

# UAV-HST 환경에서 강화학습 기반의 UAV 경로 최적화 연구

박유민, 홍충선\*  
경희대학교 컴퓨터공학과

yumin0906@khu.ac.kr, \*cshong@khu.ac.kr

## A Study on the Optimization UAV Trajectory based on Reinforcement Learning in Unmanned Aerial Vehicle-High Speed Train Environments

Yu Min Park, Choong Seon Hong\*  
Kyung Hee University

### 요 약

고속 열차에서 5G 기반의 VR/AR의 콘텐츠를 사용자들이 경험하기 위해서는 새로운 통신 인프라가 필요하다. 따라서 일반적으로 열차의 통신을 지원하는 Rail-side Units(RSUs)를 대신하여 Unmanned Aerial Vehicle(UAV)를 통신 장비로써 활용하는 것을 제안한다. 하지만 UAV 보다 빠른 고속 열차와 통신을 안정적으로 유지하기 위해서는 고도, 속도, 방향 등을 최적으로 결정하여 경로를 이동해야 한다. 이에 본 논문은 고속 열차와 UAV의 안정적인 통신을 위해 UAV의 경로 최적화 기법을 제안한다. The Air-To-Ground Path Loss Model을 토대로 가상 통신 환경을 조성하고 Unity ML-agents 라이브러리를 통해 학습 환경을 구축한다. 본 논문은 UAV 경로 최적화를 위한 방법으로 강화학습 중 Soft-Actor-Critic(SAC)을 사용한다. 결과적으로 본 연구를 통해 고속 열차의 속도에 따른 UAV의 최적 경로를 빠르고 정확하게 도출하여 통신을 안정적으로 유지할 수 있었다.

### 1. 서론

차세대 이동 통신인 5G는 밀리미터파를 이용한 무선 네트워크 기술이다. 5G는 대용량의 데이터를 초고속으로 무선 통신할 수 있게 하여 VR/AR 등 고사양의 콘텐츠를 사용자들이 무선으로 경험할 수 있게 하였다. 하지만 5G는 사용하는 밀리미터 파의 직진성으로 인해 통신 사이의 물체로 인해 손상되지 않도록 근거리로 셀을 구축해야 한다. 따라서 5G의 안정적인 통신이 가능하기 위해서는 많은 5G 통신 인프라를 구축해야 하며 이는 경제적 비효율성과 지리적 한계로 인한 음영지역을 발생시킨다.

고속 열차에서 5G 기반의 VR/AR의 콘텐츠를 사용자들이 경험하기 위해서는 새로운 통신 인프라가 필요하다. 현재 고속 열차의 통신을 담당하고 있는 Rail-side Units(RSUs)는 5G 서비스를 지원하기에 충분하지 않은 숫자이다 [1]. 충분한 수의 RSUs가 갖추어진다고 하여도 고속으로 이동하는 열차에서 잦은 Handover가 발생하게 되고 그에 따른 기기 과부하 및 통신 지연이 예측된다. 따라서 본 연구에서는 고속 열차에서의 5G 통신을 위해 RSUs를 대신하여 UAV-Base Station을 사용하는 것을 제안한다.

Unmanned Aerial Vehicle(UAV) 기반의 통신은 동적이고 예측 불가능한 서비스 수요를 충족시키는

동시에 지상 기반 기지국의 대안으로 비용을 절감할 수 있는 뛰어난 유연성을 제공할 수 있다 [2]. 현재 UAV는 이동형 기지국으로서의 활용, 캐싱 장치로의 활용 등 단순한 레저 기기, 군용 장비에서 벗어나 통신을 지원할 수 있는 장비로써 사용하는 연구가 활발히 진행 중이다. 또한, UAV의 통신 연구에서 가장 크게 문제가 되었던 비행시간, 즉 배터리 소모에 대한 이슈는 수소 배터리 기반의 UAV가 상용화가 이루어지며 점차 해당 문제가 줄어들고 있다. UAV의 유연한 이동과 저렴한 비용 등 많은 장점이 존재하지만, 여전히 효율적인 움직임, 배치, 보안 등에서 해결해야 할 문제들이 존재한다. UAV보다 빠른 고속 열차와 통신을 안정적으로 유지하기 위해서는 고도, 속도, 방향 등을 최적으로 결정하여 경로를 이동하는 것이 중요하다. 따라서 본 연구에서는 고속 열차와 UAV의 안정적인 통신을 위해 UAV의 경로 최적화 문제를 해결하고자 한다.

