Study & Analysis of various Web Pre-Fetching Techniques

Rehan Khan[1], Mr. Jitendra Dangra[2]
Dept. Of Computer Science and Engineering, Laxmi Narayan College of Technlogy Indore, (M.P.)[1, 2]
rehankhan3@gmail.com[1]

Abstract: *Increasing popularity of the net over the past few years has obligatory a serious traffic burden upon the net. the internet is additionally thought of to be an oversized distributed info systems providing access to shared information. A mass analysis has done to reinforce the latency of internet primarily based system as a result of the data is distributed over a geographical location. Web caching and pre-fetching are 2 necessary approaches accustomed scale back the noticeable latency perceived by users. During this paper we tend to study and analysis the varied previous researches & techniques supported web Pre-Fetching.*

Keywords: *Web Caching, Pre-Fetching, Web Application, Web Usage Mining.*

1. Introduction

Web could be a key resource so as to share the data on the globe. it\'s sizable amount of advertisements, stories and world connectivity between individuals and plenty of data is available for the individual research scholars. This vast use of internet around globe makes it a lot of vital within the world of analysis. Research worker always has the challenge to build online applications in such a manner that they remain economical. Several researchers work on that and provides new plan so as to allow the higher results from the previous one.

There is a large need to improve the latency of server for internet applications. Currently internet features a huge repository because of increase its use suddenly. It becomes essential to concentrate on each the amount and standard of internet contents. Even, once the speed of web has improved with the reduced costs, traffic is obtaining heavier. The large data makes it tough to seek out the relevant data quickly. This led to the trouble the effort the speed, by reducing the latency, creates the internet a lot of relevant and meaningfully connected.

The Cache pre-fetching plays a very important role so as to boost the latent period and build the application well-organized. the online pre-fetching could be a technique so as to preprocess the user requests, before they are really demanded. Therefore the time that user should watch for the requested documents will be reduced by activity the request latencies. Pre-fetching is that the technique for reducing Latencies. User invariably expects faster & interactive response, higher satisfaction and quality of output. There are numerous approaches and algorithms are projected for rising the online performance.

WEB APPLICATION: A web application is any application that uses an browser as a client for application. The appliance will be straightforward as a message board or a guest sign-in book on a web site, or as complicated as a application or a computer program.

A web application relieves the developer of the responsibility of building a client for a selected sort of pc or a selected software. Since the client runs in a web browser, the user may be using an mackintosh or a IBM-compatible or they'll be running Windows XP or Windows vista. They will even be using internet explorer or Firefox, though some applications need a selected application program.

Web applications commonly use a combination of server-side script (ASP, PHP, etc) and client-side script (HTML, Javascript, etc.) to develop the application. The client-side script deals with the

presentation of the information while the server side script deals with all the hard stuff like storing and retrieving the information.
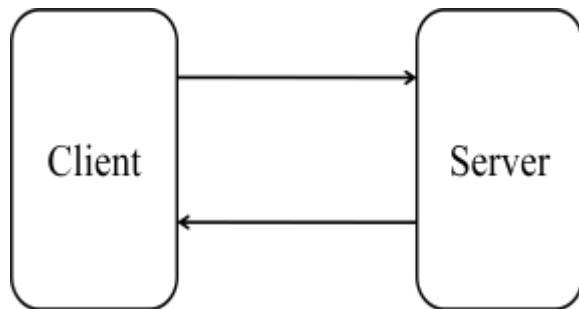

Figure 1: Client Server Architecture

A client Server design (Figure 1) during which every pc or method on the network is either a client or server. Servers are very powerful computers or processes dedicated to managing disk drives (file servers), printers (print servers), or network traffic (network servers). clients are PCs or workstations on that users run applications. clients think about servers for resources, like files, devices, and even process power.

1.1 Web Pre-Fetching

Web Pre-Fetching is another terribly effective technique that is employed to enhance the online caching mechanism. The web Pre-Fetching predicts the online object expected to be requested within the close to future; however these objects don't seem to be however requested by users. Then, the anticipated objects are fetched from the origin server and keep during a cache. Thus, the web Pre-Fetching helps in increasing the cache hits and reducing the user-perceived latency.

