

Edge Caching Strategy Based on User Preference and Game Theory

Xichen Jia

*Jiangxi Normal University, Nanchang, Jiangxi, 330022, China
1214235791@qq.com*

Abstract: *Caching video content to the mobile edge server is an effective solution to avoid multiple repeated requests from mobile terminal devices, reduce latency costs, and improve user QoE. In addition, mobile users in nearby areas tend to request the same video resource task, so reasonable deployment of video content to edge servers can effectively reduce the response time of user requests. This fact prompted us to design a caching strategy based on user preferences. The system model considered in this article contains multiple mobile users, multiple servers, and remote central servers. Based on the recommendation system predicting the user's preference for specific video resources, the recommended value ranking of the video resources in the future time period is obtained, and then based on the game theory method, in each edge cache server, each video resource is calculated for the local area and neighboring areas. The cache value of the area, based on the cache value of the video resource to be cached, minimizes the delay for users to obtain the video resource, and obtains the optimal video resource cache distribution strategy. The simulation experiment results show that compared with other caching strategies, the proposed caching strategy is better than other caching strategies in terms of response time and cache task hit rate.*

Keywords: *Edge caching, Data Caching, Caching Value, Game Theory, Recommendation System*

1. Introduction

In view of the huge potential of mobile edge caching technology in solving the growth of mobile data traffic and improving network performance, current domestic and foreign researchers have carried out extensive research on mobile edge network caching technology. Research on mobile edge caching technology is mainly focused on solving the "3W" problem, namely, where to cache, how to cache, and what to cache. Cached content refers to the objects requested by the user, such as videos, images, applications, etc.; cache location refers to the location where the user requested objects are stored, such as small cell base stations, mobile terminals, etc.; the cache strategy is used to determine which user requested objects are cached to Which locations (that is, the cache deployment plan) and how to efficiently transmit the cached content to the users by optimizing user access, resource scheduling, and other control plans (that is, the cache content distribution plan). Among them, the caching strategy is a bridge connecting the cache location and the cached content, and it is also a key issue that the mobile edge cache technology needs to solve.

Mobile edge computing can avoid repeated access to the same data by caching technology, and has outstanding advantages in reducing the delay in obtaining content services and reducing the load of the network backhaul link. This technology has attracted widespread attention in the academic and industrial circles. In recent years, there have been many researches on caching at home and abroad, mainly focusing on the research of caching entities, caching strategies that consider user mobility, system model architecture, and caching collaboration.

2. Related work

Literature [1] studied the use of big data in mobile cellular networks from the perspective of active caching. In this article, the author studies active content caching in 5G wireless networks, and proposes an architecture based on a big data platform and a BS that supports caching. First of all, a large amount of valuable data obtained in this architecture is obtained through big data The analysis platform and machine learning tools enter the content popularity evaluation, which is used as the criterion for caching decision, and finally the content is cached on the BS, and based on the user's quality of experience (QoE,

