

# MAB 기반 Access Class Barring 전략 활용 LoRaWAN 혼잡 제어 기법

김재용\*<sup>1</sup>, 박은규\*<sup>2</sup>, 구철희\*\*<sup>3</sup>, 고영욱\*<sup>3</sup>

## MAB - Driven Access Class Barring Strategy for Congestion Control in LoRaWAN

Jaeyong Kim\*<sup>1</sup>, Eun-Kyu Park\*<sup>2</sup>, Cheol Hea Koo\*\*<sup>3</sup>, and Youngwook Ko\*<sup>3</sup>

본 과제(결과물)는 2025년도 교육부 및 대전광역시의 재원으로 대전RISE센터의 지원을 받아 수행된 지역혁신중심 대학지원체계(RISE)의 결과입니다(과제관리번호: 2025-RISE-06-013)

### 요 약

LoRaWAN은 ALOHA 기반 상향링크로 인해 단말 수 증가 시 충돌이 급증하여 확장성 문제가 발생한다. 본 연구는 MAB(Multi-Armed Bandit)을 이용해 ACB 파라미터(차단 확률, 대기시간)를 실시간 최적화하는 학습구조를 제안하며, Slot-based, Window-based, Dynamic Slot-based 세가지 방식의 알고리즘을 구현하였다. 시뮬레이션 결과, 저밀도 네트워크 환경에서는 기존 고정형 ACB와 유사한 패킷성공율 ASR(약 73%)을 달성하였고, 고밀도 네트워크 환경에서는 Slot-based 방식이 고정형 대비 약 20% 높은 ASR(60% 이상)을 유지하였다. 본 연구를 통해 복잡한 트래픽 추정없이 강화학습을 통한 혼잡 제어를 통해 LoRaWAN의 확장성을 효과적으로 개선할 수 있음을 보였다.

### Abstract

LoRaWAN suffers from scalability issues because its ALOHA-based uplink causes collisions to increase sharply as the number of devices grows. This study proposes a learning framework that uses a Multi-Armed Bandit (MAB) to optimize the ACB parameters (barring probability and waiting time) in real time, and implements three algorithmic variants: slot-based, window-based, and dynamic-slot schemes. Simulation results show that in low-density network scenarios, the proposed method achieves a packet Success Ratio (ASR) of about 73%, comparable to conventional fixed ACB, while in high-density networks the slot-based scheme maintains an ASR above 60%, which is roughly 20% higher than that of the fixed ACB. These results demonstrate that LoRaWAN scalability can be effectively improved through reinforcement-learning-based congestion control without requiring complex traffic estimation.

### Keywords

LoRaWAN, access class barring, ACB, multi-armed bandit, congestion control

\* 한남대학교 정보통신공학과(\*<sup>3</sup> 교신저자)  
- ORCID<sup>1</sup>: <https://orcid.org/0009-0008-4378-8332>  
- ORCID<sup>2</sup>: <https://orcid.org/0009-0008-9046-340X>  
- ORCID<sup>3</sup>: <https://orcid.org/0000-0002-0199-652X>  
\*\* 한국항공우주연구원 책임연구원  
- ORCID: <https://orcid.org/0000-0002-7180-1476>

· Received: Dec. 10, 2025, Revised: Dec. 24, 2025, Accepted: Dec. 27, 2025  
· Corresponding Author: Youngwook Ko  
Dept. of Information and Communication Engineering, Hannam University, Daejeon 34430, Korea  
Tel.: +82-42-629-7568, Email: [youngko@hnu.kr](mailto:youngko@hnu.kr)

## I. 서론

LoRaWAN은 저전력·장거리 통신기술로 스마트 시티 및 산업 자동화 등에서 널리 활용되고 있다[1]-[3]. 그러나 상향링크가 ALOHA 기반 랜덤 액세스 방식을 따르고 있어, 단말 수가 증가하면 충돌 확률이 비선형적으로 높아지는 확장성 문제가 존재한다[4]. 특히 동일 슬롯·채널·확산지수(SF, Spreading Factor)를 사용하는 패킷 간 충돌은 대규모 환경에서 심각한 전송 실패와 네트워크 혼잡을 야기한다[5].

