RESEARCH ARTICLE OPEN ACCESS

A Review on Pattern Discovery Techniques of Web Usage Mining

M.Rekha Sundari
Dept. of CSE, GITAM University

Y.Srinivas
Dept. of IT , GITAM University

PVGD.Prasad Reddy
Dept. of CS&SE, Andhra University

Abstract---
In the recent years with the development of Internet technology the growth of World Wide Web exceeded all expectations. A lot of information is available in different formats and retrieving interesting content has become a very difficult task. One possible approach to solve this problem is Web Usage Mining (WUM), the important application of Web Mining. Extracting the hidden knowledge in the log files of a web server, recognizing various interests of web users, discovering customer behavior while at the site are normally referred as the applications of web usage mining. In this paper we provide an updated focused survey on techniques of web usage mining.
Keywords: Web Usage Mining, Pattern Discovery, Clustering, Classification.

I. INTRODUCTION

The advancement in technology has brought revolutionary strides for carrying out E-business through World Wide Web (WWW). This explosive increase in the usage of WWW and its capability of storing huge data attracted millions of visitors. As data continue to grow in size and complexity, sophisticated methods to organize the layout of the information become important. This information from the data is used in efficient and effective management of the activities related to e-business, e-education, e-commerce, personalization, Web site design, improvement and management, network traffic analysis, search engine's complexity, and to predict user's actions [40]. Nevertheless, understanding the needs of their users is vital for the owners of the Web sites in order to serve them better. This generated a need to extract useful information from huge amount of data related with web sites. This data is of many types --- the content from web documents like text and graphics, the data from web structure like HTML or XML tags, the data from web log like IP addresses, date or time of access of web pages or the data that is user specific like registration, customer profile etc.., . This user specific data is recorded in the Web access log files of Web servers and usually referred as Web Usage Data (WUD). WUM is that area of Web mining which deals with the application of data mining techniques to reveal interesting knowledge from the WUD.

1.1 Web Usage Mining(WUM)

WUM is that area of Web mining which deals with the application of data mining techniques to reveal interesting knowledge from the WUD. WUM is a three phase process [15] that includes data collection and data preprocessing, pattern discovery and pattern analysis of web data.

A. Data Preprocessing

The success of the pattern analysis phase is highly correlated to how well the data preparation task is executed. It is of utmost importance to ensure, every nuance of this task is taken care of. This process deals with loading of the data, performing accuracy check, putting the data together from disparate sources, transforming the data into required format and finally to structure the data as per the input requirements of some data mining algorithm. This involves many phases like data cleaning, feature extraction, feature reduction, user identification, session identification, page identification, formatting and finally data summarization [8].

B. Pattern Discovery

The preprocessed data is considered for the application of knowledge extraction algorithms based on AI, data mining algorithms, psychology, and information theory. Most of the systems developed for the WUM process have introduced different algorithms for finding the maximal forward reference, large reference sequence, to analyze the traversal path of a user. Different mining algorithms like path analysis, association rules, sequential patterns, clustering and classification are used for effective process of WUM (will be discussed in the subsequent sections). It totally depends on the requirement of the analyst to determine which mining techniques to make use of. When exposed to these algorithms, data in web access logs can be transformed into knowledge to uncover the potential patterns underneath the pre-processed log data and involves analyses of these patterns.

C. Pattern Analysis

The last phase in WUM process is the analysis of the obtained results in order to distinguish trivial, useless knowledge from knowledge that could be used for Web site modifications, system improvement and/or Web personalization. The common techniques used for pattern analysis are, visualization techniques, OLAP techniques [5], Data & Knowledge Querying, and Usability Analysis.

II. Pattern Discovery Techniques of WUM

The following are the various techniques identified in pattern discovery phase of web usage mining.

2.1 Clustering

Clustering aims at dividing the data set into groups (clusters) where the inter-cluster similarities are minimized while the similarities within each cluster are maximized [44]. In the context of WUM, we can distinguish two cases of clusters, user clusters and page clusters. Web page clustering is performed by grouping pages having similar content. User clustering is performed by grouping users by their similarity in navigational behavior. Clustering can be model-based or distance-based. With model-based clustering [49], the model type is often specified apriori and the model structure can be determined by model selection techniques and parameters estimated using maximum likelihood algorithms, e.g., the Expectation Maximization (EM). Distance-based clustering involves determining a distance measure between pairs of data objects, and then grouping similar objects together into clusters. The most popular distance-based clustering techniques include partitional clustering and hierarchical clustering.

