

Proxy Caching Algorithm based on Segment Group Popularity for Streaming Media

Yubin Wang¹, Yuhong Zhang² and Liwei Gu²

¹ College of Math and Information Science & Technology, Hebei Normal University of Science & Technology, Qinhuangdao, Hebei, 066004, China,

² College of Education, Hebei Normal University of Science & Technology, Qinhuangdao, Hebei, 066004, China,

qhdwyb@126.com, 13933690306@163.com, xiaxue826@sohu.com

Abstract

Multi-media is more and more popular in all levels of education for its vivid picture and video. Due to the increasing number of videos and pictures in school server, how terminal computers in classrooms access multimedia at low cost is an important problem to solve. A novel dynamic scheduling algorithm of PCASGP (Proxy caching algorithm based on segment group popularity for streaming media). Based on the popularity of the segment group, content of proxy caching would be replaced and data size would in proportion of segment popularity. The cache window size would be updated dynamically with the access time based on the algorithm. Results of simulation shows that the algorithm is quite adaptable and can gain more average number of streaming media objects with less delayed access ratio.

Keywords: multi-media education system; proxy caching; caching algorithm; segment group popularity

1. Introduction

Multi-media is media and content that uses a combination of different content forms. It is a combination of text, audio, still images, animation, video and so on. Multimedia presentations can be live or recorded and it may be made manually. For example, animation is made by person to show the original form of excise. For this reason, multimedia has been used in teaching more and more widely. Its animation vividly teaching students many more motion happened in nature.

Multi-media has been used in education for years. There are also a lot of studies on it. In the beginning, both internal and external problems bedevil most researchers and teachers [1]. With time going on, Multi-media education system has its own classification: linguistic visual, nonlinguistic visual, linguistic audible, and nonlinguistic audible [2]. Recent years, the develop direction has been studied [3]. Now, multimedia is used in nearly all the curriculums in school, Not only in computer courses, but also other courses [4-8].

Multimedia education is also used in online courses [9-11]. People with no teachers or with teachers on internet can study by themselves with the online multimedia lessons. With this condition, teaching assessment is needed to assess the effect of multimedia or online teaching. There are also lots of researches on this subject [12-14].

With development of distributed system, education system is integrated into school network and terminal computer as a conception has been used in education. In university library, there are many kinds of multimedia files and lessons. Many teachers use them directly

in the class. So, it can use steaming multimedia conception to improve efficiency of server. Computers in classrooms or other place can fetch resource from server. The system is showed in Figure 1. As the number of multimedia files becomes bigger and bigger and different files of each terminal computer access, the proxy cache capacity would be very large to keep the usual study and teaching. If we develop proper algorithm, we need not to keep the large proxy cache capacity and we would also cut cost. That would be worth to promote.

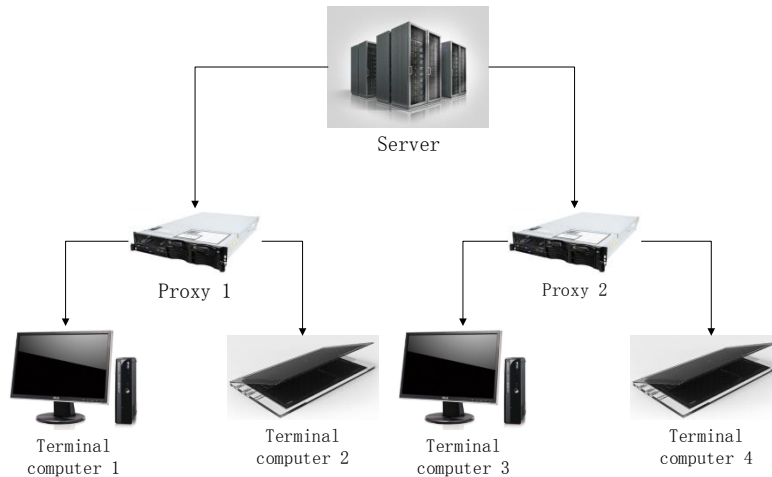


Figure 1. Teaching system in school

Many proxy cache algorithms has been proposed, such as prefix merging [15], uniform segmentation [16], exponential segmentation [17], layered media caching [18], multiple version caching [19], batching [20], patching [21], stream merging [22-25], and so on. In teaching system, proxy server pre-stores prefix part of multimedia at certain time. When class begins, terminal computer request the media file from proxy and server. Usually, proxy deliver the file to terminal computer with first delivering prefix. This method can decrease the time delay in terminal computer to keep normally teaching. In addition of patching and steaming merging technology, more terminal computers can share a same multicast streaming to effectively reduce the bandwidth and server resource consumption but to get better performance. When one video has a high access frequency, algorithms described before would still consume high bandwidth and has little dynamic adjusting capability.

