

UAV 기반 ISAC 시스템에서 통신 성능 향상을 위한 강화학습 기반 UAV 궤적 최적화 기법

송연주, 김준수, 김수민*
한국공학대학교

E-mail: {syj4700, junsukim, *suminkim}@tukorea.ac.kr

A Reinforcement Learning-Based UAV Trajectory Optimization Scheme for Improving Communication Performance in UAV-Based Integrated Sensing and Communication Systems

Youn Joo Song, Junsu Kim, Su Min Kim*
Tech University of Korea

요 약

본 논문에서는 UAV 기반 ISAC 시스템의 통신 성능을 향상시키기 위한 강화학습 기반 UAV 궤적 최적화 기법을 제안한다. 제안 기법은 강화학습 알고리즘을 통해 UAV가 최소한의 센싱 성능을 만족하면서 통신 성능을 최대화 할 수 있는 위치로 동적으로 조정된다. 시뮬레이션 결과, UAV는 강화학습을 통해 최적의 위치로 이동하여 통신 성능을 향상시킬 수 있음을 확인하였다.

I. 서 론

최근, 센싱과 통신 기능을 통합하여 주파수 자원을 효율적으로 사용할 수 있는 통합 센싱 및 통신 (integrated sensing and communication, ISAC)이 6G 핵심 기술로 대두되고 있다. 더 나아가, 높은 이동성과 강력한 line-of-sight (LoS) 채널을 확보할 수 있는 unmanned aerial vehicle (UAV) 기반의 ISAC 시스템에서는 성능이 더욱 향상된 센싱 및 통신 서비스를 제공할 수 있다 [1]. 그러나 기존의 연구들은 고정 길이의 센싱 구간을 고려하여 센싱 성능이 충분히 만족되더라도 통신 성능을 추가로 향상시키기 어렵다는 한계를 지닌다. 유동적인 ISAC 프레임 구조를 이용하여 센싱 구간 최적화를 수행한다면 추가적인 통신 성능 향상을 기대할 수 있다.

본 논문에서는 센싱 성능 요구도를 만족하면서 센싱 구간의 길이를 최소화하여 통신 성능을 향상시키는 강화학습 기반의 UAV 궤적 최적화 기법을 제안한다.

II. 강화학습 기반 UAV 궤적 최적화 기법

본 논문에서는 그림 1과 같이 UAV가 K 명의 지상 사용자들에게 통신 서비스를 제공하면서, J 개의 레이더 타겟을 감지하는 시나리오를 가정한다. UAV는 안테나 개수가 $M = M_x \times M_y$ 인 균일 평면 배열 안테나 (uniform plane array, UPA)를 탑재하고 있으며, 고정된 높이 H 에서 xy 평면을 따라 이동한다고 가정한다. 그림 2는 제안된 ISAC 시스템의 프레임 구조를 나타내고, L 개의 ISAC 프레임으로 구성된다. 전체 프레임의 지속 시간은 T 초이며, 각 ISAC 프레임은 N 개의 타임슬롯으로 구성되며 각 타임슬롯의 길이는 $\delta_t = \frac{T}{N}$ 초이다. 통신 성능을 최대화 하기 위해, 각 ISAC 프레임의 초기 타임슬롯에서는 레이더 성능 요구도를 만족될 때까지 센싱을 수행하고, 이후 타임슬롯에서는 통신 데이터를 전송한다.

본 논문에서는 UAV와 사용자 사이의 채널 모델로 LoS 확률에 따른 감쇄를 고려한 air-to-ground (A2G) path loss model을 (1)과 같이 사용한다 [2].

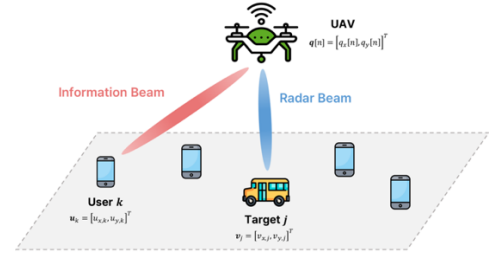


그림 1. 시스템 모델.

