

# Unlicensed Band Traffic and Fairness Maximization Approach Based on Rate-Splitting Multiple Access

Jeon Zang Woo<sup>†</sup> · Kim Sung Wook<sup>††</sup>

## ABSTRACT

As the spectrum shortage problem has accelerated by the emergence of various services, New Radio-Unlicensed (NR-U) has appeared, allowing users who communicated in licensed bands to communicate in unlicensed bands. However, NR-U network users reduce the performance of Wi-Fi network users who communicate in the same unlicensed band. In this paper, we aim to simultaneously maximize the fairness and throughput of the unlicensed band, where the NR-U network users and the Wi-Fi network users coexist. First, we propose an optimal power allocation scheme based on Monte Carlo Policy Gradient of reinforcement learning to maximize the sum of rates of NR-U networks utilizing rate-splitting multiple access in unlicensed bands. Then, we propose a channel occupancy time division algorithm based on sequential Raiffa bargaining solution of game theory that can simultaneously maximize system throughput and fairness for the coexistence of NR-U and Wi-Fi networks in the same unlicensed band. Simulation results show that the rate splitting multiple access shows better performance than the conventional multiple access technology by comparing the sum-rate when the result value is finally converged under the same transmission power. In addition, we compare the data transfer amount and fairness of NR-U network users, Wi-Fi network users, and total system, and prove that the channel occupancy time division algorithm based on sequential Raiffa bargaining solution of this paper satisfies throughput and fairness at the same time than other algorithms.

Keywords : 5G Network, Rate Splitting Multiple Access, Unlicensed Band, Game Theory, Bargaining Solution, Reinforcement Learning

## 전송률 분할 다중 접속 기술을 활용한 비면허 대역의 트래픽과 공정성 최대화 기법

전 장 우<sup>†</sup> · 김 승 옥<sup>††</sup>

### 요 약

다양한 서비스가 등장으로 인해 스펙트럼 부족 문제가 가속화됨에 따라, 면허 대역에서 통신하던 사용자들을 비면허 대역에서 통신하는 NR-U(New Radio-Unlicensed)가 등장하였다. 하지만 NR-U 네트워크 사용자로 인해 동일한 비면허 대역에서 통신하는 Wi-Fi 네트워크 사용자의 성능이 감소하게 된다. 본 논문에서는 NR-U 네트워크 사용자와 Wi-Fi 네트워크 사용자가 공존해있는 비면허 대역의 처리량과 비면허 대역의 사용에 대한 공정성을 동시에 최대화하는 것을 목표로 한다. 먼저 비면허 대역에서 전송률 분할 다중 접속 기술을 활용한 NR-U 네트워크의 합-전송 속도(Sum of Rate)를 최대화하기 위해 강화 학습의 몬테 카를로 정책 하강법(Monte Carlo Policy Gradient)을 활용한 최적의 전력 할당 기법을 제안하였다. 그 뒤, 동일한 비면허 대역에서 NR-U 네트워크와 Wi-Fi 네트워크의 공존을 위해 시스템 처리량과 공정성을 동시에 최대화할 수 있는 게임 이론의 순차적 라이파 협상 해법(Sequential Raiffa Bargaining Solution)을 활용한 채널 점유 시간 분할 알고리즘을 제안하였다. 시뮬레이션 결과에서 동일한 전력 할당 기법을 사용하였을 때, 본 논문에서 제안한 전송률 분할 다중 접속 기술이 기존의 다중 접속 기술들보다 더 빠른 합-전송속도를 보임을 확인하였다. 또한 비면허 대역 네트워크의 전송량과 공정성을 비교해본 결과 본 논문의 순차적 라이파 협상 해법을 활용한 채널 점유 시간 분할 알고리즘이 타 알고리즘보다 처리량과 공정성을 동시에 만족함을 입증하였다.

키워드 : 5G 네트워크, 전송률 분할 다중 접속, 비면허 대역, 게임이론, 협상 해법, 강화 학습

※ 본 연구는 과학기술정보통신부 및 정보통신기획평가원의 대학ICT연구센터 지원 사업(IITP-2023-2018-0-01799)과 2023년도 정부(과학기술정보통신부)의 재원으로 한국연구재단 기본연구 사업의 지원을 받아 수행된 연구임(No.2021 R1F1A1045472).

† 비 회 원 : 서강대학교 컴퓨터공학과 석사

†† 종 신 회 원 : 서강대학교 컴퓨터공학과 교수

Manuscript Received : March 2, 2023

First Revision : May 30, 2023

Accepted : May 31, 2023

\* Corresponding Author : Kim Sung Wook([swkim01@sogang.ac.kr](mailto:swkim01@sogang.ac.kr))

## 1. 서 론

이동통신 기술이 진화함에 따라 다양한 혁신 서비스에 대한 수용도가 높아지면서 모바일 기기의 수와 트래픽이 증가함에 따라 주파수 수요가 크게 증가하게 되었다. 시스코(Cisco)의 보고서에 의하면, 2023년에 전 세계 인터넷 사용자 수는 53억 명, 모바일 장치는 293억 개까지 증가할 것이라 전망하

고 있다. 또한 스마트폰으로 점점 더 고화질의 영상을 시청하고, 가전 제품들이 사물인터넷(IoT : Internet of things)으로 변모하면서 트래픽 수요를 증가시키고 있다. 고령화가 진행됨에 따라 노동인구의 감소로 인해 산업 현장에서도 산업 사물인터넷(IIoT: Industrial Internet of Things)을 사용하고, 자율주행차, AI 콜센터, 스마트팩토리, 배달 로봇이나 서빙 로봇 등의 기기로 인해 주파수 수요가 기하급수적으로 증가하는 추세이다. 또한 저출산이 지속됨에 따라 자동화를 향한 움직임은 계속 확장될 것이고 트래픽과 주파수 수요를 증가시키며 스펙트럼 부족 문제를 가속화할 것이다[1].

모바일 데이터 수요가 급증하면서 면허 대역(Licensed band)로는 모든 수요를 충족시킬 수 없게 되었다. 이를 해결하고자, 3GPP는 3GPP Release 13에서 비면허 대역을 활용하여 LTE 시스템을 동작시키는 LTE-Licensed Assisted Access (LAA) 기술을 표준화하였다. 또한 2020년 7월에 3GPP는 Release 16에서 비면허대역에서 동작하는 5G NR-U(New Radio-Unlicensed) 기술을 표준화하였다. 그러나 동일한 비면허 스펙트럼에서 운영되는 NR-U 사용자로 인해 기존에 비면허 대역에서 통신을 하던 WiFi 네트워크 사용자들의 성능은 감소하게 된다. 이 문제를 해결하기 위해 NR-U 네트워크 사용자와 Wi-Fi 네트워크 사용자들 간의 공정한 공존을 보장하면서 시스템 처리량을 최대화하는 메커니즘이 필요하다[2].

