

네트워크와 AI 기술 동향

Trends in Network and AI Technologies

김태연 (Tae Yeon Kim, tykim@etri.re.kr)

지능네트워크연구실 책임연구원/실장

고남석 (Namseok Ko, nsko@etri.re.kr)

데이터중심네트워크연구실 책임연구원/실장

양선희 (Sunhee Yang, shyang@etri.re.kr)

네트워크연구본부 책임연구원

김선미 (Sun Me Kim, kimsunme@etri.re.kr)

네트워크연구본부 책임연구원/본부장

ABSTRACT

Recently, network infrastructure has evolved into a BizTech agile autonomous network to cope with the dynamic changes in the service environment. This survey presents the expectations from two different perspectives of the harmonization of network and artificial intelligence (AI) technologies. First, the paper focuses on the possibilities of AI technology for the autonomous network industry. Subsequently, it discusses how networks can play a role in the evolution of distributed AI technologies.

KEYWORDS Autonomous Network, Distributed Intelligence Network, Network by/for AI

1. 서론

지난 30여 년간 네트워크 기술은 발전을 거듭하여 전송 및 교환, 네트워크 제어 및 관리, 유무선 접속, 인터넷 기반의 멀티미디어 서비스 기술 등에 괄목할 만한 혁신을 만들어 왔다. 그 결과로 네트워크는 수백 기가(G)급의 광액세스와 테라급 광전달망을 기반으로 SDN/NFV 기반의 유연한 인프라 구조를 지향하며, 네트워크와 컴퓨팅이 밀결합되는 인프라로 계속 진화하고 있다. 네트워크는 IoT, 클라우드, 5G 모바일망, 데이터, AI 기술 등과 밀

결합되며, ‘초연결 데이터 중심 사회’의 핵심 기반 인프라로서 미래 국가사회 경쟁력의 관건이 될 것으로 전망하고 있다.

한편, AI 기술은 2010년대 이후 ML, DML, DRL 등으로 기술혁신을 거듭하며, 분산 AI 에이전트, Swarm DRL 등으로 기술 발전이 급속히 진행 중이다[1], 향후 개인의 삶과 사회시스템의 지능화에서 더 나아가, 미션 크리티컬한 영역으로 확대될 전망이다. 이에 따라서 AI 서비스를 위한 인프라는 다중영역·조직 간에 AI 기반 협업을 지원하고, 기존 AI의 성능 한계를 극복하기 위해 분산 기반의 실시

* DOI: <https://doi.org/10.22648/ETRI.2020.J.350501>

* 이 논문은 2020년도 정부(과학기술정보통신부)의 재원으로 정보통신기획평가원의 지원을 받아 수행된 연구임[No. 2019-0-0060, 초연결 지능 인프라 원천기술 연구개발].



본 저작물은 공공누리 제4유형

출처표시+상업적이용금지+변경금지 조건에 따라 이용할 수 있습니다.

©2020 한국전자통신연구원

간 Secure AI로 진화하며 그 영향력이 계속 커지고 있다[2,3].

이처럼 네트워크는 사람들의 삶과 인터스트리를 촘촘히 연결하고, AI 기술은 연결된 사회의 빅데이터를 수집 분석함으로써 개인과 국가사회, 산업 전반의 지능화와 혁신을 가속하는 데이터 중심 사회(Data-driven Society)로 진입하고 있다.

그렇다면 네트워크 기술산업 관점에서 AI 기술은 어떤 가능성을 가졌을까? 또한, AI 기술의 진화발전에 네트워크는 어떤 역할을 할 수 있을까? 본 고에서는 이러한 관점에서 네트워크와 AI 기술의 동향과 발전 전망을 살펴보고자 한다. 많은 기술산업 영역에서 AI 기술의 활용이 폭발적으로 늘어나고 있듯이, 네트워크 영역에서도 다양한 이슈에 대한 AI 기술 활용이 모색되는 중이다. 대표적으로, ML 기반의 예방적 네트워크 보안이나[4,5], 선제적 장애 대응을 통한 고신뢰 네트워크(Resilient Network)[6,7], 네트워크 운영관리 자동화·자율화[8-10], 데이터 기반 광전송 신호처리 최적화[11] 등의 연구가 진행 중이고, 일부 기술은 실제 적용이 시작되고 있다.

이러한 다양한 시도 중에 본 고에서는 주로 ML 기반 네트워크 자동화와 자율화 및 분산 AI를 위한 네트워킹 기술 위주로 살펴보고자 한다. II장에서는 네트워크의 발전에 있어서 AI 기술의 역할과 활용 현황을 개괄하고자 한다. III장에서는 Autonomous Network를 위한 AI 기술 동향을 살펴보고, IV장에서는 차세대 AI를 위한 네트워크 기술 동향을 조사 분석한다.

II. 네트워크의 발전과 AI의 역할

1. 네트워크 진화와 챌린징 이슈

최근의 서비스 환경 변화는 초연결화¹⁾, 데이터

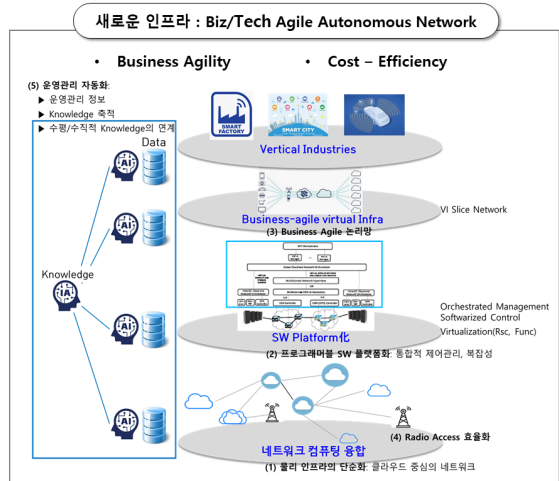


그림 1 Biz/Tech Agile Autonomous Network(BAAN) 개념도

중심 사회화, 5G에 기반한 Connected VI(Vertical Industry) 및 신뢰와 안전에 대한 요구 증대 등의 4 가지 트렌드를 중심으로 움직이고 있다. 텔코는 이처럼 동적으로 변하는 서비스 환경 변화에 신속하게 대응하여 비즈니스 민첩성을 비용 경제적으로 지원할 수 있는 인프라를 지향하고 있다. 기술적으로는 클라우드피케이션의 확산, 인프라의 가상화와 공유, 인프라의 소프트웨어라이제이션과 API 기반 프로그래머빌리티에 기반한 플랫폼화, 오픈소스 확산 등 핵심 요소기술이 무르익고 있다.

