International Journal of Engineering Applied Sciences and Technology, 2016
Vol. 1, Issue 2, ISSN 2455-2143, Pages 35-42
Published Online December – January-2016 in IJEAST (http://www.ijeast.com)

SURVEY ON WEB MINING AND ITS USAGE IN E COMMERCE SITES - PATTERN DISCOVERY, ISSUES AND APPLICATIONS

Zulqurnan Aslam
Department of Computer Science & IT
University of Lahore, Pakistan
zulqurnan.aslam@gmail.com

Rohail Shezad
Department of Computer Science & IT
University of Lahore, Pakistan
rohail.35bce@gmail.com

Abstract— **The World Wide Web is the key source of data. It is the hub of data and it is growing day by day and every company is now having its own website. Online transactions trend is also growing very fast. So World Wide Web is the best field for data mining. E-commerce has provided a very effective and cost efficient way of doing business. We can use web mining techniques to extract useful information from the web sites and also we can discover interesting user behavior patterns from the Ecommerce websites. This paper gives the general over view of the web mining and its types and then it discuss the different techniques for pattern discovery. Then this paper gives the general overview of the very famous algorithm Link Analysis Algorithm for Web Mining. And at the end this paper discusses some of the issues of web mining in e-commerce and application of web mining in e-commerce.**

Keywords— ***Web Mining, Data Mining, Ecommerce***

I. INTRODUCTION

E-Commerce Websites are always great advantage for your business. In ecommerce website we can sell, buy different kind of things without any tension of physical visit. There are many types of ecommerce solutions available for your business that can provide different advantages. The main advantages of ecommerce sites are that you can open online store with ease without any physical overhead cost of store location. It is accessible all over the world all the time. Extracting useful information from the World Wide Web (WWW) is called Web mining. Web mining can be divided into three categories.

Web Content mining: In web content mining we get knowledge from web pages, means from the content of web sites i.e. extracting useful data from the topic of different sites. This is an automatic search and retrieval of information from thousands of websites. The second category is **Web Structure Mining:** In this data is extracted from the hyperlinks and show the kind of tree that how pages are connected with each other. The third type is **Web Log Mining:** In this we get information from the logs of servers. It helps to find out different patterns and behaviors and to classify them into groups. With the advancement in web development, ecommerce sites have changed the business in cost efficient and effective way. But still unfortunately many companies think web is just a source of information where we can do transactions. All these kind of websites have high traffic but the income is very low, why? Because they are not using web mining techniques to get more output. So once you make website then you have to implement the web mining techniques to get benefits because data is useless if it does not serve any useful purpose. So by using web mining techniques we can observe the user behaviors and after seeing them we can give them customized web page look by featuring different new products of his/her interest. This will catch the interest of user when he/she will see his/her products of interests. So by doing this we can change our website also, like we can push the products back which are not much trending and can pull the information up front which is most common now a days. Web mining is very interesting. It can automatically extract information from the websites like crawlers, bot etc. Web mining gets useful information from pages hyperlinks and contents. It helps e-commerce to understand how it can be improved for some special group of customers so that average order size can be increased.

The pattern discovery techniques include algorithms to find interesting and useful patterns from web data. Some of them are association rules, clustering, sequential patterns, classification etc. Pattern analysis techniques are used to highlight overall patterns in data and to filter out uninteresting patterns. [5] [1].Pattern discovery techniques can be used to track the customers, what path they had chosen to reach particular page and what was the most visited product page. Further association rules can be used to find the similarity between the web pages.

II. WEB MINING OVERVIEW

International Journal of Engineering Applied Sciences and Technology, 2016
Vol. 1, Issue 2, ISSN 2455-2143, Pages 35-42
Published Online December – January-2016 in IJEAST (http://www.ijeast.com)

Extracting useful information that is previously unknown from the web data is called web mining. Extracting the useful information automatically from the web site pages and documents. A lot of research has been done on this and still there are huge number of researchers are doing research on this to find out new different methods and techniques to get

efficient results. Similar to Etzioni , we can decompose Web mining into these subtasks [6], namely:

- Resource finding: the task of retrieving intended Web documents.
- Information selection and pre-processing: automatically selecting and pre-processing specific information from retrieved Web resources.
- Generalization: automatically discovers general patterns at individual Web sites as well as across multiple sites.
- Analysis: validation and/or interpretation of the mined patterns.