또한 비면허 대역에서 통신하는 다수의 사용자들의 신호를 송수신하기 위해 NR-U 기지국의 한정된 전송 전력과 비면허 대역폭 하에서 NR-U 사용자들의 데이터 전송 속도를 최대화할 필요가 있다. 기지국의 한정된 자원 하에서 여러 사용자들의 신호를 송수신하기 위한 여러 다중 접속 기술이 활용되고 있다. 그중 공간 분할 다중 접속(SDMA: Space Division Multiple Access)기술과 비 직교 다중 접속(NOMA: Non-Orthogonal Multiple Access) 기술을 결합한 전송률 분할 다중 접속 기술이 미래 6G 네트워크의 핵심 기술로 여겨지고 있다. 전송률 분할 다중 접속 기술은 공간 분할 다중 접속 기술과 비직교 다중 접속 기술을 복합시켜 만든 기술로 신호를 공통 메시지와 개인 메시지로 분할하여 동시에 신호를 전송하는 다중 접속 기술이다. 하지만 불완전한 채널을 가진 네트워크 환경에서 전송률 분할 다중 접속의 합-전송 속도 최대화 문제는 여전히 도전 과제에 있다[3,4].

본 논문에서는 협조 게임 이론의 순차적 라이파 협상 해법 사용하여 NR-U 네트워크와 Wi-Fi 네트워크가 공존하는 비면허 대역의 채널 점유 시간을 공정하면서 사용자들의 데이터 전송량을 최대화할 수 있도록 분할한다. 또한 강화 학습의 몬테 카를로 정책 하강법을 사용하여 NR-U 네트워크에서 전송률 분할 다중 접속 기술을 활용한 NR-U 사용자들의 시스템 처리량을 최대화할 수 있는 최적의 전력 할당 기법을 제안한다. 본 논문이 기여하는 바는 다음과 같다.

1. 강화 학습의 대표적인 정책 기반 알고리즘인 몬테 카를로 정책 하강법을 사용하여 전송률 분할 다중 접속으로 메시지를 전달하는 NR-U 네트워크 사용자들의 합-전송 속도를 최대화할 수 있는 최적의 전력 할당 계수를 구한다.
2. 게임 이론의 순차적 라이파 협상 해법을 사용하여 NR-U 네트워크와 Wi-Fi 네트워크 간의 비면허 대역 채널 점유 시간에 대한 효용의 최적의 협상 점을 결정하여 각 네트워크의 공정성을 보장하면서 시스템의 전체 효용을 높인다.
3. 시뮬레이션 결과는 제안한 기법의 우수성을 입증한다. 기지국의 전송 전력과 대역폭이 동일한 상황에서 전송률 분할 다중 접속과 다른 다중 접속 기술들의 합-전송 속도를 비교하여 제안한 다중 접속 기술이 더 빠른 합-전송 속도를 보임을 입증한다. 그리고 순차적 라이파 협상 해법을 사용하여 비면허 대역 채널 점유 시간을 분할하였을 때의 시스템 평균 처리량과 공정성을 동시에 만족함을 증명한다.

본 논문의 구성은 다음과 같다. 2장에서는 전송률 분할 다중 접속과 순차적 라이파 협상 해법, 몬테 카를로 정책 하강법에 대한 연구를 보여주고, 3장에서는 본 논문에서 고려하는 네트워크 모델과 제안하는 기법에 대해서 설명한다. 4장에서는 제안하는 다중 접속 기술과 다른 다중 접속 기술들의 합-전송 속도를 비교하고 순차적 라이파 협상 해법을 사용했을 때의 비면허 대역의 성능 평가를 통해 본 논문에서 제안하는 기법들의 우수성을 입증한다. 5장에서는 결론으로 마무리한다.

## 2. 관련 연구

### 2.1 전송률 분할 다중 접속 기술

다중 접속 기술이란 무선 자원이 한정되어 있을 때, 증가하는 스펙트럼 요구를 감당해 내기 위해 여러 사용자들의 신호를 결합하는 메커니즘이다. 다중 접속 기술이 발전하면서 캐패시티가 향상되고 무선 시스템의 처리량이 폭발적으로 증가하였다.

기존의 다중 접속 기술들과는 달리 RSMA는 메시지를 공유 메시지와 개인 메시지로 분리하여 통신한다. 각 사용자의 공유 메시지는 하나의 공통 메시지로 결합시키고 모든 사용자가 공유하는 코드북을 사용하여 공유 스트림으로 부호화하고 모든 사용자가 복호화 할 수 있다. 반면에, 개인 메시지는 각 메시지에 해당하는 사용자만이 가지는 코드북을 통해 개인 스트림으로 부호화하고 해당 사용자만이 복호화 할 수 있다. 송신단에서 중첩 코딩을 통해서 유저의 신호들을 중첩해서 전송한

다. 수신단에서는 먼저 개인 스트림을 잡음으로 취급하고 공통 스트림을 복호화하고 공유 메시지 중에서 자신의 공유 메시지를 분리한다. 이때, 사용자는 다른 사용자의 개인 스트림으로 인한 간섭 만을 가지게 되며 이를 잡음으로 처리한다. 마지막으로 해당 사용자의 공유 메시지와 개인 메시지를 결합하여 원래의 메시지를 얻는다[5].

2.2 순차적 라이파 협상 해법

순차적 라이파 협상 해법이란 반복적인 중간 합의를 통해 최종적인 협상 해를 구해내는 협상 게임이다. 2명의 플레이어가 있을 때 순차적 라이파 협상 게임은 다음과 같이 정의된다[6].

$$\begin{aligned} f^{(s,d)}(a_1, a_2) &:= (f_1^{(s,d)}(a_2), f_2^{(s,d)}(a_1)) \\ f_1^{(s,d)}(a_2) &:= \max_{a_1 \in \mathbf{R}} (a_1 - d_1) \\ f_2^{(s,d)}(a_1) &:= \max_{a_2 \in \mathbf{R}} (a_2 - d_2) \end{aligned} \quad (1)$$

여기서  $a_i$ 는 플레이어  $i$ 에게 분배된 자원,  $d_i$ 는 협상이 결렬 되었을 때 플레이어  $i$ 의 효용 함수,  $f_1^{(s,d)}(a_2)$ 는 플레이어 1이 받는 효용을 최대화한 값이고,  $f_2^{(s,d)}(a_1)$ 는 플레이어 2가 받는 효용을 최대화한 값이다. 순차적 라이파 협상 해법에서 협상 결렬 지점  $d$ 가 주어졌을 때, 플레이어가 가장 선호하는 값은 다른 플레이어의 효용을 협상 결렬 지점에 계속 머무르게 하면서 자신의 효용이 최대의 값을 가지게 하는 결과이다 [6].

순차적 라이파 협상 해법으로 나온 협상해는 다음의 7개의 공리를 만족한다 [6].

공리1. 파레토 최적성(Pareto Optimality) : 협상 해는 협상 가능 집합의 경계 집합(boundary set)에 존재해야 한다. 즉, 협상 해는 배분할 수 있는 자원을 모두 사용하여 두 플레이어의 효용을 최대화하는 집합 안에 존재해야 한다.

공리2. 긍정적인 아핀 변환에 대한 공분산(Covariance w. r. t. positive affine transformations) : 같은 협상 단계에서 협상 결렬점과 협상 가능 집합에 대해서 양의 아핀 변환을 적용해도 협상 결과는 변하지 않는다. 이 공리를 통해 협상의 합리성과 일관성을 보장하고, 다양한 협상 상황에서의 결과를 일관되게 비교하게 된다.

