글로벌 ICT 표준 컨퍼런스 2022

Global ICT Standards Conference 2022

2022. 11.9.(수)~11.(금)
서울 양재 엘타워 오르체홀(5F)

Session 5 : AI and autonomous technology

AI/ML 기반 3GPP 5G 네트워크 지능화 표준 기술

고한얼, 조교수, 경희대학교

INDEX

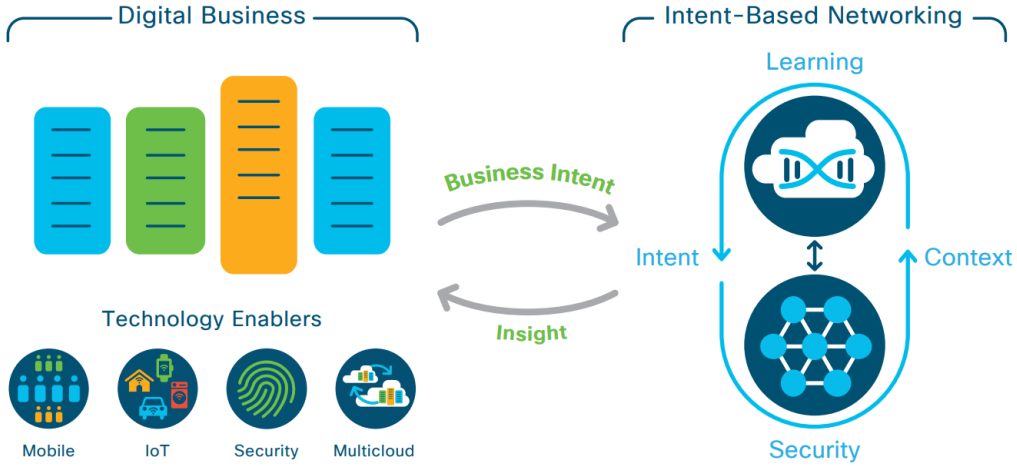
ABOUT PRESENTATION

ICT Standards

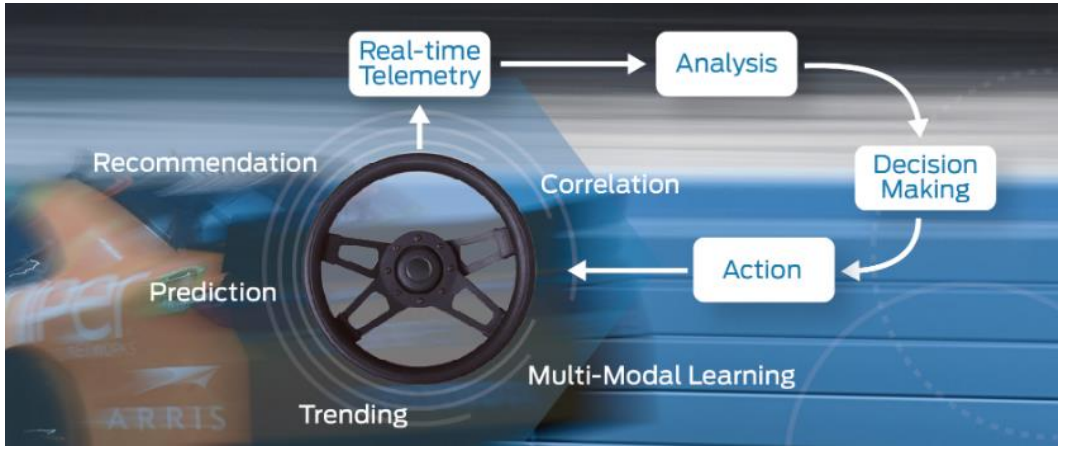
본 발표에서는 AI를 네트워크에 효과적으로 적용하기 위해 어떤 표준 기술들이 연구되고 있는지 살펴본다. 구체적으로 3GPP SA2 그룹에서 표준화가 진행되고 있는 NWDAF와 AI/ML 서비스를 위한 시스템에 관련된 내용을 살펴본다.