There are different techniques of doing web mining; I will discuss them in details latter like e.g. web content mining, web log mining etc. Web logs are always very sensitive because they contain much critical information, so accessing web logs is always very challenging task.

The dynamic nature of the Web and its increasing importance

as an economic platform creates the need of new methods and tools for business efficiency. Current Web analytic tools do not provide the necessary information of customer processes and critical paths of site visitor behavior. Such information can be used for businesses to react effectively and efficiently.

1) Web Mining and Information Retrieval (IR): In web mining we automate the information retrieval. This technique comes under the category of Web Content mining. I will explain Web Content mining in details.
Information Retrieval is the automatic retrieval of all relevant documents and some of the irrelevant documents at the same time and we try to get as least as possible this irrelevant information. Information Retrieval includes modeling, document classification and categorization, user interfaces, data visualization, filtering, etc. The task that can be considered to be an instance of Web mining is Web document classification or categorization, which could be used for indexing.

2) Web Mining and Information Extraction (IE): Automatically extracting useful structured information from unstructured or semi structured information into machine readable documents. We normally process human language texts by means of Natural Language Processing. The main difference between IR and IE is, in IR we select relevant documents where as in IE we are interested in the structure or representation of document. So we can say IE works in more depth than IR. There are two types of IE. IE from structured data and IE from semi structured data. There is very big difference from the unstructured data and results from structured or semi structured.

Web Mining Categories
Actually Web Mining can be classified into three main parts.
1) Web Content Mining
2) Web Structure Mining
3) Web Usage Mining

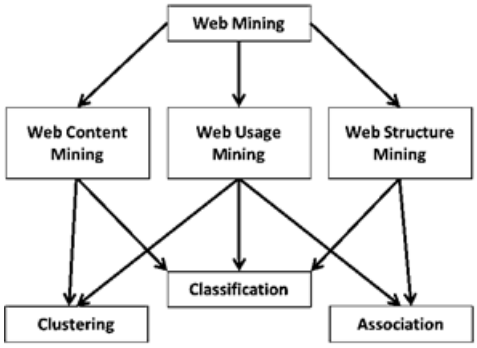

Fig. 1. Web Mining Over View

1) Web Content Mining: Mining of the web content like images, text Etc. is called web content mining. All the text, image, knowledge, multimedia in the website is its content. So when we play with this content then it is called Web Content mining. There are different tools available which help us to finding this kind of information. Extracting useful information from the documents of World Wide Web is called Web Content Mining. HTML documents contain different kind of text like text, images, multimedia, links etc. Web Content mining becomes very interesting when we extract information because HTML documents can be organized or unorganized, so some time it is very interesting and challenging task to gather information from unorganized data. Web Content Mining examines the search results of search engines. It is very hard to do things manually specially where there is huge amount of work. Same is the case here; if we extract information manually then it will be very time consuming especially where there is huge amount of data. So now technology is playing very important role in almost every field of life so same is the case here on internet. Web mining became the main boon of this magic. See Appendix A.

2) Web Structure Mining: In Web structure mining, mining is done based on the structure like hyperlinks. Web structure mining is the process of using graph theory to analyze the node and connection structure of a web site. According to the

International Journal of Engineering Applied Sciences and Technology, 2016
Vol. 1, Issue 2, ISSN 2455-2143, Pages 35-42
Published Online December – January-2016 in IJEAST (http://www.ijeast.com)

type of web structural data, web structure mining can be divided into two kinds:

1) Extracting patterns from hyperlinks in the web: a hyperlink is a structural component that connects the web page to a different location.

