



(19) 대한민국특허청(KR)
(12) 공개특허공보(A)

(11) 공개번호 10-2021-0069549
(43) 공개일자 2021년06월11일

(51) 국제특허분류(Int. Cl.)
H04W 24/02 (2009.01) H04W 84/06 (2019.01)
H04W 88/18 (2019.01)
(52) CPC특허분류
H04W 24/02 (2013.01)
H04W 84/06 (2019.01)
(21) 출원번호 10-2020-0106062
(22) 출원일자 2020년08월24일
심사청구일자 2020년08월24일
(30) 우선권주장
1020190158688 2019년12월03일 대한민국(KR)

(71) 출원인
경희대학교 산학협력단
경기도 용인시 기흥구 덕영대로 1732 (서천동, 경희대학교 국제캠퍼스내)
(72) 발명자
홍충선
경기도 용인시 수지구 상현로 30-10, 233동 101호(상현동, 상현마을성원상패빌아파트)
사르더 파쿠를 아베딘
경기도 용인시 기흥구 덕영대로 1732
이민경
경기도 수원시 영통구 청명남로 25, 1332호(영통동)
(74) 대리인
특허법인도담

전체 청구항 수 : 총 18 항

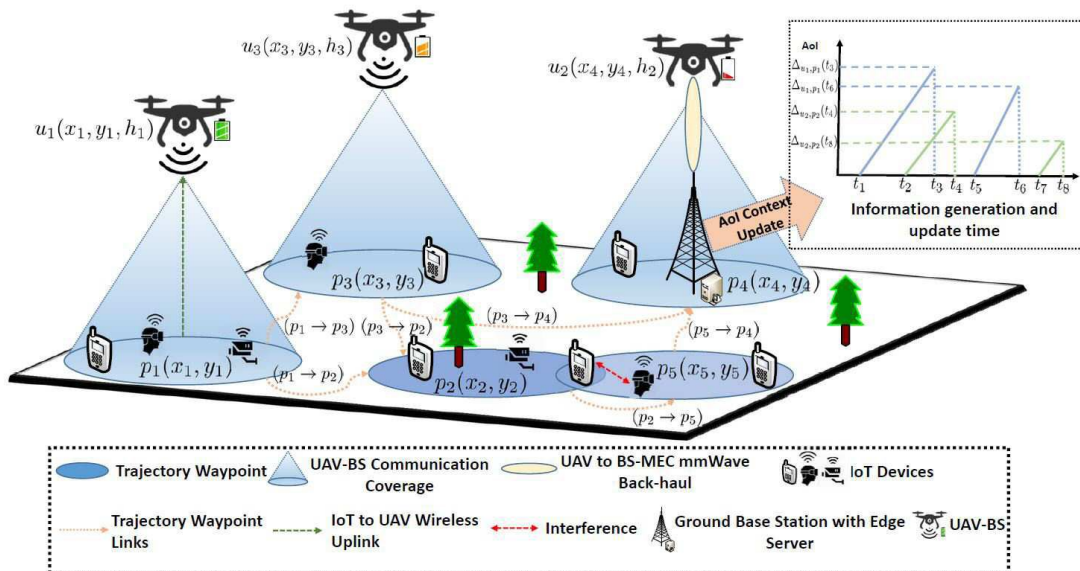
(54) 발명의 명칭 다중 무인항공기 네비게이션 최적화 방법 및 그를 이용한 무인항공기 시스템

(57) 요약

본 발명의 일 기술적 측면에 따른 다중 무인항공기 네비게이션 최적화 방법은, 배터리로 구동되어 이동하며 주어진 궤도 지점 세트를 커버하는 무인항공기기지국과 연동하여 동작하는 지상 기지국에서 수행되는 다중 무인항공기 네비게이션 최적화 방법으로서, 상기 지상 기지국이 무인항공기 기지국의 전송 범위 내에 존재하면, 통신을

(뒷면에 계속)

대표도 - 도1



통하여 상기 무인항공기 기지국으로부터 정보 업데이트를 받고 정보연령 메트릭을 계산하는 단계, 각각의 무인항공기 기지국의 궤도, 에너지 효율 및 정보연령 조건을 설정하는 단계 및 상기 에너지 효율 및 상기 정보연령의 적용 받는 무인항공기 기지국 중계망의 총 에너지 효율을 최대화 하도록, 각각의 무인항공기 기지국의 궤도 경로 정책을 찾기 위한 강화 학습을 수행하는 단계를 포함한다.

본 발명에 따르면, 데이터 최신성을 측정할 수 있는 새로운 매트릭스인 정보 연령(Age of Informatin, AoI)를 설정하고 이를 이용함으로써, 원격 클라우드 환경을 위한 에지 컴퓨팅 환경을 제공하고, 그를 이용하여 컴퓨팅 지향 통신 애플리케이션을 가능하게 할 수 있는 효과가 있다.

(52) CPC특허분류

H04W 88/18 (2019.01)

이 발명을 지원한 국가연구개발사업

과제고유번호	1711103015
과제번호	2019-0-01287
부처명	과학기술정보통신부
과제관리(전문)기관명	정보통신기획평가원
연구사업명	SW컴퓨팅산업원천기술개발
연구과제명	(SW 스타랩)분산 엣지를 위한 진화형 딥러닝 모델생성 플랫폼
기 여 율	1/1
과제수행기관명	경희대학교 산학협력단
연구기간	2020.01.01 ~ 2020.12.31

명세서

청구범위

청구항 1

배터리로 구동되어 이동하며 주어진 궤도 지점 세트를 커버하는 무인항공기기지국과 연동하여 동작하는 지상 기지국에서 수행되는 다중 무인항공기 네비게이션 최적화 방법으로서,

상기 지상 기지국이 무인항공기 기지국의 전송 범위 내에 존재하면, 통신을 통하여 상기 무인항공기 기지국으로부터 정보 업데이트를 받고 정보연령 메트릭을 계산하는 단계;

각각의 무인항공기 기지국의 궤도, 에너지 효율 및 정보연령 조건을 설정하는 단계; 및

상기 에너지 효율 및 상기 정보연령의 적용 받는 무인항공기 기지국 중계망의 총 에너지 효율을 최대화 하도록, 각각의 무인항공기 기지국의 궤도 경로 정책을 찾기 위한 강화 학습을 수행하는 단계;

를 포함하는 다중 무인항공기 네비게이션 최적화 방법.

청구항 2

제1항에 있어서, 다중 무인항공기 네비게이션 최적화 방법은,

상기 정보연령 메트릭을 계산하는 단계 이전에,

상기 무인항공기 기지국으로부터 제공된 수신전력을 계산하고, 상기 수신전력을 이용하여 상기 무인항공기 기지국과 상기 지상 기지국 간의 백홀 용량을 계산하여 백홀 링크를 생성하고, 시간 t에서 백홀 링크를 사용하는 동안 상기 무인항공기 기지국의 전송 에너지를 계산하는 단계;

를 더 포함하는 다중 무인항공기 네비게이션 최적화 방법.

청구항 3

제2항에 있어서, 상기 수신전력은

수학식

$$\hat{P}_{b,u} = P_{b,u}^{tx} \cdot G_u^{tx} \cdot G_b^{rx} \left(\frac{c}{4\pi\delta_{b,u}f_c^{mmWave}} \right).$$

에 따라 계산되고,

여기에서, 상기 $P_{b,u}^{tx}$ 는 무인항공기 기지국(UAV-BS) u로부터의 지상 기지국(BS) b로의 송신 전력, 상기 $\delta_{b,u}$ 는 무인항공기 기지국(UAV-BS) u과 지상 기지국(BS) b 간의 거리, 상기 c는 빛의 속도, 상기 f_c^{mmWave} 는 mmWave 백홀 링크의 반송파 주파수(carrier frequency), 상기 G_u^{tx} 및 상기 G_b^{rx} 는 각각 송신기 UAV-BS u와 수신기 BS b의 안테나 이득인 다중 무인항공기 네비게이션 최적화 방법.

청구항 4

제2항에 있어서, 상기 백홀 용량은

수학식

$$r_{b,u}^{mmWave}(t) = \begin{cases} \beta_{b,u}^{mmWave} \cdot \log \left(1 + \frac{\hat{P}_{b,u}}{\beta_{b,u}^{mmWave} \sigma^2} \right), \\ \quad \text{if } \delta_{u,b} \leq \bar{\alpha}, \\ 0, \quad \text{otherwise.} \end{cases}$$

에 따라 계산되고,

여기에서, 상기 $\beta_{b,u}^{mmWave}$ 는 mmWave 백홀 대역폭, 상기 σ^2 는 첨가 소음인 다중 무인항공기 네비게이션 최적화 방법.

청구항 5

제1항에 있어서, 상기 각각의 무인항공기 기지국의 궤도, 에너지 효율 및 정보연령 조건을 설정하는 단계는 각각의 무인항공기 기지국이 주어진 시간대에서 궤도 지점의 하위 집합만 다룰 수 있도록 설정하는 단계; 및 상기 궤도 지점으로 구성된 각각의 무인항공기 기지국에 대한 서브셋을 제한요건에 만족하도록 설정하는 단계; 를 포함하는 다중 무인항공기 네비게이션 최적화 방법.

청구항 6

제5항에 있어서, 상기 각각의 무인항공기 기지국의 궤도, 에너지 효율 및 정보연령 조건을 설정하는 단계는 정보 업데이트가 발생하는 지상 기지국 궤도 지점은 제외한, 무인항공기 기지국의 오버랩되지 않는 궤도를 상기 무인항공기 기지국의 궤도 조건으로 설정하는 단계; 모든 궤도 지점이 상호 의존적으로 포함된 무인항공기 기지국의 공동 궤도를 구성하는 단계; 통신 및 이동 에너지가 최소 에너지 효율 임계값 보다 크도록 무인항공기 기지국의 총 에너지 효율을 보장하는 단계; 및 정보 업데이트의 평균 최신성이 정보연령 임계값 미만을 만족하도록 보장하는 단계; 를 더 포함하는 다중 무인항공기 네비게이션 최적화 방법.

청구항 7

제1항에 있어서, 상기 지상 기지국은 무인항공기 기지국의 궤도 정책을 위한 4차원 상태 공간을 설정하고, 상기 무인항공기 기지국은 각각 궤도 및 통신제약을 충족하면서 하나의 실행 가능한 상태에서 다음 상태로의 행해를 설정하여 행동 공간을 설정하고, 상기 궤도 경로 정책을 찾기 위한 강화 학습을 수행하는 단계는, 임의의 상태에서 복수의 무인항공기 기지국이 각각 설정한 복수의 행동 공간에 대하여 경유지의 충돌 여부를 기준으로 보상을 설정하는 단계; 를 포함하는 다중 무인항공기 네비게이션 최적화 방법.

청구항 8

제7항에 있어서, 상기 경유지의 충돌 여부를 기준으로 보상을 설정하는 단계는,

무인항공기 기지국 네비게이션에 필요한 상황별 특징을 추출하는 데 사용되는 상태 공간의 차수를 줄이는 단계;

무인항공기 기지국 에이전트가 관찰하는 상태 전환을 저장하는 단계; 및

에너지 및 정보연령 지표의 적용을 받는 무인항공기 기지국 중계망의 총 에너지 효율을 최대화 하도록 무인항공기 기지국의 궤도 경로 정책을 찾기 위한 강화 학습을 수행하는 단계;

를 포함하는 다중 무인항공기 네비게이션 최적화 방법.

청구항 9

제8항에 있어서, 상기 강화 학습을 수행하는 단계는,

수학식

$$\hat{R}(s, a; t) = \sum_{t_0=0}^T \gamma \times R_{t-t_0}(a_t),$$

를 이용하여 미래 보상을 정의하는 단계;

를 포함하고,

여기에서, $\gamma \in [0, 1]$ 을 만족하는 다중 무인항공기 네비게이션 최적화 방법.

청구항 10

주어진 궤도 지점 세트에 포함되는 서로 다른 궤도 지점에 랜덤하게 위치하는 IoT(Internet on Things) 디바이스를 포함하는 IoT 디바이스 세트;

배터리로 구동되어 이동하며 상기 주어진 궤도 지점 세트를 커버하고, 상기 IoT 디바이스 세트의 릴레이 역할을 수행하는 무인항공기 기지국을 복수개 포함하는 무인항공기 세트; 및

무인항공기 세트에 의한 릴레이를 통해 상기 IoT 디바이스로부터 정보 업데이트를 받아 연산 지향 통신 애플리케이션을 지원하는 지상 기지국

을 포함하는 무인항공기 시스템.

청구항 11

제10항에 있어서, 상기 무인항공기 기지국은

유한한 시간에 걸쳐 상기 궤도 지점 세트에 포함된 서로 다른 궤도 지점을 통과하고, 업링크 통신 채널을 이용하여 임의의 궤도 지점 주변에 위치한 활성 IoT 디바이스로부터 정보 데이터 패킷을 수집하는 무인항공기 시스템.

청구항 12

제10항에 있어서, 상기 무인항공기 기지국은

상기 무인항공기 기지국의 전송 범위 내에 지상 기지국이 존재하면, 해당 지상 기지국과 백홀 통신 링크를 생성하고, 생성된 백홀 통신 링크를 통하여 수집한 정보 데이터 패킷을 상기 지상 기지국에 전송하는 무인항공기 시스템.

청구항 13

제12항에 있어서, 상기 지상 기지국은

멀티 액세스 에지 컴퓨팅 서버를 포함하고,

상기 멀티 액세스 에지 컴퓨팅 서버는

상기 무인항공기 기지국으로부터 전송된 정보 데이터 패킷을 이용하여 다른 궤도 지점에 대한 정보 업데이트를 수행하고 정보연령 메트릭을 계산하는 무인항공항 시스템.

청구항 14

제11항에 있어서,

상기 활성 IoT 디바이스와 상기 무인항공기지국 간의 공대지 경로 손실 확률은 아래의 수학식으로 정의되고,

$$\zeta_{i,p}^u = \begin{cases} \frac{1}{1+\alpha \exp(-\hat{\alpha}(\frac{180}{\pi}\Theta_u-\alpha))}, & \text{LoS channel,} \\ 1 - \left[\frac{1}{1+\alpha \exp(-\hat{\alpha}(\frac{180}{\pi}\Theta_u-\alpha))} \right], & \text{NLoS channel.} \end{cases}$$

여기에서, α 와 $\hat{\alpha}$ 는 각각 LoS 채널 및 NLoS 채널에 대한 환경의존상수이고, Θ_u 는 상기 무인항공기 기지국의 표고각인 무인항공항 시스템.

청구항 15

제11항에 있어서,

시간 경과 T에 따라 IoT 디바이스들을 서비스하기 위해 궤도지점 P를 커버하는 상기 무인항공기 기지국의 총 에너지 효율은 아래의 수학식으로 정의되고,

$$\eta(\mathcal{P}, u) = \sum_{t=1}^T \sum_{p=1}^{|\mathcal{P}|} \frac{(r_{b,u}^{mmWave}(t) + \sum_{i=1}^{|\mathcal{I}|} r_{i,p}^u(t))}{(E_u^{mmWave}(t) + E_u(t))}.$$

여기에서, $r_{b,u}^{mmWave}(t)$ 는 시간 t에서 상기 무인항공기 기지국으로부터 지상 기지국 간의 채널의 백홀 용량이고,

$E_u^{mmWave}(t)$ 는 시간 t에서 백홀 링크를 사용하는 동안 상기 무인항공기 기지국의 전송 에너지인 무인항공항 시스템.

청구항 16

제10항에 있어서, 상기 지상 기지국은

시간 t에서의 궤도 p의 정보연령을 아래의 수학식을 이용하여 계산하고,

$$\Delta_u(\mathcal{P}, t) = t - \Delta'_u(p, t), \forall p \in \mathcal{P}.$$

여기에서, $\Delta'_u(p, t)$ 는 상기 지상 기지국에서, 상기 무인항공기 기지국이 궤도 지점 p에서 가장 최근에 수신된 데이터 패킷의 시간 스탬프로 표시되는 무인항공항 시스템.

청구항 17

제10항에 있어서, 상기 지상 기지국은

무인항공기 기지국의 궤도 정책을 위한 4차원 상태 공간을 설정하고,

상기 무인항공기 기지국은 각각 궤도 및 통신제약을 충족하면서 하나의 실행 가능한 상태에서 다음 상태로의 항해를 설정하여 행동 공간을 설정하고,

상기 지상 기지국은

임의의 상태에서 복수의 무인항공기 기지국이 각각 설정한 복수의 행동 공간에 대하여 경유지의 충돌 여부를 기준으로 보상을 설정하는 무인항공항 시스템.

