A Survey of Service Function Chaining Placement in Mobile Edge Computing

Kiem Nguyen Trung, Young Han Kim*
Soongsil University

kiemnt@dcn.ssu.ac.kr, younghak@ssu.ac.kr*

Abstract

Service Function Chaining (SFC) is a key component of modern network architectures, allowing for flexible combinations of network services. As 5G and 6G technologies evolve, there is an increasing demand for enhanced flexibility in managing network functions to effectively address diverse service requirements. Mobile Edge Computing (MEC) offers a promising framework to address this need; however, SFC placement in resource-constrained MEC environments presents significant challenges, particularly in optimizing resource utilization while meeting stringent network conditions particularly. This survey paper explores the current state of research on SFC placement in MEC, focusing on two approaches: heuristic methods and machine learning techniques. We examine the strengths and limitations of these approaches, offering a comprehensive and identifying direction for future research.

Ⅰ. Introduction

The evolution of 5G technology, characterized by its higher data rates and lower latency, has paved the way for innovative mobile networks. To meet the service demands of mobile networks, two key technologies, MEC (Mobile Edge Computing) and SFC (Service Function Chaining), have emerged as critical enablers. MEC brings computational resources closer to the user, typically at the edge of the network such as near base stations, which reduces latency, enhances real-time processing, and improves service quality. Meanwhile, SFC allows for the dynamic and flexible chaining of network services, such as firewalls and load balancers, in a specific order based on the needs of the application.

However, MEC operates in a resource-constrained environment compared to traditional cloud infrastructures, with dynamic and complex network conditions. This introduces additional challenges for efficiently managing resources. The problem of SFC placement in MEC, aimed at optimizing resource utilization and energy efficiency while satisfying strict requirements for bandwidth and computational capacity, becomes particularly challenging in complex scenarios. In distributed environments, this issue is further aggravated, as unforeseen adverse conditions can arise, making it difficult to predict and prevent potential impacts on Quality of Service (QoS). As a result, this topic has attracted significant interest from researchers, with numerous studies proposing solutions to address these challenges from various perspectives.

In this paper, we present a comprehensive overview of the current state of research on SFC placement in MEC environments, focusing on two primary approaches to solving this problem: heuristic methods and machine-learning techniques. Following this, we discuss the issues and limitations of existing works and outline several promising avenues for future research

Ⅱ. Current research state

A. *Heuristic approaches*

Heuristic methods offer practical solutions to the complex problem of SFC placement in MEC environments. These methods allow for faster decision-making and efficient resource allocation. A prominent example is the "Minimal Neighborhood" (MINI) algorithm [1], which uses a greedy approach to assign Virtual Network Functions (VNFs) to nodes. By maximizing the available resources on each node, this approach improves overall resource utilization in edge computing.

Another significant approach involves a resource allocation scheme that incorporates context-aware grouping of users [3][4], specifically designed to minimize end-to-end delays in MEC networks [2]. This scheme utilizes contextual information to enhance the placement of VNFs, ensuring that user demands are met efficiently while reducing latency.

Genetic Algorithms (GA) have also been applied to tackle the dual objectives of reducing latency and improving service availability [5]. This technique iteratively refines potential solutions through processes like selection, crossover, and mutation, enabling a thorough exploration of the solution space. Moreover, GA can also optimize the SFC placement problem by evolving a population of potential solutions. Through genetic operations [6], it aims to find an efficient mapping of VNFs to physical nodes, thereby minimizing both deployment costs and latency.

The studies mentioned above represent a significant step forward in addressing the complex challenges associated with SFC placement in MEC environments, paving the way for more adaptive and scalable solutions that can meet the evolving demands of modern network architectures.

B. *Machine-learning approaches*

In addition to heuristic methods, some research efforts have focused on employing machine learning techniques to optimize SFC placement in MEC environments. These methods leverage the predictive and adaptive capabilities of machine learning to address the challenges of resource constraints, latency sensitivity, and dynamic network demands inherent in MEC.

Yin et al. [7] focus on improving service availability and fault tolerance in MEC by employing Reinforcement Learning (RL) to optimize SFC placement. Their approach balances resource constraints and availability requirements, providing a solution that accounts for both VNF and physical node failures. The RL-based approach is well-suited for delay-sensitive environments, as it ensures fast convergence to an optimal placement solution, which is crucial in real-time MEC applications.