"Web Pre-Fetching could be a technique that created efforts to resolve the matter of those access latencies". Specially, world caching ways that straddle across users work quite well. However, the increasing trend of generating dynamic pages in response to communications protocol requests from users has rendered them quite ineffective. Pre-Fetching is employed as a shot to put information near the processor before it\'s needed, eliminating as several cache misses as possible. Caching offers the subsequent benefits: Latency reduction, less bandwidth consumption, Lessens internet Server

load. Pre-Fetching is that the means that to anticipate probable future requests and to fetch the foremost probable documents, before they\'re really requested. it\'s the speculative forty retrieval of a resource into a cache within the anticipation that it is served from the cache within the close to future, thereby decreases the load time of the thing.

1.2 Types of Web Pre-Fetching

The Pre-Fetching techniques implemented enforced on server, proxy or client side. The client-based Pre-Fetching concentrates on the navigation patterns of one user across several internet servers. On another hand, the sever-based Pre-Fetching concentrates on the navigation patterns of all users accessing one web site. The proxy-based Pre-Fetching concentrates on the navigation patterns of a bunch of users across several internet servers. Thus, this approach reflect mirror a standard interest for user's community. In alternative words, the Pre-Fetching contents will be shared by several users.

2. Literature Review

In 2000 Annie P Foong, u-Hen Hu and Dennis M Heise [1] identified the existence of multi-dimensional locality in web accesses. They also proposed that design of cache strategies must account for all of these localities to be fully effective. The localities and predictors presented in this paper are by no means exhaustive. What they have achieved here is show the applicability of a simple LR model in characterizing these localities. The availability of eventful data in web re-accesses is especially suited towards such statistical methods as the LR model. They suggest that web caching algorithms must be adaptive. We have investigated but a small area in web caching and made decisions based only on what a single cache sees.

In 2000 Weifeng Zhang, Baowen Xu, William Song, Hongji Yang and Kecheng Liu [2] presents a simple WWW data model and then introduces some technologies with which the data in the cache of the browser can be mined, the knowledge is saved into the interest association repository and the users' impending activities can be predicted. Integrated with the data mining technique, the agent technology and the web technology, better quality of web service for the users is provided. The design ideas of the agent

based web pre-fetching reflect the development tendency of intelligent web browsers.

In 2001 Mas & SUZUKI, Hiroyuki YOKOYAMA and Takeshi MIZUIKE [3] have discussed caching and mirroring techniques suitable for the educational Internet. They considered equipment issues, teaching methods and problems regarding the present situation of Internet usage in the educational workplace. They described network models and traffic characteristics of the educational Internet taking these factors into account. With regard to caching and mirroring, their discussion focused on storage costs and communications costs per unit amount of data for hard disks and access circuits. For caching, they showed that limitations on storage capacities have been practically eliminated by the high capacities and lower costs of hard disks. For mirroring, they proposed that mirroring should be practiced intensively for sites that are expected to be accessed even to a small extent. Further, they touched on the possibility of incorporating pre-fetch methods with caching or mirroring. Finally, they introduced the concept of the virtual network that can exclude indiscriminate access to harmful information, SPAM mail, and nuisance mail, and expressed our conviction that such a network is useful for learning the basic techniques of the Internet.

In 2003 Xin Jin and Huanqing Xu [4] proposed the pre-fetching model based on HMM. Their scheme is based on the idea that utilizing HMM to mine the latent information requirement concepts that the user access path implies, the pre-fetching model makes semantic-based pre-fetching decisions. According to authors web pre-fetching has common feature with other web applications using prediction of user access pattern. Authors suggests that after being modified a little, this approach can he easily extended to some domains such as information filtering, information recommender system, personality web site, etc.

In 2004 Amit Thawani, Srividya Gopalan and Sridhar V [5] have proposed a system that facilitates reduction in call resolution time in a contact center through context-aware information retrieval. This is achieved by defining the context of a query to include the explicit information and implicit information and the context is constructed using customer, agent and business entities; implicit information is derived by analyzing and aggregating the past context history; and information derived by the analysis of past contexts is used in context aware information retrieval. This retrieved information aides the agent to propose quicker and better resolutions as well as to make appropriate promotional offers. According to the authors the contextual information aids the agent in such a way as to reduce the agent's need for training.

