

군집 무인기 네트워크를 위한 강화학습 기반 Slotted ALOHA 프로토콜

김상호, 유희정

고려대학교, 전자·정보공학과

{kgn7392, heejungyu}@korea.ac.kr

RL-based Slotted ALOHA protocol for group UAV networks

Sangho kim, Heejung Yu

Department of Electronics and Information Engineering, Korea University

요약

본 논문에서는 Slotted ALOHA 방식에 강화학습의 Q-learning 알고리즘을 적용하였고, 이를 UAV 통신에 사용하였다. 이를 통해 UAV간 통신에 Slotted ALOHA 방식을 사용해 발생하는 문제인 낮은 전송 성공 확률, 제한된 처리량을 해결하고, 구현이 간단하고 지연율이 낮다는 장점은 극대화 시키고자 하였다.

I. 서론

Unmanned Aerial Vehicle(UAV)은 군사 목적으로 개발되었지만 높은 비행 자유도와 간단한 구조로 빠르게 발전하였고 최근에는 다양한 민간 서비스에 이용되고 있다. 하지만 UAV는 작은 크기라는 태생적 한계로 기상의 영향을 많이 받고 적은 배터리 용량 등 다양한 문제를 가진다. 이를 극복하기 위해 UAV의 군집비행이 UAV 활용에 필수적 요소로 자리 잡았다. UAV의 군집비행에서 가장 핵심이 되는 부분은 기체 간 충돌 방지이다. 이를 위한 알고리즘은 중앙집중 제어방식과 분산 제어방식으로 나뉜다. 본 논문에서는 UAV들이 서로의 위치 정보를 주고받아 통신이 이루어지는 분산 제어방식을 연구하였다.

Additive Links On-line Hawaii Area(ALOHA) 프로토콜은 1971년 하와이 대학교에서 개발한 프로토콜로 발신 노드가 패킷을 전송하면, 수신 노드는 유효한 패킷인가를 확인하고 ACK 신호를 발송, 송신 노드는 ACK 신호가 기다리다 최대왕복 전파지연 시간 이상이 되면 해당 패킷을 재전송하는 방법으로 동일 채널을 다수의 노드가 이용하는 방법이다. 만약 동시에 여러 노드가 패킷을 보내면 충돌이 발생하고 이때 수신 노드는 해당 패킷들을 무시한다. Slotted ALOHA 프로토콜을 기존의 Pure ALOHA 프로토콜의 충돌이 자주 발생하는 문제를 해결하기 위해 하나의 프레임에 다수의 슬롯으로 나누고 슬롯의 시작 부분에서만 패킷을 전송하여 충돌 확률을 감소시킨 방식이다. 하지만 Slotted ALOHA 프로토콜 역시 각 노드가 사용할 슬롯을 랜덤하게 결정하여 패킷을 전송하기 때문에 Pure ALOHA와 마찬가지로 노드 수가 많아짐에 따라 전송 성공 확률이 낮아지는 문제가 발생한다. 그래서 이를 UAV 분산 제어방식에 사용할 경우 UAV의 충돌을 야기할 수 있다. 하여 본 논문에서는 이를 보완하고자 Q-learning의 상태 가치 함수(Q-value)를 적용했다.

Q-learning은 강화학습의 핵심 알고리즘 중 하나로, 행동의 주체가 되는 에이전트가 환경과 상호작용하며 각 상태에서 모든 행동에 대해 기대되는 보상을 추정하고, 이를 바탕으로 최적의 행동을 선택하는 과정을 반복 결과적으로 최적의 행동 정책을 학습하는 방법이다. 이를 통해 각 UAV는

자신의 위치 정보를 전송할 최적의 슬롯을 학습할 수 있다.

II. 본론

본 논문에서 가정하는 상황은 그림 1과 같다. UAV는 R의 간격으로 25기가 정사각형 형태로 배치되어있다. 각각의 UAV는 프레임당 하나의 슬롯에서 임의의 반경에 자신의 위치 정보를 전송한다.

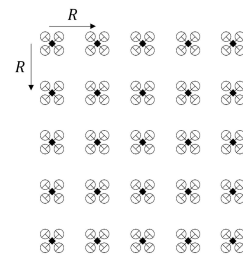


그림 1. UAV 배치

사용하는 프레임의 구조는 아래 그림 2와 같다. 한 프레임은 25개의 슬롯으로 구성되어 있으며 전체 프레임의 개수는 N개이다.

