

6G 무선접속망을 위한 AI/ML 기반 지능형 RAN 기술 동향

Trends in AI/ML-Based Intelligent Open RAN for 6G Wireless Access Networks

김민건 (M.G. Kim, mingun.kim@etri.re.kr)

지능무선액세스연구실 연구원

김준우 (J.W. Kim, jwkim@etri.re.kr)

지능무선액세스연구실 책임연구원

이훈 (H. Lee, hlee@etri.re.kr)

지능무선액세스연구실 책임연구원

배정숙 (J.S. Bae, jsbae@etri.re.kr)

지능무선액세스연구실 책임연구원/실장

김일규 (I.G. Kim, igkim@etri.re.kr)

이동통신연구본부 책임연구원/본부장

ABSTRACT

The transition from 5G to 6G is closely tied to a shift from connectivity-centric infrastructures to AI-native radio access networks. In these networks, machine learning is deeply integrated into the architecture and operation of the wireless access segment, and learning-based functions span from spectrum sensing and physical-layer signal processing to near real-time and non-real-time control applications running in cloud-native controllers. This paper surveys AI/ML-based technologies for wireless access networks, focusing on recent work in AI radio access networks (AI-RAN), open radio access networks (O-RAN), RAN intelligent controllers (RICs), real-time dApps/xApps, digital twins, and machine learning-based physical-layer receivers. Building on data and model assets such as DeepMIMO for channel modeling, LibIQ for real-time spectrum analytics, and O-RAN-driven digital twin platforms, we review online learning and reinforcement learning approaches for energy-efficient resource allocation, mobility and spectrum management as well as spiking neural network-based receivers for low-power physical-layer processing. We also highlight the roles of explainable and generative AI in making RAN intelligence trustworthy, sustainable, and easier to operate. Finally, we identify future directions for designing AI-native wireless access networks.

KEYWORDS AI-RAN, dApp/xApp/rApp, DeepMIMO, LibIQ, O-RAN, RIC, SORA, SpikingRx, XAI, 디지털 트윈, 생성형 AI

* DOI: <https://doi.org/10.22648/ETRI.2025.J.410101>

* 본고는 2024년도 정부(과학기술정보통신부)의 재원으로 정보통신기획평가원의 지원을 받아 수행된 연구임[No. RS-2024-00435652, 6GARROW: 6G Ai-native 통합 RAN-Core 네트워크].

I. 서론

5세대(5G: 5th Generation) 이동통신의 상용화 이후 무선접속망(RAN: Radio Access Network)은 초고속·초저지연·초연결을 달성했으나, 6세대(6G: 6th Generation) 시대에는 단순한 접속 인프라를 넘어 네트워크 자체가 학습하고 스스로 동작을 조정하는 지능형 인프라가 요구되고 있다. 이러한 흐름 속에서 인공지능(AI)과 머신러닝(ML: Machine Learning)을 기지국 기능 전반에 깊게 결합한 AI 기반 무선접속망(AI-RAN: AI Radio Access Network) 개념이 핵심 비전으로 제시된다[1].

한편, 개방형 무선접속망(O-RAN: Open Radio Access Network) 표준은 기지국 기능을 개방형 인터페이스를 통해 모듈화하고, O-RU, O-DU, O-CU로 기능을 분리하여 여러 벤더의 장비와 소프트웨어가 공존하는 생태계를 지향한다. 특히 무선접속망 지능형 제어기(RIC: RAN Intelligent Controller)에 탑재되는 근실시간 응용(xApp: x Application)과 비실시간 응용(rApp: r Application), 그리고 데이터 경로와 더 밀접한 다이렉트 응용(dApp: direct Application)은 AI/ML 알고리즘을 네트워크 제어 루프 내부로 끌어들이는 중요한 수단이다[2-4,9].

또한, 실제 환경을 반영한 채널 데이터셋과 스펙트럼 데이터셋, 디지털 트윈(DT: Digital Twin) 기반 시뮬레이터의 발전으로, RAN 제어 문제를 데이터 중심으로 재정의하려는 시도가 빠르게 확산되고 있다. DeepMIMO와 LibIQ 같은 데이터·모델 자산 및 O-RAN 디지털 트윈 플랫폼은 AI-RAN 연구를 뒷받침하는 대표적인 인프라로 활용되고 있다[4,7,8].

본고에서는 최신 연구를 바탕으로, 6G 지향 무선접속망을 위한 AI/ML 기술을 다음과 같은 관점에서 정리한다.