청구항 18

제17항에 있어서, 상기 지상 기지국은

심층강화학습 모델을 생성하고,

상기 심층강화학습 모델은

무인항공기 기지국 네비게이션에 필요한 상황별 특징을 추출하는 데 사용되는 상태 공간의 차수를 줄이는 심층 신경 네트워크;

무인항공기 기지국 에이전트가 관찰하는 상태 전환을 저장하는 경험 재생 메모리; 및

에너지 및 정보연령 지표의 적용을 받는 무인항공기 기지국 중계망의 총 에너지 효율을 최대화 하도록 무인항공기 기지국의 궤도 경로 정책을 찾기 위한 강화 학습을 수행하는 강화 학습 프레임워크

를 포함하는 무인항공항 시스템.

발명의 설명

기술 분야

[0001] 본 발명은 다중 무인항공기 네비게이션 최적화 방법 및 그를 이용한 무인항공항 시스템에 관한 것이다.

배경 기술

[0003] 고 세대 무선 네트워크, 예를 들어, 제5 세대 무선 네트워크의 신속한 전개는 고품질 무선 연결 및 서비스에 대하여 비할 데 없는 높은 기준을 요구한다.

[0004] 그 예로, 무선 네트워크는 향상된 모바일 광대역통신, 초신뢰성 및 저지연성 통신, 대규모 기계형 통신 애플리케이션 등의 서로 다른 5G 애플리케이션 유형에 대한 엄격한 요구사항을 충족해야 하는 어려움에 직면해 있다.

[0006] 이러한 엄격한 요구 사항은 기존의 네트워크 인프라, 즉, 기존의 기지국 인프라로서는 달성하기 어려운 한계점이 있다. 따라서, 기존의 무선 네트워크 인프라를 보완하기 위하여, 무인항공기를 공중 기지국으로서 서비스하도록 하는 네트워크 인프라에 대한 개발이 이루어지고 있다.

[0008] 무인항공기가 네트워크 애플리케이션과 사용자를 대상으로 공중 기지국으로서 서비스하는 경우, 기존의 고정식 무선 네트워크 인프라와 달리 보다 민첩하고, 커버리지가 넓으며, 지연 제약이 적은 대용량이라는 엄격한 요구를 충족시킬 수 있다.

[0010] 그러나, 이러한 무인항공기의 경우 이동 및 기지국으로서의 기능을 수행하기 위한 에너지의 용량이 제한되어 있어, 무선 네트워크에서의 복수의 무인항공기를 이용한 기지국의 구축과 자율적 탐색은 여전히 난항을 겪고 있다.

선행기술문헌

비특허문헌

[0012] (비특허문헌 0001) S. F. Abedin, M. G. R. Alam, S. A. Kazmi, N. H. Tran, D. Niyato, and C. S. Hong, "Resource allocation for ultra-reliable and enhanced mobile broadband iot applications in fog network," IEEE Transactions on Communications, vol. 67, no. 1, pp. 489-502, 2018.

발명의 내용

해결하려는 과제

[0013] 본 발명의 일 기술적 측면은 상기한 종래 기술의 문제점을 해결하기 위한 것으로써, 데이터 최신성을 측정할 수 있는 새로운 매트릭스인 정보 연령(Age of Informatin, AoI)를 설정하고 이를 이용함으로써, 원격 클라우드 환경을 위한 에지 컴퓨팅 환경을 제공하고, 그를 이용하여 컴퓨팅 지향 통신 애플리케이션을 가능하게 하는 다중 무인항공기 네비게이션 최적화 방법 및 그를 이용한 무인항공기 시스템을 제공하는 것이다.

[0014] 또한, 본 발명의 일 기술적 측면은, 에지 컴퓨팅 네트워크에서 에너지 효율을 보장하는 다중 무인항공기 네비게이션 문제에 대하여, 딥 큐 네트워크(Deep Q-Network)를 채용하여 상황별 제약하에 에너지 효율이 높은 다중 무인항공기 네비게이션 네비게이션을 제공하는 다중 무인항공기 네비게이션 최적화 방법 및 그를 이용한 무인항공기 시스템을 제공하는 것이다.

[0016] 본 발명의 상기 목적과 여러 가지 장점은 이 기술분야에 숙련된 사람들에 의해 본 발명의 바람직한 실시예로부터 더욱 명확하게 될 것이다.

과제의 해결 수단

[0018] 본 발명의 일 기술적 측면은 다중 무인항공기 네비게이션 최적화 방법을 제안한다. 상기 다중 무인항공기 네비게이션 최적화 방법은, 배터리로 구동되어 이동하며 주어진 궤도 지점 세트를 커버하는 무인항공기 기지국과 연동하여 동작하는 지상 기지국에서 수행되는 다중 무인항공기 네비게이션 최적화 방법으로서, 상기 지상 기지국이 무인항공기 기지국의 전송 범위 내에 존재하면, 통신을 통하여 상기 무인항공기 기지국으로부터 정보 업데이트를 받고 정보연령 매트릭을 계산하는 단계, 각각의 무인항공기 기지국의 궤도, 에너지 효율 및 정보연령 조건을 설정하는 단계 및 상기 에너지 효율 및 상기 정보연령의 적용 받는 무인항공기 기지국 중계망의 총 에너지 효율을 최대화 하도록, 각각의 무인항공기 기지국의 궤도 경로 정책을 찾기 위한 강화 학습을 수행하는 단계를 포함한다.

[0019] 일 실시예에서, 상기 다중 무인항공기 네비게이션 최적화 방법은, 상기 정보연령 매트릭을 계산하는 단계 이전에, 상기 무인항공기 기지국으로부터 제공된 수신전력을 계산하고, 상기 수신전력을 이용하여 상기 무인항공기 기지국과 상기 지상 기지국 간의 백홀 용량을 계산하여 백홀 링크를 생성하고, 시간 t에서 백홀 링크를 사용하는 동안 상기 무인항공기 기지국의 전송 에너지를 계산하는 단계를 더 포함할 수 있다.

[0020] 일 실시예에서, 상기 수신전력은 수학식

$$\hat{P}_{b,u} = P_{b,u}^{tx} \cdot G_u^{tx} \cdot G_b^{rx} \left(\frac{c}{4\pi\delta_{b,u}f_c^{mmWave}} \right). \quad [0021]$$

에 따라 계산되고, 여기에서, 상기 $P_{b,u}^{tx}$ 는 무인항공기 기지국(UAV-BS) u로부터의 지상 기지국(BS) b로의 송신 전력, 상기 $\delta_{b,u}$ 는 무인항공기 기지국(UAV-BS) u과 지상 기지국(BS) b 간의 거리, 상기 c는 빛의 속도, 상기 f_c^{mmWave} 는 mmWave 백홀 링크의 반송파 주파수(carrier frequency), 상기 G_u^{tx} 및 상기 G_b^{rx} 는 각각 송신기 UAV-BS u와 수신기 BS b의 안테나 이득이다.

[0022] 일 실시예에서, 상기 백홀 용량은 수학식

$$r_{b,u}^{mmWave}(t) = \begin{cases} \beta_{b,u}^{mmWave} \cdot \log \left(1 + \frac{\hat{P}_{b,u}}{\beta_{b,u}^{mmWave}\sigma^2} \right), \\ \text{if } \delta_{u,b} \leq \bar{\alpha}, \\ 0, \text{ otherwise.} \end{cases} \quad [0023]$$

에 따라 계산되고, 여기에서, 상기 $\beta_{b,u}^{mmWave}$ 는 mmWave 백홀 대역폭, 상기 σ^2 는 첨가 소음이다.

[0024] 일 실시예에서, 상기 각각의 무인항공기 기지국의 궤도, 에너지 효율 및 정보연령 조건을 설정하는 단계는, 각각의 무인항공기 기지국이 주어진 시간대에서 궤도 지점의 하위 집합만 다룰 수 있도록 설정하는 단계 및 상기 궤도 지점으로 구성된 각각의 무인항공기 기지국에 대한 서브셋을 제한요건에 만족하도록 설정하는 단계를 포함 할 수 있다.

[0025] 일 실시예에서, 상기 각각의 무인항공기 기지국의 궤도, 에너지 효율 및 정보연령 조건을 설정하는 단계는, 정보 업데이트가 발생하는 지상 기지국 궤도 지점은 제외한, 무인항공기 기지국의 오버랩되지 않는 궤도를 상기 무인항공기 기지국의 궤도 조건으로 설정하는 단계, 모든 궤도 지점이 상호 의존적으로 포함된 무인항공기 기지국의 공동 궤도를 구성하는 단계, 통신 및 이동 에너지가 최소 에너지 효율 임계값 보다 크도록 무인항공기 기지국의 총 에너지 효율을 보장하는 단계 및 정보 업데이트의 평균 최신성이 정보연령 임계값 미만을 만족하도록 보장하는 단계를 더 포함 할 수 있다.

[0026] 일 실시예에서, 상기 지상 기지국은 무인항공기 기지국의 궤도 정책을 위한 4차원 상태 공간을 설정하고, 상기 무인항공기 기지국은 각각 궤도 및 통신제약을 충족하면서 하나의 실행 가능한 상태에서 다음 상태로의 항해를 설정하여 행동 공간을 설정하고, 상기 궤도 경로 정책을 찾기 위한 강화 학습을 수행하는 단계는, 임의의 상태에서 복수의 무인항공기 기지국이 각각 설정한 복수의 행동 공간에 대하여 경유지의 충돌 여부를 기준으로 보상을 설정하는 단계를 포함 할 수 있다.

[0027] 일 실시예에서, 상기 경유지의 충돌 여부를 기준으로 보상을 설정하는 단계는, 무인항공기 기지국 네비게이션에 필요한 상황별 특징을 추출하는 데 사용되는 상태 공간의 차수를 줄이는 단계, 무인항공기 기지국 에이전트가 관찰하는 상태 전환을 저장하는 단계 및 에너지 및 정보연령 지표의 적용을 받는 무인항공기 기지국 중계망의 총 에너지 효율을 최대화 하도록 무인항공기 기지국의 궤도 경로 정책을 찾기 위한 강화 학습을 수행하는 단계를 포함 할 수 있다.

[0028] 일 실시예에서, 상기 강화 학습을 수행하는 단계는, 수학식

$$\hat{R}(s, a; t) = \sum_{t_0=0}^T \gamma \times R_{t-t_0}(a_t), \quad [0029]$$

를 이용하여 미래 보상을 정의하는 단계를 포함하고, 여기에서, $\gamma \in [0, 1]$ 을 만족 할 수 있다.

[0031] 본 발명의 다른 일 기술적 측면은 무인항공기 시스템을 제안한다. 상기 무인항공기 시스템은, 주어진 궤도 지점 세트에 포함되는 서로 다른 궤도 지점에 랜덤하게 위치하는 IoT(Internet on Things) 디바이스를 포함하는 IoT

디바이스 세트, 배터리로 구동되어 이동하며 상기 주어진 궤도 지점 세트를 커버하고, 상기 IoT 디바이스 세트의 릴레이 역할을 수행하는 무인항공기 기지국을 복수개 포함하는 무인항공기 세트 및 무인항공기 세트에 의한 릴레이를 통해 상기 IoT 디바이스로부터 정보 업데이트를 받아 연산 지향 통신 애플리케이션을 지원하는 지상 기지국을 포함한다.

[0032] 일 실시예에서, 상기 무인항공기 기지국은, 유한한 시간에 걸쳐 상기 궤도 지점 세트에 포함된 서로 다른 궤도 지점을 통과하고, 업링크 통신 채널을 이용하여 임의의 궤도 지점 주변에 위치한 활성 IoT 디바이스로부터 정보 데이터 패킷을 수집할 수 있다.

[0033] 일 실시예에서, 상기 무인항공기 기지국은, 상기 무인항공기 기지국의 전송 범위 내에 지상 기지국이 존재하면, 해당 지상 기지국과 백홀 통신 링크를 생성하고, 생성된 백홀 통신 링크를 통하여 수집한 정보 데이터 패킷을 상기 지상 기지국에 전송 할 수 있다.

[0034] 일 실시예에서, 상기 지상 기지국은 멀티 액세스 에지 컴퓨팅 서버를 포함하고, 상기 멀티 액세스 에지 컴퓨팅 서버는 상기 무인항공기 기지국으로부터 전송된 정보 데이터 패킷을 이용하여 다른 궤도 지점에 대한 정보 업데이트를 수행하고 정보연령 메트릭을 계산 할 수 있다.

[0035] 일 실시예에서, 상기 활성 IoT 디바이스와 상기 무인항공기 기지국 간의 공대지 경로 손실 확률은 아래의 수학적식으로 정의되고,

$$\zeta_{i,p}^u = \begin{cases} \frac{1}{1+\alpha \exp(-\hat{\alpha}(\frac{180}{\pi}\Theta_u-\alpha))}, & \text{LoS channel,} \\ 1 - \left[\frac{1}{1+\alpha \exp(-\hat{\alpha}(\frac{180}{\pi}\Theta_u-\alpha))} \right], & \text{NLoS channel.} \end{cases}$$

[0037] 여기에서, α 와 $\hat{\alpha}$ 는 각각 LoS 채널 및 NLoS 채널에 대한 환경의존상수이고, Θ_u 는 상기 무인항공기 기지국의 표고각이다.

[0038] 일 실시예에서, 시간 경과 T에 따라 IoT 디바이스들을 서비스하기 위해 궤도지점 P를 커버하는 상기 무인항공기 기지국의 총 에너지 효율은 아래의 수학적식으로 정의되고,

$$\eta(\mathcal{P}, u) = \sum_{t=1}^T \sum_{p=1}^{|\mathcal{P}|} \frac{(r_{b,u}^{mmWave}(t) + \sum_{i=1}^{|\mathcal{I}|} r_{i,p}^u(t))}{(E_u^{mmWave}(t) + E_u(t))}.$$

[0040] 여기에서, $r_{b,u}^{mmWave}(t)$ 는 시간 t에서 상기 무인항공기 기지국으로부터 지상 기지국 간의 채널의 백홀 용량이고, $E_u^{mmWave}(t)$ 는 시간 t에서 백홀 링크를 사용하는 동안 상기 무인항공기 기지국의 전송 에너지 이다.

[0041] 일 실시예에서, 상기 지상 기지국은, 시간 t에서의 궤도 p의 정보연령을 아래의 수학적식을 이용하여 계산하고,

$$\Delta_u(\mathcal{P}, t) = t - \Delta'_u(p, t), \forall p \in \mathcal{P}.$$

[0043] 여기에서, $\Delta'_u(p, t)$ 는 상기 지상 기지국에서, 상기 무인항공기 기지국이 궤도 지점 p에서 가장 최근에 수신된 데이터 패킷의 시간 스탬프로 표시될 수 있다.

[0044] 일 실시예에서, 상기 지상 기지국은 무인항공기 기지국의 궤도 정책을 위한 4차원 상태 공간을 설정하고, 상기 무인항공기 기지국은 각각 궤도 및 통신제약을 충족하면서 하나의 실행 가능한 상태에서 다음 상태로의 항해를 설정하여 행동 공간을 설정하고, 상기 지상 기지국은 임의의 상태에서 복수의 무인항공기 기지국이 각각 설정한 복수의 행동 공간에 대하여 경유지의 충돌 여부를 기준으로 보상을 설정 할 수 있다.

[0045] 일 실시예에서, 상기 지상 기지국은 심층강화학습 모델을 생성하고, 상기 심층강화학습 모델은, 무인항공기 기지국 네비게이션에 필요한 상황별 특징을 추출하는 데 사용되는 상태 공간의 차수를 줄이는 심층 신경 네트워크, 무인항공기 기지국 에이전트가 관찰하는 상태 전환을 저장하는 경험 재생 메모리 및 에너지 및 정보 연령 지표의 적용을 받는 무인항공기 기지국 중계망의 총 에너지 효율을 최대화 하도록 무인항공기 기지국의 궤

도 경로 정책을 찾기 위한 강화 학습을 수행하는 강화 학습 프레임워크를 포함 할 수 있다.

[0047] 상기한 과제의 해결 수단은, 본 발명의 특징을 모두 열거한 것은 아니다. 본 발명의 과제 해결을 위한 다양한 수단은 이하의 상세한 설명의 구체적인 실시형태를 참조하여 보다 상세하게 이해될 수 있을 것이다.

발명의 효과

[0049] 본 발명의 일 실시형태에 따르면, 데이터 최신성을 측정할 수 있는 새로운 매트릭스인 정보 연령(Age of Information, AoI)를 설정하고 이를 이용함으로써, 원격 클라우드 환경을 위한 에지 컴퓨팅 환경을 제공하고, 그를 이용하여 컴퓨팅 지향 통신 애플리케이션을 가능하게 할 수 있는 효과가 있다.

[0050] 또한, 본 발명의 일 실시형태에 따르면, 에지 컴퓨팅 네트워크에서 에너지 효율을 보장하는 다중 무인항공기 네비게이션 문제에 대하여 딥 큐 네트워크(Deep Q-Network)를 채용함으로써, 상황별 제약하에 에너지 효율이 높은 다중 무인항공기 네비게이션 네비게이션을 제공 할 수 있는 효과가 있다.

