

무인항공기 기반 통신-레이더 융합 시스템에서 동작수명 연장을 위한 강화학습 기반 위치제어 기법

송연주, 우수연, 김준수, 김수민*
한국공학대학교

E-mail: {syj4700, jysy2125, junsukim, *suminkim}@tukorea.ac.kr

A Reinforcement Learning-Based Location Control Method for Operational Lifetime Extension in UAV-Based Joint Communication and Radar Systems

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

요약

본 논문에서는 무선 정보 및 전력 동시전송 (simultaneous wireless information and power transfer, SWIPT) 기술이 적용된 unmanned aerial vehicle (UAV) 기반 joint communication and radar (JCR) 시스템의 동작수명을 최대화하기 위한 강화학습 기반 UAV 위치제어 기법을 제안한다. 성능평가 결과, 제안 기법은 episode 가 진행됨에 따라 UAV 중계기가 목표 데이터 전송률을 만족하면서도 동작 수명을 최대화하기 위한 위치로 UAV 를 이동시키는 것을 확인하였다. 또한, SWIPT 에너지 수확 비율이 클수록 더 긴 UAV 동작수명을 성취함을 보였다.

I. 서론

Joint communication and radar (JCR)은 하나의 신호 파형을 이용하여 통신과 레이더 기능을 동시에 수행하는 주파수 공유 시스템으로서 무선 단말의 기하급수적 증가에 따른 주파수 자원 부족 문제의 해결 방안 중 하나로 주목받고 있다 [1]. 최근에는 unmanned aerial vehicle (UAV)를 활용한 이동통신 시스템에 대한 연구가 활발히 진행되고 있고 [2], 이를 접목한 UAV 기반의 JCR 시스템에서는 UAV 가 레이더 시스템의 높은 전력과 강한 방향성 특성을 이용해 통신 신호를 증폭하여 전달하는 중계기 역할을 수행할 수 있다. [3]에서는 JCR 시스템의 성능을 최대화하기 위해 UAV 에 대한 air-to-ground (A2G) 채널모델을 고려한 강화학습 기반 UAV 궤적 최적화 기법을 제안하였다. 또한, [4]에서는 JCR 시스템에서 UAV 의 데이터 송신전력을 최소화하여 한정된 배터리 수명을 최대화하기 위한 강화학습 프레임워크를 제안하였다. 그러나 추가적인 전원공급 없이 송신전력을 줄이는 것만으로는 동작수명을 유의미하게 연장하기 어렵다는 한계를 가진다. 따라서 배터리 전원공급 제약을 가진 UAV 기반 JCR 시스템의 동작수명을 연장하기 위해서는 추가적인 전력 공급을 위한 에너지 하베스팅 기술을 고려할 필요가 있다.

본 논문에서는 무선 정보 및 전력 동시전송 (simultaneous wireless information and power transfer, SWIPT) 기술이 적용된 UAV 기반 JCR 시스템의 동작수명을 최대화하기 위한 강화학습 기반 UAV 위치제어 기법을 제안한다. 제안 기법에서는 UAV 가 에너지 하베스팅을 통해 추가적인 전원공급을 받으며, 목표 데이터 전송률을 만족하면서 동작수명을 최대화할 수 있도록 강화학습 기반으로 UAV 위치를 제어한다.

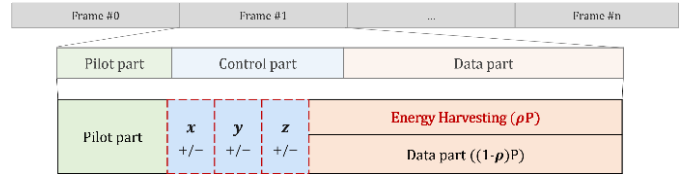


그림 1. 시스템 프레임 구조.