2) Mining the document structure: analysis of the tree-like structure of page structures to describe HTML or XML tag usage. The goal of Web Structure mining is to generate the structural summary of the website and webpages. Web Content mining focuses on the structure of inner documents whereas the web structure mining try to find the link structure of the website like using the hyperlinks, navigations etc. This technique tries to find the relationship and similarity between the different pages and websites using these navigation and links. Web Structure mining can also be used to reveal the structure of the web pages and in this way it can make this possible so that we can compare the web pages schemas. Web Structure mining has a natural relationship with Web Content mining because every page or document contain the links, it can contain links to other pages etc. So it is quite often to combine these two techniques in an application.

3) Web Usage Mining: Mining is done on web logs which contain the navigational pattern of users. And the study of this navigational pattern will trace out the interest of the users. We analyze the logs from server from which we can see how often traffic is coming and from which country we have heavy flow of traffic. Then which user is interested in which item if our system is like amazon. Web usage mining is used to discover the interesting user navigational patterns and we can apply these techniques to many real world problems. By studying the user behaviors we can do some useful things to the websites like we can improve our website, making additional topics or product recommendation. There are some basic requirements for web usage mining. It is necessary to have knowledge of what kind of features a Web Usage Mining system is expected to have in order to make the efficient use of Web Mining [2].

4) Requirements of Web Usage Mining [2]:

- Get useful data.
- Filter out the unnecessary data.
- Discovery of interesting navigational patterns.
- Display the navigational patterns.
- Analyze those navigational patterns.
- Extract the useful results from these patterns.

5) Web Usage Mining Systems Structure:

Web Usage Mining System carries out five major tasks:

- Gathering Web usage data: Web logs are the most comprehensive and detailed web usage data. These web logs contain the user activities on websites.
- Data Preparation: There is lot of noise in the logs. We have to eliminate that noise to get some useful information. This can be done by different mining algorithms. This task restores the user's activities that are recorded in the web server logs.
- Navigational Patterns Discovery: In this part we try to find some interesting and useful user navigational patterns from the logs. We use different mining algorithms, which use the method of sequential pattern generation. We apply Queries to find out different results.
- Pattern analysis and visualization: After finding navigational patterns these are still need to be processed before we get some useful results.
- Pattern applications: The navigational patterns can be useful in many ways. Like it can be used to improve the web page, making some recommendation for the users, Web personalization and learning the user behavior. As show in the Fig: 2.

III. PATTERN DISCOVERY TECHNIQUES FOR E-COMMERCE SITES

In pattern discovery technique we try to find out different interesting patterns from the web data. For this, there are different kind of algorithms are available. First of all we identified the user sessions and after identifying the user session, there are several kinds of discovering techniques that can be performed depending on the needs of analyst and some of them are written below.

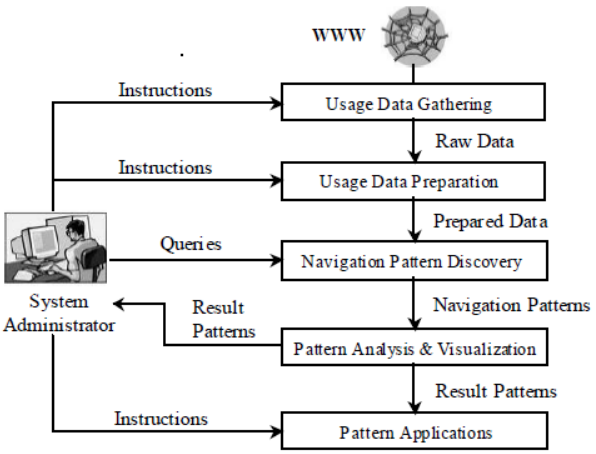

Fig. 2. Web Usage Mining System Structure

1) Path Analysis: In this we try to find out the path of the users while visiting the website. Like we find out the path of the users they followed to reach out a particular page. What pages they had visited the most frequent. Normally Graph models are used for this Path Analysis. In Graph model, graph represents relationship between websites and its pages and the link of the documents. The tree represents the whole website.

International Journal of Engineering Applied Sciences and Technology, 2016
Vol. 1, Issue 2, ISSN 2455-2143, Pages 35-42
Published Online December – January-2016 in IJEAST (http://www.ijeast.com)

Each node in a tree represents a web page and the edges between the trees represent the links between websites and the edge between nodes inside the same tree represent links between documents of the websites. These navigational patterns can give us very valuable information about the website and can help us to improve the website. For Example: what are the paths user's traverses before going to a particular URL? Some interesting example are given below

- 55% of the users who accessed the ORDER.JSP page they started the website by going to the NEWOFFER.JSP page from the HOME.JSP.
- 65% of the users left the website after visiting main page. So first observation is, 55% of the users placed the order by going to the NEWOFFER page.

