# AI의 이동통신시스템 적용

Artificial Intelligence Applications on Mobile Telecommunication Systems

예충일 (C.I. Yeh, ciyeh@etri.re.kr)

장갑석 (K.S. Chang, kschang@etri.re.kr)

고영조 (Y.J. Ko, koyj@etri.re.kr)

6G무선방식연구실 책임연구원 6G무선방식연구실 책임연구원 6G무선방식연구실 책임연구원/실장

#### **ABSTRACT**

So far, artificial intelligence (Al)/machine learning (ML) has produced impressive results in speech recognition, computer vision, and natural language processing. Al/ML has recently begun to show promise as a viable means for improving the performance of 5G mobile telecommunication systems. This paper investigates standardization activities in 3GPP and O-RAN Alliance regarding Al/ML applications on mobile telecommunication system. Future trends in Al/ML technologies are also summarized. As an overarching technology in 6G, there appears to be no doubt that Al/ML could contribute to every part of mobile systems, including core, RAN, and air-interface, in terms of performance enhancement, automation, cost reduction, and energy consumption reduction.

KEYWORDS AI/ML. E2E learning, federated learning, NWDAF, RIC

# 1.서론

AI/ML은 입력 데이터셋(Dataset)에서 기존의 기술로 구분하지 못했던 특정 패턴(Pattern)이나 피처 (Feature) 등을 정확하게 추출하여 그동안 해결하지 못했던 다음의 문제들에 대하여 신뢰할 만한 해결 방법을 제공하기 시작하였다.

- 모델링이 난해한 문제
- 비선형 특성으로 모델링이 불가능한 문제
- 실현 불가능한 구현복잡도가 요구되는 최적화 무제

기존의 이동통신시스템에서는 모델주도(Model-driven) 방식의 사건해결형(Reactive) 모델링이 사용되었으나, 5G에 이르러서는 점차로 성능, 비용 등의 측면에서 장점을 보이는 데이터주도(Data-driven) 방식의 사전예방형(Proactive) artificial intelligence (AI)/machine learning (ML) 모델링이 도입되기 시작했고, 적용분야가 더욱 확대되고 있는 추세이다. 2030년경에 서비스가 시작될 것으로 예측되는 6G에서는 AI/ML이 이동통신시스템 모든 분야에서 자율화, 비용절감, 성능향상에 기여하는 총괄적 핵

<sup>\*</sup> 이 논문은 2022년도 정부(과학기술정보통신부)의 재원으로 정보통신기획평가원의 지원을 받아 수행된 연구임[No. 2018-0-00218, 초고주파 이동통신 무선백홀 전문연구실].



<sup>\*</sup> DOI: https://doi.org/10.22648/ETRI.2022.J.370407

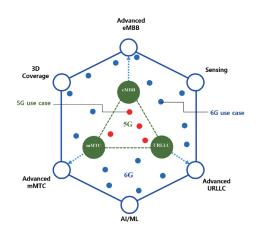


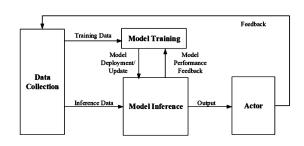
그림 1 6G에서 예상되는 AI/ML의 역할

심 기반 기술(Overarching Enabler)로서뿐만 아니라 네 트워크가 제공하는 하나의 사용시나리오(Usage Scenario)로서도 자리매김할 것으로 예상되고 있다. 그 림 1은 6G에서 AI/ML의 예상 역할을 나타낸다.

본고에서는 B5G/6G 등의 미래이동통신시스템에서 핵심적 역할이 예상되는 AI/ML의 모델링 방법, 3GPP 및 O-RAN에서의 AI/ML 관련 표준화 현황, AI/ML의 미래 기술 동향 및 향후 전개 방향에대하여 조사·분석하였다.

## II. AI/ML 모델링

이동통신시스템에 지능을 부여하기 위해 3GPP에서 제안된 AI/ML의 기능적 구조는 그림 2와 같다[1]. 데이터수집(Data Collection) 블록은 UE 등의 종단장치 또는 망에서 수집한 데이터를 모델훈련 (Model Training) 및 모델추론(Model Inference) 블록으로 제공한다. 모델훈련 블록은 데이터수집 블록이 제공하는 데이터를 이용하여 훈련, 평가, 테스팅 (Training, Validation, Testing)의 과정을 거쳐 목적 AI/ML 모델을 훈련된 AI/ML 모델로 학습시킨다. 또한 데이터수집 블록에서 입력되는 데이터에 대하



출처 Reprinted with permission from [1].

