

# 개인화된 무선 인체 영역 네트워크에서 Q-learning을 이용한 적응적 스케줄링 기법

조승현\*, 김범수\*\*

## Q-learning based Adaptive Scheduling Method In Personalized Wireless Body Area Network

Seoung-Hyeon Jo\*, Beom-Su Kim\*\*

이 연구는 2023년도 경상국립대학교 교내연구비 및 정부(과학기술정보통신부)의 재원으로 한국연구재단의 지원을 받아 수행된 연구임(No. RS-2024-00333695)

### 요약

무선 인체 영역 네트워크에서는 사용자 움직임에 의한 노드 간 가시성(Line of sight) 미확보 시 채널 페이딩으로 인해 빈번한 데이터 손실이 발생한다. 이를 해결하기 위한 다양한 링크 품질 인지 기반의 적응적 스케줄링 기법들이 제안되었으나 이들은 사용자의 개인화된 움직임을 스케줄링에 반영하지 않는다는 문제가 있다. 본 연구에서는 기기 사용자의 개인화된 움직임 패턴과 그에 따른 링크 품질 변화를 Q-learning을 통해 학습한 뒤 스케줄링에 통합하는 적응적 스케줄링 기법을 제안한다. BANSIM 네트워크 시뮬레이터에서의 시뮬레이션을 통해 사용자가 규칙적인 움직임을 보이는 경우 제안하는 기법이 기존 기법 대비 최대 41.8% 더 나은 패킷 전송 성공률을 보이는 것을 검증하였다.

### Abstract

In Wireless Body Area Networks(WBANs), frequent data loss occurs due to channel fading when the line of sight between nodes is lost due to user movement. Various link quality-aware adaptive scheduling techniques have been proposed to solve this problem, but they suffer from the problem that they do not reflect the personalized movements of users in the scheduling. In this paper, we propose an adaptive scheduling technique that learns the personalized movement patterns of WBAN device users and the resulting link quality changes through Q-learning and incorporates them into the scheduling. Through simulations on the BANSIM network simulator, we verify that our proposed method achieves up to 41.8% better packet delivery success rate than existing methods when users have regular movements.

### Keywords

wireless body area networks, WBANs, TDMA, personalized adaptive scheduling, Q-learning

\* 경상국립대학교 컴퓨터공학과 학사과정

- ORCID: <https://orcid.org/0009-0003-1940-3504>

\*\* 경상국립대학교 컴퓨터공학과 교수(교신저자)

- ORCID: <https://orcid.org/0000-0001-8035-1967>

· Received: May 04, 2024, Revised: May 27, 2024, Accepted: May 30, 2024

· Corresponding Author: Beom-Su Kim

Dept. of Computer Science, Gyeongsang National University, Korea

Tel.: +82-55-772-1383, Email: [bbk5391@gnu.ac.kr](mailto:bbk5391@gnu.ac.kr)

## I. 서 론

최근 고령화로 인한 환자 증가, 웨어러블 시장 및 의료 기기 산업의 성장으로 인체 영역 네트워크인 WBAN(Wireless Body Area Network)이 널리 사용되고 있다. WBAN은 인체 내/외부에 삽입된 생체 신호 센서와 코디네이터 간의 무선통신을 위해 개발된 3m 범위의 초 근거리 네트워크로서 저지연, 고신뢰, 저전력 무선통신을 지원한다. WBAN은 1개의 코디네이터와 센서 노드들이 원 홉 스타 토폴로지 형태로 연결되며, 코디네이터는 센서 노드들이 전송하는 데이터를 수집한 뒤 외부망(Wi-Fi, 5G 등)을 통해 원격 의료 센터로 전송하는 2-tier 아키텍처로 구성된다[1].

WBAN과 같은 낮은 전력 소모와 높은 신뢰성을 요구하는 네트워크[2]에 적용할 수 있는 스케줄링 기법으로, 높은 안정성과 우수한 에너지 효율을 보이는 TDMA(Time Division Multiple Access) 기법을 적용하는 사례가 많다[3]-[5]. TDMA 방식 스케줄링 기법의 과제는 각 타임 슬롯에 할당하는 노드의 순서를 결정하는 문제로, Q-learning을 비롯한 강화 학습을 도입하는 연구가 진행되고 있다[3][5]. WBAN 표준에서는 심전도, 뇌전도 등과 같은 중요 생체 신호 데이터의 서비스 품질 보장을 위해 중요 데이터의 우선적 스케줄링을 지원하고 있다. 데이터의 중요도에 따라 센서 노드의 우선순위를 0부터 7까지로 구분하고, 높은 우선순위를 가지는 노드에게 더 많은 타임 슬롯을 할당함으로써 우선순위에 따른 차등적인 서비스 품질을 제공하고 있다. 그러나, 이 기법은 코디네이터와 센서 노드 간 링크 품질 변화를 스케줄링에 고려하지 않으므로 신체 움직임에 따라 채널 품질 변화가 발생하여 패킷 손실이 발생한다는 문제가 존재한다[6].