So people are more interested in new exciting offers. So this page of our website should be more attracting so that we can get more users in the websites who can post the orders and it should be easily accessible. About the second observation, it tells us that mostly people are coming to our main page and see it and then leave the website. So our main page should contain all the important information so that when user comes, he can take a look of the whole website.

2) Association Rules: Association Rules are the one of the most widely used mining techniques. They reflect regularities in the co-occurrence of the same items within a set of transactions. One very typical example is what the set are of products usually purchased by the independent buyers. So E.g. A!B and B!C so A!C. It tells about what are the set of web pages often referenced together and what will be fetched next? By reading this we can observe the users buying habits.

- 45% of the clients who accessed the web page URL /COMPANY/HOME/PRODUCTS/BREAD.HTML, also accessed /COMPANY/HOME/PRODUCTS/MILK.HTML

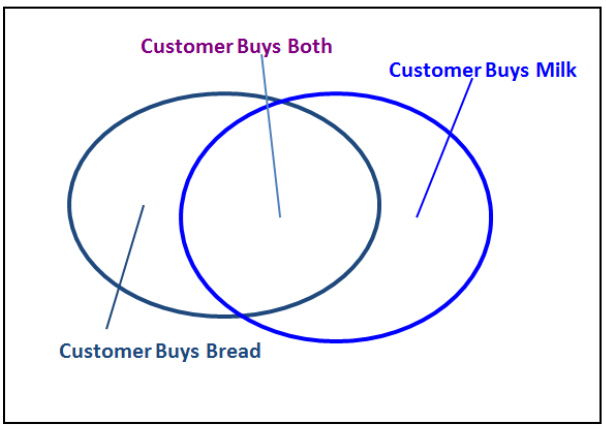

Fig. 3. Association Rule Example

3) Decision Trees: Decision trees are like flow charts of questions that ultimately lead to a decision. A decision tree is a structure contains different kind of nodes, connectors and leaf nodes. Internal nodes represent test attributes and connectors tell the outcome of the test and leaf node tells the class label. The above decision tree tells the customer behavior for buying computer or something. So assume that it tells about the computer buying behavior. So it is saying that if he is young and he is student then he buys computer. Similarly

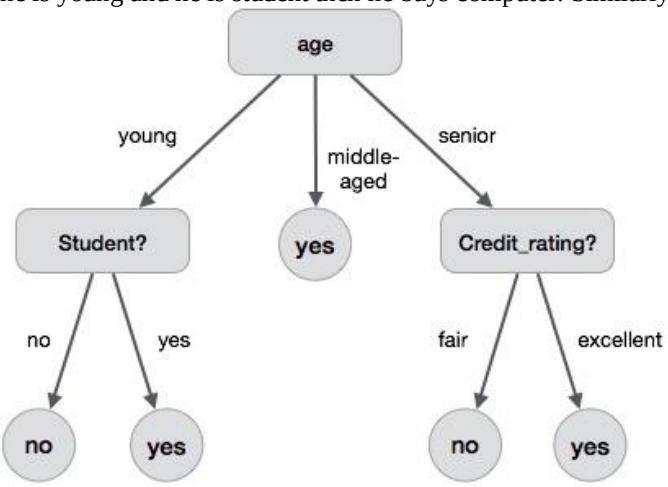

Fig. 4. Decision Tree [www.tutorialspoint.com]

if he is of middle age the he also buy computer. And if he is senior then if its credit rating is good then he buy computer. So we can see that decisions trees are very easy to find out some ultimate decision. There are many algorithms available to sort out these things. A machine researcher named J. Ross Quinlan in 1980 developed a decision tree algorithm known as ID3(Iterative Dichotomiser). Later, he presented C4.5, which was the successor of ID3. ID3 and C4.5 adopt a greedy approach.