도면의 간단한 설명

[0052] 도 1은 본 발명의 일 실시예에 따른 다중 무인항공기 네비게이션 최적화 방법이 적용된 무인항공기 시스템의 개념적인 설명을 위한 도면이다.

도 2는 멀티액세스 에지 컴퓨팅 서버를 구현하는 컴퓨팅 장치의 일 예시를 도시하는 도면이다.

도 3은 본 발명의 일 실시예에 따른 무인항공기 기지국을 설명하는 블록 구성도이다.

도 4는 지상 기지국의 멀티 액세스 에지 컴퓨팅 서버에 의하여 수행되는 심층강화학습(DEEP Q-LEARNING) 모델의 일 예를 도시하는 도면이다.

도 5는 본 발명의 일 실시예에 따른, UAV-BS 궤도 정책 최적화를 위한 경험 재생 딥 큐 네트워크 알고리즘을 설명하기 위한 도면이다.

도 6은 본 발명의 일 실시예에 따른, UAV-BS 궤도 정책 최적화를 위한 경험 재생을 통한 DQN 테스트를 위한 알고리즘을 설명하기 위한 도면이다.

도 7은 본 발명의 일 실시예에 대한 시뮬레이션을 위한 세팅을 설명하는 도면이다.

도 8은 본 발명의 일 시뮬레이션 결과에 따른, 시간 슬롯 수의 함수로 누적 보상 평가를 도시하는 그래프이다.

도 9는 시뮬레이션에 적용된 디스카운트 팩터와 정보연령(AoI) 임계값에 대한 예시를 도시하는 표를 도시하고 있다.

도 10은 본 발명에 대한 일 시뮬레이션 결과로서, 궤도 경유점의 수에 따른 평균 누적 보상을 도시하는 그래프이다.

도 11은 본 발명에 대한 일 시뮬레이션 결과로서, 다른 궤적 경유지 수에 대해 제안된 접근방식과 다른 접근방식의 평균 AoI를 비교하는 그래프이다.

도 12는 본 발명에 대한 일 시뮬레이션 결과로서, 다른 궤적 경유지 수에 대해 제안된 접근방식과 다른 접근방식의 평균 에너지 효율을 비교하는 그래프이다.

도 13는 본 발명에 대한 일 시뮬레이션 결과로서, 다른 궤적 경유지 수에 대해 제안된 접근방식과 다른 접근방식의 평균 대역폭 효율성을 비교하는 그래프이다.

도 14는 본 발명에 대한 일 시뮬레이션 결과로서, 다른 궤적 경유지 수에 대해 제안된 접근방식과 다른 접근방식의 평균 UAV 리소스 활용도를 비교하는 그래프이다.

도 15는 본 발명에 대한 일 시뮬레이션 결과로서, 상이한 AoI 임계 값 하에서 동작하는 UAV-BS의 평균 에너지 효율을 나타내는 그래프이다.

도 16은 본 발명의 일 실시예에 따른 다중 무인항공기 네비게이션 최적화 방법을 설명하는 순서도이다.

발명을 실시하기 위한 구체적인 내용

- [0053] 이하, 첨부된 도면을 참조하여 본 발명의 바람직한 실시 형태들을 설명한다.
- [0054] 그러나, 본 발명의 실시형태는 여러 가지 다른 형태로 변형될 수 있으며, 본 발명의 범위가 이하 설명하는 실시 형태로 한정되는 것은 아니다. 또한, 본 발명의 실시형태는 당해 기술분야에서 평균적인 지식을 가진 자에게 본 발명을 더욱 완전하게 설명하기 위해서 제공되는 것이다.
- [0055] 즉, 전술한 목적, 특징 및 장점은 첨부된 도면을 참조하여 상세하게 후술되며, 이에 따라 본 발명이 속하는 기술분야에서 통상의 지식을 가진 자가 본 발명의 기술적 사상을 용이하게 실시할 수 있을 것이다. 본 발명을 설명함에 있어서 본 발명과 관련된 공지 기술에 대한 구체적인 설명이 본 발명의 요지를 불필요하게 흐릴 수 있다고 판단되는 경우에는 상세한 설명을 생략한다. 이하, 첨부된 도면을 참조하여 본 발명에 따른 바람직한 실시예를 상세히 설명하기로 한다. 도면에서 동일한 참조부호는 동일 또는 유사한 구성요소를 가리키는 것으로 사용된다.
- [0056] 또한, 본 명세서에서 사용되는 단수의 표현은 문맥상 명백하게 다르게 뜻하지 않는 한, 복수의 표현을 포함한다. 본 출원에서, "구성된다" 또는 "포함한다" 등의 용어는 명세서 상에 기재된 여러 구성 요소들, 또는 여러 단계들을 반드시 모두 포함하는 것으로 해석되지 않아야 하며, 그 중 일부 구성 요소들 또는 일부 단계들은 포함되지 않을 수도 있고, 또는 추가적인 구성 요소 또는 단계들을 더 포함할 수 있는 것으로 해석되어야 한다.
- [0057] 또한, 이하에서 본 발명에 따른 시스템을 설명하기 위하여 다양한 구성요소 및 그의 하부 구성요소에 대하여 설명하고 있다. 이러한 구성요소 및 그의 하부 구성요소들은, 하드웨어, 소프트웨어 또는 이들의 조합 등 다양한 형태로서 구현될 수 있다. 예컨대, 각 요소들은 해당 기능을 수행하기 위한 전자적 구성으로 구현되거나, 또는 전자적 시스템에서 구동 가능한 소프트웨어 자체이거나 그러한 소프트웨어의 일 기능적인 요소로 구현될 수 있다. 또는, 전자적 구성과 그에 대응되는 구동 소프트웨어로 구현될 수 있다.
- [0058] 본 명세서에 설명된 다양한 기법은 하드웨어 또는 소프트웨어와 함께 구현되거나, 적합한 경우에 이들 모두의 조합과 함께 구현될 수 있다. 본 명세서에 사용된 바와 같은 "부(Unit)", "서버(Server)" 및 "시스템(System)" 등의 용어는 마찬가지로 컴퓨터 관련 엔티티(Entity), 즉 하드웨어, 하드웨어 및 소프트웨어의 조합, 소프트웨어 또는 실행 시의 소프트웨어와 등가로 취급할 수 있다. 또한, 본 발명의 시스템에서 실행되는 각 기능은 모듈 단위로 구성될 수 있고, 하나의 물리적 메모리에 기록되거나, 둘 이상의 메모리 및 기록매체 사이에 분산되어 기록될 수 있다.
- [0059] 본 발명의 실시형태를 설명하기 위하여 다양한 순서도가 개시되고 있으나, 이는 각 단계의 설명의 편의를 위한 것으로, 반드시 순서도의 순서에 따라 각 단계가 수행되는 것은 아니다. 즉, 순서도에서의 각 단계는, 서로 동시에 수행되거나, 순서도에 따른 순서대로 수행되거나, 또는 순서도에서의 순서와 반대의 순서로도 수행될 수 있다.
- [0061] 본 출원에서는 데이터 최신성과 사물인터넷(IoT) 기기와의 연결을 개선하기 위해 이동 기지국(BS, Base Station)이 배치된 다중 무인항공기(UAV, Unmanned Aerial Vehicles)에 대한 네비게이션 기술에 대하여 기술한다. 이를 기술하기 위하여 아래의 특징들을 포함하여 설명한다.
- [0062] 첫째, 다중 무인항공기 기지국(UAV-BS, Unmanned Aerial Vehicles- Base Station) 궤도 정책을 최적화하여 에너지 효율을 극대화하는 것이 목표로 하여, 다중 무인항공기 기지국(UAV-BS)의 에너지 효율 궤도 최적화 문제를 공식화한다.
- [0063] 또한 에너지 및 정보 연령(AoI, Age of Information) 제약과 같은 서로 다른 문맥 정보를 통합하여 지상 기지국(BS)에서 데이터 최신성을 보장한다.
- [0064] 둘째, 다중 무인항공기 기지국(UAV-BS) 네비게이션의 상황적 제약에 관한 공식화된 문제를 해결하기 위해 경험 재생 모델을 이용한 심층 보강 학습을 개시한다. 즉, 딥보강 학습 기법인 딥 큐 네트워크(Deep Q-Network)를 채용하여 상황별 제약하에 에너지 효율이 높은 네비게이션에 대하여 설명한다.

- [0065] 제안된 훈련 모델을 통해, 다중 무인항공기 기지국(UAV-BS)에 대한 효과적인 실시간 궤도 정책을 얻을 수 있으며, 시간이 지남에 따라 관측 가능한 네트워크 상태를 포착할 수 있다. 그 결과, 정보연령(AoI) 및 에너지 효율 제약 조건 하에서 궤도 최적화 문제를 효과적으로 해결할 수 있는 경험 재생을 통해 제안된 딥 큐 네트워크(DQN, Deep Q Network)에 대한 상태, 관측, 행동 공간 및 보상을 명시적으로 설계한다. 또한, 종래의 딥 큐 네트워크와는 달리, 제안된 딥 큐 네트워크 모델은 경험 재생 메모리의 장점을 활용하여 복수의 다중 무인항공기 기지국(UAV-BS)의 위치를 조정하면서 최적의 궤도 정책을 획득한다.
- [0066] 제안된 훈련 모델을 적용함으로써, 다중 무인항공기 기지국(UAV-BS)에 대한 효과적인 실시간 궤도 정책은 시간이 지남에 따라 관측 가능한 네트워크 상태를 포착한다.
- [0067] 마지막으로, 본 출원에 의하여 제안된 접근방식이 기존의 접근 방식 보다 더 에너지 효율적이라는 것을 설명한다. 이를 위하여, 상술한 기술적 특징의 성과를 평가하기 위해 광범위한 실험 분석에 대한 결과를 기술한다. 또한 적절한 할인율과 AoI 성능 메트릭(metrics) 임계값을 통해 학습 모델에 힘을 실어주기 위해 적절한 시스템 매개변수를 찾기 위해 광범위한 시뮬레이션 분석에 대한 결과를 기술한다. 이러한 결과를 기초로, 본 출원에서 제안된 네비게이션 기술이 기존의 접근법에 비해 상당한 에너지 효율과 데이터 최신성을 달성한다는 것을 설명한다.
- [0068] 또한, 본 발명에서는 무인정찰기 네비게이션에 관하여, 에너지 효율과 정보연령(AoI) 메트릭(metric)을 모두 고려하여 계산 지향적 통신(COC) 애플리케이션을 중심으로 설명한다.
- [0069] 또한, 본 발명에서는, 에너지 효율에 관하여, 종래기술과 달리, 에너지 효율과 정보연령(AoI) 메트릭(metric) 사이의 절충과, 계산 지향 통신(COC) 애플리케이션의 다른 궤도 제약사항을 고려하여 설정되는 무인항공기 기지국(UAV-BS)의 궤도 설정에 대하여 설명한다. 또한 무인항공기 기지국(UAV-BS)의 에너지 효율을 높이기 위해 중복 커버리지를 더욱 줄여주는 다중 무인항공기 기지국(UAV-BS) 네비게이션을 고려하는 공동 행동 공간을 설정한다.
- [0070] 또한, 본 발명에서는, 정보 연령(AoI)에 관하여, 그러나, 계산 지향 통신(COC) 애플리케이션을 위한 무인항공기 기지국(UAV-BS)의 새로운 정보 업데이트를 통합하기 위해, 무인항공기 기지국(UAV-BS)와 지상 기지국 사이의 백홀(back-haul) 통신을 가능하게 하기 위한 고주파 밀리미터 웨이브(mmWave) 무선 스펙트럼을 설정한다.
- [0072] <SYSTEM MODEL>
- [0073] 도 1은 본 발명의 일 실시예에 따른 다중 무인항공기 네비게이션 최적화 방법이 적용된 무인항공망 시스템의 개념적인 설명을 위한 도면이다.
- [0074] 도 1은 엣지 컴퓨팅을 이용한 무인항공망(Unmanned Aerial Networks) 시스템 모델을 예시하고 있다.
- [0075] 무인항공망(Unmanned Aerial Networks) 시스템은 적어도 하나의 지상 기지국(Ground Base Station)과 적어도 하나의 무인항공기 기지국(UAV-BS)을 포함한다. 또한, 주어진 궤도 지점 세트 $P = \{1, 2, 3, \dots, p, \dots, P\}$ 를 고려한다.
- [0076] P 의 궤도 지점은 배터리로 구동되는 무인항공기 기지국(UAV-BS) 세트 $\Pi = \{1, 2, 3, \dots, u, \dots, U\}$ 로 커버된다. 여기에서, 무인항공기 기지국(UAV-BS) 세트는 사용자 단말인 IoT(Internet on Things) 디바이스 $I = \{1, 2, 3, \dots, i, \dots, I\}$ 세트의 릴레이 역할을 한다.
- [0077] 본 출원에서는 IoT 디바이스 세트가 서로 다른 궤도 지점에 랜덤하게 위치한다는 점을 고려한다. 또한, 본 출원에서는 멀티액세스 엣지 컴퓨팅 서버(MEC)를 탑재하고, 무인항공기 기지국(UAV-BS)에 의한 릴레이를 통해 IoT 디바이스로부터 정보 업데이트를 받아 연산 지향 통신 애플리케이션을 지원하는 정보 융합 센터 역할을 하는 지상 기지국(BS-MEC)도 고려한다.
- [0078] 도 2는 멀티액세스 엣지 컴퓨팅 서버를 구현하는 컴퓨팅 장치의 일 예시를 도시하는 도면이다. 도 2를 참조하면, 멀티액세스 엣지 컴퓨팅 서버(MEC)는 컴퓨팅 장치(200)로 구현되며, 데이터베이스 프로그램 내의 로직 시나리오에 따라 이하에서 설명하는 기능을 수행하는 애플리케이션을 실행하는 임의의 컴퓨팅 장치로서, 적어도 프로세싱 유닛(203)과 시스템 메모리(201)를 포함할 수 있다. 컴퓨팅 장치는 프로그램을 실행할 때 협조하는 복수의 프로세싱 유닛을 포함할 수도 있다. 컴퓨팅 장치의 정확한 구성 및 유형에 의존하여, 시스템 메모리(201)는 휘발성(예컨대, 램(RAM)), 비휘발성(예컨대, 롬(ROM), 플래시 메모리 등) 또는 이들의 조합일 수 있다.

시스템 메모리(201)는 플랫폼의 동작을 제어하기 위한 적합한 운영 체제(202)를 포함하는데, 예컨대 마이크로소프트사에서부터의 WINDOWS 운영체제와 같은 것일 수 있다. 시스템 메모리(201)는 프로그램 모듈, 애플리케이션 등의 같은 하나 이상의 소프트웨어 애플리케이션을 포함할 수도 있다.

[0079] 컴퓨팅 장치는 추가적 특징 또는 기능을 가질 수 있다. 예를 들어, 컴퓨팅 장치는 자기 디스크, 광학적 디스크, 또는 테이프와 같은 추가적인 데이터 저장 장치(204)를 포함할 수 있다. 이러한 추가적 저장소는 이동식 저장소 및/또는 고정식 저장소 일 수 있다. 컴퓨터 판독가능 저장 매체는 컴퓨터 판독가능 인스트럭션, 데이터 구조, 프로그램 모듈, 또는 기타 데이터와 같은 저장정보를 위한 임의의 방법이나 기법으로 구현되는 휘발성 및 비휘발성, 이동식 및 고정식 매체를 포함할 수 있다. 시스템 메모리(201), 저장소(204)는 모두 컴퓨터 판독가능 저장 매체의 예시일 뿐이다. 컴퓨터 판독가능 저장 매체는 램(RAM), 롬(ROM), EEPROM, 플래시 메모리 또는 다른 메모리 기법, CD-ROM, DVD 또는 다른 광학적 저장소, 자기 테이프, 자기 디스크 저장소 또는 다른 자기적 저장 장치, 또는 원하는 정보를 저장하고 컴퓨팅 장치(200)에 의해 액세스될 수 있는 임의의 다른 매체를 포함할 수 있는데, 이에 한정되는 것은 아니다.

[0080] 컴퓨팅 장치의 입력 장치(205), 예컨대 키보드, 마우스, 펜, 음성 입력 장치, 터치 입력 장치, 및 비교 가능한 입력 장치를 포함할 수 있다. 출력 장치(206)는, 예컨대 디스플레이, 스피커, 프린터, 및 다른 유형의 출력 장치가 포함될 수도 있다. 이들 장치는 본 기술분야에서 널리 알려진 것이므로 자세한 설명은 생략한다.

[0081] 컴퓨팅 장치는 예컨대 분산 컴퓨팅 환경에서의 네트워크(20), 예컨대, 유무선 네트워크, 위성 링크, 셀룰러 링크, 근거리 네트워크, 및 비교가능한 메커니즘을 통해 장치가 다른 장치들과 통신하도록 허용하는 통신 장치(207)를 포함할 수도 있다. 통신 장치(207)는 통신 매체의 한가지 예시이며, 통신 매체는 그 안에 컴퓨터 판독가능 인스트럭션, 데이터 구조, 프로그램 모듈, 또는 다른 데이터를 포함할 수 있다.

