

# 공공안전융합망 자율화를 위한 강화학습기술

김대연, 김중현  
고려대학교

dk725@korea.ac.kr, joongheon@korea.ac.kr

## Reinforcement Learning Technology for the Autonomization of Public Safety Converged Networks

Daeyeun Kim, Joongheon Kim  
Korea Univ.

### 요약

공공안전융합망은 경찰, 소방, 구급 등 다기관이 동일 사건 현장에서 음성, 데이터, 영상 서비스를 상호운용 가능하도록 제공하는 임무 중심 통신 체계이다. 그러나 재난 환경에서는 트래픽 급증, 이동성 증가, 전송 구간 장애, 보안 위협이 동시 발생하며, 기존의 수동 운용 기반 절차는 조치 조합의 복잡성과 인적 오류 위험으로 인해 확장성에 한계가 있다. 본 논문은 공공안전융합망의 자동화와 자율화를 구분하여 정의하고, 종단 간 폐루프 관점에서 강화학습 (Reinforcement Learning, RL)을 자율 의사결정 엔진으로 활용하는 방법을 제시한다. 구체적으로 제로터치 네트워크 및 서비스 관리 (Zero-touch Network and Service Management, ZSM), 네트워크 데이터 분석 기능 (Network Data Analytics Function, NWDAF), 개방형 무선 접속망 (Open Radio Access Network, O-RAN) 및 무선 접속망 지능형 제어기 (RAN Intelligent Controller, RIC)를 결합한 참조 아키텍처를 제안한다. 또한 안전 RL 가드레일, 단계적 도입 로드맵 및 핵심 성능 지표 (Key Performance Indicator, KPI)를 체계적으로 정리한다.

### I. 서론

재난 및 복합 사건 상황에서는 현장 단말 수의 급증과 함께 임무 트래픽이 집중되며, 무선 접속망 (Radio Access Network, RAN) 혼잡, backhaul 제약, 코어망 기능 과부하가 연쇄적으로 발생한다. 공공안전융합망은 임무 무전 서비스인 mission critical push-to-talk (MCPTT), 임무 영상 서비스인 mission critical video (MCVideo), 임무 데이터 서비스인 mission critical data (MCData)를 기반으로 낮은 지연과 높은 가용성을 보장해야 한다 [1]. 그러나 운용자가 사건 유형과 네트워크 상태를 해석하고 매번 수동으로 서비스 품질 (Quality of Service, QoS) 정책, 우선순위, traffic steering, 장애 우회 등을 수행하는 방식은 신속성과 일관성을 동시에 담보하기 어렵다.

따라서 본 논문은 다음 기여를 제공한다. 첫째, 공공안전융합망에서 자동화와 자율화의 개념적 경계를 운영 관점에서 명확히 정의한다. 둘째, 제로터치 네트워크 및 서비스 관리 (Zero-touch Network and Service Management, ZSM) [2], 네트워크 데이터 분석 기능 (Network Data Analytics Function, NWDAF) [3], 개방형 무선 접속망 (Open Radio Access Network, O-RAN) [4] 및 무선 접속망 지능형 제어기 (RAN Intelligent Controller, RIC)를 결합한 참조 아키텍처를 제안한다. 셋째, 공공안전 특화 안전 가드레일을 포함한 안전 강화학습 (Reinforcement Learning, RL) 설계를 제안한다.

### II. 공공안전융합망 자동화 및 자율화 정의

자동화는 사람이 수행하던 반복 운용 절차를 정책 기반으로 기계화하여, 사건 발생 시 필요한 구성 요소를

일관된 절차로 신속히 수행하는 것을 의미한다. 대표 기능은 (i) 사건 단위 그룹 통화 및 서비스 provisioning, (ii) QoS 및 우선접속 설정, (iii) 장애 감지 및 복구, (iv) 단말/사용자 권한 부여 및 회수, (v) 로그 수집 및 감사 추적이다. 이때 자동화는 정해진 규칙을 빠르게 수행하는 성격이 강하다.