Yang and Balaji [49] proposed hierarchical pattern based clustering algorithm for grouping web transactions and to maximize the objective function in order to achieve good clustering of customer transactions. Sophia et al. [43] emphasized the need to discover similarities in users' accessing behavior with respect to the time and locality of their navigational acts. The two tracks of the proposed algorithms define clusters with users that show similar visiting behavior at the same time period, by varying the priority given to page or time of visiting. Raju and Sudhamani [13] proposed a novel partitional based approach for dynamically grouping Web users based on their Web access patterns using Adaptive Resonance Theory1 Neural Network (ART1 NN) clustering algorithm. Cheng et al. [7] proposed a research using both agglomerative and partitional clustering. Loyola et al. [32] proposed a novel methodology for analyzing Web user behavior based on session simulation by using an Ant Colony Optimization algorithm. Ríos et al. [42] utilized two commonly used clustering algorithms, Self Organizing Feature Maps (SOM) and K-medoids to obtain behavior patterns of the users.

Model-based clustering have been shown to be effective for high dimensional text clustering [53].Whereas, hierarchical distance-based clustering proved to be unsuitable for the vast amount of Web data. Partitioned distance-based clustering is disadvantaged by the different distance measures proposed for clustering purpose and defining a good measure is very much data dependent and often requires expert domain knowledge. Despite the variety of clustering approaches that have been used for Web usage mining, Clustering is employed to guide the predictive system and its alone cannot be an appropriate approach for web page prediction [20]. It is merely used to segment data into some homogeneous groups so that a quality model can be built on each group. Another clustering limitation is the ability to evaluate and compare their performance. The reason for this is the lack of an objective evaluation criterion that is independent of the specific application.

2.2 Association Rule mining

As proved my Mobasher et.al [27] Association Rule mining (AR mining) is a major pattern discovery technique. The association rule or frequent item sets mining algorithm was originally proposed by Agarwal et al. [1] for market basket analysis. With its significant applicability, many revised algorithms have been introduced and, AR mining is still a wide research area. Association rule discovery on usage data results in finding groups of pages that are commonly accessed. The applications of association rules are far beyond market basket applications and they have been used in various domains including Web mining.

Mobasher et al. [3] proposed an effective technique for capturing common user profiles based on association-rule discovery and usage-based clustering. They proposed techniques for combining these user profiles, with the current status of an ongoing Web activity to perform real-time personalization, taking into account both the offline tasks and the online process of automatic Web page customization. Przemysław Kazienko [36] presented a new approach by mining indirect association rules, relating them to the direct association rules, joined into one set of complex association rules which is then used for the recommendation of web pages. Yong et al. [50] gave algorithms for mining sequential association rules, based on different sequence and temporal constrains combination. The performance of these algorithms was compared on a real web log dataset by the method of variance analysis. Finally they proved that the sequence constrains, the temporal constrains and the interaction between these two constrains can affect the precision of prediction. They also concluded that temporal constrains can affect more than sequence constrains. B.Santhosh Kumar and K.V.Rukmani [38] discovered the web usage patterns of websites from the server log files using Apriori algorithm and Frequent Pattern Growth algorithm.

The main problem associated with association rule mining is the frequent item problem where the items that occur together with a high frequency will also appear together in many of the resulting rules and thus, resulting in inconsistent predictions. As a consequence, a system cannot give recommendations when the data set is large. In addition to this, AR Algorithms using multiple support thresholds results in better coverage but did not improve accuracy [24]. AR Algorithms where most frequent item sets are stored in data structure, using an algorithm to recognize most suitable items, cause scalability problem and low coverage [33]. AR Algorithms with large transactions would lead to redundant and complex rules [27].

