GNN-based Optimization VNF Deployment on Resource Awareness on the Internet of Things

Seyha Ros[1], Sa Math[2], Tharoeun Thap[2], Seokhoon Kim*[1]

[1]Software Convergence, Soonchunyang University, Asan-si, South Korea

[2]Telecommunication Regulator of Cambodia, Ministry of Post and Telecommunication of Cambodia, Phnom Penh, Cambodia

*Corresponding author's e-mail: seokhoon@sch.ac.kr

doi: https://doi.org/10.21467/proceedings.174.1

Abstract

Network functions virtualization (NFV) and software-defined networks (SDN) offer a promising system architecture with capabilities of programmable control, elastic resource placement, flexible hardware, and a dynamic service life cycle for multi-service orchestration. In the emerging era between software/virtualized networks and artificial intelligence, the deployment of data-driven mechanisms for assisting autonomous SDN/NFV management and orchestration is rapidly increasing. However, in routing problems of VNF mapping, the data can be gathered as non-Euclidean, which requires the modeling to integrate with graph-based deep learning (DL). The paper introduces a graph neural networks (GNN)-based approach tailored for optimizing graph-structured routing paths, nodes, links, and capacities in SDN/NFV environments. This approach aims to achieve delay-aware optimization and prioritize per-flow usage based on source requirements, thereby enhancing the effectiveness of NFV deployments. Key architectural components such as input graph representation, aggregation, update mechanisms, and readout strategies within SDN/NFV systems. Moreover, the proposed system models emphasize efficient utilization of both computation and communication resources, crucial for enhancing overall network performance and resource management. By leveraging GNNs, the study anticipates advancements in autonomous SDN/NFV management and orchestration, enabling more responsive and adaptive networks capable of meeting diverse service demands effectively.

Keywords: Graph Neural Network, Network Functions Virtualization, Software-defined Networks

1 Introduction

The big data generated by physical devices and machine learning (ML) techniques-based networking is becoming the mainstream for the next generation of networking in both wireless communication and conventional network model-based design schemes. There are several types of graph-structured information ranging from each physical server's node capacities to the routing path's link capacities. With efficient use of feature extraction from graph-structured data, the model can be significantly improved with sufficient data-driven input [1], [2]. Graph neural networks (GNN), which use deep learning for the graph-like environment, analyze the patterns between nodes, links, and overall capacity increment/decrement relations using the processing flow from initialization (input graph), message passing with aggregation and update, and readout (output). The input graph observes the possible network states, which are highly efficient in SDN-enabled architecture [3]. The data plane information can be gathered for the controller to obtain in the software-defined network (SDN) database. By adopting these technologies, SDN and NFV-enable the controller entities for levitating and facilitating the computing tasks and changing services in terms of network service by ruined in sequence of virtual network functions (VNFs). Through the created a sequence of ordered VNFs, which technique can be implemented by Service function chaining (SFC) to define and management the flow of network service across multiple VNFs. To perform for different characteristic perspective, namely physical resource, representee node, VNF-FG cluster and SFC cluster. For instance, the representee node is mainly represented by network devices. In capability of resources are abstracted by controller for global view. The feature extraction will be executed for normalizing the feeding information for GNN. Aggregation and update functions follow the specification of the variant GNN selection, which differs between attention, multi-layer perceptron for aggregation, graph convolution networks mechanism, etc. The comprehensive service requirements in IoT networking are faced with many obstacles in resource and routing. It is necessary to enhance the performance of core networking, which enables machine learning techniques to predict the networks required. The load resource-pooling to the physical layer is important for gathering device information to treat the functions. Furthermore, the demand for various network services requires heterogeneous network functions, which can be realized with specific hardware such as RAM, CPU, and Network when considering the conventional network modules. However, the tradition of the number of hardware modules for each of the function modules is the feasible network and cost-efficient solutions based on capital expenditure (CAPEX) and operation expenditure (OPEX). In Fig 1. An illustration of IoT devices connected to an access point for processing network service functionality on-demand is shown. After the data plane is fully negotiated and acknowledged by the IoT devices to access the network, the controller in distributed areas will gather the device

Proceedings DOI: 10.21467/proceedings.174; Series: AIJR Proceedings; ISSN: 2582-3922; ISBN: 978-81-984081-6-7

status to store in the database in the cloud. The SDN will capture the information and then interact with GNN, and enhancement of resources utilized with NFV entities. In this paper is addressed to aim provides per-flow priority for Dynamic-Resource and Shortest-Destination routing between the distributed cloud and cloud. The readout can be aimed at flow-level, link-level, and port indicators according to application service requirements [4], [5].

