

# 비직교 다중접속 지원 다중 단말 대 단말 통신을 위한 심층 강화학습 기반 자원할당 기법

유서영, 이정우\*

중앙대학교

holywillow@naver.com, \*jwlee2@cau.ac.kr

## Resource Allocation Scheme Based on Deep Reinforcement Learning for NOMA-Enabled Device-to-Device Cluster Communications

Seoyoung Yu, Jeong Woo Lee\*

Chung-Ang Univ.

### 요약

본 논문에서는 비직교 다중접속 (Non-Orthogonal Multiple Access, NOMA) 프로토콜을 지원하는 다중 단말 대 단말 직접 통신(Device-to-Device, D2D) 환경에서의 네트워크 성능 최대화를 위한 심층강화학습 기반의 주파수 자원 및 송신전력 할당 기법을 제안하였다. 제안된 기법에서 각 D2D 클러스터는 학습된 인공신경망을 활용하여 셀룰러 링크와 D2D 링크의 평균 쓰루풋의 합을 최대화시킬 수 있는 D2D 수신단말 별 송신전력과 주파수 자원을 선택한다.

### I. 서론

최근 무선통신 네트워크의 서비스를 활용하는 단말의 수가 폭발적으로 증가하고 있고, 낮은 지연시간과 높은 데이터 전송률을 보장하는 통신서비스에 대한 요구 또한 증가함에 따라 한정된 주파수 자원의 효율적인 활용을 통하여 다수의 무선 단말에 고품질의 무선통신 서비스를 제공할 수 있는 기법들에 대한 많은 연구가 이루어지고 있다. D2D통신[1]은 기지국 간의 통신을 거치지 않고 인접한 두 단말 간의 직접통신을 수행하는 기법으로 네트워크의 부하를 경감시키면서 주파수 효율성을 높여줄 수 있는 기법 중 하나로 각광받고 있다. 한편, 주파수 효율성 증가, 송신 지연시간 경감 등을 목적으로 여러 개의 통신링크가 동일한 주파수 자원을 활용하는 NOMA[2] 또한 주목받고 있는 기법 중 하나이다. 이러한 D2D통신과 NOMA의 경우 여러 개의 통신 링크가 하나의 주파수 자원을 공유하는 특성으로 인해 통신 링크 간 상호간섭이 발생하게 되며, 이를 적절히 제어하여 주파수 효율성을 극대화시킬 수 있는 자원할당기법에 대한 연구가 활발히 진행되고 있다. 이에 더해 네트워크에 참여하는 단말 수의 증가, 단말의 높은 이동성에 의해 잦은 자원할당의 변경으로 인해 발생하는 높은 시스템 복잡도를 해결하기 위한 심층강화학습[3] 기반 자원할당기법들 또한 많은 연구가 이루어지고 있다.

본 논문에서는 NOMA 프로토콜을 지원하는 다중 단말 대 단말 D2D 통신 환경에서의 네트워크 효율성을 최대화하기 위한 심층강화학습 기반의 주파수 자원 및 D2D 송신전력 할당기법을 제안하였다. 각 D2D 클러스터는 학습된 인공신경망을 활용하여 셀룰러 링크와 D2D 링크의 평균 쓰루풋의 합을 최대화시킬 수 있는 NOMA 기반 D2D통신에 필요한 D2D 수신단말 별 송신전력과 주파수 자원을 선택한다. 모의실험을 진행한 결과 제안된 기법을 통해 네트워크 성능을 향상시킬 수 있음을 확인하였다.

### II. 본론

본 논문에서는 중심에 기지국이 위치한 원형 단일 셀에 각각 고유의 주

파수 자원을 할당받아 기지국과 통신을 수행하는  $M$ 개의 셀룰러 단말과 D2D 통신을 수행하는  $K$ 개의 D2D 클러스터가 존재하는 셀룰러 언더레이 네트워크 환경을 가정한다. 각 D2D 클러스터의 내부에는 한 개의 D2D 송신단말과  $N$ 개의 D2D 수신단말들이 존재하며, 셀룰러 단말이 점유하고 있는 주파수 자원 중 하나를 활용하여 NOMA 기반의 클러스터 단위 D2D 통신을 수행한다. D2D 송신단말에서는 여러 수신단말들에 보낼 신호를 중첩하여 전송하며, 각 D2D 수신단말은 순차적 간섭 제거(SIC) 기법을 활용하여 타 D2D 수신단말과 관련된 신호를 제거하고 자신에게 할당된 신호를 수신한다. 따라서  $n'$ 번째 D2D수신단말의 신호를 검출할 때의  $k$ 번째 D2D 클러스터의  $n$  ( $n' \leq n$ )번째 D2D 수신단말과  $m$ 번째 셀룰러 단말에서의 신호 대 잡음비(SINR)는 각 다음과 같이 정의된다.

