

유한 프레임 길이에서 다중 에이전트 심층 강화 학습을 활용한 IRSA-NOMA 최적화 연구

류원재, 이재민, 김동성

국립금오공과대학교

{wj0828, ljmpaul, dskim}@kumoh.ac.kr

A Study on Optimization of IRSA-NOMA with Finite Frame Length Scenarios Using Multi-Agent Deep Reinforcement Learning

Ryu Won Jae, Lee Jae-Min, Kim Dong-Seong

Kumoh national Institute of Technology

요 약

본 연구에서는 다중 에이전트 심층 강화 학습(Multi Agent Deep Reinforcement Learning, MADRL)을 활용하여 Irregular Repetition Slotted ALOHA(IRSA)와 Non-Orthogonal Multiple Access (NOMA) 시스템의 전송 횟수 최적화를 통해 패킷 손실률(Packet Loss Rate, PLR)을 낮추는 방법을 제안한다. 제안된 방법에서는 각 노드별로 이중 심층 Q-네트워크(Double Deep Q-Network, DDQN)를 적용하여 전송 횟수를 최적화하고, 이를 통해 패킷 손실을 최소화하고자 하였다. 프레임 길이를 10으로 고정된 모의실험을 수행하였으며, 결과적으로 Random 및 Fixed Degree Distribution 방법과 비교했을 때, 제안된 방식이 더 낮은 패킷 손실을 보여, 제안된 방법이 IRSA-NOMA 시스템의 성능 향상을 할 수 있음을 확인하였다.

I. 서 론

최근 사물인터넷(Internet of Things, IoT)은 다양한 산업 분야에서 혁신을 촉진하며, 물리적 기기와 디지털 시스템 간의 상호 연결을 가능하게 하고 있다. 이러한 연결은 실시간 데이터 수집 및 분석을 통해 효율성과 생산성을 극대화할 수 있는 잠재력을 제공한다. 특히, IoT는 스마트 홈, 헬스케어, 제조업, 물류 등에서 중요한 역할을 하며, 이를 통해 더 나은 의사결정과 자동화된 운영이 가능해진다. 그러나 방대한 양의 데이터를 처리하고 복잡한 네트워크 환경에서 실시간으로 반응하기 위해서는 고도화된 분석 및 최적화 기술이 필요하다. 인공지능(Artificial Intelligence, AI)은 이러한 요구를 충족시키는 핵심 기술로, IoT 시스템의 데이터를 분석하고 패턴을 인식하며 예측을 수행하는 데 중요한 역할을 한다. AI는 IoT 디바이스 간의 상호작용을 최적화하고, 효율적인 자원 할당을 가능하게 하며, 복잡한 문제를 신속하게 해결할 수 있도록 돕는다. 따라서 IoT와 AI의 융합은 스마트 시티, 자율주행 차량, 스마트 팩토리 등 차세대 기술 발전의 중요한 기반이 되고 있다 [1].

Irregular Repetition Slotted ALOHA (IRSA)는 불규칙 반복 전송을 통해 패킷의 전송 확률을 높이는 방법으로써, 상향링크 무선 통신 시스템에서 신뢰성을 높이는 데 기여할 수 있다 [2]. Non-Orthogonal Multiple Access (NOMA)는 여러 사용자가 동일한 주파수 대역을 공유하면서도 서로 다른 전력 레벨을 활용하여 동시에 통신할 수 있도록 함으로써 스펙트럼 효율성을 극대화한다 [3]. 이러한 기술들의 결합은 IoT 환경에서 저지연, 고신뢰, 고효율 통신을 가능하게 하며, 복잡한 네트워크 문제를 효과적으로 해결할 수 있는 기반을 제공한다.

본 연구에서는 다중 에이전트 심층 강화 학습(Multi Agent Deep Reinforcement Learning, MADRL)을 활용하여 IRSA-NOMA 시스템의 전송 횟수 최적화를 통해 패킷 손실률(Packet Loss Rate, PLR)을 낮추는