이러한 혼잡 완화를 위해 3GPP는 ACB(Access Class Barring)을 도입하여 NB-IoT 등에서 성능을 입증했으나[6][7], 현재 LoRaWAN 표준에는 이와 같은 접근 제어 메커니즘이 부재하며 관련 연구 또한 매우 제한적이다. 한편, 고밀도 네트워크의 최적화를 위한 경량 온라인 학습 기술로 MAB(Multi-Armed Bandit)이 주목받고 있다. MAB는 복잡한 채널 정보 없이 성공·실패 결과만으로 학습이 가능하여 LoRa 자원 할당 등에 유용함이 확인되었으나[8]-[10], 이를 LoRaWAN의 ACB 파라미터(차단 확률 및 대기 시간) 최적화에 직접 적용한 사례는 없다.

이에 본 논문은 LoRaWAN 상향링크 환경을 위한 강화학습 MAB 기반 Access Class Barring 파라미터 최적화 기법을 제안하며, 주요 기여는 다음과 같다.

강화학습 기반 저지연 LoRaWAN 모델 제안: Slotted ALOHA 및 Rayleigh 페이딩 채널 환경에서 LoRa 노드의 성공률(ASR, Average Success Rate)을 측정하고, ACB 파라미터(차단 확률, 대기 시간) 조합을 행동으로 설정하여 이를 학습·최적화하는 모델을 설계한다.

- MAB 기반 ACB 파라미터 제어 기법 최초 제안: 네트워크 혼잡 상황에 효과적으로 대응하기 위해 새로운 파라미터 업데이트 전략인 Slot-based, Window-based, Dynamic Slot-based 세 가지 방식을 제안하고, 각 방식의 반응성, 안정성을 정량적으로 분석한다.

- 성능 검증: 시뮬레이션을 통해 제안된 전략들이 기존 LoRaWAN 및 고정형 ACB 대비 성능이 우수함을 검증하고, 각 업데이트 전략 간의 수렴성과 안정성 트레이드오프를 규명한다.

## II. 관련 연구

### 2.1 LoRaWAN MAC 구조와 확장성 한계

LoRaWAN은 저전력 구현을 위해 ALOHA 기반 상향링크를 채택했으나, 단말 수 증가 시 충돌 확률이 급증하여 확장성에 한계를 보인다[3]-[5]. 기존 연구들은 동일 슬롯·채널·확산지수를 사용하는 패킷 간 충돌이 필연적으로 성공률 저하와 네트워크 혼잡을 야기함을 지적하며, MAC 계층 차원의 효과적인 혼잡 제어 기법 도입 필요성을 강조하고 있다.

### 2.2 ACB 혼잡 제어

3GPP는 대규모 랜덤 액세스 혼잡 해결을 위해 ACB를 표준화하였으며[7], NB-IoT 환경에서 그 효과가 입증되었다[6]. 그러나 현재 LoRaWAN 표준에는 이와 유사한 접근 제어 메커니즘이 부재하며, ACB를 LoRaWAN 환경에 적용하거나 성능을 분석한 연구는 매우 미비한 실정이다.

### 2.3 MAB 기반 랜덤 액세스 및 LoRaWAN 적용

MAB 기술은 정보가 불완전한 환경에서 행동에 따른 보상을 학습하여 성능을 개선하는 경량 프레임워크로, 복잡한 MDP 대비 구현 부하가 적어 IoT 환경에 적합하다[8][10][11]. [8]은 단순 피드백만으로 자원 선택을 학습해 접속 성공률을 높일 수 있음을 보였다. 또한 [9]은 LoRaWAN SF 선택 문제에 MAB를 적용하여 제한된 피드백 상황에서도 성능 개선이 가능함을 검증했다. 하지만 기존 연구 [8]-[11]는 주로 셀룰러 IoT나 LoRa의 물리적 파라미터(SF, 채널) 선택에 집중되어 있다. ACB 파라미터(차단 확률, 대기 시간) 자체를 학습 대상으로 설정하여 LoRaWAN의 혼잡 제어 메커니즘으로 최적화하는 연구는 아직 충분히 다루어지지 않았다.

## III. 시스템 모델

### 3.1 LoRaWAN 상향링크 전송 모델

본 논문은  $N$ 개 LoRa 단말이 서비스 영역에서 랜덤 분포되어 단일 게이트웨이와 통신하는 LoRaWAN 상향링크 환경을 고려한다. 각 단말은 slotted ALOHA 방식 활용해 상향 패킷을 전송한다.

