

연합 심층 강화학습 기반 네트워크 슬라이싱의 연구 동향에 대한 조사

오준석, 이동현, 이충현, 송치현, 홍성훈, 김가현, Thanh Phung Truong, 김주영, 조성래

중앙대학교 컴퓨터공학과

{jsoh, dhlee, chlee, chsong, shhong, ghkim, tptruong, jykim}@uclab.re.kr, srcho@cau.ac.kr

A Survey on Research Trends of Federated Deep Reinforcement Learning-Based Network Slicing

Junsuk Oh, Donghyun Lee, Chunghyun Lee, Chihyun Song, Seonghun Hong,

Gahyun Kim, Thanh Phung Truong, Juyoung Kim, and Sungrae Cho

School of Computer Science and Engineering, Chung-Ang University

요약

본 조사는 차세대 이동통신 시대의 네트워크 슬라이싱을 연합 심층 강화학습(Federated Deep Reinforcement Learning, FDRL) 관점에서 간명히 정리한다. 개방형 무선접속망(Open Radio Access Network, O-RAN)의 RIC (Radio Intelligent Controller) 확산을 배경으로, 최신 연구 동향을 MDP (Markov Decision Process) 및 연합 체계 틀에서 재구성하여 서술한다. 연합 주기, 참여자 선택, near-RT (Real-Time) 및 non-RT RIC 분업이 확장성, 프라이버시, 수렴성에 미치는 효과를 통합적으로 해석하고, 상용 확장을 위한 안전성, 공정성, 설명 가능성 관련 과제를 제시한다.

I. 서론

네트워크 슬라이싱은 서비스마다 상이한 지연, 신뢰도, 처리량, 에너지 등의 요구를 SLA (Service Level Agreement) 단위로 표현하고, 이를 무선접속망(Radio Access Network, RAN)-전송-코어 전 구간에서 걸쳐 독립적으로 보장하려는 시도에서 출발한다. 초기 중앙집중형 규칙 기반 제어 및 휴리스틱 기반 연구는 트래픽과 무선 환경의 불확실성이 증가하고 xAPP이 다수 공존하는 개방형 무선접속망(Open RAN, O-RAN)이 확산되면서, 비정상적이고 부분 관측적인 환경에서도 시행착오를 통해 정책을 고도화할 수 있는 심층 강화학습(Deep Reinforcement Learning, DRL) [1-2] 기반 연구로 중심을 옮겼다. 그러나 이는 단일 셀의 데이터에만 의존할 때 일반화와 견고성의 한계를 보임에 따라, 각 위치의 에이전트가 로컬 데이터를 외부로 유출하지 않고도 전역 모델을 개선할 수 있도록 하는 연합 학습(Federated Learning, FL)과 결합하여 이후 연합 심층 강화학습(Federated Deep Reinforcement Learning, FDRL)으로 발전했다. 이는 near-RT (Real-Time) 및 non-RT RIC (Radio Intelligent Controller), 단말과 기지국, 엣지와 클라우드에 분산된 에이전트들을 하나의 학습 생태계로 묶어, 복수의 슬라이스가 공유 인프라에서 충돌 없이 진화하도록 만드는 토대를 제공한다 [3]. 또한, 공통적으로 상태-행동-보상으로 정식화된 MDP (Markov Decision Process) 위에 DRL 정책을 학습하고, FL 집계를 통해 전역 모델을 갱신한다. 본 연구는 이 틀 위에서 학습 모델 위치, 집계 주기, 참여자 선택, 그리고 업데이트 압축이 성능과 오버헤드에 어떤 식으로 반영되는지 최신 연구 동향을 조사한다.

