

IoT 네트워크에서의 심층 강화학습 기반 저전력 버퍼 관리 기법에 관한 연구

송태원*

순천향대학교 사물인터넷학과 교수

A Research on Low-power Buffer Management Algorithm based on Deep Q-Learning approach for IoT Networks

Taewon Song*

Professor, Department of IoT, Soonchunhyang University

요약 IoT 네트워크에서 클러스터와 싱크 노드 사이의 게이트웨이 역할을 하는 클러스터 헤드의 전력 관리는 IoT 단말의 수가 증가함에 따라 점점 더 중요해지고 있다. 특히 클러스터 헤드가 이동성을 가진 무선 단말인 경우, IoT 네트워크의 수명을 위하여 전력 소모를 최소화할 필요가 있다. 또한 IoT 네트워크에서의 전송 딜레이는 IoT 네트워크에서의 빠른 정보 수집을 위한 주요한 척도 중 하나이다. 본 논문에서는 IoT 네트워크에서 정보의 전송 딜레이를 고려한 저전력 버퍼 관리 기법을 제안한다. 제안하는 기법에서는 심층 강화학습 방법에서 사용되는 심층 Q 학습(Deep Q learning)를 사용하여 수신된 패킷을 포워딩하거나 폐기함으로써 전송 딜레이를 줄이면서도 소비 전력을 절감할 수 있다. 제안한 알고리즘은 비교에 사용된 기존 버퍼 관리 기법과 비교하여 Slotted ALOHA 프로토콜 기준 소모 전력 및 딜레이를 개선함을 보였다.

주제어 : 사물인터넷, 인공지능, 심층 강화 학습, 심층 Q 러닝, 버퍼 관리

Abstract As the number of IoT devices increases, power management of the cluster head, which acts as a gateway between the cluster and sink nodes in the IoT network, becomes crucial. Particularly when the cluster head is a mobile wireless terminal, the power consumption of the IoT network must be minimized over its lifetime. In addition, the delay of information transmission in the IoT network is one of the primary metrics for rapid information collecting in the IoT network. In this paper, we propose a low-power buffer management algorithm that takes into account the information transmission delay in an IoT network. By forwarding or skipping received packets utilizing deep Q learning employed in deep reinforcement learning methods, the suggested method is able to reduce power consumption while decreasing transmission delay level. The proposed approach is demonstrated to reduce power consumption and to improve delay relative to the existing buffer management technique used as a comparison in slotted ALOHA protocol.

Key Words : IoT; Artificial intelligence; Deep reinforcement learning; Deep Q learning; Buffer management

1. 서론

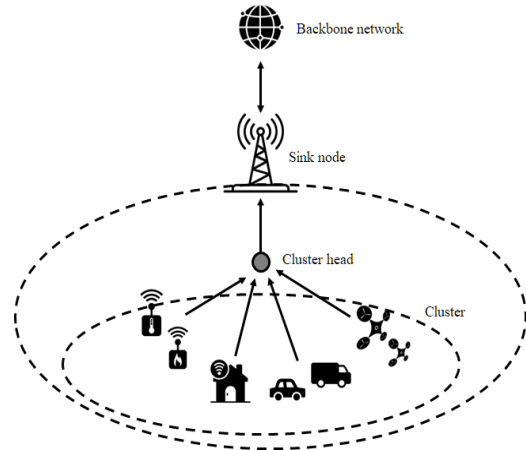
무선 센서 네트워크는 인프라스트럭처가 없고 자주적으로 운용된다는 특성으로 인해 사물 인터넷(Internet of things, IoT) 응용기술의 기반 네트워크로 주로 사용된다 [1]. <그림 1>에서 IoT 네트워크의 구조를 간략히 나타내었다. 싱크 노드는 통신 프로토콜에 따라 다를 수 있으나 3GPP에서의 eNodeB 또는 IEEE 802.11에서의 액세스 포인트 (access point, AP)에 대응될 수 있다. 하나의 싱크 노드가 담당하는 단말들은 다수의 클러스터를 형성하여 다중 홉 (multi-hop) 방식으로 싱크 노드로 메시지를 전달한다.

