An Overview on the Internet Users Behavior towards E-Commerce

K.Pariplavi[1], B. Raghuram[2]
[1]Student, Master of Technology, KITS, Warangal
[2]Assistant Professor, Department of CSE, KITS, Warangal

Abstract

Data mining has matured as a field of fundamental and applied research in computer science. The objective of this dissertation is to evaluate, propose and improve the use of some of the recent approaches, architectures and Web mining techniques (collecting personal data from customers) are the means of utilizing data mining methods to induce and extract useful data from Web data and service where data mining has been applied in the fields of e-commerce and e-business (that means User's behavior). In the context of web mining, clustering could be used to cluster comparative snap streams to determine learning behaviors on account of e-learning or general site access behaviors in e-commerce. A large portion of the calculations presented in the literature to deal with clustering web sessions treat sessions as sets of visited pages inside a time period and don't consider the sequence of the snap stream appearance. This has a huge consequence when looking at similarities between web sessions. Wang and Zaiane propose a calculation based on sequence alignment to measure similarities between web sessions where sessions are sequentially ordered sequences of page accesses.

Index Terms : User behavior,web mining, e-commerce, data mining, clustering

I. Introduction

User Behavior: notwithstanding understanding the needs of the customers, Company additionally need to understand what motivates them to purchase, and in what manner can influence the purchasing process to ensure that the items or services are on the shopping list.

Understanding the customers will help, to develop and distribute the item, and in addition getting the correct price point and developing successful special activities.

The brain science of the purchasing process has been widely studied and regardless of what size organization business, knowledge of this process can help organization become more successful.

The two businesses and consumers exhibit patterns of purchasing behavior. The business model is less open to debate as the business customers will more likely than not have some formalized process of purchasing in place. The organization undertaking is to understand the process and match the marketing activities to the different stages of the process. This means the customer will receive the correct sort of contact at the ideal time.

To provide bits of knowledge into areas, for example, pointers of customer defection, price sensitivity, segmentation, and customer needs examination, to name a few.

Our R and D work in this area has exposed numerous generally held marketing beliefs as mere mythologies with no legitimate scientific premise. We employ this unique knowledge to answer questions have, for example, What are the leading markers of customer defection in our industry? What indicates the 'health' of our image? What's more, what does this mean for us? By what means can predictive modeling help me determine the outcome of our marketing activities? What are customer's reservation prices in our industry? How might we use this to predict the effect of price changes? What are our customer's needs and how do elements, for example, age or gender affect? Which brands do we compete most closely with?

Business Buying Behavior: A run of the mill business customer will experience the accompanying steps when purchasing: Identifying a need or problem: This might be highlighted by press coverage or advertising they have seen in the trade press. Developing item specification: The customer will use whatever sources they can discover to help them specify what they need. They will give careful consideration to press releases, exhibitions, advertising, editorial comment, industry seminars and relevant direct mail. Search for items and suppliers: This is the time when the business customer is especially open to visits from your sales force and trade directory entries. Exhibitions and technical data leaflets are additionally invaluable sources. This is the time when estimating data begins to be seriously considered.

Evaluation of items and suppliers: This is a decent time to provide your potential customer with demonstration items, visits to existing customers, plant visits or outsider testimonials. You may likewise need to take a gander at special estimating packages or stocking incentives.

Evaluation of item and supplier performance: The more major the purchasing decision, the more reassurance your customer needs. Review meetings and helpline bolster provide reassurance, as does great after sales bolster and continued exposure to advertising and

press coverage - defending the purchase decision.

Take after on purchase: The principal purchase ought not be seen as the end of the process, yet the beginning of a long haul business relationship.

Consumer Buying Behavior: There are numerous models of consumer purchasing behavior, yet the steps below are genuinely normal to the majority of them.

The customer identifies a need: This is often initiated by PR coverage, including verbal. The customer may have seen a friend or celebrity utilizing an item or service, or awareness may have been sparked off by advertising.

Searching for data: At this stage the customer needs to know more and is actively seeking data. Advertising and PR are as yet essential however item demonstrations, bundling and item shows assume a role. This is the time to deploy your sales personnel, and customers discover videos and brochures are useful. Informal exchange is still very imperative.

Checking out alternative items and suppliers: The customer is currently endeavoring to choose between items, or solidify on the purchase decision. This is a place for advancing item guarantees and warranties, and augmenting bundling and item shows. Sales personnel can greatly influence the customer at this stage and sales advancement offers become of interest. Independent sources of data are still of interest, including item test reviews.

Purchase decision: This is the time to 'tip the balance'. Sales advancement offers come into their own, and if appropriate, sales force incentives need to ensure that your sales personnel are incentives to close the deal.