The algorithms described before just can be used in off line state, but how to effectively use in proxy server caching condition is still to be resolved. The paper attempts to develop a new access algorithm between terminal computers and sever. This algorithm would be called PCASGP. Existing technology analysis is presented in Section 2. The PCASGP algorithm would be presented and discussed in Section 3. Simulation and results would be showed in Section 4 and concluding remarks are then given in Section 5.

2. Existing Research Work

Traditional cache algorithm, such as LFU (least frequency used) [26], LRU (least recently used) [27], LRU-K [28], and so on. LFU algorithm often replaces the object with least frequency and value of object is high with the high frequency. This algorithm also has problem of cache pollution, which means the object still occupies the cache without being used. LRU usually replace the multimedia, which has no access for a long time, but this algorithm has problem of being replaced and got access immediately. Algorithm LRU-K

although get better performance according to combine the access frequency and access time. But it is still has problems as the same as that of LRU and LFU.

A2LS algorithm concentrates on the delay of segment dividing of multimedia object. The time length of segment is determined by the access of terminal computer. In addition, based on the utility function proposed, segments of media can be adaptively received and replaced. But the utility function doesn't refer the distance of segment being access with the beginning of initial position of multimedia file. The priority of segment cache would decrease with the distance becoming bigger [29]. The less distance, the bigger probability to be accessed. In the algorithm, all the segments would be stored in the cache, this would waste the memory. Although the algorithm delay the segment dividing as much as possible, the media once begins to divide, the length would be keep constant without variation. So, it may not meet the demand of terminal computers

3. PCASGP

In education system, there are many kinds of multimedia, such as picture, audio and video. Each kind of multimedia need different caching method. If cache stores whole file and keep the file synchronize with server, it would occupy lots of memory. This strategy are usually used to store text and picture. With large memory of video, the cache can't store the whole file due to the limit capacity of proxy cache. In an internet research, 80% people would access to visit 20% multimedia [30]. Usually, in classes, teacher would also to show some video with hot degree to students, so, we can believe that there may be sometime lots of teachers or students would choose the video to watch. At the same time, access characteristic is also in variation. For example, an important news is appeared, and many teacher would show the news to students, this would cause the network to block. But after one day or two, the news is out of date, there may be no person to visit it. Then, it needs the admission and re-placement of objects.

PCASGP includes proxy caching admission algorithm and proxy caching replacement algorithm for multimedia. If assume the block is the smallest data unit, segments is made up of different and multimedia is composed of segments. In cache, a kind of data log should be established and preserved for each multimedia object. The main parameters in the log are:

T_{first} : accessing a multimedia object for the first time;

T_{last} : last time to access the object;

T_{sum} : total accessing time length of multimedia object;

N : accessing times to multimedia object;

L_s : time length of each segment;

T : total time length of multimedia object;

B : time length of each data block;

CSN : number of segments;

$Flag$: whether to cache completely, and "1" represents total cache while "0" represents partly cache.