그림 1은 이러한 서비스 및 기술 환경 변화에 따른 미래 인프라 모델로서 'Biz/Tech Agile Autonomous Network(BAAN)'을 보여준다. BAAN은 서비스나 환경변이에 운영자의 개입을 최소화하면서 스스로 구성(Self-provisioning), 설정(Self-configuring)되며 스스로 대응(Self-monitoring, diagnosing, healing)하며, 지속적으로 스스로 최적화해 나

1) IDC는 네트워크에 연결되는 디바이스 수가 2025년에 1인당 평균 5개에 이르고, IoT가 발생하는 데이터가 79.4제타바이트에 이를 것으로 전망한다.

갈 수 있는 네트워크를 지향한다. BAAN 인프라의 주요 특성과 챌린징 이슈를 간단히 살펴보면 다음과 같다.

가. 클라우드 중심의 단순화되는 물리 인프라

물리 인프라는 전통적인 전달망에서 벗어나 클라우드를 중심으로 그들 간을 연결하는 네트워크로 단순화되는 구조를 지향한다. 클라우드 내부망, 클라우드 간을 연결하는 인터-클라우드, 시스템 내부 데이터 버스 등으로 단순화되며, 대용량 트래픽의 고속 전달, 지연 및 오류 등에서 응용별 트래픽의 QoS 차별적 지원 필요성이 커지고 있다. 클라우드는 중앙집중 및 에지 클라우드 형태의 분산 구조로 진화되며, 에지에서의 민감한 응용의 지원, 빅데이터의 스마트 데이터화, 분산 AI 응용 등을 수용하며 확대 중이다. 분산 환경에서의 취약한 보안 이슈, 분산된 에지 클라우드 간의 데이터 공유 등 해결되어야 하는 문제를 안고 있다.

나. 소프트웨어 플랫폼화 및 프로그래머블화

인프라의 컴퓨팅 및 네트워크 자원은 가상화되고 개방되며, 네트워킹 기능도 소프트웨어 모듈화, 마이크로서비스 형태로 발전 중이다. 개발단계에서부터 하드웨어와 밀접합된 시스템 소프트웨어 형태로 도입되던 네트워크 디바이스들은 범용 서버를 기반으로 API 기반의 프로그래머빌리티와 통합적 오케스트레이션을 지향한다. 새로운 서비스 요구나 환경변화에 유연하고 신속하게 대응할 수 있는 SDN/NFV 기반 개방형 플랫폼으로 진화하고 있다.

한편, 인프라 규모가 거대해지고, 가상화나 동적 프로그래머빌리티와 같은 복잡성을 최적으로 지원하기 위해서는 운용 및 제어관리에서의 AI 기술을 접목한 기술혁신이 필연적으로 요구되고 있다.

다. 비즈니스 민첩성을 지원하는 논리 인프라 지향

가상화·개방화된 인프라 플랫폼상에 서비스 요구와 환경 변이에 민첩하게 대응할 수 있는 VI 별 맞춤형 슬라이싱 인프라를 지향한다. eMBB, mMTC, uRLLC 같은 5G의 다양한 B2B 응용 유형 지원을 위해 네트워크 슬라이싱 개념구조를 기술적으로 어떻게 구체화할 것인지가 텔코 5G 성공의 관건이 될 것이다. 네트워크 슬라이스 간 격리(Isolation) 수준, 응용 유형별 요구되는 특수한 네트워크 능력의 지원 방법(예, non-IP 지원, 클록 동기화, AI를 비롯한 프로세싱 요구사항 등), 보안이나 운영관리에 대한 차별화된 요구사항을 어떻게 지원할 것인지에 대한 기술적 대안과 진화 방안에 대해 논의가 진행 중이다.

라. 운영관리 통합화 및 자동화·자율화

전통적으로 전문 지식과 경험을 가진 운영자에 의해 네트워크 인프라의 설계 및 설치, 운영 관리가 이루어져 왔다. 그러나 네트워크 규모가 커지고, 가상화 및 공유 등으로 운영 환경이 복잡해졌으며, 비즈니스 민첩성 지원을 위한 실시간 제어 관리 요구가 증대되고 있어, 운영관리 기술 혁신이 네트워크 분야의 가장 큰 화두가 되고 있다.

대규모로 수집되는 운영관리 데이터로부터 지능적 운영을 위한 지식을 분석해 내고, 이를 기반으로, 인프라의 다계층·다중도메인을 통합적으로 연계하는 지식기반 운영관리를 지향하는 자율 네트워크로 기술산업 혁신이 가속화되고 있다. 이에 더하여, 인프라 사업자들은 Autonomous Network에 대해 비즈니스, 고객, 네트워크 인프라의 구축 및 운용, 3rd 파티 협업 등 텔레콤 인더스트리 전 영역의 통합적 디지털 트랜스포메이션이라는 비전을 지향하고 있다.

2. AI 기술의 역할과 연구 방향

앞서처럼 네트워크는 인프라의 플랫폼화, 비즈니스나 기술 환경 변화에 신속하게 대응할 수 있는 Agile Infra로서의 복잡도 증가, 5G VI 수용을 위한 안정성과 신뢰성 강화에 대한 기술 혁신에 직면하고 있다. AI 기술은 이러한 네트워크 분야의 요구에 대해 기술적 가능성을 보여주고 있다. 예를 들어, 복합적 데이터의 통합적 분석에 기반한 선제적 장애 대응, 운영 지식 기반 네트워크 및 클라우드 관리의 최적화·자율화, 광송수신 신호처리에서의 알고리즘 기반 전통적 R&D의 한계를 데이터 기반 R&D로 전환하는 시도 등이 이루어지고 있다. 특히, 텔코가 관심을 갖는 주요 이슈는 다음과 같다.

- 사이버 공격에 안전한 네트워크를 위한 AI

기술

- 장애나 성능 변이 등에 유연한 대응이 가능 (resilient)하여 항상 신뢰할 수 있는 네트워크 실현을 위한 AI 기술
- 네트워크 운영관리의 복잡성과 인프라 파편화를 해결할 수 있는 AI 기술로서, 다양한 응용이나 다계층·다중 도메인 네트워크 구성 및 기술의 다양화 등에 대해 통합적인 운영 관리를 지원하는 Autonomous Network 실현 AI 기술
- 무선 액세스 자원의 최적 활용을 위해, 환경 적응적인 RAN 최적화를 지원하는 AI 기술

표 1에 AI 기반 연구가 이루어지고 있는 주요 네트워크 기술 이슈 동향을 요약하였다.