Utilizing the item: Expensive purchases can lead to what is known as cognitive dissonance - a fear that the customer has not made the correct decision. Your activity is to reassure the customer by offering great customer care, simple direction manuals and reliability schemes. They should even now be exposed to testimonial advertising to reassure them that they have made the correct decision.

Marketing does not stop at understanding the purchasing processes of the customer however, organization need to understand their purchasing patterns and the market in which they operate.

It is presently a platitude that in the times of the corner market, shopkeepers experienced no difficulty understanding their customers and responding rapidly to their needs. The shopkeepers would basically keep track of the greater part of their customers in their heads, and would comprehend what to do when a customer walked into the store. Be that as it may, the present shopkeepers face a substantially more complex circumstance. More customers, more items, more competitors, and less time to react means that understanding the customers is presently significantly harder to do. A number of forces are cooperating to increase the complexity of customer relationships:

Compressed marketing cycle times: The attention traverse of a customer has decreased significantly and dependability is a relic of times gone by. A successful organization needs to reinforce the value it provides to its customers consistently. Moreover, the time between a new desire and when customer must meet that desire is additionally contracting. In the event that organization doesn't react rapidly enough, the customer will discover someone who will.

Increased marketing costs: Everything costs more. Printing, postage, special offers (and if organization doesn't provide the special offer, the competitors will).

Streams of new item offerings: Customers need things that meet their exact needs, not things that kind of fit. This means the number of items and the number of ways they are offered have risen essentially.

Niche competitors: The best customers additionally look great to the competitors. They will center around little, profitable segments of the market and endeavor to keep the best for themselves.

The correct offer means dealing with multiple interactions with the customers, organizing what the offers will be while ensuring that irrelevant offers are minimized. The perfect person means that not all customers are fundamentally the same. The organization interactions with them need to move toward very segmented marketing efforts that target singular needs and needs. The ideal time is a result of the way that interactions with customers now happen consistently. This is essentially different from the past, when quarterly mailings were bleeding edge marketing. At last, the correct channel means that organization can interact with the customers in a variety of ways (direct mail, email, telemarketing, etc.). The organization needs to make sure that are picking the best medium for a specific interaction.

Data Mining: Data mining, by its simplest definition, automates the detection of relevant patterns in a database. For example, a pattern may indicate that married males with children are twice more likely to drive a specific games auto than married males without any children. In the event that any persons are a marketing manager for a vehicle manufacturer, this somewhat shocking pattern may be quite valuable.

Web Data Mining: Web data mining is one sort of these techniques that efficiently handle the undertakings of searching the needed data from the Internet, enhancing the Web site structure to provide better Internet service quality and discovering the informative knowledge from the Internet for advanced Web applications. Web data mining could be categorized into three types of Web content, Web structure1 and Web usage mining2,3,4. In this investigation, we center around Web usage mining: that is, discovering user

access pattern knowledge from Web log files, which contain the noteworthy going by records of users on the website.

II. Technical Issues

Benefits for the companies: Today, the enormous content of the Internet has made it hard to discover relevant data on a subject. Methods helping user route and retrieving data have become especially essential. Online shops need to offer personalized items to clients however before being able to do that they have to personalize the web sites to the clients. This is where the data mining techniques in web server logs are coming in. Companies can use the fundamental data retrieved from the data logs to analyze customer behaviors, evaluate the current usage, if the customers liked or disliked it et cetera. To create adaptable web sites to each user, to start with, the user route patterns in the web have to be found and analyzed. Data mining is a method extracting valuable data from the data for factual purpose.