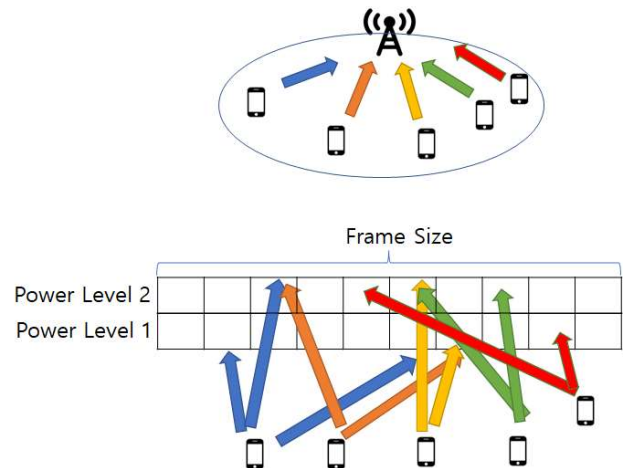


그림 1 시스템 모델

방법을 제안한다. 제안된 방법에서는 각 노드별로 이중 심층 Q-네트워크(Double Deep Q-Network, DDQN)[4,5]를 적용하여 전송 횟수를 최적화하고, 이를 통해 패킷 손실을 최소화하고자 하였다. 프레임 길이를 10으로 고정된 모의실험을 수행하였으며, Random 및 Fixed degree distribution과 비교하였다.

II. 본론

본 논문에서는 유한 프레임 길이 IRSA-NOMA를 적용한 IoT 시스템에서 각 Node별로 DDQN을 활용한 MADRL에 대한 연구를 수행하였다. 그림1의 시스템 모델과 같이 프레임길이는 10으로 하였고, 총 5개의 상향 링크 노드가 있다고 가정하여 모의실험을 실행하였다. DDQN에 대한 수식은 다음과 같다. 각 노드별로 전송할 패킷이 발생하면 전송전에 저장할 버

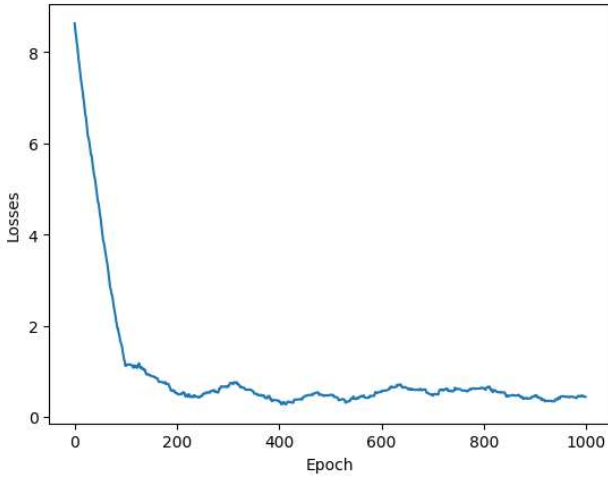


그림 2 학습 결과

퍼를 가지는데, 해당 버퍼 크기는 5로 최대 5개의 패킷을 저장할 수 있다. 전송에 성공하면 버퍼의 패킷은 하나씩 사라진다. 하지만 전송에 연속해서 실패하여 패킷이 5개를 초과하여 누적되는 경우 packet loss가 발생한다. 채널 상태는 Quasi static Rayleigh fading을 가정하였고, 발생하는 채널 이득에 따라, Orthogonal Multiple Access (OMA)로 전송할지, NOMA로 전송할지, 전송하지 않을지 결정하게 된다. OMA로 전송하는 경우는 채널이득이 h_{ti} 이상이고 h_{th} 미만일 때이고, NOMA로 전송하는 경우는 채널이득이 h_{th} 이상일 때이다. h_{ti} 보다 작을 때는 전송하지 않고 버퍼에 패킷이 저장된다.

$$Y_t = R_{t+1} + \gamma Q(S_{t+1}, \arg\max_a Q(S_{t+1}, a; \theta_t); \theta'_t). \dots (1)$$

DDQN은 action을 선택할 때 파라미터 θ_t 인 Policy network를 사용하고, action을 평가할 때는 파라미터 θ'_t 인 Target network를 사용하여 평가한다. 따라서 기존 Deep Q-learning network에서 발생할 수 있는 Overestimate 현상을 줄일 수 있다. 신경망은 3x8x8x8의 구조를 가진다. 활성화 함수로는 ReLU를 사용하였으며, 입력은 채널이득, 버퍼상태, 현재까지 발생한 Packet loss 수이다. 출력은 8개인데 1부터 8까지 각각의 반복 전송횟수를 나타낸다.

