USER PREFERENCE PROFILE BASED CACHING FOR VIDEO STREAMING IN MOBILE NETWORKS

Jimy George[1], Shinto Sebastian[2]

[1]*PG Scholar, Department of Electronics and Communication Engineering., Amal Jyothi College of Engineering, Kerala, India*

[2]*Assistant Professor, Department of Electronics and Communication Engineering., Amal Jyothi College of Engineering, Kerala, India*

Abstract

In radio access networks, mobile devices access internet content through base stations, which are connected to the internet via gateway routers. A number of studies have been carried out to improve the performance of internet access by mobile hosts. Among these, caching of popular web documents at locations close to the mobile clients is an effective method. The benefits of caching data in the core network of mobile carrier have been investigated in the literature. There has been some research in caching of videos for online video sites like YouTube, which aims at improving the performance of the mobile internet access networks. Caching of Most Popular Videos (MPV) together with LRU (Least Recently Used) - replacing the least recently used video when cache is full- algorithm is an efficient method for large sized caches which are only possible for internet content delivery networks. Caching algorithms for ad-hoc networks has also been developed. These are not applicable for radio access networks. This paper investigates the effectiveness of placing caches at base stations in cellular networks and considers the UPP (User Preference Profile) of each cell-site for caching of videos.

Keywords: *Radio Access Networks, MPV, LRU, UPP*

--***--

1. INTRODUCTION

There is a tremendous growth in the use of smart phones and tablets for accessing internet content in past few years. Recent studies revealed that the number of mobile users, streaming videos from online video sites like YouTube is increasing. When internet video is accessed by a mobile device, the video must be fetched from the servers of a Content Delivery Network (CDN) [1]. Mobile phones are primarily connected to their respective base stations for communicating with each other. In LTE (Long Term Evolution) system, the base stations in Radio Access Networks (RAN), also known as eNodeBs connect with the User Equipment (UE) directly. Videos delivered from CDN reach UE via gateway routers which interconnect RAN and the Internet. Bringing each requested video in this manner incurs video latency and may lead to congestion in the wireless carrier core network. In this context, caching of most probable videos gained relevance.

Caching of popular web documents at proxy servers in computer networks was studied extensively in the literature. The relative frequency with which web pages are requested follows Zipf's law [2]. The Zipf's law states that the relative probability of a request for a web page is inversely proportional to its statistical rank. That is for i^{th} most popular page, the probability is proportional to 1/i. A number of works have been done on the popularity of online videos. Studies by M. Cha et al. in [5] revealed that video popularity also follows a Zipf distribution. It indicates that a mere 10% of the online videos have nearly 80% views, while the remaining accounts for only 20% of views. Caching of videos based on UPP calculates the probability of each video by considering the preference of users in a cell.

This paper presents a survey on different caching techniques for online videos developed so far, taking into account the location of caches and criterion for selecting videos to be cached. The effectiveness of placing video cache at each eNodeB in the RAN and caching policy based on UPP is also studied.

2. ANALYZING CACHING BENEFITS

To improve the performance of video delivery on radio access networks various caching techniques have been proposed. Some important properties of video traffic that have an impact on a cache are: size of a video, number of views for a particular video and inter-arrival time of requests for a video [4]. As referred from the literature, video popularity follows a Zipf distribution i.e., most popular videos contribute to only 10% of the total video population. Hence caches need to consider only those videos which have most of the views i.e., by caching only 10% of the total video content, most of the requests can be served. The number of videos that can be cached depends on the size of the cache. Most of the caching strategies in the literature prefer caching of video chunks video whereas some investigates caching of full videos. If only some parts of a requested video is in the cache, the rest need to be fetched from the main server. This may cause stalling/buffering during playback and may deteriorate the user's quality of experience. The technique involving caching of full videos limits the number of videos that can be cached and may lead to low cache-hit ratio. Cache replacement policies are also important for maintaining the cache up-to-date. Zink et al. in [5] propose least recently used replacement scheme: if the cache is full and a new video needs to be cached, the video that has been used least recently is evicted to create space for the new one. Caches need to be updated with the changes in the video popularity statistics for achieving a good cache-hit ratio. The experimental results in [5] show that cache-hit ratio improves with the cache size.

3. CACHING TECHNIQUES

With the growing interest in web videos, numerous strategies for video caching have been developed. Caching of frequently requested videos reduces the user perceived latency and server/wireless network loads. Proxy server architecture for content caching in high speed local area networks has been proposed in [3]. The architecture consists of a number of proxy servers, where one of these acts as a broker or central controlling agent and others act as sibling caching proxies. All the client requests are routed through the broker which is an enhanced RTSP proxy server.