클러스터 멤버(cluster member, CM)는 어플리케이션에 따라 IoT 센서, 커넥티드 홈 디바이스, 차량, 무인기(unmanned aerial vehicle, UAV) 등이 대응될 수 있으며 클러스터 헤드를 포함한다. 클러스터 헤드(cluster head, CH)는 클러스터 멤버 중 하나로 선택되며 그 역할은 주기적으로 할당될 수 있고 고정되어 운용될 수도 있다. 하나의 CH에는 다수의 CM이 연결되고, CM에서 모니터링한 정보는 CH로 전송된다. 이러한 하나의 CH와 다수의 CM들은 하나의 클러스터를 형성하고, IoT 네트워크의 하위 구성요소를 이루게 된다. 클러스터의 수는 다수일 수 있으나 본 그림에서는 간략한 표현을 위해 하나의 클러스터만 표시하였다.

CM과 CH를 포함한 단말들에 대해 소모 전력을 최소화하는 것은 궁극적으로 IoT 네트워크의 네트워크 수명을 연장하는데 있어 일차적인 목표가 되어야 하지만, 모니터링되는 정보의 종류가 매우 다양하고 그 크기 또한 방대한 양을 갖게 되는 빅 데이터 시대에서는 저전력 설계는 매우 중요한 의미를 갖는다.

본 논문에서는 IoT 네트워크를 위한 저전력 버퍼 관리를 제안한다. 본 알고리즘을 통해 CH의 에이전트는 심층 강화 학습 (deep reinforcement learning, DRL)을 통해 버퍼링된 데이터를 전달하거나 또는 건너뛰지 여부를 결정할 수 있어 딜레이 값을 크게 증가시키지 않고 낮은 전력 소비를 달성할 수 있다. 이를 위해 먼저 마르코프 결정과정 (Markov decision process, MDP) 모델을 공식화하였다. MDP 모델에 대한 상태의 수가 많기 때문에 가치 기반 알고리즘인 DQN 알고리즘 기반의 DRL 모델을 공식화하여 최적의 정책을 얻는다. IoT 네트워크에 널리 사용되는 무선 매체 액세스 프로토콜인 슬롯형 알로하(slotted ALOHA)를 기반으로 제안 알고리즘을 시뮬레이션하였다. 광범위한 시뮬레이션 결과를 통해 제

안된 분석 기법의 정확성을 검증하고 매체 접근 프로토콜에 따른 영향을 조사한다. 성능 분석 결과는 제안 알고리즘이 적절한 딜레이 레벨을 유지하면서 전력 소비를 크게 줄일 수 있음을 보여준다.



[Fig. 1] A brief structure of IoT network

2. 관련 연구

IoT 네트워크에서의 CH의 전원 관리에 대한 여러 연구가 이루어졌다. 이는 일반적으로 1) CH의 역할을 CM들에게 주기적으로 분배하여 네트워크 수명을 연장하는 방법, 또는 2) 고정 토폴로지에서 전력 소비를 낮추는 방법으로 분류할 수 있다.

앞서 언급한 두 방법 중 전자에서 저전력 소모를 구현하는 대표적인 클러스터 기반 프로토콜은 LEACH이다 [2]. CH의 역할을 무작위로 분배하여 센서들 간의 전력 소비를 고르게 분배할 수 있어 궁극적으로 네트워크의 수명을 연장시킬 수 있다. 연구 [3, 4]에서는 퍼지 논리 기반 접근 클러스터링 방법을 채택하여 LEACH의 성능을 향상시켰다. 연구 [5]에서는 Q-LEACH라는 알고리즘을 제안하였는데, 멀리 떨어져 있는 클러스터 노드를 서비스할 수 있도록 네트워크를 분할하여 무선 센서 네트워크의 안정성, 수명 및 처리량을 향상시켰다. 연구 [6]에서는 네트워크 토폴로지 및 구축 알고리즘을 통해 원자력 발전소 모니터링 시스템에서의 무선 센서 네트워크의 수명을 향상시키는 하이브리드 접근 방식을 제안하였다. 연구 [7]에서는 O-LEACH라는 알고리즘을 제안하였는데, 에너지 사용량을 최소화하기 위해 고립된 노드의