Network Automation

- **Cisco – Intent Based Network**
  - Intent Based Network (IBN) **will understand business intent and continuously align to it**



- **Juniper – Self-Driving Network**
  - An autonomous network that is predictive and adaptive to its environment



3GPP Net4AI Milestone

	Rel-15	Rel-16	Rel-17	Rel-18	
SA 2	Introduce NWDAF	eNA	eNA Phase 2	eNA Phase 3	TR 23.700-81
				AIMLSys	TR 23.700-80

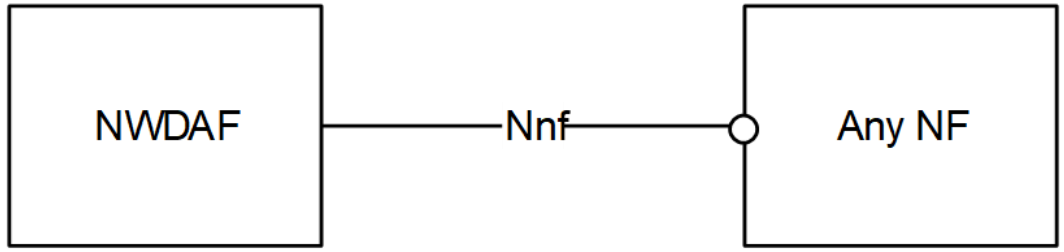
3GPP TR 23.700-81

STUDY OF ENABLERS FOR NETWORK AUTOMATION FOR 5G

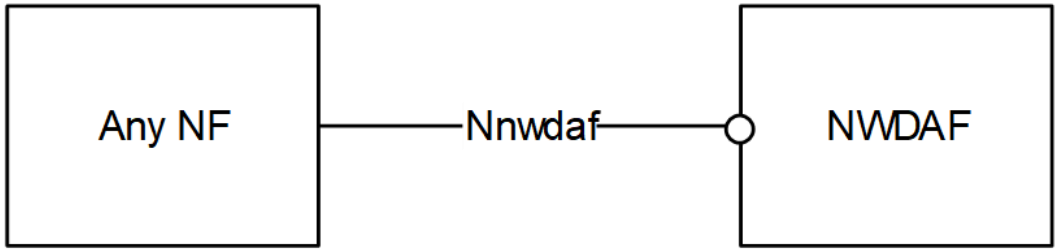
- **For further enhancement on enablers (i.e., NWDAF) for network automation**

Enablers for Network Automation

- **Architecture enhancements for 5G System (5GS) to support network data analytics services [TS23.288]**
  - Study and specify how to collect data and how to feedback data analytics to the network functions
  - **Network Data Analytics Function (NWDAF)**



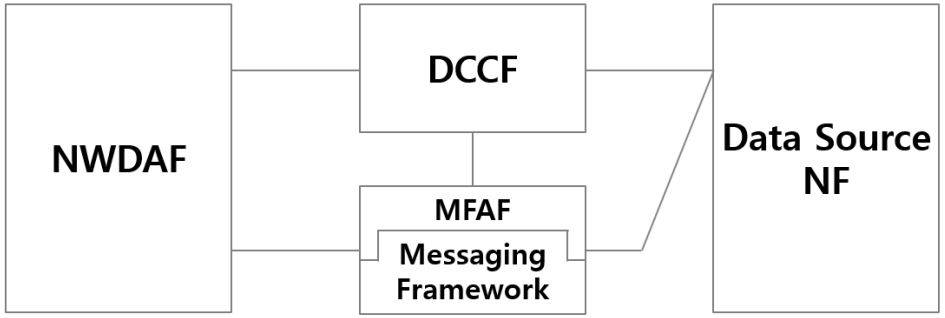
Data Collection architecture from any 5GC NF



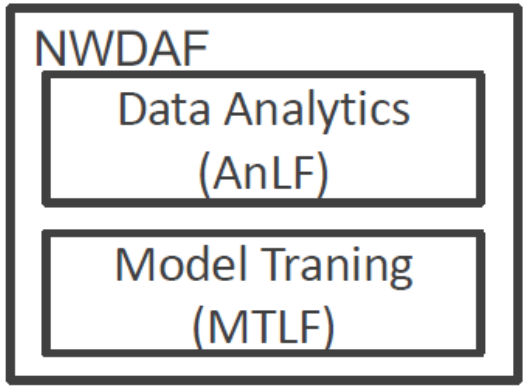
Network Data Analytics Exposure architecture