표 1 네트워크 기술 이슈 및 AI 기술 연구 방향

네트워크 기술 이슈	AI 기술 연구방향 개요
트래픽 분석 및 보안[4,5]	<ul style="list-style-type: none"> • 실시간 및 사전 예방적 Anomaly 기술 확보 필요 • 트래픽 패턴 분석으로부터 최대한 실시간으로 Anomaly 감지가 가능한 ML 기술(McGill Univ, Intel Netrolx, 네트워크 보안 전문업체 등) ※ Supervised/Unsupervised/Mixed 등 다양한 AI 기법 적용한 사전예방적 기술 확보 지향 • 최근, 분산 AI 환경을 고려한 AR/VR 등 뉴미디어 트래픽 QoS 모델링 연구시작(Rutgers Univ., Univ. of California)
선제적 장애관리[6,7]	<ul style="list-style-type: none"> • 임무지향 산업 인터넷, 5G VI 자원 등 초정밀 고신뢰 인프라를 위한 Resilient Network 필요 • 통합적 운용관리 빅데이터 분석에 기반한 장애 진단·예측의 정확도와 대응시간 단축 지향 • 구조적/비구조적 네트워크 빅데이터의 선처리 및 의미 있는 인텔리전스 추출 문제해결 • KT, 일본 NII, 학계/연구계 등 연구개발 투자 중
Optic/RF 신호전송 최적화[11]	<ul style="list-style-type: none"> • 실제 환경에서의 RF/Optical 등 물리계층 신호 처리의 정교한 최적화 필요 • 설계, 개발 및 실제 운영 등 전 주기에 대해 Data-driven 신호처리 및 최적화 연구 초기 단계(학계) • RAN/광전송 최적화(학계)
운영관리 자동화·자율화[8-10]	<ul style="list-style-type: none"> • hyper-connectivity 지원, 네트워크 제어관리의 복잡도 증가 및 네트워크의 파편화 문제, 운영자에 의한 부분적 최적화의 비효율성 심화, 운영 지식 및 스킬 부족 등의 문제해결 필요 • 사업자 주도의 ML 기반 Knowledge-Defined 자율네트워크 기술 개발 투자 시작 -5G RAN/코어/액세스/광전송, 데이터센터 망 등 최적 구성 및 자율 운용(full lifecycle automation and max utilization, zero wait/zero touch/zero trouble service 지향, best-possible user experience) -자원 가상화 최적화(자원할당, 기능배치, 최적 경로제어, 오토스케일링 등에 대해 국내외 연구 진행 중(Postech, China Unicom, Telefonica) -VI를 위한 슬라이싱 자동화·최적화 기술 개발 및 초기 적용 중(NTT, Ericsson, Huawei) -상황인지 기반 RAN/광전달망 최적화 등(SKT Tango, AI-driven OSS) • 오픈소스 기반 Autonomous Network 플랫폼화 진행 초기 단계(Acumos, Kubernetes)

한편, AI 기반 네트워크 실현을 위해 표준화와 오픈소스 커뮤니티 활동도 활발히 진행되고 있다. 네트워크 자동화·자율화를 위한 관건이 되는 개방형 인터페이스(Open Interface), 개방형 API와 오픈소스(Open Source SW & Open APIs), 개방형 참조모델(Open Reference Design)들에 대한 프로젝트들이 진행 중이다. 대표적으로, 리눅스 재단에서는 ONAP(Open Network Automation Platform), OPNFV(Open Platform for NFV) 및 AI App 개발 및 배포 프레임워크인 Acumos 등의 플랫폼 개발 프로젝트를 진행 중이다[12,13]. 한편, 표준화 영역에서는 3GPP의 5G 네트워크 데이터 분석(NWDAF), 네트워크 종단 간 관리 자동화(ETSI ZSM), Edge Automation Platform(IEEE Future Networks), ML을 적용한 Future Network/5G 자동화 프레임워크(ITU-T ML5G) 등 다양한 활동이 진행 중이다 [14].

III. Network by AI 기술

5G와 AI 기술의 발전으로 네트워크 관리/응용 기술에도 새로운 변화가 시작되었으며, 그간 수동적인 설정과 제어를 통한 운영방식에 의존해왔던 네트워크가 이제는 스스로 데이터를 분석하고 의사결정을 내리는 네트워크 자동화 및 지능화 기술을 통해 최적의 상태로 자동 제어와 관리가 가능해질 것으로 전망되고 있다.

특히 네트워크의 초공간화가 가속화되고, 서비스 및 기술과 사업자의 다양성 확대로 운영 관리 복잡성이 기하급수적으로 증가하는 6G 시대(2030년대 즈음 도래할 차세대 네트워크)에는 네트워크 상태 정보나 실시간 이벤트, 로그 등과 같은 운영 데이터가 자동으로 수집된 후, 빅데이터 분석 및 인공지능 기술 적용을 통해 최적의 의사결정 및 정

책이 자동으로 만들어지는 지능형 자율인프라로 진화될 것으로 예상된다.

이 장에서는 AI/ML 기술을 활용한 네트워크 최적화 및 자동화에 대해 크게 가상기능·자원 관리, 네트워크 슬라이싱 자동화 및 서비스 품질 최적화 등 3가지로 구분하여 소개한다.

1. 가상기능·자원 최적 배치 및 관리

다양한 네트워크 서비스를 구성하는 프로그래머블 네트워크에서 가상기능·자원 최적 배치 및 관리 기술은 서비스 요구사항과 자원의 효율성, 종단 간 트래픽 연결성을 고려하여 가상기능을 최적으로 배치하고, 지속적으로 최적화하기 위한 기술로 정의된다.

NFV, SDN, MEC 기술에 기반한 가상 네트워크 인프라에서 네트워크 서비스(NS)는 가상 자원의 할당을 통해서 인스턴스화되는 가상 네트워크 기능(VNF)의 순차적 조합을 통해서 구성되며, 다양한 NS를 생성하기 위해서는 VNF들 간의 서비스 플로우를 구성하는 서비스 기능 체이닝(SFC) 혹은 VFG(VNF Forwarding Graph) 기술이 요구된다.