요약하면, WBAN 환경에서 TDMA의 타임 슬롯 할당 문제를 다른 스케줄링 기법들이 최근 제안되었으나[8], 이들은 단순히 최근 수신한 데이터 패킷의 수신신호강도(Received signal strength) 및 우선순위가 높은 센서 노드에게 타임 슬롯을 우선 할당하는 것에 그치고 있다. 이러한 우선순위 또는 신호세기 기반의 적응적 타임 슬롯 할당 기법은 신체의 움직임이 빈번히 변하는 상황에서 환경 변화에 선

제적으로 대응하지 못한다는 문제가 있다[6][7].

이러한 문제를 해결하기 위해, 본 연구에서는 WBAN 사용자가 개인화된 생활 패턴에 따라 앉기, 일어나기, 걷기 등 주기적인 신체 움직임을 보이는 점에 착안하여[9], 개인화된 움직임 패턴에 따른 링크 품질 변화를 예측하고 이를 타임 슬롯 스케줄링에 반영하여 네트워크의 서비스 품질 개선 시키고자 한다. 구체적으로, 사용자의 개인화된 움직임 패턴에 따른 링크 품질 변화를 학습하기 위해, 매 비콘 구간마다 시행착오(Trial and error)를 통해 각 소스에 대한 전송 성공 여부에 따른 누적 보상을 Q-learning을 통해 계산한다. 이후 도출된 Q-table로부터 각 타임 슬롯에서 가장 높은 누적 보상을 가지는 소스를 우선 스케줄링하여, 신체 움직임에 따른 채널 페이딩 영향을 최소화하고자 한다.

## II. 관련 연구

WBAN은 실시간 의료 모니터링 시스템을 위한 초근거리 네트워크로서[1], 저지연, 고신뢰, 저전력 무선통신을 요구한다. 이를 만족시키기 위해 데이터 전송량 감쇄, 스케줄링 분야에서 다양한 연구들이 진행되었다. [2]에서는 WBAN의 에너지 효율을 높이기 위한 인체 데이터에 최적화된 압축 기법을 연구하였으며, LEC 등의 다른 압축 기법과 비교하여 제안하는 기법이 다른 기법보다 최대 45% 더 나은 압축률을 보임을 입증하였다. 한편 WBAN의 고신뢰, 저전력 요구사항을 달성하기 위해 TDMA 기반의 다양한 적응적 스케줄링 기법들이 제안되었다. TDMA 스케줄링 기법의 주요 과제는 타임 슬롯에 할당하는 노드의 순서를 결정하는 것이다. [3]에서는 IoT 환경, 센서 네트워크 등을 대상으로 QoS를 보장하기 위해, 패킷의 deadline, delay 정보를 바탕으로 DDQN을 활용한 강화학습 기반 타임 슬롯 할당 기법을 제안하였으며 기존 Strict Priority, Weighted Round Robin 스케줄링 기법 대비 deadline 내 패킷 도착 확률을 약 10% 향상시켰다. [4]에서 TDMA 방식 스케줄링 기법은 NP 문제에 속하는 주요 연구 분야이며, 강화 학습 기반 TDMA MAC 프로토콜 도입의 주목적은 에너지 소모율 감소와 패킷 전송률(Throughput) 향상에 있음을 보였다.

[5]에서는 5G MTC 환경에서 다수의 디바이스가 접속하는 경우 패킷 충돌 발생률을 최소화하기 위해, Q-learning을 활용한 적응적 Slotted ALOHA 기법을 제안하였다. 제안된 기법은 디바이스가 많은 환경에서 기존 기법 대비 충돌 비율을 약 30% 감소시켰다. [6]에서는 WBAN 환경에서 노드 간 가시성(LOS(Line Of Sight)), 비가시성(NLOS(Non-LOS))으로 인한 채널 페이딩 여부를 딥러닝 기반으로 분류하는 연구를 진행하였으며, 실험 시나리오와 연구 결과로부터 채널 페이딩은 수신 신호 강도에 부정적인 영향을 미친다는 것을 알 수 있다. [7]에서는 WBAN MAC의 주요 과제 중 하나로 착용자의 움직임으로 인한 토폴로지 이탈, 성능 저하를 제시하였다. [8]에서는 WBAN에서 에너지 소모를 최소화하기 위해 TDMA 기반 스케줄링 기법을 적용하였으며, 적용 전 대비 네트워크의 에너지 소모, 패킷 전송률 측면에서 우수한 성능을 보인다는 것을 입증하였다. [9]에서는 SVM, K-means clustering을 이용하여 WBAN 기기 착용자의 자세를 앉은 자세, 걷는 자세 등의 6가지 자세로 분류하는 모델을 개발하였다. 인체 센서의 에너지 소모를 최소화하기 위해, [10]에서는 WBAN 환경에서 거의 동일한 비콘 신호가 주기적으로 전송된다는 특성에 착안하여 비콘 신호를 선택적으로 수신하는 기법을 사용하였다. 그 결과 일반 비콘 모드 대비 전력 소모를 약 25% 감소시켰다. [11]에서는 다수의 센서 노드로 구성된 IoT 네트워크에서 패킷 전송 지연을 줄이기 위해, 데이터 도착률이 상한 및 하한값을 벗어나는 경우 Q-table을 초기화하는 Q-learning 기반 타임 슬

롯 스케줄링 기법을 제안하였다. 제안된 기법은 환경 변화에 적응하는 데 소요되는 타임 슬롯 수를 기존 Q-learning 대비 50% 감소시켰다.