Quality of Experience) and backhaul offloading (Backhaul Offloading). Measure the benefits of proactive caching. Sladana Josilo [2] et al. modeled the mobile backhaul network and edge cache architecture in the cache hierarchy. By modeling the backhaul network with a directed acyclic graph topology, the cache placement problem was described as a 0-1 integer Planning the problem and verifying that when the global information is available, using a distributed greedy algorithm to solve the problem can get a 2-approximate solution. Javedankhearad M [3] and others modeled the caching problem as a graph structure, and used the knowledge of graph coloring theory to solve the file strategy cached in the base station, so that when a user is covered by several base stations at the same time, the cache obtained by the graph coloring theory The strategy can enable him to obtain more files, thereby increasing the cache hit rate of the base station. Sun Y [4] and others are considering the scenario where the user sending request time is random, and using queuing theory to analyze the average request delay in a multi-base station cooperative network that supports caching, and finally adopts a heuristic greedy caching strategy to obtain both A caching strategy that ensures low complexity and can effectively reduce the average delay of the network. Huang [5] et al. developed a MEC edge cache structure and proposed a cooperative multicast-aware cache test to reduce the average delay caused by the delivery of cached content, and proved the effectiveness of this multicast model. Ren D [6] and others proposed a hierarchical collaborative caching strategy based on grouping. The utilization of grouping is to use the fuzzy C-means clustering algorithm. The basic idea is that MECs are divided into different clusters according to the user distribution and the location relationship of MECs. This allows users to obtain cached content in the MEC of the local cluster as much as possible. The author of literature [7] imagined a distributed network structure similar to that of femtocell base station assistants, and proved that the cache optimization problem under this network model is NP-complete, and finally solved it with an approximate algorithm to obtain the final cache strategy. In literature [8], an active caching layout of mobile social networks that supports caching based on a recommendation system was proposed to maximize the traffic offload rate. The article first investigates the social relationship and physical coupling between users, and then selects an important user (IU) as the assister of the base station (BS) to cache the target content, and other users can obtain the cached content from the IU and BS equipment. In [9], the author proposed a novel active caching method in view of the limitations of the current reactive network. Based on machine learning technology, Sengupta et al. [10] proposed an edge caching strategy based on reinforcement learning. From the perspective of reinforcement learning, this strategy solves the problem of distributed caching in mobile edge networks. Through the coding cache method, the cache problem is simplified to a linear programming problem related to network connectivity. Literature [11] solves the problem of designing subscriber unit associations, that is, the small base station (SBS) uses the cache function to overcome the backhaul capacity limitation and enhance the user's QoS (Quality of Service). The SBS can be based on the content availability and the data rate they can transmit. , Respectively decide which users to serve, express the problem as a one-to-many matching game between SBS and users, and propose a distributed algorithm. Although this document solves the problem of buffer allocation for users, it is only a one-to-many solution for coordinating buffer allocation to a certain user at a single base station, and cannot solve the problem of low cache hit rate for the base station itself. Literature [12] proposed an incentive mechanism, in which the base station (BS) rewards those users who share content with other users through D2D communication. The cost minimization problem of the BS and the utility maximization problem of each user are studied and expressed as In the Stackberg game, an iterative gradient algorithm (IGA) is proposed to obtain the Stackberg balance. Literature [13] investigated the problem of maximizing the profits of wireless network operators and minimizing the payment of end users, and proposed a scenario based on risk sharing between end users, in which the competitive relationship between operators and end users It is expressed as the Stackelberg game, which proves the existence and uniqueness of the Nash equilibrium of the non-cooperative subgame.

3. System model

Based on the above literature analysis, we found that the strategy of caching video files on edge devices is still to be studied, such as predicting user preferences, minimizing user delay, reducing transmission costs, collaborative caching between base stations, and so on.

First describe the existing edge computing architecture and apply the proposed caching strategy to this architecture. As shown in Figure 1, the edge computing architecture mainly includes base stations deployed in the wireless access network and edge cache servers placed in the base stations. The base stations are connected by optical fiber, and all base stations are connected to the remote cloud center. A server is placed in each base station to cache video resources, collect user information, and provide services for mobile users.

When mobile users request data resources from the cache server, the following execution strategies are adopted:

Check whether the server locally caches the video resource requested by the user. If the resource requested by the user is cached locally, the local server directly provides the video resource required by the user. If the data resource requested by the user is not cached locally, it is checked whether other servers that are not connected to the mobile user have cached the video resource requested by the mobile user. If other edge cache servers cache the data resource requested by the mobile user, the server will provide the video resource requested by the mobile user, that is, the video resource requested by the mobile user is first sent to the local server, and then sent by the local server to the mobile user. If none of the other edge servers cache the video resources requested by the mobile user, then the data resources requested by the mobile user must be obtained from the remote cloud center. The remote cloud center sends the video resource requested by the mobile user to the local server first, and then the local server sends it to the mobile user. However, when a mobile user requests video resources in a remote cloud center, in this case, the task response time will be very long, and the system will also bear a greater load on the backhaul link.

Due to the limited storage space of the edge server, the video resources with high popularity and high probability of users requesting can be preferentially cached in edge devices in advance to reduce user delay, reduce overall system response time, and improve user experience. To this end, this paper proposes to first collect local user requests for video resources, and predict each user's recommendation value for each video resource based on the recommendation system, and then each base station directly sends it to it through its cheap backhaul link. The neighbor (for example, a one-hop neighbor) base station sends the expected demand of its locally observed video content, and then reasonably caches the video resources with a larger recommended value to the base station of the corresponding edge device through game theory, so as to reach the user and the edge. The utility of the equipment is maximized.

