

# 적 UAV 위치 추정과 계층적 심층 강화학습을 활용한 다중 UAV 환경에서 빔 기반 최적 재밍 기법

조아림, 이영석, 정방철, 이호원

아주대학교

{jal9160, youngseeklee, bcjung, howon}@ajou.ac.kr

## Beam-based Optimal Jamming Technique in Multiple UAV Environments Using Malicious UAV Position Estimation and Hierarchical Deep Reinforcement Learning

Ah Rim Cho, Young-Seok Lee, Bang Chul Jung, Howon Lee

Ajou Univ.

요약

재밍을 수행하는 방법 중 지상 재머를 활용하는 방법은 비용이 저렴하고 효과적일 수 있지만, 이러한 시스템은 지상의 고정된 위치에 존재하므로 대상이 재밍 범위 내에 있을 때만 효과적인 한계가 존재한다. 따라서, 본 논문에서는 3차원 전장 환경에서 위치를 알 수 없는 적 unmanned aerial vehicle(UAV) 위치를 추정된 뒤, 해당 방향으로 이동하여 재밍을 수행함으로써 상대적으로 적은 전력을 사용하여 적 UAV를 무력화하는 것을 목표로 한다. 이에 계층적 심층 강화학습을 활용하여 재머 UAV의 전송 전력 및 빔 폭, 이동 제어를 통해 재밍을 효과적으로 수행하고자 한다.

### I. 서론

이동통신 기술의 발전으로 무선 기기가 빠르게 확산되고, 인터넷 기반 모바일 애플리케이션이 폭발적으로 증가하면서, 무선 서비스는 일상생활과 사회 통신 인프라의 핵심 요소로 자리 잡았다. 이처럼 무선 서비스의 존도가 높아짐에 따라 기밀성·무결성·가용성 손상을 유발할 수 있는 보안 위협이 심각한 문제로 대두되고 있다 [1]. 재밍은 적의 통신을 교란하는 전자공격의 주요 수단으로, 무선 통신 시스템에서 적 사용자 혹은 방해 신호원이 통신 환경을 교란하여 정상적인 신호 송수신을 방해하는 행위를 의미한다 [2],[3]. 지상 재머를 활용하는 방법은 비용 효율적이지만, 지상의 고정된 위치에 위치하므로 대상이 재밍 범위 내에 있을 때만 효과적이라는 단점이 있다. 반면, unmanned aerial vehicle(UAV)를 활용하면 목표물에 접근하여 상대적으로 적은 전력으로도 효과적인 재밍이 가능하나, 이동성을 갖는 적 UAV를 대상으로 재밍을 수행하는 시나리오에서는 적 UAV의 위치를 정확히 파악하기 어렵기 때문에 재밍 효과에 한계가 존재한다 [4]. 그럼에도 불구하고 UAV는 3차원 전장 환경에서 높은 기동성과 유연성을 제공하므로, 기존의 고정 위치 기반 재밍보다 더욱 효율적이고 유연한 운용이 가능하다. 그러나, UAV는 배터리 제약 문제가 있어 높은 전력을 지속적으로 사용하기엔 한계가 있다 [5]. 또한, 실시간으로 적 UAV의 위치를 추정할 수 있는 입사 방향(Direction of Arrival, DoA) 추정 기술은 통신·감지 시나리오 모두에서 매우 유용한 정보를 제공한다 [6]. 따라서, 본 논문에서는 적 UAV 위치 추정 및 계층적 심층 강화학습을 활용한 다중 UAV 환경에서 빔 기반 최적 재밍 기법을 제안한다.