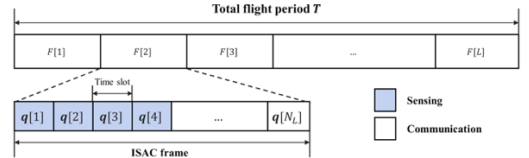


그림 2. ISAC 프레임 프로토콜.

$$\beta_k(\mathbf{q}[n], \mathbf{u}_k) = 20 \log_{10} \left(\frac{4\pi d(\mathbf{q}[n], \mathbf{u}_k) f_0}{c} \right) + P_{LoS} \cdot \eta_{LoS} + P_{NLoS} \cdot \eta_{NLoS} \quad (1)$$

이때, $\mathbf{q}[n]$ 은 n 번째 타임 슬롯의 UAV의 위치, \mathbf{u}_k 는 사용자 k 의 위치, $d(\mathbf{q}[n], \mathbf{u}_k)$ 는 UAV와 사용자 k 간의 거리이다. 또한, c 는 빛의 속도, f_0 는 반송파 주파수, P_{LoS} 와 P_{NLoS} 는 각각 LoS, non-LoS 확률로, $P_{LoS} = 1/(1 + \alpha \exp(-\beta(\theta - \alpha)))$, $P_{NLoS} = 1 - P_{LoS}$ 이다. 이때, η_{LoS} , η_{NLoS} 는 각각 LoS와 non-LoS인 경우의 추가적인 감쇄 손실이며, α, β 는 환경에 따라 달라지는 상수이다. 사용자 k 로의 UAV의 송신 배열 응답 벡터는 (2)와 같다.

$$\mathbf{a}^H(\mathbf{q}[n], \mathbf{u}_k) = \left[1, \dots, e^{-\frac{j2\pi(M_x-1)d_x\Phi(\mathbf{q}[n], \mathbf{u}_k)}{\lambda}} \right] \otimes \left[1, \dots, e^{-\frac{j2\pi(M_y-1)d_y\Omega(\mathbf{q}[n], \mathbf{u}_k)}{\lambda}} \right]. \quad (2)$$

이때, $d_x = d_y = \frac{\lambda}{2}$, λ 는 반송 파장, $\Phi(\mathbf{q}[n], \mathbf{u}_k) = \sin(\phi(\mathbf{q}[n], \mathbf{u}_k)) \cos(\theta(\mathbf{q}[n], \mathbf{u}_k))$,

$\Omega(\mathbf{q}[n], \mathbf{u}_k) = \sin(\phi(\mathbf{q}[n], \mathbf{u}_k)) \sin(\theta(\mathbf{q}[n], \mathbf{u}_k))$ 이다. 여기서 $\phi(\mathbf{q}[n], \mathbf{u}_k)$ 는 $\mathbf{q}[n]$ 에서 \mathbf{u}_k 까지의 신호의 zenith angle of departure, $\theta(\mathbf{q}[n], \mathbf{u}_k)$ 는 그에 대응하는 azimuth angle of departure 이고, \otimes 는 element-wise 벡터곱 연산자이다. UAV에서 사용자 k 까지의 채널과 센싱 타겟 j 까지의 채널은 각각 (3a), (3b)와 같이 정의된다.

$$\mathbf{h}_{c,k}^H(\mathbf{q}[n], \mathbf{u}_k) = \sqrt{\beta_k(\mathbf{q}[n], \mathbf{u}_k)} e^{-j\frac{2\pi d(\mathbf{q}[n], \mathbf{u}_k)}{\lambda}} \mathbf{a}^H(\mathbf{q}[n], \mathbf{u}_k), \quad (3a)$$

$$\mathbf{h}_{r,j}^H(\mathbf{q}[n], \mathbf{v}_j) = \frac{\mathbf{a}^H(\mathbf{q}[n], \mathbf{v}_j)}{d(\mathbf{q}[n], \mathbf{v}_j)}. \quad (3b)$$

이때, \mathbf{v}_j 는 센싱 타겟 j 의 위치이다. n 번째 타임 슬롯에서 UAV의 송신 신호는 (4)와 같다.