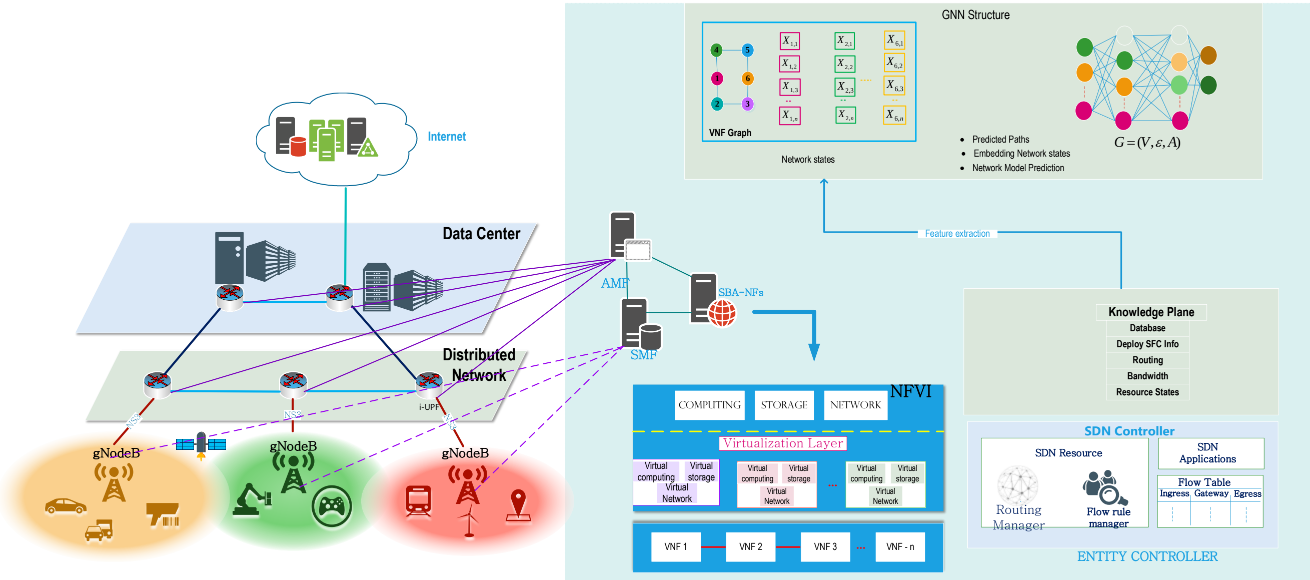

Fig. 1. The interaction of the network and graph neural network in SDN/NFV controller.

2 Proposed Approach

With the supportive systems of the data plane, control plane, and knowledge plane, adopted GNN in SDN architecture is efficient and application. In Fig 2. emerged system is proposed by (1) gathering the graph-structured topology information and PACKET_IN messages, (2) the controller is used for configuring the path, (3) the path configuration and overall topology (links, nodes, and capacities) are collected in knowledge plane-database for inputting into the GNN. The delay-aware approach aims to proactively predict the delay metrics for altering the path configuration in the controller and adjusting the resource utilized in computation capability.

The proposed GNN-based approach can activate modern delay-aware features on extending data complexity, routing, and multi-queue scheduling policies. The primary principles of relationships between flow, routing path, and link statuses have to be included. Each execution necessitates evaluating the performance accuracies compared to the conventional routing path. After a fine-grained evaluation, multi-conditions are required to be set including the different complexity levels of traffic congestion, various scheduling configurations, and the applicability in real-world topologies. The primary phases in this aspect include: (1) ***Initialization (input graph)***: directed graph of weight information between routing links, bandwidth, and patterns. For instance, GNN establishment gathers the data plane features such as communication capability, resource capacity, network state of the distributed networks. (2) ***Feature extraction and attention-based graph networks***: the fixed-dimensional vectors and the implementation of attention mechanisms are given between nodes and different neighboring. The features are extracted after multi-round aggregation and updates. (3) ***Output***: modified routing path after the delay prediction with the indication of the current node, destination node, and path forwarding paths. After completing the three phases, the GNN-proposed approach will make a model prediction of network states. GNN will create a new model of network status and provide it to the SDN/NFV controller to create rule tables and flow tables based on user requirements.

