

연속중계네트워크를 위한 계층적 강화학습 기반 자원할당 기법

이예린*, 정방철°, 이호원*

*아주대학교, °충남대학교

yerin1205@ajou.ac.kr, bcjung@cnu.ac.kr, howon@ajou.ac.kr

Hierarchical QL-Based Resource Allocation for Barrage Relay Networks

Yerin Lee*, Bang Chul Jung°, Howon Lee*

*Ajou University, °Chungnam National University

요약

본 논문은 연속중계네트워크(barrage relay networks, BRNs)를 위한 계층적 강화학습 기반 자원할당 기법 연구를 수행한다. 구체적으로, 연속중계 네트워크의 제어된 연속 지역(controlled barrage regions, CBRs)에서 견고성과 에너지 효율성을 고려한 네트워크를 구축함과 동시에 통신 오버헤드 최소화를 목표로 전송 전력을 제어하는 내부 루프 강화학습과 초과 폭 파라미터를 제어하는 외부 루프 강화학습으로 동작하는 계층적 강화학습 기법을 제안한다.

I. 서론

미래 전술 MANET(mobile ad-hoc network)의 운영 효율 극대화를 위해 연속중계네트워크(barrage relay networks, BRNs)가 제안되었다 [1][2]. 이는 TDMA(time division multiple access)와 자율협력통신체계를 기반으로 동작하며 대략적인 슬롯 수준의 동기화 조정만을 필요로 하여 오버헤드가 적고 패킷 충돌 및 중첩 문제를 회피하여 수신확률을 높인다는 이점이 있다. 본 논문에서는 전술 환경의 지형적 특성으로 인해 링크 및 노드 손실을 겪는 연속중계네트워크의 견고성 및 에너지 효율성을 보장하고 통신 오버헤드를 최소화하기 위한 계층적 강화학습 기반 최적 자원할당 기법을 제안한다.

II. 연속중계네트워크를 위한 계층적 강화학습 기반 자원할당 기법

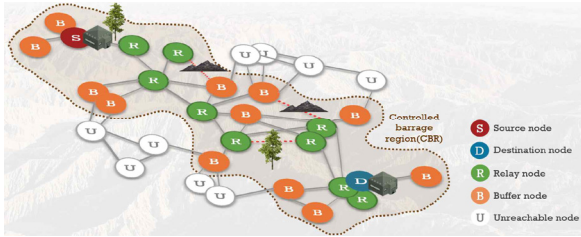


그림 1. 전술 MANET에서의 제어된 연속 지역 구성 예시

연속중계네트워크는 그림 1과 같이 제어된 연속 지역(controlled barrage regions, CBRs)을 형성하여 효율적 유니캐스트 전송 지원이 가능하다 [2]. 이는 소스, 대상, 중계, 버퍼 노드로 구성되며 노드 간 거리(홉단위)에 기반하여 $D_{RTS_i} + D_{CTS_i} \leq D_{min} + \psi$ 의 조건을 만족하는 경우 중계 노드로 선정되어 제어된 연속 지역이 결정된다. 여기서 D_{RTS_i}, D_{CTS_i} 는 각각 소스 및 대상 노드로부터 각 노드까지의 거리, D_{min}, ψ 는 각각 최단 거리와 초과 폭 파라미터를 의미한다. 전술 환경의 지형적 특성으로 링크 및 노드 손실을 겪는 연속중계네트워크에서 초과 폭 파라미터는 중계 노드의 수를 증가시켜 처리량을 높이거나 불필요한 중계 노드의 수를 줄여 에너지 효율적인 네트워크를 구축할 수 있어 매우 중요한 지표로 고려된다.