수를 최소화하도록 구성하였다. 이외에도 우수한 연구가 많이 있지만 이러한 접근 방식은 필연적으로 라우팅 테이블의 주기적인 수정을 필요로 해야 하고, 더불어 CH가 될 수 있는 CM은 송수신이 가능한 트랜시버를 장착해야 하는데, 이는 가벼운 엔트리 레벨의 IoT 칩셋만으로는 어려울 수 있다.

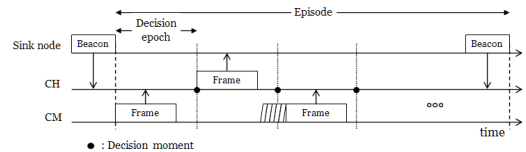
고정된 토폴로지에서의 동적 전력 관리 방법 또한 많이 연구되어 왔다. 연구 [8]에서는 대기열 임계값을 정의하고 이를 기반으로 한 버퍼 관리 프로토콜을 제안했다. 이를 통해 전송 횟수를 줄임으로써 네트워크 수명을 향상시킬 수 있다. 연구 [9]에서는 동적 전력 관리를 위하여 강화 학습을 채택하여 다양한 애플리케이션에 대해 비트 오류 확률, 채널 상태 정보, 유지 시간 등을 포함한 여러 네트워크 매개변수를 활용하여 QoS를 보장하는 알고리즘을 제안하였다. 연구 [10]에서는 에너지 효율적인 버퍼 관리 방식을 제안하였다. 제안하는 알고리즘은 다변수 데이터 축소 기법을 통해 버퍼 공간을 절약할 수 있어 클러스터 헤드의 전력 소모를 줄일 수 있다. 그러나 이러한 접근 방식에는 많은 매개변수가 필요하고 센서 노드에 대해 상당히 높은 작동 기능이 요구될 수 있다.

3. 시스템 모델

본 연구에서는 <그림 1>과 같이 일반적으로 기지국(base station, BS) 또는 액세스 포인트(access point, AP)로 여겨지는 하나의 싱크 노드, 그리고 여러 IoT 단말로 구성된 IoT 네트워크를 고려한다. IoT 단말들은 클러스터를 이루며 클러스터 멤버(cluster member, CM)로 명명된다. 싱크 노드와의 송수신을 위해 CM들 중 하나의 단말이 클러스터 헤드(cluster head, CH)로 지정된다. 각 CM은 센싱 데이터를 수집하여 해당 CH로 전송하고, CH는 버퍼링된 센싱 데이터를 싱크 노드로 전달한다. 본 연구에서는 모든 CH가 싱크 노드의 전송 범위 내에 있고 모든 CM이 해당 클러스터의 CH의 전송 범위 내에 있다고 가정한다. 클러스터의 전송 채널은 중첩되지 않으므로 각 CH는 무선 채널을 독립적으로 점유할 수 있다. 싱크 노드로의 전송은 통신 프로토콜에 따라 적절한 무선 매체 액세스 프로토콜(media access protocol, MAC)을 통해 전송한다. 본 연구에서는 IoT 네트워크에서 널리 사용되는 슬롯형 ALOHA를 고려하였다. <그림 2>는 제안 알고리즘이 동작하는 환경에서의 타이밍 다이어그램을 도시하였다. 싱크 노드는 다양한