[0082] 이러한 멀티액세스 에지 컴퓨팅 서버(BS-MEC)는 다양한 네트워크 소스에서 시기 적절한 정보 업데이트가 필수적인 특정 계산 정확도를 달성하기 위하여, 다양한 통신 자원을 이용할 수 있는 네트워크의 포인트 일 수 있다.

[0083] 멀티액세스 에지 컴퓨팅 서버(BS-MEC)를 포함하는 지상 기지국 b는 자신을 제외한 주변의 지상 기지국 b의 궤도 점 집합 $\mathcal{P}^b = \{p \in \mathcal{P} : (b, p) \in F\}$ 에 의해 표시되는 설정 P 내의 궤도점으로 간주된다. 여기서 $F \subseteq \mathcal{P} \times \mathcal{P}$ 는 궤도 지점들 사이의 궤도 연결의 집합이다.

[0084] 무인항공기 기지국(UAV-BS) u는 유한 관측 시간 T에 걸쳐 P의 서로 다른 궤도 지점을 통과한다. 임의의 궤도 지점 p로 무인항공기 기지국(UAV-BS)이 이동함에 따라, 업링크 통신 채널을 이용하여 해당 궤도 지점 p 근처에 위치한 활성 IoT 디바이스 i로부터 정보 데이터 패킷을 수집한다.

[0085] 무인항공기 기지국(UAV-BS) u는 지상 기지국(BS-MEC) b가 무인항공기 기지국(UAV-BS) u의 전송 범위 내에 있을 때 백홀 통신 링크를 사용한다. 그 결과, 지상 기지국(BS-MEC) b에 탑재된 멀티액세스 에지 컴퓨팅 서버(MEC)는 다른 궤도 지점으로부터 새로운 정보 업데이트를 받고 정보연령(AoI) 메트릭(metric)을 계산할 수 있다.

[0086] 여기에서, 지상 기지국(BS-MEC)는 일련의 mmWave 방향 안테나를 갖추고 있으며 무인항공기 기지국(UAV-BS)와 백홀 통신을 위한 전용 mmWave 스펙트럼을 제공하는 것으로 설명한다. 또한, IoT 디바이스와 무인항공기 기지국(UAV-BS)도 지향성 안테나를 포함하며, 따라서, IoT 디바이스가 비mmWave 스펙트럼을 이용해 정보 업데이트를 무인항공기 기지국(UAV-BS)에 전송할 수 있도록 했다.

[0087] 다만, 이에 한정되는 것은 아니며 다양한 다른 무선 통신 방식이 적용 가능하다.

[0088] 도 3은 본 발명의 일 실시예에 따른 무인항공기 기지국의 일 예를 설명하는 블록 구성도이다. 도 3을 참조하면, 무인항공기 기지국(UAV-BS)(100)은 구동부(110), 통신부(120), 제어부(130), 배터리(140) 및 메모리(150)를 포함할 수 있다.

[0089] 구동부(110)는 배터리(140)에 축전된 전원을 이용하여, 제어부(130)의 제어에 따라 무인항공기 기지국(UAV-BS)을 이동시키는 수단이다. 복수의 프로펠러를 가지는 드론으로 무인항공기 기지국(UAV-BS)이 구성된 예를 들면, 구동부(110)는 프로펠러와 프로펠러를 구동하는 모터, 모터를 제어하는 모터 드라이버를 포함할 수 있다. 한편, 무인항공기 기지국(UAV-BS)는 이 외에도 날개가 추진체(모터 또는 제트엔진 등)으로 구현되는 등 다양하게 구현 가능하다.

[0090] 통신부(120)는 지상 기지국(BS) 또는 IoT 디바이스와의 통신 채널을 형성할 수 있다. 상술한 예와 같이, 통신부(120)는 지향성 안테나를 구비하고, 지상 기지국과 백홀 통신을 위한 전용 mmWave 스펙트럼을 이용하여 통신하

고, IoT 디바이스와 비mmWave 스펙트럼을 이용하여 통신할 수 있다.

[0091] 배터리(140)는 전원을 축전하고, 무인항공기 기지국의 각 구성요소에 전원을 공급한다.

[0092] 제어부(130)는 배터리에 축전된 전원을 이용하여 동작하며, 무인항공기 기지국의 다른 구성요소를 제어하여 무인항공기 기지국의 구동 및 동작을 제어할 수 있다. 상술한 그리고 이하에서 설명하는 무인항공기 기지국의 각 기능적인 특징은 제어부에 의하여 수행되는 것이다.

[0093] 메모리(150)는 데이터 저장장치로서, 무인항공기 기지국의 다른 구성요소에 데이터의 읽기 또는 쓰기 기능을 제공할 수 있다.

[0095] 이하에서, 궤도 지점 p의 다른 소스 (예컨대, 활성 IoT 디바이스) i에서 지상 기지국(BS-MEC) b로 데이터 릴레이를 위한 무인항공기 기지국(UAV-BS) u의 전개, 통신 및 네비게이션을 중심으로 설명한다. 또한 IoT 디바이스와 서로 다른 궤도 지점에 있는 무인항공기 기지국(UAV-BS) 사이의 사용자 연결에 대해서는 기본 최대 신호 대 상호접속 플러스 노이즈 비율(SINR) 기반 접근법을 적용한다.

[0097] A. IoT-to-UAV-BS Communication Model

[0098] 궤도 지점 p에서, 무인항공기 기지국(UAV-BS) u와 활성 IoT 디바이스 i 간의 공대지 경로 손실 확률(air-to-ground path loss probability)은 아래의 수학적 식 1과 같이 계산된다.

[0099] [수학적 식 1]

$$\zeta_{i,p}^u = \begin{cases} \frac{1}{1+\alpha \exp(-\hat{\alpha}(\frac{180}{\pi}\Theta_u-\alpha))}, & \text{LoS channel,} \\ 1 - \left[\frac{1}{1+\alpha \exp(-\hat{\alpha}(\frac{180}{\pi}\Theta_u-\alpha))} \right], & \text{NLoS channel.} \end{cases}$$

[0100]

[0101] 여기에서, α 와 $\hat{\alpha}$ 는 각각 LoS 및 NLoS 채널에 대한 환경 의존 상수이고, 여기서 Θ_u 는 무인항공기 기지국(UAV-BS) u의 표고각이다.

[0102] 수학적 식 1에서, LoS 채널과 NLoS 채널의 경로 손실 확률에 대하여, 시골환경에서와 달리, 도시나 지하 환경에서는 무인항공기 기지국(UAV-BS)과 활성 IoT 디바이스 사이의 업링크 통신 링크는 주변 장애물(예를 들어 건물)에 의하여 방해될 수 있는 특징을 반영한다.

[0103] 데시벨(dB) 단위의 경로손실은 아래의 수학적 식 2와 같이 계산될 수 있다.

$$P_{i,p}^u = \begin{cases} 20 \log(\frac{4\pi f_c \delta_{i,p}^u}{c}) + \epsilon, & \text{LoS channel,} \\ 20 \log(\frac{4\pi f_c \delta_{i,p}^u}{c}) + \bar{\epsilon}, & \text{NLoS channel.} \end{cases}$$

[0104]

[0105] 여기에서, f_c 는 업링크 채널 주파수이고, ϵ 및 $\bar{\epsilon}$ 는 각각 LoS 및 NLoS 채널의 감쇠 계수이다.

[0106] 수학적 식 2를 이용하여, 궤도 지점 p에서 활성 IoT 디바이스 i로부터 무인항공기 기지국(UAV-BS) u 까지 수신된 신호 출력은 아래의 수학적 식 3과 같이 계산된다.

[0107] [수학적 식 3]

$$\hat{P}_{i,p}^u = \frac{\bar{P}_{i,p}^u}{P_{i,p}^u},$$

[0108]

[0109] 여기에서, $\hat{P}_{i,p}^u$ 는 데이터를 무인항공기 기지국(UAV-BS) u로 오프로드하기 위한 IoT 디바이스 i의 전송 전력이다.

[0110] 시간 t 에서 IoT 디바이스 i 가 궤도 지점 p 에서 무인항공기 기지국(UAV-BS) u 에 대해 받은 SINR은 다음의 수학적식 4와 같이 계산된다.

[0111] [수학적식 4]

$$\gamma_{i,p}^u(t) = \frac{\hat{P}_{i,p}^u (10^{\frac{\zeta_{i,p}^u}{10}})^{-1}}{I_{i,p}^u + \sigma^2}.$$

[0112]

[0113] 여기에서, $I_{i,p}^u$ 는 다른 무인항공기 기지국(UAV-BS) u' 에 의한 무인항공기 기지국(UAV-BS) u 의 수신 간섭을 의미하며, 다른 무인항공기 기지국(UAV-BS) u' 는 다른 인접지점과 겹치는 궤도지점 p' 에 위치하는 IoT 디바이스 i' 에게 서비스를 제공한다. σ^2 는 노이즈 세기를 의미한다. $I_{i,p}^u$ 는 아래의 수학적식 4-1을 만족한다.

[0114] [수학적식 4-1]

$$I_{i,p}^u = \sum_{p' \in \mathcal{P}} \sum_{u' \in \mathcal{U}} \sum_{i' \in \mathcal{I}} \hat{P}_{i',p'}^{u'} (10^{\frac{\zeta_{i',p'}^{u'}}{10}})^{-1}$$

[0115]

[0116] 수학적식 4를 이용하여, 시간 t 에서의 채널 용량은 아래의 수학적식 5와 같이 정의된다.

[0117] [수학적식 5]

$$r_{i,p}^u(t) = \begin{cases} \frac{\beta_u}{|\mathcal{I}|} \cdot \log(1 + \gamma_{i,p}^u(t)), & \text{if } \gamma_{i,p}^u(t) > \gamma_{th}, \\ 0, & \text{otherwise.} \end{cases}$$

[0118]

[0119] 여기에서, β_u 는 고정된 비 mmWave (non-mmWave) 업링크 채널 대역폭으로, 궤도 지점 p 에서 IoT 디바이스들에게 균등하게 분배된다. γ_{th} 는 IoT 디바이스와 UAV-BS 간의 성공적인 업링크 전송을 보장하기 위한 SINR 임계값이다.

[0121] B. UAV-BS-to-BS Communication Model

[0122] 무인항공기 기지국(UAV-BS) u 로부터의 지상 기지국(BS) b 의 수신 전력은 아래의 수학적식 6과 같이 계산된다.

[0123] [수학적식 6]

$$\hat{P}_{b,u} = P_{b,u}^{tx} \cdot G_u^{tx} \cdot G_b^{rx} \left(\frac{c}{4\pi\delta_{b,u}f_c^{mmWave}} \right).$$

[0124]

[0125] 여기에서, $P_{b,u}^{tx}$ 는 무인항공기 기지국(UAV-BS) u 로부터의 지상 기지국(BS) b 로의 송신 전력, $\delta_{b,u}$ 는 무인항공기 기지국(UAV-BS) u 과 지상 기지국(BS) b 간의 거리, c 는 빛의 속도, f_c^{mmWave} 는 mmWave 백홀 링크의 반송파 주파수(carrier frequency), G_u^{tx} 및 G_b^{rx} 는 각각 송신기 UAV-BS u 와 수신기 BS b 의 안테나 이득을 의미한다.

[0126] 시간 t 에서, 무인항공기 기지국(UAV-BS) u 로부터의 지상 기지국(BS) b 간의 채널의 백홀 용량(back-haul capacity)은 아래의 수학적식 7과 같이 계산된다.

[0127] [수학식 7]

$$r_{b,u}^{mmWave}(t) = \begin{cases} \beta_{b,u}^{mmWave} \cdot \log \left(1 + \frac{\hat{P}_{b,u}}{\beta_{b,u}^{mmWave} \sigma^2} \right), \\ \text{if } \delta_{u,b} \leq \bar{\alpha}, \\ 0, \text{ otherwise.} \end{cases}$$

[0128]

[0129] 여기에서, $\beta_{b,u}^{mmWave}$ 는 mmWave 백홀 대역폭, σ^2 는 첨가 소음이다. 만약 높이 h_u 의 무인항공기 기지국 (UAV-BS)과 지상 기지국(BS) b 간의 거리 $\delta_{u,b}$ 가 임계거리 $\bar{\alpha}$ 보다 짧으면, 무인항공기 기지국(UAV-BS)는 백홀 링크를 생성하고, mmWave 백홀 대역폭 즉, $\beta_{b,u}^{mmWave}$ 를 이용하여 정보 업데이트를 지상 기지국(BS)에 전송 한다.

[0130] 여기에서, $\delta_{u,b}$ 는 아래의 수학식 7-1을 만족한다.

[0131] [수학식 7-1]

$$\delta_{u,b} = \sqrt{h_u^2 + (x_u - x_b)^2 + (y_u - y_b)^2}$$

[0132]

[0133] 수학식 7을 이용하여, 시간 t에서 백홀 링크를 사용하는 동안 무인항공기 기지국(UAV-BS) u의 전송 에너지는 아래의 수학식 8과 같이 계산된다.

[0134] [수학식 8]

$$E_u^{mmWave}(t) = P_{b,u}^{tx} \times r_{b,u}^{mmWave}(t).$$

[0135]

[0137] C. UAV-BS Relay Network Energy Efficiency Metric Design

[0138] 무인항공기 기지국(UAV-BS) u는 다른 무인항공기 기지국(UAV-BS)가 서로 다른 고도를 유지할 수 있는 항상성 고도 h_u 에서 수평으로 관측 영역을 커버한다.

[0139] 또한, 시간 t에서의 UAV-BS 궤도는 다음의 수학식 8-1과 같이 정의된다.

[0140] [수학식 8-1]

$$\tau_u(t) = [x_u(t), y_u(t)]^T \in \mathbb{R}^{2 \times 1}$$

[0141]

[0142] 따라서 현재 궤도 위치에서 다음 위치까지 일정한 고도 h_u 에서 수평으로 무인항공기 기지국(UAV-BS)에 의해 적용되는 시간 변화 거리($\hat{x}_u(t)$ and $\hat{y}_u(t)$) 는 다음의 수학식 9와 같이 정의된다.

[0143] [수학식 9]

$$\delta_u(t) = \sqrt{h_u^2 + [x_u(t) - \hat{x}_u(t)]^2 + [y_u(t) - \hat{y}_u(t)]^2}, \forall u \in \mathcal{U}, \\ 0 \leq t \leq T.$$

[0144]

[0145] 시간 t에서 거리 δ_u 를 커버하기 위한 무인항공기 기지국(UAV-BS) u의 총 이동 에너지 비용은 다음의 수학식 10과 같이 계산된다.

[0146] [수학식 10]

$$E_u(t) = \delta_u(t) \times E_{prop}.$$

[0147]

여기에서, E_{prop} 는 추진 전력 소비량의 상한이고, k_1 및 k_2 는 무인항공기 기지국(UAV-BS)에 따라 설정되는 값이고, g 는 중력 가속도이다. E_{prop} 는 다음의 수학적 식 10-1을 만족한다.

[수학적 식 10-1]

$$E_{prop} = k_1 ||v||^3 + \frac{k_2}{||v||} (1 + \frac{||a||^2}{g^2})$$

따라서, 시간 경과 T 에 따른 IoT 디바이스들을 서비스하기 위해 궤도 지점 P 를 커버하는 무인항공기 기지국(UAV-BS) u 의 총 에너지 효율은 다음의 수학적 식 11과 같이 정의된다.

[수학적 식 11]

$$\eta(\mathcal{P}, u) = \sum_{t=1}^T \sum_{p=1}^{|\mathcal{P}|} \frac{(r_{b,u}^{mmWave}(t) + \sum_{i=1}^{|\mathcal{I}|} r_{i,p}^u(t))}{(E_u^{mmWave}(t) + E_u(t))}.$$

D. Age of Information Model for Ground Station

정보연령(AoI) 메트릭(metric)은 무인항공기 기지국(UAV-BS)이 IoT 디바이스의 릴레이 노드 역할을 하는 P 의 궤도 지점에서 무인항공기 기지국(UAV-BS)이 수집한 정보의 최신성을 측정하는 데 사용된다.

따라서 지상 기지국(BS) b 에서 시간 t 에서의 궤도 p 의 정보연령(AoI)은 다음의 수학적 식 12와 같이 계산된다.

[수학적 식 12]

$$\Delta_u(\mathcal{P}, t) = t - \Delta'_u(p, t), \forall p \in \mathcal{P}.$$

여기에서, $\Delta'_u(p, t)$ 는 지상 기지국(BS) b 에서 무인항공기 기지국(UAV-BS) u 에 의한 궤도 지점 p 에서 가장 최근에 수신된 데이터 패킷의 시간 스탬프로 표시된다. t 에서, 궤도 지점 p 로부터의 정보 업데이트가 가장 최근의 것이라고 가정한다. 또한, 시간 t 에서의 궤도 경유점 p 와 관련된 정보 연령(AoI)은 수학적 식 12에 의해 주어진다. 그 후, (12)에 의해 시간 t 에서의 궤도 경유점 p 와 관련된 정보 에이지(AoI)가 주어진다.