- 6G AI-RAN의 비전과 O-RAN 아키텍처
- 채널·스펙트럼·디지털 트윈 데이터 및 모델 자산
- dApp/xApp/rApp 기반 실시간·장기 RAN 제어
- 디지털 트윈 및 통합 센싱·통신(ISAC: Integrated Sensing And Communications) 응용
- 물리계층(PHY: Physical Layer) 수신기 차원의 AI/ML 적용과 에너지 효율
- 설명 가능한 인공지능(XAI: eXplainable Artificial Intelligence), 생성형 AI 및 대규모 언어모델(LLM: Large Language Model) 기반 RAN 지능화

II. 6G 지향 AI-RAN 개념과 O-RAN 진화

1. 6G 무선접속망의 요구 조건

6G에서는 5G 대비 크게 향상된 용량과 데이터율, 극저지연, 고신뢰성, 초고밀도 단말 접속, 그리고 디지털 트윈 서비스·몰입형 확장현실 등 새로운 응용을 동시에 지원해야 한다[1]. 이러한 요구는 다음과 같은 특성의 무선접속망 설계를 요구한다.

첫째, 밀리미터파(mmWave: millimeter Wave) 및 테라헤르츠(THz: Terahertz) 대역, 대규모 다중입출력(MIMO: Multiple-Input Multiple-Output), 지능형 반사표면(IRS: Intelligent Reflecting Surface), 비직교 다중접속(NOMA: Non-Orthogonal Multiple Access) 등 다양한 무선 기술을 통합적으로 운용해야 한다[1,7]. 둘째, 마이크로셀·피코셀뿐만 아니라 비지상 네트워크(NTN: Non-Terrestrial Network)까지 포함하는 초고밀도·이기종 셀 구조에서 간섭과 핸드오버를 정교하게 관리해야 한다. 셋째, 초신뢰·초저지연 통신(URLLC: Ultra-Reliable and Low Latency Communications), 고신뢰 대용량 통신(eMBB: enhanced Mobile Broadband), 초대규모 사물통신(mMTC: massive Ma-

chine-Type Communications) 등 다양한 서비스별 품질 요구(QoS: Quality of Service; QoE: Quality of Experience)를 만족시키기 위해, 슬라이스 단위의 세밀한 자원 관리와 서비스 수준 합의(SLA: Service Level Agreement) 보장이 필요하다.

이처럼 상태 공간이 방대하고 시간·공간적으로 급변하는 환경에서는, 고정 규칙 기반 설계와 수동 튜닝만으로는 효율적인 운용이 어렵다. 관측 데이터로부터 경험을 축적하고 정책을 지속적으로 학습·갱신하는 AI/ML 기반 제어가 자연스러운 해결책으로 부상하는 이유가 여기에 있다[1].

2. AI-RAN: 학습하는 무선접속망

AI-RAN은 RAN 내부의 다양한 기능, 예를 들어 채널 추정과 빔포밍, 스케줄링, 핸드오버, 슬라이싱, 에너지 관리 등 전 영역에 AI/ML을 단계적으로 주입하여, 사람이 일일이 규칙을 설계하던 구조를 학습 기반 자율 제어 시스템으로 전환하는 개념이다[1].

이를 위해 AI-RAN은 다음과 같은 계층 구조를 가질 수 있다.

- **관측 계층:** 기지국과 단말에서 수집되는 방대한 키 성능 지표(KPM: Key Performance Metric), 채널 상태 정보(CSI: Channel State Information), 로그 데이터를 통합·정제하여 AI 학습에 적합한 형태로 제공한다.
- **지능 계층:** 예측·분류·강화학습(RL: Reinforcement Learning) 모델을 활용하여 자원 할당, 셀 on/off, 핸드오버 등 다양한 의사결정 문제를 학습한다.
- **실행 계층:** 학습된 정책을 xApp/dApp/rApp 형태로 RIC 및 O-DU/O-CU에 배치하고, 초저지연·근실시간·장기 제어를 계층적으로 수행한다.

결국 AI-RAN 설계의 핵심은 어떤 위치에 어떤 모델을 배치하고, 어떤 제어 루프를 맺길 것인가라는 문제이며, 이때 O-RAN 아키텍처가 AI 모듈을 삽입할 수 있는 구조적 기반을 제공한다.

3. O-RAN 아키텍처와 RIC의 역할

O-RAN은 RAN 기능을 분리하고 개방형 인터페이스를 정의함으로써 AI/ML 기능을 소프트웨어 컴포넌트 형태로 유연하게 삽입할 수 있는 환경을 제공한다[3,4].

O-RU, O-DU, O-CU로의 기능 분리는 하드웨어와 소프트웨어, 그리고 기능 블록 간 결합도를 낮춰 특정 모듈만 교체하거나 확장하기 쉽게 만든다. 비실시간 영역에는 비실시간 RIC(Non-RT RIC)가 위치하며, 비실시간 rApp이 장기적인 데이터 분석, 정책 생성, 모델 학습·배포를 담당하고 관리 인터페이스(O1)를 통해 O-RAN 구성 요소를 제어한다[3]. 근실시간 영역에는 근실시간 RIC(Near-RT RIC)가 위치하고, 근실시간 xApp이 수 밀리초에서 1초 수준의 제어를 담당하며, E2 인터페이스를 통해 O-DU/O-CU와 상호작용을 하면서 스케줄링, 전력 제어, 간섭 조정 등 핵심 기능을 수행한다[2,4].