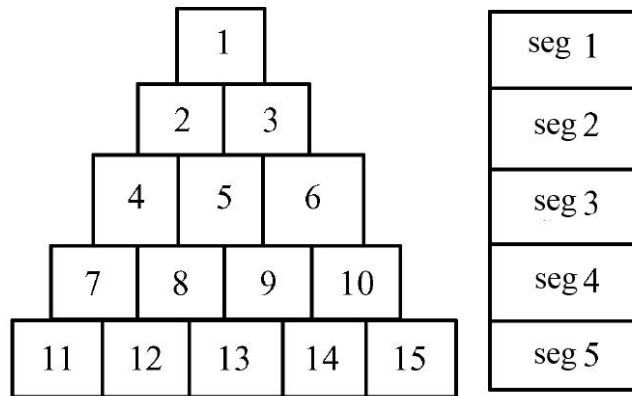


Figure 2. Segment dividing

Make some assumptions. Cache window size is w . Stream multimedia proxy cache size is PCS. Initial value of w is $T/integer$, where $integer = 1, 2, 3, 4, \dots$. In the paper, we set $integer$ to be 2. At a certain time t , $L_{avg} = T_{sum}/N$ represents statistics of access mean time length at that time. L_{avg} can do the statistics periodically or based on the access of terminal computers. Segments are the base and compose of N_s blocks. $L_s = N_s B$, and $S = N_s$, $S > 0$. This is to say that there is one block in segment 1, there are two blocks in segment 2, three in segment 3, and so on. Size of segment is relative with the location. As shown in Figure 2, further with the beginning, bigger the size would be. This is because of the beginning of the multimedia would be frequently requested and most of the postfix can't be played over. The other hand is to decrease the segments in cache.

3.1. Cache admission algorithm

In education system, courses are usually given at the same time of each year. So, the lesson would be classified with different content. If a lesson is been request, all the media prefix of the whole class lesson would be stored in proxy cache. Size of prefix of each media can be determined with network between media server and proxy. Time delay would be reduced by sending the prefix to terminal computer first. For the reason of little size of prefix, so, once it is in cache, it would be stored in proxy. If any lesson of whole class has no access for a long time, they would be replaced with the other class. If one prefix didn't get access for years, when the next round comes, the prefix would not be stored in the proxy.

Postfix is defined in this paper to be the all the content without prefix of one multimedia. We would define prefix here.

$$prefix = \beta \cdot B, \text{ where, } \beta > 0 \quad (1)$$

When each access has come, the proxy would execute the admission algorithm:

(1) If one multimedia accessed has no log before, that is to say that it is the first time to access. Time length would be calculated and sent the result to proxy. Enough space in cache would be allocated to the prefix of the accessed multimedia. Amount time of W of the multimedia would be stored in the cache. Initial value of W would be set as half of the media. After for a while, if there is terminal computer access the object, L_{avg} would make the

statistics. If $L_{avg} < T/2$, $W = 2 \cdot L_{avg}$, or else, $W = L_{avg}$. The access log would be stored and updated with the access.

(2) If one multimedia has access log, that is to say that it has been accessed before. If whole media is stored in proxy, there would be no admission on this multimedia. Then, it just need update the log.

(3) One multimedia has access log, but there are some segments has not been stored. If average access time increases and segments stored can't meet the demand of access, it should be considered that whether to store the next segments. Here, we would store all the segments of the multimedia. In fact, there are some algorithms to resolve the problems.

3.2. Cache replacement algorithm

Proxy server plays an important role in data transmission between server and terminal computers. It pre-store prefix and some segments of multimedia to reduce the time delay and keep the multimedia playing smoothly. Proper algorithm is quite important in data replacement in proxy. Once there is not enough proxy cache, cache replacement algorithm would be used in proxy. The main idea of the algorithm is:

(1) When access one multimedia object for the first time, popularity value would be calculated to determine which segments is the least popular. In all the multimedia objects without any access of terminal computers, least popular segments of multimedia would be removed from proxy. If the space is not enough, second least popular segments would be removed too. As the some reason, segments would be removed until enough space to store the new segments.

(2) If one multimedia object is accessed before and some segments without storage in proxy have higher popularity than segments with least popularity in storage, segments with least popularity would be replaced. If there is still not enough space, compare the popularity with that of second least one. If the second least one is bigger, the cache would be dropped. Or else, the second least one would be replaced. As the same method, until the cache has enough space or replacement is dropped.

The popularity of multimedia would be defined as the following:

$$popularity = \frac{\frac{T_{sum}}{T_{last} - T_{first}} \times \min\{1, \frac{(T_{last} - T_{first}) / N}{T_c - T_{last}}\}}{H} \quad (2)$$

The popularity of multimedia segments would be defined as the following:

$$U = popularity \times \frac{1}{D(s)} \quad (3)$$

Where, T_c represents current time; H is cache data size of some certain multimedia at current time, it also represents the storage cost in disk; $D(s)$ is an distance function of s . The bigger value of s , the bigger value of $D(s)$. The paper would select $\log(s)$, $\min\{1, \frac{(T_{last} - T_{first}) / N}{T_c - T_{last}}\}$ presents the possibility of subsequent access. Cache popularity function has great influence with many indexes, such as byte hit ratio, and so on.