In 2005 Hussein Fareed and Mostafa Bassiouni [6] have presented and evaluated a flow based algorithm for web caching and pre-fetching. The algorithm captures the access relationships between objects and represents them as a weighted directed graph. The weights on the edges of the graph represent the degree of access coupling between pairs of objects. According to authors the Bow-based framework can be easily integrated into existing caching algorithms with little overhead. They have applied the flow-based framework to the "LRU" algorithm and evaluated its performance using NS2 simulation with real-life traces. In all test cases, the flow algorithm gave improvement over LRU. It is expected that the proposed flow framework will perform much better in sites that use business scenarios such as reservation systems, electronic stores, online malls, and e-shopping. The scheme presented in this paper can be extended in two ways: extending the tests to more real-life applications and investigating ways to capture longer flow sequences without incurring intolerable overhead. Another area worthy of investigation is the development of predictive flow models for object access. Such models can be used to good advantage similar to predictive internet traffic models used in optimizing bandwidth provisioning. According to authors predictive flow models can help reduce the frequency of updating the flow graph and reduce the associated computational overhead.

In 2006 Ganesh Ananthanarayanan, Sean Blagsvedt and Kentaro Toyama [7] have proposed and implemented a system OWeB to improve the Internet browsing experience over slow and intermittent networks. OWeB made the browsing experience significantly independent of the network availability. The OWeB framework was made robust and resilient by employing standard techniques like queuing and re-trials. Authors observed a co-relation between the core content section of homepages and the items in the RSS feeds and devised an algorithm to automatically extract the template of a web page and stitch the incoming RSS feeds into the template

locally, thereby achieving significant savings in the data downloaded. There system implementation results validated the correctness of their observation and also illustrated the accuracy of their template identifier and stitching algorithm in addition to appreciable data savings. Algorithm presented in this paper works only when every content section is defined by a single RSS feed. They plan to extend their algorithm to work in scenarios where a single content section is defined by multiple RSS feeds and when a single RSS feed defines multiple sections of the page.

In 2007 Achuthsankar S. Nair and Jayasudha J.S [8] proposed a dynamic pre-fetching technique in which web caching and pre-fetching techniques are integrated. Web pre-fetching techniques are used for reducing latency, but it increases web traffic. In dynamic web pre-fetching technique, subsequent links are pre-fetched only if bandwidth usage of existing network is less than a predefined threshold. For each web page request, the retrieved page is parsed to identify the subsequent links and URL's corresponding to these links is searched in the hash table to get its weight information. Intelligent agents monitor the bandwidth usage, user's preferences and hash table weights to identify the number of URLs to be pre-fetched. Simulation result of this technique presented in this paper shows that dynamic pre-fetching browser maintains almost constant web traffic even if pre-fetching is done. Since dynamic pre-fetching technique increases cache hit ratio, reduces latency and maintains almost constant traffic, it is preferred to all the existing pre-fetching techniques.

In 2010 Theint Theint Shwe, Thida Myint, Theint Theint Aye, Su Su Htay, Swe Swe Nyein and Mie Mie Su Thwin [9] have proposed a framework for Web Usage Analyzer is intended to apply in many ways such as for Web Site Maintainers, Web Analyzers, Pre-fetched Systems, Web Personalized Systems and Recommender Systems because they are mining based on user's purposes, date, and site of web log data then presents the results upon different dimensions. It can also be used in statistical analysis of the information about most often used web sites to inform the particular clients and can be used in Pre-fetched system.

In 2010 Toufiq Hossain Kazi, Wenying Feng and Gongzhu Hu [10] have review and categorize different Web prefetching models and a new prefetching algorithm is proposed that is based on a modified ART1 neural network clustering. They have confidence in the new algorithm anticipating reasonably good performance in certain metrics such as user perceived delays, prefetching accuracy and overhead cost on network traffic.

In 2010 Brijendra Singh, Hemant and Kumar Singh [11] provided an evaluation and update of available web mining research. Extensive literature has been reviewed based on three types of web mining, namely web content mining, web usage mining, and web structure mining. They present year wise Summary of improvements in each type of mining. Web data mining is a fast rising research area today. As the web data and its usage will rise in future. It will prolong to generate more content, structure and usage data. So the importance of web data continues increasing. Web data is mainly semi-structured to unstructured. Due to the heterogeneity and the lack of structure of Web data, automated discovery of targeted or unexpected knowledge information still present many challenging research Problems. Data Mining is perhaps still in its infancy and much research is being carried out in the area.

