

DQN 기반 Wi-SUN 네트워크의 Border Router 배치 최적화

손현호, 고은서, 명재준, *정소이

아주대학교 AI 융합네트워크학과, *아주대학교 전자공학과

{sohn8896, rhdmstj45, as327795, *sjung}@ajou.ac.kr

DQN-based Border Router Placement Optimization for Wi-SUN Networks

Hyunho Son, Eunseo Ko, Jaejun Myung, *Soyi Jung

Dept. of Artificial Intelligence Convergence Network, Ajou University,

*Dept. of Electrical and Computer Engineering, Ajou University

요약

본 논문은 Wi-SUN 네트워크의 성능 향상을 위해 강화학습 기법인 Deep Q-Network(DQN)를 활용하여 border router(BR) 배치 최적화 프레임워크를 설계하고, 통신 세기와 속도를 결합한 보상 함수를 통해 통신 품질을 최적화하는 연구를 다룬다. RPL, CTP, LOAD 등 다양한 라우팅 프로토콜 환경에서 성능을 비교 분석하였으며, 실험 결과 RPL 환경에서 가장 우수한 성능을 달성하여 제안 프레임워크의 효과성을 확인하였다.

I. 서론

현대의 스마트 빌딩 환경에서 빌딩 관리 시스템(BMS)은 난방, 환기, 냉방(HVAC), 조명, 보안, 화재 안전 시스템의 효율적인 관리를 위해 분산된 센서와 액추에이터 간의 안정적인 무선 통신을 필요로 한다.

IEEE 802.15.4g 표준 기반의 Wi-SUN(Wireless Smart Utility Network)은 Sub-GHz 주파수 대역에서 장거리 통신, 저전력 소비, 높은 보안 기능을 제공하여 스마트 빌딩의 핵심 통신 기술로 부상하고 있다. Wi-SUN은 그림 1과 같이 메쉬 네트워크 토폴로지를 통해 다중 경로 전송이 가능하며, 모든 센서 노드는 border router(BR)를 통해 백본 네트워크에 연결된다. [1], [2].

Wi-SUN 네트워크의 성능은 BR 배치와 라우팅 프로토콜에 크게 의존한다. 본 연구는 Deep Q-Network(DQN) 기반 BR 배치 최적화 프레임워크를 제안하며, routing protocol for low-power and lossy networks(RPL), collection tree protocol(CTP), load-based routing(LOAD) 등 다양한 라우팅 프로토콜 환경에서의 성능을 비교 분석 한다. DQN 알고리즘은 round trip time(RTT)과 received signal strength indicator(RSSI)를 결합한 보상 함수를 통해 통신 품질과 연결성을 동시에 최적화하는 배치 정책을 학습한다.

II. 시스템 모델

Wi-SUN 네트워크는 다양한 라우팅 프로토콜을 지원하며, 본 연구에서는 RPL, CTP, LOAD의 성능을 비교 분석한다.

RPL은 목적지 지향 비순환 그래프(destination oriented directed acyclic graph, DODAG)를 구성하여 BR을 루트로 하는 트리 구조를 형성한다. 각 노드는 objective function(OF)에 따라 선호 부모를 선택하며, 본 연구에서는 RSSI와 RTT를 주요 메트릭으로 활용하여 통신 품질 우선 경로를 선택한다. DODAG 형성 과정은 (1) BR의 DODAG Information Object(DIO) 브로드캐스트, (2) 이웃 노드의 rank 계산, (3) OF 기반 부모 노드 선택, (4) 상향 트래픽 전송으로 구성된다.

CTP는 ETX 메트릭 기반으로 패킷 손실을 최소화하는 신뢰성 중심 프로토콜이며, LOAD는 트래픽 부하를 고려하여 네트워크 수명을 연장하는 부하 분산 프로토콜이다.

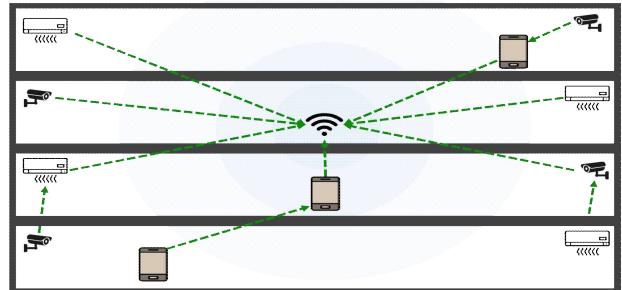


그림 1. Wi-SUN 메쉬 네트워크 구조

BR 위치는 각 라우팅 프로토콜의 토폴로지 형성과 통신 품질에 결정적 영향을 미치며, 양호한 RSSI 및 낮은 RTT를 제공하는 위치에 배치될 때 네트워크 전체 성능이 향상된다.