최근에는 dApp 개념이 도입되어, O-DU/O-CU 인근 혹은 데이터 경로 주변의 컴퓨팅 자원에서 실행되며 PHY 및 매체접속제어 계층과 더 직접적으로 상호작용하는 초경량 마이크로서비스 형태의 AI 모듈이 제안되고 있다. dApp은 수백 마이크로초에서 1밀리초 수준의 응답 시간이 필요한 스펙트럼 분류, 간섭 회피, 빠른 링크 적응 등에 적합한 구조를 제공한다[9].

정리하면 O-RAN은 AI-RAN 비전을 구현하기 위한 플러그인형 AI 제어 슬롯(xApp/rApp/dApp)을 네트워크 구조에 내장한 형태로 볼 수 있으며, 이를 통

해 다양한 AI/ML 알고리즘을 RAN 내부로 유연하게 통합할 수 있다[1,3].

III. 무선접속망 AI/ML을 위한 데이터·모델 자산

AI/ML 기반 RAN 제어가 실용화되기 위해서는, 다양한 환경을 반영한 데이터와 그 데이터를 기반으로 동작할 수 있는 모델·플랫폼이 충분히 마련되어야 한다. 대표적으로 DeepMIMO, LibIQ, O-RAN 디지털 트윈 플랫폼 등이 이러한 역할을 수행하고 있다.

1. 채널 데이터셋: DeepMIMO

DeepMIMO는 레이 트레이싱 기반 시뮬레이션으로 생성된 mmWave 및 대규모 MIMO 채널 데이터를 딥러닝 학습에 적합한 형태로 제공하는 범용 데이터셋이다[7]. 이 데이터셋의 특징은 다음과 같다.

첫째, 실제 환경을 기반으로 한 3차원 모델 위에서 레이 트레이싱을 수행하여 현실감 있는 가시선 및 비가시선 채널을 생성한다. 둘째, 기지국(BS: Base Station)과 단말(UE: User Equipment)의 위치, 안테나 수, 범 형성 방식 등 여러 파라미터를 사용자가 설정할 수 있어 다양한 시나리오를 손쉽게 구성할 수 있다. 셋째, 범 선택과 범 추정, 채널 예측, 위치 추정 등 여러 AI 문제를 동일한 데이터셋에서 다룰 수 있도록 설계되어, 연구자들이 공통 기반 위에서 알고리즘을 비교·평가하기 용이하다.

따라서 DeepMIMO는 mmWave·대규모 MIMO 환경에서의 AI 모델 연구를 위한 공용 기초 데이터셋으로 활용 가능하며, 향후 6G용 범포밍·슬라이스 설계 등에도 중요한 역할을 할 것으로 기대된다.

2. 실시간 I/Q 데이터: LibIQ와 스펙트럼 인지

LibIQ는 RAN에서 관측되는 복소 I/Q 샘플을 수집·가공하여, 실시간 스펙트럼 분류와 간섭 탐지 모델 학습에 활용할 수 있도록 만든 라이브러리이다[8]. 다양한 변조 방식과 간섭 패턴, 신호대잡음비 등에서 I/Q 시퀀스를 수집하고, 이를 합성곱 신경망 등 딥러닝 모델 학습에 사용할 수 있도록 구조화한다.

학습된 모델을 dApp 형태로 O-RU/O-DU 근처에 배치하면, 스펙트럼 상태를 빠르게 분류하고 간섭이 의심되는 상황을 신속하게 감지할 수 있다[8,9]. 이 구조는 스펙트럼 센싱, AI 추론, RAN 파라미터 조정을 하나의 짧은 페루프로 묶어 동작하게 하며, 동적 스펙트럼 공유나 간섭 회피와 같은 응용에 적합하다.

3. O-RAN 디지털 트윈 기반 데이터

O-RAN 환경에서 에너지 절감과 지연 최적화를 연구하기 위해, 디지털 트윈 기반 플랫폼을 구성한 사례도 제시된다[4]. 상용급 RAN 시뮬레이터를 사용하여 실제 O-RAN과 유사한 환경을 가상 공간에 구현하고, PRB(Physical Resource Block) 사용률, 셀·단말별 스루풋, 변조 및 부호화 방식(MCS: Modulation and Coding Scheme), RSRP/RSRQ, 지연 등 다양한 KPM을 비실시간 RIC로 수집하여 시계열 데이터베이스에 저장한다.