필드를 포함할 수 있는 비콘 프레임을 주기적으로 브로드캐스팅한다. 프레임의 이름과 프레임에 포함된 필드는 통신 프로토콜에 따라 다를 수 있고 본 연구에서는 일반적인 용어를 사용한다. 하나의 에피소드를 구성하는 비콘 인터벌 동안 CM은 애플리케이션에 따라 온도, 속도 또는 위치가 될 수 있는 값들을 모니터링하고 이를 전송하기 위한 무선 채널을 획득하기 위해 경쟁한다. CM은 채널 액세스 기회를 얻으면 모니터링된 데이터와 데이터가 관찰된 시간을 포함하는 프레임을 전송한다. 결정 구간(decision epoch)은 프레임 지속 시간 또는 이보다 약간 더 긴 시간으로 정의된다. 각 결정 구간에서 CH는 forward 또는 skip으로 이루어진 액션 공간 집합에서 하나의 액션을 결정하여 수행한다.

비콘 인터벌이 끝나면 싱크 노드는 제안 기법을 통해 계산된 보상(reward)을 포함하는 비콘 프레임을 다시 브로드캐스트하고, 수정된 보상을 토대로 알고리즘을 반복하여 수행한다.



[Fig. 2] Timing diagram for proposed algorithm

4. IoT 네트워크의 저전력 버퍼 관리 기법

마르코프 결정 과정 (Markov decision process, MDP) 모델은 수학적 의사 결정을 위해 널리 사용되는 프레임워크이다. 본 절에서는 저전력 버퍼 관리 동작을 위한 MDP 모델을 제시하고 그에 따르는 심층 Q 학습(deep Q-network, DQN) 알고리즘을 제안한다. 이를 통해 버퍼 관리에서의 최적의 정책(policy)을 도출할 수 있다. 이하에서는 설명의 편의를 위해 하나의 싱크 노드, 하나의 CH, 그리고 그 CH와 연결된 다수의 CM이 존재하는 환경을 가정하였으나 다수의 CM으로 형성된 클러스터에서도 동일하게 적용할 수 있다.

4.1 상태 공간

현재 버퍼에 존재하는 프레임의 딜레이를 정의하기 위해 무한 집합 S 를 아래와 같이 정의한다.

$$S = (D_c, D_b) \quad (1)$$

D_c 와 D_b 는 프레임의 딜레이 값을 비컨 인터벌로 나타낼 수 있는 집합이며 $[0, 1]$ 의 구간에서 연속적인 값을 가진다. $d_c^n(t)$ 와 $d_b^n(t)$ 는 특정 시간 t 에서 CM n 의 지금까지 측정된 딜레이의 최대값 및 현재 버퍼 내부에 존재하는 프레임의 딜레이 값으로 각각 D_c 와 D_b 집합에 속한 원소이며 시간에 따라 달라지는 값을 가진다.

4.2 액션 공간

3절에서 언급한 바와 같이 CH는 CM들로부터 수신하여 버퍼에 저장된 프레임을 forward하거나 skip할 수 있다. 이를 정의하기 위해 유한 집합 A 를 아래와 같이 정의한다.

$$A = \{0, 1\} \quad (2)$$

액션 $a \in A$ 가 0, 1인 경우를 각각 forward, skip으로 정의한다. $a = 0$ 일 때, CM은 버퍼에 있는 프레임 중 가장 이전에 수신했던, 즉 head에 존재하는 프레임을 싱크 노드에게 송신한다. 이 경우 CH는 결정 구간동안 CM로부터의 데이터를 수신할 수 없다. 반면 $a = 1$ 일 때, CH는 CM으로부터의 데이터를 수신하기 위해 CM 싱크 노드에게 송신하지 않는다.

4.3 보상 함수

본 연구에서는 보상 함수를 정의하기 위해 정규화된 딜레이 값과 소비된 에너지를 고려한다. 본 연구에서는 고정된 결정 구간 및 비컨 인터벌 값을 가정하기 때문에 이후부터는 소비된 에너지를 소비된 전력으로 대체하여 사용한다.