공리3. 대칭성(Symmetry) : 게임에 참여한 모든 플레이어는 대등하다라는 가정하에 플레이어의 위치가 달라 지더라도 그 해는 변하지 않아야 한다.

공리4. 관련 없는 대안의 독립성(Independence of Irrele-

vant Alternatives) : 선택 가능한 대안들이 사라져도 기존의 협상해를 포함하고 있다면 새로운 협상 문제의 해는 여전히 기존의 협상해와 동일해야 한다.

공리5. 표준 공리의 반복 적용(Repeated Application of the Standard Axioms) : 협상 주체들은 순차적으로 제안과 대응을 주고 받으면서 여러 차례 협상을 반복할 수 있다. 이를 통해 협상 주체들은 협상 과정에서 타협점을 찾아가며 합리적인 결정을 내릴 수 있다.

공리6. 라이파 과정에 의한 협상 결렬점의 변화 하의 불변성(Invariance under Change of Disagreement Points by the Raiffa Procedure) : 협상 진행 중 얻는 협상 가능 집합과 협상 결렬점들의 최종 협상 해는 변하지 않는다.

공리7. 받침 초평면 근사화(Supporting Hyperplane Approximation) : 협상 주체들은 초평면의 경계면을 이용하여 최적의 합의점을 근사하는 과정을 통해 최종적인 협상해를 구해낸다. 이를 통해, 협상 주체들은 자신의 이익을 극대화하면서도 협상 과정을 진행할 수 있게 된다.

순차적 라이파 협상 해법은 플레이어들이 단계적으로 중간 합의를 거쳐 최종 합의를 이루어내는 해법이다. 여기서 중간 합의란 2명의 플레이어가 가장 선호하는 점의 평균을 나타낸다. 중간 합의점(Midpoint)은 다음과 같이 정의된다[7].

$$m_0^{(s,d)} := (m_{0,1}^{(s,d)}, m_{0,2}^{(s,d)}) := d \quad (2)$$

$$\forall n \in \mathbf{N} : m_n^{(s,d)} = \frac{1}{2} [(f_1^{(S,d)}(m_{n-1,2}^{(s,d)}), m_{n-1,2}^{(s,d)}) + (m_{n-1,1}^{(s,d)}, f_2^{(S,d)}(m_{n-1,1}^{(s,d)}))]$$

중간 합의를  $n$  번 반복했을 때, 최종적으로 수렴한 최종 합의점이 바로 순차적 라이파 협상 해  $R(S, d)$ 가 된다 [7].

$$R(S, d) := \lim_{n \in \mathbf{N}_0} m_n^{(S,d)} \quad (3)$$

2.3 몬테 카를로 정책 하강법

몬테 카를로 정책 하강법은 목적 함수를 최대화하는 행동을 취할 확률을 최대화하여 최적의 정책을 찾는 강화 학습 기법이다. 정책 하강법은 가치 함수 대신 정책 파라미터  $\theta$ 를 사용한다. 여기서 정책 파라미터  $\theta$ 는 함수 근사화(function

approximation)에서 행동 가치 함수 대신 정책 자체를 근사화한 함수를 나타낸다. 정책 경사도 방법에서 정책은 확률적 정책(Stochastic Policy)을 사용한다. 즉, 현재 상태의 행동은 확률 분포를 기반으로 선택된다. 시간  $t$ 에 상태  $s$ 에서 정책 파라미터  $\theta$ 가 주어졌을 때 행동  $a$ 를 취할 확률은 다음과 같이 주어진다[8].

$$\pi(a|s, \theta) = \Pr\{A_t = a | S_t = s, \theta_t = \theta\} \quad (4)$$

그리고 에이전트가 확률적으로 행동을 수행하면서 얻은 보상(목적 함수)들을 저장해서, 얻은 보상이 크면 해당 행동을 취할 확률을 높이고 반대로 얻은 보상이 낮으면 취할 확률을 낮춰서 최종적으로 큰 보상들을 얻는 행동만을 행동해서 최적의 정책을 학습한다. 실제 얻은 목적 함수에 따라 근사화한 정책(특정 상태에서 행할 행동의 확률 분포)을 업데이트 한다. 즉, 정책 파라미터의 기울기를 업데이트하면서 정책을 향상시킨다. 이를 수식으로 나타내면 다음과 같다[8].

$$\theta = \theta + \alpha \nabla_{\theta} J(\theta) \quad (5)$$

여기서  $\alpha$ 는 학습률,  $J(\theta)$ 는 목적 함수(Objective Function)로 정책 파라미터  $\theta$ 가 주어진 정책  $\pi$ 를 기반으로한 trajectory  $\tau$ 에서 샘플링 했을 때의 리턴(보상의 합)의 기대값을 의미한다. 목적 함수와 목적 함수를 편미분한 값은 다음과 같이 주어진다[8].

$$J(\theta) = E_{\tau \sim \pi_{\theta}(\tau)} \left[ \sum_{t=0}^{T-1} \nabla_{\theta} \log \pi_{\theta}(a_t | s_t) G(\tau) \right] \quad (6)$$

여기에서  $G(\tau)$ 는 리턴,  $\log \pi_{\theta}(a_t | s_t)$ 는 시간  $t$ 에 상태  $s$ 에서 행동  $a$ 를 취할 로그 확률을 의미한다. 즉,  $\nabla_{\theta} \log \pi_{\theta}(a_t | s_t)$ 는  $a_t$ 를 취할 확률의 기울기를 나타낸다. 그리고 여기에 리턴  $G(\tau)$ 을 곱해줌으로써 어떤 방향으로 얼마만큼 정책을 업데이트해줘야 하는지를 결정하게 된다. 즉, 큰 리턴 값을 얻을수록 기울기를 크게 업데이트한다[8].

### 3. 제안된 기법

#### 3.1 네트워크 구조

본 논문에서 제안하는 네트워크 구조는 Fig. 1에 도시된 바와 같이 비면허 대역에서 동작하는 1개의 NR-U BS과 Wi-Fi,  $k_n$  명의 NR-U 사용자와  $k_w$  명의 Wi-Fi 사용자로 구성되어 있고 BS과 Wi-Fi에서 사용자들에게 데이터를 전송하는 하향링크만을 고려한다. NR-U BS와 WiFi AP는 백본 네트워크로 연결되어 서로의 전송해야 하는 트래픽과 전송 속도를 알고 있다고