이 플랫폼에서는 장기 메모리를 가진 순환신경망 계열 모델을 활용한 LSTM 기반 트래픽 예측 rApp과 에너지 절감 rApp을 디지털 트윈상에서 먼저 시험하고, 충분히 검증된 정책을 실제 테스트베드에 적용하는 방식으로 에너지·지연 트레이드오프를 분석한다[4]. 이러한 디지털 트윈 데이터는 실환경

과 매우 유사하지만 실험 실패에 따른 리스크가 없는 공간을 제공하여, AI 정책을 더욱 적극적으로 설계하고 튜닝할 수 있게 한다.

4. 데이터·모델 자산 비교

표 1은 앞에서 설명한 데이터·모델 자산을 용도 관점에서 비교한 것이다.

IV. AI/ML 기반 무선접속망 제어 기술

데이터와 플랫폼이 준비되면, 그 위에서 어떤 제어 알고리즘을 설계할 것인가가 핵심 질문이 된다. 최근 연구들은 dApp/xApp/rApp 계층 구조를 활용하여 자원 할당, 에너지 절감, 이동성 관리 등을 해결하는 다양한 접근을 제시한다.

1. dApp/xApp 기반 실시간 제어

dAPPs 관련 연구에서는 기존 O-RAN 구조에 dApp 계층을 추가하여, PHY 및 매체접속제어 계층과 더욱 밀접하게 연결된 초저지연 제어 루프를 구현한다[9]. dApp은 O-DU 인근의 컴퓨팅 리소스에

서 실행되며 사용자 데이터 또는 I/Q 샘플에 직접 접근할 수 있고, 마이크로서비스 구조로 설계되어 특정 기능을 독립적으로 배포·롤백할 수 있어 운영 유연성이 높다.

실험 결과에 따르면 dApp 기반 제어는 사용자 패킷 처리 시간대와 유사한 마이크로초~밀리초 수준에서 동작할 수 있어, 근실시간 RIC 기반 xApp보다 더 빠른 응답이 가능하다[9]. LibIQ 기반 스펙트럼 분류 모델을 dApp으로 구현하면, dApp이 스펙트럼 상태를 분류하고 그 결과를 O-DU 스케줄러나 xApp에 피드백하여 간접 회피 전략을 즉각 적용하는 구조를 만들 수 있다[8,9].

2. 온라인 학습 기반 자원·에너지 최적화: SORA

SORA는 O-RAN 환경에서 에너지 효율을 높이기 위해 자원 할당 문제를 온라인 학습 관점에서 다른 연구이다[10]. 사용자별 전력·PRB 할당을 컨텍스트가 있는 벤딧 문제로 보고, 각 사용자 상태(채널, 트래픽 등)를 컨텍스트로 사용하여 Thompson Sampling 기반 알고리즘을 통해 사전 지식 없이도 점진적으로 좋은 자원 할당 정책을 학습한다.

표 1 무선접속망 AI/ML 데이터·모델 자산 비교[4,7,8]

자산/플랫폼	주요 목적	데이터 유형	환경·플랫폼	특징
DeepMIMO	빔포밍, 채널·위치 예측	채널 계수, 위치, 안테나 구성	레이 트레이싱 기반 mmWave/대규모 MIMO	파라미터화된 대규모 채널 데이터 생성, 다양한 시나리오 구성 용이
LibIQ	스펙트럼 분류, 간접 팀지	I/Q 시퀀스, 레이블 (신호대집음비, 변조 등)	O-RAN dApp 환경	합성곱 신경망 학습에 최적화된 구조, 마이크로초~밀리초 수준 폐루프 스펙트럼 인지
O-RAN DT 에너지 절감 플랫폼	셀 on/off와 전력 제어를 통한 에너지 효율 최적화	셀·단말 KPM, 전력 소비, 트래픽 시계열	상용 RAN 시뮬레이터, 실제 O-RAN 테스트베드	LSTM rApp 기반 트래픽 예측, 비실시간 RIC 연계, O1 인터페이스 활용
5G DT O-RAN 드론 레이싱 플랫폼	저지연 스펙트럼 스케줄링	왕복 지연(RTD: Round-Trip Delay), 채널, 트래픽, 드론 궤적	무선 채널 에뮬레이터, 로봇 시뮬레이터, O-RAN	드론 경로·지연 조건을 통합한 디지털 트윈, xApp 평가에 적합

이 방식은 계산 복잡도가 상대적으로 낮고, 환경 모델(정확한 채널 모델 등)에 대한 가정이 약하다는 장점이 있다. 시뮬레이션 결과, SORA는 복잡한 최적화 기반 방식과 비교해도 경쟁력 있는 에너지 효율을 보이면서 구현 난이도를 낮출 수 있음을 보인다[10].