4.2절에서 언급한 바와 같이 CH는 CM들로부터 수신하여 버퍼에 저장된 프레임을 forward하거나 skip할 수 있다. 상태 s 에서 액션 a 를 선택하여 행동했을 때의 상세한 보상 함수는 아래와 같이 정의하였다.

$$R(s, a) = \alpha R_p(s, a) + (1 - \alpha) R_d(s, a) \quad (3)$$

보상 함수는 전력을 의미하는 $R_p(s, a)$ 와 딜레이 값 의미하는 $R_d(s, a)$ 의 합으로 이루어진다. 전력 보상 함수는 아래와 같이 정의하였다.

$$R_p(s, a) = \begin{cases} -P_{transmit}, & \text{if } a = 0 \\ -P_{receive}, & \text{if } a = 1 \end{cases} \quad (4)$$

여기서 $P_{transmit}$ 과 $P_{receive}$ 는 무선랜 표준인 IEEE 802.11ax simulation scenario에 정의된 common power model parameter [11]를 사용하였다.

딜레이 보상 함수는 다양한 CM들로부터 수신한 프레임을 싱크 노드에 전달했을 때 발생하는 딜레이의 최대값으로 정의하였다. 따라서 딜레이 보상 함수 $R_d(s, a)$ 는 아래와 같이 나타난다.

$$R_d(s, a) = -\max_{n=1,2,\dots,N}(d_c^n(t)) \quad (5)$$

여기서 $d_c^n(t)$ 는 CH의 버퍼 내부에 존재하는 프레임들에 대해 CM n 이 모니터링한 데이터의 딜레이 값이다. 즉, 딜레이 보상 함수 $R_d(s, a)$ 는 현실점에서 CM들이 송신한 데이터들의 딜레이 값들의 최대값으로 이해할 수 있다.

5. 심층 강화학습 알고리즘

DRL은 시행착오를 통해 동적이고, 예측이 불가능하며 불확실성이 높은 네트워크 환경에서 최적의 정책을 도출하기 위해 널리 사용되는 기계학습 방법론이다. 상태 공간이 매우 크거나 무한한 경우 최적화 작업이 복잡해지거나 불가능할 수 있는데, 이러한 환경에서 DRL 기법을 활용하여 최적화 문제를 해결한다. 본 논문에서는 DRL의 프레임워크의 일종인 DQN에 기반하여 최적의 버퍼 관리 정책을 결정한다. 이 절에서는 DQN이 CH의 버퍼 관리 기법에 어떻게 적용되는지 설명한다.

DQN은 입력이 식 (1)에서의 집합 S 의 부분 집합으로 정의되고 출력이 동작의 Q 값, 즉 각 액션의 가치인 단일 신경망을 사용한다. 알고리즘은 Q 값이 가장 높은 액션을 선택한다. DQN 알고리즘은 경험(experience) 및 교육(training)의 두 단계로 구현된다. CH는 경험 단계에서 액션을 수행하고 경험 e 를 받는다. 경험 e 는 상태, 액션, 보상, 그리고 다음 상태, 즉

$$e = \langle s, a, r, s' \rangle \quad (6)$$

의 tuple로 이루어져 있는데, 이렇게 얻어진 경험 e 와 더불어 오차 함수를 최소화하는 방향으로 목적 함수가 최적화된다. 본 절에서 설명한 과정은 아래 <알고리즘 1>과 같이 매 비컨 인터벌 마다 반복되어 실행된다.