그림 2 이동통신시스템 적용 AI/ML 모델

여 전처리/클렌징, 포맷팅, 트랜스포메이션(Transformation) 등의 작업을 수행하여 AI/ML 알고리즘에 적합한 형태의 데이터로 전환한다. 모델추론 블록은 예측(Prediction)이나 결정(Decision)과 같은 추론적 결과를 산출한다. 모델추론 블록에서 모델훈련 블록으로 제공되는 모델성능피드백(Model Performance Feedback)은 AI/ML 모델의 성능 추적에 활용될 수 있다. 동작(Actor) 블록은 모델추론 블록으로부터 산출되는 결과를 받아 상응하는 조치를 취한다. 동작 블록에서 출력되는 피드백 정보에서 모델훈련, 모델추론, 성능 추적에 필요한 정보를 얻을 수 있다. 모델추론 블록에서 출력되는 모델성능피드백은 개별 use case에 적용된 AI/ML의 성능 평가 척도로 향후 AI/ML 모델을 더욱 정교하게 개선하기 위한 재훈련 (Retraining) 결정시 활용될 수 있다.

# Ⅲ. AI/ML 표준화

5GS에서 활용 가능한 AI/ML 플랫폼으로 코어의 NWDAF와 RAN의 RIC가 있다. NWDAF는 3GPP에서 RIC는 O-RAN에서 표준화 작업이 진행되고 있다. RIC는 다시 non-RT RIC와 near-RT RIC로 구분된다. MNO 또는 개발자들은 이들 플랫폼을 활용하여 고유의 AI/ML 모델을 개발할 수 있다. 특

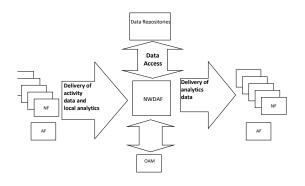
정 AI/ML 모델과 AI/ML 플랫폼 사이의 배타적 대응 관계는 없다. 예를 들어 에너지 절감을 위한 AI/ML 모델은 NWDAF, non-RT RIC, near-RT RIC 어느 AI/ML 플랫폼에서도 개발/운용 가능하다. 다만 처리 및 반응 속도는 near-RT RIC, non-RT RIC, NWDAF 순으로 빨라진다. 따라서 에너지 저감 목적으로 전력 스위치를 내릴 때, 셀을 끄는 경우는 NWDAF, 캐리어를 끄는 경우는 non-RT RIC, 프레임을 끄는 경우는 near-RT RIC 플랫폼에서 AI/ML 모델이 구현될 수 있다.

### 1. 코어 적용 AI/ML

코어는 이동통신시스템에서 네트워크(Internet, PLMN)에 가장 가깝게 위치하는 부분이다. 5GC에서 AI/ML과 관련된 기능은 NWDAF이며, 3GPP에서 관련 표준화 작업이 진행되고 있다.

NWDAF는 데이터 수집 인터페이스(Data Collection Interface)를 통해 클라우드, 엣지네트워크(Edge Network) 또는 5GC로부터 데이터를 수집하고, AI/ML 모델 등을 이용하여 수집된 데이터를 분석/가공한 다음, 데이터 제공 인터페이스(Data Exposure Interface)를 통해 소비자가 필요로 하는 데이터 애널리틱스 (Analytics)를 제공한다.

데이터가 NWDAF로 집중된다는 것은 5GS에서 AI/ML이 정착될 가장 적합한 장소가 NWDAF 임을 의미한다. 3GPP SA 규격 작성 그룹에서는 R16(Release 16)에서 NWDAF 관련 use case, key issue, solution 등을 정리하였고, 그림 3의 NWDAF 기본 골격도 확립하였다[2]. 그림 3에 따르면 NWDAF는 OAM, AF, NF로부터 데이터를 수집한다. 또한 NWDAF는 PLMN에 있는 UDR 등과 같은 저장 매체를 통해 네트워크에 저장된 데이터에도 접근할수 있다. NWDAF는 수집된 데이터를 기반으로 애널리



출처 Reprinted with permission from [2].

그림 3 5G NWDAF 기본 구조

틱스를 수행하고 그 결과를 NF 소비자(Consumer)에 게 제공한다. PCF, NSSF, AMF, SMF, NEF, UDM, AF, OAM, CEF, DCCF, NWDAF 등이 NF 소비자 가 될 수 있고, NF 소비자는 NWDAF 서비스에 가 입하거나 요청에 의하여 애널리틱스 결과를 통보받을 수 있다.