In 2011 Waleed Ali, Siti Mariyam Shamsuddin and Abdul Samad Ismail [12] reviews the principles and some existing web caching & prefetching approaches. Firstly, they have reviewed principles and existing works of web caching. This includes the conventional and intelligent web caching. Secondly, types and categories of prefetching have presented and discussed briefly. Moreover, the history-based prefetching approaches have been concentrated and discussed with review of the related works for each approach in this survey. Finally, authors presented some studies that discussed integration of web caching and web prefetching together. Web caching and prefetching are two effective solutions to lessen Web service bottleneck, reduce traffic over the Internet and improve scalability of the Web system. The Web caching and prefetching can complement each other since the web caching exploits the temporal locality for predicting revisiting requested objects, while the web prefetching utilizes the spatial locality for predicting next related web objects of the requested Web objects. Thus, combination of the web caching and the web prefetching doubles the performance compared to single caching.

In 2011 Kavita Sharma, Gulshan Shrivastava and Vikas Kumar [13] provide a survey about the research in the area of Web mining's today structure and tomorrow view. They point some confusion between data mining and web mining. Web data is growing at a significant rate. Web Mining is fertile area of research. Many Successful applications exist. They also suggest the subtask of web mining & future of web mining. They also work for the process mining and try to combine usage mining with structure mining. Authors also go for the mining from cloud. Whenever they work on mining over cloud computing that time they hesitate for the cost but that come very less by cloud mining. So, they can say that cloud mining can seen as future of web mining.

In 2012 A Vadivel, Shaila S. G, R. Devi Mahalakshmi and J. Karthika [14] have designed and developed a crawler that communicates with various components in the form of XML messages. The message processing methodology has exploited the power of web services and increased the crawling & indexing capacity of the crawler. From their experimental results, they observed that the number documents retrieved and indexed is on the higher side. By their work, the directions for the future are given to make the research more optimal in its activation with the user. Because of the large storage space that is required by the Inverted Index, a suitable technique is required to compact the indexed documents to reduce the storage space. Also, Development of an intelligent spell checker will enhance the effectiveness of web crawler to identify the wrong word in the user query.

In 2012 Sarina Sulaiman, Siti Mariyam Shamsuddin, Nor Bahiah Ahmad and Ajith Abraham [15] provides guidance to the administrator in WC regarding to selection of the best parameters to be cached and used in mobile Web pre-caching. Based on their analysis, the administrator may reorganize the parameter of log data set in proxy cache accordingly. Likewise, an empirical research has been conducted to search for the optimal classification. In addition to their work, RS classifier was implemented to optimize the performance of decision Web object to either cache or not cache in a proxy cache. The RS framework for log dataset was illustrated mutually with an analysis of reduced and derived rules, with entrenchment of their implicit properties for better classification outcomes. The actual prediction accuracy proves that the RS is capable to be used as a classifier in order to predict significant data from the main BU and EL datasets and will be used to produce a pull global Web pre-fetching on mobile applications.

In 2012 Abhay Singh and Anil Kumar Singh [16] presented an approach for prefetching the web pages for a user, on the basis of his history of browsing by using sequential data mining technique. The overall performance of Cache using the proposed approach for both Page Replacement namely LRU and LFU using pre-fetching has improved by 12% and 8 % respectively as compared to the normal caching policy of LRU, as determined by the Hit Rate and Byte Hit Rate for each sample of log files.

In 2012 Poornalatha G and Prakash S Raghavendra [17] have proposed a prediction model that yields good accuracy. Also, an integrated distance method is proposed to find the similarities between any two user sessions based on the sequence alignment. The results obtained from their work are compared with few other results available in the previous research to demonstrate the goodness of the prediction model. Since accuracy is good, the work could be useful for prefetching applications that reduce the user latency.