이러한 가상기능 및 가상자원의 최적 배치 및 임베딩 문제는 제한적인 인프라와 SLA를 고려할 경우 NP-hard 문제가 발생하며, 이를 해결하기 위하여 통상적으로 휴리스틱 방법을 사용하고 있으나, 5G 네트워크 이후 지속적으로 증가되는 다양한 네트워크 서비스가 가상 네트워크 인프라 위에서 동작하기 위해서는 이를 구성하는 복수의 가상 네트워크 기능 사이의 트래픽 전달을 위한 SFC/VFG를 최적으로 구성해야 하며, 이에 대한 최적 배치 및 경로를 결정하는 인공지능 기술을 필요로 한다.

NEC는 최적의 VNF(Virtualized Network Function)

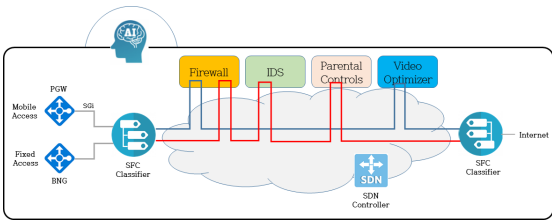


그림 2 가상기능 · 자원 배치 최적화 개념도

리소스 크기 및 배치를 개별적으로 고려하고 환경 다양성에 적응하기 위해 적응 최적화 메커니즘을 포함하는 강화학습(RL: Reinforcement Learning)을 적용하는 방법을 제안하였다. 강화학습을 가속화하기 위한 2단계 방법을 적용하여, 첫 번째 단계에서 각각의 SFC에 대한 일반적 성능 특성을 최적화하고, 2단계에서 서비스 환경의 특수성을 감안한 최적화 모델을 도출하는 방법을 통하여 기존 강화학습에서 소요되는 10,000번 이상의 탐색 부하를 1,000번으로 줄이는 성능 가속화 결과를 도출한 바 있다[15].

스페인의 Basque County 대학에서는 제한적인 인프라 자원과 다양한 네트워크 서비스 특성을 고려한 강화학습 기술과 휴리스틱 방법을 병합한 하이브리드 강화학습 기술을 제안하였다. 가상기능 · 자원 최적 배치를 위해서 Neural Combinatorial Optimization 기법을 적용한 강화학습을 적용하여 VNF의 배치와 Forwarding Graph(FG)의 임베딩 시의 실패율을 낮추고 인프라의 컴퓨팅 자원 최적화를 위한 실험을 통해서 휴리스틱 알고리즘이 적용된 하이브리드 에이전트 적용 효과를 검증하였다 [16].

ETRI에서는 응용서비스에 대한 응답시간을 줄이기 위해 선제적 예측을 통해 스케일 인/아웃을 지원할 수 있는 ML 기반 오토스케일링 기술 연구가 진행 중이며, 중국과기대 등에서 네트워크 서비스의 요청 거부 수 및 거부 비율, 처리량, 종단

간 지연, VNF 실행시간을 고려한 VNF 배치 알고리즘에 관한 연구 및 강화 학습 기반 VNF 및 서비스 체인 구성을 위한 2단계 최적화 프레임워크를 제시하여 VNF에 대한 CPU 코어의 최적 할당을 결정하기 위한 비용을 줄이고 평균 처리량을 향상시키기 위한 방안에 대한 연구를 수행하고 있다 [17,18].

2. 네트워크 슬라이싱 자동화

5G 모바일 네트워크 서비스는 신뢰성, 대기시간, 용량, 보안 등의 네트워크 성능요소가 응용에 기반한 맞춤형 연결을 요구하고 있으며, 이를 위한 네트워크 슬라이싱 기술은 동일한 물리적 인프라를 통해 여러 논리 네트워크를 제공하며 SDN과 NFV를 함께 사용하여 기존 물리적 장비 및 서비스에 영향 없이 네트워크의 논리적 분리가 가능하게 하는 기술이다.

네트워크 슬라이싱 자동화 기술은 단일 물리 인프라상에 다양한 서비스를 동시에 지원하기 위해 네트워크 및 컴퓨팅 인프라를 독립적인 논리 인프라로 구성 · 운영하기 위한 자동화 기술로서 다양한 사용자 및 서비스 요구에 따라 이종의 슬라이스들을 실시간으로 동적 할당하고, 최적으로 배치 · 운영하며, 슬라이스 간의 영향을 배제할 수 있도록

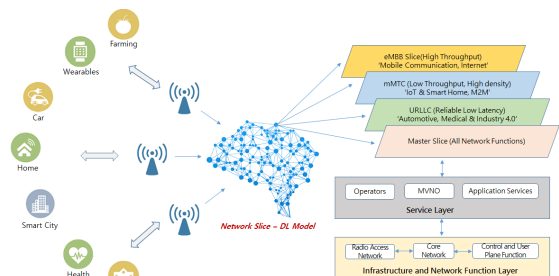


그림 3 지능형 네트워크 슬라이싱 자동화 개념도

인공지능 기반의 선제적 의사결정 및 최적화 기술이 요구된다.

AI/ML 기반의 네트워크 슬라이스 제어관리를 위한 연구 역시 대학을 중심으로 연구가 진행되고 있으며 대표적인 내용은 다음과 같다.

네트워크 연결 장치의 유형, QoS 클래스 식별자(QCI), packet delay budget, 최대 패킷 손실, 시간 및 요일 등 네트워크와 장치 모두에서 가장 관련성이 높은 KPI를 기반으로 랜덤포레스트(RF) 기법을 활용하여 최적의 네트워크 슬라이스를 추론 할당하는 DeepSlice 기술을 적용하여 슬라이스 분류의 정확도 향상에 관련한 연구가 소개되었으며[19], 5G 코어 네트워크의 서비스 기반 아키텍처와 최적의 작업 할당 알고리즘을 제안하고 SDN, NFV 및 엣지 컴퓨팅과 결합하여 네트워크 기능의 분산 및 주문형 배포, 서비스 보장 네트워크 슬라이싱, 유연한 네트워크 기능 오케스트레이션 및 최적의 워크로드 할당을 제공할 수 있으며, 시뮬레이션 결과를 통해 네트워크 운영 비용 절감 측면에서 효과적임을 보인 바 있다[20].

또한, 5G 인프라의 네트워크 슬라이싱에 대하여 SDN/NFV 기반 큐잉 전략을 기반으로 새로운 접근 방식을 제시하고 물리적 테스트 설정을 통해 성능을 검증하고 부분 과부하 시에도 네트워크 슬라이스가 가능함을 보임으로써 네트워크가 대규모로 슬라이스 된 경우를 제외하고는 성능 저하가 없는 확장성, 동적 데이터 전송률 할당, 슬라이스 재구성에 있어서 실제 조건에서 서비스가 보장됨을 증명하고 있다[21].