### III. 시스템 모델 및 문제 정의

제안하는 시스템은 각 노드가 양쪽 팔꿈치, 양쪽 손목, 양쪽 무릎, 양쪽 발목 위치에 부착되어 가슴 앞쪽에 부착된 하나의 코디네이터 디바이스와 주기적으로 통신하는 star topology 형태로 구성된다. 착용자는 사전 정의된 모빌리티 모델에 따라 주기를 가지는 행동을 반복적으로 수행하며, 코디네이터는 주기적으로 비콘 신호를 전송하고, 각 노드는 수신한 비콘 신호를 기반으로 할당받은 시간 동안 데이터를 전송한다.

그림 1과 같이, WBAN 기기 착용자는 지속적으로 양쪽 팔, 다리를 서로 교차시키는 자세를 반복하며 걷는 동작을 수행한다. 따라서 노드가 부착된 신체 부위에서 지속적인 움직임이 발생하며, 코디네이터(주황색)와 노드(회색) 간의 가시성(LOS) 확보 여부에 따라 상호 간 통신 품질은 변한다. 기존 링크 품질 인지 기반 스케줄링 정책을 따르면, 각 소스로부터 마지막으로 수신한 패킷의 수신 신호 강도가 가장 높은 순으로 스케줄링을 수행하게 된다. 하지만 신호 품질이 실시간으로 변화하는 상황에서는 LOS 확보 여부에 따른 채널 페이딩이 빈번하게 발생하기 때문에[6], 마지막으로 수신한 패킷의 수신 신호 강도는 더 이상 유효하지 않을 가능성이 높다.

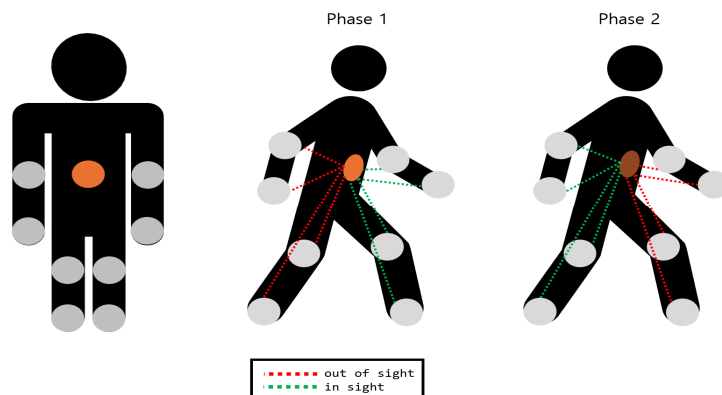


그림 1. WBAN 시스템 환경 모식도  
Fig. 1. Envisioned WBAN system model

본 연구에서는 이러한 개인화된 신체 움직임이 주기적으로 발생한다는 점 및 LOS 확보 여부가 노드와 코디네이터 간 링크 품질에 영향을 준다는 점에 착안하여, Q-learning을 통해 링크 품질 변화를 학습함으로써 신체 움직임 및 주기 변화에 선제적으로 대응할 수 있는 적응적 스케줄링 기법을 제안한다.

#### IV. 제안하는 스케줄링 기법

제안하는 기법의 핵심 아이디어는 신체 움직임에 따른 채널 페이딩의 영향을 최소화하기 위해 미래 신체 움직임을 예측하여 스케줄링을 수행하는 것이다. 이를 위해선 개인화된 WBAN 사용자의 움직임이 주기성을 가진다는 점을 이용한다. 그림 2는 본 연구에서 제안하는 스케줄링 기법의 예를 설명하며, 사용자가 현재 취하고 있는 움직임의 현재 상태와 주기 정보를 Q-learning을 통해 학습한다. 이후 도출된 Q-table을 통해 각 타임 슬롯에서 가장 높은 누적 보상을 가지는 소스를 우선 스케줄링 함으로써 신체 움직임에 따른 채널 페이딩 영향을 최소화할 수 있다.

사용자의 개인화된 움직임 패턴에 따른 링크 품질 변화를 학습하기 위해, 매 비콘 구간마다 각 소스에 대한 전송 성공 여부에 따른 누적 보상을 Q-learning을 통해 계산한다. 본 스케줄링 문제를 Q-learning 문제로 모델링하기 위한 상태 공간(State space), 행동(Action), 보상(Reward)은 다음과 같이 정의할 수 있다.

상태 공간에 사용자의 움직임 패턴을 반영하기 위해,  $t$  시점에서의 상태  $s_t$ 를 식 (1)과 같이 정의할

수 있다.

$$s_t = (\text{phase}_t, \text{slot}_i) \tag{1}$$

$\text{phase}_t$ 는 움직임 주기에 대한 인코딩된 정보이며,  $\text{slot}_i$ 는 현재 상태에서 할당해야 할 타임 슬롯이다. 에이전트는 epsilon-greedy 전략으로 탐험(Exploration)과 활용(Exploitation)을 확률적으로 선택하여 학습을 진행한다.

