

# Q-learning을 이용한 우선순위 기반 Adaptive Slotted ALOHA 기법

김지나\*, 김병찬\*, 유철우\*, 박현희<sup>o</sup>

## Priority Based Adaptive Slotted ALOHA Method Using Q-Learning

Jeena Kim\*, Byungchan Kim\*, Cheolwoo You\*, Hyunhee Park<sup>o</sup>

### 요약

본 논문에서는 기존 slotted ALOHA 기법에 Q-learning과 우선순위를 적용하는 Priority based Adaptive slotted ALOHA(PAS-ALOHA)를 제안한다. PAS-ALOHA 기법은 각 디바이스의  $Q$ 값에 따라 이전 상태에서 데이터 패킷의 전송 성공 횟수가 많은 디바이스에 우선 순위를 지정한다. 다음으로, 남은 데이터 패킷을 전송할 때 다른 디바이스의 제약을 받지 않기 때문에 성공 비율을 높이고, 충돌 비율을 감소시킨다. 또한, Jain's Fairness Index를 적용하여 음의 보상에 따라 각 디바이스의 공정성을 비교한다. 실험 결과 기존 slotted ALOHA 기법에 비해 성공 비율은 약 69% 증가하며 충돌 비율은 최대 30% 감소함을 보였다. 따라서, 5G MTC 환경에서 디바이스 개수가 증가하더라도 채널 상태에 적응하여 데이터 패킷을 성공적이고, 공평하게 전송할 수 있을 것으로 기대된다.

**Key Words** : Reinforcement learning, Q-learning, Slotted ALOHA, 5G cellular network, Machine Type Communication

### ABSTRACT

In this paper, we propose Priority-based Adaptive slotted ALOHA (PAS-ALOHA), which applies Q-learning and priorities to existing slotted ALOHA method. The PAS-ALOHA method prioritizes devices with many number of successful transmissions of data packets in the previous state according to the  $Q$ -value of each device. Next, when transmitting the remaining data packet, since it is not restricted by other devices, the success ratio is increased and the collision ratio is reduced. In addition, the fairness of each device is compared according to negative reward by applying the Jain's Fairness Index. Experiments showed that the success ratio increased by about 69% and the collision ratio decreased by up to 30%. Therefore, even if the number of devices increases in the 5G MTC environment, it is expected that data packets can be successfully and fairly transmitted by adapting to the channel state.

\* 본 논문은 2021년도 정부(과학기술정보통신부)의 재원으로 정보통신기획평가원의 지원을 받아 수행된 연구임 (No. 2021-0-00368, 6G 서비스를 위한 인공지능/머신러닝 기반 자율형 MAC 개발)

\* 본 논문은 2022년도 정부(교육부)의 재원으로 한국연구재단의 지원을 받아 수행된 연구임 (No. 2022R1A2C2005705, 분산 머신러닝 기반 지능형 플라이 지지국을 위한 AI-MAC 프로토콜)

• First Author : Myongji University, Department of Information and Communication Engineering, jnkim@mju.ac.kr, 학생회원

° Corresponding Author : Myongji University Department of Information and Communication Engineering, hhpark@mju.ac.kr, 종신회원

\* Myongji University, Department of Information and Communication Engineering, 학생회원; 종신회원

논문번호 : 202210-244-B-RE, Received October 12, 2022; Revised December 13, 2022; Accepted December 14, 2022

## I. 서론

Machine Type Communication(MTC)는 5G 이동 통신에서 다수의 기계형 통신을 지원하는 통신 형태이다<sup>[1,2]</sup>. MTC 환경에서는 다수의 디바이스를 셀룰러 네트워크에 연결시키고자 한다. MTC 환경에서 수많은 디바이스의 연결을 기존 통신 인프라에 적용해 네트워크에 연결할 수 있었다. 그러나, 데이터 패킷을 전송할 때 사용하는 기존의 slotted ALOHA<sup>[3]</sup> 기법을 5G MTC 환경에서 고려하는 다수의 디바이스의 데이터 패킷 전송을 지원하는 것은 충돌 비율이 높아진다는 문제점이 발생한다. 따라서, 주어진 시간 간격 내에서 디바이스가 데이터 패킷을 전송할 때 충돌 횟수를 낮추고 성공적인 전송을 하는 새로운 액세스 기법이 필요하다.

디바이스가 데이터 패킷을 전송하는 데 사용되는 기법 중 slotted ALOHA 기법은 이동 통신 기술 중 다중 접속을 허용하는 간단한 기법이다. Slotted ALOHA는 정해진 시간 간격 내에서 여러 개의 타임 슬롯으로 나눈다. 디바이스들은 한 타임 슬롯 당 하나의 데이터 패킷을 전송할 수 있다. 하지만, 두 대 이상의 디바이스가 동일한 타임 슬롯에서 전송을 시도할 때는 충돌이 발생한다. 따라서, 5G MTC 환경에서 기하급수적으로 증가하는 디바이스들이 방대한 양의 데이터 패킷이 전송을 시도하는 경우 충돌 횟수가 급격히 증가한다<sup>[4]</sup>.

최근 이동 통신 연구에서는 다수의 단말의 데이터 패킷의 충돌 발생 횟수 증가 문제를 해결하기 위해 기존 slotted ALOHA 기법에 딥 러닝 및 머신 러닝을 적용시켰다<sup>[5]</sup>. 따라서, 디바이스가 데이터 패킷의 전송 결과를 학습하여 다음 데이터 패킷을 보내고자 하는 타임 슬롯의 위치를 예측할 수 있다. 본 논문에서는 5G MTC 환경에서 다수의 디바이스가 동시에 데이터 패킷을 보낼 때 발생하는 타임 슬롯 내 혼잡도를 낮추는 것에 초점을 맞춘다.