지상 기지국(BS) b 가 궤도 경유점 p 로부터 새로운 정보 업데이트를 받지 못하는 동안, $\Delta_u(\mathcal{P}, t)$ 및 $\forall p \in \mathcal{P}$ 값은 시간 t 에 선형적으로 증가한다. 이는 정보가 점점 오래된 것임을 나타낸다.

지상 기지국(BS) b 가 궤도 경유점 p 로부터 새로운 정보 업데이트를 받는 즉시, 해당 시간 스탬프는 $\Delta'_u(p, t)$ 에서 $\Delta'_u(p, t+1)$ 로 즉시 업데이트되며, $\Delta'_u(p, t)$ 및 $\forall p \in \mathcal{P}$ 값을 $\Delta'_u(p, t+1) - \Delta'_u(p, t)$ 만큼 감소시키도록 업데이트 된다.

즉, $t+1$ 시점에서 무인항공기 기지국(UAV-BS)에 의하여 지상 기지국(BS) b 에서 정보 업데이트가 발생하고, $\Delta_u(\mathcal{P}, t)$ 및 $\forall p \in \mathcal{P}$ 값은 정보 업데이트 지연과 일치한다.

시간 t 가 0 일 때, $\Delta_u(\mathcal{P}, t)$ 또한 0이라고 가정한다.

타임 슬롯 T 동안의 궤도 지점 P 에 대한 지상 기지국(BS) b 에서 계산되는 평균 정보연령(AoI)은 아래의 수학적 식 13과 같이 계산된다.

[0167] [수학식 13]

[0168]

$$\hat{\Delta}_b(\mathcal{P}) = \frac{1}{T|\mathcal{P}|} \sum_{t=1}^T \sum_{p \in \mathcal{P}} \Delta_u(p, t).$$

[0170] < PROBLEM FORMULATION>

[0171] 상술한 상황적 제약 조건 하에서 UAV-BS 네비게이션 최적화 문제를 먼저 공식화하기 위해, 우리는 각 무인항공기 기지국(UAV-BS) u 가 주어진 시간대 T 에서 궤도 지점의 하위 집합만 다룰 수 있다고 설정한다. 또한, 궤도 지점으로 구성된 각 u 에 대한 서브셋을 다음의 수학식 13-1을 만족하도록 설정한다.

[0172] [수학식 13-1]

[0173] $\mathcal{P}_u \subset \mathcal{P}, \mathcal{P}_u \cap \mathcal{P}_{u'} = \emptyset \text{ where } u \neq u'.$

[0174] 결국, 에너지 효율 무인항공기 기지국(UAV-BS) 네비게이션 최적화 문제의 목적은 에너지 및 AoI 지표의 적용을 받는 무인항공기 기지국(UAV-BS) 중계망의 총 에너지 효율을 극대화하는 무인항공기 기지국(UAV-BS)의 협력 궤도 경로 구성을 찾는 것으로 설정할 수 있다. 이러한 최적화 문제 공식은 아래의 수학식 14 내지 수학식 18과 같이 나타낼 수 있다.

[0175] [수학식 14]

[0176]

$$\arg \max_{\{\mathcal{P}_u\}_{u \in \mathcal{U}}} \sum_{u \in \mathcal{U}} \eta(\mathcal{P}_u, u),$$

[0177] [수학식 15]

[0178]

$$\bigcap_{u \in \mathcal{U}} \mathcal{P}_u = \{b\}, \forall u \in \mathcal{U},$$

[0179] [수학식 16]

[0180]

$$\bigcup_{u \in \mathcal{U}} \mathcal{P}_u = \mathcal{P}, \forall u \in \mathcal{U},$$

[0181] [수학식 17]

[0182]

$$\eta(\mathcal{P}_u, u) \geq \eta_{th}, \forall u \in \mathcal{U},$$

[0183] [수학식 18]

[0184]

$$\hat{\Delta}_b(\mathcal{P}_u) \leq \hat{\Delta}_b^{th}, \forall p \in \mathcal{P}_u \setminus \{b\}.$$

[0186] 상기 수학식 15 내지 18은 각각 궤도, 에너지 효율, AoI 제약조건에 관한 것이다. 수학식 15는 정보 업데이트가 발생하는 지상 기지국(BS) 궤도 지점은 제외한, 무인항공기 기지국(UAV-BS)의 오버랩되지 않는 궤도를 나타낸다. 수학식 16은 모든 궤도 지점이 상호의존적으로 포함된 무인항공기 기지국(UAV-BS)의 공동 궤도 구성을 나타낸다. 수학식 17 및 18은 에너지 효율과 정보연령(AoI) 메트릭이 모두 $\mathcal{P}_u$ 의 함수인 의사결정 변수 $\mathcal{P}_u$ 와 결합된다. 수학식 17은 통신 및 이동 에너지가 최소 에너지 효율 임계값 η_{th} 보다 커야 하는 무인항공기 기지국(UAV-BS)의 총 에너지 효율을 보장한다. 수학식 18은 구성 $\mathcal{P}_u$ 에 의한 정보 업데이트의 평균 최신성이 정보연령(AoI) 임계값 $\hat{\Delta}_b^{th}$ 미만이어야 함을 보장한다.

[0187] 수학식 18에 의한 수학식 14의 제약 조건에 의하여, UAV-BS는 에너지 효율 제약 조건뿐만 아니라 계산 지향적

통신 애플리케이션의 성능이 다른 궤도 지점으로부터의 최신 정보 업데이트에 의존하는 AoI 제약 조건에서도 서로 다른 궤도 지점을 공동으로 탐색함을 알 수 있다.

[0188] 수학적 14의 결정 문제는 그에 대한 제약조건이며 NP-완전인 수학적 15 내지 18과 함께 정점 커버 문제 (예를 들어, Maximum Clique Problem)의 기본 문제로 축소될 수 있다. Maximum Clique Problem와 유사하게, 수학적 14는 본질적으로 결합적이다. 더욱이 수학적 14의 해법으로 볼 때, 수학적 14의 해가 주어지면 최적인지 여부를 알 수 있는 기존에 알려진 다항식 알고리즘은 존재하지 않는다. 따라서, 수학적 14의 결정문제는, NP-하드인 것으로 입증된, 정점커버 문제(problem of the vertex cover problem)의 동일한 범주에 속한다고 유추할 수 있다.

[0189] 이하에서, 딥 큐 학습(Deep Q learning) 기법을 사용하여, 상술한 수학적 14 내지 18에 대하여 보다 상세히 설명한다.

[0191] < PROPOSED TRAJECTORY POLICY ALGORITHM BASED ON DEEP Q-LEARNING>

[0192] 이하에서, 심층신경망과 강화학습 알고리즘이 결합된 심층강화학습(DEEP Q-LEARNING) 모델에 대하여 설명한다. 심층강화학습(DEEP Q-LEARNING) 모델은 멀티액세스 에지 컴퓨팅 서버(MEC)에 의하여 구현 및 수행될 수 있다.

[0193] 도 4는 지상 기지국의 멀티 액세스 에지 컴퓨팅 서버에 의하여 수행되는 심층강화학습(DEEP Q-LEARNING) 모델의 일 예를 도시하는 도면이다.

[0194] 도 4를 참조하면, 심층강화학습(DEEP Q-LEARNING) 모델(300)은 심층 신경 네트워크(310), 경험 재생 메모리(320) 및 강화 학습 프레임워크(330)를 포함한다.

[0195] 심층 신경 네트워크(310)는 무인항공기 기지국(UAV-BS) 네비게이션에 필요한 상황별 특징(예: AoI, 에너지 소비량)을 추출하는 데 사용되는 상태 공간의 차수를 줄이도록 동작한다.

[0196] 경험 재생 메모리(320)는 무인항공기 기지국(UAV-BS) 에이전트가 관찰하는 상태 전환을 저장할 수 있다.

[0197] 강화 학습 프레임워크(330)는 상술한 제약조건인 수학적 15 내지 18를 이용하여, 문제 목표인 수학적 14를 달성하는 최선의 궤도 정책을 찾기 위한 강화 학습을 수행한다.

[0198] 심층강화학습(DEEP Q-LEARNING) 모델은 자유로운 모델인 만큼 네트워크 동적 모델이 필요하지 않다. 더욱이 제안된 접근방식에서 경험 재생의 사용은 심층 신경망의 훈련에서 사용되는 관찰 사이의 시간적 의존성을 깨뜨려 안정성을 보장한다.

[0199] 먼저, 문제의 상태 및 수학적 14로 표현되는 행동 공간을 모델링 하고, 해당 제약조건이 있는 수학적 14에 근거한 보상 및 통제 정책을 모델링한다. 마지막으로, 무인항공기 기지국(UAV-BS) 궤도 정책에 대해 제안된 훈련 및 시험 모델을 설명한다.

[0201] A. State and Action Space

[0202] 무인항공기 기지국(UAV-BS)의 궤도 정책을 위한 상태 공간(state space)은 4차원 상태 공간이다. 지상 기지국(BS)은 무인항공기 기지국의 궤도 정책을 위한 4차원 상태 공간을 설정할 수 있다.

[0203] 각 시간 단계(time step) t에서, $t = \{1, 2, 3, \dots, T\}$ 를 만족하고, 지상 기지국(BS)에 구비된 멀티액세스 에지 컴퓨팅 서버(MEC)는 가상 학습 에이전트를 제공한다. 가상 학습 에이전트의 상태 또는 공동 관찰 공간은 S로 표시되며, S는 다음의 수학적 18-1과 같이 표현될 수 있다.

[0204] [수학적 18-1]

$$[0205] \mathcal{S} = \{s_t = (p_{current}^u, p_{end}, \eta, \Delta) | \eta \in [0, \eta_{th}], \Delta \in [1, \hat{\Delta}_b^{th}]\}$$

[0206] 여기에서, $p_{current}^u$ 는 높이 h_u 에서 무인항공기 기지국(UAV-BS) u의 현재 위치이고, p_{end} 는 목표 위치이고, η 는 UAV-BS 들의 평균 에너지 효율이고, Δ 는 탐색 최적화를 위한 평균 연령이다.

[0207] 또한 무인항공기 기지국(UAV-BS) u 의 탐색을 위한 궤도 위치는 $x_u \in [0, X_u]$ 와 $y_u \in [0, Y_u]$ 를 이용하며, 여기서 X_u 와 Y_u 는 특정 지리적 위치의 최대 좌표이다. 또한 각 무인항공기 기지국(UAV-BS)의 초기 위치는 IoT 디바이스의 수와 함께 무작위로 할당한다. 상태 공간에서 연속 상태 변수 η 와 Δ 의 하한과 상한은 실제 궤도 데이터에서 보정한다.

[0208] 무인항공기 기지국(UAV-BS)의 행동 공간(Action Space)은 궤도 및 통신 제약 (예를 들어, 상술한 수학식 14 내지 18로 표현되는 제약 조건)을 충족하면서 하나의 실행 가능한 상태 (예를 들어, 위치)에서 다음 상태로 각각의 각의 UAV-BS의 항해를 계획하는 궤도다. 각각의 무인항공기 기지국(UAV-BS)은 각각 궤도 및 통신제약을 충족하면서 하나의 실행 가능한 상태에서 다음 상태로의 항해를 설정하여 행동 공간을 설정할 수 있다.

[0209] 가상 학습 에이전트는 상태 S_t 에서 사용 가능한 행동 집합에서 행동 a_t 를 선택하고, a_t 는 $a_t \in \mathcal{A}_{s_t} \subset \mathcal{A}$, 을 만족한다.

[0210] $\mathcal{A} = \{a_1, \dots, a_U\} = \{\mathcal{P}_u\}_{u \in \mathcal{U}}$ 을 만족하며, $\mathcal{A}$ 는 무인항공기 기지국(UAV-BS) 네비게이션의 구성이다.

[0211] 지상 기지국(BS)는 임의의 상태에서 복수의 무인항공기 기지국(UAV-BS)이 각각 설정한 복수의 행동 공간에 대하여 경유지의 충돌 여부를 기준으로 보상을 설정할 수 있다.

[0212] 이에 대하여, 이하에서 행동공간의 디자인의 일 예시를 들어 보다 상세히 설명한다.

[0214] EXAMPLE OF THE ACTION SPACE DESIGN

[0215] 이하에서, 행동 공간(Action Space) 구성에 대하여 일 예시를 들어 설명한다.

[0216] 예를 들어, 6개의 궤도 경유점 $P = \{1, 2, 3, 4, 5, 6\}$ 과 두 개의 무인항공기 기지국(UAV-BS) u_1 및 u_2 가 있다고 가정한다. 또한, 무인항공기 기지국(UAV-BS)의 시작점은 무작위이므로 u_1 및 u_2 의 시작점은 각각 1과 2라고 가정한다. 또한, 지상 기지국(BS)이 6에 위치하고 모든 UAV-BS가 정보 업데이트를 수행하기 위해 해당 위치에 도달해야 한다고 가정한다.

[0217] 따라서, 무인항공기 기지국(UAV-BS) u_1 및 u_2 의 목표는 위치 6에 고정되어 있다. 이제, 행동 공간(Action Space) $\mathcal{A}$ 는 P 의 궤도 지점의 조합으로 정의된다.

[0218] 따라서 시간 t_1 의 궤도 경유지 조합 세트에서, u_1 및 u_2 는 조합 집합 행동(combination set action) $a_1 = \{1, 4, 6\}$ 및 $a_2 = \{2, 3, 6\}$ 에서 두 요소를 선택할 수 있고, 이는 상태 s_{t_1} 에서의 공동 행동(joint action) a_{t_1} 으로서 $\mathcal{A}_{s_{t_1}}$ 에 색인화된다. u_1 및 u_2 는, 상태 s_{t_1} 에서 시간 t_1 일 때 행동 a_{t_1} 에 대하여 경유지가 겹쳐지지 않는 것을 알 수 있다. 이 시점에서, 다른 제약조건이 위반되지 않는 경우, 시스템은 상태 s_{t_1} 에서 행동 a_{t_1} 에 대하여 보상을 설정한다.

[0219] 한편, 상태 s_{t_1} 와 시간 t_2 에서 u_1 및 u_2 는 각각 공동 행동(joint action) a_{t_2} 로서 $a_1 = \{1, 4, 6\}$ 및 $a_2 = \{2, 4, 6\}$ 을 선택했다고 하면, 경유지가 겹쳐지게 된다. 따라서, 시스템은 벌점 보상을 설정하게 된다.

[0221] B. Reward and Control Policy

[0222] 가상 학습 에이전트가 액션 a_t 을 구현하면, 환경은 새로운 상태인 s_{t+1} 로 이동한다. 전환 (s_t, a_t, s_{t+1}) 과 함께 즉각적인 보상(reward) R_{t+1} 가 연결되고 가상 학습 에이전트는 피드백을 통해 보상을 받습니다. 즉, 각 상태 전환에서, 가상 학습 에이전트는 탐색을 위한 궤도 제어 정책을 형성하는 데 사용되는 즉각적인 보상(immediate reward)을 받는다. 향후 사용을 위해, 궤도 제어 정책은 현재 상태를 최적의 제어 행동에 매핑하는 가상 학습

에이전트에서 사용된다. 즉각적인 보상은 무인항공기 기지국(UAV-BS)의 순간 에너지 효율 메트릭에 의해 공식화되고 다음의 수학적 19와 같이 정의될 수 있다.

[0223] [수학적 19]

$$R_t(a_t) = \begin{cases} \alpha_1 \eta(a_t), & \text{if constraints (15)-(18) of (14) are true} \\ -\alpha_1, & \text{if constraints (15)-(17) of (14) are violated} \\ 0, & \text{otherwise.} \end{cases}$$

[0226] 여기서 α_1 는 에너지 효율 함수에 곱한 계수이며 궤도 제약 조건을 위반할 때 가상 학습 에이전트에 페널티를 적용하는데 사용될 수 있다. 그러나 수학적 14의 수학적 18에 의한 제약 조건을 위반하면, 궤도 경유지(trjectory way-points)의 정보 업데이트가 계산 지향적 통신(COC) 애플리케이션에 대해 구식이 되면(become outdated), 시스템은 보상을 받지 못한다. 즉, 제로 보상을 받는다.

[0227] 따라서 T 시간 슬롯에 대한 가상 학습 에이전트의 목표는 다음의 수학적 20과 같이 정의되는 미래 보상을 극대화하는 것입니다.

