

효율적인 NWDAF 운영을 위한 멀티 태스크 러닝 활용 기법

정현재¹⁾, 정대영¹⁾, 고한얼²⁾, 백상현¹⁾
고려대학교¹⁾, 경희대학교²⁾

qege@korea.ac.kr¹⁾, eodud1884@korea.ac.kr¹⁾, heko@khu.ac.kr²⁾, shpack@korea.ac.kr¹⁾

Efficient NWDAF Operation using Multi-task Learning

Hyeonjae Jeong¹⁾, Daeyoung Jung¹⁾, Haneul Ko²⁾, Sangheon Pack¹⁾
Korea University¹⁾, Kyunghee University²⁾

요 약

본 연구는 코어 네트워크 기능인 network data analysis function (NWDAF)의 효율적인 운영을 위해 NWDAF에서의 machine learning (ML) 모델 프로비저닝 프레임워크를 제안한다. 해당 프레임워크는 코어 도메인 내의 컴퓨팅 비용과 네트워크 기능 (consumer network function, NF)의 지연 시간을 고려하여 추론 모델을 선택한다. 이는 NF가 요구하는 지연 시간을 만족하면서, NWDAF에서의 컴퓨팅 비용을 감소시킬 수 있도록 한다. 본 연구에서는 효율적인 NWDAF 운영을 위해 멀티 태스크 러닝 모델을 활용한 기법에 대하여 살펴본다.

I. 서 론

최근 AI/ML 기법이 고도화되고, beyond 5G (B5G) 네트워크 자동화에 대한 관심이 높아짐과 동시에 네트워크 자동화를 위한 코어 네트워크에서의 빅데이터 분석 기능인 network data analysis function (NWDAF)에 대한 사용 사례 작성 및 개발이 이루어지고 있다. 본 논문에서는 NWDAF에 대한 표준과 연구 동향을 사용 사례 중심으로 소개하고, 효율적인 NWDAF 운영 방안을 제시한다.

II. NWDAF 표준 소개

NWDAF는 3GPP TS 29.520 [1]에서 처음 규격화된 코어 네트워크 기능이고, 네트워크 슬라이스 인스턴스와 슬라이스 부하 레벨 정보 분석을 위하여 정의되었다. 이후 Release-16에서 네트워크 자동화를 위해 모바일 단말의 이동성이나 네트워크 슬라이스 부하 레벨 등을 분석하는 NWDAF의 여러 사용 사례를 추가하였다 [2]. NWDAF는 크게 두 가지 기능으로 나뉜다. 첫째는 model training logical function (MTLF)이고, 둘째는 analytics logical function (AnLF)이다. 두 기능은 각각 학습과 추론 기능을 수행하는데, NWDAF에 분석 이벤트가 발생하였을 때 모델을 새로 학습해야 할 경우에는 먼저 MTLF가 학습한 모델을 AnLF에 전달하여 추론하는 과정을 거친다. 이미 학습된 모델이 있다면 MTLF는 AnLF에 모델을 전달하고, AnLF는 이를 통해 추론을 수행한다. NWDAF는 이처럼 학습과 추론을 분리하여 사용 사례에 대해 유연하게 분석 결과를 도출할 수 있도록 하였다. Access management function (AMF), session management function (SMF), policy control function (PCF) 등의 NF들은 NWDAF 분석 결과 소비자이자, NWDAF에 데이터를 제공하는

제공자가 된다. 즉, NF는 NWDAF에 데이터를 전송하고, NWDAF는 해당 데이터를 기반으로 분석을 수행한다. 각 NF는 그림 1, 2와 같이 NWDAF에 분석을 요청하거나 NWDAF를 구독하는 방식을 통해 NWDAF에서의 분석 결과를 제공받을 수 있다.

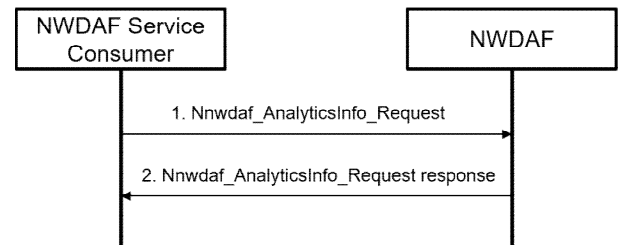


그림 1. NWDAF에 요청을 통한 분석

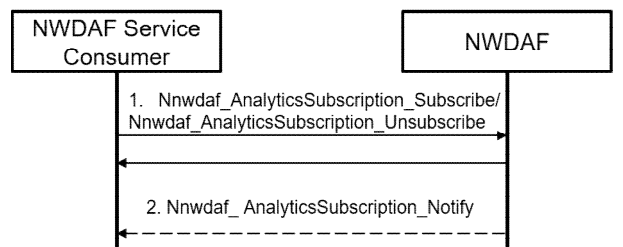


그림 2. NWDAF 구독을 통한 분석

NWDAF는 앞서 언급한 표준을 바탕으로 분석 결과를 제공하는데, 모든 분석 요청에 대해 일반적인 deep neural network (DNN)을 기반으로 추론을 수행하게 되면 NWDAF에서의 컴퓨팅 비용이 높아질 수 있다. 본 연구에서는 이를 해결하기 위해 NF 요구 사항에 따라 DNN 모델과 multi-task learning (MTL) 모델 중 추론 모델을 선택 사용하는 방법을 제안한다. MTL이란, 서로 연관 있는 과제들을 동시에 학습함으로써 모든 분석

수행의 성능을 전반적으로 향상시킬 수 있는 학습 모델 및 학습 방법이다. DNN 의 경우 수많은 계층과 파라미터들을 학습시키기 위해서는 방대한 양의 라벨링된 데이터가 필요하다. 반면 MTL 은 하나의 손실 함수로 두 개 이상의 과제 간 공통점 및 차이점을 구분하고, output 으로 여러 분석 결과를 도출하기 때문에 하나의 과제에 대해 하나의 DNN 으로 추론하는 것보다 효율적인 컴퓨팅 방안을 제공할 수 있다. 따라서 NWDAF 에서 추론 모델 프로비저닝 시 DNN 과 MTL 모델 중 선택하여 사용하는 방법을 통해 NF 의 요구 지연 시간은 만족하면서, 컴퓨팅 비용을 절감할 수 있다.