LoRaWAN에서 사용하는 채널 집합  $C=\{c_1, \dots, c_L\}$ 와 확산지수(SF) 집합  $S=\{s_1, \dots, s_Z\}$ 의 조합으로 자원 집합  $R=C \cup S$ 이 정의되며, 전체 자원 개수는  $M=L \cdot Z$ .

채널 집합  $C$ 는 동일 지역에서 사용할 수 있는 ISM 대역의 주파수 채널들로 구성되며, 각 채널은 서로 간 간섭 없이 병렬적으로 사용 가능하다. 확산지수 집합  $S$ 는 LoRa 물리계층에서 사용되는 Chirp Spread Spectrum 변조의 확산 정도를 나타내며, SF가 커질수록 패킷 전송 신뢰도 향상 시키지만 전송 시간이 길어지는 특성을 갖는다. 일반적으로 SF7은 약  $-7.5\text{dB}$ , SF12는 약  $-20\text{dB}$ 까지 수신이 가능할 정도로 최소 SNR 요구치가 다르며, 서로 다른 SF들은 서로 직교(Orthogonal)에 가까워 동일 채널에서도 간섭 없이 동시 수신이 가능하다.

각 슬롯 내 서로 다른 자원  $(c_i, s_j) \in R$ 을 통해 수신된 패킷이 최소 SNR 요구치를 만족하면 수신 성공으로 가정하며, 동일한 자원에 두 개 이상 패킷이 수신되는 경우 충돌로 수신 실패한다.

### 3.2 ACB 및 Cooldown 기반 충돌 제어

매 슬롯에서 각 단말은 확률  $p_{tx}$ 로 전송할 패킷이 생성된다고 가정한다. 실제 전송 여부는 게이트웨이가 전송한 ACB 파라미터  $(b, t_{acb})$ 에 의해 결정된다. 구체적으로, 단말은 난수  $r \sim U(0,1)$ 을 생성하여  $r < b$ 이면 barred되고,  $r \geq b$ 이면 해당 슬롯에서 전송을 시도한다. barred된 단말은 이후  $t_{acb}$  슬롯 동안 전송 불가능한 cooldown 상태에 진입하며, cooldown 종료 시 다시 Active 상태로 복귀하여 난수를 생성 이후의 과정을 반복한다. 이러한 단말의 time-evolution은 그림 1의 순환 구조를 갖는다.

위의 모델에 따르면 각 단말의 매 슬롯 전송 시도는 확률적으로 결정되며, 이때의 전송 시도 확률은  $\gamma(b, t_{acb})$ 로 표기한다. 이 값은 ACB 차단 확률  $b$ 와 쿨다운 지속시간의 기댓값  $E[K]$ 에 의해 결정된다.

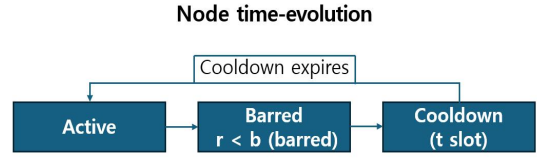


그림 1. ACB 노드 생명 주기

Fig. 1. ACB node life cycle

여기서 Active 구간의 길이를  $L$ 이라 하자. 단말은 Active 상태에서 매 슬롯 독립적으로 확률  $b$ 에 의해 차단(Barred)되므로,  $L$ 은 기하분포(Geometric distribution)를 따르며( $L \sim \text{Geom}(b)$ ), 그 기댓값은  $E[L]=1/b$ 이다. 한 사이클(Cycle)은 “쿨다운이 끝나 Active 상태로 복귀한 시점부터, 다음 차단이 발생하여 다시 쿨다운에 진입하고 끝나기 전까지”의 구간으로 정의된다. Active 슬롯에서 실제 전송이 발생할 확률은  $(1-b)$ 이므로, 한 사이클에서 발생하는 전송 시도 횟수의 기댓값은  $E[L]-1=(1-b)/b$ 가 된다. 또한 한 사이클의 전체 길이는 Active 구간과 Cooldown 구간의 합이므로  $E[L]+E[K]$ 이다. 따라서 단말의 슬롯당 평균 전송 시도 확률은 매 슬롯  $N$ 개 단말에서 전송 시도 수  $A_t$ 의 기댓값은 다음 식 (1)로 표현 가능하다.

