A Study on the Graphical Neural Network for Network Slicing for Beyond 5G Mobile Systems

M. Kowsalya, Seok-Woo Park, Hyung-Jin Kim, In-Ho Ra*

Vellore Institute of Tech., Kunsan National Univ., Jeonbuk National Univ., *Kunsan National Univ.

mkowsalya@vit.ac.in, p_sw@kunsan.ac.kr, kim@jbnu.ac.kr, *ihra@kunsan.ac.kr

5G 이후의 모바일 시스템에서 네트워크 슬라이싱을 위한 그래픽컬 뉴럴 네트워크에 관한 연구

코살라, 박석우, 김형진, 나인호*

벨로르대학교, 군산대학교, 전북대학교, *군산대학교

Abstract

Network slicing has emerged as a promising networking paradigm to provide resources tailored for Industry 4.0 and diverse services in 5G networks. The proposed graph neural network model can learn insights directly from slicing–enabled networks represented by non-Euclidean graph structures. Experimental results show that the graph neural network slicing accurately mirrors the network behavior and predicts the E2E latency under various topologies and unseen environments.

Ⅰ. Introduction

The launch of 5G and the impending arrival of 6G have ushered in a new era of hyper-connectivity based on data offerings in the current, quickly expanding telecoms market. The 5G systems have been getting a lot of attention, and they additionally offer an overview of 3GPP Release 17 as well as future releases [1]. The concept of internet slicing has been proposed as a promising approach to fulfilling the prominent paradigm shift from a one-size-fits-all solution to software and virtualized design. [2]. The recent advances in artificial intelligence provide a promising means to fulfill the demands of network development. The emergence of data-driven machine learning (ML) techniques especially deep learning has gained popularity in networking areas and led to a new breed of models that learn from data, instead of being explicitly programmed. Most of these works are based on acclaimed learning architectures such as convolutional neural networks, recurrent neural networks, and Auto Encoders and their variant. However, communication networks are fundamentally represented in the form of graphs, and most of the existing learning architectures are not designed to learn such information structure in the non-Euclidian domain, due to the irregular topology of graphs and interdependence between nodes. As a result, these models are limited in providing accurate results on graph data and are hard to achieve generalization on dynamic topologies and configurations. To solve the above challenges, we design a network based on graphical neural networks (GNN) to discover the complicated relationships and interdependence among network slicing, resource allocation, and physical infrastructure, and to generate the end-to-end metrics prediction of each slice under diverse scenarios. The network model proposed in this article achieves a vital step toward realizing the ambitious vision of autonomous management of network slicing by providing the ability to map the network status to the end-to-end quality of service performance v for complex slicing– enabled networks.

Ⅱ. Method

The problem of producing end-to-end latencies of multiple slices that share the same infrastructure is shown in Fig. 1. The left part shows a network that contains three slices sharing the same substrate network, including a RAN, an edge network, and a core network. Each slice is dedicated to a service that requires specific functionalities in a particular order. Each network function requires resources from different domains of the substrate network. To establish the graph-based virtual representation, we denote the state of a node in the graph as h_i which embodies the information of the slicing traffic traversing this node and utilization of links connected by the node. The state of a slice h_s is subject to the states of all the nodes within it, which contains the key information to obtain end-to-end latency. Next, we develop a graph-based network to produce accurate end-to-end slicing latency and have enough flexibility to adapt to the dynamic slice deployment, including the change of several slices, different physical topologies, and variations of slice utilization. Furthermore, since the traffic has to travel through a predefined ordered virtual network functioning in a network slice, the packet losses and delay on any link would affect the overall latency. We will also propose a nonlinear method to aggregate the state of nodes of slices. The proposed GNN base network slicing is implemented in TensorFlow and Stellargraph, a graph ML library. The datasets are chosen not only because of their generalized configuration but also due to the pre-source destination measurements, which can mimic the end-to-end network slicing metrics that are needed for the performance evaluation. We transform the traffic matrices under three topologies into three network slicing scenarios to fit the end-to-end slicing model. The dataset is based on the real-time network including a topology of 14 nodes and 42 links with 3 slices deployed. Fig. 2 depicts the results generated by the trained GNN applied to a scenario with various numbers of slices. Each bar in the figure refers to the median values of the generalized E2E latency predictions after running the algorithm 100 times for the scenario including the topology of 14 nodes and 42 links with three slices deployed.

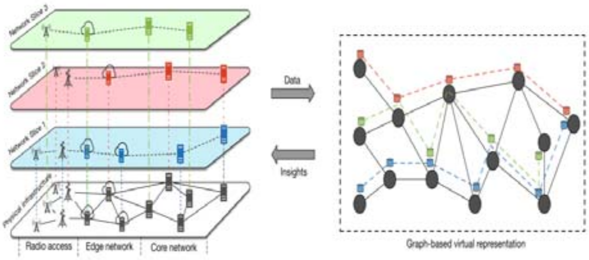

Fig. 1 Network slices and associated graphs

Fig. 2 Predicted E2E Latency vs True Latency

Ⅲ. Conclusion

In this study, we investigated the key challenges in ensuring the E2E network slicing performance in Industry 4.0 and various 5G applications and developed a network for network slicing. We exploited the state– of– the– art GNN model to solve the E2E slicing challenges, by developing a virtual representation to construct a graph from the network consisting of slices and discovering insights directly on the graph, instead of converting the network into matrices..

ACKNOWLEDGMENT

This work was supported by the National Research Foundation of Kores (NRF) grant funded by the Korea government (MSIT) (No. 2021R1A2C2014333).

REFERENCES

[1] Afraz, N, Slyne, F, Gill, H, Ruffinin, M “Evolution of access network sharing and its role in 5G networks”, Applied sciences, 9,(21), 2019, 4566.

[2] Wu, Z, Pan, S, Chen,F, Long,G, Zhang, C , Yu,P.S, “ A comprehensive survey on graph neural networks”, IEEE Trans. Neural Netw. Learn. Syst. 32,(1) 2021,4-24.