2.3 Sequential pattern Mining

Sequential patterns in Web usage data capture the web page trails that are often visited by users, in the order that they were visited. These are sequences of items that occur in a sufficiently large proportion of (sequence) transactions. The view of web transactions as sequences of pageviews paved way to a number of useful and well-studied models in discovering user navigation patterns. One such approach is to model the navigational activities in the website as a Markov Model (MM): each pageview in this model can be represented as a state and the transition probability between two states can represent the likelihood that a user will navigate from one state to the other. This representation allows for the computation of a number of useful user or site metrics. Lower order markovian model lack accuracy because of its limitation of covering enough browsing history. Higher-order Markov models generally provide a higher prediction accuracy but result in much higher model complexity due to the larger number of states. Pitkow et al. [35] proposed all-*k*th-order Markov models (for coverage improvement) and a new state reduction technique, called longest repeating subsequences to overcome the coverage and space complexity problems (for reducing model size). The use of all-*k*th-order Markov models generally requires the generation of separate models for each of the *k* orders. If the model cannot make a prediction using the *k*th order, it will attempt to make a prediction by incrementally decreasing the model order. This scheme can easily lead to even higher space complexity since it requires the representation of all possible states for each k. Deshpande et al. [10] proposed selective markov models in which they proposed three different techniques to overcome the space complexity of existing all-kth-order Markov models. The proposed schemes involve pruning the model based on criteria such as support, confidence, and error rate. Confidence pruned MM generates all the states irrespective of their frequencies. In particular, the support-pruned MM eliminates all states with low support determined by a minimum frequency threshold. Anderson et al. [2] proposed Relational Markov models, a generalization of Markov models where states can be of different types, with each type described by a different set of variables. This model tends to perform better to existing models when data about all states is available in quantity. Wang et al. and Galassi et al. [16] [47] proposed Hidden markov models. The Hidden Markov Model starts with a finite set of states. Transitions among the states are governed by a set of probabilities (transition probabilities) associated with each state. In a particular state, an outcome or observation can be generated according to a separate probability distribution associated with the state. It is only the outcome, not the state that is visible to an external observer. The states are "hidden" to outside; hence the name Hidden Markov Model. Christopher et al. [29] proposed VOGUE, Variable Order and Gapped HMM for Unstructured Elements relies on a variable gap sequence mining method to extract frequent patterns with different lengths and gaps between elements. These patterns are then used to build a variable order hidden Markov model that explicitly models the gaps.

2.4 Classification

Classification is the task of mapping a data item into one of several predefined classes. The goal of supervised learning is to build a concise model of the distribution of class labels in terms of predictor features. The resulting model is then used to assign class labels to the testing instances. A large number of methods based on the model essence have been developed, and the choice of the method always depends on the task at hand. Under this heading we describe about Decision trees a logical or symbolic technique; Naive Bayesian classifier a statistical technique, *k*-nearest neighbor classifier an instance based learning technique and a special classification technique Support Vector Machines.

2.4.1 Decision Trees

Murthy [30] provided an overview of work in decision trees and a sample of their usefulness to newcomers in the field of Data Mining. Elomaa [14] presented a comparative study of well-known pruning methods and concluded that there is no single best pruning method. Bruha, [4] proposed that not only post processing but also preprocessing algorithms for decision tree construction can be found . Zidrina Pabarskaite [55] proposes decision trees for web user behaviour analysis. This analysis predicts user future actions and typical pages that lead to browsing termination. Using Decision tree package C4.5, Olcay and Onur [31] show how to parallelize C4.5 algorithm in three ways: (i) feature based, (ii) node based (iii) data based manner. To sum up, one of the most useful characteristics of decision trees is their comprehensibility. Decision trees tend to perform better when dealing with discrete/categorical features.

2.4.2 Naïve Bayes Classifier

Bayesian networks are the most well known representative of statistical learning algorithms. The major advantage of the Naive Bayes classifier is its short computational time for training. In addition, since the model has the form of a product, it can be converted into a sum through the use of logarithms – with significant consequent computational advantages. Domingos & Pazzani [12] performed a large-scale comparison of the naive Bayes classifier with state-of-the-art algorithms for decision tree induction, instance-based learning, and rule induction on standard benchmark datasets, and found it to be sometimes superior to the other learning schemes, even on datasets with substantial feature dependencies. Deng et al. [22] proposed spy Naïve Bayes to identify the user preference pairs generated from click through data. Santra and Jayasudha [39] used Naive Bayesian Classification algorithm for classifying the interested users. They measured the performance of this algorithm on web log data with session based timing, page visits, repeated user profiling, and page depth to the site length and concluded that the memory and time taken to classify the web log files are more efficient when compared to existing C4.5 algorithm.