본 논문의 2 장에서는 UAV 통신에서 경로 최적화 문제를 다룬 관련 연구들을 살펴보고, 3 장에서는 본 논문의 실험 환경 구축을 위한 통신 모델과 강화 학습 요소를 설명한다. 4 장에서는 제안한 시스템의 성능 결과에 대한 평가를 다루며, 마지막 5 장은 논문의 결론 및 향후 연구방향을 제시한다.

## 2. 관련연구

### 2.1 고정된 사용자에게 대한 UAV 경로 최적화 [3]

논문 [3]은 고정된 사용자들에 대해 UAV 의 경로 최적화를 실시하여 평균 통신량을 최대화시켰다. 경로 최적화 기법으로는 Deep Deterministic Policy Gradient(DDPG) 라는 강화학습을 제안하였고 결과적으로 분산된 사용자들의 통신량이 전체적으로 높아질 수 있었다. 하지만 논문 [3]에서는 위치가 고정된 사용자들에 대해 경로 최적화를 실시했기 때문에 실제 움직이는 차량, 열차와의 통신에서는 한계가 존재한다. 이에 본 논문은 고속으로 움직이는 열차와 UAV 통신 상황에서 UAV 경로 최적화 방안을 제시한다.

### 2.2 움직이는 사용자에게 대한 UAV 경로 최적화 [4]

논문 [4]는 도로 위의 차량과의 UAV 통신에서 UAV 경로 최적화와 자원 할당을 시행하였다. 경로 최적화와 자원 할당을 동시에 최적화하기 위해 수치적인 방법인 Successive Convex Optimization(SCO)라는 기법을 사용하였다. 하지만 논문 [4]에서의 UAV 는 고도가 고정되어 있어 통신 범위가 달라지는 것에 대해 고려하지 않았다. 본 논문은 UAV 의 고도 또한 경로 선정의 변수로 고려하여 최적의 통신 범위를 가질 수 있도록 하였다.

## 3. 제안사항

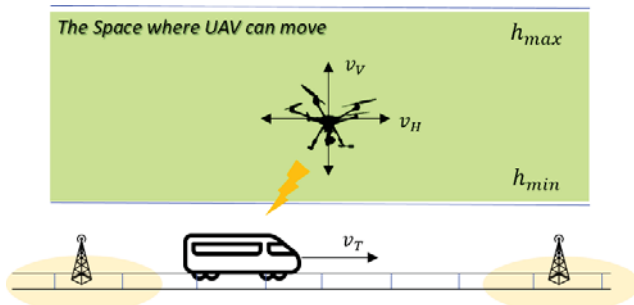


그림 1. 시스템 모델

본 논문은 그림 1 과 같이 RSUs 가 지원하지 못하는 통신 범위에 대해 UAV 를 이용하여 통신을 지원할 수 있도록 한다. 고속 열차는  $v_T$  의 속도를 가지며 UAV 는  $v_V$  의 수직 속도와  $v_H$  의 수평 속도를 가지게 된다. 또한 UAV 는 고도의 제한이 있으며  $h_{min}$  과  $h_{max}$  사이의 고도에서만 이동이 가능하다. 본 논문은 위와 같은 환경에서 UAV 가 열차의 속도에 따라 최적의  $v_H$ ,  $v_V$  를 결정하여 최적의 경로 이동을 할 수 있는 방법을 제안한다.

$$R(h_t) = h_t \cdot \tan \frac{\omega}{2} \quad (1)$$

식 (1)은 높이  $h_t$ 와 통신이 가능한 각도  $\omega$ 에 따른 통신

범위의 반경을 도출하기 위한 식이다 [5]. 통신 대상인 열차의 속도가 빠르다면 UAV 는 고도를 높여 통신 범위를 늘리는 것이 지속적인 통신에 유리할 것이다. 그렇지 않을 때는 낮은 고도에서 높은 통신량을 보장하는 것이 유리하다. 이에 본 논문은 강화학습을 적용하여 열차의 속도에 따른 UAV 의 최적 경로를 빠르고 정확하게 찾을 수 있도록 제안한다. 가상 환경 구현을 위해 3.1 절의 통신 모델과 3.2 절의 강화학습 환경을 사용한다.