Some features of NWDAF

✓DCCF/MFAF



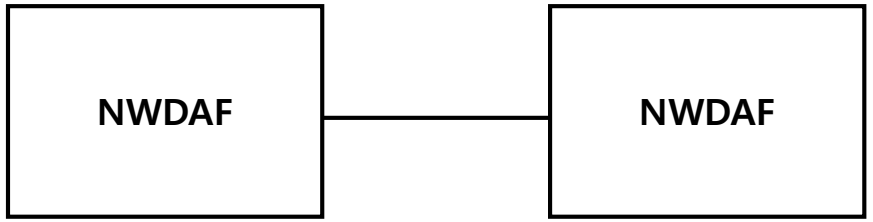
✓MTLF/AnLF



✓Cooperation of distributed NWDAF

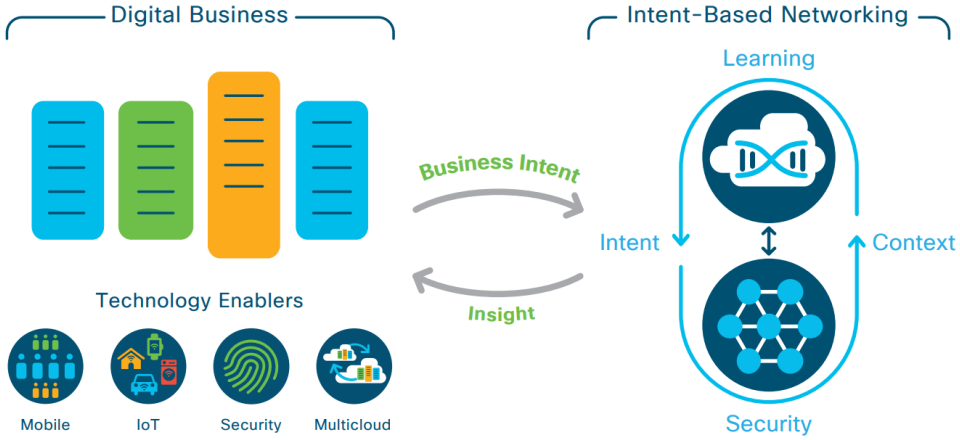
- Aggregation of analytics
- Transfer analytics

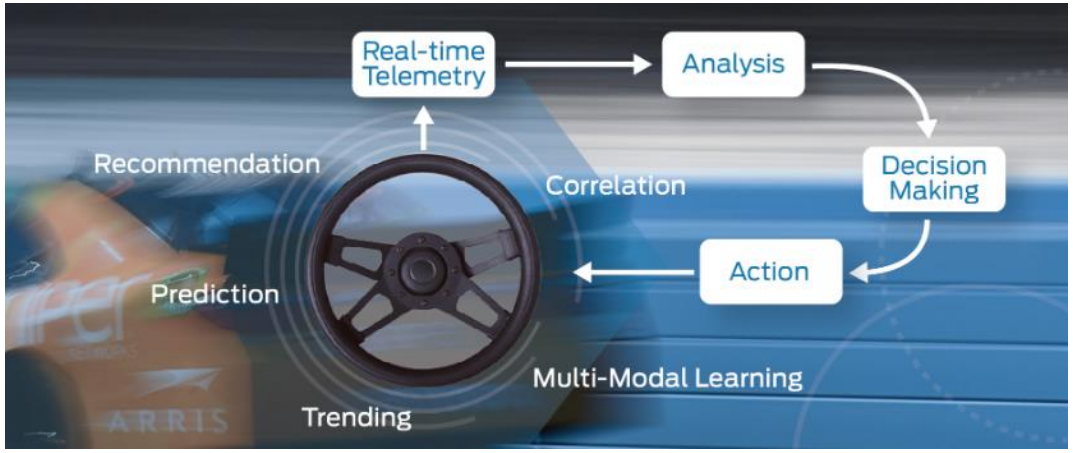
+ Federated learning (in TR phase)