2.4.3 k-Nearest Neighbour

k-Nearest Neighbour (kNN) is based on the principle that the instances within a dataset will generally exist in close proximity to other instances that have similar properties. As kNN does not make any assumptions on the underlying data distribution and does not use the training data points to do any generalization, it is called as non parametric lazy learning algorithm. Guo et al. [17] proposed a novel kNN type method for classification that is aimed at overcoming the drawback of its dependency on the selection of a “good value” for k. Yu & Liu [51] addressed the problem of determining which of the available input features should be used in modeling via feature selection because it could improve the classification accuracy and scale down the required classification time.

2.4.4 Support Vector Machines

Support Vector Machines (SVMs) are the newest supervised machine learning technique [46]. Burges, [6] gave an excellent survey of SVMs, and a more recent book about SVMs is by Cristianini & Shawe-Taylor [9]. Yuh-Jye Lee and O.L. Mangasarian [52] presented a support vector machine for pattern classification using a completely arbitrary kernel. Sung-Hae Jun [45] used SVMs to analyze web log data and estimated the dependency between the web pages overcoming the difficulty of sparsity. Satoshi Mizuno [41] proposed a method that creates user’s profile from browsing history using Term Frequency Inverse Document Frequency and then classifies the URL’s of the browsing history using SVM.

2.5 Mixture Models

Mixture Models play an important role in Classification. In order to identify the proper model to classify the data, in these models we assume that the behavior of each user in the data set is generated independently, and the behavior is generated by a mixture model with K components. In a mixture model, we are concerned with (1) the number of components; (2) the probability distribution used to assign users to the various clusters, and (3) the parameters of each model component. Once the model is estimated, we can use it to assign each user to a cluster or fractionally to the set of clusters. Yanzan Kevin Zhou and Bamshad Mobasher [48] proposed an approach for Web user segmentation and online behavior analysis based on a mixture of factor analyzers. In this framework, they modelled users’ shared interests as a set of common latent factors extracted through factor analysis, and discovered user segments based on the posterior component distribution of a finite mixture model. This measured the relationships between users’ unobserved conceptual interests and their observed navigational behavior in a principled probabilistic manner. Attenberg et al. [19] developed a generative model to mimic trends in observed user activity using a mixture of pareto distributions. Mihajlo Grbovic et al. [26] proposed, time- and memory-efficient algorithm for learning label preferences based on the Gaussian Mixture Model (GMM), this model turned to be attractive because of an intuitively clear learning process and ease of implementation. Rekha et al. [37] proposed Adaptive Gaussian Mixture model for user behavior modeling. The developed method as shown a drastic improvement in identifying the navigational pattern of user compared to GMM.

2.6 Collaborative Filtering (CF)

Collaborative filtering (CF) is a technique utilized primarily to predict individuals' preferences, has its origin in information filtering. This technique guides an active user depending on the preferences shared by like users. Once a database of preferences of like users is accumulated, a similarity measure is used to identify individuals with similar past preferences with the active user. A preference function is applied on the database to guide or recommend the active user [18].

This technique is easy to comprehend and implement, but requires a large sample to make meaningful recommendations. Erroneous recommendations result when close neighbors don't exist. Content information and customer profile or behavior information is not used for making recommendations. As database size increases, the recommendation computation becomes computationally more intensive. These also suffer from a fundamental problem, called sparsity problem. Since the set of all possible available items in a system is very large, most users may have rated very few items, and, hence, it is difficult to find the active user's neighborhood with high similarity. As a result the accuracy of the recommendations may be poor [34].

