



A Novel Hybrid Caching Algorithm for Streaming Media Network

Saravanan.C¹, Marimuthu.G², Immanuel Alphonse³, Kaliaperumal.G⁴

Assistant Professor, Department of Information Technology, Krishnasamy college of Engineering & Technology,
TamilNadu, India^{1,2}

Assistant Professor, Department of Information Technology, Tagore Engineering College, College, Chennai, India³

Assistant Software Engineer, Tata Consultancy Services, Chennai, India⁴

Abstract: In this work, a hybrid caching mechanism is implemented to improve the quality of service in online video streaming. Even though, the advancement in multimedia is streaming; the user experience latency during online video streaming. To avoid delay, a tunable victimization scheme is used to efficiently allocate the streaming video. The streams are split to minute fragments in order to provide space for number of different streams. The Fragment is that the amount of cached data is adjusted up to finer unit block. Here, both the average arrival time (AAT) and playback rate of the video are defined as hybrid scheme (ATPR). The victim object is selected based on the ATPR value. The lowest ATPR video is replaced with requested video. A log is maintained about the details of requested video. Performance evaluation demonstrates that our scheme outperforms other traditional schemes in terms of the different quality-of-service parameters.

Keywords: AAT, ATPR, Fragments, Hybrid scheme, Playback rate

I. INTRODUCTION

The major concern with regard to the streaming applications is that they consume complete bandwidth along the client-server path for the entire session. The multimedia content when maintained closer to the interested clients, it can contribute in high scalability and deliver high quality streams. The delivery of high quality streams over the internet is possible through Proxy caching, which is a client-oriented solution. The working principle behind the Proxy cache is that it stores the frequently accessed content with the hope that it can positively respond to the client requests. Doing this, it reduces the load of the server, wherein, instead of requesting the server directly, the proxy cache would respond to the request if the requested response is with it. Due to the closer location of the proxy cache to the client side, it would reduce the service response time and thereby avoid network bottleneck. The research of streaming media caching technology is a challenging.

II. RELATED WORK

Several researchers have proposed many proxy caching schemes to utilize the cache efficiently. A randomized algorithm [5] which considers the bin division for replacement strategy. Proxy server with transcoding capability [11] provides appropriate video quality for the

different network environment. Demand on the interactive request of the multimedia object is increased significantly, [14] and a DISC algorithm is introduced which satisfies the client partially. The prefix/suffix proxy-caching schemes [1], [4] split the streaming objects into two units namely, prefix and suffix. The prefix part is used to start the streaming pipeline. Client experience the delayed start because, the prefix part of the requested object is not cached. The problem here to cache the whole part or the starting part of the object in the proxy cache. Proposed [13] the caching scheme that differentiates the content-type of the documents and uses a different queue for each of them. In that way, the replacement policy that better performance is achieved for each content-type and cache size can be applied. Two-stage replica replacement algorithm [12] applies good balance between value and cost by predicting replica value to make sure which replica will be replaced, and predicting the replacement cost to make it as low as possible.

The caching schemes [8], [9] propose to determine which video to be cache for layer encoded video. Content Delivery Networks (CDN) have been used [15] on the Internet to cache media content so as to reduce the load on the original media server, network congestion, and latency. Combined the LRU and LFU Policies [16] scheme



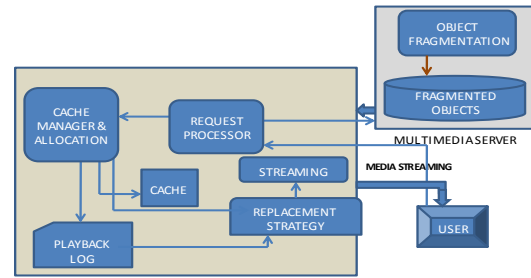
switching between the two policies according to the user pattern. These caching policies give preferential treatments to the beginning segments and aimed to minimize proxy jitter and analysing the performance of the byte-hit ratio. The framework [10] proposed with optimal cache management achieves reasonable reduction on bandwidth and also enable flexible utility management for the heterogeneous clients. The paper [6] discuss on the new BSC strategy for video-on-demand-like (VoD) applications and to cache the initial portion of the stream to avoid start up delay respectively. The algorithm [7] achieves a good balance between value and cost by predicting replica value to make sure which replica will be replaced, and predicting the replacement cost to make it as low as possible.