Key Issue #1

- **How to improve correctness of NWDAF analytics**
- **Aim to study architecture enhancement to enhance the correctness of NWDAF analytics**
  - 3GPP tries to adopt the "closed loop control"





Key Issue #5

- **Enhance trained ML Model sharing**
  - 3GPP tries to share trained **ML Model sharing between NWDAFs from different vendors**

Key Issue #8

- **Supporting Federated Learning in 5GC**
  - 3GPP tries to adopt Federated Learning technique in NWDAF containing MTLF to train an ML model

- **Aim to study architecture enhancement to support Federated Learning which allows the cooperation of multiple NWDAF containing MTLF**

3GPP TR 23.700-80

STUDY ON 5G SYSTEM SUPPORT FOR AI/ML-BASED SERVICE

Scope of TR23.700-80

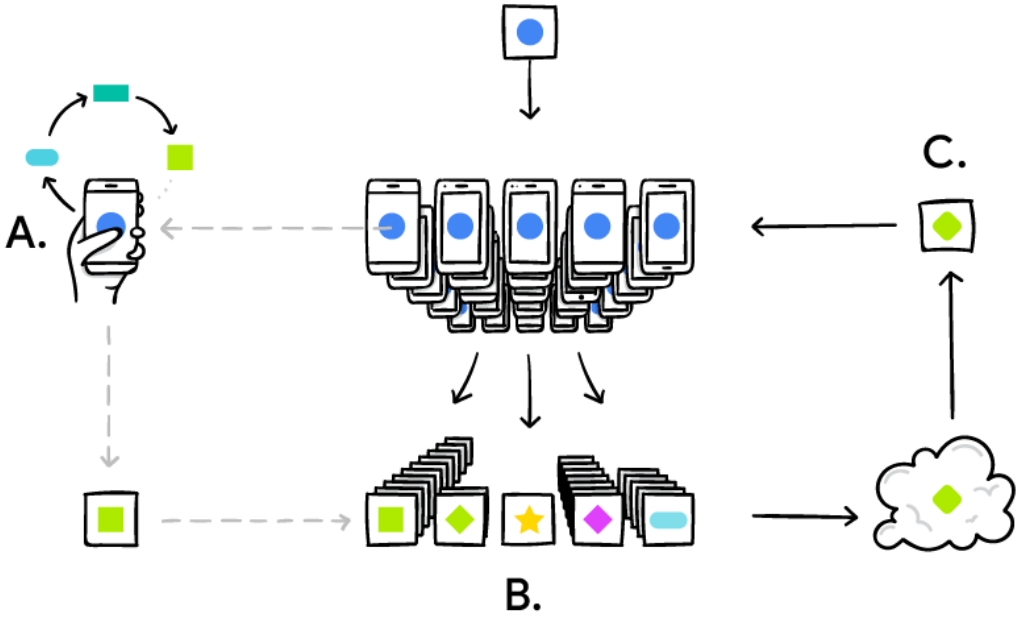
- **How to leverage 5GS as the platform to provide the intelligent transmission support for AI/ML operation (studied in TR22.874)**

- **AI/ML operation**
  - **Distributed/Federated Learning**
  - **Split computing**
  - AI/ML model distribution and sharing

- **To study the possible architecture and functional extensions to support AI/ML operations**

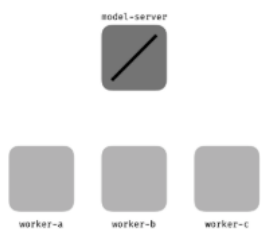
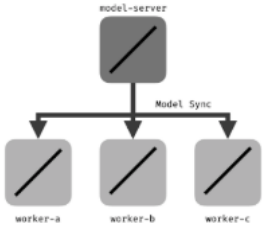
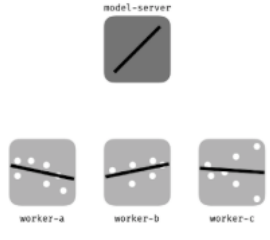
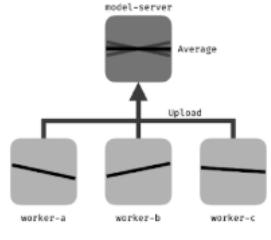
Federated Learning