To overcome the above disadvantages classification and prediction had its application in the web domain of collaborative filtering. Lin et al. [23] proposed a collaborative recommendation system using association rules. Zhong hang Xia et al. [54] proposed a collaborative filtering system with SVM. Koji Miyahara and Michael J. Pazzani [21] proposed a collaborative filtering system with Bayesian classifier. Miha Grcar et al. [25] presented experimental results of confronting the k-Nearest Neighbor (kNN) algorithm with SVM in the collaborative filtering framework using datasets with different properties. Dhruv Gupta et al. [11] emphasized on a new, principal component analysis and clustering-based linear time collaborative filtering algorithm for efficient and effective personalized information retrieval.

III. CHALLENGES

In order to navigate the user's behavioral patterns, the data stored in the web log is of crucial importance. This data generally will be in unstructured format and hence to analyze this data efficient methodologies are to be developed. The literature developed in this regard exhibits inconsistency, incorrect and missing values. Therefore advanced methodologies that can navigate the data more efficiently by minimizing the inconsistent data to retrieve the webpage's of users interests is the concern of the day. Hence efficient clustering and classification algorithms together with effective preprocessing techniques are to be developed.

IV. CONCLUSION

This paper gives an insight into the possible data mining techniques with Web usage data for achieving a synergetic effect of Web usage mining. Association rules are used to discover pages that are visited together quite often. Discovering sequential patterns from web access logs can be used for predicting future visits of the users. Clustering discovers groups of users or pages, based on their similarities. Classification classifies the new user into one of the predefined groups based on their maximum likelihood. It is hard, if not impossible, to declare that one data mining algorithm is the best in general, because the possible outcomes of WUM process always depend on the problem in hand.

References

[1] Agarwal R. and Srikant R., "Fast algorithms for mining association rules", *VLDB'94*, Chile, pp. 487–499, 1994.

[2] Anderson C., Domingos P., Weld D. S., "Relational Markov Models and their Application to Adaptive Web Navigation", *Proceedings of the 8th ACM SIGKDD Conference*, Canada, August 2002.

[3] Bamshad Mobasher, Robert Cooley, Jaideep Srivastava," Creating Adaptive Web Sites Through Usage-Based Clustering of URLs", *proceedings of the 1999 workshop on knowledge and data engineering*, pp 19, 1999.

[4] Bruha I., "From machine learning to knowledge discovery: Survey of preprocessing and postprocessing" *Intelligent Data Analysis*, Vol. 4, pp. 363-374, 2000.

[5] Buchner A. and Mulvenna M. D., "Discovering Internet Marketing Intelligence through Online Analytical Web Usage Mining", *Proceedings of the ACM SIGMOD, Intl.Conf. on Management of Data (SIGMOD'99)*, pp. 54–61, 1999.

[6] Burges C., "A tutorial on support vector machines for pattern recognition", *Data Mining and Knowledge Discovery*, Vol. 2, pp. 1-47, 1998.

[7] Cheng D., Kannan R., Vempala S. and Wang G., "A divide-and-merge methodology for clustering", *ACM SIGMOD*, pp. 196–212, 2005.

[8] Cooley R. , Mobasher B., and Srivastava J. , "Data Preparation for Mining World Wide Web Browsing Patterns", *Knowledge and Information Systems*, vol. 1(1), pp. 5–32, 1999.

[9] Cristianini N. and Shawe-Taylor J., "An Introduction to Support Vector Machines and Other Kernel-Based Learning Methods", Cambridge University Press, Cambridge, 2000.

[10] Deshpande M., Karypis G., "Selective Markov Models for Predicting Web-Page Accesses", *Proceedings of the 1st SIAM International Conference on Data Mining*, 2004.

[11] Dhruv Gupta, Mark Digiovanni, Hiro Narita, and Ken Goldberg, "Jester 2.0 : Evaluation of a New Linear Time Collaborative Filtering Algorithm", *SIGIR* '99 Berkley, CA, USA ,ACM, 1999.

[12] Domingos P. and Pazzani M. , "On the optimality of the simple Bayesian classifier under zero-one loss". *Machine Learning*, pp.103-130, 1997.

[13] Dr. G T Raju and Dr. M V Sudhamani, "A Novel Approach for Extraction of Cluster Patterns from Web Usage Data and its Performance Analysis", *IEEE*, 2011

[14] Elomaa T., "The biases of decision tree pruning strategies", *Lecture Notes in Computer Science* 1642. Springer, pp. 63-74, 1999.