To represent the drawback in all the algorithm proposed, we propose a new hybrid algorithm which takes the account of average arrival time(AAT) and playback rate of the video are defined as hybrid scheme (ATPR). The lower ATPR multimedia object is evicted from the proxy cache to house the higher ATPR object. Performance evaluation demonstrates that our scheme outperforms other traditional schemes in terms of the different quality-of-service parameters.

III. PROPOSED MODEL

The multimedia server consists of object fragmentation unit which fragments the multimedia objects to finer unit block. Figure 1 represents the multimedia proxy caching architecture. The advantage of Fragment scheme is that the amount of cached data is adjusted up to finer unit block. This improves more object is placed in cache and by reducing initial start-up delay for the users. The advantage of the fragmental scheme is when the user request for the object the request is forwarded towards the proxy server in which the incoming request is handled by the Request Processor and then forwards the request to the Cache Manager and Allocation, the cache manager is the main unit which manages and processes the request and decides how much space to be allocated for the incoming requests. The playback log is the unit which stores the history of The arrival time of each and every requested object.

The replacement strategically unit takes in account both the playback rate and the arrivaltime of each and every object to replace. The object with less playback rate and arrival time is considered as the victim. After victim the requested object is streamed from the server to the cache manager. The cache manager allocates the object to the cache and the object is streamed to use.



IV. SYSTEM IMPLEMENTATION

A. Object Division and streaming

The multimedia objects are divided into several processing units. Each processing unit consists of Cfragment and s fragment where the Cfragments are cached in proxy cache and the Sfragment are streamed towards the client end while the Cfragments are playback. The advantage of the fragmentation scheme is that the Sfragments are parallel downloaded while the Cfragments are in playback.

B. Fragmented proxy

Generally the proxy cache consists of objects which are fragmented of varying size. Figure 2 shows the objects are fragmented and the initial part of the object is prefix. Whenever the object is requested the prefix part is cached. If the same object is requested for the next time, due to caching of prefix part the initial start-up delay is totally reduced. For replacing the victim the last part of the object is replaced

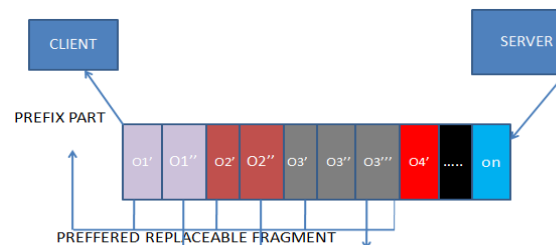


Fig. 2 Proxy Cache with Fragmented Objects

C. Calculating Average Arrival Time

The parameter arrival time (LRU) for every object is calculated for certain duration. Whenever the object is requested the corresponding time is stored in persistent storage. Several request for the same object's time is stored in corresponding persistent value. Figure 3 represents the arrival time of several multimedia objects. Let us consider the multimedia objectOk and let the several arrival time or the request time of the object Ok is tk1, tk2, tk3. The Average arrival time (AAT) for a particular object Ok is calculated as,



$$AAT_k = \frac{t_{k1} + t_{k2} + t_{k3}}{3}$$

D. Defining ATPR

By calculating the Average arrival time of each objects and defining the playback rate of corresponding object we define a Hybrid scheme (ATPR) with the combination of the arrival time and the playback rate. The victim object is selected based on the ATPR value. This improves the QOS parameters by retaining the familiar objects and reducing the network bandwidth as well as reducing the delayed starts

E. Cache space allocation

One important task is to allocate the required space in cache when the new object is requested. When the object is requested the object must be cached full or partial in the proxy cache or the object is not cached either of the reason of new request or may be the replacement victim of the previous requests. Here, according to our study the objects cached in the proxy must be proportional to the playback rate and the average arrival time of the particular object (ATPR). This scheme is termed as Hybrid scheme. When the requested object is to be cached in the proxy cache we need to find how much of the object i.e. the blocks to be cached in the proxy cache.