Flowchart of PCASGP is shown in Figure 3.

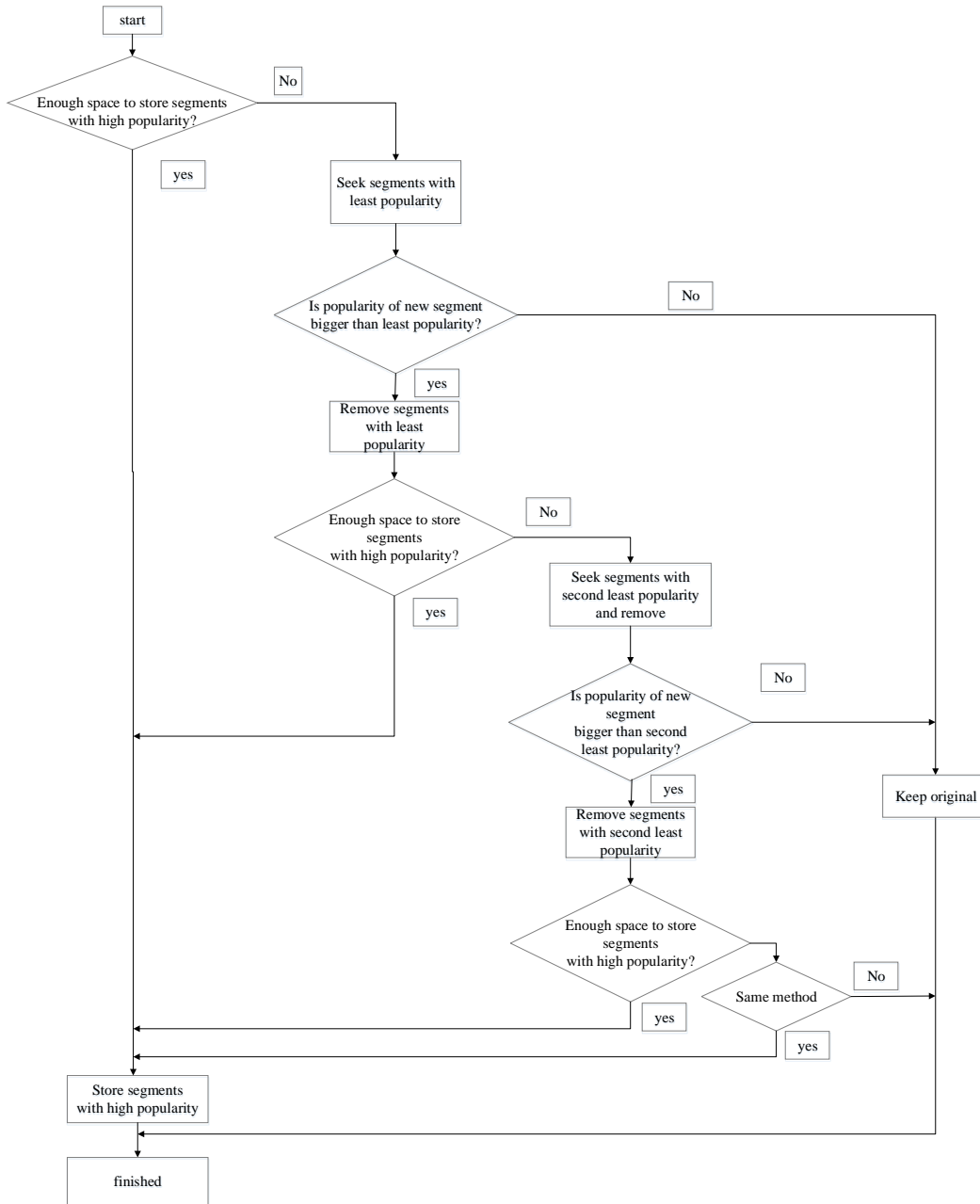


Figure 3. Flowchart of PCASGP

When multimedia object is accessed for the first time, the algorithm would set the window size to be half of object size and replace respond other multimedia object to cache half size. This is because the popularity won't change with the second access. So, in the algorithm proposed, the multimedia need at least two access. This is to say that the values of T_{last} , T_{first} , T_c . If segment S of multimedia i has been accessed for more than two times and its

