

모델 기반 강화학습을 활용한 실내 AP 배치 기법

황준규, 리진룡, 박준구*
경북대학교 전자전기공학부

junkyu891015@gmail.com, {2023000464, jgpark*}@knu.ac.kr

An Access Points Placement Method using Model based Reinforcement Learning For Indoor Environment

Hwang Jun Gyu, JinLong Li, Park Joon Goo*
Kyungpook National University, School of Electronic & Electrical Engineering

요약

실내 위치 기반 서비스(Positioning Service)의 정확성은 Wi-Fi Access Point(AP)의 배치에 따라 크게 좌우된다. 기존 연구는 무선 신호의 지문(fingerprint)을 활용하여 위치 정확도를 향상시키기 위해 다양한 최적화 기법을 적용해 왔으나, 실시간 환경 변화에 대응하는 적응성이 부족하다는 한계를 가진다. 본 논문에서는 모델 기반 강화학습(Model-Based Reinforcement Learning, MBRL)을 이용하여 AP의 배치를 최적화하는 기법을 제안한다. 제안된 기법은 환경 모델을 통해 예상되는 위치 정확도를 미리 시뮬레이션하여 실제 환경에서의 탐색 비용을 최소화하고 빠르게 최적의 AP 배치를 도출한다. 또한, 실제 실내 환경을 모델링하여 수행한 시뮬레이션 결과, 기존의 강화학습 기법에 비해 학습 효율과 위치 정확도 모두 개선됨을 확인하였다.

I. 서론

최근 스마트 빌딩, 스마트 팩토리 등 실내 환경에서 정확한 위치 정보를 요구하는 서비스가 확대되면서, 무선 신호를 활용한 실내 위치 기반 서비스(Indoor Positioning Service, IPS)에 대한 관심이 증가하고 있다. IPS는 Wi-Fi, 블루투스, UWB(Ultra Wide Band) 등의 다양한 기술을 활용할 수 있지만, 기존 인프라 활용이 용이한 Wi-Fi 기반 방식이 가장 널리 사용된다. [1]

Wi-Fi 기반 IPS의 정확도는 주로 AP(Access Point)의 배치 전략에 크게 의존한다. 특히, RSSI(Received Signal Strength Indicator) 기반 지문(fingerprint) 측위 방식에서는 AP의 개수와 위치 설정이 서비스 성능에 결정적인 영향을 미친다. 기존 연구에서는 모델프리 강화학습 기법을 이용하여 AP 배치를 최적화하는 시도가 이루어졌다. 그러나 기존의 접근 방식은 실제 환경에서의 시행착오를 통해 학습하기 때문에 높은 탐색 비용과 느린 학습 속도라는 단점을 가지고 있다.[2][3]

본 연구에서는 이러한 문제를 해결하기 위해 모델 기반 강화학습(Model-Based Reinforcement Learning, MBRL)을 제안한다. MBRL은 실제 환경에서의 시행착오를 줄이기 위해 환경에 대한 명시적인 모델을 구축하고 이를 통해 시뮬레이션된 경험을 활용하여 학습 효율을 높인다. 환경 모델을 통해 사전 시뮬레이션을 수행하므로 실제 환경에서의 탐색 비용을 크게 감소시키고, 빠르게 최적의 AP 배치를 도출할 수 있다.

본 논문은 실내 환경의 무선 신호 특성을 모델 기반 강화학습 프레임워크에 통합하여 AP 배치 최적화 문제를 해결하는 접근법을 제안한다. 실내 환경을 대상으로 한 시뮬레이션을 통해 제안한 방법의 성능을 검증하고 기존 방법과의 성능 비교 분석을 수행한다.

II. AP 배치 최적화를 위한 모델 기반 강화학습

모델 기반 강화학습을 AP 배치에 사용하기 위해 측위 성능지표가 필요하다. 환경 안에서 참조지점에서 각 AP마다 측정된 RSSI를 이용하여 무선 신호 특성을 반영한다. 그리고 DoP(Dilution of Precision)을 이용하여 AP 간의 상대위치에 따른 측위성능을 반영한다. 식 (1)은 앞서 언급한 RSSI와 DoP를 적용한 측위 성능지표이다.

$$Q_i = \{\sum_j (RSSI_j)\} * DoP_i \quad (1)$$

Q_i 는 실내 환경에서 i 번째 참조지점의 측위 성능지표이다. $RSSI_j$ 는 i 번째 참조지점에서 j 번째 AP의 RSSI를 나타낸다. DoP_i 는 i 번째 참조지점의 DoP를 나타낸다. Q_i 를 가치함수로 사용하여 강화학습을 진행한다.

$$L = \frac{1}{N} \sum_i (y_i - Q(s_i, a_i | \theta^Q))^2 \quad (2)$$

$$y_i = r_i + \gamma Q'(s_{i+1}, \mu'((s_{i+1} | \theta^{\mu'})) | \theta^Q) \quad (3)$$

식 (3), (4)에서 L 은 Critic 손실함수이며 y_i 는 타겟이다. r_i 는 보상 네트워크, γ 는 할인 계수이다. Q' , μ' 는 타겟 행동 및 비평 네트워크로 느리게 갱신되는 별도의 네트워크이다. $\theta^{\mu'}$, θ^Q' 는 각 네트워크의 가중치이다. 마지막으로 행동 네트워크 업데이트는 다음 식(5)과 같다

$$\nabla_{\theta^{\mu}} J \approx \frac{1}{N} \sum_i \nabla_a Q(s, a | \theta^Q) |_{s=s_i, a=\mu(s_i)} \nabla_{\theta^{\mu}} \mu(s | \theta^{\mu}) |_{s_i} \quad (4)$$

$Q(s, a | \theta^Q)$ 은 비평단계에서 평가한 행동 가치이며, $\mu(s | \theta^{\mu})$ 는 행동 네트워크에서 상태 s 에서 행동 a 로 결정되는 정책이다.