[0228] [수학적 20]

$$\hat{R}(s, a; t) = \sum_{t_0=0}^T \gamma \times R_{t-t_0}(a_t),$$

[0230] 여기에서, $\gamma \in [0, 1]$ 은 즉각적인 보상과 미래 보상의 중요성 사이의 상충을 반영한다. 보상 $R_t(a_t)$ 는 T 기간 동안, 서로 다른 타임 슬롯에서, 하나의 궤도 경유지(trjectory way-points)에서 다음 경유지로 가는 공동 이동(joint movement)-동작-에 의한 무인항공기 기지국(UAV-BS)의 에너지 효율 메트릭(energy efficiency metric)에 의하여 변경된다. 따라서, $R_{t-t_0}(a)$ 는 두 시간 슬롯간의 보상 차이이다.

[0231] 수학적 20은 보상함수로서, 보상 함수는 마지막 T 시간 단계 기간 동안 무인항공기 기지국(UAV-BS)의 현재 상태를 학습 한 후 시간 t에서 획득된다.

[0232] 따라서, Q 기능(Q-function) 또는 행동 가치 함수(action-value function)이 아래의 수학적 21과 같이 정의된 가상 학습 에이전트에 대해 제어 정책을 π 로 정의할 수 있다.

[0233] [수학적 21]

$$Q^\pi(s, a) = \hat{R}(s, a) + \gamma \sum_{s' \in \mathcal{S}} P_{s, s'}^a \sum_{a' \in \mathcal{A}} \pi(a' | s') Q^\pi(s', a').$$

[0235] 여기에서, $P_{s, s'}$ 는 환경 상태의 전이 확률이고, $s' = s_{t+1}$, π 는 컨트롤 정책, 액션 a는 환경 시뮬레이터를 통해 시행된다. 상태 및 보상 업데이트는 지상 기지국 BS에 의해 수신된 정보를 기반으로 한다.

[0237] 수학적 21과 Q 기능을 설명하기 위하여, 이하에서 수학적 21-1 내지 21-8을 참조하여 보다 상세히 설명한다.

[0238] 벨만 방정식(Bellman Equation)은 다른 모든 가치 함수에 비해 최대값을 제공하는 최적의 정책과 가치 함수를 찾는 데 적합하다. 즉, 상태 s의 값은 가상 학습 에이전트가 해당 상태를 떠날 때 얻은 보상과, 더불어 가상 학습 에이전트가 착지(도착)한 상태의 할인된 값 s0에 가상 학습 에이전트가 이전할 전환 확률을 곱한 값이다.

[0239] 즉, 상태 s의 값은 가상 학습 에이전트가 해당 상태를 떠날 때 받은 보상과, 가상 학습 에이전트가 도착한 상태의 할인된 값 s'에 가상 학습 에이전트가 이동할 전환 확률을 곱한 값으로 산출된다.

[0240] 따라서, 가치 함수에 대한 벨만 방정식을 다음의 수학적 21-1과 같이 수립할 수 있다.

[0241] [수학식 21-1]

$$V(s) = \mathbb{E}[R_{t'} + \gamma V(s_{t'} \mid s_t = s)],$$

[0242]

여기에서, $t' = t+1$ 이다.

[0243]

한편, 수학식 1을 사용하여, 가치 함수 (예를 들어, 상태-가치 함수)에 대한 벨만 기대 방정식에서, 가상 학습 에이전트는 다음과 같이 정의된 일부 정책 π 의 적용을 받는 특정 상태의 값을 아래의 수학식 21-2로 찾을 수 있다.

[0245] [수학식 21-2]

$$V^\pi(s) = \mathbb{E}^\pi[R_{t'} + \gamma V^\pi(s_{t'}) \mid s_t = s]$$

[0246]

유사하게, 상태-동작 값 함수 (Q-함수)는 다음의 수학식 21-3과 같이 수립할 수 있다.

[0247]

[수학식 21-3]

$$Q^\pi(s, a) = \mathbb{E}^\pi[R_{t'} + \gamma V^\pi(s_{t'}, a_{t'}) \mid s_t = s, a_t = a]$$

[0249]

따라서, 에이전트가 Q-값의 평균을 산출할 때, 상태에 있는 값은 수학식 21-4와 같이 수립할 수 있다.

[0250]

[수학식 21-4]

$$V^\pi(s) = \sum_{a \in \mathcal{A}} \pi(a \mid s) Q^\pi(s, a).$$

[0252]

수학식 21-2은 상태-값 함수와 상태-동작 함수 사이의 연결을 제공한다. 그러나, 가상 학습 에이전트는 행동 a 에 대한 장점을 평가하여야 한다. 따라서, 가상 학습 에이전트는 상태 s 에서 행동 a 의 장점을 다음의 수학식 21-5와 같이 계산합니다.

[0253]

[수학식 21-5]

$$Q^\pi(s, a) = \hat{R}(s, a) + \gamma \sum_{s' \in \mathcal{S}} P_{s, s'}^a V^\pi(s'),$$

[0255]

여기서, P 는 전이 확률이다.

[0256]

특정 정책 π 를 따르는 특정 행동 a 의 효과는 기 상술한 수학식 21과 같이정의될 수 있다. 수학식 21을 모델의 목표는 최상의 제어 정책 π^{opt} 를 얻는 것이며, 따라서 최대 Q-함수는 다음의 수학식 21-6과 같이 정의될 수 있다.

[0257]

[수학식 21-6]

$$Q^{\pi^{opt}}(s, a) = \mathbb{E}[\hat{R}(s, a) + \gamma \max_{a'} Q^{\pi^{opt}}(s', a') \mid s, a],$$

[0259]

여기서, 할인된 누적 상태 함수는 다음의 수학식 21-7과 같이 정의될 수 있다.

[0260]

[수학식 21-7]

$$V^{\pi^{opt}}(s) = \max_{a'} [Q^{\pi^{opt}}(s, a)].$$

[0262]

최적의 제어정책 π^{opt} 를 도출해내기 위하여, Q-함수는 다음의 수학식 21-8과 같이 업데이트 될 수 있다.

[0263]

[0264] [수학식 21-8]

$$Q_{t'}(s, a) = Q_t(s, a) + \psi(\hat{R}(s, a) + \gamma[\max_{a'} Q_t(s', a')] - Q_t(s, a)),$$

[0266] 여기에서, $t'=t+1$, $a'=a_{t+1}$ 을 만족하고, Q -함수는 채귀 메커니즘을 사용하여 업데이트되며 ψ 는 학습률이다.

[0268] C. Training with Experience Replay

[0269] 본 발명에 따른 딥 큐 네트워크(DQN) 접근법은, 시뮬레이션 중에 탐색을 위해 UAV-BS의 궤도 구성을 최적으로 제어하는 방법을 학습한다.

[0270] 따라서, 각 궤도 관측 환경 상태에 대한 목표값이 아래의 수학식 22와 같이 주어지는 Q-네트워크를 훈련하는 것이 필요하다.

[0271] [수학식 22]

$$y_k = \hat{R}(s, a) + \gamma \max_{a'} Q(s', a'; \theta_k).$$

[0273] 여기에서, θ_k 는 K 번째 반복 동안 훈련을 통해 얻은 네트워크 가중치이다. 수학식 22를 이용하면, 훈련 네트워크의 손실 함수(loss function)은 아래의 수학식 23과 같이 나타낼 수 있다.

[0274] [수학식 23]

$$\mathcal{L}_k(\theta_k) = \mathbb{E}_{(s,a) \sim \rho(\cdot)} \left[y_k - Q(s, a; \theta_k) \right]^2.$$

[0276] 여기에서, $\rho(s, a)$ 는 시퀀스 s와 행동 a에 대한 확률 분포이고, y_k 는 수학식 22에서 파생된 훈련 네트워크의 목표 값이다. 또한, 훈련을 통해 최적의 네트워크 가중치 θ^{opt} 를 얻을 수 있다.

[0277] 또한 딥 큐 네트워크(DQN)의 훈련을 강화하고 안정화하기 위해, 미니 배치 메소드(Mini-batch method)를 적용하여, 고정된 크기의 재생 메모리 $M_t = \{e_1, \dots, E\}$ 에서 모든 교육 에피소드 단계 $e_t = (s_t, a_t, R_t(a_t), s_{t'})$ 로부터 예제를 무작위로 수집한다. 따라서, 하나의 샘플이 학습에서 여러 번 사용되어 데이터 효율성이 크게 향상될 수 있다.

[0278] 결국, 따라서 수학식 22를 사용하여 수학식 23의 손실 함수는, $\mathcal{M}$ 에 대한 균일 분포로 다음의 수학식 24와 같이 표시됩니다.

[0279] [수학식 24]

$$\mathcal{L}(\theta) = \mathbb{E}_{(s,a,r,s') \sim U(\mathcal{M})} \left[(R + \gamma \max_{a'} Q^{\pi^{opt}}(s', a'; \theta^-) - Q(s, a; \theta))^2 \right].$$

[0281] 여기서 $U(\mathcal{M})$ 은 경험 재생 메모리 $\mathcal{M}$ 에 대한 균일 분포이고 θ^- 는 딥 큐 네트워크(DQN)의 저장된 가중치 매개 변수입니다. 수학식 24의 손실함수는 기울기 하강법을 적용하여 아래의 수학식 25로 표현할 수 있다.

[0282] [수학식 25]

$$\nabla_{\theta_k} \mathcal{L}_k(\theta_k) = \mathbb{E}_{(s,a,r,s') \sim U(\mathcal{M})} \left[(R + \gamma \max_{a'} Q^{\pi^{opt}}(s', a'; \theta_k^-) - Q(s, a; \theta_k)) \nabla_{\theta_k} Q(s, a; \theta_k) \right].$$

[0283]

[0285] 도 5는 본 발명의 일 실시예에 따른, UAV-BS 궤도 정책 최적화를 위한 경험 재생 딥 큐 네트워크 알고리즘을 설명하기 위한 도면이다.

[0286] 도 5를 참조하여, 딥 큐 네트워크 알고리즘에 대하여 보다 상세히 설명한다.

[0287] 1 단계(Step 1)에서, 먼저 네트워크 매개 변수를 무작위로 초기화하고, 오리지널 딥 큐 네트워크(DQN)와 동일한 네트워크 구조를 가지는 대상 딥 큐 네트워크 (DQN)를 설정한다(1-2 lines).

[0288] 이후, 각 훈련 에피소드 단계 e_t 에서, 에너지 효율 메트릭은 수학식 11을 이용하여 계산되며, 여기에서, 네비게이션 및 통신 파라미터는 수학식 1 내지 11을 이용한다(7 line).

[0289] E-greedy 정책 프레임워크의 일부로서, 탐사 단계에서 에너지 효율 메트릭은 보상 함수를 계산하는 데 사용되며, 행동은 탐험 확률 ϵ 를 가진 현재 딥 큐 네트워크(DQN)에서 파생된다(8 line).

[0290] 보상 함수는 수학식 14의 다른 제약과 함께 정보 연령(AoI) 컨텍스트를 고려하고, $t' = t + 1$ 인 다음 상태 $S_{t'}$ 로 전이한다(9 line).

[0291] 또한 탐사 단계를 통해, UAV-BS는 최고의 에너지 효율로 적절한 조치를 선택하는 데 도움이 되는 더 나은 보상 가치를 달성하기 위한 모든 공동 조치를 탐구 할 수 있다.

[0292] 그 후, 관찰 시퀀스의 상관 관계를 제거하기 위해, 리플레이 메모리 버퍼의 경험을 무작위로 섞는 미니 배치 방식을 적용하여, 관찰 데이터 분포의 변화를 평활화할 수 있다(10-11 lines).

[0293] 이후, 딥 큐 네트워크(DQN)를 훈련시키기 위해 훈련 손실 함수를 사용하는 확률적 경사 하강법 (SGD) 알고리즘을 채택하고, 네트워크 매개 변수 P와 네트워크 편향을 갱신할 수 있다(12-13 lines).

[0294] UAV-BS가 터미널 궤도에 도달하고 DQN 네트워크가 테스트를 위해 마지막으로 저장되면 훈련 프로세스를 종료한다(14 line).

[0296] D. Testing UAV-BS Trajectory Policy

[0297] 도 6은 본 발명의 일 실시예에 따른, UAV-BS 궤도 정책 최적화를 위한 경험 재생을 통한 DQN 테스트를 위한 알고리즘을 설명하기 위한 도면이다.

[0298] 도 6을 참조하면, 훈련 단계에서 저장된 UAV 궤도 정책 네트워크는, 시간 슬롯 t 에서 UAV-BS의 R_t 가 검색되는 곳에서 사용된다(1-3 line).

[0299] 테스트 단계에서의 공격의 일부로서, 활동 a_t 는 UAV-BS의 궤도와 DQN의 목표 값이 모두 업데이트되는 공동 UAV-BS 네비게이션을 위해 선택된다.

[0301] <PERFORMANCE EVALUATION>

[0302] 이하에서, 다양한 핵심 지표를 통해 본 발명의 일 실시예에 대한 성능 분석 실험 환경을 설명한다. 이후, 실험에서 얻은 결과를 설명하고 마지막으로 시뮬레이션 결과로부터의 관찰 결과를 개시한다.

[0303] 심층신경망(DNN, deep neural network)을 양성하기 위해서는 100개의 숨겨진 노드를 가진 2개의 완전히 연결된 (FC, fully connected) 숨은 층을 가진 신경망 구조를 가정한다. 또한 경험 재생 메모리 크기 M은 200으로 설정

한다. 시뮬레이션 결과는 100회 이상의 평균값과 정상값을 이용하여 산출된다.

[0305] A. Experiment Setting

[0306] 도 7은 본 발명의 일 실시예에 대한 시뮬레이션을 위한 세팅을 설명하는 도면이다. 도 7에 도시된 바와 같이, 시뮬레이션 파라미터를 설정하고 시뮬레이션을 수행한다.

[0308] B. Convergence Analysis of the Neural Network

[0309] 도 8은 본 발명의 일 시뮬레이션 결과에 따른, 시간 슬롯 수의 함수로 누적 보상 평가를 도시하는 그래프이다.

[0310] 도 8을 참조하여, 본 발명에 따른 보상 함수의 몇 가지 통계적 특성을 관찰한다. 접근 기울기는 수학적 19에서 설정된 보상으로 제안된 접근 방식이 안정화 된 후 UAV의 궤도 정책의 효과를 나타낸다. 또한 곡선의 최소 지점 (예를 들어, 타임 슬롯이 9인 지점)은 제안된 접근 방식이 개선되기 시작하기 전에 얼마나 많은 보상이 요구되는지를 도시한다. 이러한 현상의 이유는 UAV-BS가 고정 된 지점에서 시작하면 모든 UAV-BS가 무작위로 탐색하게 되고, 그에 따라 에너지 효율이 낮거나 겹치는 궤도 경유지를 쉽게 선택하게 되어 페널티가 증가하기 때문이다.

[0311] 타임 슬롯 10 이후 누적 된 보상이 플러스가 되면서 보상 설정은 증가 된 누적 되는 현상을 가진다. 이는 본 발명에 따른 접근 방식이 내비게이션 환경 학습 비용을 회수 할 때까지 UAV-BS가 공동으로 걸리는 시간을 나타낸다.

[0312] 한편, 도 5의 최소 및 제로 교차점은 긍정적인 보상과 부정적인 보상을 모두 고려하는 수학적 19의 보상 설계를 통해 제안된 알고리즘이 균형 잡힌 합리적 동작을 수행함을 나타낸다.

[0313] 누적 보상은 총 보상을 고려하여 측정되기 때문에 제안된 알고리즘은 마지막 시간 단계까지 각 단계에서 할인 보상을 최적화하여 실행 가능한 솔루션 (예를 들어, 실행 가능한 궤도 정책)으로 수렴함을 알 수 있다.

[0315] C. Finding Appropriate Discount Factor and AoI Threshold

[0316] 이하에서는, 재생 메모리 접근 방식을 사용하여 본 발명의 일 실시형태에 따라 제안된 딥 큐 네트워크(DQN)에 대한 적절한 할인 계수 및 정보연령(AoI) 임계값을 찾기 위한 시뮬레이션 결과를 설명한다.

[0317] 도 9는 시뮬레이션에 적용된 디스카운트 팩터와 정보연령(AoI) 임계값에 대한 예시를 도시하는 표를 도시하고 있다.

[0318] 도 9에 도시된 표 2 및 3에서, 평균 에너지 효율 메트릭을 고려하여 적절한 할인 계수와 AoI 임계 값을 평가하고 궤도 포인트 수 $|\mathcal{P}| = 14$ 를 설정한다.

[0319] 표 2에서, 할인 계수가 에너지 효율에 어떤 영향을 미치는지 도시한다. 보다 구체적으로, 리플레이 메모리가 있는 제안된 DQN의 경우 할인율이 증가하면, 에너지 효율 메트릭은 할인율 γ 이 0.5 일 때 1.73 %까지 올라간다. 반면에, 기존 DQN에 대한 에너지 효율 메트릭은 할인 계수 γ 이 0.5 일 때 0.23 %까지 약간 떨어지고 할인 계수 γ 이 0.6이 될 때까지 추세가 계속된다.