Algorithm 1 DQN for proposed algorithm

```

Initialize
for episode = 1 to N, do
    for iteration = 1 to T, do
        Select action  $a$  with greedy manner
        Execute action  $a$ 
        Observe reward  $r$  & next state  $s'$ 
        Store experience  $e$ 
        Proceed a gradient step
    endfor
endfor
    
```

6. 실험 결과

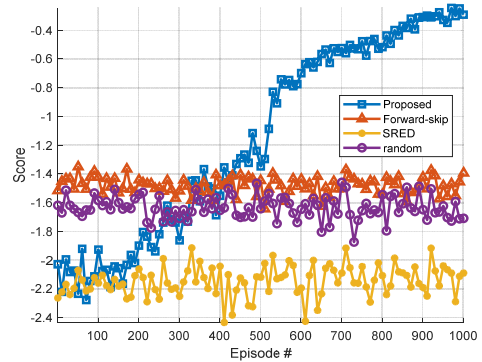
본 제안 기법의 성능 분석을 위해 Python 3.9를 기반으로 하여 강화 학습을 위한 표준 API인 gym [12], 과학 컴퓨팅을 위한 파이썬 패키지인 NumPy [13], 오픈 소스 머신 러닝 프레임워크인 PyTorch [14]를 사용하여 광범위한 시뮬레이션을 수행하였다. 제안 기법은 1) forward-skip 기법, 2) 무작위 액션 기법, 그리고 3) stabilized random early drop (SRED) (SRED) [15]이다. Forward-skip 기법은 CM이 forward와 skip을 번갈아가며 선택하는 기법이고, 무작위 액션 기법은 매 결정 구간마다 액션을 랜덤하게 결정하는 기법이다. 그리고 SRED 기법은 잘 알려져 있는 고전적 능동 버퍼 관리 기법 중 하나로, 연결된 CM의 수와 무관하게 버퍼의 대기열 크기를 안정화하도록 설계된 기법이다.

〈Table 1〉 Simulation parameters

Parameter	Value
Number of neurons	128
Optimizer	Adam [15]
Activation function	ReLU
Learning rate	0.0005
Discount rate	0.98
Beacon interval	100 ms
Decision epoch	300 us
Frame duration	270 us
Number of CMs	10
Buffer size	100 frames
Balance parameter	0.1
MAC protocol	slotted ALOHA with $p=1/N$

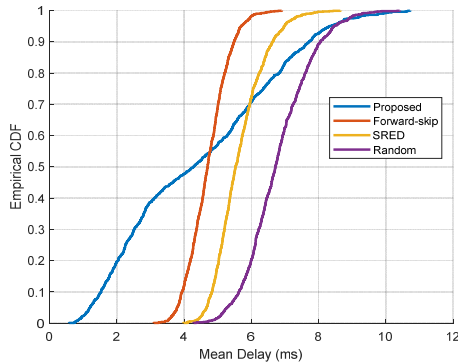
DRL을 위한 파라미터는 gym 및 PyTorch에서 널리 사용되는 값을 사용하였다. 비컨 인터벌은 IEEE 802.11 무선랜 표준에서 널리 사용되는 값인 100 ms를, 하나의 CH에 연결된 CM의 수는 10개로 가정하였다. CM의 무선 매체 접근 프로토콜은 slotted ALOHA에서 접근 확률이 $1/N$, 즉 0.1의 확률로 전송을 시도하도록 설정하였다. 그리고 수식 (3)에서 사용된 조정 파라미터는 0.1로 설정하였다. 〈표 1〉에서 DQN 및 IoT 네트워크의 시뮬레이션 파라미터를 나열하였다.

아래 〈그림 3〉에서는 에피소드의 진행에 따른 보상의 값을 제안 기법 및 비교 기법에 대해 도시하였다. 에피소드가 진행됨에 따라 시행 착오 과정을 거쳐 심층 신경망 (deep neural network, DNN)의 파라미터가 현재의 상태에 대한 가치를 잘 설명할 수 있도록 학습되어가고, 이를 통해 제안 기법의 보상이 점차적으로 증가하는 것을 볼 수 있다. IoT 네트워크의 경우 단기간으로 운용되기보다는 장기간으로 운용되는 경우가 많기 때문에 이러한 DQN 기반 학습 기법은 IoT 네트워크의 특성 상 매우 적합한 프레임워크임을 알 수 있다.