본 논문에서는 연속중계네트워크의 통신 견고성과 에너지 효율성을 향상시키는 동시에 강화학습의 계산 복잡도를 줄이기 위한 계층적 다중 에이전트 큐리닝 기반 최적 자원할당 기법을 제안한다. 제안 기법은 최적 초과 폭 파라미터 결정을 위한 중앙 집중형 외부 루프 강화학습과 최적 전송 전력 결정을 위한 분산형 내부 루프 강화학습으로 구성된다. 초과 폭 파라미터의 변동은 제어된 연속 지역의 재구축으로 네트워크 관점에서 큰 오버헤드를 유발하여 외부 루프 강화학습의 에이전트인 소스 노드(S_i)가 일정 주기(T^{Out})마다 제어한다. 이는 범위 $[\psi^{min} : \psi^{max}]$ 내에서 $\pm \Delta \psi$ 만큼의 초과 폭 파라미터를 조정하거나 유지할 수 있다. 따라서 외부 루프 강

화학습의 행동 집합은 $A_S^{Out} = \{+\Delta \psi, -\Delta \psi, 0\}$ 로 정의되며 상태는 현재의 초과 폭 파라미터로 정의된다. 비교적 적은 양의 오버헤드가 발생하는 전송 전력은 매 이터레이션(t)마다 내부 루프 강화학습을 통해 제어되며 네트워크를 구성하는 다수의 노드(N_i)들이 에이전트가 되어 범위 $[P_i^{min} : P_i^{max}]$ 내에서 $\pm \Delta P_i$ 만큼의 전송 전력을 조정하거나 유지한다. 따라서, 내부 루프 강화학습의 행동은 $A_{N_i}^In = \{+\Delta P_i, -\Delta P_i, 0\}$ 로 정의되며 상태는 자신의 현재 전송 전력으로 정의된다. 제안 기법의 공유 보상(R)은 성공적으로 도달한 패킷 수($|S_{su}|$)와 중계 참여 노드 수($|N_{rel}|$)를 고려하여 $\frac{|S_{su}|}{|N_{rel}|}$ 로 정의한다.

III. 시뮬레이션 결과 및 결론

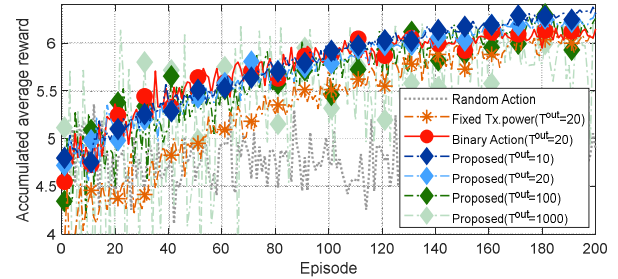


그림 2. Accumulated average reward vs. Episode

1km×1km 영역 내 포아송 분포에 따라 랜덤하게 분포된 18개의 노드가 하나의 제어된 연속 지역을 구성하는 시나리오를 고려한다. 또한, 전술 환경의 지형적 특성을 반영하고자 [3]의 $\Phi_{block} = \{\rho, l, \theta\}$ 의 시퀀스 형태로 이루어진 germ-grain blockage 모델을 고려한 네트워크 토폴로지를 구현하였다. 여기서 ρ, l, θ 는 각각 세그먼트의 중점, 길이 방향을 의미한다. 그림 2는 에피소드에 따른 각 기법의 누적 평균 보상을 나타낸다. 제안 기법은 초과 폭 파라미터와 전송 전력을 계층적으로 제어함으로써 계산 복잡도를 감소시키고 동시에 벤치마크 알고리즘 대비 최대 성능으로 수렴함을 확인할 수 있다.

ACKNOWLEDGMENT

이 논문은 2022년도 정부(과학기술정보통신부)의 재원으로 정보통신기획재정의 지원(No.2022-0-00704, 초고속 이동체 지원을 위한 3D-NET 핵심 기술 개발, 50%)과 정부(과학기술정보통신부)의 재원으로 한국연구재단의 지원을 받아 수행된 연구임(No. 2022R1A2C1010602, 50%).

참고 문헌

- [1] H. Seong, J. Kim, W. -Y. Shin and H. Lee, "FiFo: Fishbone Forwarding in Massive IoT Networks," in IEEE Internet of Things Journal, vol. 10, no. 5, pp. 4339-4352, Mar., 2023.
- [2] Ying Ding, Y. Wang and J. Yuan, "A blackhole attack analysis for Barrage Relay Networks," SNS & PCS, Nangang, 2013, pp. 13-17.
- [3] C. Saha and H. S. Dhillon, "Millimeter Wave Integrated Access and Backhaul in 5G: Performance Analysis and Design Insights," in IEEE Journal on Selected Areas in Communications, vol. 37, no. 12, pp. 2669-2684, Dec. 2019.