The performance metric used to evaluate various caching techniques is cache-hit ratio. Whenever the requested video can be found in the cache, we say there is a cache-hit whereas its absence results in a cache-miss. If the number of videos stored in cache is large enough so that most of the requests can be served from the cache, high cache-hit ratio can be achieved. Hence for large sized caches the hit ratio will be large compared to small micro-caches.

Caching of Most Popular Videos (MPV) is a proactive caching policy, where most popular videos according to the available video popularity distribution are cached. In this case, cache update occurs only when the video popularity distribution changes. MPV technique is used in CDNs where caches are of large size.

Cache replacement policies play an important role in achieving high cache-hit ratio. When cache is full and a new content which is not in the cache and became popular within a short span of time needs to be cached, some content in the cache should be replaced. Cache replacement policies determine the content to be replaced to create space for new content. There are two categories of replacement policies: single-factor policies and multiple factor policies [6]. First-In-First-Out (FIFO) policy, random replacement policy, LRU, and LFU (Least Frequently Used) are some examples of single-factor replacement policies. Single factor replacement policies consider only a single factor such as popularity, age or cost for selecting the content to be replaced while multiple factor policies take into account a combination of these factors for selecting the content to be evicted.

LFU policy maintains a reference count for each object in the cache. When cache is full, the object to be replaced is selected as the one with the least reference count. If a tie occurs, LRU is applied which replaces the least recently used object. Another replacement policy discussed in the literature is Greedy Dual Size (GDS) [7] which combines temporal locality, size and other cost information. In [6], a new rank value based replacement algorithm is proposed. By considering the size, cost factor and age of the video along with the Zipf-like law for video popularity distribution, they have obtained better results compared with other algorithms.

A collaborative content caching algorithm for mobile ad-hoc networks is proposed in [9]. The algorithm consists of: initial selection of backbone nodes, cache placement policy and cache replacement policy. Selection algorithm selects between different nodes and forms a network of virtual access points. These nodes are responsible for caching video segments. They jointly decide which of the nodes should cache the requested video segments. Replacement policy used is LRU.

Video content delivery using coded distributed caching for wireless networks is discussed in [10]. The key idea is to use helper stations to cache video files and to deliver a requested video to UE via short-range wireless links. The files are encoded such that distributed storage is made possible and hence improving robustness and storage capacity. Helpers are placed uniformly in a square grid spanning the cellular region. The range of each helper is 100m.

Content caching at the base station of RAN is a way to reduce backhaul transmission and to improve the quality of experience. Reference [11] introduces an edge caching mechanism where caches are located at the edge base stations. These caches are fully interconnected and are able to disseminate contents via direct links. Thus all interconnected caches are considered as a single entity and duplication of contents in these caches can be avoided. This facilitates placement of other unavailable contents in the cache.

In [12], video caching problem in mobile networks is addressed where the caching nodes are distributed along with the mobile gateways. That is, caches are located in the core network of the wireless carrier. With collaborative caching, the cached nodes at gateways are interconnected. When a request for a video arrives at a gateway node, the node checks whether it has the video in its local cache. If the video is available in the cache, it is directly delivered to the client and if it is not there, the video is fetched from any one of its neighboring nodes. Each serving gateway has a cache associated with it and they are interconnected via Packet Data Network (PDN) gateways and possibly by other network routers.

4. UPP BASED CACHING

The popularity distribution of videos at national level does not reflect the local video popularity [5]. So in a mobile network the local popularity distribution of videos must be calculated on a cell-by-cell basis. The active users in a cell at any given time determine the probability distribution of videos in that cell. Hence a User Preference Profile (UPP) can be assigned with each active user which indicates the probability that a user request a particular video given all video categories. Active users hereby mean the mobile users who have watched a video when present in the cell or those who are having an active video session [1]. Since in LTE networks, the connection statistics of each UE is known to their respective eNodeBs the number of active users in a cell can be determined. Thus placing caches at each eNodeB in the RAN is beneficial in the sense that the cache serving each cell-site contains the most probable videos according to the UPP of users in that cell**.**

Overall video popularity distribution can be used to find the popularity of each video in different video categories. Once individual user preferences are available, the probability that a video being requested can be calculated. Based on the probabilities videos can be classified as Most Likely Requested(MLR) and Least Likely Requested(LLR) sets. Cache update is performed based on these probability values. The caching algorithm ensures that the cache memory space is always loaded with most probable videos.

5. SIMULATION RESULT

Simulations are performed using MATLAB. A reactive caching policy which updates cache only upon receiving requests from clients is simulated. The cache hit ratio obtained for cache sizes from 10GB to 200GB is evaluated and the result is compared against two conventional cache replacement policies – LRU and LFU. The results indicates that UPP based caching has better performance for all the cache sizes.