3. 디지털 트윈을 활용한 에너지·지연 제어

디지털 트윈을 활용하여 에너지 절감 및 지연 성능을 최적화하는 사례도 있다. O-RAN 디지털 트윈 연구에서는 coverage 셀과 capacity 셀로 구성된 환경에서, LSTM 기반 트래픽 예측 rApp이 수요를 예측하고 에너지 절감 rApp이 capacity 셀 on/off를 결정하여 에너지 절감을 도모한다[4]. 드론 레이싱 시나리오에서는 드론의 경로와 속도, 무선 채널 상태를 디지털 트윈에서 재현하고, RTD를 최소화하는 스펙트럼 스케줄링 xApp을 시험하여 근실시간 RIC가 지연 요구를 만족하는 자원 배분을 수행할 수 있음을 보인다[4].

이러한 방식은 디지털 트윈에서 위험 없이 정책을 충분히 검증한 뒤, 실제 O-RAN에 적용하는 운영 프로세스를 가능하게 함으로써, AI 기반 제어 도입에 따른 리스크를 줄이는 효과가 있다.

4. 이동성·스펙트럼 관리

AI-RAN 튜토리얼에서는 이동성 관리와 스펙트럼 관리 역시 AI/ML의 핵심 적용 분야로 제시된다 [1]. 혼잡 셀과 비혼잡 셀 사이의 단말 분배를 강화 학습 문제로 모델링하고, 장기적인 QoE와 부하 균형을 동시에 고려하는 정책을 탐색하는 연구가 보고된다. 또한 NOMA, IRS, mmWave/THz를 활용하는 복잡한 스펙트럼 공유 환경에서, 학습 기반 간섭 관리와 채널 예측을 활용해 자원 사용 효율을 높이는 방안이 제안된다[1,7].

실제 O-RAN 구조에서는 이러한 알고리즘을 xApp/rApp 형태로 구현하여, 핸드오버 파라미터, 스케줄링 가중치, 빔 선택 정책 등을 동적으로 조정할 수 있다.

5. AI 기반 RAN 제어 기술 분류

표 2는 RAN 제어 분야를 기능별로 정리한 것이다.

V. 디지털 트윈 및 ISAC 기반 지능형 RAN

디지털 트윈과 ISAC는 6G에서 자주 언급되는 키

표 2 AI/ML 기반 무선접속망 제어 기술 분류[1,3,4,7-10]

범주	대표 사례	주요 학습 기법	기대 효과
자원·스펙트럼 관리	AI-RAN 자원 관리, DeepMIMO 기반 빔포밍	딥러닝, 강화학습	스펙트럼 효율 향상, 사용자 QoE 개선, 간섭 완화
에너지 절감	SORA, O-RAN 디지털 트윈 기반 에너지 절감 rApp	Thompson sampling, LSTM 기반 예측, 정책 최적화	셀 on/off, 전력 제어를 통한 에너지 및 운영비용(OPEX: Operational Expenditure) 절감
초저지연 제어	dApp·LibIQ 기반 스펙트럼 제어	합성곱 신경망, 마이크로서비스 기반 배포	마이크로초~밀리초 폐루프 제어, 간섭 회피, 링크 안정성 향상
이동성·핸드오버	AI-RAN 이동성 관리	강화학습, 예측 분석	핸드오버 실패 및 핑퐁 감소, 부하 균형
DT-ISAC 제어	O-RAN 디지털 트윈, 로보틱 ISAC 디지털 트윈	LSTM, 대체 모델(Surrogate Model), 생성형 모델	에너지·지연·센싱 성능 동시 최적화

워드로, AI-RAN과도 밀접하게 연결된다. 디지털 트원은 현실 세계의 네트워크 상태를 가상 공간에 투영하는 역할을 하고, ISAC는 통신과 센싱을 통합함으로써 네트워크가 주변 환경을 더욱 정교하게 인지할 수 있게 한다.

1. 6G 디지털 트원과 생성형 모델

생성형 AI 튜토리얼 논문에서는 6G 디지털 트원을 구현하기 위한 도구로 생성형 적대 신경망(GAN: Generative Adversarial Network), 변분 오토인코더(VAE: Variational AutoEncoder), 흐름 기반 모델(Flow-Based Model), 확산(Diffusion) 모델 등을 언급한다[6]. 이들 모델은 다음과 같은 방식으로 활용될 수 있다.

첫째, 불규칙한 채널 상태나 트래픽 패턴을 데이터로부터 학습한 뒤 유사한 조건을 인공적으로 생성하여 테스트 시나리오를 확장할 수 있다. 둘째, 드문 이벤트(장애, 극단적인 혼잡 상태 등)를 합성해 실제로 자주 발생하지 않는 상황에 대해서도 제어 정책을 미리 훈련할 수 있다. 셋째, 디지털 트원 환경에서 다양한 가상 하루를 빠르게 생성하여 정책의 장기간 성능을 평가할 수 있다.