III. 강화학습 설계

BR 배치 최적화 문제는 markov decision process(MDP)로 정식화되며, 상태(State), 행동(Action), 보상(Reward)으로 구성된다.

1) State(상태) : $s \in \{BR, U, J\}$ 으로 정의한다. $BR = (x_{BR}, y_{BR})$ 는 현재 BR의 위치 좌표를 나타내고, $U = (x_{n_1}, y_{n_1}), (x_{n_2}, y_{n_2}), \dots, (x_{n_i}, y_{n_i})$ 는 네 $(R_{m_1}, x_{n_1}, y_{n_1}), (R_{m_2}, x_{n_2}, y_{n_2})$ 트위크에 합류하지 못한 노드들의 위치 좌표를, $J = , \dots, (R_{m_{15-i}}, x_{n_{15-i}}, y_{n_{15-i}})$ 는 네트워크에 합류한 노드들의 rank와 위치좌표를 나타낸다.

2) Action(행동) : $a = (x_{br}, y_{br}) \in \{(x_1, y_1), (x_2, y_1), \dots, (x_{100}, y_5)\}$ 로 정의되며, BR의 새로운 배치 위치를 선택한다. 복도 길이 100m를 1m 간격으로 구분하고, 5개 층으로 구성되어 총 500개의 가능한 배치 위치가 존재한다.

3) Reward(보상) :

$$r = \begin{cases} -penalty & \text{if disconnected,} \\ \alpha \times R_{rtt}(x_{br}, y_{br}) + \beta \times R_{rss}(x_{br}, y_{br}) & \text{otherwise} \end{cases}$$

로 정의한다. 네트워크가 연결되지 않으면 큰 페널티를 부여하고, 그렇지 않으면 RTT와 RSSI 기반 보상의 가중합으로 계산한다.

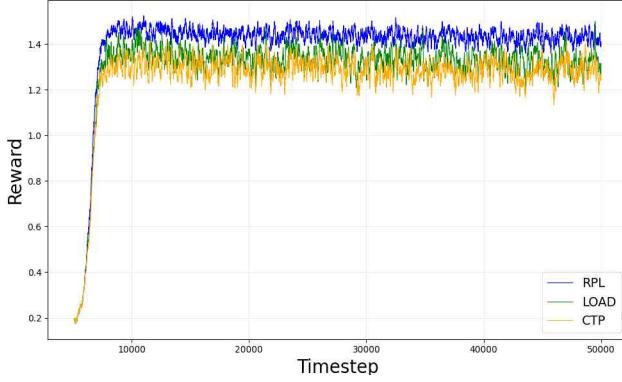


그림 2. 강화학습 누적 보상 그래프

RTT 기반 보상은 식 (1)과 같이 각 노드에서 BR까지의 전송 시간을 역수 형태로 변환하여 낮은 지연을 선호하도록 설계되었다. RSSI 기반 보상은 식 (2)와 같이 모든 노드의 평균 수신 신호 세기를 반영한다. 각 노드의 전송 시간은 식 (3)에서 Shannon capacity를 기반으로 산출되며 p 는 패킷 크기, B 는 채널 대역폭을 의미한다. 가중치 α 와 β 는 RTT와 RSSI의 상대적 중요도를 조절하며, 이러한 복합 보상 함수를 통해 DQN 에이전트는 통신 지연과 신호 품질을 동시에 최적화하는 BR 배치를 학습한다. SNR은 식 (4)와 같이 송신 전력, 안테나 이득, 경로 손실로부터 계산되며, 경로 손실 PL은 식 (5)의 log-distance 모델을 따른다. 식 (6)은 노드와 BR 간의 유clidean 거리를 나타낸다.

$$R_{rtt}(x_{br}, y_{br}) = \frac{1}{nodes} \sum_{n=1}^{nodes} \left\{ \frac{1}{T_{rtt}(n, br)} \times 1000 \right\} \quad (1)$$

$$R_{rss}(x_{br}, y_{br}) = \frac{1}{nodes} \sum_{n=1}^{nodes} \left\{ P_{tx} - PL_{(n, br)} \right\} \quad (2)$$

$$T_{rtt}(n, br) = \frac{p}{B \times \log_2(1 + SNR(n, br))} \quad (3)$$

$$SNR_{(n, br)} = P_{tx} + G_t + G_r - PL_{(n, br)} - noise_{floor} \quad (4)$$

$$PL = 20 \log f + N \log d_{(n, br)} + L_{f(n, br)} - 28 \quad (5)$$

$$d_{(n, br)} = \sqrt{(x_n - x_{br})^2 + 4 \times (y_n - y_{br})^2} \quad (6)$$

IV. 성능 분석

그림 2는 각 라우팅 프로토콜 환경에서 DQN 에이전트의 학습 과정을 나타낸다. 모든 프로토콜에서 약 10,000 timestep 이후 보상이 수렴하는 것을 확인할 수 있으며, 이는 에이전트가 각 프로토콜 특성에 적합한 BR 배치 정책을 성공적으로 학습했음을 의미한다. 제안하는 RPL 기반 방식이 가장 높은 누적 보상(약 1.4)을 달성하였으며, LOAD와 CTP가 그 뒤를 따른다. RPL의 우수한 성능은 DODAG 구조가 BR 중심의 계층적 토플로지를 형성하여 DQN의 배치 최적화와 높은 시너지를 발휘하기 때문으로 분석된다.