5G 네트워크 슬라이싱의 다중 테넌시 특성에 관한 연구로는 슬라이스 다중 테넌트 네트워크에서 제약 조건의 일부로 할당된 대역폭과 전송 전력으로 최대 용량을 달성하는 데 필요한 자원 할당 문제를 시뮬레이션을 통해 슬라이스의 개수,

슬라이스 내의 사용자 수와 전송전력이 MVNO의 용량에 미치는 영향에 관한 연구가 진행 중이다[22].

3. 네트워크 서비스 품질 최적화

5G 네트워크에서는 mMTC, eMBB, uRLLC와 같이 서로 상이한 QoS 요구사항을 가지고 있는 서비스에 맞는 네트워크의 성능 제공이 요구되고 있다. 이를 위해서 QoS 관련 이벤트에 대한 추적 관리가 요구되며, 현재 이벤트에 대한 정확한 감지, 자동 완화와 즉시 처리 그리고 향후 발생할 수 있는 QoS 관련 이벤트에 대한 신뢰성 있는 예측 기능이 필요하다.

하지만 기존의 하드웨어 장비 기반의 정적인 사결정이 아닌 유동성을 지닌 프로그래머블 인프라에서 동적으로 변화되는 상황을 고려한 QoS 보장 문제는 전문가 시스템으로 해결하기 어려운 문제로 인식되고 있으며, 특히 복잡다단한 네트워크의 이벤트와 QoS와의 상관관계는 수학적 분석의 한계에 직면하게 되었다.

이에 따라서 최근 네트워크 인프라 및 사용자로부터 발생하는 위치 정보, 이동성 및 통화 패턴과 같은 유용한 정보를 포함하는 많은 양의 데이터

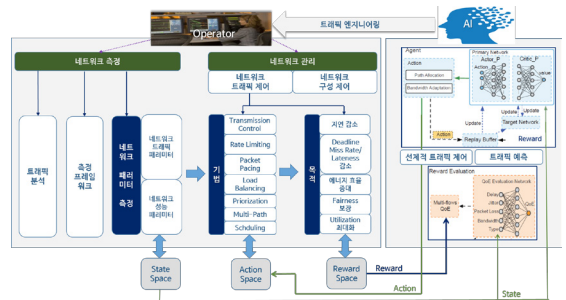


그림 4 네트워크 서비스 품질 최적화 개념도

를 기계 학습 방법과 결합하여 네트워크 및 통신 구성 요소의 품질 최적화에 대한 필요성이 증대되고 있다.

AI/ML 기반의 네트워크 서비스 품질 최적화를 위한 연구로서 명시적인 프로그래밍에 의한 메커니즘이 아닌 축적된 데이터를 통해서 과거의 QoS 관련 정보 및 이상 현상으로부터 학습하고 이들 간의 관계를 자동으로 정확하게 재구성하는 지도 학습 기반의 QoS 보증을 위한 아키텍처에 관한 연구가 소개된 바 있다[23].

또한, 분산 클라우드 환경에서 네트워크 서비스(NS) 장애 및 성능 문제는 가상 자원 간에 복잡한 연계요소로 인하여 기존의 규칙 기반 시스템으로 효과적인 처리가 불가하므로, 저수준 학습모델과 고수준 학습모델의 조합형 구조 기반의 네트워크 서비스 장애 및 품질 감지/예측 및 국부화(Localization) 방안에 관한 연구 등이 진행 중이다[24].

5G 네트워크 슬라이스에서 강화학습기반 QoS/QoE 인식 SFC에 관한 연구로서 SDN 컨트롤러의 하향 인터페이스에서 piggyback 방식으로 작동하여 QoS를 인식할 수 있는 경량 정보 수집기를 구현하여 DQN(Deep Q Network)기반 에이전트 프레임워크를 제시한 바 있다. DQN 에이전트는 QoS 제한 조건을 존중하면서 QoE를 최대화하기 위한 정책에 기반하여 동적 네트워크 환경에서 SFC에 대한 QoE 프로비저닝 및 QoS 요구사항 유지 관리를 위한 실험 결과를 도출한 바 있다[25].

IV. Network for AI 기술

이 장에서는 네트워크에서 어떻게 효율적으로 AI를 지원할 것인지를 관점에서 살펴본다. AI 기술이 보편화됨에 따라 컴퓨팅 자원(에지 컴퓨팅, 포그

컴퓨팅 등) 또는 디바이스 자체에 AI 엔진이 내장된 형태로 네트워크 전체에 널리 분산되어 적용되고 있다. 이러한 분산된 개별 AI는 자체 데이터를 기반으로 학습하고 추론하며 스스로 진화하게 될 것이다.

문제는 분산화되어 있는 AI의 자체적 학습 및 추론만으론 로컬 데이터가 갖는 정보의 지역적 특성, 컴퓨팅 파워의 제약 등에 의해서 총체적 지능화(Collective Intelligence)를 갖는 데 한계가 있다는 것이다. 이는 분산된 지능이 서로 협력하여 좀 더 강력한 Super Intelligence를 생성하는 분산 협력 AI 패러다임으로 전환을 고려할 수밖에 없는 이유가 된다.

네트워크 관점에서 분산 AI 협력을 달성하기 위한 핵심 기술 동향을 분산 AI의 협업을 조율하기 위한 분산 AI 오케스트레이션 기술과 분산 AI 네트워크 연결기술로 구분하여 소개한다.

1. 분산 AI 오케스트레이션 기술

향후의 네트워크는 소규모의 분산 클라우드 환경이 더욱 확산되고, 인공지능의 발전에 따른 다양한 AI 기반 지능형 서비스들이 분산된 네트워크에 탑재되어 기계 지능(Machine Intelligence) 간 상호연결 및 산업 간 융합 서비스가 확산될 것으로 전망되고 있다[26].

이러한 분산 지능 협업에 관한 연구로서 다계층 분산 클라우드 네트워크 환경에서 사용자의 개인 정보를 보호하고 다양한 자원의 효율적인 사용뿐만 아니라, 요구되는 서비스 품질, 위치, 오프로딩 수준 등을 통합적으로 고려한 분산 협업 오케스트레이션 기술에 대한 관심이 증대되고 있다.

특히 디바이스, 에지, 클라우드 간의 협업형 인공지능 모델의 필요성이 증대됨에 따라서 응용에 따른 중앙집중형, 분산형, 혼합형 구조와 같은 다양한 학습 및 추론 모델에 관한 연구뿐만 아니라,

분산 협업 학습 과정에서 원시 데이터 또는 중간 연산 결과를 주고받으면서 발생하는 비용을 최소화하기 위한 효율적인 방법 등에 관한 연구가 진행 중이다[27].