3GPP SA 규격 작성 그룹에서는 R16/17에서 데이터 애널리틱스 서비스 제공을 위하여 5GS 기준 구조(Reference Architecture)를 개선하였고, NWDAF 와 관련된 통신 절차 등을 정의하였다. 특히 R17에서는 ML 모델 훈련 및 훈련된 모델을 NF 소비자에게 제공하기 위해 NWDAF 내에 MTLF와 관련된입출입인터페이스를 정의하였다[3]. 현재의 규격은 AI/ML을 위한 데이터 수집 인터페이스, 절차 등을 규정하고 있으며, AI/ML의 세부적인 알고리즘이나 구체적 적용에 대해서는 규정하지 않고 있다.

MNO는 규정된 인터페이스와 절차를 통하여 수집된 데이터에 기반하여 AI/ML 모델을 개발하고이를 통해 고유한 AI/ML 적용 성공 사례들을 만들어나갈 수 있다. NWDAF라는 AI/ML 플랫폼을 통해 MNO는 스스로 학습하여 환경 변화나 사용자 요구에 자율적으로 대응하는 인간의 개입이 필요 없는(Zero Human-touch) 5GN을 추구해 나갈 수 있다.

예를 들어, NWDAF는 자체에 구현된 다양한 AI/ML 모델을 이용하여 망에 걸린 부하 정도, UE 이동 방향, 트래픽 패턴, 특정 지역 혼잡 상황 등을 예측하고 자율적 네트워크 슬라이스 구성을 포함한 5G네트워크 오토메이션을 실현할 수 있다.

### 2. RAN 적용 AI/ML

RAN이 개방화, 가상화, 지능화를 추구하는 이유 는 CAPEX, OPEX 등의 비용 절감 및 사용자 요구에 신속하게 대응하기 위함이다. Open RAN은 RAN 을 기능적으로 분할하고, 분할된 기능체 간의 인터 페이스를 표준화하여 MNO가 다양한 공급자들로 부터 RAN을 구성하는 기능들을 상용제품(COTS) 형태로 구입하여 자신의 구미에 맞는 RAN 구성을 가능하게 한다. 또한 RAN 구성 기능체들을 소프트 웨어로 구현하여 RAN을 가상화(Virtualized RAN)함 으로써 유연성(Flexibility, Programmability)을 높여 필 요에 따라 신속한 재구성(Reconfiguration)이 가능하 도록 하고 있다. 최근에는 RAN에 지능을 부여하 여 환경의 변화 및 사용자의 요구에 인간의 관여 없 이 자율적으로 대응하는 RAN(Automated RAN)을 실 현하기 위한 노력이 경주되고 있다. O-RAN에서는 RIC 개념을 도입하여 지능을 부여할 수 있는 Open RAN 규격을 개발하고 있다. 그림 4는 3GPP RAN 과 O-RAN의 구조를 비교한 그림이다. O-RAN은 RAN을 SMO, 가상화된 다수의 RAN 기능체셋(RAN Functions) 및 O-Cloud의 3개 부분으로 나누었다.

SMO는 MNO 또는 텔코클라우드(Telco Cloud)에서 제공되는 서비스 형태로 구현될 수 있으며, RAN을 총괄 관리하는 콘트롤러(Controller) 및 RAN 오토 메이션을 위하여 데이터를 수집/가공하는 플랫폼역할을 수행한다.

O-RAN에서 정의한 RAN 기능체로는 near-RT

RIC와 가상화된 O-CU, O-DU 등이 있다. O-RU 의 가상화는 아직 시기상조이다. O-CU는 O-CU-CP와 O-CU-DP로 구성되며, O-CU-CP는 PDCP의 CP및 RRC 프로토콜을 지원하고, O-CU-DP는 PDCP의 UP및 SDAP 프로토콜을 지원한다. O-DU는 RLC, MAC, High-PHY에 해당하는 프로토콜을 지원한다. O-RU는 송수신과 관련된 하위레벨 PHY 기능을 지원한다. 이러한 O-RAN의 RAN기능체셋은 O-Cloud라고 불리는 컴퓨팅 플랫폼을통하여 제어된다. O-Cloud는 RAN인프라 구축및관리서비스를 제공한다. O-RAN에서 정의한 RAN의주요인터페이스로 다음과 같은 것들이 있다.