$$\gamma(b, t_{acb}) = \frac{E[L](1-b)}{E[L] + E[K]} = \frac{1-b}{1+b \cdot t_{acb}} \quad (1)$$

### 3.3 Offered Load 정의 및 충돌 성능 분석

$$E[A_t] = N p_{tx} \gamma(b, t_{acb}) \quad (2)$$

고밀도 네트워크에서 노드-to-자원 접속 비율을 offered load라고 정의하자. 자원 수  $M$  대비  $N$ 개 노드가 접속이면, 식 (2)를 이용하여 평균 offered load per resource는 다음으로 표현 가능하다.

$$G = \frac{N p_{tx} \gamma(b, t_{acb})}{M} \quad (3)$$

대규모 단말 환경에서는 각 자원에 대한 도착 패킷 수가 Poisson( $G$ )으로 근사될 수 있다. 충돌만 고려하는 본 모델에서 자원에 단일 패킷 도착 확률은  $\geq -G$ 이다.  $(b, t_{acb})$  주어지면, 한 슬롯 수신 성공 패킷 수의 기댓값은 다음으로 표현 가능하다.

$$T_{sys}(b, t_{acb}) = E[A_t] \cdot \exp\left(-\frac{E[A_t]}{M}\right) \quad (4)$$

또한 전송 시도 대비 성공률(시스템 ASR)은

$$ASR_{sys}(b, t_{acb}) = \exp\left(-\frac{E[A_t]}{M}\right) \quad (5)$$

식 (5)에서 접속 성공률(ASR)은 부하  $G$ 의 증가에 따라 지수적으로 감소한다. 고밀도 환경에서 식 (3)에 따라 노드 수  $N$ 의 증가는  $G$ 값 증가시킨다.  $G$ 의 증가는 식 (4) 처리량이 concave 성능을 가진다.

#### IV. 강화학습 모델 및 알고리즘

알고리즘 1. Acc-greedy 기반 ACB 제어

Algorithm 1. Accelerated greedy-based ACB control

---

**Algorithm 1** Proposed Accelerated Greedy-based ACB control

---

```

1: Input & Hyperparameters:
2: Action space  $\mathcal{A} = \{(b, t_{acb}) \mid b \in \mathcal{B}, t_{acb} \in \mathcal{T}\}$ 
3:  $M$ : Total number of resources
4:  $\alpha \in (0, 1]$  (Learning rate),  $\beta > 0$  (ASR weight)
5: Definitions:
6:  $S_k, A_k$ : Number of successes and attempts at epoch  $k$ 
7:  $D_k$ : Update slot duration interval (Strategy-dependent)
8: Initialize:
9:  $Q(a) \leftarrow 0, N(a) \leftarrow 0, \forall a \in \mathcal{A}$ 

10: procedure RUN-OPTIMIZATION
11:   while Active do
12:     ▷ Step 1: Action Selection
13:     Identify untried actions
14:      $\mathcal{A}_{untried} = \{a \in \mathcal{A} \mid N(a) = 0\}$ .
15:     if  $\mathcal{A}_{untried} \neq \emptyset$  then
16:        $a_k \leftarrow \text{Random}(\mathcal{A}_{untried})$ 
17:     else
18:        $a_k \leftarrow \arg \max_{a \in \mathcal{A}} Q(a)$ 
19:     end if
20:     Broadcast parameter  $a_k$  to the node.
21:     ▷ Step 2: Observation
22:     Wait for duration  $D_k$ .
23:     Measure  $S_k$  and  $A_k$ .
24:     Compute reward  $r_k \leftarrow \sqrt{\frac{S_k}{M \cdot D_k}} \cdot \left(\frac{S_k}{A_k}\right)^\beta$ .
25:     ▷ Step 3: Update
26:      $N(a_k) \leftarrow N(a_k) + 1$ .
27:     if  $N(a_k) = 1$  then
28:        $Q(a_k) \leftarrow r_k$ .
29:     else
30:        $Q(a_k) \leftarrow Q(a_k) + \alpha \cdot (r_k - Q(a_k))$ .
31:     end if
32:   end while
33: end procedure