II. 본론

1. 연구 동향

접속제어와 단말 연관 연구는 FL의 이점을 가장 직접적으로 드러낸다. 단말은 로컬에서 관측한 무선 환경과 부하 정보를 바탕으로 접속 및 핸드오버 정책을 RL로 학습하고, 기지국 혹은 RIC는 수평 연합을 통해 분산된

파라미터를 집계하여 전역 정책을 정제한다. 동시에 민감한 원시 데이터를 외부로 내보내지 않기 위해, 암호화된 연합 파티에서 수직 연합을 수행해 특정 수준의 정보를 통합함으로써 다양한 셀과 슬라이스의 패턴을 포착한다. 이러한 하이브리드 FL 구조는 처리량 개선과 통신 오버헤드 절감을 함께 달성하면서, 다수 단말의 이질적 경험을 효율적으로 한데 모으는 방식으로 작동할 수 있음을 보인다 [4-5]. 산업용 사물인터넷 시나리오에서는 DFQL (Deep Federated Q-Learning)을 활용한다. 간헐적 트래픽과 서로 다른 KPI (Key Performance Indicator) 요구를 전제로 하여 전력 및 SF (Spreading Factor) 등 무선 제어 변수를 슬라이스별 보상에 정합되도록 조정하고, 각 단말 및 셀의 네트워크 모델은 로컬에서 학습되어 파라미터만 연합 집계된다. 이 구조는 중앙집중식과 유사한 수렴 특성을 유지하면서도 연합 통신량을 줄이고, 부하 변동이 큰 환경에서 SLA 위반율을 억제하는 경향을 보인다 [6].

자원할당과 xAPP 조정에서는 다중 에이전트의 협업이 핵심이다. 전력 제어 xAPP과 슬라이스 자원 xAPP이 독립적으로 최적화될 경우 정책 충돌로 이어지기 쉬운데, FDRL은 두 에이전트를 연합 체계 안에서 공동 훈련하여 조화를 유도한다. [7]은 eMBB (enhanced Mobile Broadband) 처리량 상승과 URLLC (Ultra-Reliable Low Latency Communications) 지연 감소를 보고하며, 이는 분산된 관찰의 다양성과 프라이버시 보존이라는 FL의 장점이 현실의 O-RAN 운영 문제에 직접적인 성과로 연결될 수 있음을 시사한다. 동시에 연합 주기와 참여자 선택, 그리고 업데이트의 압축은 통신 오버헤드를 좌우하므로, 트래픽 유사도 기반의 에이전트 군집화나 선택적 참여가 제안되어 확장성과 안정성을 도모한다 [8].

VNF (Virtual Network Functions) 분할과 엣지 오케스트레이션의 맥락에서 CU (Centralized Unit) 및 DU (Distributed Unit) 사이에서 기능 블록을 동적으로 재배치하는 문제가 중요하다. 엣지에 배치된 에이전트가 로컬 지표를 바탕으로 분할 결정을 내리고, near-RT RIC에서 연합으로 정책을 집계하면, 재구성 비용을 줄이면서 지연 한계를 만족시키는 방향으로 진화할 수 있다. 보상 설계는 마이그레이션 비용과 성능 이득 사이의

균형을 요구하며, 연합은 현장별 이질성을 보존한 채 정책의 일반화를 촉진한다 [9]. 차량망과 같은 고속 이동 시나리오에서는 서비스 클래스, 차량 유형, 응용 계층으로 분해된 계층형 O-RAN 위에 FDRL을 결합하여, 계층 간 자원 조율을 통해 불안정한 채널과 급변하는 부하에서의 견고한 성능을 달성하려는 시도를 보고한다 [10]. 하나의 애플리케이션 인스턴스와 인프라를 여러 파티가 공유하는 멀티 테넌트 환경에서는 인프라 제공자와 MVNO (Mobile Virtual Network Operator), 그리고 단말 사이의 상호작용을 두 계층으로 나누어, 상위 계층의 자원 결정을 FDRL로 탐색하고 하위 계층의 자원 배분을 게임 이론과 결합해 유틸리티를 극대화하는 혼합 모델이 제안된다 [11]. 이 접근은 프라이버시를 유지하면서도 시장을 정책 학습에 반영할 수 있다는 장점이 있다. 한편 대규모 RAN에서의 통신 비용 부담과 데이터 이질성 문제를 완화하기 위해, 트래픽 양상에 따라 에이전트를 군집화하고 필요한 군집만 선별적으로 연합에 참여시키는 특화 전략도 유용한 대안으로 부상한다 [8].