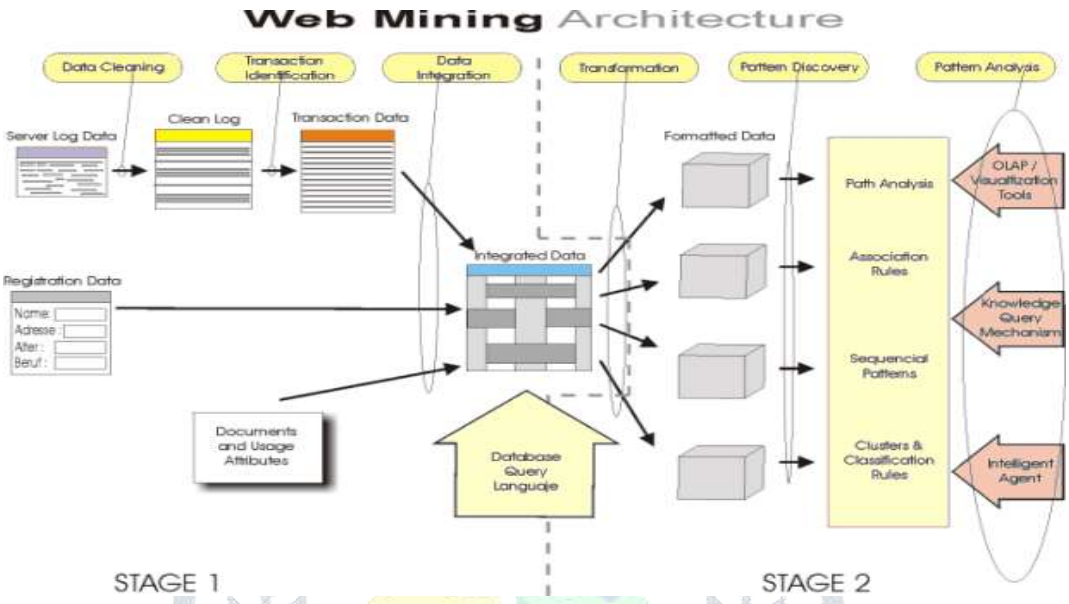

Figure-1 Architecture of Web mining

How does it work? Data mining techniques: One of the most known data mining techniques is WEBSOM (web based self-organizing map), which organizes documents into two-dimensional map according to their content rather than by keyword.
Unlocking the usage patterns of the web users hidden in the log files has therefore been a challenge for several researchers. Fu, Perkowitz and Etzioni[5] have demonstrated that web users can be grouped into meaningful clusters, which help the web designers to provide high-value customized services as the data mining of the web server logs provides them with the information needed to understand users better. These systems can also be used to improve current WebPages.
Mobasher[6,7] created a Web Personalization system that organizes web usage data not the content of the data mentioned above into clusters. The system analyzes the web server logs, it identifies to which user group the current user belongs to and makes suggestions to links that would interest the user. These suggestions are based on the past experience of a particular user group.
LOGSOM (log based self-organizing map) combines the benefits of the both of these systems. It keeps a track and organizes the web pages according to the user navigation behaviors and interest not to the web content[8].

More detailed view how do they actually do it? Data caching algorithms: The general idea is to make the full use of the web log data utilizing data mining applications. Data mining is aimed to discover the models, structure, content, usage patterns et cetera of users and web pages. Intelligent web storing calculations are the apparatuses to predict the web requests. "These web-storing calculations are able to adjust their behavior based on the user access patterns, which in step are extracted from the authentic access data recorded in the log files by means of data mining techniques."

The objective of these calculations is to increase the number of web pages that are retrieved directly from the cache instead requesting them from the server. There are several web storing calculations, we list three of them:

Frequent patterns In the case of frequent patterns, we extract from the web logs the patterns that take after the shape A□B (if A, then B). In the event that A has been requested then B is likely to be requested next.

Decision trees In the case of decision trees, we develop a decision tree in a premise of the chronicled data in the web logs yet on this case concentrating on the time needed until the next request.

Clickstream data is the way that the user creates when steering through the sites and following connections. It can be used to evaluate the activity and prevalence of the page

Shopping outline can provide data in e-business where the purchases were made and where the customer left the order unfinished.

Psychographic data would include data on user's attitudes towards subjects, items etc., purchasing behavior and beliefs.

Social issues: What Do Users Think? As data mining apparatuses and calculations become more sophisticated and widely available, customer's protection concerns are always increasing. These concerns are especially high due to chance of World Wide Web to easily naturally collect consumer data and add it to databases. With associations increasingly assembling comprehensive consumer databases and applying sophisticated data-mining techniques protection and ethics issues become more pressing.

The user's consider the accompanying four questions1,9, Does Internet Data Mining Violate Users' Privacy? What is the User Persistence of Internet Data Mining? What would marketers be able to do about Internet Data Mining Privacy? What would users be able to do about Internet Data Mining Privacy.

From the results above 4 hints for online marketers can be suggested to improve user trust and data mining techniques' security:

Explain the purpose of data mining: Data mining dependably has a certain objective. "The survey found that in any event some segments of online users lose their negative way to deal with Internet data mining when they better understand its purpose".

Control of data appropriation: Users are in dominant part unequivocally against the sharing of mined data among different companies.

Provide key trust focuses to improve e-commerce practice: That implies clear statements of protection strategy. Likewise security software and description or correspondence about that.

Why mine e-commerce and snap stream data? Improve conversion rate through personalization. Optimize marketing efforts (banners, email, and other media) that convey guests to your site by measuring return on investment (ROI). Improve basket size through strategically pitches and up-sells. Streamline route ways through the site. Dodge content delivery issues (inadequately formatted for AOL, excessively rich for low data transfer capacity users, redundant or confounding content). Identify customers segments that you can target offline. Experiment rapidly. The Web is a research center. Understand what works rapidly.