행동은 상태 공간에 정의된 타임 슬롯  $\text{slot}_i$ 에 탐색한 노드를 할당하는 것으로 식 (2)와 같이 정의하였다.

$$a_t = \text{slot}_i \leftarrow n_k \tag{2}$$

보상은  $\text{slot}_i$ 에 노드  $n_k$ 을 할당하여,  $n_k$ 에서 코디네이터로 20B의 패킷을 전송하였을 때, 노드  $n_k$ 의 우선순위 값인  $p_i$ 로 정의하였다. 이때 전송을 성공적으로 마치고 ACK 메시지 수신까지 완료한 경우 양의 값을, ACK 메시지 수신을 받지 못한 경우 음의 값을 취한다. 식 (3)은 보상 매커니즘을 정리한 것이다.

$$r_t = \begin{cases} 1 \times p_i & (\text{success}) \\ -1 \times p_i & (\text{fail}) \end{cases} \tag{3}$$

Q-table은  $t$  시점에서의 상태  $s_t$ 와,  $s_t$ 에서 취한 행동인  $a_t$ 의 순서쌍 ( $s_t, a_t$ )를 키로 두는 해시테이블 구조로 정의하였다.

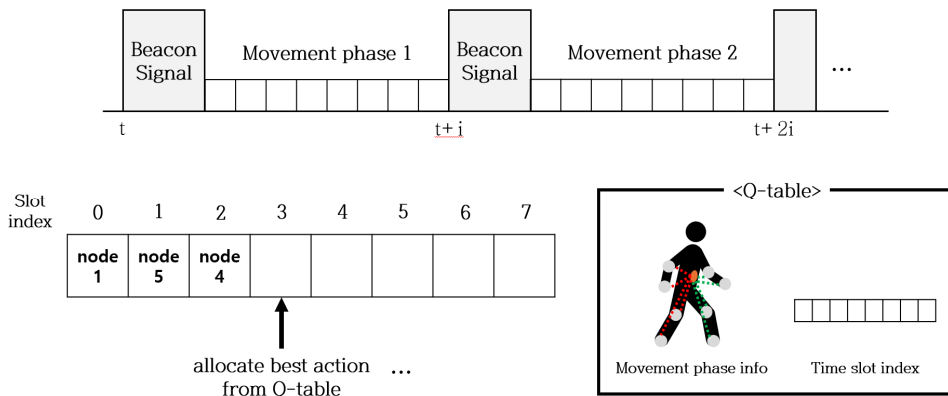


그림 2. 제안하는 스케줄링 기법의 예  
Fig. 2. Example of Q-learning based scheduling

학습은 비콘 인터벌 타임아웃 시 코디네이터가 저장하는 통신 성공 여부 및 패킷 전송률 데이터를 기반으로 진행된다. 즉 새 비콘 신호를 보내기 전 각 노드별 통신 성공 여부를 기반으로 보상을 산출하여 Q-table에 기록한 다음, 후술할 알고리즘을 동작시켜 다음 비콘 신호에 반영할 타임 슬롯 스케줄링을 진행하는 방식으로 동작한다.

알고리즘 1. 제안하는 스케줄링 알고리즘

Alg. 1. Proposed scheduling algorithm

---

Algorithm 1 Priority-based scheduling

---

Initialization:

1:  $current\_state \leftarrow S_0$

Function: Choose\_Action(current\_state)

1:  $r_i \leftarrow$  random value in range 0, 1

2: if  $r_i < exploration\_rate$  then

3:     return not\_selected\_random\_action

4: else

5:     return  $argmax(Q\_table[current\_state])$

Function: Time\_Slot\_Allocation()

1: for each slot  $t$  do

2:      $current\_state \leftarrow S_{i+1}$

3:      $t \leftarrow$  Choose\_Action(current\_state)

4: end for

5:  $exploration\_rate \leftarrow 0.0005$

---

위 학습 과정을 통해 각 비콘 구간에서 타임 슬롯별 소스 노드의 Q-value를 도출할 수 있으며, 학습이 완료되면 코디네이터는 각 움직임 단계별 타임 슬롯 할당 시 가장 높은 누적 보상을 가지는 소스를 우선 스케줄링한다. 알고리즘 1은 제안하는 스케줄링 기법을 pseudo code 형태로 표현한 것이다.

## V. 성능 평가

제안하는 스케줄링 기법의 성능 평가를 위해, Python 기반의 네트워크 시뮬레이터인 BANSIM[12]을 사용하였으며 정적/동적 네트워크 시나리오 하에서 우선순위 기반의 TDMA 스케줄링 기법과 성능을 비교 분석하였다.

## 5.1 실험 환경 정의

부착된 노드 디바이스는 비콘 신호 인터벌 기간 동안 하나의 새 데이터를 샘플링 및 전송 대기 큐에 저장하여, 비콘 신호 수신 후 할당받은 타임 슬롯 기간동안 반드시 하나의 데이터를 코디네이터로 전송한다. 이때 각 전송은 수신자로부터의 ACK 메시지 수신을 포함한다.