[0320] 마찬가지로, 제안된 접근 방식에 대한 에너지 효율 메트릭도 할인 계수 γ 이 0.6 일 때 약간 변동된다. 그러나, 이 시점까지 재생 메모리가 있는 제안된 DQN은 여전히 기존 DQN 접근 방식을 능가한다. γ 이 0.7 일 때 에너지 효율 메트릭은 재생 메모리가있는 제안된 DQN의 최고점에 있으며 기존 DQN의 성능도 크게 향상된다.

[0321] 그러나 γ 값이 높을수록 두 접근 방식의 성능이 크게 저하 된다. 이러한 현상의 이면에는 할인 계수 값이 낮으면 (예를 들어, $\gamma < 0.8$) 미래 보상은 그다지 중요하지 않으며 에이전트는 더 적은 단계로 진행하여 즉각적인 보상을 높이기 위한 동작을 수행하기 때문이다. 이것은, 합리적인 에너지 효율에 도달하는 데 적합하지 않다.

반면에 할인 계수 값이 증가함에 따라 (예를 들어, $\gamma > 0.7$) 에이전트는 시스템의 일시적인 동작을 포착하는 향후 조치가 가져올 수 있는 이점을 추구하도록 동작한다. 따라서, 할인 요소가 $\gamma > 0.7$ 이면 UAV-BS 에이전트는 미래의 보상에 너무 많은 관심을 기울여 결국 즉각적인 보상을 무시할 우려가 있다. 따라서 성능이 저하되고 UAV-BS의 배터리가 빠르게 소모되므로 페널티가 유발될 수 있다. 따라서 시뮬레이션에서 할인 계수 γ 은 0.7로 설정한다.

[0322] 표 3에서 할인 계수 γ 을 0.7로 고정한 후, UAV-BS에 대한 궤도 정책을 찾기 위해 적절한 AoI 임계 값을 선택하도록 추론하는 정규화 된 평균 보상 및 평균 AoI에 대한 다양한 AoI 임계 값의 효과를 설정한다.

[0323] AoI 임계 값 $\hat{\Delta}_b^{th}$ 을 4가지 수준, 즉, [0.3, 0.5, 0.7, 0.9]로 설정하여 두 성능 메트릭 간의 균형을 확인한다.

[0324] AoI 임계 값 $\hat{\Delta}_b^{th}$ 가 0.3일 때, 엄격한 AoI 임계 값으로 정규화 된 평균 보상이 최소 인 반면 AoI는 최고점에 있음을 알 수 있다. 이는, 엄격한 AoI 임계 값으로 인해 페널티가 증가하여 평균 누적 보상 측면에서 시스템 성능이 저하되기 때문에 정상적인 것이다. AoI 임계 값을 완화하면 (즉, $\hat{\Delta}_b^{th} > 0.5$), 평균 AoI가 크게 감소하고 평균 보상이 증가한다.

[0325] 한편, 임계값 $\hat{\Delta}_b^{th}$ 가 0.9 일 때, 보상 값은 최대이지만이 시점에서 AoI 임계 값의 영향은 사소해진다. 구체적으로, 이 시점에서 UAV-BS는 바람직하지 않은 지상 BS에서 정보 업데이트의 최신성을 무시하고 시스템을 작동할 수 있다.

[0326] 따라서 평균 보상과 AoI 메트릭 간의 균형을 맞추기 위해 허용 가능한 AoI 임계값 $\hat{\Delta}_b^{th}$ 을 0.7로 설정할 수 있다.

[0328] D. Experiment Results

[0329] 도 10은 본 발명에 대한 일 시뮬레이션 결과로서, 궤도 경유점의 수에 따른 평균 누적 보상을 도시하는 그래프이다.

[0330] 도 10에서 우리는 궤도 경유점 수가 증가함에 따라 평균 누적 보상 측면에서 각 방식을 비교할 수 있다. 적은 수의 궤도 지점 (예를 들어, 최대 $|\mathcal{P}| = 8$)의 경우 지리적 환경에서 경유점의 밀도가 더 적기 때문에 각 방식 간의 성능 차이가 뚜렷하지 않다. 그러나 궤도 지점의 수가 증가하면 접근 방식 간의 성능 격차가 크게 증가합니다.

[0331] 보다 구체적으로, 재생 메모리 접근 방식을 사용하는 제안된 딥 큐 네트워크(DQN)는 기존의 그리디 DQN 및 기준선 DQN 보다 2.84 % 및 1.91 %까지 상승한 성능을 발휘한다.

[0332] 고밀도 궤도 경유점 (예를 들어, $|\mathcal{P}| = 14$)인 경우, 재생 버퍼가 있는 제안된 DQN이 탐욕스러운 DQN과 기준선 DQN을 각각 5.08 % 및 4.02 % 이상 증가하여 접근 방식 간의 성능 격차가 더욱 증가한다.

[0333] 제안된 모델의 일 특징 중 하나는 UAV-BS 탐색을 위한 정책을 찾는 동안 정보연령(AoI) 메트릭을 최소화하는 것이다.

[0334] 도 11은 본 발명에 대한 일 시뮬레이션 결과로서, 다른 궤적 경유지 수에 대해 제안된 접근방식과 다른 접근방식의 평균 AoI를 비교하는 그래프이다. 도 11을 참조하여, 그리디 DQN 및 기준선 DQN 접근 방식과, 제안된 재생 메모리를 사용하는 DQN의 성능을 비교한다.

[0335] 도 11에서 볼 수 있듯이 제안된 접근 방식은 평균 정보연령(AoI) 를 1.21 % 및 1.17 %까지 줄이는 효과를 가지고 있다. 서로 다른 접근 방식 간의 성능 차이는 $|\mathcal{P}| = 8$ 까지 점진적으로 증가한다. $|\mathcal{P}| = 10$ 에서 접근 방

식의 성능은 실험에서 서로 다른 수의 궤도 점으로 시뮬레이션을 독립적으로 실행하기 때문에 약간 변동함을 알 수 있다. 그러나 궤도 경유점 수가 증가함에 따라 AoI 메트릭을 줄이기 위해 제안된 접근 방식의 추세가 계속되며, $|P| = 14$ 에서 제안된 DQN의 성능이 그리디 DQN 및 기준선 DQN 접근 방식에 대하여 0.72 % 및 1.29 % 증가된 격차를 가진다.

[0337] 도 12는 본 발명에 대한 일 시뮬레이션 결과로서, 다른 궤적 경유지 수에 대해 제안된 접근방식과 다른 접근방식의 평균 에너지 효율을 비교하는 그래프이다.

[0338] 도 12를 참조하여, 평균 에너지 효율에 관한 기준 접근 방식과 재생 메모리와 함께 제안된 DQN의 성능을 비교한다.

[0339] 재생 메모리가 있는 제안된 DQN이 기준 접근 방식보다 점차적으로 더 나은 성능을 발휘하는 것이 관찰된다. 궤도 경유점의 수가 상대적으로 높을 때 (예를 들어, $|P| = 10$), 재생 메모리가 있는 제안된 DQN은 그리디 DQN과 기준선 DQN보다 각각 0.32 % 및 1.97 % 에너지 효율이 높은 것을 알 수 있다.

[0340] 반면에 mTSP는 모든 접근 방식에 비해 최악의 성능을 보입니다. 그 이유는 궤도 경유점의 수가 많을수록 UAV-BS의 궤도 공간이 증가함에 따라 mTSP의 경우 UAV-BS의 평균 에너지 소비가 크게 증가하기 때문이다.

[0341] 각 시간 슬롯 t 에서 각 UAV-BS는 탐색을위한 에너지 비용이 발생하고 배터리가 다른 모든 접근 방식에 비해 훨씬 빠르게 소모되는 IoT 디바이스에 서비스를 제공하기 위해, 하나의 궤도 경유점에서 인접한 경유점으로 이동해야 한다.

[0342] 궤도 경유점이 많을수록, 예를 들어, $|P| = 14$, 제안된 접근 방식이 보다 에너지 효율적임을 알 수 있다. 즉, 비교 대상에 대해 최대 3.6 %, 3.13 %, 및 7.87 % 효율적이다.

[0344] 궤도 지점에서 IoT 디바이스의 수가 증가함에 따라 프론트 홀(Front-haul) 용량이 제한되므로 대역폭을 효율적으로 활용해야 한다. 도 13는 본 발명에 대한 일 시뮬레이션 결과로서, 다른 궤적 경유지 수에 대해 제안된 접근방식과 다른 접근방식의 평균 대역폭 효율성을 비교하는 그래프로서, 도 13을 참조하여, 그리디 DQN 및 기준선 DQN을 비교하여 제안된 접근 방식의 효율성을 다양한 궤도 지점 수와 다양한 IoT 디바이스 밀도와 비교한다.

[0345] 리플레이 버퍼가 있는 제안된 DQN은 서로 다른 궤도 경유점을 가로 질러 이동하면서 프론트 홀 및 백홀 대역폭을 효율적으로 활용한다. 평균 대역폭 효율성은 환경에서 궤도 지점의 수가 적고 지점 사이의 거리가 클 때 매우 유사하다. 따라서 다른 UAV-BS에서 제공하는 IoT 디바이스에서 수신 된 간섭 수준은 모든 방식에서 적다.

[0346] 그러나 IoT 디바이스 수가 증가함에 따라 (예를 들어, $|P| = 10$) 조밀한 궤도 경유점 네트워크에서는 모든 접근 방식이 간섭에 직면하고 대역폭 효율성이 감소하는 경향을 가진다. 그럼에도 불구하고 재생 버퍼가 있는 제안된 DQN은 여전히 대역폭 효율성을 보장하는 측면에서 그리디 DQN과 기준 DQN을 2.41 % 및 2.87 % 증가하는 효율을 가진다.

[0348] 도 14는 본 발명에 대한 일 시뮬레이션 결과로서, 다른 궤적 경유지 수에 대해 제안된 접근방식과 다른 접근방식의 평균 UAV 리소스 활용도를 비교하는 그래프이다.

[0349] 도 14를 참조하면, 제안된 접근 방식과 다른 두 가지 기준 접근 방식 하에서 UAV-BS 또는 네트워크 자원의 활용을 도시한다. 궤적 웨이 포인트 수가 증가함에 따라 UAV-BS 당 연결 수가 증가하여 UAV-BS의 네트워크 리소스를 사용하는 IoT 디바이스 수도 증가한다. 고정 된 수의 UAV-BS (예를 들어, $|U| = 3$)의 경우 다른 궤도 웨이 포인트의 IoT 디바이스는 UAV-BS가 제공하는 최대 네트워크 리소스를 활용하는 경향이 있다. 그러나 재생 버퍼가 있는 제안된 DQN은 기준선 접근 방식보다 궤적 웨이 포인트를 더 효율적으로 다루기 때문에, UAV-BS에서 제공하는 네트워크 리소스는 탐욕스러운 DQN 및 기준선 DQN에 비해 9.26 % 및 4.71 % 더 효율적으로 활용됨을 알 수 있다.

- [0351] 도 15는 본 발명에 대한 일 시뮬레이션 결과로서, 상이한 AoI 임계 값 하에서 동작하는 UAV-BS의 평균 에너지 효율을 나타내는 그래프이다.
- [0352] 도 15를 참조하면, 상대적으로 완화 된 AoI 임계 값 $\Delta_b^{th} = 7$ 에서, 제안된 접근 방식을 사용하는 UAV-BS의 에너지 효율이, 그리디 DQN과 기준선 DQN보다 각각 1.63 % 및 0.95 % 더 에너지 효율적임을 알 수 있다. 즉, 제안된 접근 방식을 사용하는 UAV-BS의 에너지 효율은 그리디 DQN과 기준선 DQN에 대하여 1.63 % 및 0.95 % 더 에너지 효율적이다.
- [0354] 이하에서는, 본 발명의 일 실시예에 따른 다중 무인항공기 네비게이션 최적화 방법에 대하여 설명한다. 이하에서 설명하는 본 발명의 일 실시예에 따른 다중 무인항공기 네비게이션 최적화 방법은 도 1 내지 도 15를 참조하여 상술한 무인항공기 시스템에 포함된 지상 기지국에서 수행된다. 따라서, 도 1 내지 도 15를 참조하여 상술한 설명을 참조하여 쉽게 이해할 수 있다.
- [0355] 도 16은 본 발명의 일 실시예에 따른 다중 무인항공기 네비게이션 최적화 방법을 설명하는 순서도이다.
- [0356] 도 16을 참조하면, 지상 기지국(BS)은 지상 기지국이 무인항공기 기지국의 전송 범위 내에 존재하면, 통신을 통하여 상기 무인항공기 기지국으로부터 정보 업데이트를 받고 정보연령 메트릭을 계산 할 수 있다(S1610).
- [0357] 지상 기지국(BS)은 각각의 무인항공기 기지국의 궤도, 에너지 효율 및 정보연령 조건을 설정 할 수 있다(S1620).
- [0358] 지상 기지국(BS)은 에너지 효율 및 정보연령의 적용 받는 무인항공기 기지국 중계망의 총 에너지 효율을 최대화 하도록, 각각의 무인항공기 기지국의 궤도 경로 정책을 찾기 위한 강화 학습을 수행 할 수 있다(S1630).
- [0359] 일 실시예에서, 지상 기지국(BS)은, 정보연령 메트릭을 계산하는 단계 S1610이전에, 무인항공기 기지국으로부터 제공된 수신전력을 계산하고, 수신전력을 이용하여 무인항공기 기지국과 상기 지상 기지국 간의 백홀 용량을 계산하여 백홀 링크를 생성하고, 시간 t에서 백홀 링크를 사용하는 동안 상기 무인항공기 기지국의 전송 에너지를 계산할 수 있다. 이에 대해서는, 수학적 1 내지 수학적 7-1을 참조하여 상술한 내용을 참조하여 이해할 수 있다.
- [0360] 단계 S1620에 대한 일 실시예에서, 지상 기지국(BS)은, 각각의 무인항공기 기지국이 주어진 시간대에서 궤도 지점의 하위 집합만 다룰 수 있도록 설정하고, 궤도 지점으로 구성된 각각의 무인항공기 기지국에 대한 서브셋을 제한요건에 만족하도록 설정할 수 있다.
- [0361] 추가로, 지상 기지국(BS)은, 정보 업데이트가 발생하는 지상 기지국 궤도 지점은 제외한, 무인항공기 기지국의 오버랩되지 않는 궤도를 상기 무인항공기 기지국의 궤도 조건으로 설정하고, 모든 궤도 지점이 상호 의존적으로 포함된 무인항공기 기지국의 공동 궤도를 구성할 수 있다. 지상 기지국(BS)은, 통신 및 이동 에너지가 최소 에너지 효율 임계값 보다 크도록 무인항공기 기지국의 총 에너지 효율을 보장할 수 있고, 정보 업데이트의 평균 최신성이 정보연령 임계값 미만을 만족하도록 보장할 수 있다.
- [0362] 단계 S1630에 대한 일 실시예에서, 지상 기지국(BS)은 무인항공기 기지국의 궤도 정책을 위한 4차원 상태 공간을 설정하고, 무인항공기 기지국은 각각 궤도 및 통신제약을 충족하면서 하나의 실행 가능한 상태에서 다음 상태로의 항해를 설정하여 행동 공간을 설정할 수 있다. 지상 기지국(BS)은 임의의 상태에서 복수의 무인항공기 기지국이 각각 설정한 복수의 행동 공간에 대하여 경유지의 충돌 여부를 기준으로 보상을 설정할 수 있다.
- [0363] 단계 S1630에 대한 일 실시예에서, 지상 기지국(BS)은 경유지의 충돌 여부를 기준으로 보상을 설정하기 위하여, 무인항공기 기지국 네비게이션에 필요한 상황별 특징을 추출하는 데 사용되는 상태 공간의 차수를 줄이고, 무인항공기 기지국 에이전트가 관찰하는 상태 전환을 저장할 수 있다. 지상 기지국(BS)은 에너지 및 정보연령 지표의 적용을 받는 무인항공기 기지국 중계망의 총 에너지 효율을 최대화 하도록 무인항공기 기지국의 궤도 경로 정책을 찾기 위한 강화 학습을 수행할 수 있다. 이에 대해서는, 수학적 18-1 내지 수학적 25를 참조하여 상술한 내용을 참조하여 이해할 수 있다.