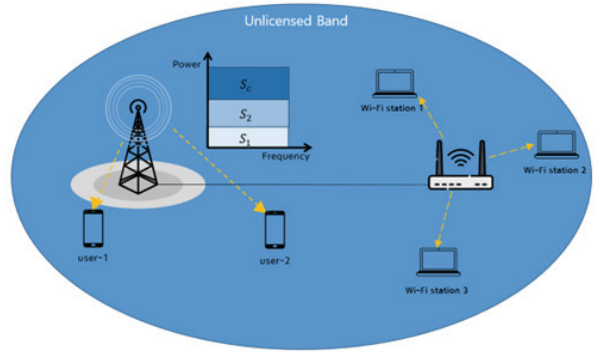


Fig. 1. Network Structure

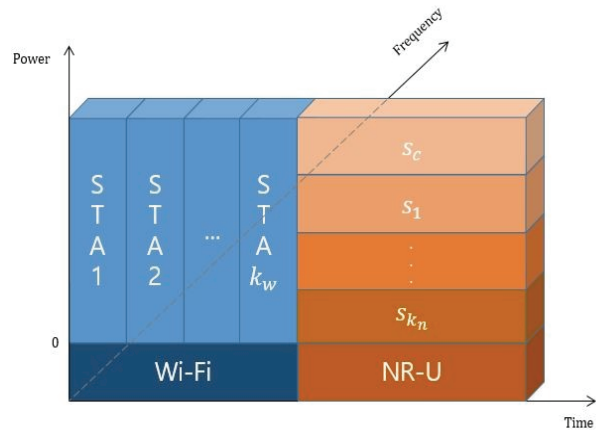


Fig. 2. Structure of Time and Frequency and Power Resources in Unlicensed Band

가정한다. 각 네트워크의 전송 속도와 전송해야 할 트래픽을 고려한 채널 점유 시간은 상대적으로 더 큰 계산 능력을 가진 NR-U BS에서 계산한다. Fig. 2에 도시된 바와 같이 NR-U 네트워크에서 NR-U 사용자들은 해당 대역의 대역폭  $B$ 를 사용하고, NR-U 채널 점유 시간 전체를 공유해서 통신한다. 이때, NR-U 사용자들의 스트림은 전송률 분할 다중 접속 기반의 하향링크 전송을 지원하기 위해  $k$ 개의 스트림으로 분할된다. 반면에, Wi-Fi 네트워크의 사용자들은 동일한 전력을 할당받고 전체 Wi-Fi 채널 점유 시간을 동일한 비율로 분할하여 통신한다.

본 장에서는 이전 장에서 소개한 몬테 카를로 정책 하강법을 활용해 공통 메시지와 NR-U 사용자들의 개인 메시지에 대한 최적의 전력 할당을 통해 전송 속도를 최대화한다. 그 뒤, 순차적 라이카 협상 해법을 활용해 비면허 대역의 NR-U 네트워크와 Wi-Fi 네트워크의 공평성과 데이터 전송량을 최대화한다. 즉, 본 논문에서는 전체 시스템의 데이터 전송량과 공평성을 동시에 최대화하는 것을 목표로 한다. 이를 수식으로 나타내면 다음과 같다.

$$\max_{t_w, t_n, \mu_c, \mu_k} S_{tot} * F(S_k^n, S_k^w) \quad (7)$$

$$S_{tot} = S_k^m + S_k^w$$

$$F(S_k^m, S_k^w) = \frac{\left(\sum_{i=1}^n S_i\right)^2}{n \sum_{i=1}^n S_i^2}$$

여기서  $S_{tot}$ 는 전체 시스템 전송량,  $F(S_k^m, S_k^w)$ 는 데이터 전송량에 대한 공정성을 의미하고  $S_k^m$ ,  $S_k^w$ 는 각각 NR-U 네트워크와 WiFi 네트워크의 처리량을 의미한다. 는 각 네트워크의 전송량을 의미한다.

비면허 대역에서 사용자들의 처리량의 대한 효용 함수를 직관적으로 파악하기 위해, Equation (8)의 데이터 전송량에 대한 효용 함수를 로그 함수의 형태로 나타낸다. 이를 수식으로 나타내면 다음과 같다[9].

$$S_{tot} = E\left(\alpha \log S_k^m + (1-\alpha) \log S_k^w\right)$$

$$= E\left(\alpha \log R_{sum} t_n + (1-\alpha) \sum_{j=1}^{k_w} \log R_j^w t_{sta}\right) \quad (8)$$

여기서  $E(\cdot)$ 는 기대값을 의미하고,  $\alpha$ 는 NR-U 사용자와 Wi-Fi 사용자가 만족스러운 자원 분배 결과를 얻도록 조정될 수 있는 가중치, 그리고  $R_{sum}$ ,  $t_n$ 는 NR-U 네트워크의 합 전송 속도와 채널 점유 시간,  $R_j^w$ ,  $t_{sta}$ 는 WiFi 네트워크의 전송 속도와 채널 점유 시간을 의미한다. 제한 조건의  $\mu_c$ ,  $\mu_k$ 는 전송률 분할 다중 접속에서 공통 메시지와 개인 메시지의 전력 할당 계수를 의미한다.

### 3.2 RSMA 전력 할당 기법

RSMA는 주파수를 분할하는 것이 아닌 전력 할당을 통해 여러 사용자와 통신하는 다중 접속 기술이다. RSMA에서 전력 할당을 하는 방법은 Fig. 3과 같다. NR-U 네트워크에서 각 사용자 단말기의 메시지  $W_k$  ( $k=1,2$ )는 공통 메시지  $W_k^c$ 와 개인 메시지  $W_k^p$ 로 각각 분할된다. 이후 각 단말기의 공통 메시지  $W_k^c$ 는 하나의 공통 메시지  $W^c$ 로 합쳐진다.

결합된 공통 메시지  $W^c$ 와 개인 메시지  $W_k^p$ 는 각각 스트림  $s_c, s_1, s_2$ 로 인코딩되어 전송된다. 기지국의 송신 신호는 다음과 같이 정의된다 [10].

$$\mathbf{x} = \sqrt{\mu_c P_t} \mathbf{f}_c s_c + \sqrt{\mu_k P_t} \mathbf{f}_k s_k \quad (9)$$

여기서  $\mathbf{f}_c \in \mathbb{C}^{M \times 1}, \mathbf{f}_k \in \mathbb{C}^{M \times 1}, k=1,2$ 는 프리코딩 벡터,  $P_t$ 는 총 전송 전력,  $\mu_c$ 와  $\mu_k$ 는 전력 할당 계수를 의미한다. 여기서  $\mu_c + \mu_k = 1$ 이 된다.

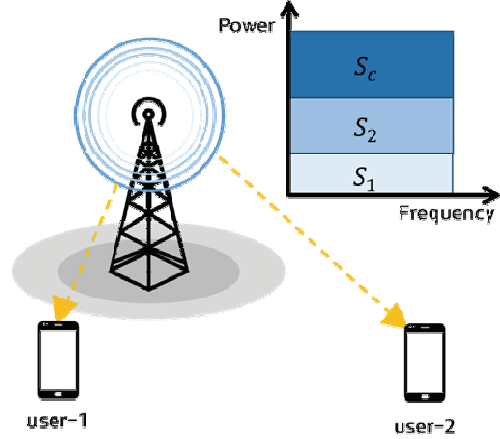


Fig. 3. RSMA Network

단말기가 받는 수신 신호는 다음과 같이 정의된다 [10].

$$\mathbf{y} = \mathbf{h}_k^H \mathbf{x} + n_k \quad (10)$$

여기서  $\mathbf{h}_k^H \in \mathbb{C}^{M \times 1}$ 는 기지국과 단말기  $k$ 간의 채널 이득,  $n_k$ 는  $\mathcal{N}(0, \sigma^2)$ 인 AWGN을 의미한다.