In 2012 Gou-feng Zhao, Bing Li and Tong Hong [18] propose a novel user aware model to provide personalized pre-fetching for VIP users while guaranteeing the common pre-fetching for ordinary users based on dynamic Markov chain. Markov chain has been applied in webpage pre-fetching for a long history: traditional Markov, multi-markov, and hybrid-markov. However, these methods cannot be applied to MSN pre-fetching for some weaknesses. This model may have a large application space: it can not only be applied in mobile social network ecosystem, but also may be applied in some Mobile Internet websites, in which distribution of user clicks and traffic flow polarized.

In 2012 K. Ramu, Dr. R. Sugumar and B. Shanmugasundaram [19] have given brief introduction about the different Web prefetching techniques namely Prediction by partial match, Predictive web prefetching Model based predictive pre-fetching, Semantic web prefetching, Link pre-fetching, Domain top approach, Data prefetching, Content prefeching, Context based prefetching, Proxy cache prefetching, Dynamic web pre-fetching and

Greedy-Dual-Size pre-fetching are analyzed and discussed. The web prefetching scheme focus on the property spatial locality of web objects. These techniques are applied to reduce the network traffic and improve the user satisfaction. Web prefetching and caching can also be integrated to get better performance.

In 2013 Yogish H K and G. T. Raju [20] proposes a novel approach that makes use of ART1NN Clustering and Prefetching technique, to fetch the pages for the user community before request them. Most of the research works in pre-fetching concentrates on individual user's requests according to their previous access patterns. Authors presented CPF approach that showed its usefulness in reasonable utilization of network resources through prefetching of Web pages for a community of users instead of a single user with an average prediction accuracy of 93.16%. Though the CPF approach results in substantial increase of network traffic, it effectively reduces the user perceived latency.

In 2013 N Nandini, H K Yogish and G T Raju [21] have been presented Various Pre-fetching techniques along with the parameters and applications. They show the comparison of various pre-fetching techniques. Most of the research works in pre-fetching concentrates on individual user's requests according to their previous access patterns. Although these methods are efficient for pre-fetching, they may considerably overload the network with unnecessary traffic when pre-fetching for a large number of users. Authors suggest Future research directions in this regard concern with the development of a novel pre-fetching approach that pre-fetch requests for a large community of users instead of individual user's request based on frequent sequential patterns and cluster patterns and development of adaptive predictive systems that use hybrid approach such as use of statistical, neural, and Bayesian learning algorithms.

In 2014 Partha Ghosh and Soumya Sen [22] have proposes a new methodology to compute the pagerank based on the existing parameters as well as newly identified parameters based on location and time. Location is identified in terms of local and global. Time is identified in terms of current and long history. All of these values are considered and summarized to obtain the pagerank. In this method a number of threshold values are considered. These

threshold values could be specific to some applications or these could be set to a constant value. Authors suggest that this work could be extended to compute the threshold values of different variables based on experimental result or statistical analysis. Thus this research work could be exploited in mathematical or statistical domain for further enhancement.

In 2014 Gowtham S, Mausumi Goswami, Balachandran K and Bipul Syam Purkayastha [23] puts forth the concepts of document pre-processing, which is achieved by extraction of keywords from the documents fetched from the web processing and generating a term-document matrix, TF-IDF and the different approaches of TF-IDF (term frequency Inverse document frequency) for each respective document. The techniques mentioned by those are necessary for every web or text mining techniques, as it affects the efficiency and effectiveness of any text related areas like document clustering, web mining and data mining. It is mandatory to perform any of these techniques to improvise the quality of mining process. The proposed algorithm is simple to implement and also most effective in terms of performance. As a result, use of these method and algorithms can increase the overall performance in the field of search engines, document classification, document clustering and most text oriented fields.

In 2014 Avadh Kishor Singh, Ajeet Kumar and Ashish K. Maurya [24] have find results that FP-Growth algorithm is used for finding the most frequently access pattern generated from the web log data. By using the concept of web usage mining they can easily find out the user's interest and they can modify their web site more valuable and more easily accessible for the behavior users. The main focus of the empirical analysis and comparison is to identify factors. In their comprehensive study and defined different factors, they find results that the Apriori and FP-growth method is efficient and scalable for mining both long and short frequent patterns. In future the both algorithm can be extended to web content mining, web structure mining.

