논문 23-48-09-17 The Journal of Korean Institute of Communications and Information Sciences '23-09 Vol.48 No.09
https://doi.org/10.7840/kics.2023.48.9.1190

RAN Intelligent Controller의 Data-Driven Control Loop를 위한 네트워크 데이터 분석

조 영 준•, 방 석 영*, 유 현 민*, 홍 인 기°

Network Data Analytics for Data-Driven Control Loop of RAN Intelligent Controller

Young-Jun Cho•, Seok-Yeong Bang*, Hyeon-Min Yoo*, Een-Kee Hong°

요 약

5G 기술이 본격적으로 상용화됨에 따라, 기존 네트워크의 문제점을 해결하고 성능을 향상하기 위한 새로운 radio access network (RAN) 구조에 대한 연구가 다수 진행되고 있다. 대표적인 연구는 O-RAN alliance가 주관하는 open RAN으로, 네트워크에서 발생하는 데이터를 기반으로 지능적으로 RAN을 제어하는 RAN intelligent controller (RIC)라는 새로운 entity를 도입하였다. 본 논문에서는 RIC에 탑재되는 인공지능 모델의 학습을 위한 오픈소스 데이터를 소개하고, 이를 분석하여 활용 방안을 제시한다.

Key Words : 5G, Open RAN, RAN intelligent controller, Closed-loop control, Colosseum

ABSTRACT

As 5G technology deploys on a commercial scale, many researchers studied the new radio access network (RAN) structures to solve some problems and improve the performance of existing networks. The O-RAN Alliance suggests open RAN as a next-generation cellular network. Also, the O-RAN alliance defines a new entity called RAN intelligent controller (RIC), which controls RAN leveraging artificial intelligence (AI) and machine learning (ML). In this paper, we introduce and analyze open-source datasets used to train AI models for RAN control by RIC and suggest practical ways to utilize them.

※ 이 논문은 2023년도 정부(과학기술정보통신부)의 재원으로 정보통신기획평가원의 정보통신방송기술국제공동연구사업 지원을 받아 수행된 연구임 (No.RS-2022-00207387, 5G Open RAN 기반 지능형 네트워크 슬라이싱 기술 개발)
※ 본 연구는 과학기술정보통신부 및 정보통신기획평가원의 대학ICT연구센터육성지원사업의 연구결과로 수행되었음 (IITP-2021-0-02046*)
※ 이 논문은 2023년도 정부(과학기술정보통신부)의 재원으로 정보통신기획 평가원의 지원을 받아 수행된 연구임 (No. 2018-0-01659, 5G NR 기반 지능형 오픈 스몰셀 기술 개발)
• First Author : Kyunghee University Department of Electronics and Information Convergence Engineering, rkdtmd34@khu.ac.kr, , 학생회원
° Corresponding Author : Kyunghee University Department of Electronics and Information Convergence Engineering, ekhong@khu.ac.kr, 종신회원
* Kyunghee University Department of Electronics and Information Convergence Engineering, {qkdtjr97, yhm1620}@khu.ac.kr, 학생회원
논문번호 : KICS202305-113-0-SE, Received May 31, 2023; Revised June 19, 2023; Accepted June 19, 2023

Ⅰ. Introduction

In the 5G era, new radio access networks (RANs) have emerged to support various complex use cases. Unlike conventional RAN, which makes it difficult to change functions as desired owing to black-box equipment, open RAN promotes virtualized RAN, where software-defined functions are deployed in programmable white-box equipment [1]. Software-defined RAN enables the dynamic control of RAN, and intelligent controllers (RICs) intelligently control multiple RANs [2].

The RIC controls multiple base stations (BSs) and collects data generated by BSs. The RIC monitors the current network state and determines the appropriate policies using data-driven approaches based on the collected data. In March 2023, the O-RAN Alliance Working Group 1 released specifications with 23 open RAN use cases [3]. These include radio resource management (RRM), MIMO, RAN slicing, and industrial Internet of Things (IoT) optimization. In most of these cases, the RIC plays a key role in closed-loop control. Closed-loop control refers to the logical loop of data collection and policy decisions made by the RIC to control BSs. The RIC can train artificial intelligence (AI) models using data collected from BSs, and the trained models are deployed on the RIC in the form of software called xApp and rApp [4].

To improve AI-based control-loop control functions implemented through xApp and rApp, various types of data generated by real BSs are essential. The dataset for RAN control requires not only basic network metrics such as uplink/downlink throughput but also control plane data indicating the channel state. However, mobile network operators are the only institutions that can collect data, and industries and universities have difficulty obtaining datasets for research purposes. In this paper, we introduce an open network testbed implemented to solve these problems and a project that releases the dataset collected from the testbed as an open source. We analyze the dataset and propose a new research direction for the commercialization of open RAN.

