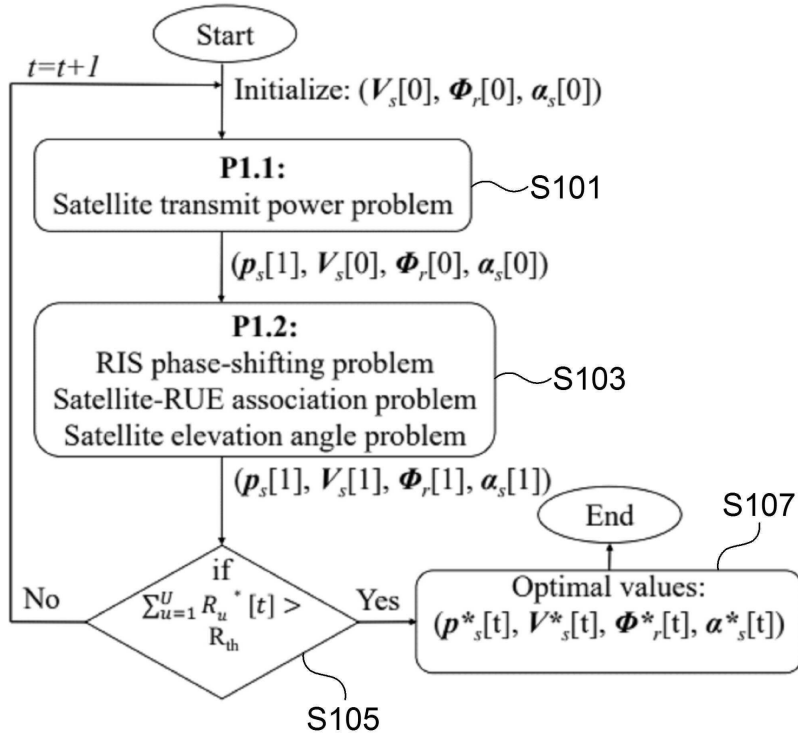
도면
도면1



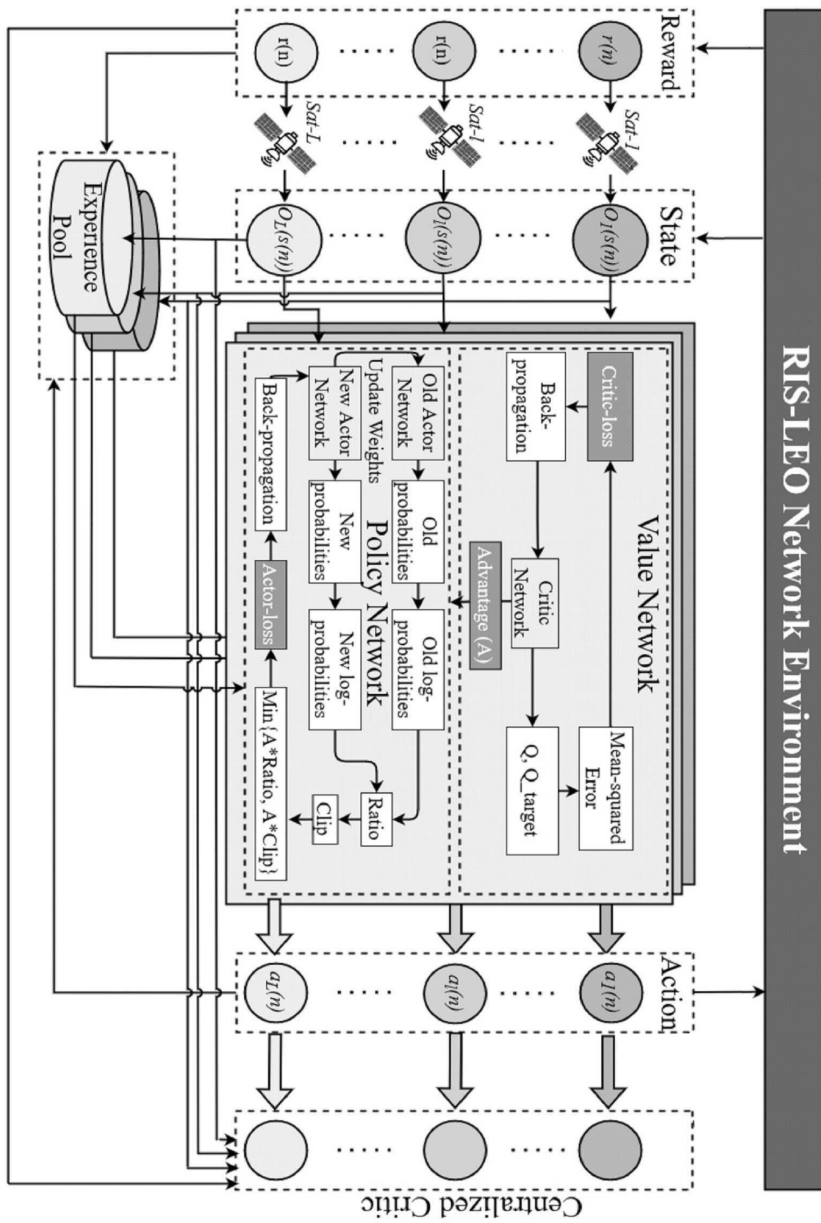
도면2



도면3



도면4



도면5

10