본 연구에서는 강화학습의 Action으로 AP의 위치 변화, Reward는 측위 성능 향상, State는 AP의 위치와 환경요소들을 사용한다.

III. 시뮬레이션 및 결과

시뮬레이션 환경은 $20m \times 20m$ 공간으로 가정하였다. 모델 기반 강화학습 AP 위치조정은 Python을 통해 진행하였고 Fingerprint-Map을 이용한 측위 시뮬레이션은 Matlab을 사용하였다.

그림 1.은 기준 기법별 최적 AP를 찾는 과정을 나타낸다. 초기에 램덤하게 AP들이 배치되다 보니 처음 학습 과정 중에 이동이 크다가 점점 수렴하는 움직임을 보인다. 표 1에서는 최종 측위 성능과 학습시간이다. 최종 측위 성능은 DDPG과 뛰어나지만 총 학습시간이 약 1.9 배 길어짐을 볼 수 있다.

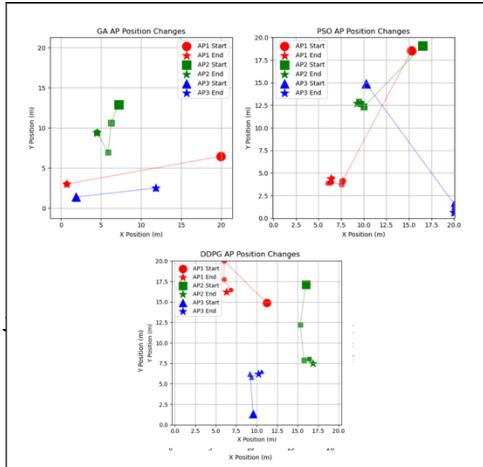


그림 2.GA, PSO, DDPG 최적 AP 배치 과정

표 1. 기법별 측위성능 및 학습시간

기법	측위성능	학습시간
GA	2.19 m	753.8 s
PSO	2.07 m	822.4 s
DDPG	1.71 m	1534.1 s
제안한 방법	1.76 m	789.5 s

그림 2.는 2 장에서 제안한 모델 기반 강화학습을 적용했을 때 최적 AP 위치를 찾는 과정이다. 기존 모델 브리 강화학습 기법 보다 다양한 시도를 하며 최적의 AP 위치를 찾음을 보인다. 이는 보다 빠르게 최적의 AP를 찾는 결과를 나타낸다.

표 1.에서 보이듯이 제안한 방법으로 측위성능은 DDPG를 사용했을 때보다는 0.05 m 증가하지만 학습시간이 750 초 가량 감소한다.

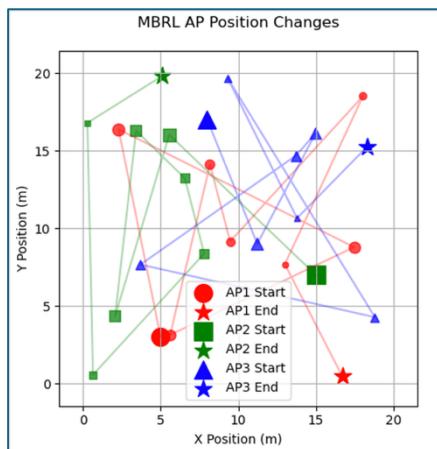


그림 3. 제안한 방법 최적 AP 배치 과정

IV. 결론

본 논문에서 모델 기반 강화학습을 활용한 실내 AP 배치 기법을 제안했다. 유전 알고리즘, 입자 군집 최적화, 심층 결정론적 정책 경사 기법인 모델 브리 강화학습과 성능비교를 진행했으며 모델 기반 강화학습이 더 빠르고 정확한 AP의 최적 위치를 찾음을 확인했다. 향후 다양한 환경요소들을 고려하고 각각의 요소들을 입력하면 최적기지국 위치가 출력으로 나오는 함수(Net)를 연구할 예정이다.

ACKNOWLEDGMENT

이 논문은 정부(과학기술정보통신부)의 재원으로 정보통신기획평가원-지역지능화혁신인재양성사업의 지원을 받아 수행된 연구임(IITP-2025-RS-2022-00156389).

참 고 문 헌

- [1] Farahsari, Pooyan Shams, et al. "A survey on indoor positioning systems for IoT-based applications." *IEEE Internet of Things Journal* 9.10 (2022): 7680-7699.
- [2] Liu, Jiandong, et al. "Unmanned Aerial Vehicle Path Planning in Complex Dynamic Environments Based on Deep Reinforcement Learning." *Machines* 13.2 (2025): 162.
- [3] Wang, Na, et al. "Placement of Access Points for Indoor Positioning based on DDPG." *Proceedings of the 2020 the 4th International Conference on Innovation in Artificial Intelligence*. 2020.