Ⅱ. Colosseum Dataset

Colosseum is the world's largest wireless network emulator at Northeastern University in the United States and consists of USRP acting as a software-defined radio module and srsRAN acting as an open-source RAN software. Colosseum can be used to investigate radio frequency performance by modeling various network scenarios, such as densely populated cities and shopping malls, and conducting AI research by leveraging the data generated by multiple network nodes. The colosseum also includes open RAN scenarios, enabling researchers to experiment with massively deployed softwarized-RAN and near-real-time RIC (near-RT RIC).

In this paper, we present a 7 GB dataset from the Colosseum testbed released by Northeastern University as an open source that includes communication metrics from BSs and user equipment (UEs)[5]. The dataset was collected from partial areas consisting of four BSs and 40 UEs located in the colosseum testbed and was produced in the form of a CSV file. The UEs in the network were categorized as stationary UEs and UEs moving at 3 m/s. The downlink and uplink frequencies were 0.98 and 1.02 GHz, respectively. The channel bandwidth is set to 3 MHz. The dataset consists of information about the BSs and UEs as well as network metrics indicating

Table 1. Metrics of Colosseum dataset

IMSI	International Mobile Subscriber Identity
RNTI	Radio Network Temporary Identifier
PRB	Physical Resource Block
CQI	Channel Quality Indicator
PMI	Precoding Matrix Indicator
RI	Rank Indicator
MCS	Modulation Coding Scheme
RSSI	Received Signal Strength Indicator
SINR	Signal Interference Noise Ratio
PHR	Power Headroom Report
eMBB	enhanced Mobile Broad Band
mMTC	massive Machine Type Communication
URLLC	Ultra-Reliability and Low Latency Communication

the status of the communication link. The metrics of the dataset are summarized in Table 1.

Timestep: collected every 500 ms for about 15 minutes.

IMSI, RNTI: A metric for distinguishing the UEs, which refers to a unique identifier stored in the USIM and an identifier defined only within the connected cell.

Slice enable, id, PRB: The services used by the UE are classified into three slices: eMBB, mMTC, and URLLC, each labeled with an id of 0, 1, and 2.

Buffer, bit rate: This refers to the size of the buffer inside the BS and the bit rate serviced to the UE. Service latency can be measured by leveraging the buffer size.

Number of packets: Defined differently for each use case, the eMBB transmits 125 bytes of packets several times to achieve a transmission rate of 1 Mbps. mMTC transmits 125 bytes of packets 30 times per second, while URLLC transmits packets of the same size 10 times per second. The transmitted packets followed a Poisson distribution.

CQI, PMI, RI: Information on the channel states between the BSs and UEs that change over time.

MCS: Modulation and coding scheme based on channel state defined by CQI.

RSSI, SINR, PHR: Metrics representing the strength of the signal received by the UEs.

Sum requested PRB, Sum granted PRB: PRB requested by the UE and PRB granted by the BS; the ratio of the two metrics is the ratio granted requirement. This can be used as a performance evaluation metric for URLLC.

Fig. 1 shows the ratio of granted requirements of the raw dataset. In many cases, UEs requesting URLLC do not achieve guaranteed performance. This suggests the need for sophisticated policy-based network management and that performance can be improved by leveraging AI/ML models through RIC. To develop an efficient AI/ML model, it is necessary to reduce dimensionality by analyzing the correlation of the dataset and selecting only metrics with independent information (Fig.2). For example, CQI and MCS have a high correlation because better channel conditions naturally lead to a high MCS; therefore, we can exclude one of them to decrease the dimensionality of the state of the network.

The authors of [6] develop deep reinforcement learning (DRL) agents that determined the scheduling policies for each BS slice. They trained DRL agents using a buffer, bit rate, ratio grant requirement, slice ID, and slice PRB. They deployed the DRL agents on the O-RAN software community's RIC. Similarly, a dataset can also be used to develop a data-driven cellular network system. Researchers can use the appropriate data to train AI models according to their purposes.