### 3.1 통신모델

**Path Loss Model:** UAV 와 열차 간의 통신 과정에서 거리( $d$ )에 따른 신호 감쇄 효과로 인한 SNR(신호 대 잡음 비)을 계산하기 위해 Air-To-Ground Pass Loss 모델을 사용했다 [6]. 식 (2)은 거리( $d$ ), 빛의 속도( $c$ )와 UAV 의 통신 주파수( $f_0$ )을 통해 UAV 와 열차 간의 통신 손실을 구하는 식이다.  $P(LoS)$ 는 통신 간에 장애물이 없는 line-of-sight 가 될 확률이고  $P(NLoS)$ 는 통신 간에 장애물이 있을 확률이다. 따라서  $P(LoS) = 1 - P(NLoS)$ 라는 식을 얻을 수 있다. 그리고  $\eta_{LoS}$ 와  $\eta_{NLoS}$ 는 각각,  $LoS$ 와  $NLoS$ 에서 추가적인 감쇄 수치이다. 식 (6)의  $a$ 와  $b$ 는 설정 환경에 따라 변하는 상수이다.

$$PL(dB) = 20 \log_{10} \left( \frac{4\pi d f_0}{c} \right) + P(LoS)\eta_{LoS} + P(NLoS)\eta_{NLoS} \quad (2)$$

$$P(LoS) = \frac{1}{1 + a \exp(-b(\theta - a))} \quad (3)$$

이렇게 구한 통신 손실( $PL$ )을 식 (4)와 식 (5)에 차례로 대입하여 통신 세기( $P_r$ )와 SNR을 각각 구할 수 있다.

$$P_r(d) = P_t + G_t - PL + G_r - L \quad (4)$$

$$SNR = \frac{\text{Signal Power}}{\text{Noise Power}} = \frac{P_r(d)}{N_0} \quad (5)$$

**Data Rate Model:** Shannon-Hartley 이론에 기반하여 UAV 와 열차 간의 통신량에 대한 식 (6)을 구할 수 있다. 따라서 대역폭( $B$ )에 대해 앞서 구한 SNR을 식 (6)에 대입하여 통신량( $D$ )을 도출할 수 있다.

$$D_t = B \log_2(1 + SNR) \quad (6)$$

### 3.2 강화학습 환경

강화학습을 위해 Unity 의 ML-Agents 를 사용하였다 [7]. 해당 라이브러리는 다양한 강화학습 환경을 구축하는데 유용한 라이브러리로서 본 논문의 학습 환경 구축에 사용되었다. 강화학습은 Tensorflow 을 사용하였으며 Soft-Actor-Critic 이라는 방법을 사용하여 학습하였다. 환경은 그림 1 과 같이 하나의 고속열차와 한대의 UAV 와의 통신에서 다루었다. 열차는 직선의 선로를 이동한다고 가정하였고 최저 속도(70 m/s)에서 최대 속도(90 m/s) 사이의 속도로 이동한다. UAV 는 허용 고도(100m ~ 200m)에서 최대 25m/s의 수직 속도  $v_V$ 와 최대 60m/s의 수평 속도  $v_H$ 를 가진다. 하나의 에피소드

T는 일정한 시간 간격으로 나누어 진다.

$$T = \{t_0, t_1, t_2, \dots, t_i, \dots\} \quad (7)$$

**State:** 시간  $t_i$ 에서 상태  $s_i$ 는 식 8 과 같이 정의한다.

$$s_i = \{p_u, p_t, v_T\} \quad (8)$$

$p_u$ 와  $p_t$ 는 각각 UAV 의 위치와 열차의 위치이며  $v_T$ 는 열차의 속도이다.

**Action:** 시간  $t_i$ 에서 행동  $a_i$ 는 식 9 과 같이 정의한다.

$$a_i = \{v_v, v_H\} \quad (9)$$

$v_v$ 와  $v_H$ 는 방향의 가진 속도으로써 UAV 의 방향과 속력을 결정할 수 있다. 학습 초반에는 행동은 최대 속도를 벗어나지 않는 선에서 무작위로 실행되며 후반으로 갈수록 보상이 커지는 행동을 하게 된다.