$$\gamma_{n',n(k)}^D = \frac{P_{n'(k)} |h_{k,n(k)}|^2 d_{k,n(k)}^{-\alpha}}{I_n^{CU} + I_n^{RES} + I_n^{NOMA} + I_n^{D2D} + \sigma_w^2}, \quad (1)$$

$$\gamma_m^C = \frac{P_m^C |h_{m,B}|^2 d_{m,B}^{-\alpha}}{I_m^{D2D} + \sigma_w^2}. \quad (2)$$

이 때  $I_n^{CU}, I_n^{RES}, I_n^{NOMA}, I_n^{D2D}$  는 D2D 수신단말에서 측정되는 셀룰러 단말로부터의 간섭, 불완전한 SIC로 인해 발생하는 잔여 간섭, NOMA 중첩신호로 인해 발생하는 간섭, 다른 D2D 클러스터로부터의 간섭을 의미하며,  $I_m^{D2D}$  는 기지국에서 측정되는 D2D 클러스터로부터 발생하는 간섭을 의미한다. 식(1)과 (2)로부터 셀룰러 링크 및 D2D 링크의 아웃티지 확률을 각각  $\rho_{n(k)}^D = \Pr\{\gamma_{n,n(k)}^D \leq 2^{R^D} - 1\}$  와  $\rho_m^C = \Pr\{\gamma_m^C \leq 2^{R^C} - 1\}$  로 정의할 수 있으며, 이를 통해 네트워크 내 존재하는 셀룰러 링크와 D2D 링크의 평균 쓰루풋의 합  $\lambda$ 를 다음과 같이 정의할 수 있다.

$$\lambda = \frac{1}{M} \sum_{m=1}^M R^C (1 - \rho_m^C) + \frac{1}{KN} \sum_{k=1}^K \sum_{n=1}^N R^D \rho_n^D. \quad (3)$$

제안된 심층강화학습 기반 자원할당기법에서 각 D2D 클러스터는 에이전트로 동작하여 셀룰러 링크와 D2D 링크의 평균 쓰루풋의 합을 최대화 할 수 있는 동작을 학습한다. 네트워크 에이전트가 관찰하는 상태  $s^k$  는 네트워크에 존재하는 단말들의 위치정보와 자원할당 정보 및 D2D 수신단말들에서의 주파수 자원별 간섭량으로 정의되며, 각 D2D클러스터는 에이전트로 동작하여 각 클러스터의 NOMA 지원 D2D 통신에 활용할 셀룰러 주파수 자원과 송신전력을 선택한다. 선택한 행동에 따른 보상  $r^k$  는 D2D 클러스터들의 자원할당으로부터 얻어지는  $\lambda$  를 활용한다.

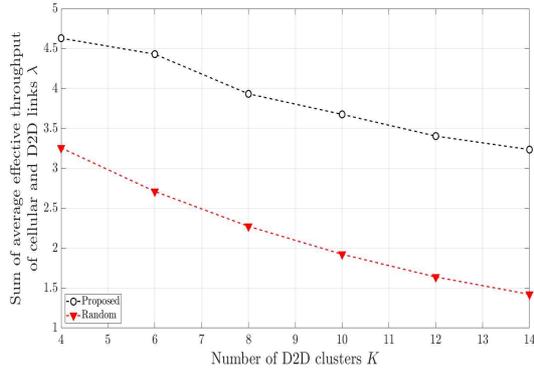


그림1. 네트워크 내 D2D 클러스터의 참여 개수 변화에 따른 제안된 기법과 무작위 자원할당기법의 성능변화 추이

### III. 모의실험 결과 및 결론

그림1과 같이  $R^C=4$ ,  $R^D=2$ ,  $N=2$ , 셀 반경이 100[m] 인 셀룰러 언더레이 네트워크 환경에서 네트워크에 존재하는 D2D 클러스터의 수  $K$ 를 증가시켜가며 모의실험을 진행한 결과, 제안된 강화학습 기반 자원할당 기법이 무작위 자원할당 기법에 비해 성능적 우위를 지님을 확인할 수 있었으며 기법간의 성능차는 네트워크에 참여하는 D2D 클러스터의 개수가 증가할수록 더욱 증가함을 관찰 할 수 있었다.

### ACKNOWLEDGMENT

“본 연구는 과학기술정보통신부 및 정보통신위원회의 대학 ICT 연구센터육성 지원사업의 연구결과로 수행되었음”(IITP-2024-RS-2022-00156353).

### 참 고 문 헌

- [1] K. Doppler, M. Rinne, C. Wijting, C. B. Ribeiro and K. Hugl, “Device-to-device communication as an underlay to LTE-advanced networks,” in IEEE Communications Magazine, vol. 47, no. 12, pp. 42-49, Dec. 2009.
- [2] Y. Saito, Y. Kishiyama, A. Benjebbour, T. Nakamura, A. Li, and K. Higuchi, “Non-orthogonal multiple access (NOMA) for cellular future radio access,” in Proc. IEEE 77th Veh. Technol. Conf., Jun. 2013, pp. 1 - 5.
- [3] V. Mnih et al., “Human-level control through deep reinforcement learning,” Nature vol. 518, no. 7540, pp. 529 - 533, Feb. 2015.