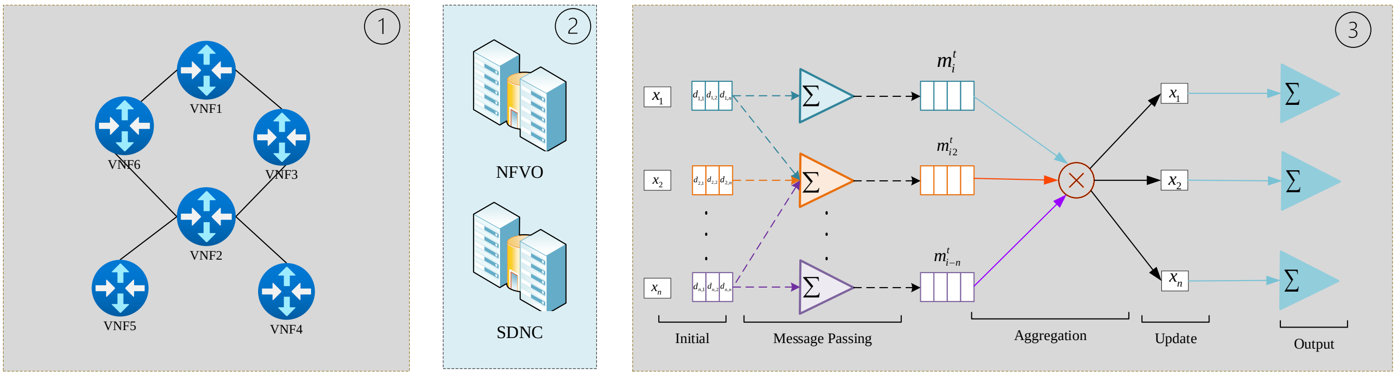

Fig. 2. GNN emerging in SDN-enabled architecture.

3 Conclusion

This paper proposes an early approach to applying GNN-based modeling for delay-aware routing and resource optimization in SDN/NFV-enabled networks. The network states information, output metrics, and plane interactions for supporting GNN-based deployment are presented. The flows between the data plane, controllers, and modeling mechanisms are given as deployment cooperation. In future studies, in-depth communication and computation models for fine-grained resource consideration will be extensively studied. It is expected that GNN will be a player with increasing importance for the next generation in wireless networking and core network as well. Future research directions outlined in the paper include exploring advanced communication and computation models to incorporate fine-grained resource considerations. This will involve developing more nuanced approaches that integrate GNNs with traditional methods, such as message passing and graph attention mechanisms, to achieve deeper insights into network behavior and performance. Furthermore, the study anticipates GNNs becoming increasingly pivotal in shaping the future of wireless networking and core network architectures. As these technologies evolve, there is a growing expectation that GNNs will play a prominent role in enhancing network intelligence and operational efficiency

4 Declarations

4.1 Acknowledgment

This work was supported by the Institute of Information & communications Technology Planning & Evaluation (IITP) grant funded by the Korean government (MSIT) (No. RS-202200167197, Development of Intelligent 5G/6G Infrastructure Technology for The Smart City).

4.2 Competing Interests

The authors declare no conflict of interest.

4.3 Publisher's Note

AIJR remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

How to Cite

Seyha Ros, Sa Math, Tharoeun Thap, Seokhoon Kim (2025). GNN-based Optimization VNF Deployment on Resource Awareness on the Internet of Things. *AIJR Proceedings*, 1-3. https://doi.org/10.21467/proceedings.174.1

References

[1] Xie, Y., Huang, L., Kong, Y., Wang, S., Xu, S., Wang, X., Ren, J.: Virtualized Network Function Forwarding Graph Placing in SDN and NFV-Enabled IoT Networks: A Graph Neural Network Assisted Deep Reinforcement Learning Method. IEEE Transactions on Network and Service Management. 19, 524–537 (2022).

[2] Rusek, K., Suarez-Varela, J., Almasan, P., Barlet-Ros, P., Cabellos-Aparicio, A.: RouteNet: Leveraging Graph Neural Networks for network modeling and optimization in SDN. IEEE Journal on Selected Areas in Communications. 1–1 (2020).

[3] He, S., Xiong, S., Ou, Y., Zhang, J., Wang, J., Huang, Y., Zhang, Y.: An Overview on the Application of Graph Neural Networks in Wireless Networks. IEEE Open Journal of the Communications Society. 2, 2547–2565 (2021).

[4] Ruiz, L., Gama, F., Ribeiro, A.: Graph Neural Networks: Architectures, Stability, and Transferability. Proceedings of the IEEE. 109, 660–682 (2021).

[5] Suarez-Varela, J., Almasan, P., Ferriol-Galmes, M., Rusek, K., Geyer, F., Cheng, X., Shi, X., Xiao, S., Scarselli, F., CabellosAparicio, A., Barlet-Ros, P.: Graph Neural Networks for Communication Networks: Context, Use Cases and Opportunities. IEEE Network. 1–8 (2023).