III. Material and Methods

Clustering Algorithm: Clustering investigation is a widely used data mining calculation for some data management applications. Clustering is a process of parceling a set of data objects into a number of object clusters, where each data object shares the high comparability with the other objects inside the same cluster however is quite unlike objects in other clusters. Different from grouping calculation that appoints a set of data objects with different labels previously defined through a supervised learning process, clustering examination is to segment data objects objectively based on measuring the shared similitude between data objects, i.e. by means of an unsupervised learning process. Due to the way that the class labels are often not known before data investigation, for example, in case of being difficult to allocate class labels in large databases, clustering examination is sometimes an efficient approach for dissecting such sort of data. To perform clustering investigation, closeness measures are often utilized to assess the distance between a couple of data objects based on the feature vectors describing the objects, thusly, to help allocating them into different object classes/clusters. There are a variety of distances capacities used in different scenarios, which are really dependent on the application foundation. For example, cosine work and Euclidean distance work are two normally used distance works in data retrieval and pattern recognition10. Then again, assignment strategy is another vital point involved in apportioning the data objects. Therefore, distance capacity and assignment calculation are two core research focuses that pull in a ton of efforts contributed by different research area experts, for example, from database, data mining, insights, business intelligence and machine learning etc.

User Profile Algorithms: Web clustering is one of the generally used techniques in the context of Web mining, which is to aggregate comparative Web objects, for example, Web pages or users session, into a number of object bunches by means of measuring their shared vector distance. The resulting Web user session bunches are considered as representatives of user navigational behavior patterns, while Web page clusters are used for generating errand oriented usefulness aggregations of Web associations. Moreover, the mined usage knowledge in terms of Web usage patterns and page aggregates can be utilized to improve Web site structure designs.

Latent Usage Information Algorithm: In this section, we present a calculation called latent Usage Information (LUI) for clustering Web sessions and generating user profiles based on the discovered clusters. This calculation comprises of two steps, the initial step is a clustering calculation, which is to cluster the converted latent usage data into a number of session gatherings; and the next step is tied in with generating a set of user profiles, which are derived from figuring the centroids of the discovered session clusters.

Building User Profile: As we mentioned above, each user session is represented as a weight-based page vector. Along these lines, it is reasonable to derive the centroid of the cluster obtained by the described clustering calculation as a user profile. In this work, we

compute the mean vector to represent the centroid.

Experimental Results: In order to evaluate the effectiveness of the proposed LUI calculation, which comprises of the Web clustering calculation and the user profile generating calculation, and evaluate the discovered user access patterns, we direct experiments on two real world data sets and make correlations with the previous work.

Results of User Profiles: We initially utilize LUI calculation to direct Web usage mining on the selected two usage datasets respectively. We tabulate some results in below table-1 and table-2. In these tables, each user profile is represented by a sequence of noteworthy pages together with corresponding weights. As we indicated before, the calculated weight is expressed in a normalized frame, that is, the biggest value of them is set to be 1 while others are the relatively corresponding values, which are constantly less than 1.

Table-1 depicts 2 user profiles generated from KDD dataset utilizing LUI approach. Each user profile is listed in an ordered page sequence with corresponding weights, which means the greater weight a page contributes, the more likely it is to be visited. The main profile in table-1 represents the activities involved in online shopping behaviors, for example, login, shopping basket, and checkout operation etc, especially occurred in acquiring leg-wear items, whereas the second user profile reflects the customers' concern with regard to the department store itself.

Table-1
Example of generated user profiles from KDD dataset

Page #	Page content	weight
29	Main-shopping_cart	1.00
4	Products-productDetailleagwear	0.86
27	Main-Login2	0.67
8	Main-home	0.53
44	Check-express_Checkout	0.38
65	Main-welcome	0.33
32	Main-registration	0.32
45	Checkout-confirm_order	0.26
Page #	Page content	weight
11	Main-vendor2	1.00
8	Main-home	0.40
12	Articles-dpt_about	0.34
13	Articles-dpt_about_mgmtteam	0.15
14	Articles-dpt_about_broadofdirectors	0.11

Table-2
Example of generated user profiles from CTI dataset

Page #	Page content	weight
29	Main-shopping_cart	1.00
4	Products-productDetailleagwear	0.86
27	Main-Login2	0.67
8	Main-home	0.53
44	Check-express_Checkout	0.38
65	Main-welcome	0.33
32	Main-registration	0.32
45	Checkout-confirm_order	0.26
Page #	Page content	weight
11	Main-vendor2	1.00
8	Main-home	0.40
12	Articles-dpt_about	0.34
13	Articles-dpt_about_mgmtteam	0.15
14	Articles-dpt_about_broadofdirectors	0.11