II. 강화학습 기반 위치제어 기법

제안 기법에서 UAV 에서 수신신호는 에너지 하베스팅을 위한 신호와 AF (amplify-and-forward) 중계 전송을 위한 신호로 전력 분할되어 에너지 저장과 정보 전송을 동시에 수행하며, 시스템 프레임 구조는 그림 1 과 같다. Pilot part 는 UAV 의 위치 파악을 위한 레이더 센싱 신호 부분, control part 는 UAV 의 3 차원 공간 상의 좌표 값을 제어하는 신호 부분, data part 는 실제 데이터를 전송하는 부분이다. 이때, UAV 는 ρ ($0 < \rho < 1$) 만큼의 비율로 수신 전력을 변환하여 에너지로 저장한다.

본 논문에서는 UAV 와 기지국 및 사용자의 무선채널 환경에 대해서 LoS 확률과 거리 (d)에 따른 감쇄 효과를 고려한 A2G path loss model 을 (1)과 같이 사용한다 [5].

$$PL_{dB} = 20 \log_{10} \left(\frac{4\pi d f_0}{c} \right) + P_{LoS} \cdot \eta_{LoS} + P_{NLoS} \cdot \eta_{NLoS} [dB]. \quad (1)$$

여기서 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 LoS 와 non-LoS 인 경우의 추가적인 감쇄 손실이며, α, β 는 환경에 따라 달라지는 상수이다.

에너지 하베스팅과 AF 중계를 위해 UAV 가 수신한 신호는 각각 (2), (3)과 같다.

$$y_{UAV}^{EH} = \sqrt{\rho}(\sqrt{P_{BS}}h_{bu}x_{BS} + n), \quad (2)$$

$$y_{UAV}^{AF} = \sqrt{1-\rho}(\sqrt{P_{BS}}h_{bu}x_{BS} + n). \quad (3)$$

(2), (3)으로부터 UAV 에서 수신전력과 수확된 에너지, 데이터 전송에 사용되는 에너지는 다음과 같이 표현된다.

$$\begin{aligned} P_{UAV} &= P_{BS}|h_{bu}|^2, \\ E_{eh} &= T_d\eta\rho P_{UAV}, \\ E_{AF} &= T_d(1-\rho)P_{UAV}. \end{aligned}$$

여기서 P_{BS} 는 기지국의 송신전력, h_{bu} 는 기지국과 UAV 사이의 채널이득, T_d 는 데이터 전송시간, η 는 에너지 변환 효율이다.

본 논문에서는 목표 데이터 전송률 R_{target} 을 만족하면서, 수확된 에너지 E_{eh} 를 최대화하는 것을 목표로 강화학습을 수행한다. 제안하는 강화학습 프레임워크의 agent는 UAV, state는 UAV의 3차원 공간상 좌표, action은 UAV의 위치이동 값으로 x, y, z 축의 +/- 변화량의 조합으로 총 8가지이다. 한 episode가 실행될 때 마다 UAV는 잔여 배터리 에너지 (E_{res})가 0이 될 때까지 R_{target} 을 만족하면서 에너지를 수확하여 동작수명 (*lifetime*)이 최대가 되는 위치로 이동하는 과정을 반복한다.

III. 성능 검증

시뮬레이션을 위한 파라미터는 표 1과 같다 [5].

표 1. 시뮬레이션 파라미터.

Parameters	Values
f_c	2.8 [GHz]
P_N	-90 [dBW]
P_{BS}	30 [dBW]
B	20 [MHz]
η_{LoS}, η_{NLoS}	0.1, 21
α, β	4.8810, 0.4290
η	0.5
$E_{UAV-init}$	5 [J]
R_{target}	250 [Mbps]
τ_{slot}	1 [ms]

그림 2는 episode 진행에 따른 UAV의 *lifetime*의 이동평균 값을 나타낸 것이다. UAV의 초기 에너지는 5[J], R_{target} 은 250Mbps로 설정하였다. 그림 2에서 보는 바와 같이, 에너지 수확 비율 ρ 의 값이 클수록 *lifetime*이 더 큰 값으로 수렴하는 것을 확인할 수 있다.