실험 환경은 착용자의 가슴 앞쪽에 부착된 1개의 코디네이터(에이전트)와 8개의 노드로 구성되며, 8개의 노드는 각각 양쪽 팔꿈치, 양쪽 손목, 양쪽 무릎, 양쪽 발목에 부착되어 있다. 착용자는 시뮬레이션 시작과 동시에 걷는 동작을 진행하며, 코디네이터는 현재 동작 패턴의 페이즈를 식별한 다음 Q-learning을 통해 노드별 최적의 타임 슬롯 기간을 선택 및 할당한다. 복잡성을 줄이기 위해 타임 슬롯의 수는 노드의 수와 같은 8개로 고정하였으며, 각 노드는 타임 슬롯 할당 기간에 반드시 데이터를 전송하는 상황을 가정하였다. 시뮬레이션 파라미터는 시뮬레이션 시간, 착용자의 움직임 주기 간격으로 두었으며 실험값은 표 1과 같다. Q-learning 파라미터인 학습률, 할인율은 각 시행에서 가장 높은 성능을 보인 값을 기준으로 두었다.

본 연구에서 제안하는 기법을 테스트하기 위한 시나리오는 WBAN 기기 착용자가 규칙적인 움직임을 보이는 경우를 대표하는 시나리오 1과 무작위 움직임을 보이는 시나리오 2의 2가지로 두었다.

표 1. 시뮬레이션 파라미터

Table 1. Simulation parameter

	range	description
simulation time(s)	600, 1800, 3600	10 minutes, 30 minutes, 60 minutes
scenario	1. walking, 2. random movement	1. walking: regular movement with repetition of phase 1 and phase 2  2. random movement: causes random channel fading

시나리오 1은 그림 1과 같은 페이지를 반복하며 걷는 자세를 취한다. 따라서 코디네이터 디바이스와 통신이 불가능한 노드가 규칙적인 패턴을 띠며 발생한다. 시나리오 2에서 WBAN 기기 착용자가 취하는 불규칙한 움직임은 별도의 주기를 갖지 않으며, 따라서 무작위로 통신 불가능한 노드가 발생하는 경우를 상정하였다. 시뮬레이션 시간은 신체 검진이나 운동량 측정에 사용되는 WBAN 기기의 특성을 고려하여 600초(10분), 1800초(30분), 3600초(1시간)로 구분하였다. 착용자는 시뮬레이션 시작과 동시에 각 시나리오별 자세를 취하며, 시뮬레이션 종료 시까지 반복한다. 비콘 신호는 0.5초를 주기로 전송되며, 비콘 신호 전송 이후 Epsilon-Greedy 전략에 따라 Q-learning 학습을 진행한다. 탐험률은 1.0에서 매 학습 진행 시 0.0005만큼 고정적으로 떨어트려, 시뮬레이션 진행에 따라 학습한 전략을 우선 이용하도록 유도하였다.

## 5.2 목표 변수 정의

시뮬레이션 목표 변수는 패킷 전송 시도 횟수  $T$ 를 ACK 메시지 수신까지 완료한 전송 성공 횟수  $R$ 로 나눈 값으로 식 (4)와 같이 정의하였다.

$$\eta = \frac{T}{R} \quad (4)$$

$\eta$ 는 모든 전송 시도에 대한 ACK를 수신하는 경우 1이고, ACK를 하나도 받지 못하는 경우 0이다. 즉  $\eta$ 가 1에 가까워질수록 높은 성능의 스케줄링 기법이라고 판단할 수 있다. 그러나 목표 변수  $\eta$ 는 비율이므로 각 노드가 타임 슬롯을 충분히 할당받지 못하는 경우, 패킷 전송 횟수인  $T$ 가 줄어들어 절대적인 전송 수가 줄어도  $\eta$ 는 오히려 증가할 수 있다는 문제가 있다. 따라서 보다 정확한 성능 비교를 위해 성능 평가시 패킷 전송률(Throughput)을 함께 비교하고자 한다. Throughput은 전송 성공한 패킷의 바이트 수를 전체 시뮬레이션 시간으로 나눈 값으로 정의하였다.  $T$ 가 줄어들어  $\eta$ 가 증가하는 상황에서, throughput은 일반적으로  $T$ 와  $R$  모두에 비례하므로 정확한 성능 비교가 가능하다.

## 5.3 실험 결과

성능 평가 결과, 전체적으로 기존 정적 타임 슬롯 할당 기법 대비 성능 향상을 관찰할 수 있었으며, 시뮬레이션 시간이 길어질수록 성능 향상 폭 역시 비례하여 목표 변수  $\eta$ 는 기존 기법 대비 최대 41.8% 증가하였다. 특히 착용자가 불규칙한 움직임을 보이는 경우에도 기존 정적 할당 기법 대비 패킷 전송률이 최대 24.5% 감소하였으며 제안하는 기법의 안정성도 확인할 수 있었다.