4) Clustering: Clustering is the process where we separate the objects of similar or same classes. Similar objects are grouped in one cluster and other in other cluster. It identifies visitors who share common characteristics. Other than the information from the web logs, customer profiles are also very important. For example we often fill the online surveys in which they ask our email address, age, gender hobbies etc. So that information will be stored in the company's customer profile database and then that information can be used for future data mining purpose. For example

- 45% of the clients who applied for VISA credit card in domain.com/visa/newcarreq, were in the 20-30 age group and their annual income were in between 30; 00040,000

5) Grouping: Grouping same kind of objects and information and then drawing the high level conclusion. For example grouping all Google Chrome browser together and all the Mozilla Firefox browser and then show which browser is more

International Journal of Engineering Applied Sciences and Technology, 2016
Vol. 1, Issue 2, ISSN 2455-2143, Pages 35-42
Published Online December – January-2016 in IJEAST (http://www.ijeast.com)

popular among the users, regardless the other minor version of the browsers. Similarly grouping all the users who came from the Google search and those who came from Bing search.

6) Filtering: Filtering useful information from the complex and detailed information is very challenging task. Since there is lot of noise which can lead us to different kind of results, so we have to filter out the relevant data and discard the irrelevant data. Simple reporting require only very simple

analysis but when company's website becomes more integrated with other functionalities like Customer Service, Human Resource, Marketing Etc. then if company launches some kind of marketing campaign then Print and Television ads are more important than calling 100 phone numbers individually. So now online marketing is no longer a small minor issue. [8] Database filtering gives you the flexibility when you design your mining structure and data sources because you can create a single mining structure based on comprehensive data source view. For example, you define the data source view on the Customers table and related tables. Next, you define a single mining structure that includes all the fields you need. Finally, you create a model that is filtered on a particular customer attribute, such as Region. For example

- SELECT MODEL_NAME, [FILTER] FROM Table

- SELECT * FROM Table WHERE (Age > 30 AND EXISTS (SELECT * FROM Products WHERE ProductName=Milk AND Quantity>2))

7) Cookies: Cookies are other very useful method that can be used for pattern analysis. Cookies are the random IDs which are assigned by the web server when customer visits the website first time. Now when same customer visits the same website again then server sends the same old ID and browser recognizes it. We often use this feature in online shopping cart. Cookies also benefit visitors by allowing Web sites to recognize repeat visits. For example, Amazon.com uses cookies to enable their "oneclick" ordering system. Since Amazon already has your mailing address and credit card on file, so you don't need to re-enter this information, making the transaction faster and easier.

IV. LINK ANALYSIS ALGORITHM FOR WEB MINING

Website can be a link of connected pages. We can move from one page to other through hyperlinks. So we can view the web as a directed labeled graph whose nodes are the pages and edges are the hyperlinks. And this is called directed graph. So there are lots of algorithms based on link analysis and I will discuss only one PAGE RANK which is very important and is base of the other algorithms. One of the very popular search engine Google is using this algorithm.

PageRank Algorithm

[7]PageRank is one of the algorithms Google uses to find out the importance a page or website. If we add the Google toolbar then on the top of the browser, page rank is calculated by Google. Its range is from 0-10. So if we see it in logarithmic scale:

Toolbar Page Rank	Real Page Rank
0	0 —10
1	10 —1,000
2	1,000 —10,000
3	10,000 —100,000
4	And so on

Definition: What Google Says

We assume page A has pages T1...Tn which point to it (i.e., are citations). The parameter d is a damping factor which can be set between 0 and 1. We usually set d to 0.85. There are more details about d in the next section.

Also C (A) is defined as the number of links going out of page A. The PageRank of a page A is given as follows:

PR(A) = (1-d) + d (PR(T1)/C(T1) + ... + PR(Tn)/C(Tn))

Note that the Page Ranks form a probability distribution over web pages, so the sum of all web pages' PageRanks will be one. PageRank or PR (A) can be calculated using a simple iterative algorithm, and corresponds to the principal eigenvector of the normalized link matrix of the web.