**Reward:** 시간  $t_i$ 에서 보상  $r_i$ 는 식 10 과 같이 정의한다.

$$r_i = \begin{cases} 0.1 \cdot \frac{D}{i} & \text{out coverage} \\ -1 & \text{out altitude} \\ D_t & \text{else.} \end{cases} \quad (10)$$

$$D = \sum_{t=0}^i D_t \quad (11)$$

보상은 3 가지의 경우로 나뉘게 된다. 먼저 열차가 UAV 의 통신 범위를 벗어나게 되면 에피소드를 종료하며  $0.1 \cdot \frac{D}{i}$ 을 보상하게 된다. 식 11 의  $D_t$ 는 해당 시간의 통신량을 의미하며 결과적으로 에피소드 동안의 평균 통신량을 보상한다. 다음 경우는 UAV 가 허용 고도를 넘어 이동했을 때는 에피소드를 종료하며 -1의 보상을 주어 해당 행동이 나오지 않게끔 유도한다. 그 외의 통신이 안정적으로 이루어지는 경우는 해당 시간에 대한 처리량  $D_t$ 을 보상하며 에피소드를 그대로 진행한다.

#### 4. 실험결과

표 1. 변수 정의 및 사용 값

Notion	Description	Value
$\omega$	The Angle of the sensing cone	90°
$f_0$	Transmitter frequency	5 GHz
$P_t$	Transmitting power	10 dBW
$G_t, G_r$	Antenna gains for the transmitter and receiver	8 dB, 0 dB
$L$	Total System Losses	8 dB

$B$	Bandwidth	20 MHz
$N$	Noise Power	107 dBW
$a, b$	Constants on the environment	4.88, 0.49
$\eta_{LoS}, \eta_{NLoS}$	Additional losses in LoS and NLoS	0.1, 21

표 2. 강화학습 서버 주요 사양

Part	Spec
CPU	Intel® Core™ i5-8500 @ 3.00GHz
GPU	GeForce GTX 1660 Ti
RAM	16.0GB

표 1 은 실험에 사용된 변수들의 정의와 수치를 나타낸다. UAV 가 지원하는 통신을 5G 서비스라고 가정하였기 때문에 주파수와 대역폭을 표 1 에서와 같이 정의하였다. 또한, 교외 상황에서의 통신이라고 가정하여 그에 따른 환경 변수( $a, b$ ), 추가 감쇄 수치( $\eta_{LoS}, \eta_{NLoS}$ )가 정의되었다. 표 2 에서는 강화학습을 실행한 서버의 사양을 확인할 수 있다.

