

UAVs-as-Relay 환경에서 통신 음영 지역 해소를 위한 Actor Critic 알고리즘 기반 UAV 최적의 배치 방안 연구

이민경⁰, 홍충선*
 경희대학교 컴퓨터공학과
 {minkyung0110, cshong}@khu.ac.kr

Optimal Deployment of UAVs based on Actor Critic Algorithm for Solving Communication Shading Problem in UAV-as-Relay Environment

Minkyung Lee⁰, Choong Seon Hong*
 Department of Computer Science and Engineering, Kyung Hee University

요 약

다양한 모바일 디바이스들의 등장과 함께 이를 사용하는 사용자들의 수 역시 급증하고 있는 추세이다. 급증하는 사용자 수 대비 저지연 고신뢰 통신을 제공하기 위해 5G 네트워크가 상용화되기 시작하였다. 그러나 모바일 디바이스 사용자 수에 비해 5G 네트워크의 상용화는 더딘 추세이고, 산간지역, 교외지역과 같이 도심에서 먼 지역이나 야외 공연장, 축구장 등과 같이 사람들이 많이 몰리는 지역에서 발생하는 통신 장애 또는 고지연 문제가 여전히 발생하고 있는 상황이다. 따라서 본 논문은 이를 해결하고자 Relay로서의 역할을 수행하는 UAV 군집 환경에 Actor Critic 강화학습 알고리즘을 적용하여 UAVs간의 최적의 거리를 유지하며 군집을 이루는 배치 방안을 제안하고자 한다.

1. 서론 및 관련 연구

5G 네트워크가 상용화 되고 있지만, 여전히 네트워크의 안정성과 품질을 높이기 위해서는 다량의 기지국 건설이 필요하다. 이러한 문제를 보완하기 위해 Unmanned Aerial Vehicle(UAV)가 기지국 또는 Relay 역할을 수행하는 연구가 진행되고 있다[1]. 즉, UAVs가 기지국 또는 릴레이로서 역할을 수행하는 연구와 관련하여 자원할당, 서비스 처리량, 통신 지연 부분의 문제들을 해결하려는 연구들이 진행되는 추세이다[1][2]. UAV가 기지국 또는 Relay로서의 역할을 수행하기 위해 UAV 배터리 문제, 보안 문제, 최적의 배치 시스템 문제 등 각종 문제들 역시 대두되고 있다. 논의되고 있는 다양한 문제들 중에서 본 논문은 UAV 최적의 배치 시스템 부분에 대한 기존 연구를 분석하고 새 방안을 제안한다.

논문[3]은 현재 5G 기술의 한계점을 논하며 지상 기지국을 대신하여 기지국으로서의 역할을 수행하는 UAV가 사용자들의 요구를 고려하여 최적의 서비스 처리량을 가질 수 있는 위치를 찾는 방안을 Proximal Policy Optimization(PPO) 강화학습 방법을 적용하여 제안 및 해결하였다. 논문[4]은 Relay 기지국으로서의 역할을 수행하는 UAV 최적의 고도를 찾기 위한 연구로, UAV가 이동 없이 공중에 떠있는 경우와 비행하는 경우 총 두 가지 상황을 연구 제안하였다. 논문[5]은 기지국 서비스 범위 내 기기들의 평균 데이터율을 최대화하고 인접한 기지국의 커버리지 안에 들지 않는 고립된 기기들에 대해 안정적인 통신을 제공하기 위해 다중 UAV

Relay 시스템을 위한 UAV 배치 방안을 제안하였다. 이와 같이 기존 연구들의 경우, UAV와 기지국과의 지속적인 통신을 통해 계산된 최적의 위치를 찾아가는 방안들을 제안하였다.

그러나 본 연구의 경우, 기지국과 UAV-Relay 군집이 최초 단 한 번만 통신을 하고 그 후 이루어지는 최적의 배치 방안은 UAV에 부착된 에지 컴퓨팅에서의 반복적인 학습 훈련을 통해 이루어진다.

2. 제안 사항

2.1 시스템 모델

본 논문은 Relay 역할을 수행하는 다수의 UAV들을 각각 최적의 지점에 배치하여 지연 시간이 낮은 통신을 제공하는 방안을 제안한다. 이때 최초 단 한 번을 제외하고 최적의 지점을 찾기 위한 반복적인 학습 훈련은 각 UAV마다 부착된 에지컴퓨팅 내 Actor Critic 알고리즘을 통해 진행되며 UAV 간의 지속적인 통신을 통한 최적의 배치 방안을 연구한다.