popularity value is more than multimedia Q , last segment of Q would be replaced. If last segment of Q is bigger than that of segment s of multimedia i , segment s would be cached, or else, it would not be cached. After multimedia object has been replaced, it still may be accessed by terminal computers, so, we would make a decision that once the multimedia object has been cached, first segment of object would be cached in the proxy to keep the popularity value. At the same time, the system would store the access log of multimedia objects. If the log has not been kept, the algorithm can't judge whether the first time to access the multimedia. Because with the development of time, the popularity would decrease, if it is judged to be the first access, the cache storage efficiency would be very low. When multimedia object is accessed for the first time, the algorithm would set the window size to be half of object size and replace respond other multimedia object to cache half size. This is because the popularity won't change with the second access. So, in the algorithm proposed, the multimedia need at least two access. This is to say that the values of T_{last} , T_{first} , T_c . If segment s of multimedia i has been accessed for more than two times and its popularity value is more than multimedia Q , last segment of Q would be replaced. If last segment of Q is bigger than that of segment s of multimedia i , segment s would be cached, or else, it would not be cached. After multimedia object has been replaced, it still may be accessed by terminal computers, so, we would make a decision that once the multimedia object has been cached, first segment of object would be cached in the proxy to keep the popularity value. At the same time, the system would store the access log of multimedia objects. If the log has not been kept, the algorithm can't judge whether the first time to access the multimedia. Because with the development of time, the popularity would decrease, if it is judged to be the first access, the cache storage efficiency would be very low.

4. Simulation Results

The so called byte hit ratio is a ratio between user get mean byte array from high speed proxy cache and that from server. The ratio becomes higher with the flow between network and server decreases. If one request is delayed, the main reason is the segment to access is not cached in proxy. The ratio between delayed request and total request can be used to assess the delay degree of initial access. If this ratio is very little, users would find it nearly has no delay. Number of cached objects in unit time means total cached and partly cached objects in unit time. It is also an important index in assessing whether it is good for caching prefix of big amount of multimedia objects or caching popular segments of little amount of objects. Simulation results of PCASGP would compared with A²LS. The two algorithm would give same condition. Both algorithms demand the multimedia accessed at least two times and if there is not enough space and with just one access, the new segment would not been cached. This is because of one time access of multimedia would fail all the access logs and not enough space can't cache the new segments.

Firstly, we should make some assumptions as the following:

- (1) There are 100 streaming multimedia objects in server, access possibility is followed by Zipf distribution with parameter $\theta = 0.271$ [31];
- (2) The streaming multimedia objects code with MPEG-1 with the transmission speed of 1.5Mbit/s, space size is in 300MB~800MB and time length is 60min. Play end position is in evenly distributed and VCR operating influence is neglected;

(3) Access would adopt two kinds of low intensity and high intensity. Low intensity access arriving follows the Poisson distribution of mean speed λ equal to 10 times per minute and high intensity access arriving follow the passion distribution of mean speed λ equal to 60 times per minute;

(4) Total access number would be 3000 or 18000 with the access time of 300 minute.

The simulation experimental device is listed in Table 1.

Table 1. Simulation environment

CPU	Memory	Hard disk	Operating system	Simulation too
Xeon 2.4GHz	8GB	750GB	Windows 7 × 64bit	Matlab 2010b

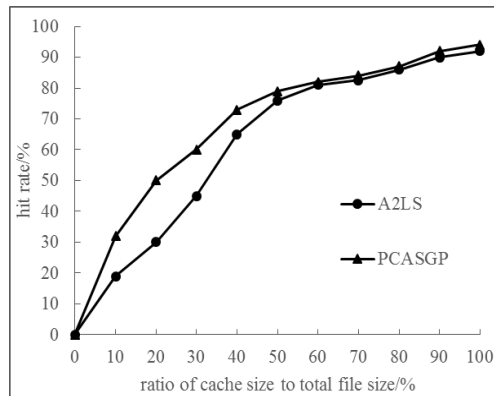


Figure 4. Hit ratio ($\lambda = 10$)