본 논문에서는 기존의 slotted ALOHA 기법에 강화학습 기법의 하나인 Q-learning을 적용한다<sup>[6,7,8,9]</sup>. Q-learning은 현재 정책에 따라 결정된 행동에 대해 즉각적인 보상을 받는다. 이 보상은 다음 상태에서 사용할 정책에 반영한다. 이를 통해 Q값을 업데이트하고 보상을 최대화할 수 있는 정책을 찾도록 학습한다.

일반적인 통신 환경에서 디바이스가 데이터 패킷을 보내는 과정은 지속해서 변화하며 채널 내 환경이 혼잡하다. 따라서, 전송하는 데이터 패킷의 전송 결과를 디바이스가 Q-learning을 사용하여 학습하도록 한다.

각 디바이스는 데이터 패킷 전송 결과에 따라 보상을 받는다. 보상은 다음 상태에서 사용할 정책에 반영되며, 다음 상태에서의 행동을 결정한다. 각 디바이스는 개별의 Q값을 가지고 있으며, Q-table을 통해 이전 상태의 행동을 기록한다. 따라서, 디바이스는 Q값을 통해 보상을 최대화할 수 있는 정책을 찾아 자신의 고유한 슬롯을 점유할 수 있다.

반면에 제안하는 기법은 동시에 데이터 패킷을 전송하는 디바이스가 상대적으로 많은 5G MTC 환경을 고려한다. 본 논문에서 제안하는 Priority based Adaptive slotted ALOHA(PAS-ALOHA) 기법은 Q-learning을 적용한 Adaptive slotted ALOHA 기법에서 Q값에 따라 디바이스의 우선순위를 지정하도록 학습한다. PAS-ALOHA 기법은 기존의 slotted ALOHA 기법과 비교했을 때 데이터 패킷 전송 성공 비율은 최대 69% 증가하였고, 충돌 발생 비율은 최대 30% 감소했다. 또한 Jain's Fairness Index<sup>[10]</sup>를 통해 우선순위를 지정받은 디바이스를 제외한 모든 디바이스가 공평하게 데이터 패킷을 전송하는지 공정성을 검증한다.

본 논문의 2장에서는 관련 연구에 관해서 기술하며, 3장에서는 본 논문에서 제안하는 PAS-ALOHA 기법에 관해서 기술한다. 4장에서는 실험을 통해 PAS-ALOHA 기법의 성능을 검증한다.

## II. 본론

### 2.1 관련 연구

#### 2.1.1 ALOHA 기법

ALOHA는 각 디바이스가 데이터 패킷을 전송할 때 사용되는 기법으로 pure ALOHA와 slotted ALOHA 기법이 있다. Pure ALOHA 기법은 전송해야 할 데이터 패킷이 생기면 바로 전송을 시도한다. 데이터를 전송한 뒤 일정 시간 동안 ACK을 기다린다. 또한 pure ALOHA 기법은 전송하는 데이터 패킷이 1bit만 겹쳐도 충돌이 발생하여 재전송을 시도한다. 따라서, 전송량에 비해 채널 효율이 18%로 낮은 효율을 보인다. 충돌에 의한 채널 효율의 문제를 해결하기 위해 설계된 기법이 slotted ALOHA 기법이다.

Slotted ALOHA 기법은 정해진 시간 간격 내에서 여러 개의 타임 슬롯으로 나눈다. 각 디바이스는 한 타임 슬롯 당 하나의 데이터 패킷을 전송할 수 있다. 그러므로, pure ALOHA 기법보다 정해진 타임 슬롯에 데이터 패킷을 전송하는 slotted ALOHA 기법의

충돌 비율이 감소할 수 있다. 따라서, pure ALOHA 기법보다 채널 효율이 2배 더 높은 36%로 더 높아진다. 하지만 두 대 이상의 디바이스가 같은 타임 슬롯에 전송을 시도할 때 충돌이 발생한다.

Pure ALOHA와 slotted ALOHA 기법의 동일한 문제는 데이터 패킷을 전송할 때 현재 채널의 상태를 고려하지 않는 것이다. 데이터 패킷이 충돌한 디바이스들은 임의의 대기 시간만큼 기다린 후 재전송을 시도한다. 따라서, 다수의 디바이스가 동시에 데이터 패킷을 전송하는 5G MTC 환경에서 기존의 pure ALOHA와 slotted 기법을 사용하면, 충돌 비율이 높아지거나 지연 시간이 증가한다는 문제점이 발생한다.

### 2.1.2 Q-learning based slotted ALOHA

강화학습 기법 중 하나인 Q-learning은 의사결정자가 현재 상태  $s^t \in State$ 에서 정책에 따라 결정된 행동  $a^t \in Action$ 에 대한 보상  $r^t \in Reward$ 를 받는다. 이를 통해 다음 상태의 Q값을 업데이트한다. Q-learning은 현재 time step  $t \in T$ 에서 다음 상태까지 고려하기 때문에 지속해서 변화하는 환경에 적용할 수 있다. 의사결정자는 총 에피소드를 학습하면서 현재 상태에서 보상을 최대화할 수 있는 행동을 한다. 이때, 보상 값은 미래의 받을 보상의 가치를 현재 상태에서 환산하여 보상의 가치가 감소하는 것을 줄여야 한다. 따라서, 0과 1 사이의  $\gamma$ 값 할인율(discount factor)을 사용하여 현재의 가치로 환산할 수 있다. 식(1)은 Q-learning에서 의사결정자가 가진 Q-value를 업데이트하는 식을 나타낸다<sup>[10]</sup>.