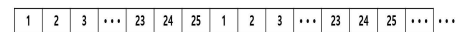


그림 2. 프레임의 구조

실험이 시작되면 각 UAV는 모든 슬롯에 0이라는 Q-value를 가지고 아래와 같은 수식으로 자신의 행동을 선택한다. 행동은 어떤 슬롯에 위치 정보를 전송할 것인가이다.

$$ACTION = \begin{cases} \max Q\text{-value} & \text{with probability } 1-\epsilon \\ \text{random} & \text{with probability } \epsilon \end{cases}$$

한 프레임이 종료되면 각 슬롯에서 2개 이상의 UAV에 위치 정보를 전송 받은 UAV가 있는지를 확인하고 2개 이상의 위치 정보를 전달받은 UAV가 존재할 경우 충돌이 발생했다고 가정 $-1 * (\text{충돌이 발생한 UAV의 수})$ 의 reward가 해당 슬롯을 사용한 UAV들에 업데이트된다. 반대로 한 슬롯에서 UAV들이 하나의 위치 정보만을 전송 받은 경우 해당 슬롯을 사용한 UAV들에 $+1$ 의 reward가 업데이트된다. 그리고 업데이트된 reward를 통해 Q-value를 계산하고 다음 프레임으로 넘어간다. Q-value를 계산하는 수식은 아래와 같다.

$$Q\text{-value} \leftarrow \text{learning rate} * (\text{기존 } Q\text{-value} + \text{reward})$$

learning rate란 reward, 기존의 Q-value 등의 정보가 새로운 Q-value에 얼마나 영향을 줄지 결정하는 요소이다. learning rate가 클수록 모델은 새로운 정보에 빠르게 적응하지만 Q-value가 수렴하는데 방해될 수 있다. 반대의 경우로 작은 경우 학습의 속도가 더딜 수 있다.

MATLAB을 통해 실험을 진행하였고 $N = 500$, learning rate = 0.8, epsilon의 경우는 epsilon greedy policy를 사용하여 1에서 시작해 서서히 값이 작아지고 $N = 150$ 에서 완전히 0이 되도록 설정하였다. 반경의 경우는 2R, 3R, 4R, 5R로 4차례 진행하였다.

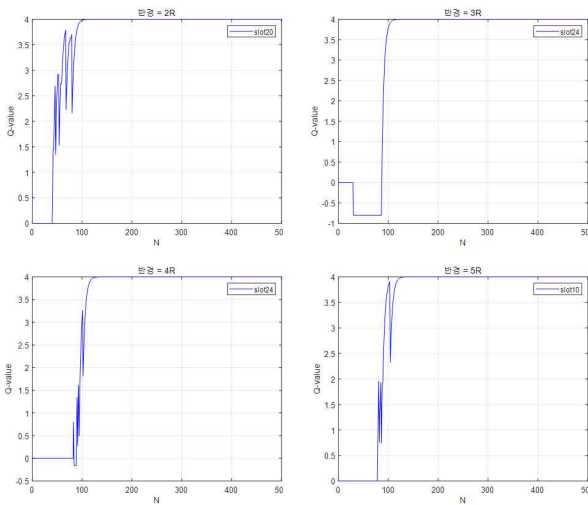


그림 3. 각 반경에서 1번 UAV의 Q-value 수렴과 수렴한 슬롯

실험 결과 그림 3과 같이 각 반경에서 모든 UAV가 하나의 슬롯에 Q-value가 4로 수렴하는 모습을 볼 수 있었다. Q-value가 4로 수렴하는 이유는 learning rate를 0.8로 설정하여 그 값이 등비수열의 형태로 계산되기 때문이다. 추가로 위 실험을 각 반경에서 100회 반복한 결과 반경 2R의 경우 평균적으로 20개의 슬롯을 반경 3R 이상의 경우에는는 대부분 25개의 슬롯을 모두 사용하는 결과를 얻었다. 수렴시간이나 수렴 값이 달라지는 차이는 보이지 않았다.

III. 결론

본 논문에서는 Slotted ALOHA 프로토콜에 Q-learning을 적용하여 UAV 통신에 Slotted ALOHA 방식을 사용했을 때 나타나는 문제인 낮은 전송 성공률, 제한된 처리량 등의 문제를 해결하고자 했다. 현재 실험에서 각각의 UAV들이 자신의 배치를 모두 알고 있고 자신은 어느 UAV에 신호를 보내야 하는지 또 슬롯의 수는 어떻게 되는지 모두 알고 있다고 가정

하고 실험을 진행하였다. 추후에는 모든 UAV가 이러한 정보를 모르는 상태로 ACK 신호를 활용 UAV들이 스스로 자신들의 배치와 거리, 최적의 슬롯 수, 최적의 반경을 학습하는 모델을 만들고자 한다.

ACKNOWLEDGMENT

This work was partly supported by “Regional Innovation Strategy (RIS)” through the National Foundation of Korea (NRF) funded by the Ministry of Education (MOE) (2021RIS-004), Basic Science Research Program through the NRF funded by the MOE (2021R111A3041887), and Institute of Information & communications Technology Planning & Evaluation (IITP) grant funded by the Korea government (MSIT) (2022-0-00704, Development of 3D-NET Core Technology for High-Mobility Vehicular Service).

참고 문헌

- [1] 장현준, 황경호, “Q-learning 기반 Dynamic Framed Slotted ALOHA 프로토콜”, 한국통신학회 동계종합학술발표회, pp. 959-960, 2021