```

---

#### 4.1 MAB 모델링 및 보상 함수

본 논문은 LoRaWAN 상향링크의 혼잡 제어 문제를 비정상(Non-stationary) 환경에서의 MAB 최적화 문제로 정의한다. 에이전트인 네트워크 서버는 매 시점  $t$ 마다 행동 집합  $\mathcal{A}$ 에서 차단 확률과 대기 시간의 이산적 조합을 선택하여 수행하고, 일정 관측 슬롯이 경과한 후 해당 행동에 대한 보상을 측정한다. 알고리즘 1에 21번째 줄에 정의된 보상 함수는 시스템의 데이터 처리량(Throughput)과 접속 신뢰성(Reliability)을 동시에 고려하여 설계되었다. 특히, 성공률(ASR) 항에는 선형 가중치가 아닌 지수 가중치( $\beta$ , Exponential Weight)를 적용하였다. 이는 배터리 용량이 제한적인 LoRaWAN 단말 특성상, 재전송으로 인한 에너지 소모가 네트워크 수명에 치명적이기 때문이다. 따라서  $\beta > 1$  (본 논문에서는  $\beta=4$ )로 설정하여 성공률이 소폭 감소하더라도 보상값이 급격히 하락하는 비선형 페널티(Steep penalty) 구조를 적용했다. 이는 에이전트가 단순히 처리량은 높지만 충돌이 빈번한 ‘불안정 상태’를 최적점으로 오판하는 것을 방지하고, 높은 성공률이 보장되는 파라미터 영역 내에서만 탐색을 수행하도록 유도하는 소프트 게이팅(Soft gating) 역할을 수행한다.

#### 4.2 정적/동적 관측 전략 ACB 및 보상

알고리즘 1의 관측 및 업데이트 단계에서 관측 기간  $D_k$ 을 사용하는 세 가지 경우를 제안한다: slotted-, windowed-, and dynamic-strategy. 각 경우는 수집 데이터의 범위와 행동 빈도에 영향을 준다.

(1) Slot-based Strategy: 가장 세밀한 단위의 업데이트 방식이다. 매 슬롯마다( $D_k = 1$ ) 즉시 행동에 따른 보상을 계산하고 행동의 Q-값을 갱신한다. 매 슬롯의 성공/실패 여부를 관측하여 다음 슬롯의 파라미터를 결정한다.

(2) Window-based Strategy: 사전 정의된 고정 길이  $W$ 를 관측 주기로 사용한다( $D_k = W$ ). 에이전트는  $W$ 개의 슬롯 동안 행동 파라미터를 유지하며 데이터를 누적하고, 관측 기간이 종료되는 시점에 관측 구간 전체의 평균 처리량과 성공률을 기반으로 행동 보상 값을 계산한다.

(3) Dynamic(-Slotted) Strategy: Epoch  $k$ 에서 행동의 차단 시간  $t_{acb}$ 를 관측 기간으로 설정한다 ( $D_k = t_{acb}$ ). 즉, Epoch  $k$ 에서 특정 행동  $(b, t_{acb})$ 이 선택되면, 해당  $D_k = t_{acb}$  슬롯 기간이 경과 될 때까지 대기하며 데이터 수집한 후 보상을 측정하여 학습을 진행한다. 이는 행동 제어 주기와 학습 알고리즘의 갱신 주기를 동기화하는 방식이다.

### V. 실험 결과 및 분석

본 장에서는 제안된 가속-Greedy 기반 ACB 제어 알고리즘 활용하여 저지연 LoRaWAN 랜덤 접속 성능을 컴퓨터 모의실험 하였다. ACB 행동 학습 알고리즘은 세 가지 정적/동적 관측 전략을 사용한다. 주요 시스템 성능 지표로 접속 성공률(ASR), 처리량(Throughput)을 분석하였다.

성능 비교를 위해 3가지 LoRa 랜덤 접속 기술을 실험하였다: (1) NO ACB: ACB 없는 기존 LoRa 랜덤 접속; (2) Fixed-ACB: (임의의  $N = 30$  환경에) 정적으로 튜닝된 파라미터 ( $b=0.45, t_{acb}=8$ )를 사용하는 Fixed-ACB 기반 LoRa 랜덤 접속; (3) Dynamic-ACB: 제안된 MAB 기반 ACB와 slotted-, windowed-, dynamic-strategy 전략 활용하는 LoRa 랜덤 접속을 실험하였다. 시뮬레이션 설정은 표 1과 같다.