$$Q'(s^t, a^t) = Q(s^t, a^t) + \alpha(r^{t+1} + \gamma \max_a Q(s^{(t+1)}, a) - Q(s^t, a^t)) \quad (1)$$

Q-learning based slotted ALOHA 기법은 기존 slotted ALOHA 기법에 Q-learning을 적용했다<sup>[11]</sup>. Q-learning based slotted ALOHA 기법은 slotted ALOHA와 같이 정해진 시간 간격 내에서 여러 개의 타임 슬롯으로 나눈다. 한 타임 슬롯당 하나의 디바이스가 데이터 패킷을 전송할 수 있다. 데이터 패킷 전송에 참여하는 각 타임 슬롯은 개별의 Q값을 가지고 있으며, 초기 접속 시 Q값은 0으로 초기화된다. Q값은 디바이스가 타임 슬롯에 데이터 패킷을 전송한 결과를 바탕으로 이전의 보상을 받는다. 디바이스가 타임 슬롯에 데이터 패킷을 성공적으로 전송했을 때는 +1의 양의 보상을 받는다. 하지만, 같은 타임 슬롯에

데이터 패킷 전송을 시도한 디바이스가 존재하여 충돌이 발생했을 때는 -1의 음의 보상을 받는다.

$$R_{n,i} = \begin{cases} +1, & \text{if TX success} \\ -1, & \text{otherwise} \end{cases} \quad (2)$$

$$Q_{n,i}^{t+1} = Q_{n,i} + \alpha(R_{n,i} - Q_{n,i}^t) \quad (3)$$

식(2)은 디바이스의 데이터 패킷 전송 결과에 따른 이전 보상<sup>[11]</sup>을 나타낸다.  $R_{n,i}$ 는  $n \in N$ 번째 노드가  $i \in T$ 번째 슬롯에 전송을 시도했을 때 받을 수 있는 보상을 나타낸다. 각 디바이스는 받은 보상을 바탕으로 다음 상태의 Q값을 식(3)과 같이 업데이트한다.  $Q_{n,i}^t$ 는 현재 상태의 디바이스의 Q값이며,  $\alpha$ 는 learning rate(학습률)이다. Learning rate는 시행 횟수에 따라 늘어나거나 줄어들도록 조절해서 학습이 잘 되는 방향으로 이끈다.

그림 1은 Q-learning based slotted ALOHA의 Q값 업데이트 예시이다. 정해진 시간 간격이 4개의 타임 슬롯으로 나누어져 있다고 가정한다. Case 1은 디바이스가 첫 번째 타임 슬롯에 데이터 패킷을 전송한다. 하나의 디바이스가 전송을 시도하여 전송을 성공한 예시이다. 따라서, 다음 상태에서 해당 타임 슬롯은 다음 타임 스템에서 양의 보상 +1과 함께 업데이트된 Q값을 갖는다. Case 2는 첫 번째 타임 슬롯에 두 대 이상의 디바이스가 전송을 시도한다. 따라서 충돌이 발생하여 다음 타임 스템에서 해당 타임 슬롯은 음의 보상 -1과 함께 업데이트된 Q값을 갖는다.

Q-learning based slotted ALOHA는 기존의 slotted ALOHA 기법에 Q-learning을 적용하였다. Q-learning based slotted ALOHA는 낭비되는 타임 슬롯을 줄이고 처리량을 높이는 데 성공하였다. 그러나, 3GPP 표준에서 제안하는 5G MTC 환경의 디바이스 개수를 모두 적용하기에 한계가 있을 것으로 판단된다<sup>[12]</sup>. 또

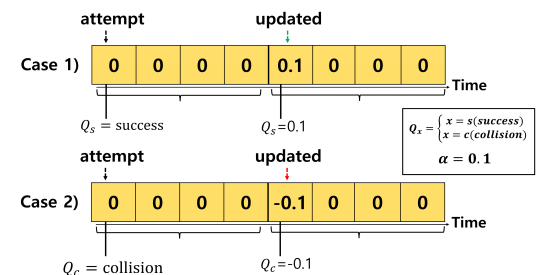


그림 1. Q값 업데이트 예시  
Fig. 1. Example of Q-value update

한, 실험 결과에서 볼 수 있듯이 이를 극복할 필요가 있다.

따라서, 다수의 디바이스가 동시에 데이터 패킷을 전송할 때 충돌 비율이 급격히 증가하는 문제를 해결하기 위해 PAS-ALOHA 기법을 제안한다.

### 2.2 Priority based Adaptive slotted ALOHA (PAS-ALOHA) 기법

본 논문에서 제안하는 기법은 기존 slotted ALOHA 기법에 Q-learning을 적용한다. 따라서, 데이터 패킷 전송에 참여하는 디바이스는 의사결정자가 된다. 각 디바이스는 개별의 Q값을 가지며, 데이터 패킷 전송 결과를 바탕으로 디바이스가 가진 Q값을 업데이트한다. 각 디바이스가 가진 Q값의 Q-table에는 이전 상태의 행동을 기록한다. 디바이스가 타임 슬롯에 데이터 패킷 전송을 시도하고 성공한 경우 +1의 양의 보상을 받는다. 두 대 이상의 디바이스가 같은 타임 슬롯에 데이터 패킷을 전송 시도하면 충돌이 발생한다. 따라서 디바이스는 -1의 음의 보상을 받는다. PAS-ALOHA 기법은 같은 타임 슬롯에 데이터 패킷을 전송 시도하여 충돌이 발생한 디바이스의 총 개수를 고려하여 음의 보상을 조절한다.