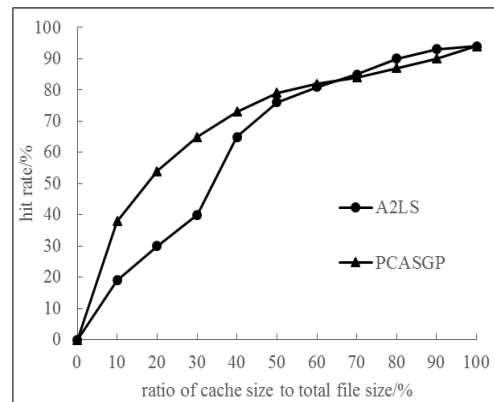


Figure 5. Hit ratio ($\lambda = 60$)

Figure 4 and Figure 5 shows the byte hit ratio with different access intensity. Ratio of the two methods would be bigger with the size increasing of proxy cache. Ratio under PCASGP increases faster than that of under A2LS. When proxy cache size is less than half of total objects size, byte hit ratio under is much higher under PCASGP than that under A2LS. When proxy cache size is more than half of total objects size, ratio under both algorithm would be quite close.

Figure 6 and Figure 7 shows the delayed request under different access intensity with both algorithms of PCASGP and A2LS. Initial request delayed ratio under A2LS would decrease

with the proxy cache size becomes bigger. In algorithm of PCASGP, time delay would be diminished for the reason of all the prefix were stored in proxy cache.

Figure 8 and Figure 9 show the cached mean number of multimedia objects in different request intensity. Under algorithm of A2LS, mean number of cached objects would be increase with the cache size becomes bigger; Under PCASGP, because of all the prefix of multimedia objects are cached in proxy, mean number of cached media object would keep constant.

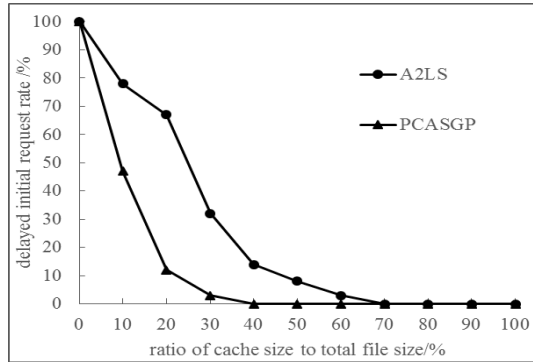


Figure 6. Delayed initial request rate ($\lambda = 10$)

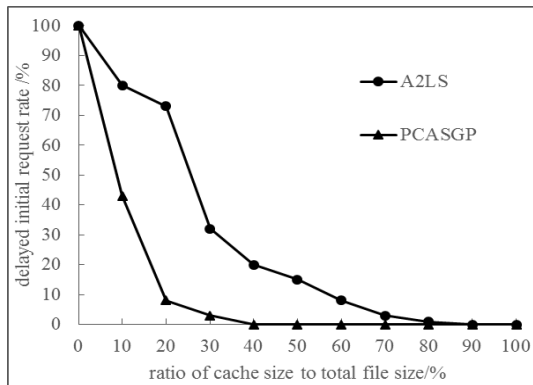


Figure 7. Delayed initial request rate ($\lambda = 60$)

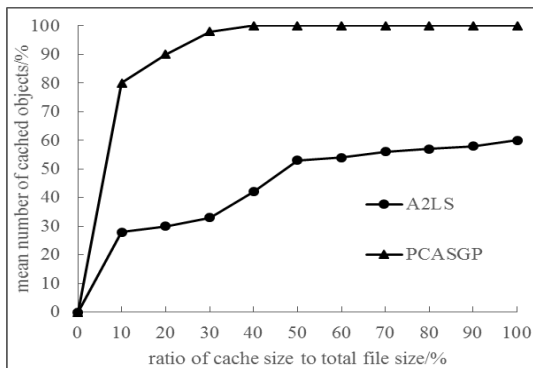


Figure 8. Mean number of cached objects ($\lambda = 10$)

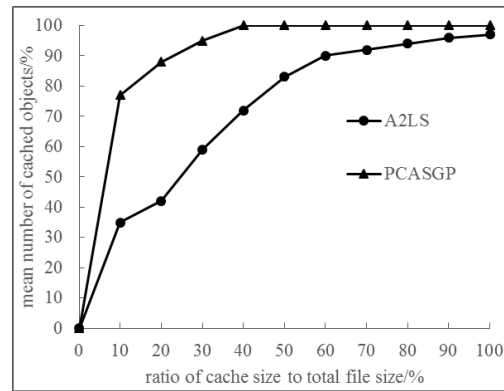


Figure 9. Mean number of cached objects ($\lambda = 60$)