그림 3은 100개의 랜덤 노드 배치 시나리오에서 각 라우팅 프로토콜의 RSSI와 RTT 성능을 box plot으로 나타낸다. RSSI 측면에서 RPL이 평균 -82.2 dBm으로 가장 우수한 신호 품질을 보였으며, LOAD(-82.8 dBm), CTP(-83.1 dBm) 순으로 나타났다. RPL의 RSSI 성능이 우수한 이유는 objective function 기반 부모 선택 과정에서 신호 품질이 양호한 경로를 우선적으로 선택하기 때문이다. RTT 측면에서도 RPL이 평균 3.15 ms로 가장 낮은 지연을 달성하였으며, CTP(3.44 ms), LOAD(3.50 ms)는 유사한 수준의 RTT를 보였다. RPL의 낮은 RTT는 DODAG 구조를 통해 BR

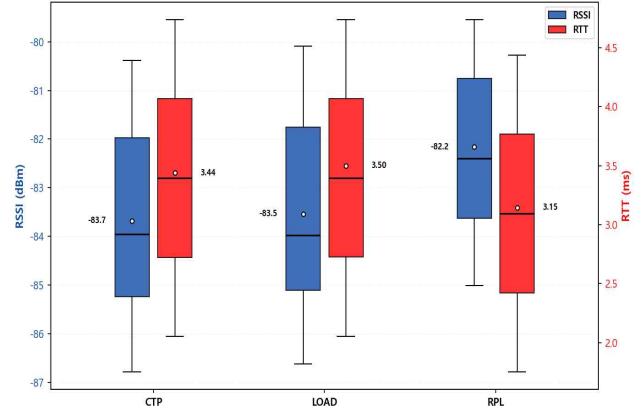


그림 3. 라우팅 알고리즘에 따른 성능 분석

표 1. 강화학습 시뮬레이션 파라미터

Parameter	Value
Batch size	1024
Buffer size	10240
Learning rate	0.001
Beta	0.0005
Epsilon	0.2
Lambda	0.99
Number of epochs	50000

까지의 경로가 사전에 최적화되어 있어 패킷 전송 시 추가적인 경로 탐색 오버헤드가 발생하지 않기 때문에 분석된다. 종합적으로, 제안된 DQN 기반 BR 배치 최적화 프레임워크는 RPL 환경에서 가장 효과적으로 동작하며, RSSI와 RTT 모두에서 다른 프로토콜 대비 우수한 성능을 달성함을 확인하였다.

V. 결론

본 논문에서는 Wi-SUN 네트워크의 BR 배치 최적화를 위한 DQN 기반 프레임워크를 제안하였다. RTT와 RSSI를 결합한 보상 함수를 설계하여 통신 지연과 신호 품질을 동시에 최적화하는 배치 정책을 학습하였으며, RPL, CTP, LOAD 환경에서 성능을 비교 분석하였다.

실험 결과, RPL 기반 환경에서 제안 프레임워크가 가장 우수한 성능을 보였으며, 평균 RSSI -82.2 dBm, RTT 3.15 ms를 달성하여 다른 프로토콜 대비 신호 품질과 지연 측면에서 모두 개선된 결과를 확인하였다. 이는 RPL의 DODAG 구조가 BR 중심 최적화와 높은 시너지를 발휘하기 때문으로 분석된다.

ACKNOWLEDGMENT

이 성과는 정부(과학기술정보통신부)의 재원으로 한국연구재단의 지원을 받아 수행된 연구임(RS-2024-00358662)

참 고 문 현

- [1] H. Kharrufa, H. A. A. Al-Kashoash and A. H. Kemp, "RPL-Based Routing Protocols in IoT Applications: A Review," *IEEE Sensors Journal*, vol. 19, no. 15, pp. 5952–5967, Aug, 2019
- [2] R. Hirakawa, K. Mizutani, and H. Harada, "Specification and Performance Analysis of Wi-SUN FAN," *IEEE Internet of Things Journal*, vol. 10, no. 20, pp. 18120–18134, Oct, 2023