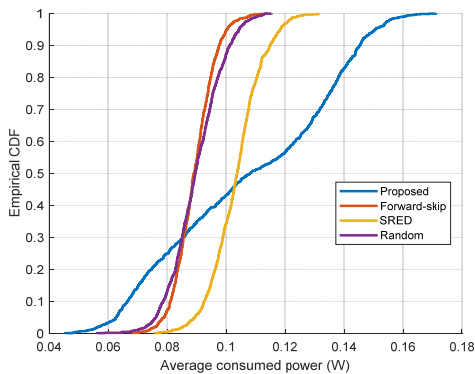


[Fig. 3] The evolutions of average reward

〈그림 4〉와 〈그림 5〉에서는 제안 기법 및 비교 기법에 대한 소모된 전력과 정보의 딜레이를 누적 분포 함수의 형태로 도시하였다. 두 그림 공통으로 제안 기법의 값 분포가 비교 기법들의 값의 분포에 비해 넓게 분포하는 것을 관찰할 수 있는데, 이는 학습이 진행됨에 따라 소모 전력과 딜레이가 점차 최적화되기 때문인 것으로 관찰된다. 에피소드가 진행됨에 따라 보상 함수를 최대화하기 위한 방향으로 학습이 진행되면서 소모 전력 및 딜레이 모두 줄어드는 것을 관찰할 수 있었다.



[Fig. 4] Mean delay



[Fig. 5] Average power consumption

7. 결론

본 논문에서는 심층 강화 학습의 프레임워크 중 하나인 심층 Q 학습 기법을 사용하여 IoT 네트워크에서 데이터의 중계기 역할을 수행하는 CH의 저전력 버퍼 관리기법을 제안하였다. 시뮬레이션을 통해 시간이 흐름에 따라 심층 신경망이 수렴하는 것을 보여 심층 Q 학습 기법은 장기간 운용되는 IoT 네트워크의 특성에 적절한 프레임워크임을 보였다. 제안하는 알고리즘을 현재 배치되어 운용되고 있는 IoT 네트워크에 적용한다면 CH의 전력 소모를 줄이면서 전송 딜레이 또한 감소시킬 수 있어 IoT 네트워크의 실질적인 수명을 늘릴 수 있고 더불어 모니터링되는 정보의 딜레이 또한 감소시킬 수 있어 IoT 네트워크를 기반으로 하여 양질의 어플리케이션 서비스를 제공할 수 있을 것으로 보인다. 특히 정보의 적시성이 요구되는 차량 네트워크 및 정보의 신선도가 요구되는 동적 IoT 네트워크에서 본 기법을 적용한다면 낮은 전력으로도 해당 어플리케이션의 요구 조건을 만족시킬 수 있을 것으로 보인다.