### II. 적 UAV 위치 추정과 계층적 심층 강화학습을 활용한 빔 기반 최적 재밍 기법

본 논문에서는 지상의 적 기지국과 통신하는 적 UAV가 존재하고, 재머 UAV는 적 UAV를 대상으로 air-to-air(A2A) 재밍을 수행하는 시나리오를 고려한다. UAV의 제한된 배터리 문제를 고려하면서 재밍 성능을 극대화하기 위해 재머 UAV의 이동 제어와 빔 폭 및 전송 전력 최적화를 수행하는 것을 목표로 한다. 이에 재머 UAV는 매 타임 스텝마다 적 UAV의 위치를 추정된 후, 해당 방향으로 이동하여 재밍을 수행한다. 위치 추정을

위해 소스 신호  $s$ 에 대한 조향 벡터  $a^{3d}$ 를 이용하여  $N$ 개의 모든 안테나에서의 수신 신호 벡터  $X$ 를 생성하며, 수신 신호에 대해 자기 상관 행렬을 계산한 뒤 고유값 분해를 수행한다.  $X$ 의 상관 행렬  $R_{XX} = UAU^H$ ,  $U = [U_s, U_n]$ 을 통해 잡음 공간  $U_n$ 을 추출할 수 있으며,  $P_{MUSIC}(\hat{\theta}, \hat{\phi})$ 의 피크 값을 갖는 방향을 추정 방향으로 결정한다 [7].

$$P_{MUSIC}(\hat{\theta}, \hat{\phi}) = \frac{a^H(\theta, \phi) a(\theta, \phi)}{a^H(\theta, \phi) U_n U_n^H a(\theta, \phi)}. \quad (1)$$

해당 추정 방향  $\hat{\theta}_{MU}^t, \hat{\phi}_{MU}^t$ 으로 이동한 뒤, 적 UAV와 재머 UAV 간의 거리를 RSSI 기반으로 추정하여 두 UAV 간 거리가 멀어진 경우 추정 방향으로 한 번 더 이동한 뒤 재밍을 수행하며, 가까워진 경우 즉시 재밍을 수행한다.

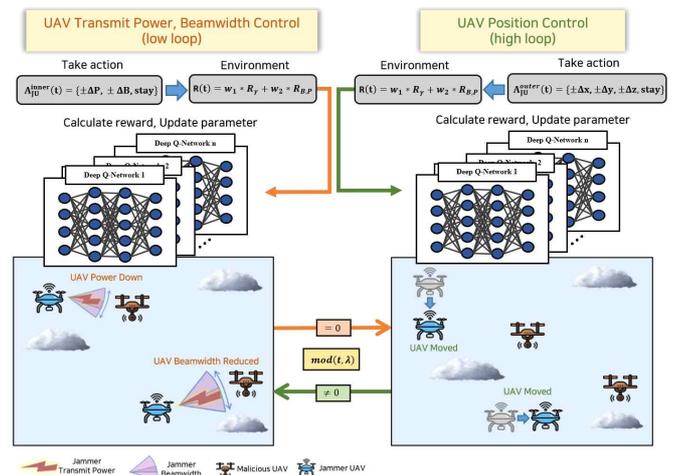


그림 1. Proposed framework

그림 1과 같이 재머 UAV를 에이전트로 하여 재밍 성능 최대화를 위한 이동 제어 및 재밍 기법 최적화 문제를 계층적 심층 강화학습으로 풀고자 Markov decision problem(MDP)로 정의한다.

- State: 두 UAV 간의 거리 비교를 위한 이전 타임스텝과 현재 타임 스텝의 RSSI, 적 UAV의 추정 방향  $\hat{\theta}_{MU}^t, \hat{\phi}_{MU}^t$ , 제머 UAV의 3차원 위치, 전송 전력과 빔 폭으로 구성된다.

$$S_{JU} = \{RSSI_{JU,MU}^{t-1}, RSSI_{JU,MU}^t, \hat{\theta}_{MU}^t, \hat{\phi}_{MU}^t, x_{JU}^t, y_{JU}^t, z_{JU}^t, P_{JU}^t, B_{JU}^t\}.$$

- Action: 제머 UAV는 계층적 학습 주기  $\lambda$ 마다 상위 계층 루프를 통해 3차원 공간 내에서 이동 제어를 수행하며,  $\lambda$ 를 제외한 모든 타임스텝에서 하위 계층 루프를 통해 제머의 빔 폭과 전송 전력 제어를 수행한다.  $A_{JU}^{High} = \{\pm x, \pm y, \pm z, stay\}$ ,  $A_{JU}^{Low} = \{\pm \Delta B_{JU}, \pm \Delta P_{JU}, stay\}$ .