1) (Tn) - Each page has a notion of its own self importance. That's PR(T1) for the first page in the web all the way up to PR(Tn) for the last page

2) C(Tn) - Each page spreads its vote out evenly amongst all of its outgoing links. The count, or number, of outgoing links for page 1 is C (T1), C(Tn) for page n, and so on for all pages.

3) PR(Tn)/C(Tn) - so if our page (page A) has a backlink from page n the share of the vote page A will get is PR(Tn)/C(Tn)

4) d(... - All these fractions of votes are added together but, to stop the other pages having too much influence, this total vote is damped down by multiplying it by 0.85 (the factor d)

5) (1 - d) - The (1 d) bit at the beginning is a bit of probability math magic so the sum of all web pages' PageRanks will be one: it adds in the bit lost by the d(.... It also means that if a page has no links to it (no backlinks) even then it will still get a small PR of 0.15 (i.e. 1 0.85).

International Journal of Engineering Applied Sciences and Technology, 2016
Vol. 1, Issue 2, ISSN 2455-2143, Pages 35-42
Published Online December – January-2016 in IJEAST (http://www.ijeast.com)

(Aside: the Google paper says the sum of all pages but they mean the normalized sum otherwise known as the average to you and me.

Now let's take a simple example to understand this magical code. It seems very tough in definition but it is very simple when we will discuss its example. Assume a very simple web page with two pages. Page A has one out going link to Page B and similarly Page B has one out going link back to A.

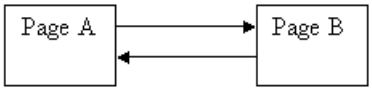

Fig. 5. Page Rank Example

So

C(A) = 1 and C(B) = 1 , d = 0.85
PR(A) = (1 d) + d(PR(B)/1)
PR(B) = (1 d) + d(PR(A)/1)
PR(A) = 0.15 + 0.85 * 1 = 1
PR(B) = 0.15 + 0.85 * 1 = 1

So this value is one which is ideal. So we will change the values and find out the resulting values. We will repeat this process until we start getting same values. Now lets change the value again and see the results.

PR(A) = 0.15 + 0.85 * 0.1 = 0.235
PR(B) = 0.15 + 0.85 * 0.235 = 0.34975

Now change the value again and we get

PR(A) = 0.15 + 0.85 * 0.34957 = 0.4471345
PR(B) = 0.15 + 0.85 * 0.4471345 = 0.530064325

Again

PR(A) = 0.15 + 0.85 * 0.530064325 = 0.60055467
PR(B) = 0.15 + 0.85 * 0.60055467 = 0.6604714

So by repeating this process we are getting closer towards 1. So while repeating this process we will stop till we get very minor difference. That will be the page rank of the page.

Findings

If a website having PageRank of more than 3 then it is said to be good website. But if a website having PageRank of more than 5 than the website is getting great traffic and the overall performance or the structure of website is good enough.

V. ISSUES OF WEB MINING IN E-COMMERCE SITES

Websites are the great field for data mining. It contains huge amount of data where we can apply different data mining techniques to get useful information. And especially E-Commerce site is the very perfect domain for data mining. But still there are some issues that we can discuss here.

1 - Web server logs are very rough to read and understand. Web server logs are plain text files and a normal person will get no useful information from it.

2 - Web server logs don't store the user session information. HTTP is a stateless protocol, and our web server logs don't contain information relating to the user sessions. User sessions are the key in data mining.

3 - Web Server logs contain some redundant information. Web server logs contain many noisy data. Some information is repeating that can lead to different results. So for example a web page contains many images, and log files contain request for each image. So if we open same page multiple times then it will contain redundant data.

4 - Web Server Logs lack information of Dynamic websites. Modern dynamic websites contain smaller generic URLs which make it very harder to get the useful information. Like E.g. www.domain.com/product.jsp?id=1234 , now same page will call for each new product only id will be changing.

5 - Mining data at the right level of granularity is essential else results will be different.

6 - Web Server logs are for debugging the server. Web Server Logs are mainly for the debugging the web server application. So they contain very minor details of useful information and they have no information regarding the customer related transactions.