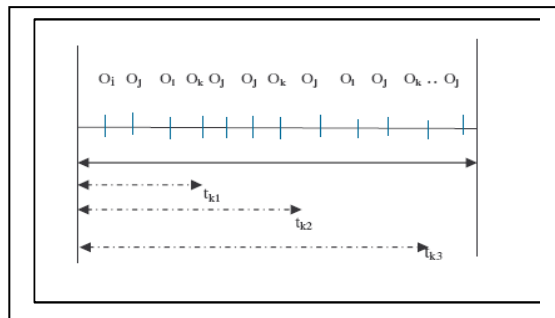


Fig. 3 ArrivalRequestPattern forMultimediaObjects

A simple and effective approach is been used to identify the number of C-blocks required for the requested object O_i , we specify an object O_j in the playback file that $ATPR_j - ATPR_i$ is minimum. Let both the objects O_i and O_j in the proxy cache and the number of C-blocks is C_i and C_j respectively. If the number of C-blocks in O_i is lesser than the C-blocks in O_j , the number of C-blocks for object O_i is gradually added to equal the number of C-blocks in object O_j and if the number of C-blocks in O_i is greater than the number of C-blocks in O_j , a single C-block is added to the O_i in the proxy cache. The C-blocks to be increment in object O_i for each request is calculated as

$$NCI = (C_j - C_i) / PUS$$

In the actual implementation requested object's C-blocks are increased by the proxy server during the playback instead of downloading them completely before playback. The advantage of this implementation is that it saves the network bandwidth, reduces the user request latency and also it reduces the power consumption of the Proxy server's CPU.

V. HYBRID VICTIMIZATION APPROACH

The cache space (CS) is been allocated for the requested object. In the worst case scenario, the absence of free cache space would lead to the drop of objects. To overcome it, based on the object's ATPR some space is been freed. The objects with larger ATPR contribute us in keeping more C-blocks of the same in the proxy cache. Hereby it increases the byte-hit ratio. On the contrary, the larger the object's ATPR so would be the less delay in starting. The requests sent to the object with lower ATPR causes delayed start.

Begin

Let objects $i = 1$ to k in proxy cache

Find CS for requested object.

To find ATPR min values for objects 1 to k Specify object O_j in the playback, where $ATPR_j - ATPR_i$ is minimum

If $C_i \leq C_j$ then, gradually add C-blocks to C_i equivalent to C_j

If $C_i > C_j$ then, add one C-block to C_i

Remove suffix part of victim object

If space not enough then Remove suffix part, prefix part of victim object.

Add calculated CS to proxy server

VI. EXPERIMENTAL RESULTS

The proxy simulator model consists of a user request generator. Proxy server simulator is used to simulate the streaming and caching actions as it happens in the real proxy server in various caching schemes. The user request characterization is created by the user request generator. Several parameters have been considered for the performance studies. Following parameters are considered such as cache size, number of distinct media objects, the size of prefix size and the total number of request.



TABLE I
 EXPERIMENTAL PARAMETERS

PARAMETERS	UNITS
Size of the Cache	7,000 blocks
Total number of Multimedia objects	200
Average size of the object	250
Total number of request	200
Prefix size	3 blocks
Block size	1MB
Number of blocks per processing units	10

For performance study, 200 distinct multimedia objects, whose sizes are uniformly distributed between 100 and 300 blocks, with a mean object size of 200 blocks are taken.