그림 2는 디바이스의 타임 슬롯 내 충돌 예시를 나타낸다. 디바이스 1은 첫 번째 타임 슬롯에 전송을 시도했다. 첫 번째 타임 슬롯에 데이터 패킷 전송을 시도한 디바이스는 한 대 이므로 데이터 패킷 전송에 성공한다. 하지만, 디바이스 3과 디바이스 4는 두 번째 타임 슬롯에 동시에 데이터 패킷 전송을 시도하여 충돌이 발생한다. 두 번째 타임 슬롯에 전송을 시도한 디바이스의 수는 총 두 대이다. 따라서, 충돌이 발생한 디바이스의 총개수인 N은 2가 된다.

$$R_{n,i}^{ADP} = \begin{cases} +1, & \text{if transmission success in time slot } i \\ \psi_{n,i}, & \text{if collision occurs in timeslot } i \end{cases} \quad (4)$$

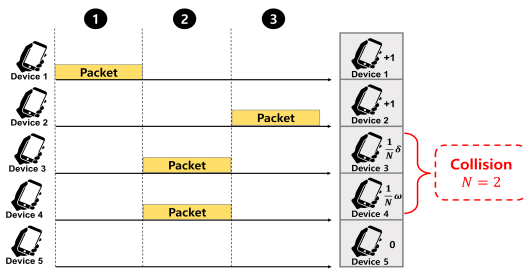


그림 2. 디바이스의 타임 슬롯 내 충돌 예시  
Fig. 2. Example of device collision in time slot

$$\psi_{n,i} = \begin{cases} \frac{1}{N}\delta, & \text{where } \max(Q_1, Q_2, \dots, Q_{n-1}, Q_n), n \in N \\ \frac{1}{N}\omega, & \text{otherwise} \end{cases} \quad (5)$$

식(4)는 제안된 기법에서 디바이스의 데이터 패킷 전송 결과에 따른 보상을 나타낸다. 디바이스가 데이터 패킷 전송에 성공했을 때는 +1의 양의 보상을 받는다. 하지만 디바이스가 충돌이 발생했을 때는 식(5)에서 충돌한 디바이스들의 Q값을 비교한다. 충돌이 발생한 각 디바이스 Q값을 비교한 후 충돌한 디바이스 중 가장 높은 Q값을 가진 디바이스는 완화된 음의 보상을 받는다.

그림 3은 제안하는 PAS-ALOHA 기법에서 보상 업데이트 예시를 나타낸다. 같은 타임 슬롯에 데이터 패킷 전송을 시도하여 충돌이 발생한 디바이스 3과 디바이스 4의 Q값을 비교한다. 디바이스 3은 이전 상태에서의 전송 성공 횟수가 많아 디바이스 4보다 많다. 따라서 디바이스 4와 비교했을 때 더 높은 Q값을 가지고 있다. 따라서, 그림 3과 같이 디바이스 3은 완화된 음의 보상 1/2δ를 받게 된다. 디바이스 4는 이전 상태에서 디바이스 3보다 충돌 발생 횟수가 많아 디바이스 3보다 더 낮은 Q값을 가지고 있다. 따라서, 디바이스 4는 1/2ω의 음의 보상을 받게 된다. 여기서 δ와 ω는 각각 다른 상수를 갖는다고 가정한다. 이때, δ를 받는 디바이스는 ω를 받는 디바이스 보다 항상 높은 Q값을 가져야 한다.

디바이스가 가진 Q값을 비교하여 우선순위를 지정한 후, 음의 보상을 다르게 받는 의사코드를 알고리즘 1에 작성했다.

각 디바이스는 전송 결과에 따라 보상을 받은 후 식(6)과 같이 Q값을 업데이트한다. 디바이스는 데이

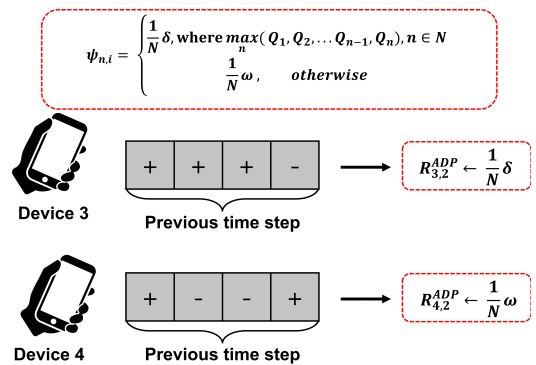


그림 3. PAS-ALOHA 기법의 보상 업데이트 예시  
Fig. 3. Example of PAS-ALOHA method reward update

**알고리즘 1: Q값 비교 알고리즘**

```

Initialize  $Q_{n,i}^t = 0, \forall n \in N, \forall i \in I$ 
1: for  $n = 1 : N$ 
2:   if collision occurs then
3:     Compare Q-value according to Eq. (5)
4:     if  $\max_n Q_n$  then
5:        $R_{n,i}^{ADP} \leftarrow \frac{1}{N} \delta$ 
6:     else
7:        $R_{n,i}^{ADP} \leftarrow \frac{1}{N} \omega$ 
    
```

터 패킷 전송 결과에 따라 받은 보상을 바탕으로 지속적으로 Q값을 업데이트한다. PAS-ALOHA 기법은 우선순위를 지정할 때 충돌이 발생한 각 디바이스의 Q값을 비교한다. 따라서, 충돌이 발생한 디바이스 중 가장 Q값이 높은 디바이스는 이전 상태에서 전송 성공 횟수가 많으므로 다른 디바이스보다 빠르게 고유한 타임 슬롯을 점유할 수 있다. 따라서, 남아있는 데이터 패킷들을 다른 디바이스에 의해 제약받지 않고 데이터 패킷 전송에 성공할 수 있도록 우선순위를 지정한다. PAS-ALOHA는 기존의 Q-learning based slotted ALOHA와 비교했을 때 이진으로 구성된 음의 보상을 받는 것과 달리 타임 슬롯 내 혼잡도와 각 디바이스가 가진 Q값을 고려하여 개별적으로 음의 보상을 받는다. 따라서 다음 데이터 패킷 전송 시 최적의 타임 슬롯을 선택하도록 학습한다.