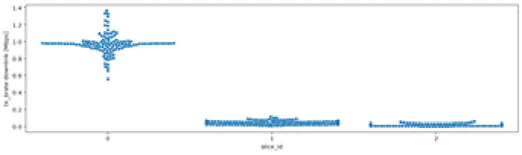

Fig. 1. Downlink bit rate distribution by slice among data generated by Colosseum testbed

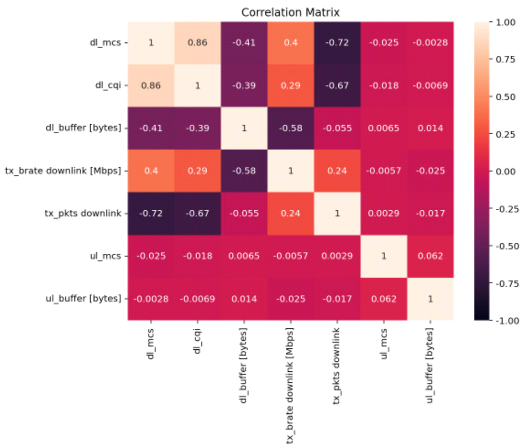

Fig. 2. Correlation plot of data generated

Ⅲ. Conclusions

In this study, we introduced a Colosseum testbed for collecting data for AI-based RAN control and analyzed the dataset in detail. Unlike existing open-source cellular network datasets that only disclose the aggregated throughput of the BS, this dataset has significant academic value because it discloses information on control plane signals. We hope that the dataset presented in this paper will contribute to advances in open RAN technology.

References

[1] L. Bonati, M. Polese, S. D'Oro, S. Basagni, and T. Melodia, "Open, programmable, and virtualized 5G networks: State-of-the-art and the road ahead," *Comput. Netw.*, vol. 182, pp. 1-28, Dec. 2020. (https://doi.org/10.1016/j.comnet.2020.107516)

[2] B. Balasubramanian, et al., "RIC: A RAN intelligent controller platform for AI-Enabled cellular networks," *IEEE Internet Computing*, vol. 25, no. 2, pp. 7-17, Mar.-Apr. 2021. (https://doi.org/10.1109/MIC.2021.3062487)

[3] *O-RAN Use Cases Analysis Report-R003-V10.00*, O-RAN Working Group 1, Alfter, Germany, 2023.

[4] A. Arnaz, J. Lipman, M. Abolhasan, and M. Hiltunen, "Toward integrating intelligence and programmability in open radio access networks: A comprehensive survey," *IEEE Access*, vol. 10, pp. 67747-67770, 2022. (https://doi.org/10.1109/ACCESS.2022.3183989).

[5] *https://github.com/wineslab/colosseum-oran-commag-dataset*

[6] L. Bonati, S. D'Oro, M. Polese, S. Basagni and T. Melodia, "Intelligence and learning in O-RAN for data-driven NextG cellular networks," *IEEE Commun. Mag.*, vol. 59, no. 10, pp. 21-27, Oct. 2021. (https://doi.org/10.1109/MCOM.101.2001120)

조 영 준 (Young-Jun Cho)

2022년 2월 : 경희대학교 전자공학과 학사 졸업
2022년 3월~현재 : 경희대학교 전자정보융합공학과 석사과정
<관심분야> User Association, Load Balancing, 5G
[ORCID:0009-0002-3798-1546]

방 석 영 (Seok-Yeong Bang)

2022년 2월 : 경희대학교 전자공학과 학사 졸업
2022년 3월~현재 : 경희대학교 전자정보융합공학과 석사과정
<관심분야> Small cell on/off, Traffic Classification, 5G
[ORCID:0009-0003-1258-658X]

유 현 민 (Hyeon-Min Yoo)

2021년 2월 : 경희대학교 전자공학과 학사 졸업
2023년 2월 : 경희대학교 전자정보융합공학과 석사 졸업
2023년 3월~현재 : 경희대학교 전자정보융합공학과 박사과정
<관심분야> Mobile Communication, O-RAN, 5G
[ORCID: 0000-0001-6385-2655]

홍 인 기 (Een-Kee Hong)

1989년 2월 : 연세대학교 전기공학과 학사 졸업
1991년 2월 : 연세대학교 전기공학과 석사 졸업
1995년 8월 : 연세대학교 전기공학과 박사 졸업
1995년~1999년 : SKT 선임연구원
1999년~현재 : 경희대학교 전자공학과 교수
2012년~현재 : 미래창조과학부 주파수 정책 자문위원
2013년~현재 : 5G 포럼 주파수 위원회 위원장
2014년~현재 : 국무조정실 주파수 심의위원
2018년~현재 : 한국통신학회 회장
2018년~현재 : 과기정통부 전파정책 자문위원
2021년~현재 : 위성통신포럼 주파수 위원회 위원장
<관심분야> Mobile Communication, O-RAN, 5G
[ORCID:0000-0001-6777-7058]