- R1: Non-RT RIC와 rApps 사이의 인터페이스
- O1: SMO가 O-RAN에서 정의한 RAN 기능 체셋을 관리할 목적으로 정의된 인터페이스, RAN의 기능으로 FCAPS/소프트웨어/파일 관 리 등이 있다.
- O2: SMO와 O-Cloud 사이의 인터페이스
- A1: Non-RT RIC와 near-RT RIC 사이의 인터 페이스, non-RT RIC는 정책지침을 near-RT

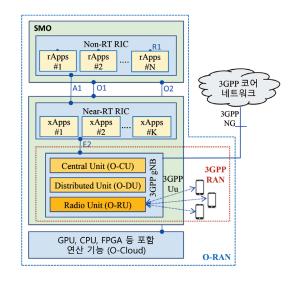


그림 4 3GPP RAN과 O-RAN의 비교

RIC로 보내고 near-RT RIC로부터 관련 정책 피드백을 받을 수 있다.

• E2: near-RT RIC와 가상화된 RAN 기능체셋 사이의 인터페이스

Non-RT RIC 및 near-RT RIC는 RAN 지능화를 담당하며 non-RT RIC는 1s 이상의 시간 프레임, near-RT RIC는 10ms-1s 정도의 시간 프레임으로 동작한다. Non-RT RIC 및 near-RT RIC는 AI/ML을 활용하여 RAN 인프라/자원 관리를 수행한다. 3rd party 개발자는 MNO가 제공하는 개방형 API 환경을 활용하여 non-RT RIC 및 near-RT RIC에 접근하여 플랫폼이 제공하는 다음과 같은 각종 기능 및 능력을 활용하여 다양한 AI/ML 모델을 개발할수 있다.

- 데이터 관리(Data Management)
- 자율화(Automation)
- 애널리틱스(Analytics Insights)
- ML 수행 및 훈련
- 서비스/자원 관리
- 보안(Security)
- 애플리케이션 라이프사이클 관리
- 콘트롤러 프레임워크(Controller Framework)

rApp/xApp은 non-RT RIC 및 near-RT RIC 상에서 동작하는 3rd party가 개발한 애플리케이션 프로그램으로 다수의 마이크로 서비스로 구성된다. rApps의 예제로 에너지 저감이나 셀 혼잡 사전 예방 등의 애플리케이션을 고려해 볼 수 있다. 특정 셀의 트래픽 양을 예측하는 AI 알고리즘으로 LSTM이 사용될 수 있다. SMO는 O1/O2 인터페이스를통해 수집된 RAN 데이터를 non-RT RIC에 제공한다. rApps의 요청에 따라 SMO와 결부되어 있는 AI 서버는 최근에 수집된 RAN 데이터를 활용하여

LSTM을 훈련시켜 모델 파라미터를 갱신한 다음 이를 non-RT RIC에게 전달하면 이를 기반으로 non-RT RIC는 에너지 절감, 셀 혼잡 사전 예방 등의 조치를 취할 수 있다. 예를 들어, 예측되는 트래픽의 양이 적은 경우, CoMP를 작동시켜 현재 셀의 로드를 인접 셀로 오프로딩시킨 다음 현재 셀의 스위치를 내려 에너지 저감을 실현할 수 있다.

3rd party 개발자는 사용자 이동 정보, 트래픽 행태, 셀 로드 변화 등의 RAN에서 수집된 데이터에 기반한 애널리틱스 또는 AI/ML 모델을 통해 다음과 같은 rApps/xApps 서비스를 창출할 수 있다.

### rApps 적용 분야

- 서비스 및 정책 관리
- 상위계층 절차 최적화
- near-RT AI/ML 모델 훈련
- 에너지 저감
- 네트워크 슬라이싱

#### xApps 적용 분야

- 전파자원관리
- 핸드오버제어
- QoS 관리
- 연결성관리(Connectivity Management)
- 트래픽 조정(Traffic Steering)

### 3. 에너지 저감을 위한 AI/ML

5G NR 망에서 최대 전력 소비처는 기지국이고, 기지국 사용 전력의 90%는 AAU에서 소비되는 것으로 알려져 있다[4]. 기지국 에너지 저감은 사용 비율이 낮은 슬롯, 프레임, 채널, 캐리어 또는 기지국 자체의 스위치를 내리는 기술을 의미한다. 단말측에너지 저감은 주로 슬립 모드 기술에 의존하고 있으며, 주요 목표는 배터리 재충전 기간을 늘려 사용

자의 불편을 해소하는 데 있다. 조건에 따라 셀의 스 위치를 내리고 다시 셀을 활성화시키는 에너지 저감 기술은 5G 최초 시스템인 R15에 이미 도입되었다. R17에 이르자 에너지 저감에 AI/ML을 활용하기 시작했다. AI/ML를 에너지 저감에 활용하려면데이터가 필요하다. 3GPP RAN은 R17에서 다음과같이 데이터 수집과 관련된 WI와 SI를 수행하였다.