$$Q_{n,i}^{t+1} = Q_{n,i}^t + \alpha (R_{n,i}^{ADP} - Q_{n,i}^t) \quad (6)$$

**III. 실험**

**3.1 실험 환경**

본 논문에서는 제안하는 기법의 데이터 패킷 전송 성공 비율과 충돌 발생 비율을 기존의 slotted ALOHA 기법과 비교하여 성능을 검증한다. 전송 성공 비율과 충돌 발생 비율을 learning rate를 각각 {0.5, 0.1, 0.01}로 변화시켜 가며 성능을 검증한다. 또한, PAS-ALOHA 기법의 음의 보상 파라미터가 성능에 어떠한 영향을 미치는지 검증하기 위해 파라미터에 변화를 주어 성능에 어떠한 영향을 미치는지 검증한다. 본 논문에서는 음의 보상 파라미터를 각각 {-0.01, -0.02}, {-0.1, -0.2}, {-1, -2}로 설정하여 파라미터에 따른 정규화된 처리량을 구한다. 또한, 가장

높은 Q값을 가진 디바이스를 제외한 모든 디바이스들이 공평하게 데이터 패킷을 검증하기 위해 정규화된 처리량을 통해 Jain's Fairness Index를 계산한다.

본 논문에서는 데이터 패킷을 전송하고자 하는 디바이스의 개수가 상대적으로 적은 환경의 디바이스 개수를 5대부터 25대로 설정한다. 또한, 상대적으로 디바이스 개수가 많은 5G MTC 환경을 고려하기 위해 디바이스 개수를 150대부터 300대로 설정하여 디바이스가 데이터 패킷을 보내는 환경의 성능을 검증한다. 각 디바이스는 보내고자 하는 데이터 패킷을 100개씩 가지고 있으며 디바이스가 100개의 데이터 패킷을 모두 보낼 때까지 반복한다.

다음 표 1은 본 실험 환경을 나타낸다. 성공 비율과 충돌 비율의 성능 검증의 기준을 얻고자 본 실험에서는 동일한 환경을 가진 slotted ALOHA 기법과 디바이스의 증가에 따른 성공 비율과 충돌 비율을 비교한다.

표 1. 실험 환경  
Table 1. The experiment environment

Contents	Value
The number of devices	{5;25} {150;300}
The number of packet per devices	100
The number of slots	300
Learning rate	{0.5, 0.1, 0.01}
Negative rewards parameter	{-0.01, -0.02}, {-0.1, -0.2}, {-1, -2}

**3.2 실험 결과**

그림 4는 동시 접속 채널에 접속하는 디바이스 개수가 상대적으로 적은 환경을 고려하기 위해 디바이스 개수를 5대부터 25대까지로 설정하였다. 본 환경에서 PAS-ALOHA 기법의 데이터 패킷 전송 성공 비율을 약 97%로 기존 slotted ALOHA 기법보다 최대 70% 증가함을 보였다.

그림 5는 5G MTC 환경을 고려하기 위해 동시 접속 채널에 참여하는 디바이스 개수를 150대부터 300대까지로 설정하였다. 본 환경에서 PAS-ALOHA 기법의 전송 성공 비율은 약 63%으로 기존의 slotted ALOHA 기법보다 최대 68% 증가함을 보였다.

그림 6은 디바이스 개수 증가에 따른 충돌 비율을 나타낸다. 디바이스의 개수가 5대에서 25대까지 증가하는 PAS-ALOHA 기법의 충돌 발생 비율은 약 0.04%으로 기존 slotted ALOHA 기법보다 최대 29%

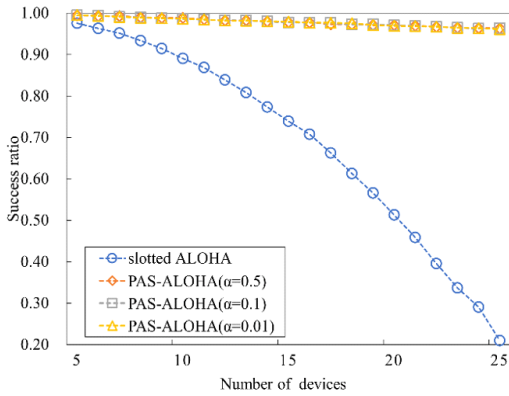


그림 4. 디바이스 개수 증가에 따른 전송 성공 비율(5-25대 사이)  
Fig. 4. Ratio of success with increasing number of devices(between the number of devices 5-25)

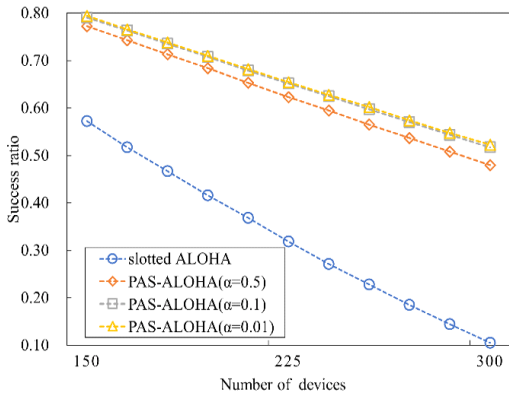


그림 5. 디바이스 개수 증가에 따른 전송 성공 비율(150-300대 사이)  
Fig. 5. Ratio of success with increasing number of devices(between the number of devices 150-300)