Similarly, some informative discoveries can be obtained in table-2, which is derived from CTI dataset. In this table, three profiles are generated: the first reflects the fundamental point of international students concerning issues regarding applying for confirmation, and the second one involves in the online applying process for graduation, whereas the last one indicates the most widely recognized activities happened amid students perusing the university website, especially while they are determining course selection, i.e. selecting course, searching syllabus rundown, and after that experiencing specific syllabus. Taking a gander at the generated user profile examples, it is demonstrated that the vast majority of them do reflect one specific navigational intention, however some may represent more than one access themes.

From the definition of WAVP, it is realized that the higher the WAVP value is, the better the nature of obtained session cluster possesses.

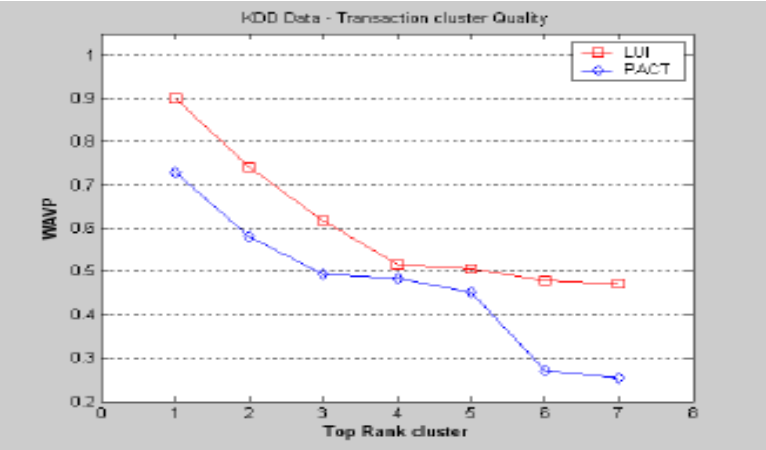

Figure-2 User cluster quality analysis results in terms of WAVP for KDD dataset

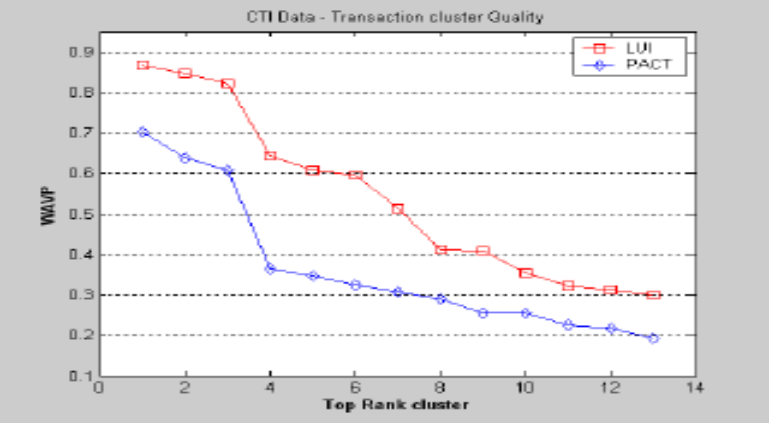

Figure-3 User cluster quality analysis results in terms of WAVP for CTI dataset

To compare the effectiveness and efficiency of the proposed calculation with existing calculations, here we use the PACT calculation. We lead data recreations upon two real world datasets by utilizing these two approaches. Figure-2 and figure-3 depict the correlation results in terms of WAVP values for KDD and CTI datasets with PACT respectively. In each figure, the obtained user profiles are arrayed in a descending rank as indicated by their WAVP values, which reflect the nature of different clustering calculations. From these two curves, it is easily concluded that the proposed LUI-based technique overweighs the standard k-means based calculation in term of WAVP parameter.

This is for the most part due to the unmistakable latent examination ability of LUI calculation. In other words, LUI approach is capable of catching the latent relationships among Web exchanges and discovering user profiles representing the genuine navigational patterns more effectively and accurately.

IV. Related Work and Discussion

In the context of Web usage mining, there are two types of clustering methods performed on the usage data Web exchange clustering and Web page clustering7. One successful use of Web page clustering is the adaptive Web site. For example, the calculation called PageGather5 is proposed to synthesize index pages that don't exist at first, based on parceling Web pages into different gatherings. The generated index pages are conceptually representing the different access interests of users as indicated by their navigational histories. Another example is that clustering user rating results has been successfully adopted in collaborative filtering applications as a data preparing step to improve the versatility of recommendation utilizing k-Nearest-Neighbor (kNN) algorithm11. Mobasher et al utilize Web exchange and page clustering techniques, which is employing the customary k-means clustering calculation to characterize user access patterns for Web personalization based on mining Web usage data6. These proposed clustering-based techniques have been proven to be efficient from their experimental results since they are really capable of identifying the inherent normal attributes revealed from their noteworthy clickstream data. Generally, these usage patterns are explicitly captured at the level of user session or page. They, however, don't reveal the underlying characteristics of user navigational activities and additionally Web pages.