구체적으로, 그림 3과 같이 규칙적인 움직임을 보이는 시나리오 1에서는 시뮬레이션 시간이 길어질수록 움직임 패턴을 학습하여 전송 성공한 패킷의 비율 증가 폭이 (600s) 11.1%, (1800s) 33.5%, (3600s) 41.8%로 점차 증가하는 것을 확인할 수 있었다. 앞서 언급하였듯이  $\eta$ 는 실제 전송 성공 횟수를 대표하지 않기 때문에, 보조 지표인 패킷 전송률도 함께 비교하였다.

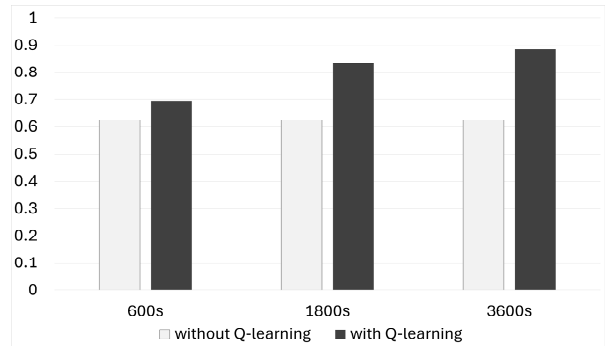


그림 3. 시나리오 1에서의 목표 변수  $\eta$  비교

Fig. 3. Comparison of the target variable  $\eta$  in scenario 1

그림 4와 같이, 주기적인 움직임이 발생하는 시나리오 1에서는 제안하는 기법이 시나리오의 모든 시뮬레이션 파라미터(시간)에서 좋은 성능을 보였다. 목표 변수  $\eta$  비교 때와 동일하게, 시뮬레이션 시간이 길어질수록 패킷 전송률 증가 폭이 (600s) 1.8%, (1800s) 6.3%, (3600s) 7.8%로 증가하는 경향도 확인할 수 있었다.

그림 5는 전원 사용량 비교 결과를 보여준다. 전원 사용량은 BANSIM[12]에서 도출하는 전원 사용 비율 지표를 활용하였으며, 전원 사용량 증가 폭 또한 기존 기법 대비 (600s) -8.4%, (1800s) -20.9%,

(3600s) -24.5%로, 시뮬레이션 시간이 길어질수록 전원 사용량 증가율이 누적 감소하는 경향을 확인할 수 있다.

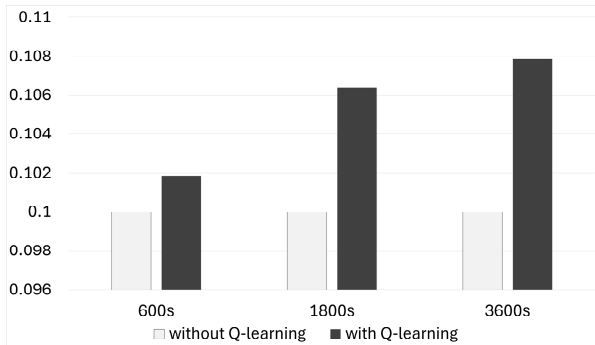


그림 4. 시나리오 1에서의 throughput 비교  
Fig. 4. Comparison of the throughput in scenario 1

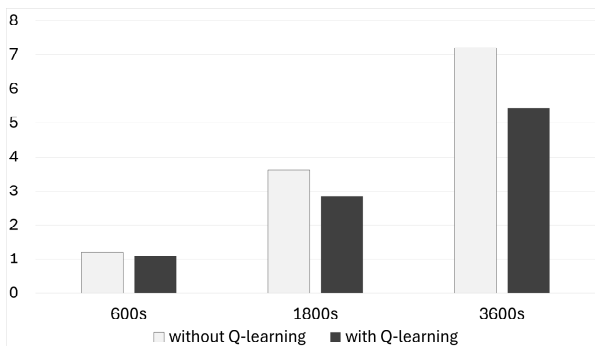


그림 5. 시나리오 1에서의 전원 사용량 비교  
Fig. 5. Comparison of the power consumption in scenario 1

시나리오 2에서는 WBAN 기기 착용자가 불규칙한 움직임을 보이는 경우를 가정하였다. 제안하는 스케줄링 기법은 사용자의 움직임이 개인화된 주기를 가진다는 것을 가정하여 설계되었으므로, 주기를 갖지 않는 무작위적 움직임을 보이는 경우는 제시하는 기법이 마주할 수 있는 최악의 상황이라고 할 수 있다. 두 번째 시나리오는 제안하는 기법이 적어도 어느 정도의 성능을 보장하는지 확인해 보기 위해 계획하였으며, 무작위 노드를 선택하여 20%의 확률로 해당 노드의 패킷을 드랍시키는 방법으로 시뮬레이션을 진행하였다.

그림 6과 같이, 목표 변수  $\eta$  비교 시 모든 시뮬레이션 파라미터의 경우(600s, 1800s, 3600s)에서 0.8 대로 기존 정적 스케줄링 기법과 같은 수준의 성능을 보인다는 것을 확인할 수 있다. 하지만 그림 7의 throughput 비교의 경우 제안하는 기법이 기존 기법

대비 (600s) -8.3%, (1800s) -20.8%, (3600s) -24.5%의 성능 감소를 보였다.