이와 같이 생성형 모델은 디지털 트원의 데이터 엔진 역할을 하며, 부족한 실측 데이터를 보완하는데 기여한다[6].

2. O-RAN 디지털 트원 사례

O-RAN과 디지털 트원을 결합한 연구에서는 다양한 사례가 소개된다[4]. 에너지 절감 시나리오에서는 coverage 셀과 capacity 셀 구조를 디지털 트원에 구성하고, rApp이 수집한 트래픽 · KPM 데이터를 기반으로 LSTM 모델을 학습해 capacity 셀의 on/off 정책을 최적화한다. 드론 레이싱 시나리오

에서는 로봇 시뮬레이터를 활용해 3차원 환경과 드론 궤적을 모델링하고, 무선 채널 에뮬레이터에서 채널 · 트래픽을 합성하여 xApp이 지연 요구를 만족하도록 스펙트럼을 배분하는 전략을 실험한다[4].

이러한 사례는 디지털 트원을 실험실에 존재하는 또 하나의 RAN으로 활용하는 방법을 보여준다. 운영자는 디지털 트원에서 다양한 xApp/rApp 조합을 시험한 뒤, 성능과 안정성을 확인하고 실제 네트워크에 적용할 수 있다.

3. 로보틱 ISAC 디지털 트원

로보틱 ISAC 디지털 트원 연구에서는 통신과 센싱을 동시에 수행하는 로봇 시스템에서 디지털 트원을 어떻게 활용할 수 있는지 보여준다[5]. RGB-D 카메라로부터 얻은 3차원 포인트 클라우드를 기반으로 환경을 재구성하고, 이 위에서 레이 트레이싱을 수행해 채널을 추정한다. 로봇의 이동, 장애물 위치, 반사 특성 등 환경 변화를 실시간으로 반영하며 채널 모델도 함께 개선한다.

또한, LLM 기반 음성 인터페이스를 도입하여, 사용자가 자연어로 로봇과 디지털 트원의 동작을 제어할 수 있는 데모를 구현한다[5]. 이러한 접근은 향후 RAN이 차량, 로봇, 드론 등과 결합할 때 통신 인프라가 환경 인지와 제어 루프에도 직접 참여할 수 있음을 시사한다.

4. 전파 모델링 가속을 위한 AI

AI 기반 기민 전파 모델링 연구에서는 디지털 트원에서 전파 모델을 빠르게 개선하기 위한 방법을 제안한다[12]. 계산량이 큰 레이 트레이싱 결과 일부를 학습 데이터로 사용하여 신경망 기반 대체 모

델을 학습한 뒤, 나머지 위치의 채널을 빠르게 예측하는 방식이다. 환경 변화가 발생하면 제한된 수의 레이 트레이싱 샘플로 대체 모델을 업데이트하여 디지털 트윈을 민첩하게 유지할 수 있다.

이와 같은 접근은 AI-RAN에서 채널 예측, 빔포밍, 자원 최적화를 실시간에 가깝게 수행할 수 있는 기반을 마련한다.

VI. 물리계층 AI/ML: 신경망 및 스파이킹 수신기

AI/ML은 상위 계층 제어뿐만 아니라 PHY에도 깊이 들어오고 있다. 특히 뉴로모픽 하드웨어와 결합 가능한 스파이킹 신경망(SNN: Spiking Neural Network) 수신기는 에너지 효율 측면에서 많은 관심을 받고 있다.

1. SpikingRx: 스파이킹 신경망 수신기

SpikingRx 연구는 기존 딥러닝 기반 수신기(NeuralRx)를 스파이킹 신경망으로 변환하여 에너지 소비를 크게 줄이면서도 성능을 유지하는 방법을 제안한다[11]. 기존 인공신경망(ANN: Artificial Neural Network)으로 설계된 수신기를 기반으로 스파이크 기반 표현으로 변환하는 절차를 제시하고, 양자화 인지 학습과 스파이크 코딩을 결합해 낮은 정밀도의 정수 연산으로도 충분한 성능을 확보한다.

이 수신기의 블록 오류율(BLER: Block Error Rate)은 채널에 대한 완전한 지식(Genie-aided)을 가정한 하한선에 근접하면서도, 뉴로모픽 하드웨어 상에서 높은 에너지 효율을 달성하는 것으로 보고된다[11]. 이러한 특성은 향후 O-RU나 단말 내 베이스밴드 처리에 저전력 AI 수신기를 탑재하는 현실적인 방향을 제시한다.