또한, 중앙집중식 인공지능(AI)을 학습하는 기존 방식과 달리, 분산 노드에서 직접 사용하는 데이터를 처리하고 모델을 강화해, 이 모델을 한곳에 모아 더 정교한 모델을 만들어 다시 배포하는 방식인 연합학습(Federated Learning)에 관한 다양한 연구가 진행 중이다.

이때 발생하는 통신 비용의 문제를 최소화하기 위해서 계층적 연합학습 구조에서 단계별로 근집화하는 방법을 적용 시 기존의 클라우드 기반의 연합학습 방식보다 속도가 빨라지고 학습 과정에서 소모되는 클라이언트의 에너지양과 통신 비용도 함께 줄어드는 효과가 있음을 실험을 통하여 보인 바 있다[28].

그 밖에 분산된 개별 인공지능 기반의 에이전트들이 상호 협력하여 단일 시스템으로 해결하기 어렵거나 불가능한 문제를 해결하는 멀티 에이전트 방법, 추론 대기시간과 에너지 소모량을 개선하기 위해, DNN 모델을 레이어 단위로 분할하여 추론을 수행하는 단말 디바이스와 에지 노드 간 분산 추론 방법, DNN 모델의 추론 대기시간을 개선하기 위해, 일부 레이어까지만 연산을 하여, 목적 클래스에 대한 확률 벡터를 기반으로 신뢰도를 계산하고, 이를 임계치 값과 비교하여 분류 결과 반환 여부를 결정하는 로컬-반환 방법 등의 AI 응용을 위한 분산 협업 오케스트레이션 기술과 관련한 연구가 활발히 진행되는 중이다[29].

2. 분산 AI 네트워크 연결 기술

분산된 AI의 네트워크 연결 기술은 AI 엔진의 학

습에 필요한 원격 데이터와 AI 엔진이 분산 협력을 위해 서로 주고받는 AI 학습 모델의 안전한 공유·교환을 제공하기 위한 기술을 의미한다. 2020년 1월 9일 국회 본회의를 통과한 데이터 3법(개인정보 보호법·정보통신망법·신용정보법) 개정안 시행 등으로 앞으로 좀 더 활발한 데이터 공유가 가능해질 전망이다[30]. 다른 데이터를 보유한 회사들이 다양한 형태로 협업이 이루어질 수 있게 되었다. 예를 들어, AI 기술 전문 업체가 개인 데이터 또는 일반 기업의 데이터를 이용해 AI 서비스를 제공할 수 있게 되었다.

네트워크의 연결과 가장 쉽게 비교가 되는 현실 세계의 개념은 도로이다. 과학기술정보통신부 역시 최근 우리나라를 AI 강국으로 만들고자 하는 정책의 하나로 데이터 경제 활성화를 위한 일명 “데이터 고속도로” 건설을 추진하고 있다[31]. 그 실현 기술로는 많은 데이터를 빠르게 전달할 수 있는 5G가 핵심으로 이야기되고 있다. 많은 데이터를 잘 전달할 수 있는 넓은 대역의 도로가 필요하다는 측면에서 중요하게 고려되어야 하는 요소이다.

얼마나 데이터를 잘 전달하느냐 역시 넓은 대역만큼이나 중요하게 고려되어야 한다. 데이터를 실어 나르는데 목적지에 대한 정확한 정보나 내비게이션 없이 무작정 운전하는 “구식 화물차”가 아닌 아주 똑똑한 내비게이션을 장착하고, 보안도 잘 갖추고, 필요한 곳에 필요한 시간에 잘 전달을 할 수 있는 “스마트 화물차”가 제공되어야 한다.

연결될 수 있는 모든 것이 연결되는 초연결 네트워크의 복잡 다변화되는 환경에서의 분산 AI 엔진 간 그리고 분산 AI 엔진을 위한 데이터의 안전하고 효율적인 데이터 연결제공을 위해 다음과 같은 특성을 지원하는 스마트한 네트워킹 기술이 고려되어야 한다.

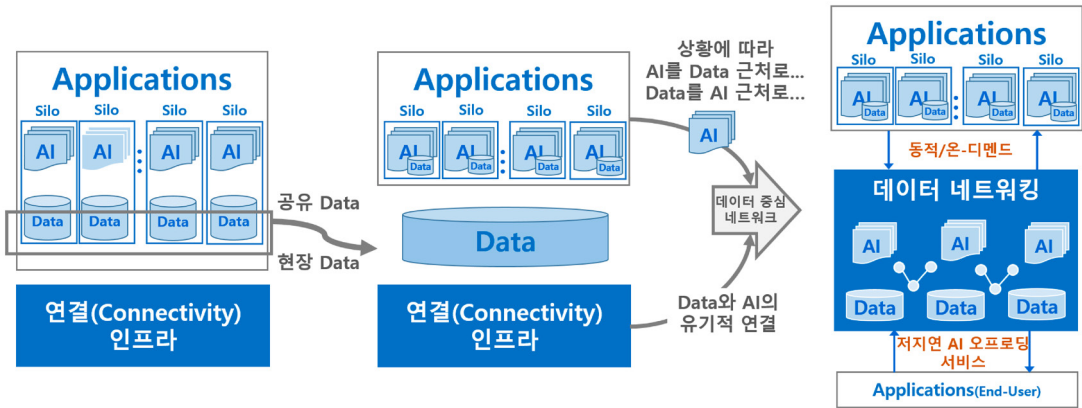


그림 5 분산 AI 네트워크 연결 기술 개념도

- (안전성) AI 학습 데이터 및 모델의 출처 (provenance) 및 불변성(immutability) 보장을 통한 AI 학습 신뢰성 제고
- (효율성) 다수의 AI 엔진 간 또는 AI 엔진과 AI 학습 데이터 및 모델 간 복잡한 연결 방식의 효율화 및 단순화

그림 5는 분산 AI 네트워크 연결을 위한 비전 및 개념을 보여주고 있다. 기존 AI 응용별로 사일로(Silo)화되어 공유되지 못하는 데이터를 안전하고 효율적으로 공유할 수 있는 환경을 제공하여, AI 엔진은 동적 필요에 따라 데이터에 대한 위치에 대한 별도의 고민 없이 데이터를 활용할 수 있다. 이러한 개념은 좀 더 확장되어 데이터가 있는 가까운 곳에서 AI 엔진이 필요한 컴퓨팅 기능이 수행되도록 할 수 있다. 네트워크 내에서 동적으로 컴퓨팅이 이루어지는 것이다.