Fig. 3에서 기지국이 단말기 user-1과 user-2에게 RSMA로 통신하기 위해서는 각 단말기와의 채널 상태를 확인하여야 한다. 채널에 다른 간섭이 없다고 가정한다면 user-1이 user-2보다 가까이 있으므로 user-1의 채널 상태가 user-2보다 더 좋은 상태일 것이다. 그러므로 더 먼 거리의 단말기에게 신호를 전달할 확률을 증가시키기 위해서 더 높은 송신 전력을 할당해 줘야 한다. 또한 RSMA에서 사용자들은 공통 스트림을 먼저 복호화 한 뒤, 자신의 개인 스트림을 복호화 하기 때문에 공통 스트림에 더 큰 전력을 할당해 줘야 한다.

공통 스트림과 개인 스트림의 SINR은 다음과 같이 계산된다[10].

$$\text{SINR}_k^c = \frac{\mu_c P_t \left| \mathbf{h}_k^H \mathbf{f}_c \right|^2}{\sum_{j=1}^2 \mu_j P_t \left| \mathbf{h}_k^H \mathbf{f}_j \right|^2 + \sigma^2} \quad (11)$$

$$\text{SINR}_k^p = \frac{\mu_k P_t \left| \mathbf{h}_k^H \mathbf{f}_k \right|^2}{\sum_{j \neq k} \mu_j P_t \left| \mathbf{h}_k^H \mathbf{f}_j \right|^2 + \sigma^2}$$

수신단에서는 개인 스트림을 잡음으로 취급하고 공통 스트림을 복호화하고 SIC를 통해 수신 신호에서 공유 스트림을 제거하고 개인 스트림을 추출한다. 개인 스트림은 다른 단말기의 개인 스트림으로 인한 간섭만을 가지게 된다.

공통 스트림의 전송 속도와 개인 스트림 전송 속도는 다음과 같이 계산된다[10].

$$R_c = B \times \log_2 \left( 1 + \min_{k \in \{1,2\}} \text{SINR}_k^c \right) \quad (12)$$

$$R_k = B \times \log_2 \left( 1 + \text{SINR}_k^p \right), \forall k \in \{1,2\}$$

여기서  $B$ 는 대역폭을 의미한다. 공통 스트림의 모든 사용자들이 공유 스트림을 성공적으로 복호화할 수 있도록 하기 위해서 가장 SINR이 좋지 않은 사용자를 기준으로  $R_c$ 가 정해진다.

최종적인 합 전송 속도는 공통 스트림의 전송 속도와 개인 스트림의 전송 속도의 합으로 이루어지게 된다 [10].

$$R_{sum} = R_c + \sum_{i=1}^2 R_k \quad (13)$$

즉, 특정 스트림이 다른 스트림의 간섭으로 작용하는 환경에서 합 전송 속도를 최대화하기 위해서는 각 스트림의 적절한 전력을 할당해 줘야 한다. 본 논문에서는 비면허 대역의 NR-U 사용자들의 합 전송 속도를 최대화하기 위해 몬테 카를로 정책 하강법을 활용한 최적의 전력 할당 기법을 제안한다.

상태를 각 스트림의 전력 계수로 설정하고, 행동은 각 스트림에 할당되는 전력 계수로 설정한다. 그리고 해당 전력 계수 일 때의 합 전송 속도를 보상으로 설정한다. 현재 각 스트림의 전력 계수에 행동을 통해 얻은 새로운 전력 계수를 더해주고 그 때의 합 전송 속도를 계산한다. 그 뒤, 그 때의 보상(합 전송 속도)를 기존의 리턴에 더해주고 이를 통해 목적 함수의 기울기  $\nabla_{\theta} J(\theta)$ 를 계산한다. 그리고 계산된  $\nabla_{\theta} J(\theta)$ 를 통해 파라미터  $\theta$ 를 업데이트한다. 해당 과정을 반복해서  $\theta$ 가 최종 수렴하게 되면, 그 때 나오는  $R_{sum}$ 이 최대의 합 전송 속도  $R_{sum}^*$ 가 된다.

해당 과정을 의사 코드로 나타내면 Algorithm 1과 같다.

### 3.3 비면허 대역 채널 점유 시간 분할 기법

전체 채널 점유 시간  $T$ 는 Wi-Fi 사용자들이 점유하는 시간  $t_w$ 와 NR-U 사용자가 점유하는 시간  $t_n$ 만을 고려하며 충돌과 유희 시간은 고려하지 않는다. Wi-Fi 네트워크에서 단 한 명의 Wi-Fi 사용자만이 전체 스펙트럼 대역을 사용한다. 그리고 모든 Wi-Fi 사용자들은 동일한 경합 방식을 따르기 때문에 각 사용자들은 동일한 시간 점유율을 가지게 된다. 그러므로, Wi-Fi 사용자들의 트래픽은 다음과 같이 주어진다 [9].

$$S_k^w = \sum_{j=1}^{k_w} R_j^w t_{sta}^w \quad (14)$$

$$t_{sta}^w = \frac{t_w}{k_w}$$

#### Algorithm 1 Reinforce를 활용한 RSMA 전력 할당 알고리즘

```

Initialize:
State  $S = [\mu_c, \mu_1, \mu_2, \dots, \mu_k] = [0, 0, 0, \dots, 0]$ 
Action
 $A = [a_c, a_1, a_2, \dots, a_k] = [\nabla \mu_c, \nabla \mu_1, \nabla \mu_2, \dots, \nabla \mu_k]$ 
Reward  $R = R_{sum}$ 
Return  $G = 0$ 
Discount factor  $\gamma = 0.98$ 
Learning rate  $\alpha = 0.0002$ 
Iteration = 500000
Do
  For i = 1 to Iteration Do
    State ← State + Action
    Calculate R
     $G = \gamma G + R$ 
     $J(\theta) = \frac{1}{N} \sum_{i=1}^N \left[ \sum_{t=0}^{T-1} \nabla_{\theta} \log \pi_{\theta}(a_t | s_t) G(\tau) \right]$ 
     $\theta = \theta + \alpha \nabla_{\theta} J(\theta)$ 
  End For
End
Output converge Reward =  $R_{sum}^*$ 
    
```

여기서  $R_j^w$ 는 Wi-Fi 사용자의 전송 속도,  $t_{sta}$ 는 각 Wi-Fi 사용자의 채널 점유 시간을 나타낸다.