V. Experimental Result

In this chapter, we have proposed a LSI-based approach, named LUI, for gathering Web exchanges and generating user profiles.

Right off the bat, we model the relationships among the co-occurrence observations (i.e. user sessions) into a usage data model as a session-page grid. Then, a dimensionality reduction calculation based on the SVD calculation has been employed on the usage lattice to capture the latent usage data for apportioning user sessions. Based on the decomposed latent usage data, we propose a k-means clustering calculation to generate user session clusters. Moreover, the discovered user bunches are utilized to develop user profiles expressed as a weighted page collection, which represent the regular usage pattern associated with one specific user access pattern. The constructed user profiles corresponding to different errand oriented behaviors are represented as a set of page-weight sets, in which each weight reflects the significance contributed by the page. Experiments have been conducted on two real world datasets to validate the effectiveness and efficiency of the proposed LUI calculation. Meanwhile, an evaluation metric is adopted to assess the nature of the discovered clusters in correlation with existing clustering calculations. The experimental results have demonstrated that the proposed approach is capable of effectively discovering user access patterns and revealing the underlying relationships among user going by records.

VI. **Results and Discussion**

The result and dialog of the paper will be clearly indicate what techniques (Web Mining), approaches and architectures is the fastest of using data mining methods to induce and extract useful data from Web data, services and merchandise online increases, data mining activities can expand quickly enabling firms to retrieve exceptionally personalized data about customers, which too implies high security infringement and concerns. The two marketers and users ought to take after protection strategy rules. Marketers should give careful consideration to level of user trust and couple their data mining efficiency with respect to user security. In this dissertation, Kurt Thearling provides a business and technological overview of data mining and outlines how, alongside sound business processes and complementary technologies, data mining can reinforce and redefine customer relationships.

VII. Conclusion

The result of this paper will be clearly indicated that how web mining (in an expansive sense, Data Mining applied to ecommerce) is applicable to enhancing the services provided by e-commerce based enterprises. Specifically, we initially discussed some recent approaches and techniques used in data mining. We now present some manners by which web mining can be extended for further research. With the developing interest in the idea of semantic web, an increasing number of sites use structured semantics and space ontologisms as a component of the site design, creation, and content delivery.

References

1.Shian-Hua Lin, Chi-Sheng Shih, Meng Chang Chen and et al., Extracting Classification Knowledge of Internet Documents with Mining Term Associations: A Semantic Approach. In Proceedings of 21st Annual International ACM SIGIR Conference on Research and Development in Information Retrieval, Melbourne, Australia (1998)

2.Cooley R., Mobasher B. what's more, Srivsatava J., Web Mining: Information and Pattern Discovery on the Word Wide Web, Technical Report TR 97-027, University of Minnesota, Dept. of Computer Science, Minneapolis, (1997)

3.Bray T., Measuring the Web, In Proceedings of the fifth Intl. WWW Conference, Paris, France (1996)

4.Chen M.S., Han J. what's more, Yu P.S., Data Mining: An Overview from a Database Perspective, IEEE Transaction on Knowledge and Data Engineering, 8, 866-833 (1996)

5.Perkowitz M. what's more, Etzioni O., Adaptive Web Sites: Conceptual Cluster Mining. In Proceeding of sixteenth International Joint Conference on Articial Intelligence, 264-269, Stockholm, Sweden: Morgan Kaufmann, (1999)

6.Mobasher B. what's more, et al., Discovery and Evaluation of Aggregate Usage Profiles for Web Personalization. Data Mining and Knowledge Discovery (2002)

7.Mobasher B., Web Usage Mining and Personalization, in Practical Handbook of Internet Computing, CRC Press, 15, 1-37 (2004)

8.Ellen, Spertus and ParaSite, Mining Structural Information on the Web, In proceedings of sixth International WWW Conference, April, (1997)

9.Wee-Keong Ng, Ee-Peng Lim, Chee-Thong Huang, Sourav Bhowmick, Fengqiong Qin. Web Warehousing: An Algebra for Web Information. In Proceedings of the IEEE Advances in Digital Libraries Conference, Santa Barbara, U.S.A., April, (1998)

10.Wang K. what's more, Liu H., Schema Discovery for Semistructured Data. In Proceedings of International Conference on Knowledge Discovery and Data Mining, Newport Beach, AAAI, Aug. (1997)

11.Nestorov S., Abiteboul S. what's more, Motwani R., Inferring Structure in Semi structured Data, In Proceedings of International Workshop on Management of Semistructured Data, (1997)

12.