- Reward: 보상은 크게 두가지 요소로 구성되며, 최소화를 목적으로 하는 적 UAV의 데이터 달성률에 대한 보상과 제머 UAV의 빔 폭 및 전송 전력과 제밍 성능 간의 trade-off 관계를 고려한 보상으로 구성된다.  $\Gamma_M$ 은 적 기지국과 적 UAV 간의 signal-to-jamming-plus-noise ratio(SJNR)을 의미하며, 여기서  $G(B_{JU}^t)$ 는 제머 UAV의 빔 폭에 따른 안테나 이득을 의미한다. 제밍 효과가 클수록  $\Gamma_M$ 이 감소하므로 negative의 보상 형태를 가지며, 각 보상 항목에는 trade-off 관계를 고려하기 위한 가중치  $w_1, w_2$ 를 적용한다.

$$R_{JU} = w_1 \times R_{\Gamma} + w_2 \times R_{B,P}. \quad (2)$$

$$R_{\Gamma} = -BW \times \log_2(1 + \Gamma_M), \quad R_{B,P} = -(P_{JU}^t + B_{JU}^t). \quad (3)$$

$$\Gamma_M = \frac{P_{MB} * PL_{A2G}^{-1}}{P_{JU}^t * G(B_{JU}^t) * PL_{A2A}^{-1} + \sigma^2}. \quad (4)$$

### III. 시뮬레이션 결과 및 결론

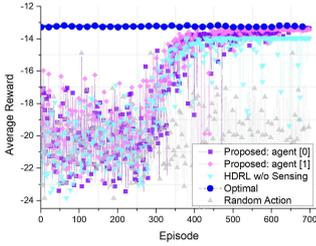


그림 2. Average reward vs. episode.

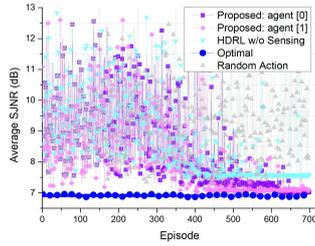


그림 3. Average SJNR vs. episode.

본 논문에서는 3차원 이동성으로 인한 제밍 성능의 불안정성을 해결하기 위해 위치 추정을 수행하며, 이를 바탕으로 동적 환경에서 최적의 제밍 전략을 학습 및 적용하기 위해 계층적 강화학습을 활용하였다. 아군 영역에서 [80, 150][m] 사이 고도에 2대의 적 UAV와 2대의 제머 UAV가 랜덤 배치된 시나리오를 고려한다. 아래의 벤치마크 알고리즘을 고려하여 제안 방안과의 성능을 비교 분석한다.

- Proposed: 적 UAV에 대한 위치 추정을 수행하며, 계층적 심층 강화학습을 통해 전송 전력 및 빔 폭, 이동 제어를 수행하는 방안이다.
- Random Action: 매번 랜덤한 행동을 수행하는 방안이다.
- Hierarchical deep reinforcement learning(HDRL) w/o Sensing: 적 UAV에 대한 위치 추정을 수행하지 않고, 계층적 심층 강화학습을 수행하는 방안으로 위치 추정을 수행하지 않기 때문에 제머 UAV가 적 UAV 근처에서 제밍을 수행하지 못할 가능성이 있다. 이로 인해 제밍 성능과 안정성이 저하될 수 있어 위치 추정을 수행하는 것이 얼마나 효과적인지를 제안 방안과 비교하기 위한 방안이다.
- Optimal: 현재 상태에서 모든 가능한 행동을 탐색하여 최적의 행동을 수행하는 방안이다. 탐욕 정책을 기반으로 가장 높은 성능을 보장할 수 있지만, 계산량이 매우 크기 때문에 실시간 적용이 어려울 수 있다. 따라서, 제안 방안과 이론적 최적 성능을 비교하기 위한 기준으로 활용한다.