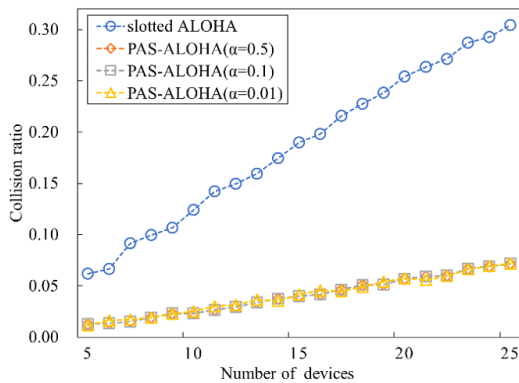


그림 6. 디바이스 개수 증가에 따른 충돌 발생 비율(5-25대 사이)  
Fig. 6. Ratio of collision with increasing number of devices(between the number of devices 5-25)

감소함을 보였다.

그림 7은 디바이스 개수 증가에 따른 충돌 비율을 나타낸다. 디바이스의 개수가 150대부터 300대까지증가하는 PAS-ALOHA 기법의 충돌 발생 비율은 약 52%으로 기존 slotted ALOHA 기법보다 최대 30% 감소함을 보였다.

그림 8은 음의 보상에 따른 정규화된 처리량을 나타낸다. 디바이스의 개수는 5대부터 25대로 증가한다. PAS-ALOHA 기법의 정규화된 처리량을 구하기 위해, learning rate를 가장 학습이 잘 되었던 0.5로 고정하였다. PAS-ALOHA 기법은 음의 보상 파라미터를 {-1, -2}로 설정하였을 때 0.5509로 가장 높은 정규화된 처리량을 보였다.

그림 9는 음의 보상에 따른 정규화된 처리량을 나타낸다. 디바이스의 개수는 150대부터 300대로 증가

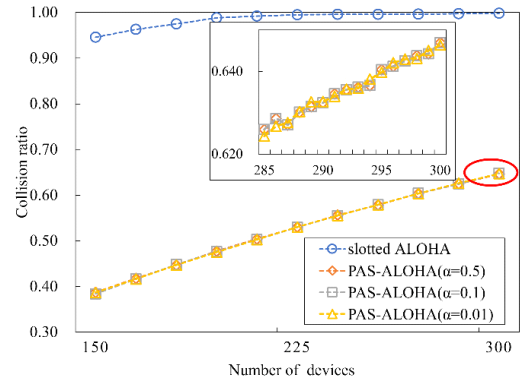


그림 7. 디바이스 개수 증가에 따른 충돌 발생 비율(150-300대 사이)  
Fig. 7. Ratio of collision with increasing number of devices(between the number of devices 150-300)

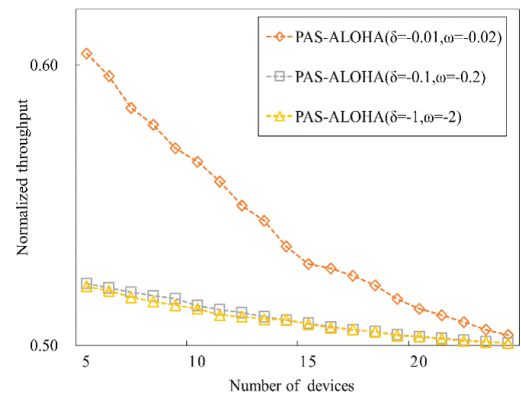


그림 8. 음의 보상에 따른 정규화된 처리량(5-25대 사이)  
Fig. 8. Normalized throughput based on negative reward parameter(between the number of devices 5-25)



한다. PAS-ALOHA 기법의 정규화된 처리량을 구하기 위해 learning rate를 가장 학습이 잘 되었던 0.01로 고정하였다. PAS-ALOHA 기법은 음의 보상 파라미터를  $\{-0.01, -0.02\}$ 로 설정하였을 때 0.427로 가장 높은 정규화된 처리량을 보였다.

그림 10은 디바이스 개수가 상대적으로 적은 5대에서 25대까지인 환경에서 PAS-ALOHA 기법의 공정성을 비교했다. PAS-ALOHA 기법은 디바이스가 가진  $Q$ 값에 따라 우선순위를 지정한다. 따라서, 충돌이 발생했을 경우  $Q$ 값이 가장 높은 디바이스를 제외한 다른 디바이스들이 공평하게 데이터 패킷을 전송하는지 검증했다. 제안하는 기법의 공정성을 검증하기 위해 그림 8에서 구한 정규화된 처리량을 이용하여 Jain's Fairness Index를 계산했다. Jain's Fairness Index는 0과 1의 사이의 값으로 1에 가까울수록 공정성이 높다고 판단한다. PAS-ALOHA 기법은 약 0.87의 공정성을 나타냈으며, 음의 보상 파라미터를  $\{-1, -2\}$ 로 설정하였을 때 0.907로 가장 높은 공정성을 보였다.