Zhou Y., Jin X. what's more, Mobasher B., A Recommendation Model Based on Latent Principal Factors in Web Navigation Data, In Proceedings of the third International Workshop on Web Dynamics, ACM Press, (2004)

13.Bhowmick Sourav S., Madria S.K., Ng W.K. what's more, Lim E.P., Web Bags, Are They Useful in Web warehouse? In proceedings for fifth International Conference on Foundation of Data Organization, Japan, Nov. (1998)

14.O'Malley M.J. what's more, et al., Fuzzy Clustering of Children with Cerebral Palsy Based on Temporal-Distance Gait Parameters, IEEE Trans, ON Rehab.

Eng., 5(4), 300-309 (1997)

15.Han J. what's more, Kamber M., Data Mining: Concepts and Techniques (2007)

16.Bhowmick Sourav S., Ng W.K. what's more, Lim E.P., Information Coupling in Web Databases, In Proceedings of the seventeenth International Conference on Conceptual Modeling (ER'98), Singapore, November 16-19, (1998)

17.World Wide Web Consortium. Document Object Model (DOM) Level 1 Specification. http://www.w3.org/TR/, (1998)

18.Wang K. what's more, Liu H., Discovering Typical Structures of Documents: A Road Map Approach, ACM SIGR, August, (1998)

19.Florescu D., Levy A. what's more, Mendelzon A., Database Techniques for the World Wide Web, A Survey, SIGMOD Record (1998)

20.Inman V.T., Ralston H.J. what's more, Todd F., Human Walking, Baltimore (1981)

21.H. Vernon Leighton and Srivastava J., Precision among WWW Search Services (Search Engines): Alta Vista, Excite, Hotbot, Infoseek, Lycos (1997)

22.Han J. what's more, Fu Y., Discovery of Multi-level Association Rules. In Proceedings of International Conference on Very Large Databases, pages 420-431, Zurich, Switzerland, Sept, (1995)

23.Backman D. what's more, Rubbin J., Web log examination: Finding a Recipe for Success, (1997)

24.Pitkow J., In Search of Reliable Usage Data on the WWW, In Proceedings of the sixth International World Wide Web Conference, Santa Clara, California, April, (1997)

25.Baeza-Yates R. what's more, Ribeiro-Neto B., Modern Information Retrieval, Addison Wesley, ACM Press (1999)

26.O'Conner M. what's more, Herlocker J., Clustering Items for Collaborative Filtering. In Proceedings of the ACM SIGIR Workshop on Recommender Systems, Berkeley, CA, USA: ACM Press (1999)

27.Madria Anjay, Bhowmick Sourav S., Ng W.K. what's more, Lim E.P., Center for Advanced Information Systems, School of Applied Science, Nanyang Technological University, Singapore 639798 {askumar, p517026, awkng, aseplim}@ntu.edu.sg

28. Petrovskiy Ikhail, Faculty of Computational Mathematics and Cybernetics, Moscow State University Vorobjevy Gory, Moscow, Russia michael@cs.msu.su

29.By Juan D. Velásquez, PhD University of Tokyo, Assistant Professor, Department of Industrial Engineering University of Chile jvelasqu@dii.uchile.cl http://wi.dii.uchile.cl/

30. Heikki Mannila Nokia Research Center, P.O. Box 407 (Itamerenkatu 11) FIN-00045 Nokia Group, Finland Heikki.Mannila@nokia.com

31. B. Srinivas, Shoban Babu Sriramoju, "A Secured Image Transmission Technique Using Transformation Reversal" in "International Journal of Scientific Research in Science and Technology", Volume-4, Issue-2, February-2018, 1388-1396 [Print ISSN: 2395-6011 | Online ISSN: 2395-602X]

32. B. Srinivas, Gadde Ramesh, Shoban Babu Sriramoju, "A Study on Mining Top Utility Itemsets In A Single Phase" in "International Journal for Science and Advance Research in Technology (IJSART)", Volume-4, Issue-2, February-2018, 1692-1697, [ISSN(ONLINE): 2395-1052]

33. B. Srinivas, Gadde Ramesh, Shoban Babu Sriramoju, "An Overview of Classification Rule and Association Rule Mining" in "International Journal of Scientific Research in Computer Science, Engineering and Information Technology", Volume-3, Issue-1, February-2018, 643-650, [ISSN : 2456-3307]