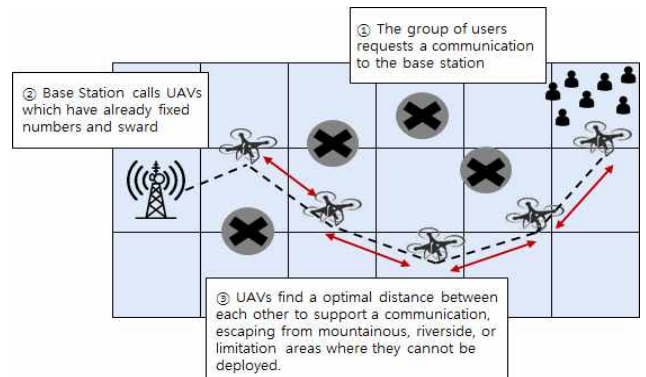


그림 1. UAV as a Relay proposed model

본 연구는 과학기술정보통신부 및 정보통신기술진흥센터의 Grand ICT연구센터 지원사업의 연구결과로 수행되었음 (IITP-2019-2015-0-00742) 또한 이 논문은 2019년도 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

그림 1과 같이 교외 지역에서 다수의 사용자들이 먼 거리에 위치한 기지국에 통신을 요청하는 경우 저지연 네트워크 지원 제공이 어렵다. 따라서 기지국에서 통신 음영 지역까지 Relay로서의 역할을 수행하는 다수의 UAV를 보낸다. 군집을 이루는 UAVs는 서로 간에 단절 없는 통신거리를 확보하기 위해 최적의 거리를 유지하며 다수의 사용자들이 위치한 지역까지 도달하게 된다. 이때 UAVs는 정해진 일정한 고도에 위치하며, UAVs의 배터리 수명시간은 충분하기에 배터리 교체 또는 UAVs 교체 문제는 발생하지 않는다. 본 논문은 UAVs간의 최적의 거리 기반 배치 방안을 고려하여 원거리 저지연 통신 제공 향상 방안을 연구한다.

2.2 UAV 통신모델

본 논문은 UAVs간의 최적의 거리 기반 배치 방안을 고려하여 원거리 저지연 통신 제공 향상 방안을 연구한다.

$$P_{LoS} = \frac{1}{1 + \epsilon \exp(-\beta[\theta - \epsilon])} \quad (1)$$

$$P_{NLoS} = 1 - P_{LoS} \quad (2)$$

수식(1)은 air to ground 채널 모델로 Line-of-Sight(LoS)를 고려한 값이다. 수식(1)에서 구한 P_{LoS} 값을 토대로 수식(2)의 Non-Line-of-Sight(NLoS) 값을 구할 수 있다. 이때 NLoS는 장애물로 회절 또는 반사되어 전파되는 비직진파를 의미한다[1]. 본 연구의 경우 장애물에 의해 회절 또는 반사되어 통신 저지연이 발생하는 경우에 한하여 통신지원을 하는 UAVs 군집의 최적 거리를 구하고자 하기에 수식(2)에서 구한 P_{NLoS} 값이 요구된다. 수식 파라미터에 대한 설명은 표1을 참고한다.

표 1. Notion Definitions

| Notion | Description |
|-------------------|--|
| θ | elevation angle |
| ϵ, β | constant which represent environments |
| γ | discount factor which is a binary function |
| V | value function approximator |

2.3 Actor-Critic 모델 기반 최적 배치 방안

2.2절의 수식(2)를 통해 구한 통신 음영 지역 네트워크 통신 비가시 거리와 UAVs 간의 채널 모델을 바탕으로 UAVs 간의 최적의 거리를 학습을 통해 찾는 방안을 연구한다. 이때 UAVs들의 고도는 일정하다는 가정 하에 UAVs 간의 거리만을 고려하여 통신 지연을 낮추는 방안을 Actor Critic 모델 기반 Reward가 최상의 값인 경우의 Policy를 찾을 수 있다.

- **State** State는 두 가지 요소, UAVs 위치 $\{x_i^i, y_i^i\}$ 와 UAVs 간의 거리 d_i^i 로, $s_t^i = \{x_i^i, y_i^i, d_i^i\}$ 구성된다. 따라서 State는 $S = \{s_t^i | i \in N, t = 1, \dots, T\}$ 로 나타낼 수 있다.