A. Effect Of Cache Space On BHR And Delayed Starts

To study the impact of cache size on the performance of the streaming proxy server by varying the cache size from 2×10^3 blocks to 7×10^3 blocks. The performance of the cache size is shown in Figure 4. The system was run for both the caching schemes to study the influence of the cache size on the BHR. Increasing cache size allows more data to be cached in the proxy server, thereby reducing the percentage of delayed starts and increasing the byte-hit ratio for user requests. As a result, the hybrid proxy-caching scheme which involves both APR and the LRU terms outperforms the existing APR based approach in terms of the percentage of delayed starts and byte hit ratio for the user requests.

When proxy-cache size is relatively small, the hybrid proxy-caching scheme shows less improvement in Byte Hit Ratio. For instance, when the cache size is 2000 blocks, the novel caching scheme experiences 17% byte hit ratio more than the existing scheme. When the cache size is 4000 blocks, the byte-hit ratio using the hybrid proxy-caching scheme is 13% better than the existing proxy caching scheme. With the cache size of 6000 blocks the hybrid scheme outperforms 15% more than the existing approach. When the cache space proxy increases the novel scheme performs better than the existing approaches.

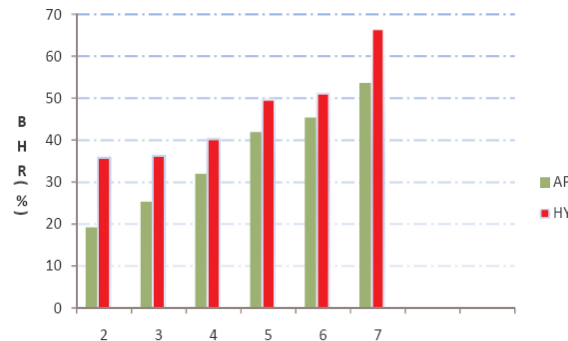


Fig. 4 Byte Hit Ratio under Various Cache Size

By studying the impacts of cache size on delayed start, the hybrid approach has the lowest fraction of requests with delayed starts. Figure 5. Presents the impact of cache size on delayed starts with the requested objects. The result shows that the delayed starts for hybrid scheme decreases as the size of the cache size increases. As the cache size increases more and more object prefixes are placed in the cache thereby experiencing less delayed starts by the user. For instant of 3000 blocks the percentage of delayed starts for hybrid scheme decreases by about 2% compared to the existing scheme.

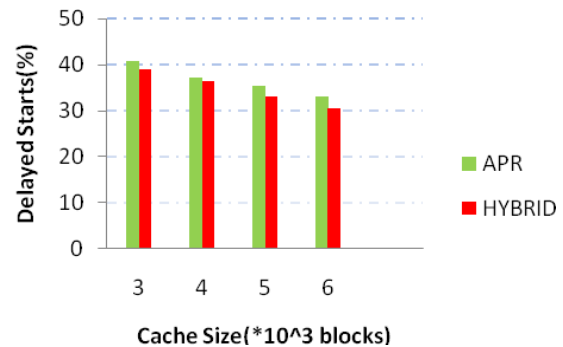


Fig. 5 Delayed Starts under Various Cache Size

When the cache consists of 6000 blocks the hybrid scheme performs better than the existing scheme by 3%. From the analysis, the delayed start is lesser for increasing cache size.

B. Effect Of Numbers Of Distinct Objects On BHR And Delayed Starts

In this study the impact of the number of the distinct multimedia objects on the proxy server performance is analysed by varying the number of the distinct objects from 50 to 200. The results are depicted in Figure 6. As the number of the distinct objects increases, larger cache space is needed to keep the increased number of prefixes in the cache. Thus, the byte-hit ratio decreases for both proxy-caching schemes.



The hybrid proxy-caching scheme outperforms the other caching schemes in terms of the Byte Hit Ratio. For instance, when the number of the distinct objects is 50, the hybrid caching scheme experiences about 1% increase in Byte Hit Ratio which is caused by initial requests for the objects.