34. B. Srinivas, Shoban Babu Sriramoju, "Managing Big Data Wiki Pages by Efficient Algorithms Implementing In Python" in "International Journal for Research in Applied Science & Engineering Technology (IJRASET)", Volume-6, Issue-II, February-2018, 2493-2500, [ISSN : 2321-9653]

35. Shoban Babu Sriramoju, "Analysis and Comparison of Anonymous Techniques for Privacy Preserving in Big Data" in "International Journal of Advanced Research in Computer and Communication Engineering", Vol 6, Issue 12, December 2017, DOI 10.17148/IJARCCE.2017.61212 [ISSN(online) : 2278-1021, ISSN(print) : 2319-5940]

36. Shoban Babu Sriramoju, " Review on Big Data and Mining Algorithm" in "International Journal for Research in Applied Science and Engineer

37. Dr. Shoban Babu Sriramoju, Prof. Mangesh Ingle, Prof. Ashish Mahalle "Trust and Iterative Filtering Approaches for Secure Data Collection in Wireless Sensor Networks" in "International Journal of Research in Science and Engineering" Vol 3, Issue 4, July-August 2017 [ISSN : 2394-8299].

38. Dr. Shoban Babu, Prof. Mangesh Ingle, Prof. Ashish Mahalle, "HLA Based solution for Packet Loss Detection in Mobile Ad Hoc Networks" in "International Journal of Research in Science and Engineering" Vol 3, Issue 4,July-August 2017 [ISSN : 2394-8299].

39. Shoban Babu Sriramoju, "A Framework for Keyword Based Query and Response System for Web Based Expert Search" in "International Journal of Science and Research" Index Copernicus Value(2015):78.96 [ISSN : 2319-7064].

40. Sriramoju Ajay Babu, Dr. S. Shoban Babu, "Improving Quality of Content Based Image Retrieval with Graph Based Ranking" in "International Journal of Research and Applications" Vol 1, Issue 1,Jan-Mar 2014 [ISSN : 2349-0020].

41. Dr. Shoban Babu Sriramoju, Ramesh Gadde, "A Ranking Model Framework for Multiple Vertical Search Domains" in "International Journal of Research and Applications" Vol 1, Issue 1,Jan-Mar 2014 [ISSN : 2349-0020].

42. Mounika Reddy, Avula Deepak, Ekkati Kalyani Dharavath, Kranthi Gande, Shoban Sriramoju, "Risk-Aware Response Answer for Mitigating Painter Routing Attacks" in "International Journal of Information Technology and Management" Vol VI, Issue I, Feb 2014 [ISSN : 2249-4510]

43. *Mounica Doosetty, Keerthi Kodakandla, Ashok R, Shoban Babu Sriramoju, "Extensive Secure Cloud Storage System Supporting Privacy-Preserving Public Auditing" in "International Journal of Information Technology and Management" Vol VI, Issue I, Feb 2012 [ISSN : 2249-4510]*

44. *Shoban Babu Sriramoju, "An Application for Annotating Web Search Results" in "International Journal of Innovative Research in Computer and Communication Engineering" Vol 2,Issue 3,March 2014*
[ISSN(online) : 2320-9801, ISSN(print) : 2320-9798]

45. *Shoban Babu Sriramoju, "Multi View Point Measure for Achieving Highest Intra-Cluster Similarity" in "International Journal of Innovative Research in Computer and Communication Engineering" Vol 2,Issue 3,March 2014 [ISSN(online) : 2320-9801, ISSN(print) : 2320-9798]*

46. *Shoban Babu Sriramoju, Madan Kumar Chandran, "UP-Growth Algorithms for Knowledge Discovery from Transactional Databases" in "International Journal of Advanced Research in Computer Science and Software Engineering", Vol 4, Issue 2, February 2014 [ISSN : 2277 128X]*

47. *Shoban Babu Sriramoju, Azmera Chandu Naik, N.Samba Siva Rao, "Predicting The Misusability Of Data From Malicious Insiders" in "International Journal of Computer Engineering and Applications"*
Vol V,Issue II,Febrauary 2014 [ISSN : 2321-3469]

48. *Ajay Babu Sriramoju, Dr. S. Shoban Babu, "Analysis on Image Compression Using Bit-Plane Separation Method" in "International Journal of Information Technology and Management", Vol VII, Issue X, November 2014 [ISSN : 2249-4510]*

49. *Shoban Babu Sriramoju, "Mining Big Sources Using Efficient Data Mining Algorithms" in "International Journal of Innovative Research in Computer and Communication Engineering"*
Vol 2, Issue 1,January 2014 [ISSN(online) : 2320-9801, ISSN(print) : 2320-9798]