- **Action** 각각의 UAVs들이 서로 간의 최적의 거리를 찾기 위한 과정을 의미한다. 즉, $a_t^i = \{d_t^i\}$ 가 된다.
- **Reward** Reward는 UAVs의 Action에 대한 통신 지연 시간의 변화이다. 이 때 군집을 이루어 UAVs 경우기에 Reward가 최대가 된다고 해서 반드시 최적의 UAVs 거리를 야기한다고 가정할 수 없다. 즉, UAVs 군집 전체의 통신 지연 시간도 계산되어야한다.

$$\theta \leftarrow \theta + \theta_\alpha Q_\omega(s, a) \nabla_\theta \ln \pi_\theta(a|s) \quad (3)$$

$$\delta_t = r_t + \gamma Q_\omega(S', a') - Q_\omega(s, a) \quad (4)$$

앞서 언급한 State, Action, Reward 모델을 기반으로 수식(3)은 최적의 Policy를 찾기 위해 gradient descent 방식으로 가중치를 업데이트 하는 것을 의미한다. 또한 수식(4)를 통해 action-value을 위한 Temporal Difference error의 값을 구하여 action-value의 가중치를 업데이트 할 수 있고 결과적으로 action과 state값을 업데이트하며 학습과정을 반복하게 된다[6].

Algorithm: Process of proposed approach

- 1 Initialize s, θ, ω at random; sample $a \sim \pi_\theta(a|s)$.
- 2 For $t = 1, \dots, T$:
- 3 Sample reward $r_t \sim R(s, a)$ and next state $s' \sim P(s'|s, a)$;
- 4 Then sample the next action $a' \sim \pi_\theta(a'|s')$;
- 5 Update the policy parameters: $\theta \leftarrow \theta + \alpha_\theta Q_\omega(s, a) \nabla_\theta \ln \pi_\theta(a|s)$;
- 6 Compute the correction for action-value at time t :
- 7 $\delta_t = r_t + \gamma Q_\omega(S', a') - Q_\omega(s, a)$
- 8 and use it to update the parameters of action-value function:
- 9 $\omega \leftarrow \omega + \alpha_\omega \delta_t \nabla_\omega Q_\omega(s, a)$
- 10 Update $a \leftarrow a'$ and $s \leftarrow s'$.
- 11 Break

위의 알고리즘은 본 논문에서 적용하는 actor critic 알고리즘의 일련의 동작 과정을 나타낸다. 앞서 정의한 State, Action, Reward 모델과 수식(3)과(4)를 기반으로 통신 지연 시간을 최소화할 수 있는 UAVs 간의 최적의 거리 기반 배치 방안을 학습하는 과정이다.

3. 성능평가

3.1 성능평가 환경 구성

본 연구에서 제안한 방안의 성능은 UAVs 간의 거리 대비 통신 지연으로 구하였다. 이를 위해 구성한 환경은 N대의 군집을 이루는 UAVs이며 통신 음영지역과 기지국간의 거리는 최대 270km(서울에서 포항까지의 거리 기반)에 위치한다고 전제하는 환경 속에서 성능평가를 진행하였다.

3.2 성능평가 결과

4.1 성능평가 환경 구성을 기반으로 진행한 성능평가에 대한 결과 그래프를 말한다.

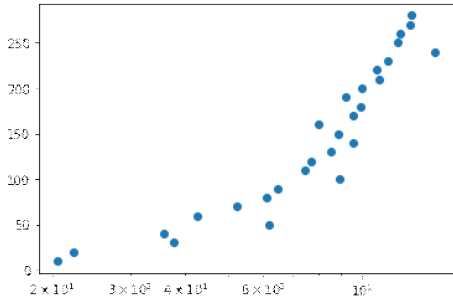


그림 1. 거리 대비 UAV-Relay 배치 분포도

총 500번의 에피소드를 실행하여 학습 훈련을 진행하였을 때 (0,0)에 위치한 지상 기지국에서부터 통신 음영 지역인 280km까지 UAV-Relay가 배치되는 분포도를 그림 1을 통해 알 수 있다. 지상 기지국 근처에서는 기지국 커버리지 내에 들기에 UAV-Relay 분포가 간격을 보이는데 반면, 해당 커버리지가 벗어난 120km 이상부터는 UAV-Relay의 분포가 급증하는 것을 알 수 있다.