우리는 스마트한 분산 AI 네트워크 연결 기술로서 미래 인터넷의 하나로 연구되어 온 『정보(데이터) 중심 네트워킹(ICN: Information-Centric Networking)』 기술에 주목하고 있다[32,33]. Palo Alto PARC에서 시작되어 미국, 유럽, 일본, 중국 등 세계 각국에서 꾸준한 연구가 지속되었다. 인터넷이

용 패턴 자체가 점차 웹을 기반으로 하는 콘텐츠/데이터 수요기반으로 바뀌에 따라 데이터 전달 및 공유에 효율적이며 데이터 자체에 보안이 내재된 새로운 구조로 제안된 것이다[34]. 이러한 개념은 NFN(Named-Function Networking), CFN(Compute First Networking) 등과 같이 데이터 자체가 아닌 데이터를 기반으로 하는 컴퓨팅 처리에 대한 요청도 가능하도록 확장되어 진행 중이다[35].

AI 학습 데이터(또는 학습 모델)의 전달에도 데이터 중심 네트워킹 기술을 사용하게 되면 AI 데이터 각각에 이름을 부여하고 그 데이터를 만들어 내는 생성자가 시그니처를 붙여 누구나 쉽게 누가 만들었는지 그리고 그 데이터가 위변조되지 않았는지의 여부를 검증하여 AI 데이터의 전달과정의 안전성을 보장한다. 또한, AI 데이터가 현재 어디에 있는지에 대한 정보 없이 그리고 물론 보유하고 있는 시스템 또는 저장소로의 별도의 연결설정도 없이 AI 데이터의 이름을 기반으로 AI 데이터를 획득할 수 있다.

AI 데이터에 대한 요청이 있는 경우에만 그 응답으로 요청이 전달된 경로의 역(逆)으로 AI 데이터를 돌려보내 주는 풀(pull) 모드 방식의 전달 방식을 사용하여 동일한 AI 데이터에 대한 요청이 여러 번

오더라도 AI 데이터 생성자에게는 한 번만 전달되도록 하여 AI 데이터 생성자에 대한 DDoS(Distributed Denial of Service) 공격이 원천적으로 차단되며, 네트워크 노드에서의 AI 데이터 캐싱으로 AI 데이터의 중복 전송을 방지하도록 하여 분산 AI 네트워크 전송으로 인한 네트워크 트래픽 폭증을 방지할 수 있다.

연결설정 절차가 필요 없어 다중 인터페이스를 갖는 AI 엔진 시스템은 필요에 따라서 자유롭게 다중 인터페이스 사용을 사용할 수 있으며, AI 데이터 생산자나 AI 엔진이 이동하더라도 쉽게 이동성 보장 지원이 가능한 장점이 있다[36]. 완전히 새로운 개념의 네트워크 구조임에도 링크 계층으로서 기존 TCP/IP 계층을 포함함으로써 IP 네트워크를 대체할 필요 없이 데이터 중심 AI 네트워크의 실현이 가능하다.

V. 결론 및 시사점

본 고에서는 네트워크와 AI 기술의 진화발전에 있어서, 네트워크 혁신을 위한 AI 기술(AI for Network)과 AI 기술의 발전을 위한 네트워크 기술(Network for AI)의 두 가지 관점에서 연구동향과 전망을 살펴보았다.

최근의 5G Connected VI 확산 등 서비스 및 기술 환경의 급속한 변화에 대응하기 위하여 인프라는 ‘Biz/Tech Agile Autonomous Network’로 진화하고 있다. 이러한 네트워크의 진화에 있어서 Agile Infra로서의 복잡도 증가, 플랫폼화에 따른 관리 자동화, 5G VI 수용을 위한 품질 최적화 등 여러 가지 기술적 현안들을 해결하기 위해 AI를 활용하는 연구가 진행 중이다. 특히, 최근에 가장 큰 화두가 되는 기술은 AI 기반 운영관리의 최적화 및 자율화 기술이며, 이를 Ⅲ장에서 고찰하였다. 가상기능과 가

상자원 최적 배치관리, 네트워크 슬라이싱 자동화, 네트워크 서비스 품질 최적화 이슈는 지도학습, 비지도학습, 세미지도학습, 강화학습 등의 다양한 형태의 학습모델이 연구 적용되고 있다. 현재는 주로 텔코 입장에서 접근하고 있으나, 향후 지자체나 공공기관, 민간기업과 같은 엔터프라이즈급 자가망의 운영관리 자동화·자율화 기술로 확산될 수 있을 것으로 기대된다.

한편, AI 기술 측면에서는 다양한 영역으로의 활용 확산과 AI 기반 기술의 보편화에 힘입어 AI 엔진이 네트워크 전체에 널리 분산되고, 그들 간에 서로 협력함으로써 자원의 제약을 극복하면서, 더 막강한 능력을 갖추는 분산 협력 AI 패러다임으로 발전할 것으로 전망되고 있다. 이러한 분산 협력 AI를 위한 핵심기술로서 분산 AI 간의 협업을 조율하는 분산 AI 오케스트레이션 기술과 분산 AI 네트워크 연결기술에 대해 Ⅳ장에서 고찰하였다.

결론적으로, 네트워크 기술과 AI 기술은 당분간 독자적인 발전을 계속하게 될 것이나, 초연결 지능화 사회가 될수록 AI 서비스와 네트워크 서비스의 경계가 모호해지게 될 것이다. 즉, AI 서비스의 네트워크를 통한 확산이 크게 대두되는 시점에는 두 가지 기술이 밀결합되어 ‘AI 응용서비스를 위한 자율 네트워킹 기술’로 진화될 것으로 전망된다. 아울러, 앞으로 네트워크 도메인에서 AI 기술의 활용은 계속 늘어날 것이다. 이를 위해서는 각 네트워킹 이슈에 관한 개별적 연구뿐만 아니라 기반 기술로서 네트워크 도메인에 특화된 데이터 레포지토리와 네트워크 AI 플랫폼 기술 확보가 필요할 것이다.