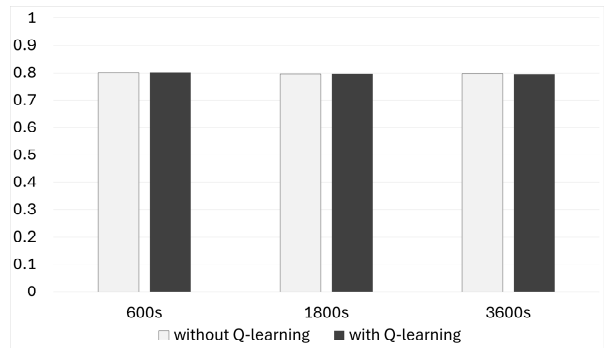


그림 6. 시나리오 2에서의 목표 변수  $\eta$  비교  
Fig. 6. Comparison of the target variable  $\eta$  in scenario 2

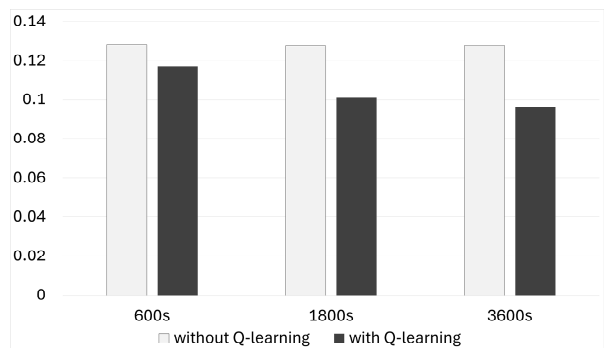


그림 7. 시나리오 2에서의 throughput 비교  
Fig. 7. Comparison of the throughput in scenario 2

그림 8은 시나리오 2에서의 전원 사용량 비교 결과이며, 시뮬레이션 시간에 비례하여 제안하는 기법의 전원 사용량 증가 폭이 기존 기법 대비 (600s) -8.4%, (1800s) -20.8%, (3600s) -24.5%로, 시뮬레이션 시간에 비례하여 전원 사용량 증가율이 감소하는 경향을 보이는 것을 확인할 수 있었다.

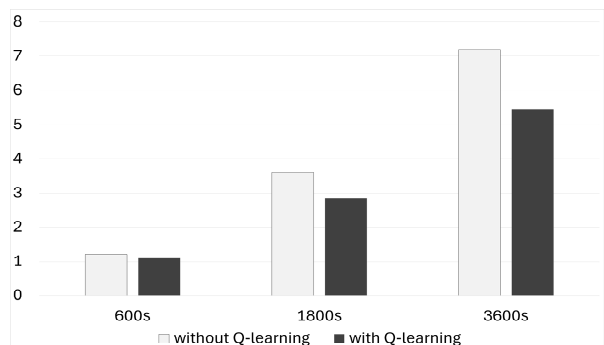


그림 8. 시나리오 2에서의 전원 사용량 비교  
Fig. 8. Comparison of power consumption in scenario 2

시나리오 1, 2의 전원 사용량 비교와 시나리오 2의 throughput 비교에서 증감 폭이 (600s)  $\pm 8\%$ , (1800s)  $\pm 20\%$ , (3600s)  $\pm 24\%$ 로 비슷하다는 점과, 시나리오 2는 제안하는 기법이 맞닥뜨릴 수 있는 최악의 경우를 모델링한 것임을 미루어 보아, 제안하는 기법은 최악의 경우 전원 사용량 감소 폭만큼의 throughput 저하를 보인다는 점을 알 수 있다.

## VI. 결론 및 향후 과제

본 연구에서는 WBAN 사용자가 개인화된 주기적 움직임 패턴을 보인다는 것에 착안하여, 패턴의 반복을 학습하는 Q-learning 기반의 최적 타임 슬롯 할당 스케줄링 기법을 제시하였다. 주기적인 움직임을 보이는 경우, 기존 노드 우선순위 기반 정적 할당 기법에 비해 목표 변수  $\eta$ 이 최대 약 41% 증가하였음을 확인하였으며 전원 사용량 또한 약 21% 감소하였다. 제시하는 기법이 마주할 수 있는 가장 최악의 상황에서는 기존 스케줄링 기법보다 성능이 최대 24.5% 감소하였다. 착용자가 완전한 무작위 움직임을 보이는 경우 기존 할당 기법보다 패킷 전송률이  $10^{-2}$  수준에서 약 20% 하락하였으며, 전원 사용량 또한 약 21% 감소하였다. 본 연구의 한계점은 앓기, 기구 운동 등의 다양한 시나리오에 대한 안정성 및 효용성을 제시하지 못하였다는 점이다. 향후 연구에서는 다양한 자세에 대한 성능 검증을 진행하고, 딥러닝 기반 강화 학습 전략을 이용하여 제안하는 기법을 고도화하고자 한다.