- WI: SON/MDT를 위한 데이터 수집 개선[5]
- SI: 데이터 수집 추가 개선에 대한 연구[6]

SON/MDT WI에서 취급된 주요 내용은 UE 에서 측정한 데이터 수집과 이를 활용한 SON 등에 대한 것이다. 따라서 SDN/MDT WI에서 중점적으로 논의된 사항은 RAN self-optimization 및 self-configuration 등과 관련된 사항이었고, 아울러RAN 데이터 활용 에너지 저감 기술도 SDN/MDT의 한 use case로써 언급되었다. 그러나 아직 SON/MDT WI에서 AI/ML 적용은 논의되지 않았다. 이후 데이터 수집 추가 개선 관련 SI는 AI 기반의에너지 저감에 대한 연구 결과를 보고하였다[7]. R18에서는 "NR 네트워크 에너지 저감에 관한 연구(Study on network energy savings for NR)"라는 제목의SID가 승인되었으나 AI/ML이 언급되지는 않고 있다[8].

에너지 저감을 위해 특정 셀의 트래픽 로드 크기가 규정된 값 이하일 경우, 셀 스위치를 내리고 새로운 인접 타켓 셀을 UE에게 지정할(UE offloading) 수있다. 지금까지는 현재의 트래픽 로드 크기에 기반하여 에너지 저감을 위한 결정이 수행되었으나, AI/ML 알고리즘을 사용할 경우 미래의 에너지 효율이나 로드 상태를 예측할 수 있으므로 이를 활용하여 매우 동적인 방법으로 균형 잡힌 에너지 저감 전략을 구사할 수 있다. 이는 세밀하게 제어되는 스위치온/오프, UE 오프로딩, 셀 커버리지 및 RAN 구성

제어 등의 다양한 기술을 활용하여 시스템 성능과 에너지 효율이라는 상반된 문제를 균형 있게 조절 함을 의미한다.

#### 4. 물리계층 적용 AI/ML

3GPP에서 AI/ML을 물리계층에 적용하기 위한 표준 활동은 R18부터 시작되었다. 3GPP는 위치측 정 정밀도 개선, CSI 예측 및 정밀도 개선, 빔 관리· 예측 성능 개선의 3개 분야를 우선적으로 선정하였 고, 적용 가능성 여부를 결정하기 위한 논의 시작을 합의하였다[9].

#### 가. 위치측정 정밀도 개선(Positioning)

변화무상하고 임의성(Randomness)의 정도가 매우 심한 채널 특성은 단말의 위치 측정을 위한 정교한 수학적 모델링을 매우 어렵게 한다. PRS/SRS 등을 이용하여 측정된 값들은 코어의 LMF에게 보고되고 LMF는 LOS 가정하에서 단말의 위치를 결정한다. 위치 측정을 위한 기존의 알고리즘은 LOS/NLOS 구분을 할 수 없으므로 빈번하게 발생되는 NLOS 환경은 정교한 측정을 방해하는 요소 중의 하나이 다. AI/ML은 단말/기지국이 제공하는 채널매트릭 스(채널 측정값)에서 LOS/NLOS를 구분하여 정밀하 게 단말 위치를 판별해내는 데 기여할 것으로 기대 된다(그림 5 참고).

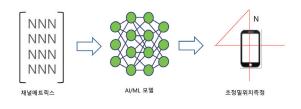


그림 5 AI/ML 적용 위치 측정

#### 나. CSI 예측 및 정밀도 개선

CSI 피드백이란 특정 목적을 달성하기 위해 UE가 채널 측정 정보를 기지국(gNB)에게 보고하는 것이 다. CSI 피드백에 AI/ML을 활용하는 목적은 오버헤 드 감소 및 CSI 정밀도 개선에 있다. 여기서 UE 측의 AI/ML은 측정된 채널 정보를 압축하여 전송 데이터 의 양을 줄이는 목적으로 활용되고 기지국 측의 AI/ ML은 압축된 채널 정보에서 원래의 채널 정보를 복 원하는 목적으로 활용된다. AI/ML 알고리즘을 활 용하여 UE는 CIS를 압축하고, gNB는 압축된 CIS에 서 원래의 CIS를 복원할 수 있다(그림 6 참고). 이 분 야에서 AI/ML 적용은 CSI 전송에 소모되는 주파 수/공간 자원을 절약하여 오버헤드를 줄이거나 정 밀한 CSI를 적은 오버헤드로 전송 가능하게 하여 고 주파(mmWave) 영역에서 massive MIMO의 성능을 개 선시킬 수 있을 것으로 기대된다. 충분히 다양한 지 리적 위치에서 수집된 데이터를 활용하여 오프라 인으로 학습된 CSI-AI/ML 모델이 UE에서 측정된 CSI와 기지국에서 복원된 CSI가 상당한 정도로 유 사함을 보여준다는 연구 결과가 제시되고 있다[10].