그림 11은 디바이스가 개수가 상대적으로 많은 150

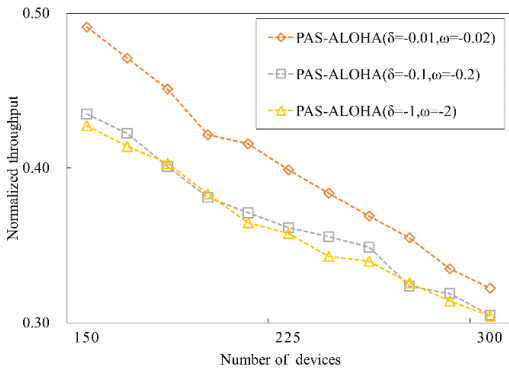


그림 9. 음의 보상에 따른 정규화된 처리량(150-300대 사이)  
Fig. 9. Normalized throughput based on negative reward parameter(between the number of devices 150-300)

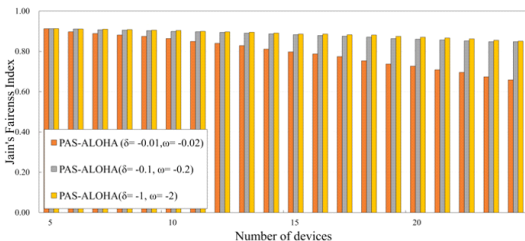


그림 10. 디바이스 개수 증가에 따른 Jain's Fairness Index (5-25대 사이)  
Fig. 10. Jain's Fairness with increasing number of devices(between the number of devices 5-25)

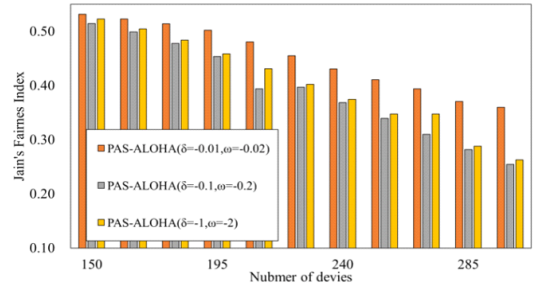


그림 11. 디바이스 개수 증가에 따른 Jain's Fairness Index (150-300대 사이)  
Fig. 11. Jain's Fairness with increasing number of devices(between the number of devices 150-300)

대에서 300대 사이의 환경에서 제안하는 기법의 공정성을 검증했다. 그림 9의 정규화된 처리량을 이용하여 본 환경에서 Jain's Fairness Index를 계산했다. PAS-ALOHA 기법은 약 0.4의 공정성을 보이며, 음의 보상 파라미터를  $\{-0.01, -0.02\}$ 로 설정하였을 때 0.45로 가장 높은 공정성을 보였다.

### 3.2.1 Learning rate에 따른 성능 비교

본 논문에서는 PAS-ALOHA 기법의 디바이스 개수의 증가에 따른 learning rate 별 전송 성공 비율과 충돌 발생 비율의 성능을 검증하였다. Learning rate는 각각  $\{0.5, 0.1, 0.01\}$ 로 설정하였으며 성능은 다음과 같다.

- 평균 전송 성공 비율 (5-25대)={98.7%, 98.5%, 97.5%}
- 평균 전송 성공 비율 (150-300대)={62.4%, 65.5%, 65.6%}
- 평균 충돌 발생 비율 (5-25대)={0.403%, 0.404%, 0.409%}
- 평균 충돌 발생 비율 (150-300대)={52.47%, 52.44%, 52.43%}

따라서, PAS-ALOHA는 learning rate를 변화시키더라도 전송 성공 비율이 크게 감소하지 않으며 충돌 발생 비율이 크게 증가하지 않음을 보였다.

### 3.2.2 음의 보상에 따른 성능 비교

본 논문에서는 PAS-ALOHA 기법의 디바이스 개수 증가에 따른 음의 보상 파라미터별 정규화된 처리량과 공정성을 검증하였다. 파라미터는 각각  $\{-0.01, -0.02\}$ ,  $\{-0.1, -0.2\}$ ,  $\{-1, -2\}$ 로 설정하였으며 성능은 다음과 같다.

- 평균 정규화된 처리량 (5-25대)={0.54, 0.509,

- 0.508}
- 평균 정규화된 처리량 (150-300대)={0.40, 0.365, 0.361}
- 평균 Jain's Fairness Index (5-25대)={0.79, 0.9057, 0.9051}
- 평균 Jain's Fairness Index (5-25대)={0.45, 0.37, 0.40}

따라서, PAS-ALOHA 기법은 음의 보상 파라미터를 {-0.01, -0.02}로 설정하였을 때 가장 높은 정규화된 처리량을 나타냈다. 이는 근소한 값인 음의 보상을 받았을 때 가장 처리량이 높다고 판단된다. 또한, Jain's Fairness Index의 파라미터도 {-0.01, -0.02}로 설정하였을 때 가장 높은 공정성을 나타냈다. 그러나, 디바이스 크게 증가하면서 공정성이 감소한다. 이는 모든 디바이스가 공평하게 데이터 패킷을 보내고 있지 않음을 나타낸다.

#### IV. 결 론

본 논문에서는 기존의 Q-learning을 이용한 Adaptive slotted ALOHA 기법에 우선순위를 적용한다. 각 디바이스가 가진 Q값을 통해 데이터 패킷의 전송 결과와 전송 채널 내 혼잡도를 학습하며 우선순위를 지정한다. PAS-ALOHA 기법은 slotted ALOHA와 비교하여 실험을 진행했다. 제안하는 기법은 기존의 slotted ALOHA 기법과 비교했을 때 디바이스 개수가 적은 환경에서는 성공 비율은 최대 70% 증가하였고 충돌 비율은 최대 29% 감소했다. 디바이스 개수가 많은 환경에서는 성공 비율은 68% 증가하였고 충돌 비율은 30% 감소했다. 따라서, 디바이스 개수가 증가할수록 성공 비율은 점차 낮아지고 충돌 비율은 높아진다.