$$\mathbf{x}[n] = \mathbf{w}_c[n]s_k[n]. \quad (4)$$

여기서 $\mathbf{w}_c[n] = \sqrt{P_{max}} \frac{\mathbf{h}_{c,k}}{\|\mathbf{h}_{c,k}\|}$ 는 통신 신호를 위한 송신 빔포밍 벡터, P_{max} 는 최대 송신 전력, $s_k \sim \mathcal{CN}(0, 1)$ 는 정보 신호이다. 이에 따라, n 번째 타임 슬롯에서 사용자 k 가 수신 신호는 (5)와 같다.

$$y_k[n] = \mathbf{h}_{c,k}^H(\mathbf{q}[n], \mathbf{u}_k)(\mathbf{w}_c[n]s_k[n]) + n_k[n]. \quad (5)$$

$n_k[n] \sim \mathcal{CN}(0, \sigma^2)$ 은 사용자 k 의 수신기에서의 가우시안 백색 잡음을 나타낸다. (5)로부터, n 번째 타임 슬롯에서 사용자 k 의 signal-to-noise ratio (SNR), 데이터 전송률은 각각 (6), (7)과 같다 [3].

$$\gamma_k[n] = \frac{|\mathbf{h}_{c,k}^H(\mathbf{q}[n], \mathbf{u}_k)\mathbf{w}_c[n]|^2}{\sigma^2}, \quad (6)$$

$$R_k[n] = \log_2(1 + \gamma_k[n]). \quad (7)$$

레이더 성능 지표로는 레이더 위치 추정 오차의 하한 값인 Cramer-Rao lower bound (CRLB)을 사용한다. 타겟 j 의 SNR과 CRLB는 각각 (8), (9)와 같이 주어진다 [4].

$$\gamma_j[n] = \frac{|\mathbf{h}_{r,j}^H(\mathbf{q}[n], \mathbf{v}_k)\mathbf{w}_r[n]|^2}{\sigma^2}, \quad (8)$$

$$\sigma_{CRLB}^2[n] \equiv \frac{c^2}{32\pi^2 B_s^2 \tau_s \gamma_j[n]}. \quad (9)$$

이때, $\mathbf{w}_r[n] = \sqrt{P_{max}} \frac{\mathbf{h}_{r,j}}{\|\mathbf{h}_{r,j}\|}$, $B_s = \frac{B_s}{\sqrt{12}}$ 는 제곱 평균 제곱근 대역폭, τ_s 는 센싱 지속 시간이다.

본 논문에서는 최소 센싱 성능을 CRLB의 임계값 $\sigma_{CRLB,th}^2$ 로 정의하고, 이를 만족하면서 통신 성능지표인 평균 데이터 전송률을 최대화하는 것을 목표로 강화학습을 수행한다. 제안하는 강화학습 프레임워크의 agent는 UAV, state는 UAV의 x, y 좌표, action은 UAV의 x, y 축의 +/- 변화량으로 총 4가지이다. 강화학습 에피소드가 진행됨에 따라 UAV는 매 ISAC 프레임의 초기 타임슬롯에서는 $\sigma_{CRLB,th}^2$ 를 만족하면서 센싱 시간을 최소화하는 위치로 이동한 후, $R_k[n]$ 이 최대화 되는 곳으로 이동하는 과정을 반복한다. 강화학습 알고리즘으로는 Q-learning과 신경망을 결합하여 에이전트가 상태-행동 쌍의 Q값을 학습하여 학습의 안정성을 높일 수 있는 deep Q-network (DQN)을 사용하였다.

III. 성능 검증

시뮬레이션에서, $H = 40$ m, $\sigma^2 = -100$ dB, $P_{max} = 1$ W, $B = 2.16$ GHz, $T = 80$ s, $\delta_t = 0.25$ s, $\sigma_{CRLB,th}^2 = 12e^{-5}$ 로 설정하였다. 성능 비교를 위해, 다음의 2가지 비교방안을 선정하였다.

- i) **Straight flight (SF)**: UAV가 일정 속도로 시작 위치 \mathbf{q}_I 에서 최종 위치 \mathbf{q}_F 까지 일직선으로 이동한다.
- ii) **Fly-Hover-Fly (FHF)**: UAV가 시작 위치 \mathbf{q}_I 에서 중간 위치 \mathbf{q}_H 에서 정지한 후 최종 위치 \mathbf{q}_F 로 이동한다.

