

# 우주-항공-지상 통합 네트워크에서 데이터 전송 성공률 향상과 에너지 효율을 위한 FL-MAPPO 기반 다중 UAV 운용 기법

이충녕, 채승호\*

한국공학대학교

lc9902130509@tukorea.ac.kr, \*shchae@tukorea.ac.kr

## FL-MAPPO-Based Multi-UAV Operation Scheme for Enhancing Transmission Success Rate and Energy Efficiency in Space-Air-Ground Integrated Networks

Chungnyeong Lee, Seong Ho Chae\*

Tech University of Korea

요약

본 논문에서는 6G 핵심 기술인 우주-항공-지상 통합 네트워크(SAGIN) 환경에서, 유한한 배터리를 갖는 다중 무인항공기(UAV)를 활용하여 전체 시스템 데이터 전송 성공률과 UAV 에너지 효율 동시 최적화를 하는 FL-MAPPO 기반 UAV 운용 기법을 제안한다. 제안 기법은 연합 학습(Federated Learning)과 다중 에이전트 근접 정책 최적화(MAPPO)를 결합하여 UAV 간 협력 정책을 학습하고, 이를 기반으로 UAV는 지역 정보를 활용해 이동 경로와 서비스 노드 수를 동적으로 조정한다. 시뮬레이션 결과, 제안된 기법은 기존 스킴 대비 데이터 전송 성공률을 약 9.4% 향상시키고, 임무 종료 후 UAV 잔여 배터리를 4~6% 수준으로 유지함으로써 에너지 효율과 데이터 전송 성공률의 균형을 효과적으로 달성하였다.

### I. 서론

6세대 이동통신(6G)은 지상 네트워크의 한계를 극복하고 유비쿼터스 연결성 제공을 위해 우주-항공-지상 통합 네트워크(Space-Air-Ground Integrated Network; SAGIN)를 핵심 구성요소로 고려하고 있다[1]. SAGIN에서 위성은 광역 커버리지를 제공하지만, 넓은 빔 커버리지는 주파수 재사용으로 인한 심각한 사용자 간 간섭을 야기하며, 긴 신호 경로는 불안정한 채널 품질의 원인이 되어 다수의 이용자에게 안정적인 서비스를 제공하기 어렵다. 이를 보완하고자, 최근 무인항공기(UAV)를 이동형 릴레이로 활용하여 더 작은 셀을 구성하고 협력 통신을 통해 다중 접속 환경에서의 통신 성능을 향상시키는 연구가 활발히 진행되고 있다[2]. 그러나, UAV 운용에는 제한된 배터리 용량의 물리적 제약이 존재하므로 통신 성능과 에너지 효율의 균형이 중요하다.

본 연구에서는 이러한 문제를 효과적으로 해결하기 위해 연합 학습(FL, Federated Learning)과 다중 에이전트 심층 강화학습(MADRL, Multi-Agent Deep Reinforcement Learning)을 결합한 FL-MAPPO 기반 UAV 운용 기법을 제안한다.

### II. 시스템 모델

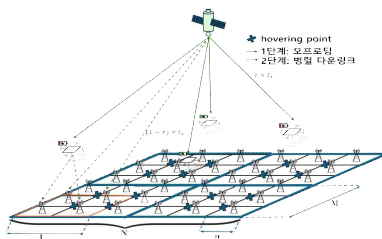


그림 1. 시스템 모델

본 논문에서는 그림 1과 같이 1개의 위성과  $K$ 개의 UAV가 지상에 각 셀의 가로, 세로 길이가  $n$ 인  $N \times N$  개의 인접 셀을 총  $T$  시간 동안 서비스하는 환경을 고려한다. 위성과 UAV는 서로 다른 주파수 대역을 사용함에 따라 상호간 간섭은 존재하지 않음을 가정한다. 이때, 전체 시간  $T$

는 시간 간격  $t_r$ 의 시간 슬롯(time slot)으로 분할되며, 각 시간 슬롯은 준정적(Quasi-static) 상태를 가정한다. 각 격자 셀의 중앙에는 지상 노드가 배치되어 있으며, 위성은 전체 지역에 다운링크 서비스를 제공하고,  $K$ 개의 UAV는 이를 보조하기 위해 각 담당 서비스 영역을 분할하여 이동형 복호후 전달(DF, Decode-and-Forward) 릴레이 역할을 수행한다.