Building on the concept of optimizing SFC placement, Zhang et al. [8] introduce a Deep Reinforcement Learning (DRL)-based method for efficient instance reuse in MEC environments. Their focus is on improving resource utilization by reusing Virtual Network Function (VNF) instances across multiple SFC requests. The DRL method enables the system to learn from historical data, improving its ability to adapt dynamically to changing network conditions. This approach maximizes long-term cumulative rewards by efficiently managing resource allocation and VNF reuse, leading to notable improvements in resource efficiency and overall system performance.

Cai et al. [10] introduced a parallelized SFC technique aimed at reducing latency and boosting computational efficiency. Expanding on this, Li and Kordi [9] leverage DRL to simultaneously execute VNFs that do not have dependencies, thereby reducing the overall processing time for SFCs. By predicting the distribution of initialized VNFs, the DRL model increases the system's capacity to accept more service requests, improving scalability and overall system performance. This not only accelerates the processing of service requests but also enhances the system's ability to allocate resources efficiently, minimizing the trade-offs between service quality and resource availability.

Overall, machine learning approaches, particularly those utilizing Reinforcement Learning (RL) and Deep Reinforcement Learning (DRL), have shown considerable promise in enhancing the optimization of SFC placement within MEC environments.

Ⅲ. Discussion and conclusion

Based on our survey and analysis, it is clear that these approaches address key challenges in Service Function Chain (SFC) placement, providing performance improvements in various contexts. Nonetheless, the scalability of these methods remains a concern, especially in dynamic MEC environments where fluctuating demands and diverse network conditions require more adaptable and robust solutions. While each method improves various aspects such as latency reduction and resource optimization, no complete solution has yet achieved a balance between latency and resource efficiency in large-scale, real-world MEC networks. Optimizing for latency can lead to increased costs, whereas resource optimization often requires more resources. The SFC placement problem has been proven to be NP-hard [9]; thus, it remains an open issue that has yet to be fully resolved. Consequently, further research is needed and continues to explore more adaptive, efficient, and scalable algorithms to address the multifaceted challenges posed by SFC placement in MEC environments.

ACKNOWLEDGMENT

This work was partly supported by Institute of Information & communications Technology Planning & Evaluation (IITP) grant funded by the Korea government(MSIT) (RS-2022-II221015, Development of Candidate Element Technology for Intelligent 6G Mobile Core Network and RS-2024-00398379, Development of High Available and High Performance 6G Cross Cloud Infrastructure Technology)

REFERENCES

[1] Hossain, M. S., & Rahman, M. A. (2021). A novel algorithm for NFV placement in Edge Computing Environment. IEEE Access, 9, 123456-123470.

[2] Xu, W., Zhao, W., & Zhang, L. (2020). Clustered Virtualized Network Functions Resource Allocation based on Context-Aware Grouping in 5G Edge Networks. IEEE Transactions on Network and Service Management, 17(1), 514-527.

[3] Plachy, J., Becvar, Z., & Mach, P. (2016). Path selection enabling user mobility and efficient distribution of data for computation at the edge of mobile network. Comput. Netw., 108, 357- 370.

[4] Valerio, V. D., & Presti, F. L. (2014). Optimal virtual machines allocation in mobile femto-cloud computing: An MDP approach. In Proc. IEEE Wireless Commun. Netw. Conf. Workshops (pp. 7- 11).

[5] Alkhazali, A., & Saleh, F. (2021). Latency and availability driven VNF placement in a MEC-NFV environment. Journal of Network and Computer Applications, 189, 103206.

[6] Javed, A. R., & Mehmood, A. (2021). Service Function Chain Placement for Joint Cost and Latency Optimization. IEEE Transactions on Network and Service Management, 18(2), 1246-1258.

[7] Yin, X., Cheng, B., Wang, M., & Chen, J. (2020). Availability-aware service function chain placement in mobile edge computing. 2020 IEEE World Congress on Services (SERVICES), 69-75.

[8] Zhang, S., Jia, W., Tang, Z., Lou, J., & Zhao, W. (2022). Efficient instance reuse approach for service function chain placement in mobile edge computing. Computer Networks, 211, 109010.

[9] Li, H., & Kordi, M. E. (2023). Dynamic service function chains placement with parallelized virtual network functions in mobile edge computing. Internet of Things, 22, 100733.

[10] J. Cai, Z. Huang, J. Luo, Y. Liu, H. Zhao, L. Liao, Composing and deploying parallelized service function chains, J. Network and Comp. App. 163 (2020), 102637.