반면에, NR-U 네트워크에서는 모든 NR-U 사용자들이 중첩 코딩을 통해 전체 스펙트럼 대역을 함께 사용하며, 동시에 데이터들을 전송한다. 그러므로, NR-U 사용자들의 처리량은 전에 구한 합-전송 속도와 NR-U 네트워크의 채널 점유 시간의 곱으로 나타낸다.

$$S_k^N = R_{sum} t_n \quad (15)$$

채널 점유 시간 분할 알고리즘의 참여자는 NR-U 네트워크 사용자들과 Wi-Fi 네트워크 사용자들로 나누어진다. 네트워크 참여자들이 점유할 수 있는 총 시간은  $T$  만큼 주어진다. NR-U 네트워크 사용자들이 점유하는 시간은  $t_n$ , Wi-Fi 네트워크 사용자들이 점유하는 시간은  $t_w$ 로 나타내지며, 각 사용자들이 점유하는 시간을 모두 합하면 총 시간  $T$ 가 된다 ( $T = t_w + t_n$ ). 두 그룹 간에 협상이 결렬되었을 때의 결렬 점은  $d = (d_1, d_2)$ 로 나타내고,  $d_1$ 은 협상이 결렬되었을 때의 NR-U 사용자의 점유 시간,  $d_2$ 는 협상이 결렬되었을 때의 Wi-Fi 사용자의 점유 시간을 나타낸다. 각 네트워크 사용자들

은 협상 결렬 점에서 상대방의 점유 시간을 고정시키고 자신의 점유 시간을 최대화하려 한다.

$$t_n = \left( \max_{x_n \in \mathbf{R}} (\alpha_n - d_1), d_2 \right) \quad (16)$$

$$t_w = \left( d_1, \max_{x_w \in \mathbf{R}} (\alpha_w - d_2) \right)$$

여기서  $\alpha_n$ 은 NR-U 사용자들에게 분배된 점유 시간,  $\alpha_w$ 는 Wi-Fi 사용자들에게 분배된 점유 시간을 나타낸다.

협상 결렬 점을 기준으로  $t_n$ 과  $t_w$ 의 중간 점을 새로운 협상 결렬 점으로 업데이트한다. 그 뒤, 새로운 협상 결렬 점을 기준으로 다시 자신의 점유 시간을 최대화하고 그 중간 점을 새로운 협상 결렬 점으로 업데이트하는 것을 반복한다. 마지막으로 협상 결렬 점이 경계 집합에 도달하면 해당 결렬 점이 순차적 라이파 협상 해가 된다. 해당 과정을 수식으로 나타내면 Equation (17), 최종 협상 해는 Equation (18)으로 나타낸다.

$$\forall n \in \mathbf{N} : m_n^{(S,d)} = \frac{1}{2} \left[ \left( \sqrt{T^2 - t_w^2}, d_2 \right) + \left( d_1, \sqrt{T^2 - t_n^2} \right) \right] \quad (17)$$

$$R(S, d) := \lim_{n \in \mathbf{N}_0} m_n^{(S,d)} \quad (18)$$

해당 과정을 의사 코드로 나타내면 Algorithm 2와 같다.

<b>Algorithm 2</b> 순차적 라이파 협상 해법을 활용한 채널 점유 시간 분할 알고리즘	
Initialize:	
Total Time = $T$	
Disagreement Point = $d = (d_1, d_2)$	
NR-U user's time occupation	
$t_n = \left( \max_{x_1 \in \mathbf{R}} (\alpha_n - d_1), d_2 \right)$	
Wi-Fi user's time occupation	
$t_w = \left( d_1, \max_{x_2 \in \mathbf{R}} (\alpha_w - d_2) \right)$	
While $T^2 > t_n^2 + t_w^2$	
$t_n \leftarrow \left( \sqrt{T^2 - t_w^2}, d_2 \right)$	
$t_w \leftarrow \left( d_1, \sqrt{T^2 - t_n^2} \right)$	
$d \leftarrow d + \frac{1}{2} (t_n, t_w)$	
Endwhile	
Output $R(S, d) = d$	

#### 4. 성능 평가

본 장에서는 본 논문이 제안하는 순차적 라이파 알고리즘의 성능을 분석하고, 몬테 카를로 정책 하강법을 적용한 RSMA의 성능을 평가하기 위해 제안한 알고리즘을 실행했을 때 다른 다중 접속 기술들(NOMA, OMA)과의 성능을 비교하고 한다. 먼저 성능 평가를 위한 환경 구성에 관해 설명을 한 뒤, 알고리즘의 성능 분석 및 타 기법들과의 성능을 비교한다.

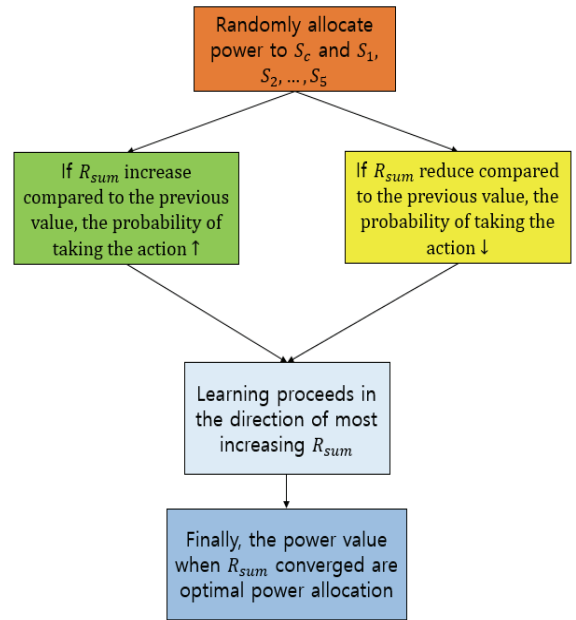


Fig. 4. Steps in Simulation

Table 1. System parameters

Notation	Value	Description
BS	1	No. of NR-U base station
$k_n$	5	No. of NR-U users
$P_t$	40dbm, 1~10dbm	Transmit Power of NR-U BS
B	10MHz	Bandwidth
$D_1, D_2, D_3, D_4, D_5$	5, 4, 3, 2, 1 km	Distance between BS and NR-U users
N	-114dbm	AWGN

본 논문에서 고려하는 네트워크의 형태는 Table 1과 같다. 기지국의 수는 1개이며, NR-U 사용자의 수는 5명으로 설정하였고 기지국과의 거리는 각각 5km, 4km, 3km, 2km, 1km가 떨어져 있으며 채널 상태는 무작위로 설정하였다. 기지국의 전송 전력은 40dBm이고 후에 전송 전력을 0~10dBm 일 때의 변화량을 관찰한다. 대역폭은 10MHz, AWGN은 -114dBm으로 설정하였다.

본 논문에서 제안하는 RSMA 기술의 성능 분석을 위해, 동일한 환경에서 RSMA와 다른 다중 접속 기술인 NOMA와 OMA(FDMA)의 평균 합 전송 속도를 비교하였다. 그 뒤, 비면허 대역의 채널 점유 시간을 분할할 때, 본 논문에서 제안한 순차적 라이파 협상 해법의 성능 분석을 위해 본 논문에서 제안한 순차적 라이파 협상 해법과 채널 점유 시간을 공평하게 반으로 나누는 동일 시간 할당 방식(STAM)과 처리량에 따른 전송 속도를 고려하여 비례하게 나눠주는 비례 공평성 분할 방식(PFWA)들을 통해 각 방식에 따른 처리량과 공평성을 비교하였다.