그림 3은 강화학습을 통해 최적화된 UAV의 이동 궤적을 나타낸다. UAV는 초기 타임 슬롯에서 $\sigma_{CRLB,th}^2$ 값을 만족하는 위치로 이동한 후, 통신 성능을 최대화 할 수 있는 경로를 따라 이동한다. 그림 4는 최적화된 궤적에 따른 성취한 데이터 전송률 결과이다. 세가지 방안을 비교했을 때, 제안하는 강화학습 방식을 따른 경우 성취한 데이터 전송률이 가장 높게 나타났다. 또한 그림 5에서 보이는 바와 같이, $\sigma_{CRLB,th}^2$ 값이 클수록 센싱 시간을 줄여 더욱 높은 데이터 전송률을 성취할 수 있다.

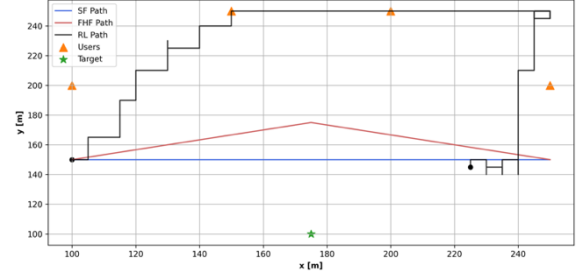


그림 3. UAV의 이동 궤적 ($\sigma_{CRLB,th}^2 = 12e^{-5}$).

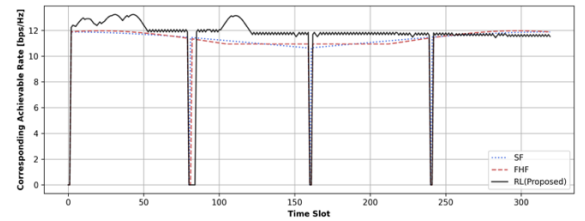


그림 4. 성취 데이터 전송률 ($\sigma_{CRLB,th}^2 = 12e^{-5}$).

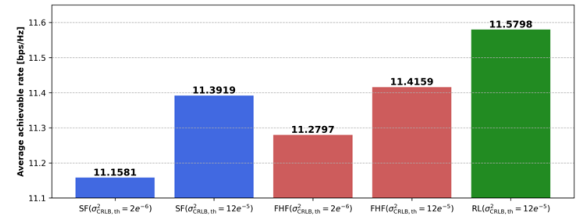


그림 5. 평균 데이터 전송률

IV. 결론

본 논문에서는 UAV 기반 ISAC 시스템에서 통신 성능을 최대화하기 위한 강화학습 기반 UAV 궤적 최적화 기법을 제안하였다. 제안 기법에서 UAV는 강화학습을 통해 최소한의 레이더 센싱 성능을 만족하면서, 통신 성능이 최대화 되는 최적의 위치로 이동하여 성능 향상을 성취하였다.

ACKNOWLEDGMENT

본 연구는 과학기술정보통신부 및 정보통신기획평가원의 ICT 혁신인제 4.0 사업 (IITP-2024-RS-2022-00156326) 과 과학기술정보통신부의 재원으로 한국연구재단의 지원을 받아 수행된 연구임 (2022R1F1A1074556).

참고 문헌

- [1] K. Meng et al., "UAV-Enabled Integrated Sensing and Communication: Opportunities and Challenges," *IEEE Wireless Communications*, pp. 1-9, Apr. 2023.
- [2] S. u. Rahman, G. Kim, Y. Cho, and A. Khan, "Positioning of UAVs for throughput maximization in software-defined disaster area UAV communication networks," *Journal of Communications and Networks*, vol. 20, no. 5, pp. 452-463, Oct. 2018.
- [3] K. Meng, Q. Wu, S. Ma, W. Chen, K. Wang, and J. Li, "Throughput Maximization for UAV-Enabled Integrated Periodic Sensing and Communication," *IEEE Transactions on Wireless Communications*, vol. 22, no. 1, pp. 671-687, Jan. 2023.
- [4] J. M. Park, J. Cho, S. Noh, and H. Yu, "Optimal Pilot and Data Power Allocation for Joint Communication-Radar Air-to-Ground Networks," *IEEE Access*, vol. 10, pp. 52336-52342, May. 2022.