각 UAV는 자신의 담당 서비스 지역인  $M \times M$ 개의 인접한 셀 내에서, 미리 정의된 유한한 수의 호버링 후보 위치로만 이동할 수 있으며, 특정 위치에 호버링시 최대  $L \times L$ 개의 인접 셀들을 커버함을 가정한다. 여기서,  $L \leq M (= N/K)$ 을 만족한다. 각 지상 노드는  $R$ 개의 트래픽 요구량 레벨 중 하나에 해당하는 트래픽을 요구하며, 요구량은 시간 슬롯마다 독립적이고 무작위로 변화함을 가정한다. 각 시간 슬롯마다 UAV는 담당 서비스 지역 내 지상 노드들의 데이터 요구량과 채널 상태 정보를 얻을 수 있으며, 이를 기반으로 호버링 위치와 서비스 노드의 수  $l$  ( $l \leq L \times L$ )을 결정한다. 서비스 노드의 수  $l$ 이 결정되면, UAV는  $L \times L$ 개의 커버하는 셀 중에서 트래픽 요구량이 높은 순서로 우선적으로 서비스함을 가정한다. 위성은 모든 지상 노드의 트래픽 요구량과 각 UAV가 서비스하기로 결정한 지상 노드들에 대한 정보를 가지고 있다. 결정된 호버링 위치로 UAV들의 이동이 완료되면 다운링크 통신이 시작되며, 이는 TDD(Time Division Duplexing) 기반의 2단계 구조로 동작한다.

• **1단계(오프로딩):** 주어진  $\tau \times t_r$  시간 동안 위성은 전체 대역폭을 균등 분할하여  $K$ 개 UAV에 UAV를 통해 지상 노드로 전송할 데이터를 전송한다. 여기서,  $0 < \tau \leq 1$ 의 임의의 상수 비율이다. 각 UAV는 자신이 서비스하기로 결정한 지상 노드들 중 요구량이 높은 노드의 요청 데이터부터, 위성으로부터 우선 수신 후 복호(Decode)하며, 이후 병렬 다운링크 단계에서 이를 재전송한다.

• **2단계(병렬 다운링크):** 다음  $(1 - \tau) \times t_r$  시간 동안 UAV가 커버리지 내 선택된 지상 노드들에 위성으로부터 수신 받은 트래픽을 전송하고, 위성은 나머지 노드들의 요구 트래픽을 직접 트래픽을 전송한다. 이때, 위성과 UAV는 각기 서비스하는 노드 수만큼 자신의 대역폭을 직교 균등 분할하여 트래픽을 전송한다.

위성-지상노드와 위성-UAV 채널은 쉐도우드 라이시안 페이딩(Shadowed Rician Fading)[1]으로, UAV-지상 노드 채널은 레일레이 페이딩(Rayleigh Fading)[2]으로 모델링 된다. 또한, 모든 송수신 링크는 자유 공간 경로 손실(Free Space Path Loss)을 겪음을 가정한다.

각 UAV는 유한한 초기 배터리 용량  $B_{\max}$ 을 가지며, 매 시간 슬롯마다 호버링( $E_h$ ), 이동( $E_m(t)$ ), 데이터 송신( $E_{\text{down}}(t)$ ) 행위에 따라 에너지를 소모한다. 이때  $E_m(t) \gg E_h, E_{\text{down}}(t)$ 이며, 이때 이동 에너지 소모는 이동하는 거리에 비례하여 증가하고, 데이터 송신 에너지 소모는 서비스하는 지상 노드 수에 따라 선형적으로 증가한다.