Fig. 5는 기지국의 전송 전력이 40dBm 일 때, RSMA와 NOMA, OMA(FDMA)의 학습이 진행됨에 따른 평균 총 전송 속도의 변화를 보여준다. NOMA에서도 RSMA와 마찬가지로 5명의 사용자가 있고 기지국으로부터 각각 5, 4, 3, 2, 1km씩 떨어져 있으며, 1000MHz의 대역폭을 가지고 있다고 설정하였다. OMA는 FDMA를 기준으로 설정했으며 가드 대역폭은 25MHz, 각 사용자들이 가지는 대역폭은 180MHz씩 가지고 있다고 설정하였다. OMA는 각 사용자들이 전력을 동일하게 소유하고 주파수 범위에 영향을 받기 때문에 일정한 평균 전송 속도를 가지는 것을 확인하였다. 또한 정책이 수렴하여 평균 전송 속도가 최대점에 도달하였을 때, RSMA가 다른 다중 접속 기술들보다 더 빠른 전송 속도를 나타내는 것을 확인할 수 있다.

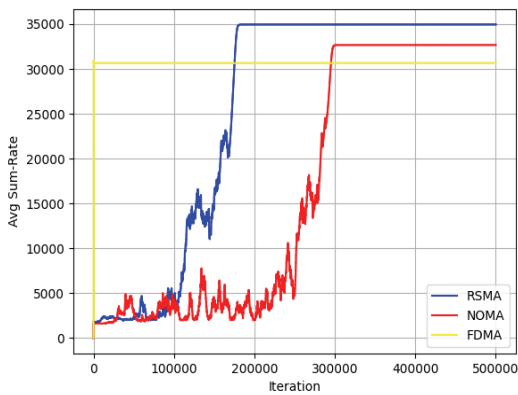


Fig. 5. Average Sum of Rate According to Iteration

Fig. 6은 기지국의 전송 전력이 1~10dBm 일 때, RSMA와 NOMA, OMA(FDMA)의 수렴했을 때의 합 전송 속도의 변화를 보여준다. NR-U 사용자와 기지국 사이의 거리, 기지국의 대역폭 등의 다른 조건들은 위의 실험과 동일하다. 모든 구간에서 최종적으로 합-전송 속도가 수렴하였을 때, RSMA가 다른 다중 접속 기술들보다 높은 속도를 나타내는 것을 확인할 수 있었다.

RSMA는 상대적으로 큰 전송 전력을 할당받는 공통 스트림과 작은 전송 전력을 할당받는 개인 스트림으로 분리된다. 그 뒤, 수신단에서 공통 스트림이 먼저 디코딩이 되고, 개인 스트림은 다른 개인 스트림에 의한 간섭만을 받게 된다. 따라서 RSMA는 NOMA보다 더 적은 간섭을 받기 때문에 SNR이 상대적으로 높아지므로 더 높은 합-전송 속도를 가지게 된다.

Fig. 7은 채널 점유 시간 분할 알고리즘에 따른 NR-U 사용자와 WiFi 사용자, 시스템의 평균 처리량을 나타낸다. NR-U 사용자와 WiFi 사용자의 수는 각각 5명으로 동일하고 총 채널 점유 시간은 24시간으로 설정하였다. 또한 NR-U 네트워크는 RSMA 방식을 사용하였을 때의 최대의 합 전송 속도인 35 Kbps, WiFi 네트워크의 전송 속도는 25Kbps로 설정하였다. 순차적 라이파 협상 해법과 STAM은 시스템의 평균 처리량에서 큰 차이가 없으나 PFWA에서는 다른 알고리즘에 비해 시스템의 평균 처리량이 떨어지는 것을 확인하였다.

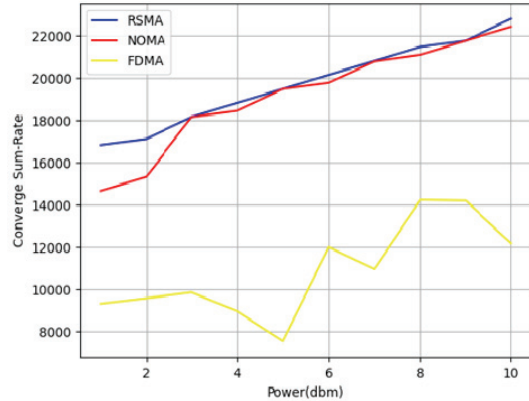


Fig. 6. Converge Sum of Rate from 1 to 10 dbm

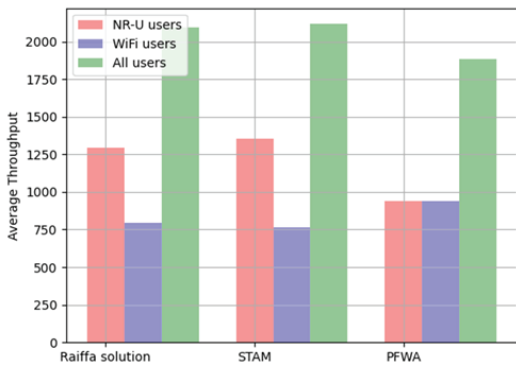


Fig. 7. Average Traffic According to Channel Occupancy Time Division Scheme

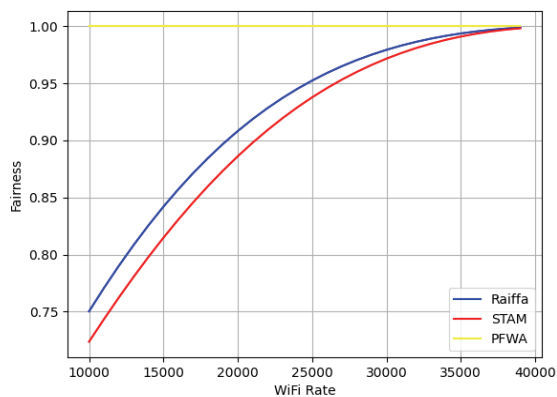


Fig. 8. Fairness Index Between NR-U and WiFi Network



Fig. 8은 NR-U 사용자와 WiFi 사용자 간의 처리량에 대한 공평성을 나타낸다. 공평성을 평가하기 위해 네트워크 엔지니어링에서 사용자 또는 응용 프로그램이 시스템 자원의 공정한 공유를 받고 있는지를 결정하는 데 주로 사용되는 Jain의 공평성 지수(Jain's Fairness Index)를 사용하였다. Jain의 공평성 지수는 다음 수식으로 정의된다[11].

$$J(x_1, x_2, \dots, x_n) = \frac{\left(\sum_{i=1}^n x_i\right)^2}{n \sum_{i=1}^n x_i^2} \quad (19)$$

본 논문에서  $x_i$ 는 제안알고리즘을 포함한 3가지의 알고리즘들의 NR-U 네트워크와 WiFi 네트워크의 처리량,  $n$ 은 네트워크 종류의 수로 정의하였다. NR-U 네트워크는 Fig 6에서와 마찬가지로 RSMA 방식에서의 전송 속도를 내고, WiFi 네트워크의 전송 속도가 10Kbps에서 40Kbps로 점차 증가할 때의 각각의 알고리즘들이 각 네트워크 사용자들의 작업을 얼마나 공평하게 처리할 수 있도록 채널 점유 시간을 할당해 주었는지를 확인하였다.