5. Conclusion

In education network system, vast amount of multimedia would be stored in server. And terminal computers in classes or other places would access to the files. In order to improve the efficiency and decrease the time delay, a proxy would be set between server and terminal computers. The paper proposes an improved algorithm of PCASGP to improve the efficiency of multimedia objects distribution, reduce the occupied bandwidth. The algorithm is based on segment group popularity. The cache size of multimedia objects would increase with increasing of popularity value. The replacement would be executed or dropped by whether to have enough space to store the segment with high popularity. Space would be free by comparing the popularity value. If popularity of new segment is bigger than that of least popular segment in cache, the memory which store least popular segment would be free, if the space is not enough, the second least would be compared with the same method until the space is big enough. If the free space is not enough and there is no less popular segment than the new segment, the replacement would be dropped. There is also problem with the algorithm. If there are not enough space to store the new segment and cached segment with least popularity is removed from the cache, the total efficiency of the proxy cache would be decreased. So, it is a problem to resolve.

References

- [1] P. R. Voumard, "Multi-media education - student and staff issues", IEEE International Conference on Multi-Media Engineering Education – Proceedings, (1996) July 3-5, pp. 319-324, Melbourne, Australia.
- [2] M. Tanaka, "Study on the multi-media education system", Proc 1988 Int Conf Syst Man Cybern, (1988), pp. 934-937.
- [3] Y. Hu and W. Gui, "Multi-media Technology in education: where are they come and where are they go? Proceedings of the 2007 1st International Symposium on Information Technologies and Applications in Education, (2007) November 23-25, Kunming, China.
- [4] H. Bin and L. Jian "Research on application of multi-media technology in colleges' table tennis course teaching" 2008 IEEE International Symposium on Knowledge Acquisition and Modeling Workshop Proceedings, (2008), pp. 1136-1137.
- [5] Y. Yao and J. Wang and Z. Li, "A study of the curriculum setup of web-page designing in multi-media major" 2011 IEEE 3rd International Conference on Communication Software and Networks, ICCSN 2011, (2011), pp. 254-255.
- [6] Z. Z. Feng, "On organic combination of multi-media information technology and college physical education teaching", Lecture Notes in Electrical Engineering, vol. 163, (2013), pp. 255-260.