표 1의 파라미터를 기반으로 시뮬레이션을 진행하였으며, 시뮬레이션을 통해 제안 방안의 성능을 평가한 결과 제안 방안의 경우 그림 2에서 에피소드가 진행됨에 따라 ‘HDRL w/o Sensing’, ‘Random Action’ 대비 높은 값에 수렴함을 보이며, ‘Optimal’에 수렴하는 것을 통해 최적의 방안을 학습한 것을 확인하였다. 또한, 그림 3에서 적 UAV의 통신 품질을 SJNR로 평가하였으며, 벤치마크 대비 가장 낮은 값으로 수렴하였다. 이를 통해 제안 방안이 적 UAV의 통신 품질을 효과적으로 무력화시키는 것을 확인하였다.

Parameter	Value
Network size	500x500x150 [m]
Jammer UAV velocity( $v_{JU}$ )	3 [m/iteration]
Malicious UAV velocity( $v_{MU}$ )	3 [m/iteration]
Thermal noise power( $\sigma^2$ )	-101 [dBm]
Malicious BS location	100,100.0
Malicious BS transmit power( $P_{MB}$ )	30 [W]
Jammer UAV initial beamwidth	60 °
Jammer UAV beamwidth angle interval	10 °
Jammer UAV initial transmit power	5[W]
Jammer UAV transmit power interval	1[W]
Bandwidth(BW)	20 [MHz]
Carrier frequency	2.4 [GHz]
Discount factor	0.99
Learning rate	0.0001

표 1. Simulation parameters

### ACKNOWLEDGMENT

이 논문은 한국연구재단의 지원(RS-2025-00563401, 3차원 공간에서 에너지 효율적 멀티레벨 AI-RAN 구현을 위한 AI-for/and-RAN 핵심 원천기술 연구), 정부(과학기술정보통신부)의 재원으로 정보통신기획평가원의 지원(RS-2024-00300002, 저궤도 위성통신 핵심 기술 기반 큐브 위성 개발), 2022년도 정부(과학기술정보통신부)의 재원으로 정보통신기획평가원의 지원(No.2022-0-00704, 초고속 이동체 지원을 위한 3D-NET 핵심 기술 개발)을 받아 수행된 연구임.

### 참고 문헌

- [1] H. Pirayesh and H. Zeng, "Jamming Attacks and Anti-Jamming Strategies in Wireless Networks: A Comprehensive Survey," in *IEEE Communications Surveys & Tutorials*, vol. 24, no. 2, pp. 767-809, Secondquarter 2022.
- [2] P. Sharma, K. K. Sama and N. E. Mastorakis, "Artificial Intelligence Aided Electronic Warfare Systems- Recent Trends and Evolving Applications," in *IEEE Access*, vol. 8, pp. 224761-224780, 2020.
- [3] Y. Fan, B. Jiu, W. Pu, Z. Li, K. Li and H. Liu, "Sensing Jamming Strategy From Limited Observations: An Imitation Learning Perspective," in *IEEE Transactions on Signal Processing*, vol. 72, pp. 4098-4114, 2024.
- [4] S. Papaioannou, P. Kolios and G. Ellinas, "Distributed Estimation and Control for Jamming an Aerial Target With Multiple Agents," in *IEEE Transactions on Mobile Computing*, vol. 22, no. 12, pp. 7203-7217, Dec. 2023.
- [5] D. Guo, L. Tang, X. Zhang and Y. -C. Liang, "Joint Optimization of Trajectory and Jamming Power for Multiple UAV-Aided Proactive Eavesdropping," in *IEEE Transactions on Mobile Computing*, vol. 23, no. 5, pp. 5770-5785, May 2024.
- [6] J. Cong, X. Wang, C. Yan, L. T. Yang, M. Dong and K. Ota, "CRB Weighted Source Localization Method Based on Deep Neural Networks in Multi-UAV Network," in *IEEE Internet of Things Journal*, vol. 10, no. 7, pp. 5747-5759, 1 April, 2023.
- [7] B. Vikas and D. Vakula, "Performance comparison of MUSIC and ESPRIT algorithms in presence of coherent signals for DoA estimation," 2017 International conference of Electronics, Communication and Aerospace Technology (ICECA), Coimbatore, India, 2017, pp. 403-405.