WiFi 네트워크의 전송 속도가 변화하는 모든 구간에서 순차적 라이파 협상 해법이 STAM보다 더 공평하게 채널 점유 시간을 할당해 주는 것을 확인하였다. PFWA는 WiFi 네트워크의 전송 속도에 상관없이 각 네트워크가 동일한 처리량을 처리하였다. 순차적 라이파 협상 해법은 시스템 처리량 면에서 STAM과 비슷한 성능을 보이지만 공평성 면에서 좀 더 뛰어난 성능을 보이고, PFWA는 공평성 면에서 다른 알고리즘보다 뛰어난 성능을 보이지만 시스템 처리량이 크게 떨어지는 것을 확인하였다.

Fig. 9는 Equation (7)의 결과로, 각 기법간의 전송량과 공평성을 고려한 효용을 나타낸다. 실험 결과, 제안하는 순차적

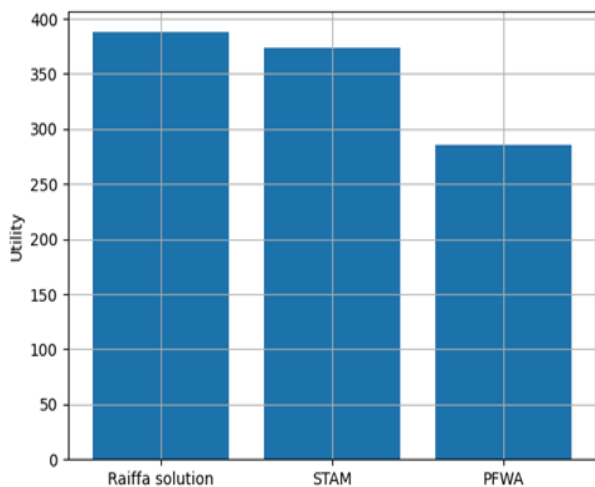


Fig. 9. Utility Between Traffic and Fairness

라이파 협상 해법이 다른 기법들보다 데이터 전송량과 공평성 면에서 뛰어난 성능을 보였다. 순차적 라이파 협상 해법은 앞의 3.1장에서 소개한 파레토 최적을 비롯한 7개의 공리를 만족한다. 즉, 최대 효율성을 보장함과 동시에 공평성을 고려하기 때문에 각 면에서 최고의 성능은 아닐지라도 종합적인 면에서 다른 알고리즘보다 뛰어난 성능을 보인다.

### 5. 결론

본 논문에서는 NR-U 기지국의 전송 전력과 비면허 대역 대역폭 등의 무선 자원이 한정되어 있는 상황에서 증가하는 모바일 데이터 수요를 충족시키기 위해 전송률 분할 다중 접속 기술을 활용하여 5G NR-U 네트워크의 데이터 전송 속도를 최대화하고, 비면허 대역의 NR-U 네트워크와 WiFi 네트워크의 처리량과 공정성을 최대화하기 위해 강화 학습의 몬테 카를로 정책 하강법과 게임 이론의 순차적 라이파 협상 해법을 제안하였다.

먼저 몬테 카를로 정책 하강법을 사용하여 RSMA를 사용하는 NR-U 사용자들의 합-전송 속도를 최대화하도록 전력을 할당해 주었다. 그 뒤, NR-U 네트워크와 WiFi 네트워크의 처리해야 할 작업량과 데이터 전송 속도를 고려해 순차적 라이파 협상 해법을 사용하여 비면허 대역의 채널 점유 시간을 할당하였다.

실험 결과 제안된 강화 학습 알고리즘에서 기지국의 전송 전력이 동일할 때, 전송률 분할 다중 접속 기술이 기존의 다중 접속 기술들인 비직교 다중 접속, 직교 다중 접속 기술보다 더 높은 합-전송 속도를 보임을 입증하였다. 또한 제안한 순차적 라이파 협상 해법이 전체 시스템의 평균 처리량과 공정성 면에서 우수한 성능을 보임을 확인하였다.

### References

- [1] Cisco, Cisco Annual Internet Report. Cisco Annual Internet Report (2018-2023) White Paper [Internet], <https://www.cisco.com/c/en/us/solutions/collateral/executive-perspectives/annual-internet-report/white-paper-c11-741490.html>
- [2] 3GPP TS 36.213 V13.0.1, Evolved universal terrestrial radio access (E-UTRA) physical layer procedures (Release 13), 3GPP, 2016.
- [3] M. Le, Q. V. Pham, H. C. Kim, and W. J. Hwang, "Enhanced resource allocation in D2D communications with NOMA and unlicensed spectrum," *IEEE System Journal*, Vol.16, No.2, pp.2856-2866, 2022.
- [4] M. Alsabah et al., "6G wireless communications networks: A comprehensive survey," *IEEE Access*, Vol.9, pp.148191-148243, 2021.

- [5] S. Gamal, M. Rihan, S. Hussin, A. Zaghoul, and A. A. Salem, "Multiple access in cognitive radio networks: From orthogonal and non-orthogonal to rate-splitting," *IEEE Access*, Vol.9, pp.95569-95584, 2021.
- [6] W. Trockel, "Axiomatization of the discrete Raiffa solution," *Economic Theory Bulletin*, Vol.3, No.1, pp.9-17, 2014.
- [7] W. Trockel, "An exact non-cooperative support for the sequential Raiffa solution," *Journal of Mathematical Economics*, Vol.47, No.1, pp.77-83, 2011.
- [8] R. S. Sutton and A. G. Barto, "Reinforcement learning: An introduction," Cambridge, MA: MIT Press, 2018.
- [9] H. He, H. Shan, A. Huang, L. X. Cai, and T. Q. Quek, "Proportional fairness-based resource allocation for LTE-U coexisting with Wi-Fi," *IEEE Access*, Vol.5, pp.4720-4731, 2016.
- [10] N. Q. Hieu, D. T. Hoang, D. Niyato, and D. I. Kim, "Optimal power allocation for rate splitting communications with deep reinforcement learning," *IEEE Wireless Communications Letters*, Vol.10, No.12, pp.2820-2823, 2021.
- [11] R. Jain, D. Chiu, and W. Hawe, "A quantitative measure of fairness and discrimination for resource allocation in shared computer systems," Digital Equipment Corporation, Sep. 1984.



### 전 장 우

<https://orcid.org/0009-0000-1106-0664>  
e-mail : zangwoo@sogang.ac.kr  
2020년 인천대학교 전자공학과(학사)  
2023년 서강대학교 컴퓨터공학과(석사)  
관심분야 : 게임이론을 이용한 네트워크 자원관리



### 김 승 욱

<https://orcid.org/0000-0003-1967-151X>  
e-mail : swkim01@sogang.ac.kr  
1993년 서강대학교 전자계산학과(학사)  
1995년 서강대학교 전자계산학과(석사)  
2003년 Syracuse University,  
Computer Science 박사

2005년 중앙대학교 컴퓨터공학부 조교수  
2006년 ~ 현 재 서강대학교 컴퓨터공학과 교수  
관심분야 : 게임이론을 이용한 네트워크 자원관리