- [7] Z. Sun, M. Yang and Y. Song, "A case study of multi-media in college english teaching. Proceedings - 2009 2nd IEEE International Conference on Computer Science and Information Technology, ICCSIT 2009, (2009), pp. 94-97.
- [8] F. Shao, "A research on the College English teaching pattern based on the multi-media & the internet", 2nd International Workshop on Education Technology and Computer Science, vol. 3, (2010), pp. 640-632..
- [9] C. Luman, "Study of online multi-media learning environment design", 2011 IEEE 3rd International Conference on Communication Software and Networks, ICCSN 2011, (2011), pp. 754-757.
- [10] G. Sorin, and R. Terry, "Teaching online", Association for Information Systems - 12th Americas Conference on Information Systems, AMCIS 2006, vol. 4, (2006), 2198-2206.
- [11] C. Baiyun, V. Dale and R. D. Weaver, "Improving professional development for teaching online", vol. 2, no. 1, (2006), pp. 303-308.
- [12] D. Yuanwei, H. Chen, Y. Hongjuan and D. Wanchun, "All-round evaluation method for multi-media teaching", Journal of Software, vol. 8, no. 3, (2013), pp. 701-708.
- [13] S. Xiaoying, W. Zhongjie, S. Jiafeng and T. Jingna, "Discussion on multi-media teaching quality appraisal system", 1st International Conference on Information Science and Engineering, ICISE 2009, (2009), pp. 3287-3290.
- [14] W. Xiaojing, "On how to efficiently carry out multi-media teaching", Advances in Intelligent and Soft Computing, vol. 109, (2011), pp. 319-323.
- [15] S. Sen, J. Rexford and D. Towsley, "Proxy Prefix caching for multimedia streams", In Proceeding of IEEE INFOCOM, vol. 3, (1999), pp. 1310-1319.
- [16] R. Rejaie, M. Handley, H. Yu, *et al.* "Proxy caching mechanism for multimedia Playback streams in the internet", In Proceedings of International web Caching Workshop. (1999) March, San Diego, CA.
- [17] K. Wu, P. S. Yu, J. Wolf, "Segment-based proxy caching of multimedia streams", In Proceeding of WWW, (2001) May, Hong Kong, China.
- [18] R. Rejaie, M. Handley, H. Yu, *et al.*, "Multimedia proxy caching mechanism for quality adaptive streaming applications in the internet", In Proceeding of IEEE IN FOCOM, (2000) March, Tel-Aviv, Israel.
- [19] T. Kim and M. H. Ammar, "A comparison of layering and stream replication video multicast schemes", In Proceedings of ACM NOSSDAV 2001, (2001) June, NY.
- [20] C. Aggarwal, J. Wolf and P. Yu, "On optimal batching policies for video on demand storage servers", In Proc. IEEE Int. Conf. Multimedia Computing and systems, (1996), pp. 253-258.
- [21] K. Hua, Y. Cai and S. Sheu, "Patching: a multicast technique for true video on demand services", In Proc. ACM Multimedia, (1998), pp. 191-200.
- [22] C. C. Aggaral, J. L. Wolf and P. S. Yu, "On Optimal piggyback merging policies for video-on-demand systems", Proc. ACM SIGMETRICS Conf., (1996) May, Philadelphia, PA.
- [23] D. L. Eager, M. K. Vermon and J. Zahorjan, "Minimizing bandwidth requirements for on-demand data delivery", In Proceedings of the 5-th. International Workshop on Advances in Multimedia Information System (MIS'99), (1999), pp. 80-87.
- [24] B. N. Amotz, Ladnerre, "Efficient algorithms for optimal stream merging for media-on-demand", SISIM J. Comput., vol. 33, no. 5, (2004), pp. 1011-1034.
- [25] D. Eager, M. Vernon and J. Zahojan, "Optimal and efficient merging schedules for video-on-demand servers", In Proc. ACM Multimedia, (1999), pp. 199-202.
- [26] J. T. Robinson and M. V. Devarakonda, "Data cache management using frequency-based replacement", Proceedings of SIGMETRIC on Measuring and Modeling of Computer Systems, (1990), pp. 134-142.
- [27] J. Alghazo, A. Akaaboune and N. Botros, "SF-LRU Cache Re-Placement Algorithm", Records of the 2004 International Workshop, (2004), pp. 19 -24.
- [28] W. S. Seung, Y. K. Ki and S. J. Jong, "LRU based small latency first re-placement (SLFR) algorithm for the proxy cache", Proceedings of IEEE International Conference on Web Intelligence, (2003), pp. 499-502.
- [29] S. Q. Chen, B. Shen, S. Wee, *et al.* "Adaptive and lazy segmentation based proxy caching for streaming media delivery", Proceedings of the 13th International Workshop on Network and Operating Systems Support for Digital Audio and Video, (2003), pp. 22-31.
- [30] K. L. Wu, P. S. Yu and J. L. Wolf, "Segmentation of multimedia streams for proxy caching", [J]. IEEE Transactions on Multimedia, vol. 6, no. 5, (2004), pp. 770- 780.
- [31] B. Wang, S. Sen, M. Adler, *et al.* "Optimal proxy cache allocation for efficient streaming media distribution", Proceedings of IEEE INFOCOM 2002, (2002), pp. 1726-1735.

Authors



Yubin Wang

He received his bachelor's degree of engineering in Yanshan University, Qinhuangdao, Hebei. (2000) and master's degree of engineering in Yanshan University (2009) , Now he is a lecturer in Hebei Normal University of Science & Technology, Qinhuangdao,Hebei. His major fields of study are computer network technology.



Zhang Yuhong

She received her bachelor's degree of education in Hebei Normal University of Science & Technology, Qinhuangdao, Hebei. (2001) and master's degree of education in Hebei Normal University, Baoding, Hebei. (2007), Now she is a lecturer in Hebei Normal University of Science & Technology, Qinhuangdao,Hebei. Her major fields of study are education technology, vocational education, and information education.



Gu Liwei

She received her bachelor's degree of education in Hebei Normal University Shijiazhuang, Hebei, (2002) and master's degree of Software Engineering in Beijing University of Technology, Beijing, (2010) , Now she is a lecturer in Hebei Normal University of Science & Technology, Qinhuangdao, Hebei. Her major fields of study are education technology, distance learning, software engineering and education.