#### 다. 빔 관리 및 예측

다운링크(Downlink)에서 UE는 최적의 "gNB-송신 빔/UE-수신빔"을 지속적으로 추적할 필요가 있다. 이와 같은 "송수신 빔짝(Tx-Rx Beam Pair)" 추적에는 상당한 양의 CSI-RS 전송이 요구되므로 오버헤드

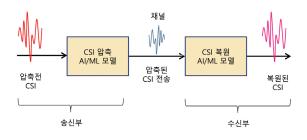


그림 6 AI/ML 적용 CSI 피드백 개선

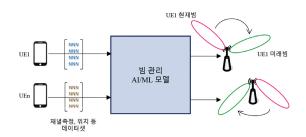


그림 7 AI/ML 적용 빔관리

의 증가가 필연적이다. 현재 사용 중인 빔과 좌우에 있는 빔만을 고려하는 간단한 경우에도 송수신 빔짝을 정렬하려면 9개의 CSI-RS 자원이 소모된다. 500km/h로 달리는 초고속열차는 특정 빔에 머무르는 시간(Dwelling Time)이 7ms 정도로 매우 짧은 경우도 발생하므로, CSI-RS를 매우 촘촘하게 할당해야할 필요성도 대두된다[11].

AI/ML은 관련된 UE들의 과거 이동 경로에서 특정 UE의 움직이는 방향을 추론할 수 있으므로 빔관리(Beam Management)에 활용될 수 있다. 이와 같은 AI/ML 적용 빔관리 기술은 CSI 피드백을 필요로 하지 않으므로 기존 빔관리 기술보다 통신 절차가 단순해진다. 따라서 레이턴시(Latency) 및 오버헤드 감소의 효과가 기대되고, 고속으로 이동하는 UE에게도 폭이 좁은 빔을 할당하여 커버리지 확장 및 성능개선에 활용될 수 있다. UE는 측정에 따라 최고의빔을 gNB에게 보고하고 gNB는 UE에게 다음에 어떤 빔을 사용할 것인지를 알려준다. 그림 7은 gNB의빔 관리 AL/ML 모델이 UE에게 가장 적합한 미래빔을 예측하고 할당하는 모습을 보여준다.

# Ⅳ. AI/ML 미래 기술 동향

# 1. 연합학습(Federated Learning)

중앙집중형(Centralized) AI/ML 모델에서, 네트워

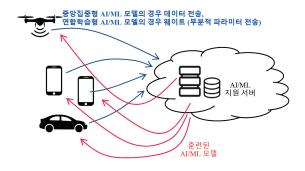


그림 8 중앙집중형/연합학습형 AI/ML 모델

크 내의 적절한 장소에 위치한 AI/ML 지원 서버는 단말 등의 종단장치들이 전송하는 로컬 데이터셋에 기반하여 AI/ML 모델을 훈련시켜 사용자에게 웨이트(Weight) 등을 포함한 AI/ML 모델 파라미터를 제공한다. 이와 같은 중앙집중형 방식의 단점으로 종단장치들의 데이터셋 전송에 수반되는 통신량의 양적 증가에 따른 자원 소모 부담과 이에 따른 개인프라이버시 노출의 문제점 등이 있다. 이런 문제점을 해결하기 연합학습형, FL 모델이 연구되고 있다.

FL에서는 각 종단장치가 로컬 데이터셋에 기반하여 모델 파라미터를 계산한 다음 이를 전송하면 AI/ML 지원 서버는 수신된 모델 파라미터들을 집성하여 원하는 결과를 얻는다. FL을 사용할 경우, 전체데이터셋을 한꺼번에 이용할 수 없으므로 학습 성능(정확도) 측면에서 약간의 성능 저하가 있지만, 통신 오버헤드를 현저하게 줄이는 효과를 거둘 수 있다. AI/ML의 물리계층 적용의 경우, 수집된 로컬데이터셋을 바탕으로 단말이 부분적으로 모델 파라미터를 계산하고 이를 기지국을 통해 AI/ML 지원서버로 전송하는 절차가 보편적인 형태일 것이므로, 이를 지원하기 위한 네트워크 구조 연구가 필요하다. 경우에 따라 중앙집중형과 연합학습형을 섞어 사용하는 혼재 방식이 유용할 수도 있다. 그림 8은 두가지 방식을 비교한 그림이다.