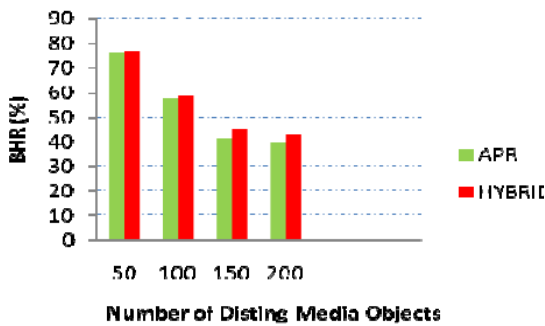


Fig. 6 BHR under Various Numbers of Distinct Objects

However, when the number of the distinct objects increases to 200, the number of Byte Hit Ratio using the hybrid caching scheme outperforms the existing scheme by 4%. In addition, as shown in the Figure 7, when the number of the distinct objects increases, the hybrid caching scheme performs still better, in terms of the byte-hit ratio, compared to the other caching schemes. When the number of the distinct objects is 150, the fragmental proxy-caching scheme outperforms the existing scheme By 5%. The result shows that the delayed starts for the Hybrid scheme are lower when compared to the existing scheme. The delayed starts increased for increase in the number of distinct media objects due to constant cache size. For this analysis the cache size is kept constant with 10,000 blocks. For instance of 50 multimedia objects the delayed starts for the hybrid scheme is 25% and for existing scheme is 28%, resulting in the reduction of delayed starts by 3%. When the number of multimedia objects is 200, the delayed starts for the novel scheme is 31.5% and for the existing scheme is 35% resulting in an improvement of 3.5%. When more and more different objects are requested, the number of delayed starts increases since the number of requested objects in the cache is less.

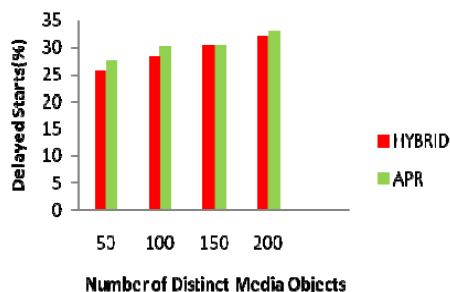
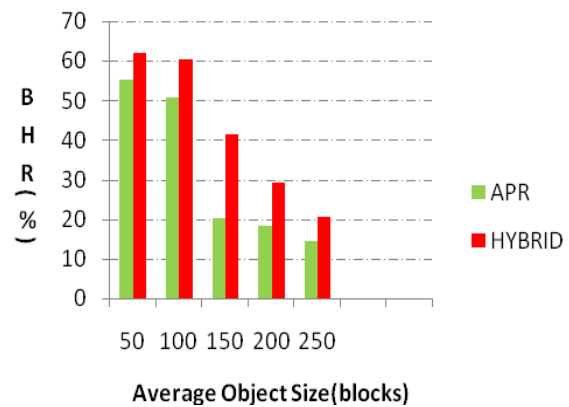


Fig. 7 Delayed Starts under Number of Distinct Media objects Thus the above results prove that for various cache sizes the Byte Hit Ratio and the Delayed Starts are better than the existing system. Moreover the BHR and Delayed Starts are better for the various number of distinct media objects.

C. Effect Of Average Object Size On BHR And Delayed Starts

In this work, we study the impact of the average object size on the proxy server performance. We vary the average object size from 50 to 250 blocks. Because the prefix size of a multimedia object is usually fixed and does not increase when the object size increases, varying the average object.



The Size should not have an obvious impact on the number of delayed starts as long as the total cache size is large enough to store the prefixes of all multimedia objects. As demonstrated in Figure 8, the fragmental proxy-caching scheme outperforms the other caching schemes by large margins in terms of the Byte Hit Ratio. On the other hand, the byte-hit ratio decreases for all proxy-caching schemes when the average object size increases. The simulation shows that Hybrid scheme outperforms the existing scheme. For this simulation the cache size is fixed to 10000 blocks and the user request is limited to 200 requests. For the instance for average object size of 50 the fragmental caching scheme under Hybrid scheme is 10% more than the existing APR. For the instance of average object size of 150 blocks the novel scheme outperforms about 20% than the existing fragmental caching scheme. To analyse the delayed starts, the average object size is varied from 50 blocks to 200 blocks and the cache size is fixed. As from Figure 9, for the instance of initial 50 blocks the Hybrid scheme outperforms by around 4% than the existing scheme. When number of Average object size increases the delayed starts also increases due to the fixed cache size. This reflects in the replacement of prefix object. And for the other instance of 200 blocks the delayed starts for Hybrid scheme is 25.84% and the existing scheme is 26.483%. Thus from the study Hybrid scheme outperforms by around 1%.