## References

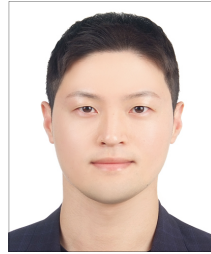
- [1] H. Taleb, A. Nasser, G. Andrieux, N. Charara and E. M. Cruz, "Wireless technologies, medical applications and future challenges in WBAN: a survey", *Wireless Networks* Vol. 27, pp. 5271-5295, Sep. 2021. <https://doi.org/10.1007/s11276-021-02780-2>.
- [2] W. Lee, D. Yu, and N. Kim, "Lossless Compression Schemes to Increase Energy Efficiency in Wireless Body Sensor Networks", *The Journal of Korean Institute of Information Technology*, Vol. 13, No. 7, pp. 67-74, Jul. 2015. <https://doi.org/10.14801/jkiit.2015.13.7.67>.
- [3] J. Ryu, J. Kwon, and J. Joung, "Timeslot Scheduling with Reinforcement Learning Using a Double Deep Q-Network", *The Journal Of Korean Institute of Communications and Information Sciences*, Vol. 47, No. 7, pp. 944-951, May 2022. <https://doi.org/10.7840/kics.2022.47.7.944>.
- [4] M. Kherbache, O. Sobirov, M. Maimour, E. Rondeau, and A. Benyahia, "Reinforcement Learning TDMA-Based MAC Scheduling in the Industrial Internet of Things: A Survey", *IF AC-PapersOnLine*, Vol. 55, No. 08, pp. 83-88, Aug. 2022. <https://doi.org/10.1016/j.ifacol.2022.08.014>.
- [5] J. Kim, B. Kim, C. You, and H. Park, "Priority Based Adaptive Slotted ALOHA Method Using Q-Learning", *The Journal Of Korean Institute of Communications and Information Sciences*, Vol 48, No. 3, pp. 350-358, Dec. 2022. <https://doi.org/10.7840/kics.2023.48.3.350>.
- [6] K. K. Cwalina, A. Olejniczak, O. Blaszkiwicz, P. Rajchowski, and J. Sadowski, "Fast Fading Influence on the Deep Learning-Based LOS and NLOS Identification in Wireless Body Area Networks", *2021 XXXIVth General Assembly and Scientific Symposium of the International Union of Radio Science (URSI GASS)*, Rome, Italy, pp. 1-4, Aug. 2021. <https://doi.org/10.23919/URSIGASS51995.2021.9560557>.
- [7] A. S. Rajasekaran, L. Sowmiya, A. Maria, and R. Kannadasan, "A survey on exploring the challenges and applications of wireless body area networks (WBANs)", *Cyber Security and Applications*, Vol. 2, pp. 1-29, Jan. 2024. <https://doi.org/10.1016/j.csa.2024.100047>.
- [8] N. Bilandi, H. K. Verma, and R. Dhir, "An Energy Efficient TDMA based MAC Protocol for Wireless Body Area Networks", *First International Conference on Secure Cyber Computing and*



Communication (ICSCCC), Jalandhar, India, pp. 545-549, Dec. 2018. <https://doi.org/10.1109/ICSCCC.2018.8703340>.

- [9] S. Jeong and J. G. Ko, "Body Area Network Sensor Based Posture Classification System", Proc. of the Korean Society of Computer Information Conference, Jeju, Korea, pp. 2146-2148, Jun. 2018.
- [10] B. Kang and J. Kim, "Low Power Beacon Listening Scheme Using Periodic Signal Characteristic in WBAN", The Journal of Korean Institute of Communications and Information Sciences, Vol. 44, No. 6, pp. 1121-1123, Apr. 2019. <https://doi.org/10.7840/kics.2019.44.6.1121>.
- [11] D.-H. Kim, H.-H. Lim, B.-J. Lee, K.-T. Kim, and H.-Y. Youn, "Adaptive Packet Scheduling Algorithm in IoT environment", Proc. of the Korean Society of Computer Information Conference, Jeju, Korea, Vol. 26, No. 2, pp. 15-16, Jul. 2018.
- [12] B.-S. Kim, K.-I. Kim, and B. Shah, "BANSIM: A new discrete-event simulator for wireless body area networks with deep reinforcement learning in Python", Journal of Systems Architecture, Vol. 126, pp. Apr. 2022. <https://doi.org/10.1016/j.sysarc.2022.102489>.

김 범 수 (Beom-Su Kim)



2009년 3월 ~ 2016년 2월 :  
경상국립대학교(학사)  
2016년 3월 ~ 2018년 2월 :  
경상국립대학교(석사)  
2018년 3월 ~ 2022년 2월 :  
충남대학교(박사)  
2022년 7월 ~ 2022년 12월 :

펜실베이니아 주립대학교 연수연구원  
2023년 9월 ~ 현재 : 경상국립대학교 컴퓨터공학과  
조교수  
관심분야 : 5G/6G, 에드혹 네트워크, 사물인터넷,  
기계학습

## 저자소개

조 승 현 (Seoung-Hyeon Jo)



2020년 3월 ~ 현재 :  
경상국립대학교 컴퓨터공학과  
학사과정  
관심분야 : 5G, 사물인터넷,  
기계학습