본 논문의 목표는 UAV의 에너지 효율을 향상시키면서, 전체 시스템의 데이터 전송 성공률을 최대화하는 것이다. 이를 위해, 해당 문제를 다음의 POMDP (Partially Observable Markov Decision Process)로 모델링한다.

- **관측**: 각 UAV는 담당 구역 내 지상 노드의 트래픽 요구량, 채널 상태, 정규화된 배터리 잔량을 관측한다.
- **행동**: UAV는 호버링 위치와 커버리지 내 서비스할 지상 노드의 수를 동시에 결정한다.
- **보상**: 보상 함수는 데이터 전송 성공률과 에너지 효율을 동시에 최대화하도록 다음과 같이 정의한다.

$$R_t = \frac{S_t}{(N \times N)} - \beta \left( \frac{E_t}{E_{\max}} \right).$$

여기서,  $S_t$ 는 요구한 데이터를 성공적으로 수신한 지상 노드의 수,  $E_t$ 는 소모된 에너지,  $E_{\max}$ 는 시간 슬롯 내에 소모 가능한 최대 에너지,  $\beta$ 는 에너지 페널티 가중치이다.

### III. FL MAPPO를 활용한 UAV 운용 기법

본 연구는 다중 UAV의 협력적 정책 학습을 위해 MAPPO(Multi-Agent Proximal Policy Optimization) 알고리즘을 적용하였다[1]. 각 UAV는 로컬 관측에 기반해 행동을 결정하고(Decentralized Execution), 학습 단계에서는 모든 에이전트의 정보를 활용하는 중앙집중형 Critic을 통해 안정적인 학습을 수행한다(Centralized Training).

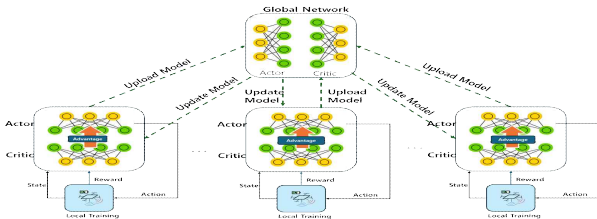


그림 2. 제안하는 FL-MAPPO 구조 및 학습 과정

또한, UAV 간 데이터 공유 시 발생하는 통신 오버헤드와 프라이버시 문제를 완화하기 위해 FL을 결합하였다. 그림 2와 같이 각 UAV는 로컬 환경에서 Actor-Critic 모델을 독립적으로 학습하고, 학습된 모델 가중치만 중앙 서버로 전송한다. 서버는 이를 FedAvg 알고리즘[3]으로 병합 및 갱신한 뒤 UAV들에게 재배포하며, 반복 과정을 통해 글로벌 정책으로 수렴한다.

### IV. 성능 검증

제안 기법의 성능 검증을 위해, 시간 슬롯 간격  $t_r = 5s$ , 총 30개의 시간 슬롯 동안 각 셀의 한 변이 100m인  $6 \times 6$ 개의 인접 셀을 고도 550km 위성과 고도 300m의  $K=4$ 개의 UAV가 각기  $3 \times 3$  담당 구역 내  $2 \times 2$ 의 커버리지를 가지고 서비스하는 시나리오를 고려한다. 시스템 파라미터는 다음과 같다:  $\tau=0.3$ , 위성의 송신 전력 300W, 대역폭 300MHz, 전송 주파수 28GHz, UAV의 각 링크별 전송 전력 1W, 대역폭 300MHz, 전송 주

파수 2GHz. 에너지 페널티 가중치  $\beta=0.27$ . 지상 노드는  $R=3$ 개 레벨의 트래픽 요구량이 있음을 가정했다(20Mbps, 40Mbps, 60Mbps).

효과적인 성능 분석을 위해 다음과 같은 비교 스킴을 설정하였다.