그림 3은 $\rho = 0.5$ 일 때, UAV의 이동 궤적을 나타낸 것이다. Episode가 진행될수록 *lifetime*이 연장되는 최적 위치로 UAV가 지속적으로 이동함을 확인할 수 있다.

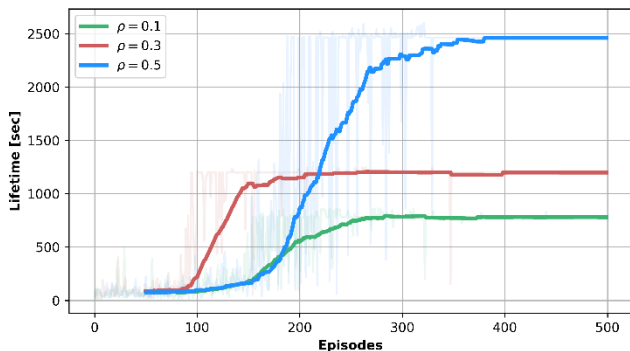


그림 2. Episode에 따른 *lifetime* ($\rho = 0.1, 0.3, 0.5$).

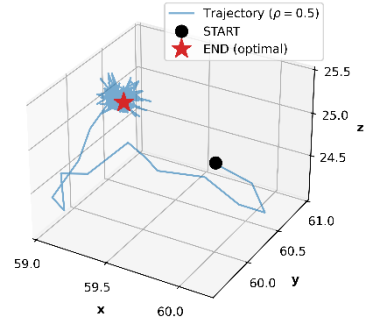


그림 3. UAV 중계기의 이동 궤적 ($\rho = 0.5$).

IV. 결론

본 논문에서는 SWIPT 기술이 적용된 UAV 기반 JCR 시스템의 동작수명을 최대화하기 위한 강화학습 기반의 위치제어 기법을 제안하였다. 제안 기법에서 UAV 중계기는 SWIPT를 통해 추가적인 전원공급을 받으며, 강화학습을 통해 목표 데이터 전송률을 만족하면서 *lifetime*이 최대화되는 위치로 이동하는 것을 확인하였다. 또한, SWIPT 에너지 수확 비율 ρ 값이 클수록 *lifetime*이 더 큰 값으로 수렴하였다. 목표 데이터 전송률을 충분히 만족시킬 수 있을 때 에너지 수확 비율을 키우는 것이 UAV 기반 JCR 시스템의 동작수명 연장에 도움이 된다고 볼 수 있다. 향후, 본 연구 결과를 바탕으로 UAV 위치와 SWIPT 에너지 수확 비율을 동시에 최적화하는 강화학습 기법에 대한 연구로 확장할 계획이다.

ACKNOWLEDGMENT

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

참고 문헌

- [1] S. Mazahir, S. Ahmed, and M.-S. Alouini, "A survey on joint communication-radar systems," *Frontiers in Communications and Networks*, vol. 1, p. 619483, Feb. 2021.
- [2] S. Zhang, Y. Zeng, and R. Zhang, "Cellular-Enabled UAV Communication: A Connectivity-Constrained Trajectory Optimization Perspective," *IEEE Transactions on Communications*, vol. 67, no. 3, pp. 2580-2604, Mar. 2019.
- [3] S. Y. Woo, S. M. Kim, and J. Kim, "Optimization of UAV based Joint Communication and Radar System using Reinforcement Learning," in Proc. *KICS Winter Conf.*, pp.1446-1447, Feb. 2023.
- [4] S. Y. Woo, S. M. Kim, and J. Kim, "Operation Life Optimization Method of UAV Based Joint Communication and Radar System," in Proc. *KICS Summer Conf.*, pp.28-29, Jun. 2023.
- [5] 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.