III. Multi-task Learning 모델을 활용한 NWDAF

본 절에서는 NWDAF 에서 DNN 과 MTL 모델을 사용한 경우의 장단점과 효율적인 모델 프로비저닝 방법에 대해 알아본다. NWDAF 는 NF 데이터로 학습된 DNN 혹은 MTL 을 저장하고 있는 것으로 가정한다. 학습된 모델이 없는 경우, 본문에서 언급한 것처럼 NWDAF 내 MTLF 에서 먼저 학습 후 학습 모델을 AnLF 에 전달하는 과정을 거친다.

표 1. 추론 모델에 따른 장단점 비교

모델	MTL	DNN
장점	낮은 컴퓨팅 비용	짧은 대기 시간
단점	긴 대기 시간	높은 컴퓨팅 비용

먼저 MTL 모델을 사용하는 경우에 대해 소개한다. NWDAF 는 요구 지연 시간이 긴 분석 요청이 발생하면, 다른 분석 요청이 발생하길 기다린다. 일정 대기 시간이 지나면 MTL 모델을 사용하여 여러 분석을 한 번에 추론하여 컴퓨팅 비용을 낮추면서, 요구 지연 시간은 만족할 수 있다. 반면에 NWDAF 가 요구 지연 시간이 상대적으로 짧은 분석 요청에 대해서는 긴 대기 시간을 가지지 않고 바로 DNN 으로 추론을 수행하여 NF 의 요구 지연 시간을 만족할 수 있도록 한다. 따라서 모든 분석 요청에 대해 MTL 모델을 사용하게 되면 긴 대기 시간을 가지지만 낮은 컴퓨팅 비용을 필요로 하고, DNN 모델을 사용하게 되면 요청 당 추론을 수행하기 때문에 높은 컴퓨팅 비용을 필요로 하나, 짧은 대기 시간을 가진다. 표 1 은 이러한 추론 모델에 따른 장단점을 비교한 표이다.

NWDAF 에서의 효율적인 모델 프로비저닝을 위해서는 어떤 요청들을 MTL 로 한 번에 추론할 수 있는지 먼저 정의하는 것이 필요하다. 우선 모바일 단말의 이동성 예측 등의 짧은 지연 시간을 요구하는 분석 요청에 대해서는 DNN 을 활용하나, 네트워크 슬라이싱과 같은 상대적으로 긴 지연 시간을 요구하는 분석 요청에 대해 다른 분석 요청을 기다리는 대기 시간을 가진다. 대기 시간을 가지는 경우, NWDAF 는 이미 들어왔던 분석에 필요한 데이터와 관련된 input 데이터를 활용하는 분석을 기다린다. 즉, 분석 요청이 발생한 이후 대기 시간에 또 다른 분석 요청이 발생하였을 때, 앞서 발생한 분석과 관련이 없는 분석이라면 두 요청 모두 긴 대기 시간을 수용하는 분석 요청이라도, 같은 MTL 모델을 사용하지 않는다. 이와 같은 경우 관련 데이터로 학습된 다른 MTL 모델을 활용할 수 있다. 결과적으로 학습 데이터셋에 따라 다른 학습 모델을 유지하고, 이후 분석

요청에 대해 적합한 MTL 및 DNN 모델을 선택하면 NF 의 요구 지연 시간을 만족하면서 NWDAF 에서의 컴퓨팅 비용을 낮춰 효율적인 NWDAF 운용이 가능하다.

IV. 결론

본 논문에서는 코어 네트워크 기능인 NWDAF 에 대한 표준을 소개하고 NWDAF 의 효율적인 운영을 위해 ML 모델 프로비저닝 프레임워크를 제안하였다. 제안 프레임워크는 NWDAF 에서의 컴퓨팅 비용과 소비자 네트워크 기능의 지연 시간을 고려하여 추론 모델을 선택한다. 이는 NF 가 요구하는 지연 시간을 만족하면서, NWDAF 의 컴퓨팅 비용을 감소시킬 수 있도록 한다. 향후 연구에서는 본 논문에서 기술한 추론 모델의 장단점을 바탕으로 최소한의 비용으로 NWDAF 를 운영하는 알고리즘을 만들어 성능 평가를 수행할 계획이다.

ACKNOWLEDGMENT

본 연구는 2022 년도 정부(과학기술정보통신부)의 재원으로 정보통신기획평가원의 지원을 받아 수행된 연구임 (No.2022-0-01015, 지능형 6G 모바일 코어 네트워크 후보 요소기술 개발)

참 고 문 헌

- [1] 3GPP TS 29.520 v18.0.0, 3rd Generation Partnership Project; Technical Specification Group Core Network and Terminals; 5G System; Network Data Analytics Services (Release 15), December, 2022.
- [2] 3GPP TS 23.288 v18.0.0, 3rd Generation Partnership Project; Technical Specification Group Services and System Aspects; Architecture enhancements for 5G System (5GS) to support network data analytics services (Release 18), December, 2022.