2. 향후 과제

실제 구현에서는 학습과 집계의 위치가 가장 먼저 결정된다. 단말과 기지국 수준의 배치는 원시 데이터 보존과 빠른 반응을 보장하지만 연산 자원이 제한적이다. near-RT RIC은 수십 ms에서 수백 ms 수준의 제어 주기를 담당하며, 다수 xAPP 간의 조정을 실질적으로 수행하기에 적합하다. non-RT RIC과 클라우드르는 장기적 모델 관리와 정책 종류, 지식 이전의 중심이 된다. 이들 사이의 경계에서는 모델 버저닝과 롤백, A/B 검증과 드립트 탐지가 중요하며, 연합 참여자 선정과 업데이트 빈도, 그리고 통신량을 줄이기 위한 업데이트 압축이 통신 비용을 좌우한다 [7-8]. 프라이버시와 보안의 관점에서는 차등 프라이버시 및 안전한 집계와 같은 기법이 필수적이며, 비잔틴 참여자에 대한 강건 집계나 신뢰 점수 기반의 참여 통제가 요구된다. 무엇보다 제약을 수반하는 SLA 환경에서는 보상에 라그rangian 승수를 도입한 제약 DRL이 안정적 학습과 안전한 배포를 위해 현실적인 해법이 된다 [9].

향후 연구는 이기종 상태 및 행동 공간을 가진 에이전트 간의 지식 이전을 보다 자연스럽게 만들 필요가 있다. 공통 표현을 학습하는 표현 학습과 다중 작업 학습, 정책 종류는 유력한 도구이지만, 서로 다른 관측과 제약을 지닌 xAPP 사이에서 어느 수준까지 지식이 재사용 가능한지에 대한 체계적인 검증이 부족하다 [8]. 공정성과 격리 보장은 멀티 테넌트 운영의 핵심이며, 공정성 인식 보상과 제약 DRL, 그리고 정책의 설명 가능성을 결합한 도구가 필요하다 [4-5]. 확장성과 통신 비용 절감을 위한 에이전트 군집 선택, 적응형 연합 주기, 이벤트 기반 동기화, 모델 업데이트의 압축은 실제 배포에서 필수적인 기술 요소이다 [8]. 마지막으로 시장 및 기술 통합의 관점에서, 단말-MVNO-InP (Infrastructure Provider) 간 시장 메커니즘과 기술적 자원 제어를 공동으로 최적화하는 프레임워크가 요구되며, 이 과정에서 디지털 트윈을 활용한 시나리오 기반 검증은 안전한 온라인 학습의 현실적 경로이다 [11].

III. 결론

본 조사는 접속제어와 단말 연관, 자원할당과 xAPP 조정, VNF 분할과 엣지 오케스트레이션, 계층형 O-RAN, 멀티 테넌트 자원거래, 그리고 에이전트 특화 전략의 흐름을 하나로 연결하여, 학습 및 집계의 위치와 연합 체계, 그리고 운영 파이프라인의 결정이 성능과 비용, 안전성에 어떤 형태로 반영되는지를 정리하였다. 다음 단계의 과제는 제약 하에서의 안전한 온라인 학습과 설명 가능한 운영, 통신 비용과 탄소 발자국을 고려한 연합

최적화, 그리고 시장과 기술을 연결하는 통합된 오케스트레이션의 구현이다. 실제 배포의 제약을 고려한 벤치마크와 공개 데이터 및 코드의 축적은 그 출발점이 될 수 있다.