### 2. 종단 간 학습(E2E Learning)

최근까지 통신시스템에서 송수신기는 기능별 블록을 순차적으로 연결하는 방식으로 구현되었다. 그러나 종전의 블록 단위의 해석 및 최적화 방식은 전체 시스템 최적화 관점에서 볼 때 최선이 아닌 차선의 선택이었다. 최근 AI/ML 기술은 P2P 전체 통신 구간의 E2E 관점에서 채널은 물론 하드웨어 불완전성(Impairment)까지 고려하여 송신기와 수신기를 동시 학습시킬 수 있다는 가능성을 보여주고 있다. PHY-MAC 등 여러 프로토콜 계층을 망라하는 cross-layer E2E 학습은 향후 시스템 설계는 물론 표준화에도 상당한 영향을 미칠 것으로 예측된다.

종단 간 학습의 한 예로 송수신의 전체 물리계층체인을 AI/ML의 한 형태인 오토인코더(Autoencoder)로 구현할 수 있다. 오토인코더는 인코더와 디코더로 구성되며, 송신단의 인코더는 전송하고자 하는데이터에서 핵심적인 특징을 추출한 다음 이에 상응하는 코드(Code)로 전환하여 채널을 통해 전송하고 수신단의 디코더는 수신데이터에서 원래데이터를 복원한다. P2P 물리계층 통신시스템을 NN을적용한 인코더와 디코더로 모델링하고 E2E 관점에서 학습시킨 오토인코더를 통해 기존 방식으로 구현된 통신 시스템에 필적하는 BLER 성능을 얻은 연구결과가 있다[12].

#### 3. AlaaS

지금까지 살펴본 바에 따르면 AI/ML은 코어, RAN에서는 물론 무선접속 분야의 물리계층에까지 적용되어 자율화, 비용 및 에너지 저감, 성능 향상에 기여하는 총괄 인에블러(Enabler) 역할을 수행하고 있다(AI for Network). 이 밖에도 AI/ML은 그 자체가 서비스로 제공될 수 있다(AIaaS). 개인 개발자나

회사는 큰 비용 투자나 위험 부담 없이 AIaaS 플랫폼 운영자가 제공하는 API를 활용하여 클라우드에 설 치되어 있는 AI/ML 관련 알고리즘, 소프트웨어, 하 드웨어, 컴퓨팅 능력, 데이터 등의 자원을 활용하여 자신의 AI/ML 모델 등을 개발/개선할 수 있다. 또 한 AIaaS를 위해서는 광대역폭, 고신뢰, 저지연 등 의 측면에서 네트워크의 능력도 더욱 강화될 필요 성이 있다(Network for AI).

### V. 결론

본 논문을 통해 AI/ML을 이동통신시스템에 적용하기 위해 3GPP/O-RAN에서 수행하고 있는 표준화 진행 상황을 조사/분석하였고, AI/ML의 미래 기술 동향에 대해서도 살펴보았다. 5GC의 NWDAF 및 RAN의 RIC는 5GS/네트워크에서 수집되거나 저장된 데이터에 AI/ML 알고리즘을 적용하여 사용자가 필요로 하는 다양한 데이터 애널리틱스 또는데이터주도형 애플리케이션을 제공할 수 있다. 무선접속 분야에서도 위치추정, CSI 피드백, 위치추정, 빔 관리/예측 분야에서 AI/ML이 도입되어 성능개선에 기역할 것으로 기대된다.

컴퓨팅 능력, 보유 AI/ML 알고리즘과 더불어 데이터 수집, 저장, 교환을 위한 인터페이스 및 프로토콜 정립은 AI/ML 적용 활성화를 위해 네트워크가 갖추어야 할 가장 필수적인 요소이다. 현재 연구되고 있는 연합학습, 종단 간 학습 등의 AI/ML 기술은 이동통신시스템의 설계 방법론에도 상당한 영향을 미칠 것으로 예측된다.

AI/ML을 활용하는 데이터주도형 애플리케이션 의 점차적 확산은 인간의 개입 없이 환경 및 사용자 의 요구에 최적으로 자율 대응해 나가는 이동통신 시스템을 가능하게 할 것으로 기대된다.