자율화는 자동화 위에서 환경 변화(트래픽, 이동성, 장애, 위협)에 대응하여 의사결정 자체를 적응적으로 최적화 하는 능력이다. 예를 들어 동일한 사건 유형이라도 현장 단말 구성, 무선 품질, backhaul 용량이 다르면 최적 QoS 분배와 트래픽 제어는 달라져야 하며, 자율화는 이러한 맥락 의존 결정을 폐루프로 수행한다. 공공안전 환경에서는 평균 성능 향상보다 임무 서비스 하한 성능 보장과 정책 준수가 더 중요하므로, 자율화는 안전 제약 하 최적화로 모델링되어야 한다.

### III. 표준 기반 폐루프 참조 아키텍처

본 연구에서는 공공안전융합망 자율화를 위한 3 계층 구조를 제안한다. 서비스 계층은 MCPTT-MCVideo-MCData 및 사건 관리, 위치/상황판을 포함한다. 지능 계층은 정책 엔진, 오피스트레이터, 관측 및 분석을 포함하며, 이 영역이 폐루프의 의사결정 중심이 된다. 인프라 계층은 RAN, 코어망, 전송망, multi-access edge computing (MEC) 및 network functions virtualization (NFV) 기반 자원 풀로 구성된다.

NWDAF 자율 의사결정의 전제는 고품질 관측 데이터이다. NWDAF는 네트워크 데이터 분석을 통해 혼잡 예측, 이상 징후 탐지, 서비스 경험 추정 등을 제공하며 폐루프의 Observe 단계에 해당한다. 공공안전 시나리오에서는 구역/셀별 부하, 지연/손실, 임무 세션 수,

backhaul 상태, 이동형 기지국 투입 여부, 보안 이벤트를 통합 지표로 수집·정규화하는 것이 핵심이다.

O-RAN RIC 무선 구간의 혼잡 및 이동성 제어는 임무 서비스 품질에 직접적인 영향을 미친다. O-RAN RIC는 near-real-time 제어 루프에서 정책 기반 제어 및 인공지능/머신러닝 기반 최적화를 수행하는 구조를 제공하며 [4], 폐루프의 Act 단계에서 핵심 실행점이 된다.

#### IV. 안전 강화학습 기반 단계적 도입 로드맵

공공안전용합망의 자율화 과정에서 RL의 오작동을 방지하고 신뢰성을 확보하기 위해서는 무엇보다 안전성 검증이 선행되어야 한다. 이를 위해 본 논문에서는 다음과 같은 4 가지 핵심 가드레일을 필수 요건으로 정의한다. (1) Hard constraints의 정책화: MCPTT의 최저 성능 보장 및 핵심 역할에 대한 우선권은 학습의 대상이 아닌, 고정된 정책으로 설정하여 절대적인 안전을 확보한다. (2) Safety filter: 정책망이 산출한 행동은 실제 망에 적용되기 전 규정 위반 여부를 검사받으며, 위반이 감지될 경우 즉시 사전에 정의된 보수적 대체 행동으로 치환된다. (3) Shadow mode: 도입 초기에는 RL이 직접 제어하지 않고 추천만 수행하며, 운영자의 검토 및 승인을 거쳐 적용하는 방식으로 안전성을 검증한다. (4) Fail-safe 및 rollback: 자동화 프로세스에 장애가 발생할 경우, 즉시 음성 서비스 우선 중심의 안전 모드로 복귀하여 기본 통신 기능을 유지한다.

이러한 가드레일을 바탕으로, 자율화 기술의 안정적인 정착을 위해 다음과 같은 4 단계의 순차적 도입 로드맵을 제안한다. 우선 1 단계에서는 관측 데이터의 정규화와 사건 템플릿을 정의하고, NWDAF 및 OAM 이벤트 기반의 데이터 파이프라인을 구축하여 자율화의 기초를 마련한다. 이어지는 2 단계는 반자동 오케스트레이션 단계로, RL은 시뮬레이션 및 운영자 의사결정 지원을 위한 추천 정책으로만 활용한다. 3 단계로 나아가면 제한적 자율을 적용하여, 사전 승인된 특정 조건과 행동 범위 내에서만 자동 실행을 허용한다. 마지막 4 단계에서는 축적된 데이터와 신뢰성을 바탕으로 ZSM 기반의 종단 간 완전 자율화를 달성하는 것을 최종 목표로 한다 [2].