REFERENCES

- [1] M. T. Lazarescu, "Design of a wsn platform for long-term environmental monitoring for iot applications," IEEE Journal on emerging and selected topics in circuits and systems, Vol.3, No.1, pp.45-54, 2013.
- [2] W. R. Heinzelman, A. Chandrakasan, and H. Balakrishnan, "Energy-efficient communication protocol for wireless microsensor networks." IEEE Proceedings of the 33rd annual Hawaii international conference on system sciences, p.223, 2000.
- [3] I. Gupta, D. Riordan, and S. Sampalli, "Cluster-head election using fuzzy logic for wireless sensor networks," IEEE 3rd Annual Communication Networks and Services Research Conference (CNSR'05), 2005.
- [4] A. K. Dwivedi and A. K. Sharma, "EE-LEACH: energy enhancement in LEACH using fuzzy logic for homogeneous WSN," Wireless Personal Communications, Vol.120, pp.3035-3055, 2021.
- [5] B. Manzoor, N. Javaid, O. Rehman, M. Akbar, Q. Nadeem, A. Iqbal, and M. Ishfaq, "Q-LEACH: A new routing protocol for WSNs," Vol.19, pp.926-931, 2013.
- [6] M. E. Haque, T. Hossain, M. R. Sarker, M. Paul, M. S. Hoque, S. Uddin, A. A. Suman, M. H. M. Saad, and T. U. Huque, "A hybrid approach to enhance the lifespan of wsns in nuclear power plant monitoring system," Scientific Reports, Vol.12(1), pp.1-14, 2022.
- [7] G. A. Senthil, A. Raaza, and N. Kumar, "Internet of things energy efficient cluster-based routing using hybrid particle swarm optimization for wireless sensor network," Wireless Personal Communications, Vol.122, pp.2603-2619, 2022.
- [8] R. Maheswar, P. Jayarajan, S. Vimalraj, G. Sivagnanam, V. Sivasankaran, and I. S. Amiri, "Energy efficient real time environmental monitoring system using buffer management protocol: energy efficient real time environmental monitoring system using buffer management protocol," IEEE 2018 9th International Conference on Computing, Communication and Networking Technologies (ICCCNT), 2018.
- [9] D. Hosahalli and K. G. Srinivas, "Enhanced reinforcement learning assisted dynamic power management model for internet-of-things centric wireless sensor network," IET Communications, Vol.14, pp.3748-3760, 2020.
- [10] H. A. Alwasef, "An energy-efficient buffer management scheme based on data integrity and multivariate data reduction for wireless sensor networks," Journal of Control Engineering and Applied Informatics, Vol.23, No.3, pp.53-61, 2021.
- [11] S. Merlin, G. Barriac, H. Sampath, L. Cariou, T. Derham, J.-P. L. Rouzic, R. Stacey, M. Park, C. Ghosh, R. Porat, N. Jindal, Y. Inoue, Y. Asai, Y. Takatori, A.

Kishida, A. Yamada, R. Hedayat, S. Choudhury, K. Doppler, J. Knecht, E.-H. Rantala, D. X. Yang, Y. (Ross), Z. Lan, J. Zhang, Y. Li, Y. Li, J. Pang, H. Su, Y. Lin, W. Lee, H. Cho, S. Kim, H. Choi, J. Levy, F. L. Sita, J. Jiang, L. Chu, Y. Sun, F. Mestanov, G. Li, S. Marin, E. Sakai, W. Carney, B. Sun, K. Lv, Y. Ke, H. Zhiqiang, C.-C. Wang, R. Huang, C. Yu, J. Yee, E. Wong, J. Kim, and X. Wang, "TGax Simulation Scenarios." [Online]. Available: <https://mentor.ieee.org/802.11/dcn/14/11-14-0980-16-00ax-simulation-scenarios.docx>

- [12] "Gym Documentation." [Online]. Available: <https://www.gymnasium.ml>
- [13] "NumPy." [Online]. Available: <https://numpy.org>
- [14] "PyTorch." [Online]. Available: <https://pytorch.org>
- [15] T. J. Ott, T. V. Lakshman, and L. H. Wong, "Sred: Stabilized red," IEEE INFOCOM'99, Vol.3, pp.1346-1355, 1999.
- [16] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," 2014. [Online]. Available: <https://arxiv.org/abs/1412.6980>

송 태 원(Taewon Song)

[정회원]



- 2010년 2월 : 고려대학교 공과대학 전기전자전파공학부 (공학사)
- 2017년 2월 : 고려대학교 공과대학 전기전자전파공학과 전자전기 컴퓨터공학전공 (공학박사)
- 2017년 2월 ~ 2020년 9월 : LG 전자 CTO 차세대표준연구소 선임연구원
- 2020년 9월 ~ 2021년 2월 : 순천향대학교 SCH미디어 랩스 연구조교수
- 2021년 3월 ~ 현재 : 순천향대학교 SCH미디어랩스 사물인터넷학과 조교수

〈관심분야〉

사물인터넷, 협력네트워킹, 차세대무선랜