- 1) No-UAV: UAV 없이 위성만을 사용하여 통신한 경우
- 2) Random UAV: UAV가 위치와 서비스 개수를 무작위로 결정
- 3) Heuristic (Max-Demand): UAV가 항상 트래픽 수요가 가장 높은 곳으로 이동하여 최대 지상 노드를 서비스
- 4, 5) Fixed-Relay-1 / 4: 제안 모델의 위치 결정 정책에 따르되, 서비스 노드 개수를 1개 또는 4개로 고정
- 6) FL-MAPPO: 제안하는 기법

	No-UAV	Random UAV		Heuristic UAV		Fixed-Relay-1		Fixed-Relay-4		FL-MAPPO	
평균 데이터 전송 성공률	49.4%	57.7%		55.5%		63.6%		64.0%		70.1%	
UAV 당 평균 서비스 노드 수	2.5		4		1		4		2.3		
최종 배터리 잔량/방전된 시간슬롯	UAV_0	0%	16	0%	14	36%	-	0%	21	4%	-
	UAV_1					36%			21	5%	
	UAV_2					36%			19	5%	
	UAV_3					37%			20	6%	

표 1. 기법에 따른 성능표

표 1은 제안한 FL-MAPPO 기법과 기존 스킴 간의 성능 비교 결과를 나타낸다. No-UAV 환경은 49.40%로 가장 낮은 데이터 전송 성공률을 보였으며, 이는 UAV를 이동형 릴레이로 활용할 경우 네트워크 전체 데이터 전송 성공률이 향상됨을 확인할 수 있다.

Fixed-Relay-4는 64.04%로 비교 스킴 중 가장 높은 데이터 전송 성공률을 기록했으나, 과도한 에너지 소모로 조기 방전이 발생하였다. 반면, 제안한 FL-MAPPO는 70.05%로 Fixed-Relay-4 대비 약 9.4% 향상된 성공률을 달성하면서도, 모든 UAV의 잔여 배터리를 4~6% 수준으로 유지하였다. 이는 제안 기법이 고정된 서비스 노드 수나 트래픽 패턴에 의존하지 않고, 환경 변화에 따라 UAV의 위치와 서비스 노드 수를 동적으로 결정했기 때문이다. 결과적으로, FL-MAPPO는 데이터 전송 성공률과 에너지 효율 모두에서 기존 스킴을 상회하며, 지속 가능한 UAV 운용 정책을 효과적으로 학습했음을 확인할 수 있다.

### V. 결론

본 논문에서는 FL-MAPPO 기반 다중 UAV 운용 기법을 통해 SAGIN 환경에서 데이터 전송 성공률과 UAV 에너지 효율을 동시에 향상시킬 수 있음을 보였다. 제안 기법은 학습된 협력 정책을 통해 데이터 전송 성공률 약 9.4% 향상과 지속 가능한 임무 수행 능력을 달성하였으며, 이는 복잡한 SAGIN 환경에서의 실용적 적용 가능성을 시사한다.

### ACKNOWLEDGMENT

이 논문은 2022년도 정부(방위사업청)의 재원으로 국방기술진흥연구소의 지원을 받아 수행된 연구임 (KRIT-CT-22-047 우주계층 지능통신망 특화연구실).

### 참 고 문 헌

- [1] C. Lee et al., "Handover minimization scheme using multi-agent deep reinforcement learning in multi-beam low earth orbit satellites", *J. Korean Institute of Communications and Information Sciences (J-KICS)*, vol. 50, no. 8, pp. 1196-1206, 2025.
- [2] L. Cheng et al., "Cooperative transmission for space-air-ground integrated networks: A multi-agent cooperation method," *IEEE Trans. Veh. Technol.*, vol. 74, no. 8, pp. 12879-12894, Aug. 2025.
- [3] F. Tang et al., "Federated learning for intelligent transmission with space-air-ground integrated network toward 6G," *IEEE Network*, vol. 37, no. 2, pp. 198-204, Mar. 2023.