7 - User Interface is very important in websites. User interface plays a very important role so data mining issues should be kept in mind while designing user interfaces. Like E.g. we should avoid giving default values to the attributes like Gender, Age and Marital Status etc. So each time if user will enter original data, that would give us more chance to get the real data for our analysis.

8 - Generating logs for several million transactions is a very costly exercise. It would be wise if we generate appropriate logs by conducting random sampling.

9 - In several data mining algorithms they assume that if user remain inactive for certain time then its mean he has left the site, so for this, session time out technique is implemented on the Ecommerce sites. So due to this sometime many important clients or customers loses their cart.

10 - Web server logs don't store the web form information. Web server logs do not contain the web form information, like what they have entered. E.g. what they have entered for search

International Journal of Engineering Applied Sciences and Technology, 2016
Vol. 1, Issue 2, ISSN 2455-2143, Pages 35-42
Published Online December – January-2016 in IJEAST (http://www.ijeast.com)

of a particular item. So these information needs to be logged for our analysis.

VI. APPLICATION OF WEB MINING IN E-COMMERCE SITE

1 - Web Mining can be very help full in improving the web site design. [4] Web Mining can be very helpful in updating the website and improving the design of websites. E.g. we can see the behavior of users on the website and then can modify the website. We can put the main deals of the products or important products on the main page.

2 - Targeting potential customers. We can target special age group customers. E.g. if our website is visiting by age group of like 20 30 then we can get the analysis.
3 - Can be very useful to make website personalized.
4 - Can be used for enhancing the web server performance.
5 - Fraud detection.
6 - Predicting user actions.
7 - Can attract more customers.

VII. FUTURE DIRECTION

[3]Websites are growing day by day and this trend is likely to continue as Web services continue to flourish. As the Web and its usage grow, it will continue to generate evermore content, structure, and usage data, and the value of Web mining will keep increasing. In future Web Mining will give direction to some of these new fields like Web metrics and measurements, Process mining, temporal evolution of the Web, Web services optimization and Fraud Detection.

VIII. CONCLUSION

The World Wide Web is growing very fast and now everyone is making sure to make its electronic existence. By the rapid growth in WWW now every business function executed fast and efficiently. E-commerce sites reduced the physical distance gap and now we can do online payment easily without bothering our location. So WWW is a huge hub of data, and if we can't make a good use of data then it is useless. So data mining gives us power to play this data in a meaningful manner. Using Web Mining, we can extract different interesting patterns of users and similarly we can observe the behavior of the customers and by observing these behaviors, companies can improve their websites, so that more and more traffic of user can be obtained. So to make all this possible we have to overcome different challenges that we have to face in Web Mining, like noisy data, to make the site more user friendly, effective and easy to search. In this paper we have tried to discuss the web mining in Ecommerce sites in detail, their types, their challenges and issues in web mining applications and pattern discover techniques. We also discussed the Link Analysis Algorithm for Web Mining and little over view of Page Rank algorithm.

APPENDIX A

Google Trends [www.google.com/trends] Google Trends is the great website service from google. You can see the overall trends of the word currently running. So we just search the word Game of Thrones. It is an American TV Series which is getting attention day by day all over the world. So here are the screen shots of the results we got from the google trends in fig: 6.

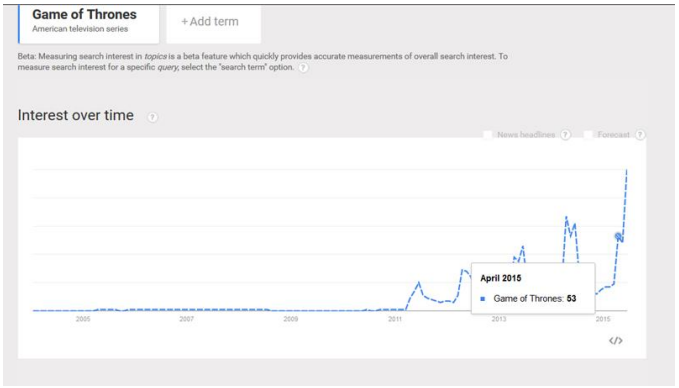

Fig. 6. Google Trends for Game of Thrones

So in before 2011 its trend was 0 because series started in 2011. So then it gets some attention. And you can see in April it was 53% and in May and June 2015 its graph goes very high. Because it's new season came out in 2015 and that season is getting more fame, as you can see in the graph. Now you can see the trends Region wise, countries wise.