#### V. 평가 지표 및 향후 연구 방향

자율화 시스템의 성능을 검증함에 있어, 단순한 평균 처리량보다는 임무 성공 여부와 직결된 지표가 우선시되어야 한다. 이에 따라 본 논문에서는 사건 생성 후 통신 체계 완비 시간, MCPTT 호출 설정 시간 및 성공률, MCVideo 업링크 사용 시간 등을 핵심 성능 지표 (Key Performance Indicator, KPI)로 선정할 것을 제안한다. 아울러 시스템의 안정성을 대변하는 평균 복구 시간, 정책 위반률, 그리고 기관 간 상호운용 성공률 또한 중요하게 다루어져야 한다.

본 연구를 빌전시키기 위한 향후 연구 방향으로는 크게 네 가지를 제안한다. (i) 디지털 트윈 (Digital Twin)을 활용하여 다양한 재난 시나리오를 생성하고 정책을 사전에 검증하는 연구; (ii) 현실망 적용 전 시행착오를 최소화하기 위한 오프라인 강화학습 기반의 초기 정책 학습; (iii) 재난 현장의 예측 불가능한 통신 환경 하에서도 안정적으로 동작하는 학습 기법; (iv) 마지막으로 보안 위협 대응 시 비용과 효용의 균형을 보장할 수 있는 정교한 제약 설계에 대한 연구가 지속적으로 수행될 필요가 있다.

#### VI. 결론

본 논문은 기존 수동 운영의 한계를 극복하기 위해 공공안전용합망의 자동화와 자율화를 명확히 구분하여 정의하고, ZSM-NWDAF-O-RAN RIC을 유기적으로 결합한 표준 정합적 폐루프 참조 아키텍처를 제시하였다.

나아가 안전 가드레일, 단계적 도입 로드맵, 그리고 KPI 중심의 평가 체계를 체계적으로 제안하였다. 종합해보면, RL은 공공안전용합망의 자율화를 위한 의사결정 엔진으로서 충분한 유효성을 가지나, 실제 현장 적용을 위해서는 hard constraints 및 엄격한 검증 절차를 포함한 설계 및 연구가 필수적으로 수행되어야 함을 확인하였다.

#### ACKNOWLEDGMENT

이 성과는 정부(과학기술정보통신부)의 재원으로 한국연구재단의 지원을 받아 수행된 연구임(RS-2025-00561377). 본 논문의 교신 저자는 김중현임.

#### 참 고 문 헌

- [1] M. H. Qutqut, A. Al-Alajullah, and H. S. Hassanein, "A survey on public safety networks: Protocols, standards, and architectures," *IEEE Communications Surveys & Tutorials*, vol. 22, no. 4, pp. 2633– 2664, Fourthquarter 2020.
- [2] C. Benzaid and T. Taleb, "AI-driven zero touch network and service management in 5G and beyond: Challenges and research directions," *IEEE Network*, vol. 34, no. 2, pp. 186– 194, March/April 2020.
- [3] S. Sevgican, M. Turan, K. Gökarslan, H. B. Yilmaz, and T. Tugcu, "Intelligent network data analytics function in 5G cellular networks using machine learning," *J. Commun. Netw.*, vol. 22, no. 3, pp. 269– 280, Jun. 2020.
- [4] M. Polese, L. Bonati, S. D'Oro, S. Basagni, and T. Melodia, "Understanding O-RAN: Architecture, interfaces, algorithms, security, and research challenges," *IEEE Communications Surveys & Tutorials*, vol. 25, no. 2, pp. 1376– 1411, Secondquarter 2023.