표 1. 실험 파라미터 세트

Table 1. Experimental parameter set

Parameter	Value
$N$	30, 90
$P_{tx}$	0.8
$M$	18
Slot	2000
Avg SNR	10 dB
$\alpha$	0.1
$b, t_{acb}$	0.45, 8
scenario	10

#### 5.1 Offered Load에 따른 성능 이론 분석

그림 2는 이론적인 ASR과 처리량의 Offered load ( $G$  in node packets per resource)에 따른 상관관계 곡선을 보여준다. 그림 2의 높은  $G$  범위에서 급격한

성능 감소 보임. 예로, 처리량은  $G = 1.0$ 에서 최대화되지만, 이때 ASR은 36%에 불과하다. 이는 재전송 비용을 증가시킨다.

이후 모의실험에서는 목표 ASR = 60%-70% 수준을 요구하는 시스템을 고려한다. 이 요구를 유지하기 위해서는 ACB 파라미터의 적절한 제어가 필요하여  $G \approx 0.3-0.4$  구간 ‘최적 운영 영역’-을 유지해야 한다. 이후 실험 시나리오는 네트워크 노드 수  $N$ 이 증가해도  $G$ 를 최적 운영 영역으로 유지하기 위한 최적의 ACB 행동 전략 탐색을 실험한다.

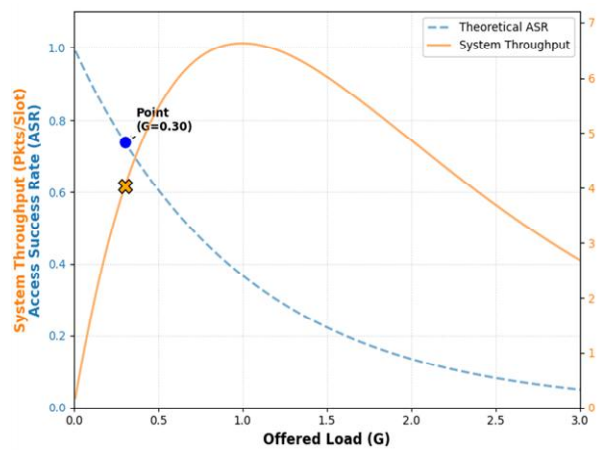


그림 2. Offered Load에 따른 ASR, Throughput 곡선

Fig. 2. ASR and Throughput curves versus offered loads: As  $G$  grows, congestion risk increases

#### 5.2 Scenario 1: 중간부하 ( $N=30$ ) 환경 성능

첫 번째 시나리오는 네트워크 자원 대비 moderate-dense 단말 수  $N=30$  환경이다. 표 2는 해당 환경에서의 성능 지표를 요약한 것이다.

Fixed ACB의 안정성: 시스템 요구(목표 ASR = 60-70%)에 맞춰 튜닝된 Fixed ACB( $b=0.45$ ) 활용한 시뮬레이션은 ASR 0.7201(약 72%)을 기록하며 가장 우수한 성능을 보였다. 이는 설정된 ACB 파라미터가 트래픽 부하를 최적 운영 영역( $G \approx 0.3$ ) 내로 적절히 제어했기 때문이다. Proposed(slotted-, windowed-, dynamic-strategy) MAB의 수렴: 제안하는 Slot-based MAB 기법은 초기 탐색(Exploration) 과정으로 인해 초반 진동이 있었으나, 학습이 진행됨에 따라 ASR 0.7149로 Fixed ACB와 유사한 성능에 도달하였다.

표 2. 시뮬레이션 결과 표 - 네트워크 성능에 대한 노드 수( $N$ )의 영향Table 2. Simulation results table - Impact of the number of nodes ( $N$ ) on network performance