본 논문은 5G MTC 환경 중 데이터 패킷을 전송하는 환경을 고려하였다. 따라서, 디바이스가 기하급수적으로 증가하더라도 본 기법을 통해 성공 비율과 충돌 비율이 지속해서 성능을 유지해야 한다. 또한, 디바이스 개수가 증가하더라도 모든 디바이스가 공평하게 데이터 패킷을 전송할 기회를 지속하여 가질 수 있는 연구가 필요하다.

#### References

[1] P. Popovski, K. F. Trillingsgaard, O. Simeone, and G. Durisi, "5G wireless network slicing

for eMBB, URLLC, and mMTC: A communication-theoretic view," *IEEE Access*, vol. 6, pp. 55765-55779, Sep. 2018.

(<https://doi.org/10.1109/ACCESS.2018.2872781>)

[2] C. Bockelmann, N. Pratas, H. Nikopour, K. Au, T. Svensson, C. Stefanovic, and A. Dekorsy, "Massive machine-type communications in 5G: Physical and MAC-layer solutions," *IEEE Commun. Mag.*, vol. 54, no. 9, pp. 59-65, Sep. 2016. (<https://doi.org/10.1109/MCOM.2016.7565189>)

[3] A. Laya, C. Kalalas, F. Vazquez-Gallego, L. Alonso, and J. Alonso-Zarate, "Goodbye, aloha!," *IEEE Access*, vol. 4, pp. 2029-2044, Apr. 2016. (<https://doi.org/10.1109/ACCESS.2016.2557758>)

[4] D. Vukobratovic, "Massive machine type communications and revival of ALOHA," in *Proc. INFOTEH-JAHORINA*, vol. 16, pp. 228-231, Mar. 2017.

[5] H. S. Jang, H. Lee, T. Q. Quek, and H. Shin, "Deep learning-based cellular random access framework," *IEEE Trans. Wireless Commun.*, vol. 20, no. 11, pp. 7503-7518, Nov. 2021. (<https://doi.org/10.1109/TWC.2021.3085303>)

[6] S. Sharma and X. Wang, "Collaborative distributed q-learning for RACH congestion minimization in cellular IoT networks," *IEEE Commun. Lett.*, vol. 23, no. 4, pp. 600-603, Apr. 2019. (<https://doi.org/10.1109/LCOMM.2019.2896929>)

[7] C. K. Lee and S. H. Rhee, "Performance improvement of reinforcement learning based slotted ALOHA," *J. KICS*, no. 11, pp. 1886-1892, Nov. 2020.

[8] G. M. F. Silva and T. Abrao, "Throughput and latency in the distributed q-learning random access mMTC networks" *Computer Netw.*, vol. 205, Apr. 2022. (<https://doi.org/10.1016/j.comnet.2022.108787>)

[9] J. Kim and H. Park, "Adaptive slotted ALOHA based on q-learning considering device activation states," in *Proc. Symp. KICS*, pp. 1012-1013. Jun. 2022.

[10] C. J. C. H. Watkins and P. Dayan,



“Q-learning,” *Mach. Learn.*, vol. 8, no. 3, pp. 279-292, May 1992.

- [11] R. Jain, D. Chiu, and W. Hawe, “A quantitative measure of fairness and discrimination for resource allocation in shared computer system,” *Eastern Research Laboratory, Digital Equipment Corporation, Hudson*, Sep. 1984.
- [12] A. Reddy, M. Kumari, V. Dhanwani, A. Backhaniwala, N. Kumar, and K. Vasudevan, “5G new radio key performance indicators evaluation for IMT-2020 radio interface technology,” *IEEE Access*, vol. 9, pp. 112290-112311, Jul. 2021.  
 (<https://doi.org/10.1109/ACCESS.2021.3099845>)

**유철우 (Cheolwoo You)**



1993년 : 연세대학교 전자공학과  
 학사 졸업  
 1995년 : 연세대학교 전자공학과  
 학사 졸업  
 1999년 : 연세대학교 전자공학과  
 박사 졸업  
 1999년~2003년 : LG전자 책임  
 연구원

2003년~2004년 : EoNex 책임 연구원  
 2004년~2006년 : 삼성전자 책임 연구원  
 2006년~현재 : 명지대학교 정보통신공학과 교수  
 <관심분야> 5G mobile communications system, IoT,  
 M2M, new multiple access schemes, multiple  
 antenna transmission, advanced FEC  
 [ORCID:0000-0003-3519-3490]

**김지나 (Jeena Kim)**



2022년 2월 : 명지대학교 정보통신공학과 졸업  
 2022년 2월~현재 : 명지대학교 정보통신공학과 석사과정  
 <관심분야> 5G Mobile Communication, slotted ALOHA, 강화학습

[ORCID:0000-0002-0361-4396]

**박현희 (Hyunhee Park)**



2011년 : 고려대학교 전자컴퓨터공학과 공학박사  
 2011년~2012년 : 고려대학교 정보기술사업단 연구교수  
 2012년~2014년 : 프랑스 INRIA Research Center Postdoctoral researcher

2014년~2017년 : LG전자 차세대표준연구소 선임연구원  
 2017년~2020년 : 한국성서대학교 컴퓨터소프트웨어학과 조교수  
 2020년~현재 : 명지대학교 정보통신공학과 부교수  
 <관심분야> 무선통신 표준화, 통신시스템, 데이터 분석 및 알고리즘

[ORCID:0000-0003-3810-7367]

**김병찬 (Byungchan Kim)**



2022년 2월 : 명지대학교 정보통신공학과 졸업  
 2022년 2월~현재 : 명지대학교 정보통신공학과 석사과정  
 <관심분야> 무선통신, EDCA, 이동통신, Random Access

[ORCID:0000-0001-5597-4353]