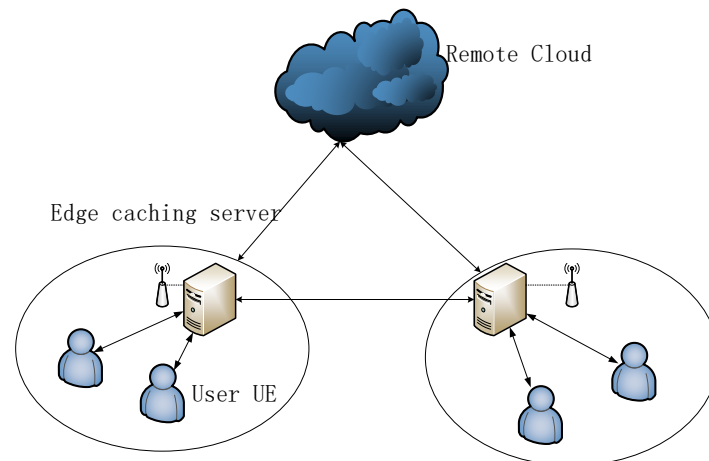


Figure 1: Edge computing architecture

4. Solution

This article intends to use collaborative filtering to predict user preferences for video resources, as follows:

First of all, by analyzing the user's preference behavior, to dig out the relationship between the video and the video, or between people. Secondly, by performing certain calculations on these associated relationships, a guess about the degree of likeness between the person and the video, that is, the recommended value, is obtained. Finally, use this recommended value as a parameter to participate in the game between edge cache devices.

The basis of the algorithm can be summarized in three aspects:

- 1) Pre-filter video resources that users are relatively unwilling to open, that is, irrelevant resources;
- 2) Through the collaborative filtering algorithm, the recommendation matrix of the user's preferred video resources is obtained;
- 3) Since the recommendation matrix is not accurate enough, it is possible to improve the

recommendation matrix through potential factors to obtain the final recommendation matrix, and perform certain transformations on the recommendation matrix to obtain a recommendation value.

Then, each base station sends the expected demand of its locally observed video content to its direct neighbor (for example, a one-hop neighbor) base station through its cheap backhaul link, and then reasonably divides the video based on the higher recommended value through the game method. The resources are cached in the base station of the corresponding edge device to maximize the utility of the user and the edge device.