In 2014 Neha Sharma and Dr. Sanjay Kumar Dubey [25] have used SPRINT- a decision tree induction classification technique, instead of naive bayes. Initially anchor text is found and using lexical analyser, tokens are counted. Then after, SPRINT is applied. On the results of SPRINT, patterns are

compared with the threshold. If they are greater than the threshold, they are prefetched and stored in the cache, else discarded. Results of the proposed approach is compared with the previous ones using Matlab 5.0 and found that proposed work provides more accurate results.

3. Conclusion

In this paper, we discussed various techniques that are implemented for Web Pre-Fetching. We had also analyzed data model techniques, Caching & Mirroring techniques, Information Retrieval techniques, Web Caching techniques, Web Usage Analyzer, Web Mining, Crawler, Sequential Data Mining Technique, Prediction Model, Markov Models, Clustering, Apriori & FP-Growth Method, SPRINT (a decision tree induction classification technique). A short survey of various Web Pre-Fetching techniques has been given. The previous approaches that rely solely on caching offer limited performance improvement because it is difficult for caching to handle the large number of increasingly diverse files, also the benefit of caching is limited.

Reference

[1] Annie P Foong, u-Hen Hu and Dennis M Heise, "Web Caching: Locality of References Revisited", at 0-7695-0777-8/0$01 0.00 0 2000 EEE.

[2] Weifeng Zhang, Baowen Xu, William Song, Hongji Yang and Kecheng Liu, "Data Mining Algorithms for Web Pre-Fetching", at 0-7695-0577-5/00 $10.00 0 2000 IEEE.

[3] Mas & SUZUKI, Hiroyuki YOKOYAMA and Takeshi MIZUIKE, "Study on Caching, Mirroring, and Pre-fetch Methods for Educational Internet", 0-7695-0945-2/01$10.0 00 2001 IEEE.

[4] Xin Jin and Huanqing Xu, "An Approach to Intelligent Web Pre-fetching Based on Hidden Markov Model", Proceedings afthe 42nd IEEE, Conierenre on Decision and Control, Mani, Hawaii USA, December 2003.

[5] Amit Thawani, Srividya Gopalan, and Sridhar V, "Web-based Context Aware Information Retrieval in Contact Centers", Proceedings of the IEEE/WIC/ACM International Conference on Web Intelligence (WI'04), 0-7695-2100-2/04 $ 20.00 IEEE.

[6] Hussein Fareed and Mostafa Bassiouni, "Adaptive Object Cache Pre-fetching Scheme based on Object Flow". 0-7803-8991 -3/05/$20.00 0 2005 IEEE.

[7] Ganesh Ananthanarayanan, Sean Blagsvedt, Kentaro Toyama, "OWeB: A Framework for Offline Web Browsing", Proceedings of the Fourth Latin American Web Congress (LA-WEB'06), 0-7695-2693-4/06 $20.00 © 2006 IEEE.

[8] Achuthsankar S. Nair and Jayasudha J. S., "Dynamic Web Pre-fetching Technique for Latency Reduction", 0-7695-3050-8/07 $25.00 © 2007 IEEE, DOI 10.1109/ICCIMA.2007.303.

[9] Theint Theint Shwe, Thida Myint, Theint Theint Aye, Su Su Htay, Swe Swe Nyein and Mie Mie Su Thwin, "FRAMEWORK FOR MULTI-PURPOSE WEB LOG ACCESS ANALYZER", 978-1-4244-6349-7/10/$26.00 _c 2010 IEEE.

[10] Toufiq Hossain Kazi, Wenying Feng and Gongzhu Hu, "Web Object Prefetching: Approaches and a New Algorithm" 2010 11th ACIS International Conference on Software Engineering, Artificial Intelligence, Networking and Parallel/Distributed Computing, 978-0-7695-4088-7/10 $26.00 © 2010 IEEE, DOI 10.1109/SNPD.2010.28.

[11] Brijendra Singh and Hemant Kumar Singh, "WEB DATA MINING RESEARCH: A SURVEY", 978-1-4244-5967-4/10/$26.00 ©2010 IEEE.