ACKNOWLEDGMENT

이 논문은 정부(과학기술정보통신부)의 재원으로 정보통신기획평가원-대학ICT연구센터(ITRC)의 지원을 받아 수행된 연구임 (IITP-2025-RS-2022-00156353, 50% / IITP-2025-RS-2024-00436887, 50%)

참 고 문 헌

- [1] W. J. Yun, S. Park, J. Kim, M. Shin, S. Jung, D. Mohaisen, "Cooperative Multiagent Deep Reinforcement Learning for Reliable Surveillance via Autonomous Multi-UAV Control," *IEEE Trans. Ind. Informat.*, vol. 18, no. 10, pp. 7086-7096, Oct. 2022.
- [2] T. T. H. Pham, W. Noh, and S. Cho, "Multi-Agent Reinforcement Learning Based Optimal Energy Sensing Threshold Control in Distributed Cognitive Radio Networks with Directional Antenna," *ICT Express*, vol. 10, no. 3, pp. 472-478, Jun. 2024.
- [3] A. Abouaomar, A. Taik, A. Filali and S. Cherkaoui, "Federated Deep Reinforcement Learning for Open RAN Slicing in 6G Networks," *IEEE Commun. Mag.*, vol. 61, no. 2, pp. 126-132, Feb. 2023.
- [4] Y.-J. Liu, G. Feng, J. Wang, Y. Sun and S. Qin, "Access Control for RAN Slicing based on Federated Deep Reinforcement Learning," in *Proc. IEEE Int. Conf. Commun.*, Montreal, QC, Canada, 2021, pp. 1-6.
- [5] Y.-J. Liu, G. Feng, Y. Sun, S. Qin and Y.-C. Liang, "Device Association for RAN Slicing Based on Hybrid Federated Deep Reinforcement Learning," *IEEE Trans. Veh. Technol.*, vol. 69, no. 12, pp. 15731-15745, Dec. 2020.
- [6] S. Messaoud, A. Bradai, O. B. Ahmed, P. T. A. Quang, M. Atri and M. S. Hossain, "Deep Federated Q-Learning-Based Network Slicing for Industrial IoT," *IEEE Trans. Ind. Informat.*, vol. 17, no. 8, pp. 5572-5582, Aug. 2021.
- [7] H. Zhang, H. Zhou and M. Erol-Kantarci, "Federated Deep Reinforcement Learning for Resource Allocation in O-RAN Slicing," in *Proc. IEEE GLOBECOM Conf.*, Rio de Janeiro, Brazil, 2022, pp. 958-963.
- [8] F. Rezazadeh, L. Zanzi, F. Devoti, H. Chergui, X. Costa-Pérez and C. Verikoukis, "On the Specialization of FDRL Agents for Scalable and Distributed 6G RAN Slicing Orchestration," *IEEE Trans. Veh. Technol.*, vol. 72, no. 3, pp. 3473-3487, Mar. 2023.
- [9] E. Amiri, N. Wang, M. Shojafar and R. Tafazolli, "Edge-AI Empowered Dynamic VNF Splitting in O-RAN Slicing: A Federated DRL Approach," *IEEE Commun. Lett.*, vol. 28, no. 2, pp. 318-322, Feb. 2024.
- [10] B. Hazarika, P. Saikia, K. Singh and C.-P. Li, "Enhancing Vehicular Networks With Hierarchical O-RAN Slicing and Federated DRL," *IEEE Trans. Green Commun. Netw.*, vol. 8, no. 3, pp. 1099-1117, Sep. 2024.
- [11] R. Ou, G. Sun, D. Ayepah-Mensah, G. O. Boateng and G. Liu, "Two-Tier Resource Allocation for Multitenant Network Slicing: A Federated Deep Reinforcement Learning Approach," *IEEE Internet Things J.*, vol. 10, no. 22, pp. 20174-20187, Nov. 2023.