2. 채널 데이터셋과 PHY-AI 결합

DeepMIMO와 같은 채널 데이터셋, 그리고 AI 기반 전파 모델링을 활용한 디지털 트윈 기반 채널 생성 기법은 SpikingRx와 같은 수신기를 학습·검증하는 환경을 제공한다[7,11,12]. DeepMIMO를 활용하면 다양한 빔포밍 조건과 채널 상황에서 NeuralRx/SpikingRx를 사전 학습하고, 특정 환경에 맞게 미세 조정할 수 있다[7,11]. 또한, 대체 모델 기반 전파 모델링을 사용하는 디지털 트윈에서는 위치·환경 변화에 따라 채널이 동적으로 생성되므로, 실제 환경과 유사한 조건에서 수신기 성능을 장기간 평가할 수 있다[12].

이와 같이 PHY-AI 모델과 채널 데이터/모델 자산이 결합되면, 상위 계층 AI-RAN 제어와 더불어 PHY까지 포함하는 종단간 지능형 무선 링크 설계가 가능해진다.

VII. XAI를 통한 신뢰 가능한 AI-RAN

AI가 네트워크 제어에 깊게 관여할수록 왜 이런 결정을 내렸는가를 설명하는 능력이 중요해진다. 특히 O-RAN과 같이 여러 벤더의 xApp/rApp이 공존하는 환경에서는 설명 가능성과 검증 가능성이 핵심 요구가 된다.

1. XAI의 필요성

XAI 튜토리얼에서는 6G O-RAN에서 XAI가 필요한 이유를 운영·튜닝, 규제·책임성, MLOps 관점에서 정리한다[2]. 운영자는 xApp/rApp이 내리는 결정을 이해해야 적절히 파라미터를 조정하고 문제 상황에서 원인을 진단할 수 있으며, 공공 안전과 산업 제어, 의료 등 분야에서는 AI 의사결정을 설명할 수

있어야 규제와 책임 소재를 명확히 할 수 있다. 또한, 모델 드리프트나 편향을 탐지하고 적절한 재학습·배포 전략을 수립하기 위해, 모델의 행동이 어떤 특징에 의해 좌우되는지 파악할 필요가 있다[2].

2. O-RAN 아키텍처와 XAI 통합

해당 튜토리얼은 LIME, SHAP, integrated gradients 등 널리 알려진 XAI 기법을 O-RAN에 통합하는 방법을 제안한다[2]. 비실시간 RIC에 XAI 모듈을 두고 xApp/rApp의 입력·출력과 KPM 변화를 수집해 정책이 어떤 입력 패턴에 민감한지 분석하며, Global 설명(정책 전반의 일반적 행동 양상)과 Local 설명(특정 단말·셀·시간 구간에서의 상세한 결정 근거)을 병행해 대시보드를 구성하는 구조를 예시로 제시한다.

이러한 구조는 AI-RAN의 블랙박스를 투명하게 만들 수 있는 도구 상자에 해당하며, 향후 상용망 적용 시 필수 기능으로 자리 잡을 가능성이 크다. 또한, XAI 지표를 이용해 편향이나 이상 행동이 감지되면 rApp이 자동으로 재학습·정책 롤백을 수행하는 MLOps 시나리오도 구상할 수 있다[2,3].

VIII. 생성형 AI 기반 RAN 지능화

생성형 AI와 LLM은 무선접속망에서도 새로운 활용 가능성을 열어가고 있다[6].

1. 데이터 증강과 시나리오 생성

GAN, VAE, 확산 모델 등은 희귀한 상황이나 복잡한 트래픽 패턴을 모사하는 데 유용하다. 평상시에는 잘 관찰되지 않는 장애 상황을 합성해 RL 기반 제어 정책을 사전에 훈련할 수 있고, 트래픽 부하와 채널 환경, 사용자 이동 패턴을 조금씩 바꿔가며 다

양한 가상 시나리오를 생성하여 xApp/rApp의 일반화 성능을 평가할 수 있다. 디지털 트윈과 결합하면 현실에서는 실험하기 부담스러운 극단 조건을 대량으로 시험하는 것도 가능하다[4,6,7].

2. LLM 기반 운영자 인터페이스 및 자동화

LLM은 네트워크 운영과 설계 과정에도 자연스럽게 스며들 수 있다[5,6]. 운영자는 자연어 질의를 통해 특정 셀의 핸드오버 실패율이 높은 이유를 요약해 달라고 요청하거나, 특정 기간의 KPM 변화를 설명할 것을 요구할 수 있다. LLM은 xApp/rApp 설정 템플릿, 테스트 시나리오, KPI 대시보드 구성을 자동 제안함으로써 설계·운영 업무를 보조할 수 있다. 로보틱 ISAC 디지털 트윈 사례처럼 음성 인터페이스를 통해 사람-네트워크-로봇의 상호작용을 자연어 기반으로 구성하는 것도 가능하다[5].

이러한 흐름은 AI가 AI를 관리하는 방향으로 진화할 수 있으며, XAI와 결합하면 설명 가능한 자동화 어시스턴트로 발전할 여지가 있다[2,6].