[12] Waleed Ali, Siti Mariyam Shamsuddin and Abdul Samad Ismail, "A Survey of Web Caching and Prefetching", Int. J. Advance. Soft Comput. Appl., Vol. 3, No. 1, March 2011, ISSN 2074-8523; Copyright © ICSRS Publication, 2011.

[13] Kavita Sharma, Gulshan Shrivastava and Vikas Kumar, "Web Mining: Today and Tomorrow", 978-1-4244-8679-3/11/$26.00 ©2011 IEEE.

[14] A Vadivel, Shaila S. G, R. Devi Mahalakshmi and J. Karthika, "Component based Effective Web Crawler and Indexer using Web Services", IEEE-International Conference On Advances In Engineering, Science And Management (ICAESM -2012)

March 30, 31, 2012, ISBN: 978-81-909042-2-3 ©2012 IEEE.

[15] Sarina Sulaiman, Siti Mariyam Shamsuddin, Nor Bahiah Ahmad and Ajith Abraham, "Meaningless to Meaningful Web Log Data for Generation of Web Pre-caching Decision Rules Using Rough Set", 2012 4th Conference on Data Mining and Optimization (DMO), 02-04 September 2012, Langkawi, Malaysia, 978-1-4673-2718-3/12/$31.00 ©2012 IEEE.

[16] Abhay Singh and Anil Kumar Singh, "Web Pre-fetching at Proxy Server Using Sequential Data Mining", 2012 Third International Conference on Computer and Communication Technology, 978-0-7695-4872-2/12 $26.00 © 2012 Crown Copyright, DOI 10.1109/ICCCT.2012.14.

[17] Poornalatha G and Prakash S Raghavendra, "Web Page Prediction by Clustering and Integrated Distance Measure", 2012 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining, 978-0-7695-4799-2/12 $26.00 © 2012 IEEE, DOI 10.1109/ASONAM.2012.231.

[18] Gou-feng Zhao, Bing Li and Tong Hong, "Pre-fetching Webpages on Mobile Social Network: User-aware Dynamic Markov Chain", 2012 8th International Conference on Mobile Ad-hoc and Sensor Networks, 978-0-7695-4961-3/13 $26.00 © 2013 IEEE, DOI 10.1109/MSN.2012.30.

[19] K.Ramu, Dr. R. Sugumar and B. Shanmugasundaram, "A STUDY ON WEB PREFETCHING TECHNIQUES", Journal of Advances in Computational Research: An International Journal, Vol. 1 No. 1-2 (January-December, 2012).

[20] Yogish H K and G. T. Raju, "A Novel ART1NN Clustering and Pre-fetching Technique for Reducing Web Latency", 2013 5th International Conference on Computational Intelligence and Communication Networks, 978-0-7695-5069-5/13 $26.00 © 2013 IEEE, DOI 10.1109/CICN.2013.74.

[21] N Nandini, H K Yogish and G T Raju, "Pre-fetching Techniques for Effective Web Latency Reduction - A Survey", 978-1-4673-5943-6/13/$31.00 ©2013 IEEE.

[22] Partha Ghosh and Soumya Sen, "Time and Location Based Summarized PageRank Calculation of Web Pages", 978 2014 IEEE International Conference on Industrial Technology (ICIT), Feb. 26 - Mar. 1, 2014, Busan, Korea, 978-1-4799-3939-8/14/$31.00 ©2014 IEEE.

[23] Gowtham S, Mausumi Goswami, Balachandran K and Bipul Syam Purkayastha, "An approach for document Pre-Processing and K Means Algorithm Implementation", 2014 Fourth International Conference on Advances in Computing and Communications, 978-1-4799-4363-0/14 $31.00 © 2014 IEEE, DOI 10.1109/ICACC.2014.46.

[24] Avadh Kishor Singh, Ajeet Kumar and Ashish K. Maurya, "An Empirical Analysis and Comparison of Apriori and FP- Growth Algorithm for Frequent Pattern Mining", 2014 IEEE International Conference on Advanced Communication Control and Computing Technologies (lCACCCT), ISBN No. 97S-1-4799-3914-5/14/$31.00 ©2014 IEEE.

[25] Neha Sharma and Dr. Sanjay Kumar Dubey, "Semantic Based Web Prefetching Using Decision Tree Induction", 978-1-4799-4236-7/14/$31.00_c 2014 IEEE.