[15] Fayyad U. M. , Piatetsky-Shapiro G., and Smyth P. "From Data Mining to Knowledge Discovery: An Overview", *Advances in Knowledge Discovery and Data Mining*, AAAI/MIT Press, pp. 1–34, 1996.

[16] Galassi U., Botta M., and Giordana A., "Hierarchical hidden markov models for user process profile learning", *Fundamenta Informatica* 78, vol. 4, pp. 487–505, 2007.

[17] Guo G., Wang H., Bell D., Bi Y., and Greer K. "KNN Model-Based Approach in Classification", *Lecture Notes in Computer Science*, Vol 2888, Pages 986 – 996, 2003.

[18] http://www.fico.com/en/Communities/AnalyticTechnologies/Pages/CollaborativeFiltering.aspx.

[19] Josh Attenberg, Sandeep Pandey, Torsten Suel, "Modeling and Predicting User Behavior in Sponsored Search" *KDD*'09, June 28–July 1, ACM 978-1-60558-495-9/09/06, 2009.

[20] Kim D., Adam N., Alturi V., Bieber M. and Yesha, Y., "A click stream based collaborative filtering personalization model: Towards a better performance",*WIDM* '04, pp. 88–95, 2004.

[21] Koji Miyahara and Michael J. Pazzani, "Collaborative Filtering with the Simple Bayesian Classifier", *Proceedings of the 6th Pacific Rim International conference on Artificial intelligence*, pp. 679-689 ,Springer-Verlag Berlin, Heidelberg, 2000.

[22] Lin et al., "Spying Out Real User Preferences for Metasearch Engine Personalization", *Proceedings of the WEBKDD'04 in conjunction with KDD'04*, August 22, 2004.

[23] Lin W., Alvarez S. A., and Ruiz C., "Efficient Adaptive-Support Association Rule Mining for Recommender Systems", *Data Mining and Knowledge Discovery*, vol. 6, pp. 83–105, 2002.

[24] Liu B., Hsu W. and Ma Y., "Mining association rules with multiple minimum support", *KDD*, San Diego, pp. 337–341, 1999.

[25] Miha Grcar, Miha Grcar, and Dunja Mladenic, "kNN Versus SVM in the Collaborative Filtering Framework", WebKDD '05, August 21, Chicago, Illinois, ACM 1-59593-214-3, 2005.

[26] Mihajlo Grbovic and Nemanja Djuric and Slobodan Vucetic, "Learning from Pairwise Preference Data using Gaussian Mixture Model", *Preference Learning Workshop, European Conference on Artificial Intelligence*, 2012.

[27] Mobasher B., Dai H., Luo T. and Nakagawa M., "Discovery of aggregate usage profiles for web personalization", *WebKDD*'00, USA pp. 61–82, 2000.

[28] Mobasher B., Dai H., Luo T. and Nakagawa M., "Effective personalization based on association rule discovery from web usage data", *WIDM'01*, USA, 2001.

[29] Mohammed J. Zaki Christopher ,D. Carothers and Boleslaw K. Szymanski, "VOGUE: A Variable Order Hidden Markov Model with Duration based on Frequent Sequence Mining ", *ACM Transactions on Knowledge Discovery from Data*, vol.4(1),article 5, January, 2010.

[30] Murthy, "Automatic Construction of Decision Trees from Data: A Multi-Disciplinary Survey", *Data Mining and Knowledge Discovery* pp. 345–389, 1998.

[31] Olcay Taner Yıldız and Onur Dikmen , "Parallel univariate decision trees", *Pattern Recognition Letters*, Vol. 28 , Issue. 7, pp. 825-832, May 2007.

[32] Pablo Loyola, Pablo E. Rom´an and Juan D. Vel´asquez, "Clustering-Based Learning Approach for Ant Colony Optimization Model to Simulate Web User Behavior",*IEEE*, 2011.

[33] Park J. S., Philip S. Y. and Chen M. S., "Mining association rules with adjustable accuracy", *CIKM'97*, pp. 151–160, 1997.

[34] Paul et al., "GroupLens: an open architecture for collaborative filtering of netnews", CSCW '94 *Proceedings of the 1994 ACM conference on Computer supported cooperative work*, pp 175-186 , 1994.