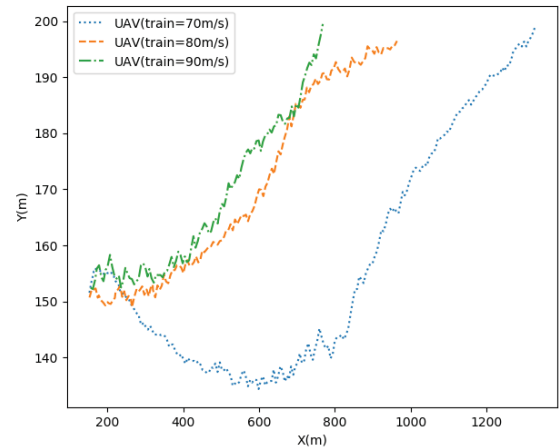


그림 2. 열차 속도에 따른 UAV 경로

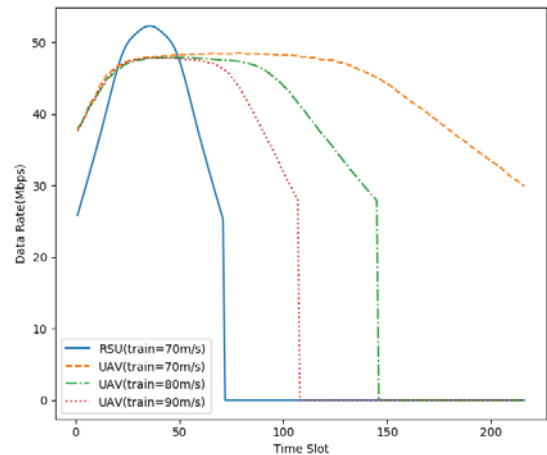


그림 3. 열차 속도에 따른 통신량

그림 2 는 열차의 속도에 따른 UAV 경로를 시간 순으로 나타낸 것이다. 열차의 속도가 빨라짐에 따라 UAV 가 통신을 지원할 수 있는 시간이 줄어드는 것을 알 수 있다. UAV 는 공통적으로 빠른 열차와의 통신을 안정적으로 유지하기 위해 고도를 높이는 것을 확인할 수 있다. 상대적으로 열차가 느린 경우 즉 그림 2 에서 열차 속도가 70m/s 일 때 UAV 는 고도를 낮추어 통신량을 높이려는 움직임 또한 관찰할 수 있다.

그림 3 은 열차의 속도에 따른 통신량을 시간 간격으로 나타낸 것이다. 그림 3 을 통해 일반적인 지상 RSUs 와 UAV 를 비교하면 더욱 오래 통신 서비스를 지원할 수 있다는 것을 알 수 있다. 따라서 UAV 통신은 통신 지원 범위가 넓힐 수 있고 따라서 Handover 를 줄여 기기 과부하 또는 지연 시간을 줄일 수 있을 것으로 기대된다. 그리고 열차의 속도가 빨라짐에 따라 UAV 는 통신을 지속하기 위해 더욱 높은 고도를 유지해야 한다. 이것으로 인해 결과적으로 평균 통신량은 줄어드는 것을 알 수 있다.

## 5. 결론 및 향후 연구

본 논문에서는 고속 열차와 UAV 와의 통신 상황에서 고속 열차에 속도에 따른 UAV 의 경로 최적화 방안에 관해 연구하였다. 최적의 경로를 찾기 위해 강화학습을 Unity ML-Agents 환경에서 사용하였다. 결과적으로 제안한 방법을 통해 기존의 RSUs 와 통신을 할 때보다 UAV 가 열차에 따라 이동을 하면서 통신 지원 범위를 넓힐 수 있어 효율이 좋았다. 또한, 열차의 속도에 따라 UAV 가 고도를 조정하며 최대한 길고 안정적으로 통신을 지원할 수 있었다. 앞으로 본 논문에서 다루지 않았던 다수의 UAV 통신을 통해 통신 지원 범위를 더욱 넓히는 방안과 배터리 효율 및 충전에 대한 변수까지 고려하여 본 논문을 발전시킬 것이다.

## ACKNOWLEDGMENT

이 논문은 2019 년도 정부(과학기술정보통신부)의 재원으로 정보통신기획평가원의 지원을 받아 수행된 연구임 (No.2019-0-01287, 분산 엣지를 위한 진화형 딥러닝 모델생성 플랫폼)  
\*Dr. CS Hong is the corresponding author.

## 참 고 문 헌

- [1] Zhou, Yuzhe. "Future communication model for high-speed railway based on unmanned aerial vehicles." arXiv preprint arXiv:1411.3450 (2014).
- [2] Yu Min Park, Min Kyung Lee, and Choong Seon Hong. "Multi-UAVs Collaboration System based on Machine Learning for Throughput Maximization." 2019 20th Asia-Pacific Network Operations and Management Symposium (APNOMS). IEEE, 2019.
- [3] Qi, Hang, et al. "Energy Efficient 3-D UAV Control for Persistent Communication Service and Fairness: A Deep Reinforcement Learning Approach." IEEE Access 8 (2020): 53172-53184.
- [4] Samir, Moataz, et al. "Joint Optimization of UAV Trajectory and Radio Resource Allocation for Drive-Thru Vehicular Networks." 2019 IEEE Wireless Communications and Networking Conference (WCNC). IEEE, 2019.
- [5] Trotta, Angelo, et al. "Joint coverage, connectivity, and charging strategies for distributed UAV networks." IEEE Transactions on Robotics 34.4 (2018): 883-900.
- [6] Al-Hourani, Akram, Sithamparanathan Kandeepan, and Simon Lardner. "Optimal LAP altitude for maximum coverage." IEEE Wireless Communications Letters 3.6 (2014): 569-572.
- [7] Juliani, Arthur, et al. "Unity: A general platform for intelligent agents." arXiv preprint arXiv:1809.02627 (2018).