References

- [1] Zeydan E, Bastug E, Bennis M, et al. Big data caching for networking: Moving from cloud to edge [J]. *IEEE Communications Magazine*, 2016, 54(9): 36-42.
- [2] Jošilo S, Pacifici V, Dán G. Distributed algorithms for content placement in hierarchical cache networks [J]. *Computer Networks*, 2017, 125: 160-171.
- [3] Javedankherad M, Zeinalpour-Yazdi Z, Ashtiani F. Cache placement phase based on graph coloring[C]2018 9th International Symposium on Telecommunications (IST). IEEE, 2018: 187-191.
- [4] Sun Y, Chen Z, Liu H. Delay analysis and optimization in cache-enabled multi-cell cooperative networks [C] 2016 IEEE Global Communications Conference (GLOBECOM). IEEE, 2016: 1-7.
- [5] Zhang K, Leng S, He Y, et al. Cooperative content caching in 5G networks with mobile edge computing [J]. *IEEE Wireless Communications*, 2018, 25(3): 80-87.
- [6] Ren D, Gui X, Lu W, et al. Ghcc: Grouping-based and hierarchical collaborative caching for mobile edge computing [C] 2018 16th International Symposium on Modeling and Optimization in Mobile, Ad Hoc, and Wireless Networks (WiOpt). IEEE, 2018: 1-6.
- [7] Tran T X, Hajisami A, Pompili D. Cooperative hierarchical caching in 5G cloud radio access networks [J]. *IEEE Network*, 2017, 31(4): 35-41.
- [8] Wang Y, Ding M, Chen Z, et al. Caching placement with recommendation systems for cache-enabled mobile social networks [J]. *IEEE Communications Letters*, 2017, 21(10): 2266-2269.
- [9] Bastug E, Bennis M, Debbah M. Living on the edge: The role of proactive caching in 5G wireless networks [J]. *IEEE Communications Magazine*, 2014, 52(8): 82-89.
- [10] Gu J, Wang W, Huang A, et al. Distributed cache replacement for caching-enable base stations in cellular networks [C] 2014 IEEE International Conference on Communications (ICC). IEEE, 2014: 2648-2653.
- [11] Wang S, Zhang X, Zhang Y, et al. A survey on mobile edge networks: Convergence of computing, caching and communications [J]. *Ieee Access*, 2017, 5: 6757-6779.
- [12] Robbins H, Monro S. A stochastic approximation method [J]. *The annals of mathematical statistics*, 1951: 400-407.
- [13] Pantisano F, Bennis M, Saad W, et al. Cache-aware user association in backhaul-constrained small cell networks[C]2014 12th International Symposium on Modeling and Optimization in Mobile, Ad Hoc, and Wireless Networks (WiOpt). IEEE, 2014: 37-42.
- [14] Osborne M J, Rubinstein A. A course in game theory [M]. MIT press, 1994.
- [15] Chen Z, Liu Y, Zhou B, et al. Caching incentive design in wireless D2D networks: A Stackelberg game approach [C] 2016 IEEE International Conference on Communications (ICC). IEEE, 2016: 1-6.
- [16] Hosny S, Alotaibi F, El Gamal H, et al. Towards a mobile content marketplace [C] 2015 IEEE 16th International Workshop on Signal Processing Advances in Wireless Communications (SPAWC). IEEE, 2015: 675-679.
- [17] Wu D, Huang J, He J, et al. Toward cost-effective mobile video streaming via smart cache with adaptive thresholding [J]. *IEEE Transactions on Broadcasting*, 2015, 61(4): 639-650.
- [18] Luo Z, LiWang M, Lin Z, et al. Energy-efficient caching for mobile edge computing in 5G networks [J]. *Applied sciences*, 2017, 7(6): 557.
- [19] Tran T X, Hajisami A, Pandey P, et al. Collaborative mobile edge computing in 5G networks: New paradigms, scenarios, and challenges [J]. *IEEE Communications Magazine*, 2017, 55(4): 54-61.
- [20] Rusek K, Cholda P. Message-passing neural networks learn little's law [J]. *IEEE Communications Letters*, 2018, 23(2): 274-277.
- [21] Shen F, Hamidouche K, Bastug E, et al. A stackelberg game for incentive proactive caching mechanisms in wireless networks [C] 2016 IEEE Global Communications Conference (GLOBECOM). IEEE, 2016: 1-6.
- [22] Jiang Y, Chen S Z, Hu B. Stackelberg games-based distributed algorithm of pricing and resource allocation in heterogeneous wireless networks [J]. *Journal of China Institute of Communications*, 2013, 34(1): 61-68.

- [23] Amer R, Butt M M, Bennis M, et al. Delay analysis for wireless D2D caching with inter-cluster cooperation [C] GLOBECOM 2017-2017 IEEE Global Communications Conference. IEEE, 2017: 1-7.
- [24] Srivastava N, Hinton G, Krizhevsky A, et al. Dropout: a simple way to prevent neural networks from overfitting [J]. *The journal of machine learning research*, 2014, 15(1): 1929-1958.
- [25] Goodfellow I, Bengio Y, Courville A, et al. *Deep learning* [M]. Cambridge: MIT press, 2016.
- [26] Ma N, Tian G D, Zhou X. A lip-reading recognition approach based on long short-term memory [J]. *J. Univ. Chinese Acad. Sci*, 2018, 35(1): 109-117.
- [27] Wang X, Han Y, Wang C, et al. In-edge ai: Intelligentizing mobile edge computing, caching and communication by federated learning [J]. *IEEE Network*, 2019, 33(5): 156-165.
- [28] Ramasubbareddy S, Ramasamy S, Sahoo K S, et al. CAVMS: Application-Aware Cloudlet Adaption and VM Selection Framework for Multicloudlet Environment [J]. *IEEE Systems Journal*, 2020.
- [29] Bastug E, Bennis M, Debbah M. Living on the edge: The role of proactive caching in 5G wireless networks [J]. *IEEE Communications Magazine*, 2014, 52(8): 82-89.
- [30] Wang X, Chen M, Taleb T, et al. Cache in the air: Exploiting content caching and delivery techniques for 5G systems [J]. *IEEE Communications Magazine*, 2014, 52(2): 131-139.
- [31] Su Z, Qi Q, Xu Q, et al. Incentive scheme for cyber physical social systems based on user behaviors [J]. *IEEE Transactions on Emerging Topics in Computing*, 2017, 8(1): 92-103.
- [32] Li Y, Xiao L, Dai H, et al. Game theoretic study of protecting MIMO transmissions against smart attacks [C] 2017 IEEE International Conference on Communications (ICC). IEEE, 2017: 1-6.