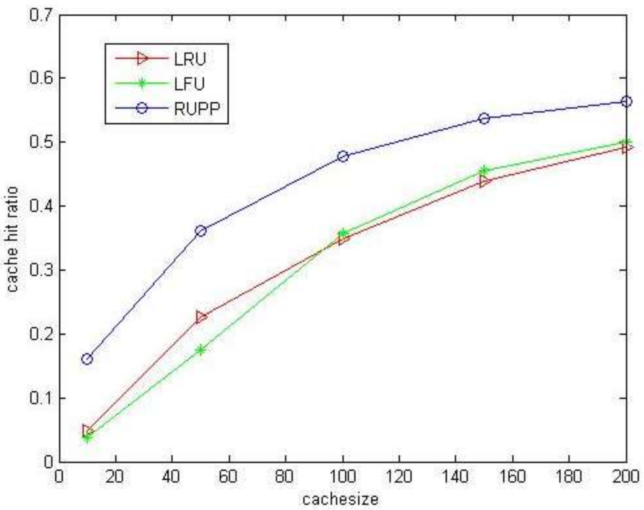

Fig-1: Cache hit ratio v/s cache size

6. CONCLUSION

With the increasing popularity of video streaming on mobile devices, new methods should be formulated to efficiently manage the RAN backhaul bandwidth and to provide the users with better quality of experience. Caching of most frequently requested videos at locations close to mobile clients is found to be an effective method to reduce the video latency. This paper surveyed on different caching policies adopted in the literature. Different criteria for the selection of videos to be cached and various cache replacement algorithms are discussed. User Preference Profile based caching will be useful as it considers the popularity of videos within each cell to select the videos to be cached and those to be evicted when the cache is full. If cache locations are at each edge nodes, video latency can be significantly reduced.

REFERENCES

[1] H. Ahlehagh and S. Dey, “Video Aware Scheduling and Caching in the Radio Access Networks,” IEEE/ACM Trans. Network., vol. 22, no. 5, Oct. 2014.

[2] Lee Breslau, Pei Cao, Li Fan, Graham Phillips, Scott Shenker,“Web Caching and Zipf-like Distributions: Evidence and Implications”,IEEE INFOCOM, vol. 9, no. 5, 1999.

[3] N. Laoutaris, “A closed form method for LRU replacement under generalized power law demand,” in *Proc. MMNC*, San Jose, CA, USA, Jan. 2008.

[4] S. S. Krishnan and R. K. Sitaraman, “Video stream quality impacts viewer behavior: Inferring causality using quasi- experimental designs,” in *Proc. ACM Conf. Internet Meas., Conf*., 2012, pp. 211-224.

[5] M. Cha, H. Kwak, and P. Rodrigues, “Analyzing the Video Popularity Characteristics of large scale user generated content systems,” IEEE/ACM Trans. Network., vol. 17, no. 5, pp. 1357-1370,Oct. 2009.

[6] G. Nair and P. Jayarekha, “A Rank Based Replacement Policy for Multimedia Server Cache Using Zipf-like Law,” *J. Comput.*, vol.2, no.3, Mar.2010.

[7] Breslau, L., Cao, P., Fan, L., Phillips, G., and Shenker, “Web Caching and Zipf-like Distributions: Evidence and Implications,” in Proc., IEEE/INFOCOM, USA,vol.1, pp. 126-134, Mar.1999.

[8] N.Wakamiya,M.Murata, and H.Miyahara,, “Video Streaming System with Co-operative Caching Mechanisms,”in*Proc.,SPIE Int. Symp.*,2002, pp.305-315.

[9] Y. Abdelmalek, A. Abad El Al, and T. Saadawi, “Collaborative Content Caching Algorithms in Mobile Ad-Hoc NetworkEnvironment,” *IFIP/IEEE Int. Symp. on Integrated Network Mgmt.*, 2009.

[10] NeginGolrezaei, KarthikeyanShanmugam, “Wireless Video Content Delivery through Coded Distributed Caching,” IEEE Comm. Theory, 2012.

[11] Kai Dong, Jun He, and Wei Song, “QoE-Aware Adaptive Bit-Rate Video Streaming over Mobile Networks with Caching Proxy,” *Int. Conf. on Computing, Networking and Communications.*, 2015.

[12] Yang Li, YuemeiXu, “A Novel Co-ordinated Edge Caching with Request Filtration in Radio Access Networks,” *The Scientific World Journal.*, vol.2, no.3, Nov. 2013.

BIOGRAPHIES

Jimy George is a PG scholar in Communication Engineering.

Mr. Shinto Sebastian is working as Assistant Professor in Electronics and Communication Engg. His area of specialization is Communication Engineering.