- **Distributed learning method**
  - Network (e.g., edge cloud)
  - Selected mobile devices



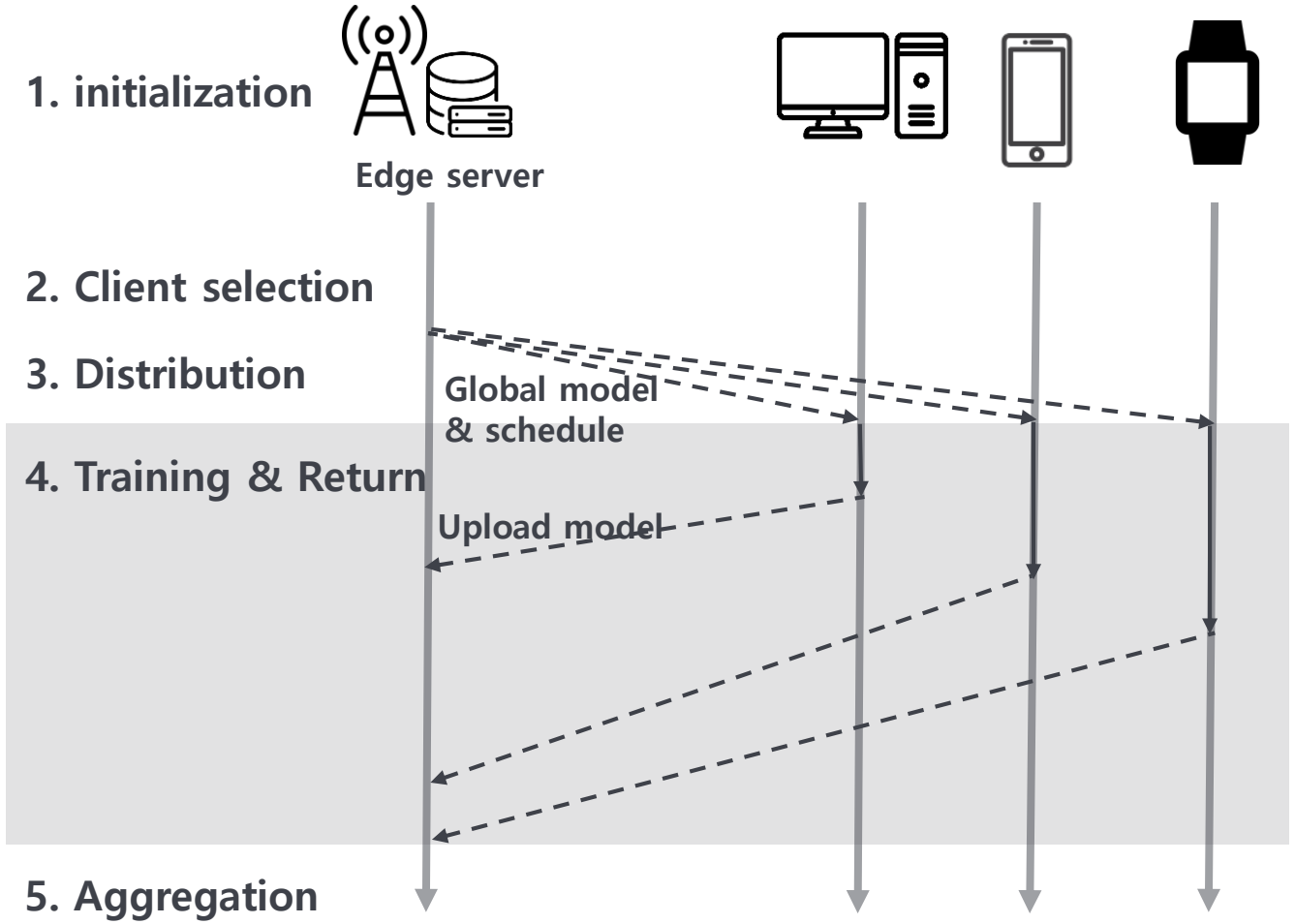
Advantage

- **Data privacy enhancement**
- **Low communication cost**

Step 1	Step 2	Step 3	Step 4
			
Central server chooses a statistical model to be trained	Central server transmits the initial model to several nodes	Nodes train the model locally with their own data	Central server pools model results and generate one global mode without accessing any data