Node	Mode	ASR	Throughput (pps)	Collision rate(pps)	Avg $b$	Avg $t_{acb}$
$N=30$	No-ACB	0.2657	6.3739	0.7326	N/A	N/A
	Fixed-ACB	0.7201	2.9358	0.2749	0.45	8
	Slot-based ACB	0.7149	2.1889	0.2799	0.4228	13.45
	Window-based ACB	0.5397	1.7365	0.4565	0.4186	15.88
	Dynamic Slot ACB	0.4740	1.4213	0.5229	0.2488	38.74
$N=90$	No-ACB	0.0814	3.6626	0.9181	N/A	N/A
	Fixed-ACB	0.4545	6.2072	0.5414	0.45	8
	Slot-based ACB	0.6285	3.7033	0.3674	0.5443	19.64
	Window-based ACB	0.4445	3.7377	0.5528	0.5160	13.93
	Dynamic Slot ACB	0.3980	3.2497	0.5993	0.3274	38.98

이는 기지국이 노드 수  $N=30$  정보를 모르는 상태(Unknown state)에서도, 보상 기반 학습을 통하여 최적의 제어 파라미터를 스스로 찾아냈음을 의미한다. 마지막, No-ACB의 제어가 없는 경우 ASR은 0.2657(약 26%)에 불과하여,  $N=30$  수준에서도 심각한 충돌이 발생하고 있음을 확인하였다.

### 5.3 Scenario 2: 고부하 ( $N=90$ ) 환경 적응성

다음 시나리오는 단말 수가 3배로 급증한 과부하( $N=90$ ) 환경이다. 이 실험은 고정된 파라미터 방식의 한계와 제안 기법의 적응성(Adaptability)을 검증하는 핵심 시나리오이다. 성능 값은 표 2에 나타난다.

Fixed ACB의 성능 감소(Lack of adaptability): 단말 수 3배( $30 \rightarrow 90$ ) 증가 경우, Fixed ACB는 기존 파라미터( $b=0.45$ )를 유지하였다. 식 (3)에 의해  $N$  증가는 부하  $G$ 의 증가로 이어졌으며, 그 결과 ASR은 0.7201(when  $N=30$ )  $\rightarrow$  0.4545(when  $N=90$ )로 급락하였다. 충돌률 또한 0.2749 pps에서 0.5414 pps로 2배 가까이 증가로 네트워크 신뢰성이 감소하였다.

반면, 제안된 Slot-based ACB 기법은 충돌 증가에 따른 보상 하락을 매 슬롯 감지하였다. 에이전트는 더 높은 차단 확률(Avg  $b \approx 0.5443$ )을 스스로 학습하여 적용하고, 그 결과 ASR 0.6285(약 63%) 달성. Fixed ACB(45%) 대비 약 20% 높은 성공률이며, 성능 감소폭 최소화하여 고밀도 트래픽 변동 상황에서 네트워크 성능의 성공적 방어를 입증한다.

## VI. 결론

본 논문에서는 LoRaWAN 상향링크 혼잡을 효과적으로 제어하기 위해 Acc-Greedy 기반 MAB-ACB 최적화 기법을 제안하였다. 제안 기법은 처리량과 접속 성공률(ASR)을 동시에 고려한 보상 함수를 통해 네트워크 신뢰성을 보장하는 최적 파라미터를 스스로 학습한다. 시뮬레이션 결과, 광범위한 고밀도 환경에서 제안 기법이 기존 LoRaWAN 상향링크 대비 우수 성능 보임. 특히, 제안 기법은 과부하( $N=90$ ) 경우 파라미터를 동적으로 조정하여 Fixed ACB 대비 약 20% 높은 ASR(60% 이상)을 달성하며 우수한 적응성과 강건성을 입증하였다. 결론적으로, 본 연구는 복잡한 트래픽 추정 없이도 강화학습을 통해 LoRaWAN의 확장성 한계를 극복할 수 있음을 보였으며, 대규모 IoT 환경의 신뢰성 보장을 위한 핵심 기술로 활용될 수 있다.

향후 연구에서는 기존 MAB 기반 ACB 혼잡 제어를 확장하여 Q-러닝 및 고도화된 강화학습 기반으로 Offered Load를 학습 최적화하고, 트래픽 변동이 큰 고밀도 환경에서도 ASR을 목표 수준으로 안정적으로 유지할 수 있는 적응형 혼잡 제어 기법으로 발전시킬 계획이다.