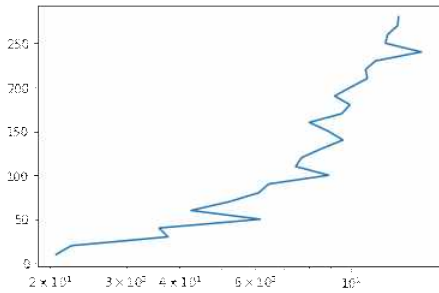


그림 2. UAV-Relay 배치 분포도 평균값

그림 2는 그림 1에서 보여준 500번의 에피소드 기반 학습 훈련 데이터를 토대로 1000번의 에피소드로 학습 훈련을 진행한 결과 나타난 UAV-Relay 분포도의 평균을 연결한 그래프이다. 이를 통해 통신 음영 지역에 이룰수록 많은 양의 UAV-Relay가 요구됨을 알 수 있다.

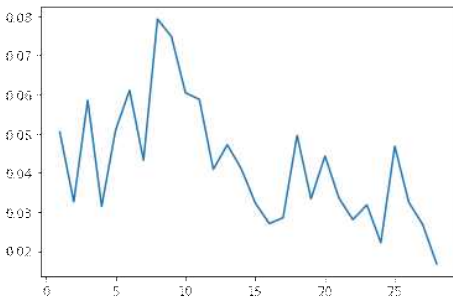


그림 3. UAV-Relay 배치에 따른 지연시간 변화

그림 3은 통신 음영 지역까지 UAV-Relay가 배치되는 분포도 대비 통신 지연시간 관련 그래프이다. UAV-Relay가 배치될 때, UAV 에너지, 통신을 요청하는 정확한 사용자 분포도에 대한 변화율을 고려하지 않아 통신 지연을 감소가 일정하지는 않지만, UAV-Relay를 배포함으로써 기지국이 있는 (0,0)의 위치보다 통신 지연시간이 현저히 감소하였음을 알 수 있다.

4. 결론 및 향후 연구

본 논문은 그림 1과 같이 릴레이 역할을 수행하는 UAVs들이 통신 음영 지역의 네트워크 자원 할당을 도울 때 군집을 이루는 UAVs간의 최적의 거리 모델을 예측 및 학습하여 저지연 통신을 제공하는 UAVs 군집의 최적의 배치를 찾는 방안을 연구하였다. 그러나 제한 조건으로는 군집을 이루는 다수의 UAVs가 동일한 고도에 위치하였고, UAVs간의 직교성 통신으로 간섭이 없었으며, UAVs 배터리 수명은 충분하다는 전제를 통해 진행된 연구라는 부분에서 한계가 존재한다. 따라서 향후 연구에서는 군집을 이루는 UAVs에서 각각의 UAV가 통신 환경 대비 커버리지를 고려한 최적의 고도 배치 및 UAVs 간 비직교성 통신으로 인해 발생하는 간섭을 고려하여 통신 음영 지역까지의 UAVs 군집의 최적의 배치 방안을 연구하고자 한다.

참 고 문 헌

- [1] Yong Zeng, Rui Zhang, and Teng Joon Lim, "Wireless Communications with Unmanned Aerial Vehicels: Opportunities and Challenges", IEEE Communications Magazine, vol.54, Issue.5, paget.36-42, May 2016
- [2] Bin Li, Zesong Fei, and Yan Zhang, "UAV Communications for 5G and Beyond: Recent Advances and Future Trends", IEEE Internet of Things Journal, vol.6, no.2, April, 2019
- [3] Yu Min Park, Minkyung Lee, Choong Seon Hong, "Multi-UAVs Collaboration System based on Machine Learning for Throughput Maximization", The 20th Asia-Pacific Network Operations and Management Symposium(APNOMS), 2019
- [4] Yunfei Chen, Wei Feng, Gan Zheng, "Optimum Placement of UAV as Realys", IEEE Communications Letters, vol.22, Issue.2, Feb, 2018
- [5] Sang Ik Han, Jaeuk Baek, and Youngnam Han, "Deployment of Multi-layer UAV Relay System", 2018 IEEE Wireless Communications, vol.22, no.1, 2018
- [6] Ian Goodfellow, Yoshua Bengio and Aaron Curville, "Deep Learning", [Online]. Available: <http://www.deeplearningbook.org/> (downloaded 2019, Oct. 29)