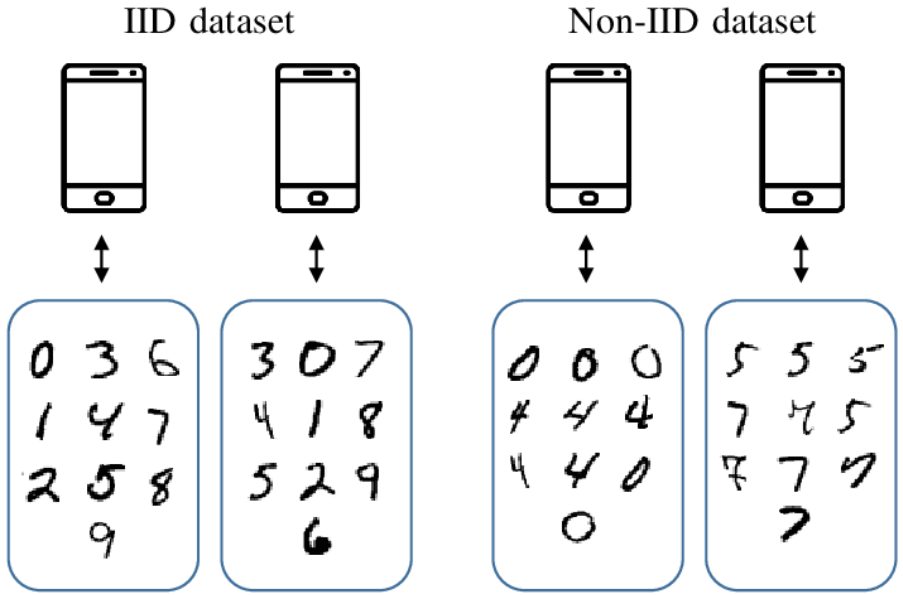
Issues on Federated Learning (1/2)

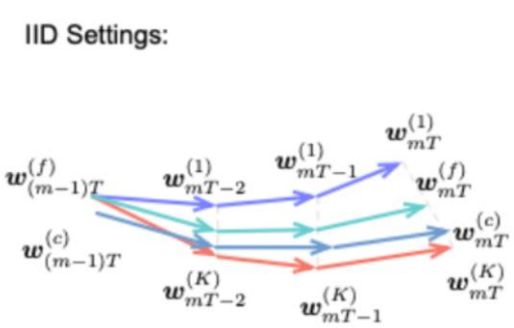
- **Straggler issue**
  - Mobile devices do not have the same computational power and network condition
  - Learning can be slower

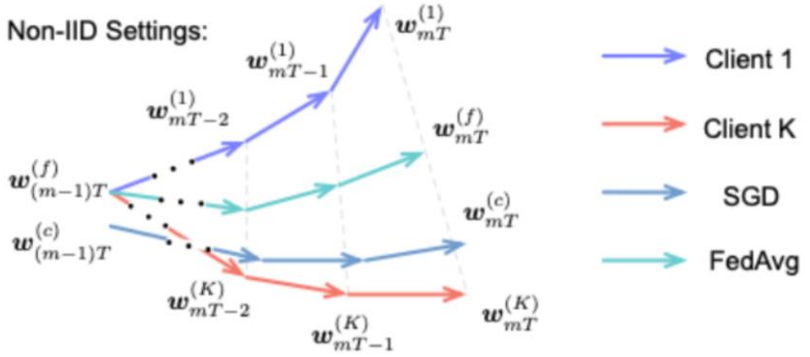


Issues on Federated Learning (2/2)

- **Data distribution issue**
  - Data from different sources will not be representative of the global distribution (i.e., non-iid)
  - Cannot easily be converged