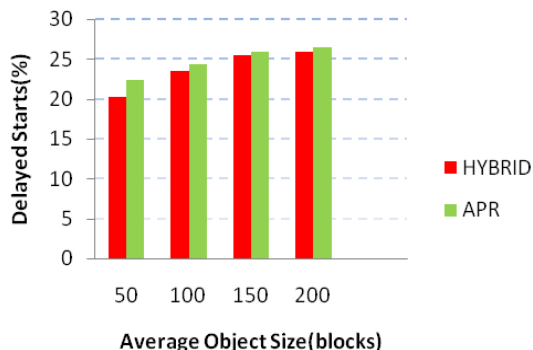


Fig. 9 Delayed Starts under Average object sizes

VII. CONCLUSION

Proxy Cache is to improve the Multimedia Streaming Service's quality. The effective Cache replacement is backed up by the finer granularity of caching units. With it, a new cache replacement algorithm is proposed in which the user request arrival time and the playback rates of the multimedia objects are considered in making the cache replacement decisions. In the proposed cache replacement algorithm, the dynamic user request is tracked using the playback log; and the replacement victims are selected using a hybrid tunable victimization procedure and it's also used to determine the amount of cache data to be replaced for the victims. A detailed comparison of the APR based Proxy caching scheme and Hybrid proxy caching scheme is done. The study results shows that the fragmental proxy-caching scheme with Hybrid scheme significantly outperforms the existing schemes in terms of two commonly used QoS metrics, the byte-hit ratio and the number of delayed starts by 10% and 3% respectively with the existing system. The experimental study further demonstrates that the fragmental proxy-caching scheme can be tuned to either minimize the number of delayed starts or maximize the byte hit ratio based on different QoS requirements of the streaming multimedia applications.

REFERENCES

- [1] S. Sen, J. Rexford, and D. Towsley, "Proxy prefix caching for multimedia streams," in Proc. IEEE INFOCOM, New York, Mar. 1999.
- [2] Z. Miao and A. Ortega, "Scalable proxy caching of video under storage constraints," IEEE J. Select. Areas Commun., vol. 20, no. 7, pp. 1315–1327, Sep. 2002.
- [3] S. H. Park and K. D. C. E. J. Lim, "Popularity-based partial caching for VOD systems using a proxy server," in Workshop on Parallel and Distributed Computing in Image Processing, Video Processing and Multimedia, Apr. 2001.
- [4] Eun-Ji Lim, Seong-Ho park, Hyeoin-Ok Hong, Ki-Dong Chung; "A proxy caching scheme for continuous media streams on the Internet" Information Networking, Proceedings, 15th International Conference on pp: 720-725, 27-31 January 2001
- [5] Bhattacharjee, A, Debnath, B.K, "A new Web Cache Replacement Algorithm" Communications, Computers and signal Processing, 2005.

PACRIM. 2005 IEEE Pacific Rim Conference, pp: 420 – 423, 24-26, August. 2005.