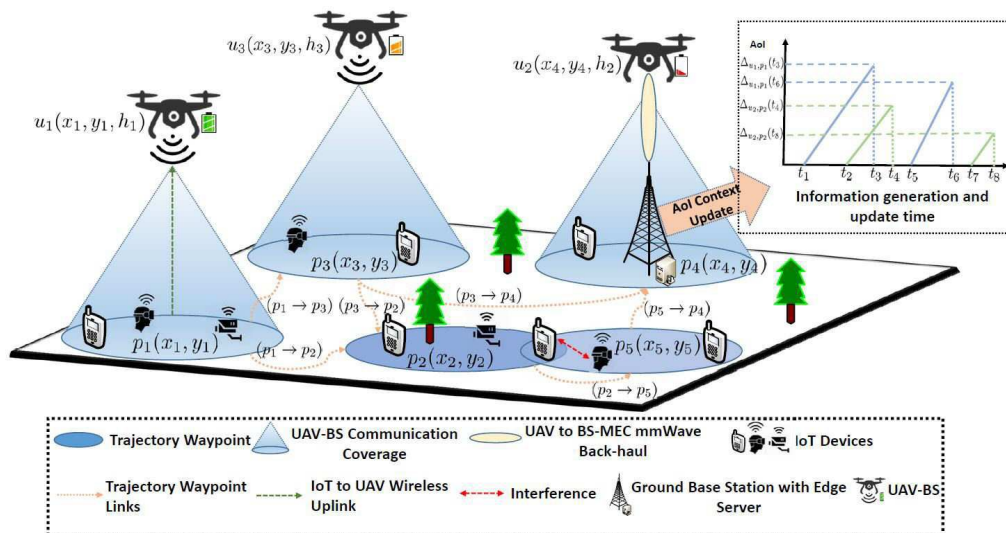
[0365] 이상에서 설명한 본 발명은 전술한 실시예 및 첨부된 도면에 의해 한정되는 것이 아니고 후술하는 특허청구범위에 의해 한정되며, 본 발명의 구성은 본 발명의 기술적 사상을 벗어나지 않는 범위 내에서 그 구성을 다양하게 변경 및 개조할 수 있다는 것을 본 발명이 속하는 기술 분야에서 통상의 지식을 가진 자는 쉽게 알 수 있다.

부호의 설명

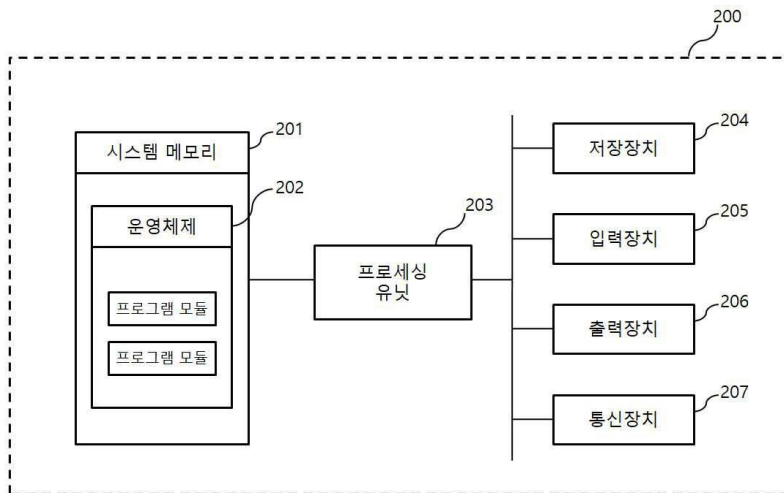
[0367] 100 : 무인항공기
 110 : 구동부 120 : 통신부
 130 : 제어부 140 : 배터리
 150 : 메모리
 200 : 컴퓨팅 장치
 201 : 시스템 메모리 202 : 운영체제
 203 : 프로세싱 유닛 204 : 저장소
 205 : 입력장치 206 : 출력장치
 207 : 통신장치
 300 : 심층강화학습 모델
 310 : 심층 신경 네트워크 320 : 경험 재생 메모리
 330 : 강화 학습 프레임워크

도면

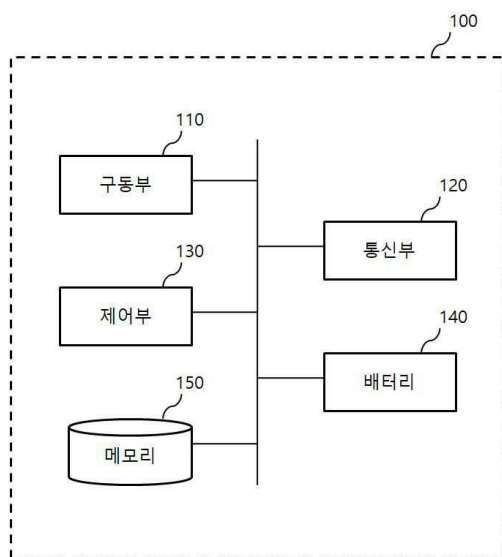
도면1



도면2

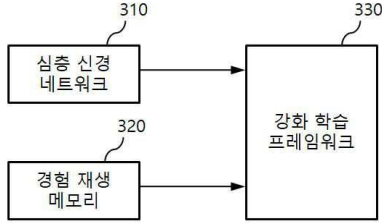


도면3



도면4

300



도면5

Algorithm 1: Training DQN with experience replay for UAV-BS Trajectory Policy Optimization in Navigation

```

1 Step 1: Initialization
2 Initialize  $Q(s, a; \theta)$ ,  $\mathcal{M}$ , target DQN parameters  $\theta^-$ 
   and construct DQN
3 Step 2: Training DQN with experience replay
4 for  $e = 1, \dots, E$  do
5     Initialize  $\mathcal{S}$ 
6     for  $t = 1, \dots, T$  do
7         Calculate the energy efficiency metric of the
           UAV-BSs using (11)
8         Select action  $a_t$  with given probability  $\epsilon$  and
           calculate instant reward  $R_t(a_t)$  using (19)
9         Observe instant reward  $R_t(a_t)$  and next state
            $s_{t'}$ 
10        Store experience  $(s_t, s_{t'}, a_t, R_t(a_t))$  in the
           experience replay memory  $\mathcal{M}$ 
11        Randomly sample minibatch of experiences
           from  $\mathcal{M}$ 
12        Adopt stochastic gradient descent (SGD) to
           train the DQN using loss function in (24)
13        Update  $\theta$ ,  $\theta^-$ ,  $Q(s, a; \theta)$ , and  $Q^{\pi^{opt}}(s, a; \theta^-)$ 
14 Store the Q-network

```

도면6

Algorithm 2: Testing DQN with experience replay for UAV-BS Trajectory Policy Optimization in Navigation

```

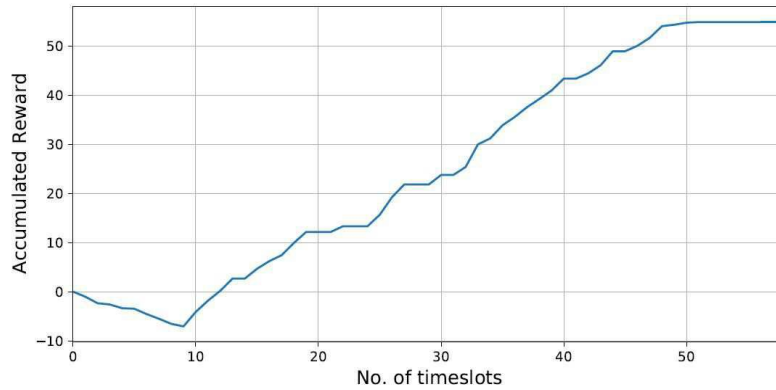
1 Load the stored Q-network of Alg. 1
2 for  $t = 1, \dots, T$  do
3   Retrieve  $R_t$  of the UAV-BSs at time slot  $t$ 
4   Retrieve and select joint UAV-BS action
      $a_t = \arg \max_a Q^{\pi^{opt}}(s_t, a; \theta)$ 
5   Update trajectory of UAV-BSs based on joint
     action index and target values of DQN

```

도면7

Simulation Parameters	Values
No. of UAV-BS	3
No. of MEC-BS	1
No. of IoT devices	100
No. of Trajectory points	[6, 14]
Max UAV-BS heights	[140, 250] (m)
Maximum UAV-BS velocity and acceleration	100 m/s and , 5 m/s ²
Radius of UAV-BS	300 (m)
f_c^{mmWave}	28 GHz [36]
$\beta_{b,u}^{mmWave}$	20 * 100 MHz[36]
f_c, β_u	2 GHz, 20 MHz [36]
$P_{b,u}^{tx}, \sigma^2, \gamma_{th}$	20 dBm [36], -100 dBm, 5 dB
For urban scenario, $\alpha, \hat{\alpha}, \epsilon, \bar{\epsilon}, k_1, k_2$	9.61, 0.16, 1, 20, 9.26 × 10 ⁻⁴ , 2250 [36] [38]
Normalized AoI threshold Δ_b^{th}, α_1	[0.3, 0.9], 1

도면8



도면9

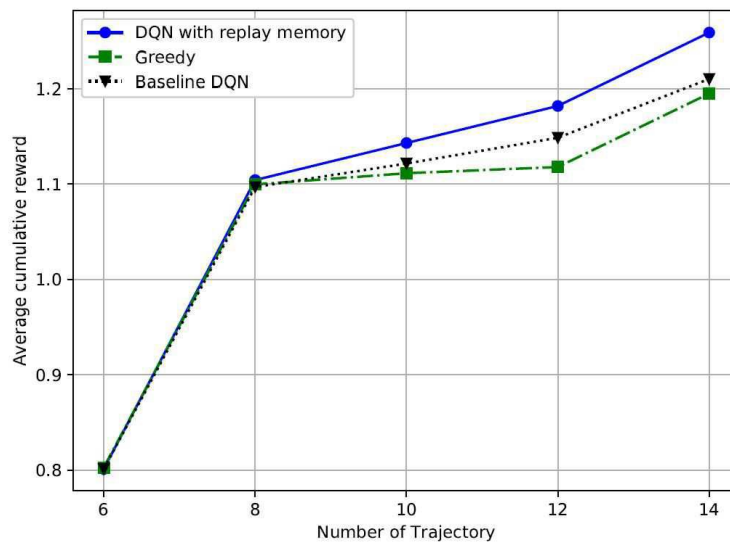
TABLE II: Effects of different discount factors over the average energy efficiency (EE) of the proposed and the baseline DQN approaches.

Discount factor γ	DQN with replay memory (Average EE)	Baseline DQN (Average EE)
0.4	0.483663	0.487285
0.5	0.492033	0.486141
0.6	0.489998	0.480957
0.7	0.4927502	0.489459
0.8	0.490233	0.486083
0.9	0.486811	0.484898

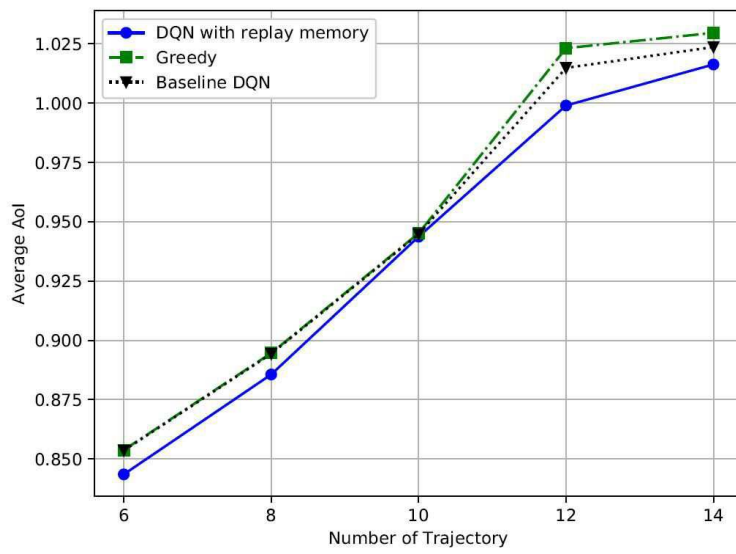
TABLE III: Trade-off analysis between normalized average reward and normalized average AoI for different AoI thresholds.

AoI threshold $\hat{\Delta}_b^{th}$	Normalized average reward (Proposed)	Normalized average AoI (Proposed)
0.3	0.011284	0.449802
0.5	0.116763	0.420419
0.7	0.226587	0.385290
0.9	1.0	1.0

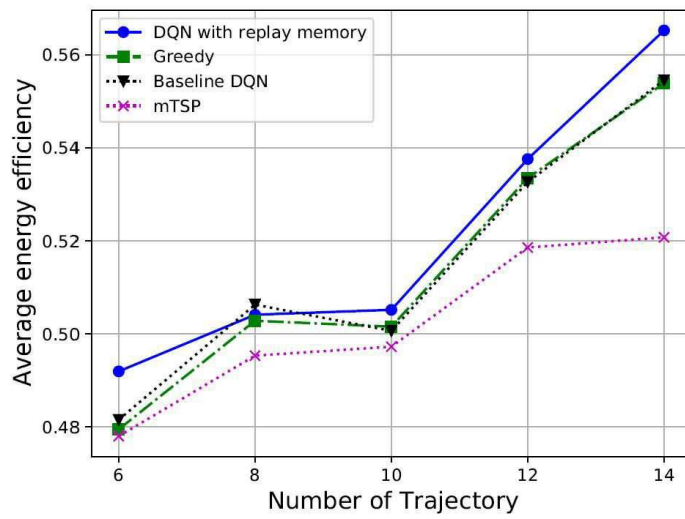
도면10



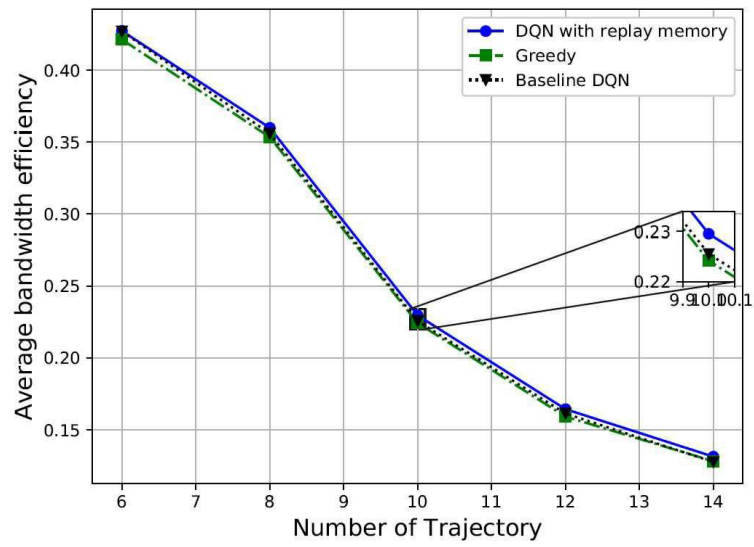
도면11



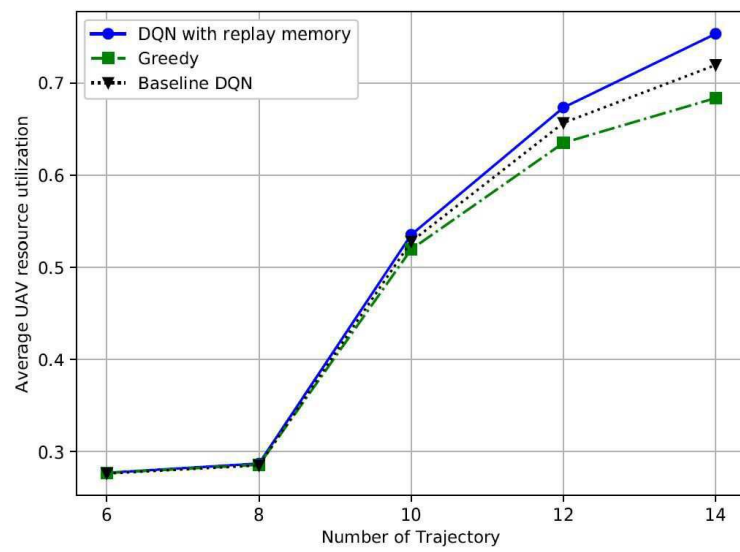
도면12



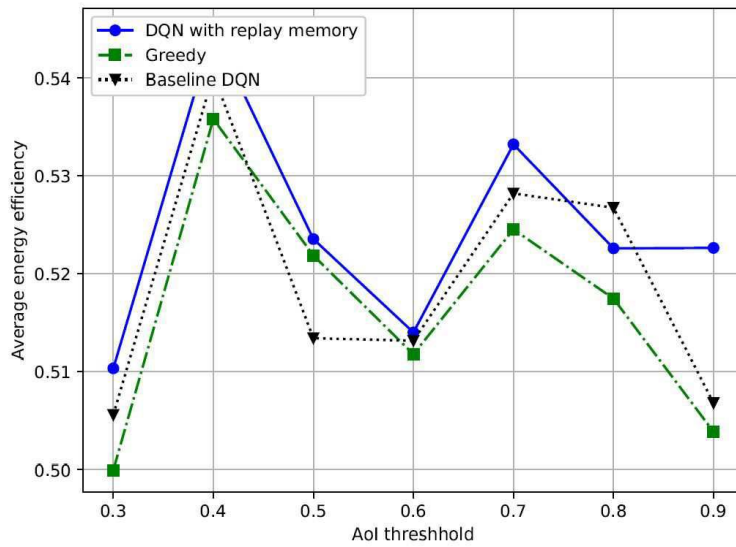
도면13



도면14



도면15



도면16