[35] Pitkow J. and Pirolli P., "Mining Longest Repeating Subsequences to Predict WWW Surfing", *Proceedings of the 2nd USENIX Symposium on Internet Technologies and Systems*, 1999.

[36] Przemyslaw Kazienko, "Mining Indirect Association Rules For Web Recommendation", *Int. J. Appl. Math. Comput. Sci*., Vol. 19, No. 1, pp. 165–186, 2009.

[37] Rekha Sundari M., Prasad Reddy PVGD and Srinivas y., "User Behavior Modeling based on Adaptive Gaussian Mixture Model", *International Journal of Computer Applications* 60(4):1-3, December 2012.

[38] Santhosh Kumar B. and Rukmani K.V., "Implementation of Web Usage Mining Using APRIORI and FP Growth Algorithms", *International Journal of Advanced Networking and Applications*, Vol.01, Issue.06, pp. 400-404, 2010.

[39] Santra A. K. and Jayasudha S., "Classification of Web Log Data to Identify Interested Users Using Naïve Bayesian Classification", *IJCSI International Journal of Computer Science* Issues, Vol. 9, Issue 1, No 2, January 2012.

[40] Sasa Bosnjak ,Mirjana Maric, and Zita Bosnjak, "The Role of Web Usage Mining in Web applications Evaluation", *Management Information Systems,*Vol. 5, No. 1, pp. 031-036, 2010.

[41] Satoshi Mizuno, "Personalized Web Search System with Categorization using SVM", *University of Aizu, Graduation Thesis*,March, 2009.

[42] Sebastian A. Rios, Roberto A. Silva, and Felipe Aguilera., "A dissimilarity measure for automate moderation in online social networks", *WI&C*, 3, ACM, 2012.

[43] Sophia G. Petridou, Vassiliki A. Koutsonikola, Athena I. Vakali, and Georgios Papadimitriou, "Time Aware Web Users Clustering" 1041-4347, IEEE, 2007.

[44] Srivastava J., Cooley R., Deshpande M. and Tan P., "Web usage mining: Discovery and applications of usage patterns from web data", *SIGDD Explorations*, pp. 12–23, 2000.

[45] Sung-Hae Jun, "Web usage mining using support vector machine", WANN'05 *Proceedings of the 8th international conference on Artificial Neural Networks: computational Intelligence and Bioinspired Systems*, pp. 349-356, 2005.

[46] Vapnik. V, "The Nature of Statistical Learning Theory", *Springer Verlag*, 1995.

[47] Wang et al., "Mining complex time-series data by learning markovian models", *6th IEEE International Conference on Data Mining*, 2006.

[48] Yanzan Kevin Zhou and Bamshad Mobasher, "Web User Segmentation Based on a Mixture of Factor Analyzers EC-Web 2006", *LNCS* 4082, pp. 11–20, 2006.

[49] Yinghui Yang and Balaji Padmanabhan, "A Hierarchical Pattern-Based Clustering Algorithm For Grouping Web Transactions", *IEEE Transactions On Knowledge And Data Engineering*, Vol. 17, No. 9, September 2005.

[50] Yong Wang, Zhanhuai Li and Yang Zhang, "Mining Sequential Association-Rule For Improving Web Document Prediction", *ICCIMA*'05, pp. 146–151, 2005.

[51] Yu L. and Liu H. , "Efficient Feature Selection via Analysis of Relevance and Redundancy", *JMLR*, 5(Oct):1205-1224, 2004.

[52] Yuh-Jye Lee and O.L. Mangasarian, "SSVM: A Smooth Support Vector Machine for Classification", *Computational Optimization and Applications*, pp 5–22, 2001.

[53] Zhong S. and Ghosh J., "A unified framework for model-based clustering", *Machine Learning Research* 4, 1001–1037, 2003.

[54] Zhonghang Xia Yulin Dong and Guangming Xing, "Support Vector Machines For Collaborative Filtering", *ACM SE'06*,Melbourne, Florida, USA., pp 10-12, March 2006.

[55] Zidrina Pabarskaite, "Decision trees for web log mining", *Journal*

[56] *Intelligent Data Analysis*, Vol. 7, Issue. 2, pp. 141 - 154, April 2003.