- [6] M. Kim and J. A. Copeland, "Bandwidth sensitive caching for video streaming application," in Proc. IEEE Int. Conf. Communications, May 2003, vol. 3, pp. 1557–1562.
- [7] Lim, S. Park, H. Hong, and K. Chung, "A proxy caching scheme for continuous media streams on the internet," in Proc. 15th Int. Conf. Information Networking, 2001, pp.720–725.
- [8] J. Kangasharju, F. Hartanto, M. Reisslein, and K. W. Ross, "Distributing layered encoded video through caches," in Proc. IEEE INFOCOM 2001, 2001, pp. 1791–1800.
- [9] M. Zink, J. Schmitt, and C. Griwodz, "Layer-encoded video streaming: A proxy's perspective," IEEE Commun. Mag., vol. 42, no. 8, pp. 96–103, Aug. 2004.
- [10] J. Liu, X. Chu, and J. Xu, "Proxy cache management for fine-grained scalable video streaming," in Proc. IEEE INFOCOM'04, Hong Kong, Mar. 2004.
- [11] B. Shen, S.-J. Lee, and S. Basu, "Caching strategies in transcoding-enabled proxy systems for streaming media distribution networks," IEEE Trans. Multimedia, vol. 6, no. 2, pp. 375–386, Apr. 2004.
- [12] Junzhou Luo, Tian Tian, "A Prediction-based Two-Stage Replica Replacement Algorithm", Computer Supported Cooperative Work in Design, 2007. CSCWD2007. 11th International Conference, pp: 594 – 598, 26-28 April 2007
- [13] Gonzalez-Canete, F.J., Casilari, E, TRivino-Cabrera, "Evaluation of a Multi-Queue Web Caching Scheme that Differentiates the Content-Type of Documents", Internet Surveillance and Protection, 2006. ICISP '06. International Conference pp:56 – 60, 26-28 August. 2006
- [14] Lei Guo, Songqing Chen, Zhen Xiao, Xiaodong Zhang, "DISC: Dynamic Interleaved Segment Caching for Interactive Streaming", Distributed Computing Systems, 2005. ICDCS 2005. Proceedings. 25th IEEE International Conference, pp:763 – 772, 10-10 June 2005
- [15] K. Y. Leung, E. W. M. Wong, and K. H. Yeung, "Designing efficient and robust caching algorithms for streaming-on-Demand services on the internet," World Wide Web: Internet and Web Inform. Syst., vol. 7, January 2004, pp. 297–314.
- [16] Zhan-sheng Li, Da-wei Liu, Hui-juan Bi, "CRFP: A Novel Adaptive Replacement Policy Combined the LRU and LFU Policies", Computer and Information Technology Workshops, 2008. CIT Workshops 2008. IEEE 8th International Conference, pp:72 – 79, 8-11 July 2008

ACKNOWLEDGEMENT

The authors would like to thank the anonymous referees for their valuable comments and suggestions.

BIOGRAPHY



Saravanan. C is an Assistant Professor in the Department of Information Technology in Krishnasamy College of Engineering and Technology, Cuddalore, affiliated to Anna University, India. He obtained his M.Tech in Information Technology from Madras Institute of Technology, Anna University and finished his under graduation in Electrical and Electronics Engineering in Annai Teresa College of Engineering (Villupuram District), Anna University. He has published three articles in National Conferences. His research interest is in the area of Cloud Computing and Web Services.



Marimuthu. G is an Assistant Professor in the Department of Information Technology in Krishnasamy College engineering & Technology, Cuddalore, affiliated to Anna University, India. He obtained his M.Tech in Distributed Computing Systems (CSE) from Pondicherry Engineering College, Pondicherry University, India in 2009. He has published three articles National and International journal/Conferences. His research includes Cloud Computing and Web Services, Distributed database.



Immanuel Alphonse is an Assistant Professor in the Department of Information Technology in Tagore Engineering College, Vandalur, affiliated to Anna University. His area of research is in Assistive Technology and Renewable Energy. He finished his post-graduation from College of Engineering, Guindy Campus, Anna University in Electrical Drives and Embedded Control. He finished his Bachelors of Technology in DMI College of Engineering, Poonamalle, affiliated to Anna University. He has published three articles in National and International Journals/Conferences.



Kaliaperumal. G is an Assistant Software Engineer in Tata Consultancy Service, Chennai, India. He obtained his M.Tech in Information Security (IT) from Pondicherry Engineering College, Pondicherry University, India in 2009. He has published two journals. His research includes Information security, Web Services and database security.