약어 정리

AI	Artificial Intelligence
BAAN	Biz/Tech-Agile Autonomous

	Network
CFN	Compute First Networking
DDoS	Distributed Denial of Service
DNN	Deep Neural Networking
DML	Distributed ML
DRL	Deep Reinforcement Learning
DQN	Deep Q Network
eMBB	enhanced Mobile BroadBand
FG	Function Graph
ICN	Information-Centric Networking
KPI	Key Performance Indicator
ML	Machine Learning
MTC	Massive Machine Type Communication
NFN	Named-Function Networking
NFV	Network Function Virtualization
NS	Network Service
ONAP	Open Network Automation Platform
OPNFV	Open Platform for NFV
QCI	QoS Class Identifier
QoE	Quality of Experience
QoS	Quality of Service
RL	Reinforcement Learning
SDN	Software Defined Networking
SFC	Service Function Chaining
SLA	Service Level Agreement
uRLLC	ultra Reliable Low Latency Communication
VFG	VNF Forwarding Graph
VI	Vertical Industry
VNF	Virtualized Network Function
VNFI	VNF Instance

참고문헌

- [1] Mc Carty & Minsky, "Artificial intelligence: Allow machine to accomplish task human execute with intelligence," https://atos.net/wp-content/uploads/2018/07/IA-infography_mini.pdf
- [2] Northstar, "The Arm 2020 Global AI Survey: AI Today, AI Tomorrow," <https://www.arm.com/resources/report/ai-today-ai-tomorrow>
- [3] Stanford University, AI100 Project 보고서, "ARTIFICIAL INTELLIGENCE AND LIFE IN 2030," <https://ai100.stanford.edu>.
- [4] Satoru Kobayashi et al., "Causal analysis of network logs with layered protocols and topology knowledge," CNSM 2019.
- [5] AWAKE, White Paper, "The Advent of Advanced Network Traffic Analysis and Why it Matters," <https://awakesecurity.com/white-papers/wp-the-advent-of-advanced-network-traffic-analysis-why-it-matters-white-paper/>
- [6] Ali Hussein et al., "Machine Learning for Network Resilience: The Start of a Journey," 2018 Fifth International Conference on Software Defined Systems(SDS).
- [7] Cisco, "AI and Machine Learning: A White Paper for Technical Decision Makers," <https://www.cisco.com/c/en/us/solutions/collateral/enterprise-networks/nb-06-ai-nw-analytics-wp-cte-en.html>
- [8] Nashid Shahriar et al., "Reliable Slicing of 5G Transport Networks with Dedicated Protection," CNSM, 2019.
- [9] TM Forum, "A white paper of Autonomous Networks: Empowering Digital Transformation For e Telecoms Industry," Rel 1.0, 2020.
- [10] GSMA Future Networks, "AI & Automation," <https://www.gsma.com/futurenetworks/wiki/ai-automation-an-overview/>
- [11] Javier Mata et al., "Artificial intelligence (AI) methods in optical networks: A comprehensive survey," Optical Switching and Networking Volume 28, April 2018, pp. 43-57.
- [12] The Linux Foundation Projects ONAP Homepage, <https://www.onap.org/>
- [13] The Linux Foundation Projects AcumosAI Homepage, <https://www.Acumos.org/>
- [14] 신명기 외, "5G 네트워크 자동화 및 지능 기술 표준화 동향," 전자통신동향분석, 제34권 제2호 2019년 4월.
- [15] Manabu Nakanoya et al., "Environment Adaptive Sizing and Placement of NFV Service Chains with Accelerated Reinforcement Learning," IFIP/IEEE Symposim on INM, 2019.
- [16] Ruben Solozabal et al., "Virtual Network Function Placement Optimization with Deep Reinforcement Learning," IEEE Journal of Selected Areas in Comm., 2020.

- [17] Jianing Pei et al., "Optimal VNF Placement via Deep Reinforcement Learning in SDN/NFV-Enabled Networks," *IEEE Journal of Selected Areas in Comm.*, 2020.
- [18] Jiaqi Zheng et al., "Optimizing NFV Chain Deployment in Software-Defined Cellular Core," *IEEE Journal of Selected Areas in Comm.*, 2020.
- [19] Anurag Thantharate et al., "DeepSlice: A Deep Learning Approach towards an Efficient and Reliable Network Slicing in 5G Networks," *IEEE UEMCON*, 2019.
- [20] L. Ma et al., "An SDN/NFV based framework for management and deployment of service based 5G core network," in *China Communications*, vol. 15, no. 10, Oct. 2018, pp. 86-98.
- [21] F. Kurtz et al., "Network Slicing for Critical Communications in Shared 5G Infrastructures-An Empirical Evaluation," *NetSoft*, 2018, pp. 393-399.
- [22] S. O. Oladejo et al., "5G network slicing: A multi-tenancy scenario," *Global Wireless Summit (GWS)*, 2017, pp. 88-92.
- [23] Guosheng Zhu et al., "A Supervised Learning Based QoS Assurance Architecture for 5G Networks," *IEEE Access*, Vol 7, 2019.
- [24] Lav Gupta et al., "Fault and Performance Management in Multi-Cloud Based NFV using Shallow and Deep Predictive Structures," *Journal of Reliable Intelligent Environments*, Vol. 3, No. 4, Dec. 2017, pp. 221-231.
- [25] Xi Chen1 et al., "Reinforcement Learning based QoS/QoE-aware Service Function Chaining in Software-Driven 5G Slices," *Transactions on Emerging Telecommunications Technologies*, April 2018.
- [26] Noah Schwartz, "The future of AI is distributed: Peer-to-peer learning and multi-agent AI at the edge," *AI Conference*, Sep. 2018.
- [27] Zhi Zhou et al., *Edge Intelligence: Paving the Last Mile of Artificial Intelligence with Edge Computing*, IEEE, 2019.
- [28] L. Liu, J. Zhang, S. H. Song, and K. B. Letaief, "Client-Edge-Cloud Hierarchical Federated Learning," *arXiv:1905.06641v2*, Oct. 2019.
- [29] S. Ponomarev et al., "Multi-agent systems and decentralized artificial superintelligence," *arXiv2017*, Feb. 2017.
- [30] <http://www.korea.kr/special/policyCurationList.do>
- [31] <https://www.gov.kr/portal/ntnadmNews/2208623>
- [32] <https://irtf.org/icnrg>
- [33] <https://named-data.net>
- [34] Zhiyi Zhang, Yingdi Yu et al., "An Overview of Security Support in Named Data Networking," *NDN community*, April. 2018, <https://named-data.net/wp-content/uploads/2018/04/ndn-0057-2-ndn-security.pdf>
- [35] Michał Król et al., "Compute First Networking: Distributed Computing meets ICN," *ACM ICN*, 2019.
- [36] Yue Zhang et al., "A Survey of Mobility Support in Named Data Networking," *the third Workshop on Name-Oriented Mobility: Architecture, Algorithms and Applications*, 2016.