#### 약어 정리

3D

3GPP	Third Generation Partnership Project
5GC	5G Core
5GN	5G Network
5GS	5G System
AAU	Active Antenna Unit
AF	Application Function
AI	Artificial Intelligence
AIaaS	AI as a Service
AMF	Access and Mobility Management
	Function
AnLF	Analytics Logical Function
API	Application Programming Interface
B5G	Beyond 5G
BLER	Block Error Rate
CAPEX	Capital Expenditures
CEF	Charging Enablement Function
CoMP	Coordinated MultiPoint
	transmission/reception
COTS	Commercial Off-The-Shelf
CP	Control Plane
CSI	Channel State Information
DCCF	Data Collection Coordination
	Function
DP	Data Plane
E2E	End to End
eMBB	enhanced Mobile Broadband
FCAPS	Fault, Configuration, Accounting,
	Performance and Security
FL	Federated Learning
LMF	Location Management Function
LOS	Line Of Sight
LSTM	Long Short-Term Memory
MAC	Medium Access Control
MDT	Minimization Drive Test
ML	Machine Learning
mMTC	massive Machine Type
	Communication
MNO	Mobile Network Operator

Three Dimensional

MTLF	Model Training Logical Function
NEF	Network Exposure Function
Near-RT	Near-Real Time
NF	Network Function
NLOS	Non Line Of Sight
NN	Neural Network
Non-RT	Non-Real Time
NR	New Radio
NSSF	Network Slice Selection Function
NWDAF	Network Data Analytics Function
OAM	Operation, Administration, and
	Maintenance
O-CU	Open Central Unit
O-DU	Open Distributed Unit
OPEX	OPerating EXpenses
O-RAN	O-RAN Alliance
O-RU	Open Radio Unit
P2P	Point to Point
PCF	Policy Control Function
PDCP	Packet Data Convergence Protocol
PHY	Physical
PLMN	Public Land Mobile Network
PRS	Positioning Reference Signal
RAN	Radio Access Network
RIC	RAN Intelligence Controller
RLC	Radio Link Control
RRC	Radio Resource Control
RRH	Remote Radio Head
SA	Service and System Aspects
SDAP	Service Data Adaptation Protocol
SI	Study Item
SID	Study Item Description
SMF	Session Management Function
SMO	Service Management and
	Orchestration
SON	Self-Organizing Network
SRS	Sounding Reference Signal

ToA Time of Arrival
UDM Unified Data Management
UDR Unified Data Repository
UE User Equipment
URLLC Ultra Reliable and Low Latency
Communication
WI Work Item

#### 참고문헌

- [1] 3GPP TR 37.817 V1.2.0, E-UTRA and NR; Study on Enhancement for Data Collection for NR and EN-DC(Release17), 2022. 1.
- [2] 3GPP TR 23.791 V16.2.0, Study of Enablers for Network Automation for 5G, 2019. 6.
- [3] 3GPP TS 23.288 V17.3.0, Architecture Enhancements for 5G System(5GS) to Support Network Data Analytics Services(Release 17), 2021. 12.
- [4] 3GPP RWS-210153, NR Network Energy Saving Enhancement for Rel-18, TSG RAN Release 18 Workshop, June 28-July 2, 2021.
- [5] 3GPP RP-193255, New WID on Enhancement of Data Collection for SON/MDT in NR, TSG RAN#86, Sitges, Barcelona, Dec. 9-12, 2019.
- [6] 3GPP RP-201304, New SID: Study on Further Enhancement for Data Collection, TSG RAN#88e, June 29-July 3, 2020.
- [7] 3GPP TR 37.817 V1.2.0, Study on Enhancement for Data Collection for NR and EN-DC(Release 17), 2022. 1.
- [8] 3GPP RP-213554, New SI: Study on Network Energy Savings for NR, TSG RAN#94e, Dec. 6-17, 2021.
- [9] 3GPP RP-213599, New SI: Study on Artificial Intelligence(AI)/Machine Learning(ML) for NR Air Interface, TSG RAN#94e, Dec. 6-17, 2021.
- [10] 3GPP RP-213387, Support of Artificial Intelligence Applications for 5G Advanced, TSG RAN#94e, Dec. 6-17, 2021.
- [11] 3GPP RWS-210481, Enhancements on Predictable Mobility for Beam Management, TSG RAN WG Meeting REL-18 Workshop, June 28-July 2, 2021.
- [12] S. Dörner et al., "Deep learning based communication over the air," IEEE J. Sel. Top. Signal Process., vol. 12, no. 1, Feb. 2018, pp. 132–143.