IX. 결론

본고에서는 AI-RAN, O-RAN, 디지털 트윈, PHY-AI, XAI, 생성형 AI와 관련된 최신 연구를 바탕으로, 6G 지향 무선접속망을 위한 AI/ML 기술 동향을 정리하였다. 무선접속망을 위한 AI/ML 기술은 개별 알고리즘이나 특정 계층에 국한되지 않고, 데이터-모델-플랫폼-운영-설명성을 아우르는 전체 생태계를 포괄하는 방향으로 진화하고 있다. 이러한 관점에서, AI-RAN과 O-RAN의 결합, 데이터·모델 자산의 체계화, 온라인 학습과 지속가능성, PHY-AI와 뉴로모픽 수신기, 디지털 트윈·ISAC·생성형 AI의 융합, XAI 및 LLM 기반 운영자

동화는 향후 6G RAN 연구 및 실증 프로젝트의 핵심 축이 될 것으로 전망된다.

본고에서 정리한 내용은 향후 6G 무선접속망 연구와 실증 사업에서 어떤 데이터·플랫폼·알고리즘을 우선적으로 준비해야 할지 전략을 세우는데 기초 자료로 활용될 수 있을 것으로 기대한다.

용어해설

AI-RAN 무선접속망의 자원관리·핸드오버·에너지 제어 등을 인공지능·머신러닝으로 수행하여, 망이 스스로 상태를 학습하고 최적화하도록 설계한 차세대 지능형 RAN 구조

O-RAN 기지국 기능을 O-RU·O-DU·O-CU로 분리하고 개방형 인터페이스와 RIC를 통해 여러 벤더 장비와 AI 기반 제어 응용을 유연하게 수용하는 개방형 무선접속망 아키텍처

디지털 트윈 무선네트워크 실제 무선망의 구성·채널·트래픽 상태를 가상 공간에 복제하여, 에너지 절감이나 지연 최적화 같은 제어 알고리즘을 안전하게 시험·검증할 수 있는 무선 네트워크 기술

참고문헌

- [1] N. A. Khan and S. Schmid, "AI-RAN in 6G Networks: State-of-the-Art and Challenges," *IEEE Open J. Commun. Soc.*, vol. 5, 2023, pp. 294-311.
- [2] B. Brik et al., "Explainable AI in 6G O-RAN: A tutorial and survey on architecture, use cases, challenges, and future research," *IEEE Commun. Surv. Tutor.*, vol. 27, no. 5, 2024, pp. 2826-2859.
- [3] H. Ahmadi et al., "Towards Sustainability in 6G and beyond: Challenges and Opportunities of Open RAN," *IEEE Commun. Stand. Mag.*, vol. 9, no. 3, 2025, pp. 126-135.
- [4] L. Zhou et al., "Digital Twins Meet Open RAN: Case Studies, Implementation, and Opportunities," *IEEE Commun. Mag.*, vol. 63, no. 8, 2025, pp. 162-168.
- [5] V. C. Andrei et al., "Demonstration of a continuously updated, radio-compatible digital twin for robotic integrated sensing and communications," in *Proc. IEEE Int. Symp. Joint Commun. Sens.*, (Oulu, Finland), Jan. 2025, pp. 1-2.
- [6] A. Celik and A. M. Eltawil, "At the dawn of generative AI era: A tutorial-cum-survey on new frontiers in 6G wireless intelligence," *IEEE Open J. Commun. Soc.*, vol. 5, 2024, pp. 2433-2489.
- [7] A. Alkhateeb, "DeepMIMO: A generic deep learning dataset for millimeter wave and massive MIMO applications," *arXiv preprint*, 2019. doi: 10.48550/arXiv.1902.06435
- [8] F. Olimpieri et al., "LibIQ: Toward Real-Time Spectrum Classification in O-RAN dApps," *arXiv preprint*, 2025. doi: 10.48550/arXiv.2505.10537
- [9] A. Lacava et al., "dApps: Enabling real-time AI-based Open RAN control," *Comput. Netw.*, vol. 269, 2025.
- [10] A. Tuinxun and A. Nakao, "SORA: Energy Efficient Resource Allocation in Open Radio Access Network with Online Learning," *IEEE Access*, vol. 13, 2025, pp. 113686-113701.
- [11] A. Gupta et al., "SpikingRx: From neural to spiking receiver," *arXiv preprint*, 2024. doi: 10.48550/arXiv.2409.05610
- [12] A. Saeizadeh et al., "AI-assisted Agile Propagation Modeling for Real-time Digital Twin Wireless Networks," in *Proc. IEEE Int. Workshop Comput. Aided Model. Des. Commun. Links Netw.*, (Athens, Greece), Oct. 2024, pp. 1-6.