A Market Mechanism for Edge- Cloud Task Offloading under Price Competition and Capacity Constraints

PENG ZHANG, RUOYU ZHU, PIAOHUI ZHE

Hanyang Univ., Hanyang Univ., Hanyang Univ.
zp010807@hanyang.ac.kr, zhryu888@gmail.com, park0912@hanyang.ac.kr.

가격 경쟁 및 용량 제약 하에서 엣지-클라우드 작업 오프로드를 위한 시장 메커니즘

장청명, 주루오위, 박철위
한양대학교, 한양대학교, 한양대학교

Abstract

We study an edge- cloud MEC market with InPs and an SP under multi-hop transmission delays and finite compute capacities. For given wholesale prices, the SP jointly decides task admission/assignment and CPU frequency provisioning via a multi-cut generalized Benders decomposition (MGBD) solver. InPs then update prices by best responses with convex operational costs, enabling evaluation of profit, admission ratio, and utilization.

Ⅰ. Introduction

Edge- cloud MEC reduces service latency by moving computation closer to users, but it also creates a market where infrastructure owners and service providers interact strategically under capacity and delay constraints. We consider a three-tier setting in which InPs announce wholesale prices for computing capacity and an SP procures resources to admit and serve tasks. Our focus is a solver-consistent simulation workflow: shortest-path multi-hop transmission delay modeling, an SP-layer MINLP solved by multi-cut GBD for joint admission/assignment and frequency allocation, and InP best-response pricing with convex cost- load coupling. This design connects MEC system modeling with principled MINLP optimization tools used in wireless resource allocation research.

Ⅱ. Method

We consider an edge- cloud MEC market with a set of tasks and InPs (one cloud and multiple edge nodes). Each InP announces a wholesale price and provides limited CPU frequency capacity. The physical network is modeled as a weighted graph, where link weights represent hop transmission delays. For each task- InP pair, the end-to-end transmission delay is obtained by the shortest-path delay in the graph, which captures multi-hop routing effects commonly emphasized in edge computing systems.

Ⅲ. Conclusion

This study builds a solver-consistent simulation framework for an edge- cloud MEC market with multi-hop delays. The SP layer is solved via multi-cut GBD for joint admission/assignment and frequency allocation, and the InP layer follows best-response pricing with convex operational costs. The framework enables reproducible comparisons on profits, admission ratios, and capacity utilization, and can be extended to multi-SP competition and reliability-aware reservations.

REFERENCES

[1] W. Shi, J. Cao, Q. Zhang, Y. Li, and L. Xu, "Edge Computing: Vision and Challenges," *IEEE Internet of Things Journal*, vol. 3, no. 5, pp. 637- 646, Oct. 2016.

[2] Y. Mao, C. You, J. Zhang, K. Huang, and K. B. Letaief, "A Survey on Mobile Edge Computing: The Communication Perspective," *IEEE Communications Surveys & Tutorials*, vol. 19, no. 4, pp. 2322- 2358, 2017.

[3] P. Mach and Z. Becvar, "Mobile Edge Computing: A Survey on Architecture and Computation Offloading," *IEEE Communications Surveys & Tutorials*, vol. 19, no. 3, pp. 1628- 1656, 2017.

[4] D. T. Nguyen, L. B. Le, and V. K. Bhargava, "Price-Based Resource Allocation for Edge Computing: A Market Equilibrium Approach," *IEEE Transactions on Cloud Computing*, vol. 9, no. 1, pp. 302- 317, Jan.- Mar. 2021.