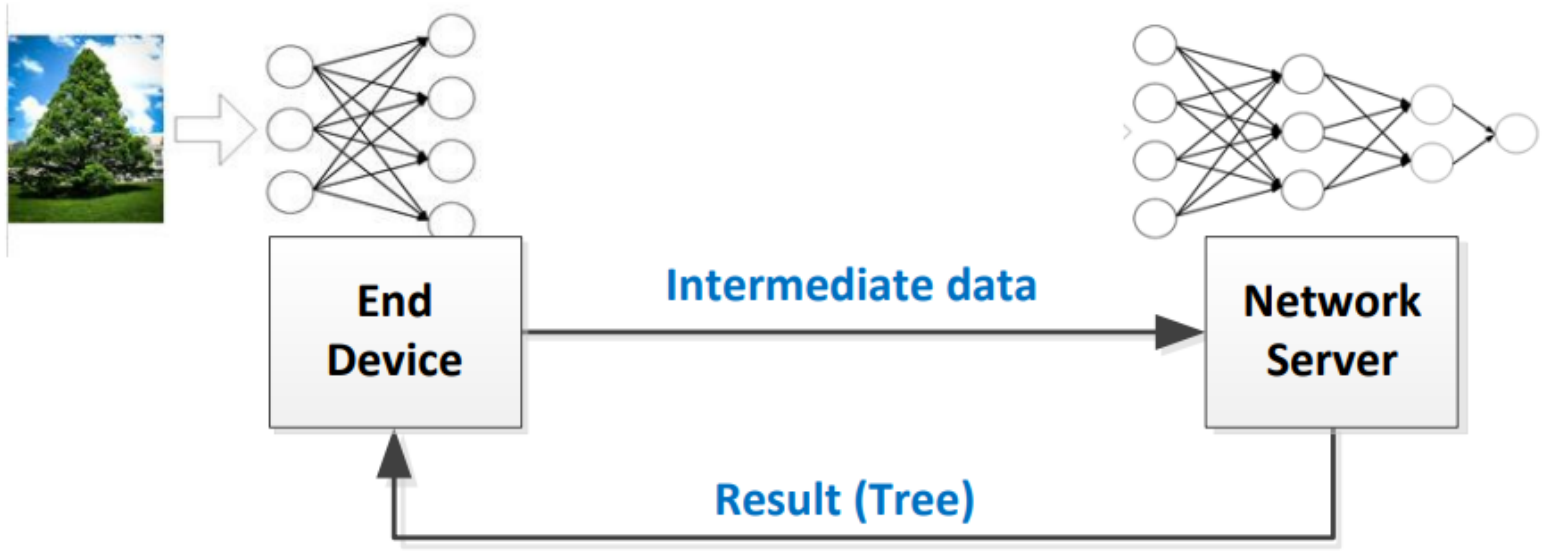
Issues on Split Computing (1/2)

- **How to obtain the result within a desired deadline**
  - Basically, we have to decide the splitting point by considering the status of opponent entity

- **Even with the optimal splitting point, we cannot obtain the result within the deadline**
  - Due to the network congestion
  - Due to the overload at a specific entity (e.g., edge cloud)

Issues on Split Computing (2/2)

- **Asymmetric traffic characteristic (e.g., size)**



Key Issue #1

- **Monitoring network resource utilization for AI/ML operation (e.g., federated learning and split computing)**
  - How to monitor network resource utilization relevant to UE's performance (e.g., data throughput provided to UE)
  - How to capture and/or predict resource utilization relevant to UE's performance

Key Issue #2, 3

- **5GC information exposure to UE and authorized 3rd party**
  - Whether and how to extend 5GC to expose information to UE and authorized 3rd party

- **UP-based solution & CP-based solution**

Key Issue #5

- **5GC Enhancements to enable AI/ML traffic transport**
  - Different levels of interactions are expected between UE and AF to exchange AI/ML traffic
  - AI/ML traffic: AI/ML model, intermediate data, inference results, and model performance

- **Whether and how the existing 5GC data transfer/traffic routing mechanisms are re-used or enhanced to support the transmission of AI/ML traffic**

Conclusion

- For the network automation, NWDAF is defined and upgraded
- For efficiently supporting AI/ML service, 5G system should be upgraded
- By focusing on the issues on AI/ML service (i.e., FL and split computing), we can derive the intuition how to update 5G system

Thank you

고한얼, 조교수, 경희대학교
heko@khu.ac.kr