## References

- [1] E. Seo and J. Jang, "Analysis of transmission rate

- according to LoRaWAN communication distance", The Journal of IIBC, Vol. 19, No. 3, pp. 207-211, Jun. 2019. <https://doi.org/10.7236/IIBC.2019.19.3.207>.
- [2] S. Mah and B. Kim, "LoRa Technology Analysis and LoRa Use Case Analysis By Country," The Journal of IIBC, Vol. 19, No. 1, pp. 15-20, Feb. 2019. <https://doi.org/10.7236/IIBC.2019.19.1.15>.
- [3] S. Devalal and A. Karthikeyan, "LoRa Technology - An Overview", 2018 Second International Conference on Electronics, Communication and Aerospace Technology (ICECA), Coimbatore, India, pp. 284-290, Mar. 2018. <https://doi.org/10.1109/ICECA.2018.8474715>.
- [4] F. Adelantado, et al., "Understanding the limits of LoRaWAN", IEEE Commun. Magazine, Vol. 55, No. 9, pp. 34-40, Sep. 2017. <https://doi.org/10.1109/MCOM.2017.1600613>.
- [5] G. H. Park, et al., "Energy-efficient ADR mechanism considering collision probability in LoRa networks", J. of KIISE, Vol. 47, No. 12, pp. 1464-1473, Dec. 2020. <https://doi.org/10.5626/KTCP.2020.26.12.535>.
- [6] C. Ramineni, et al., "Impact of access class barring in narrow band Internet of Things", IET Wireless Sen. Sys., Vol. 9, No. 6, pp. 407-414, Dec. 2019. <https://doi.org/10.1049/iet-wss.2019.0036>.
- [7] 3GPP, "TS 22.011: Service Accessibility, Release 13", 3GPP, Dec. 2016.
- [6] N. Jiang, et al., "A decoupled learning strategy for massive access optimization in cellular IoT networks", IEEE J. on Selec. Areas in Comm., Vol. 39, No. 3, pp. 668-685, Mar. 2021. <https://doi.org/10.1109/JSAC.2020.3018806>.
- [9] R. E. Navas, et al., "Energy-aware spreading factor selection in LoRaWAN using delayed-feedback bandits", Proc. IFIP Networking, Barcelona, Spain, Jun. 2023. <https://doi.org/10.23919/IFIPNetworking57963.2023.10186444>.
- [10] X. Wu, Y. Ko, and A. Tyrrell, "Distributed

multi-agent reinforcement learning for heterogeneous NOMA-ALOHA systems", IEEE Trans. Cogn. Comm and Net., Vol. 11, No. 3, pp. 1902-1912, Jun. 2025. <https://doi.org/10.1109/TCCN.2024.3474709>.

- [11] Y. Ko and J. Choi, "Reinforcement learning for NOMA-ALOHA under fading", IEEE Trans. Commun., Vol. 70, No. 10, pp. 6861-6873, Oct. 2022. <https://doi.org/10.1109/TCOMM.2022.3198125>.

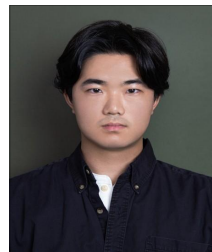
## 저자소개

김재용 (Jeayong Kim)



2025년 3월 ~ 현재 : 한남대학교  
정보통신공학과 학부연구원  
관심분야 : 머신러닝, 지능형 네트워크, 강화학습, LoRaWAN

박은규 (Eun-Kyu Park)



2025년 3월 ~ 현재 : 한남대학교  
정보통신공학과 학부연구원  
관심분야 : 머신러닝, 초밀도 네트워크, 분산 강화학습, LoRaWA

구철회 (Cheol Hea Koo)



1997년 2월 : 충남대학교 전자공학과(공학사)  
1999년 2월 : 충남대학교 의용전자공학(공학석사)  
2021년 2월 : 충남대학교 컴퓨터공학과(공학박사)  
2002년 3월 ~ 현재 : 한국항공우주

연구원 책임연구원  
관심분야 : 내장형 소프트웨어, 위성 통신, 심우주통신, CFDP, DTN

고 영 욱 (Youngwook Ko)



1999년 2월 : 한남대학교 정보통신  
공학과(공학사)

2002년 12월 : 아리조나주립대학교  
전기공학과(공학석사)

2006년 12월 : 아리조나주립대학교  
전기공학과(공학박사)

2025년 3월 ~ 현재 : 한남대학교

정보통신공학과 부교수

관심분야 : DTN, UDN, 6G TN/NTN, 자율 지능형 Ad-  
hoc 네트워크, 위성 IoT