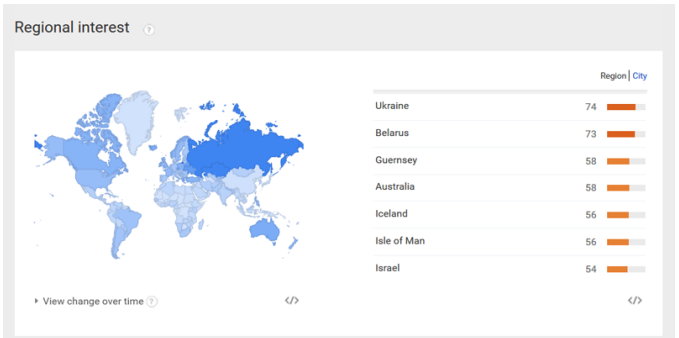

Fig. 7. Google Trends for Game of Thrones Region Wise

Now see the related searches. Many people search this topic with intension of Game. Like you can see in the figure its trend is 100%. The people searched its latest season 4 which came out in 2015. And then people searched out for its sub titles.

International Journal of Engineering Applied Sciences and Technology, 2016
Vol. 1, Issue 2, ISSN 2455-2143, Pages 35-42
Published Online December – January-2016 in IJEAST (http://www.ijeast.com)

As you can see you can use this application of google to find out the interesting findings. Actually this is the good example of Web mining applications. Google used these techniques to filter out the results. They used the mining techniques we had discussed above to make this application. Similarly there are lot of other software's who can help us in web mining. Very good software we often use is called WebLog Expert. We normally use its lite version because that version is free. This software takes the web log file and give you the very nice results, like who visited the website, how many, what were there IP, Regions and Browsers etc.

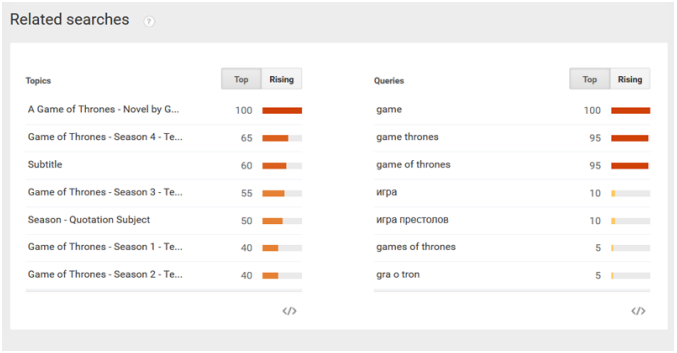

Fig. 8. Google Trends related Search

IX. REFERENCES

[1] Y. Fu and M. Shih. A Framework for Personal Web Usage Mining.

[2] Wen-Chen Hu, Xuli Zong, Chung wei Lee, and Jyh haw Yeh. World Wide Web Usage Mining Systems and Technologies. Grand Forks, ND 58202.

[3] Vipin Kumar Jaideep Srivastava, Prasanna Desikan. Web Mining - Accomplishments and Future Directions. University of Minnesota, Minneapolis, MN 55455, USA.

[4] S.Yadav K.Ahmad and J.Shekar. Analysis of Web Mining Applications and Beneficial Areas. Meerut-250002, India.

[5] Jigar D. Patel Ketul B. Patel, Jignesh A. Chauhan. Web Mining in E-Commerce. Kherva, India.

[6] Raymond Kosala and Hendrik Blockeel. Web Mining Research: A Survey. Heverlee, Belgium.

[7] Ian Rogers IPR Computing Ltd. Filters for mining models (analysis services - data mining). http://www.cs.princeton.edu/_chazelle/courses/BIB/pagerank.htm/.

[8] Microsoft. Filters for mining models (analysis services – data mining). https://msdn.microsoft.com/en-us/library/bb895167.aspx/, 2015. [SQL Server 2016].