III. 모의실험 및 분석

본 절에서는 모의실험 및 결과 등에 대해서 다룬다. 모의실험 환경으로는 Intel(R) Core(TM) i7-10700K CPU @ 3.80GHz ~ 3.79 GHz, RAM 16GB, RTX 2060 Super, Windows 10 Pro, Pytorch 2.4.0, Cuda 12.4이다. 프레임 사이즈는 10으로 하였고, 전송 노드 수는 5이다. 최대 1000번의 전송을 실시하였고, 전송 횟수에 대해서 1~8을 무작위적으로 선택하는 Random 방식, 기존 연구에 사용된 Degree Distribution $0.5112x^2 + 0.266x^3 + 0.2228x^8$ 인 Fixed, 본 연구에서 제안하는 DDQN기반 MADRL방식을 비교하였다. 그림 2에서는 제안하는 DDQN

표 1 전송 방식 및 노드별 Packet Loss 결과

Node No	Random	Fixed	Proposed
1	245	154	130
2	238	158	129
3	247	181	136
4	246	175	134
5	239	162	119
Average	243	166	129.6

기반 MADRL 방식의 학습 결과를 보여준다. 반복횟수 약 200회까지 Loss가 줄어들고 그 뒤로는 유지되는 것을 볼 수 있다. 이를 통해 제안방식을 통해 IRSA-NOMA의 전송 반복에 대한 학습이 이루어짐을 확인할 수 있다. 표1은 Random 방식, Fixed 방식, 제안 방식을 비교한 결과이고 1000회 전송에 따른 노드별 Packet loss 발생 횟수와 그 평균값이다. Random이 가장 높은 Packet loss를 보여주고 그 다음이 Fixed, 가장 낮은 것은 제안 방식이다. 이를 통해 제안 방식이 가장 낮은 Packet loss를 보여주어 성능이 가장 좋음을 알 수 있다.

IV. 결론

본 연구에서는 유한 프레임 길이에서 IRSA-NOMA 시스템의 전송 횟수 최적화를 위해 다중 에이전트 심층 강화 학습(MADRL)과 이중 심층 Q-네트워크(DDQN)를 적용한 방식을 제안하였다. 모의실험 결과, 제안한 방법이 기존의 Random 및 Fixed 방법과 비교하여 더 낮은 Packet Loss Rate (PLR)을 보여주었다. 특히, 제안 방식은 전송 반복 횟수에 대한 최적화를 통해 패킷 손실을 최소화하는데 효과적임을 확인할 수 있었다. 이를 통해 IRSA-NOMA 시스템의 성능을 더욱 향상시킬 수 있음을 확인하였다. 추후 연구로는 시나리오들을 추가하여 제안 방식에 대한 성능 평가를 하는 것과 NOMA user pair간 전력 할당 부분에 대한 최적화를 이루는 부분이 될 수 있다.

ACKNOWLEDGMENT

이 논문은 2024년도 정부의 재원으로 한국연구재단의 지원을 받아 수행된 연구임(2022R1I1A1A01069334, 20%, 2018R1A6A1A03024003, 20%), 이 논문은 과학기술정보통신부 정보통신기획평가원의 지원을 받아 수행된 연구임 (IITP-2024-2020-0-01612, 20%), 본 연구는 산업통상자원부와 한국산업기술진흥원, 경북지역산업진흥원이 지원하는 지역혁신클러스터 육성사업으로 수행된 연구결과(P0025729, 20%)임. 본 연구는 과학기술정보통신부 및 정보통신기획평가원의 대학ICT연구센터사업의 연구결과로 수행되었음(IITP-2024-RS-2024-00438430, 20%).

참 고 문 헌

- [1] H. Tran-Dang and D. -S. Kim, "A Survey on Matching Theory for Distributed Computation Offloading in IoT-Fog-Cloud Systems: Perspectives and Open Issues," in IEEE Access, vol. 10, pp. 118353-118369, 2022.
- [2] W. J. Ryu, J. H. Lee, J. H. Yu and D. -S. Kim, "Enhancing Reliability and Energy Efficiency for IndustrialIoT Using IRSA-PSO," in The Journal of Korean Institute of Communications and Information Sciences, vol. 48, no. 10, pp. 1203-1210, 2023.
- [3] W. J. Ryu, J. W. Kim and D. -S. Kim, "Resource allocation in downlink VLC-NOMA systems for factory automation scenario," in Sensors, vol. 22, no. 23, 2022.
- [4] W. J. Ryu, J. W. Kim and D. -S. Kim, "Deep Reinforcement Learning Based Cooperative Retransmission in Downlink NOMA Systems," 2022 13th International Conference on Information and Communication Technology Convergence (ICTC), Jeju Island, Korea, Republic of, 2022, pp. 883-885.
- [5] S. Bhardwaj, J. -M. Lee and D. -S. Kim, "Double Deep Q-Learning Based Channel Estimation for Industrial Wireless Networks," 2020 International Conference on Information and Communication Technology Convergence (ICTC), Jeju, Korea (South), 2020, pp. 1318-1320.