



Web Mining for Information Retrieval

Shadab Irfan¹, Subhajit Ghosh²
Research Scholar (SCSE)¹, Professor (SCSE)²
Galgotias University Greater Noida, Uttar Pradesh, India

Abstract:

The present era is engulfed in data and it is quite difficult to churn the emergence of unstructured data in order to mine relevant information. In this paper we present a basic outline of web mining techniques which changed the scenario of today world and help in retrieval of information. Web Mining actually referred as mining of interesting pattern by using set of tools and techniques from the vast pool of web. It focuses on different aspects of web mining referred as Web Content Mining, Web Structure Mining and Web Usage Mining. The overview of these mining techniques which help in retrieving desired information are covered along with the approaches used and algorithm employed.

Keywords: Web Mining; Web Content Mining, Web Structure Mining, Web Usage Mining

I. INTRODUCTION

The present era has changed a lot during the current years and transformed the outlook of the peoples. They view the information from different perspective which helps them to find the valuable information by aggregating various techniques and tools. The World Wide Web has emerged as vast and popular repository of hyperlinked documents containing enormous information which helps in satisfying the needs and requirements of every human being. Web is a repository of data which contain majority of documents in HTML form. The World Wide Web is expanding at a rapid pace and with the expansion of data various mining methods are being used which help in discovering hidden information and relevant patterns.

This hidden information help the user in different set of tasks which span from taking feedback from users, making decisions, customer relationship management, restructuring websites etc. The mining methods are being clubbed with artificial intelligence, natural language processing and other statistical techniques which efficiently improve the mining tasks. Web Mining extracts relevant information from the web pages by using different set of tools and algorithms from semi-structured and unstructured data found in the web pages.

Web Content Mining focuses on the content of the web pages to retrieve the information, similarly web structure mining keep tracks of the links among pages and provide the most relevant pages to the user as per the query need. Web Usage mining take into account web log information about web pages which is helpful in tracking user information.

II. OVERVIEW OF MINING

Mining basically covers a wide area where documents of all types are handled by using statistical techniques and uncover hidden pattern which help in decision making tasks. It uses intelligent methods to discover new information. Basically mining task can be categorized into following types-

A. TEXT MINING

Text mining focuses on unstructured texts. It is a set of processes which employs statistical techniques to monitor unstructured text and thereby convert those documents into valuable structured information. With rise of social networks it has been realized that information is a strategic asset made of text and sentiment analysis is a new focus point in the field of text mining.

B. DATA MINING

Data mining deals primarily with structured data, which is properly organized and various algorithms are used to analyze and extract useful information from the data. It is useful in discovering hidden patterns and helpful in predicting outcomes from large datasets. It focuses on data-dependent activities and defined algorithms can be quickly deployed for providing better solutions.

C. WEB MINING

Web Mining includes discovery and analysis of documents from World Wide Web. It works on semi-structured and unstructured data which is quite difficult to handle. Web mining basically deals with providing solution to different problems and finding relevant information from the World Wide Web by using suitable algorithms. It performs this task by using content, structural links or some log statistics which help in fulfilling the information needs. In spite of using only data mining technique, it also incorporate certain other techniques which help in the retrieval task like natural language processing, machine learning, and artificial intelligence. Most of the web mining technique is associated with information retrieval process which is applied to semi-structured and unstructured data [1]. The query generated by the user can be changed during the course of searching and it is the initial step which thereafter provides relevant ranked items [2]. Web Mining finds its role in respective all the areas ranging from e-commerce to customer relationship management and can be depicted in Fig. 1[3].

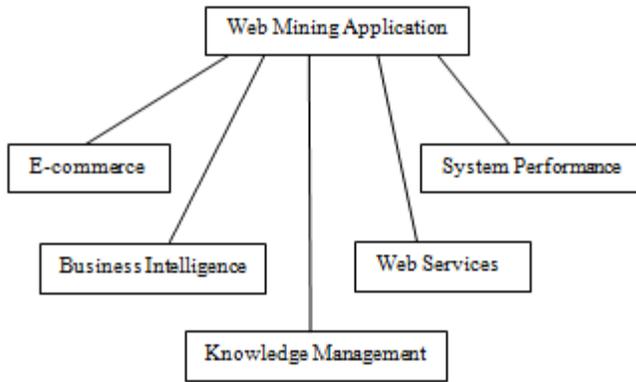


Figure.1. Web Mining Application

D. INFORMATION RETRIEVAL

Information retrieval focuses on finding the relevant information by filtering out the non-relevant from the relevant. The main target is on indexing and searching useful documents from the collection. Information retrieval relies on classification and categorization of data, user interfaces and modeling. Focused crawlers can be used for extracting desired information from the web [4].

E. INFORMATION EXTRACTION

Information extraction considered as extracting specific documents which are retrieved from IR steps. The main task of Information Extraction relies on transforming information so that it can be analyzed properly by extracting relevant facts and figures. It works at a finer level by studying the structure of documents.

III. CLASSIFICATION OF WEB MINING TECHNIQUES

The Web Mining techniques on the basis of extracting information from the web can be categorized into three types and is depicted in Fig .2.

- Web Content Mining (WCM)
- Web Structure Mining (WSM)
- Web Usage Mining (WUM)

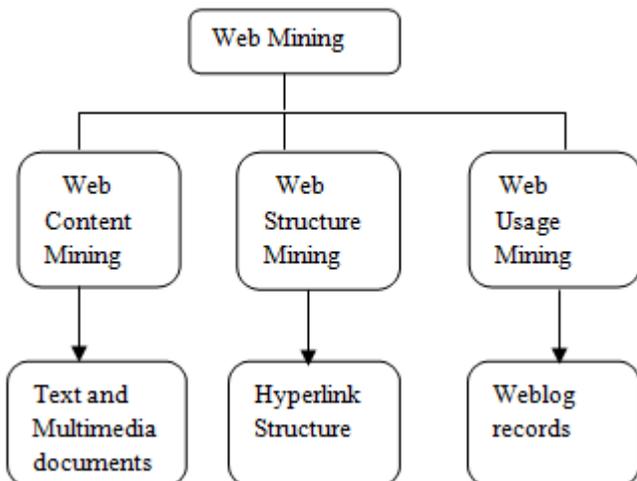


Figure. 2. Web Mining Structure

IV. WEB CONTENT MINING

Web Content Mining as the name implies deal with the content

of the documents. It is the automatic searching of information resources from the web repository. It is differentiated from two points of view-i.e. Agent-based approach which deal with information finding and filtering while Database approach use to model the data on the web in a more structured format by applying querying mechanism. The type of content present in the web can be in different forms like text, audio, video, image.

A. FORMS OF DATA

The information present on the web can be of different types – structured, semi-structured, unstructured and multimedia. Various types of techniques are present which help to handle different forms of data. When the data is unstructured then Information Extraction can be applied which help in converting the unstructured into structured format thereby help in applying pattern matching over it. The Topic Tracking technique helps the user in finding the relevant pages as per the query entered and help to apply certain prediction methods for extracting information. The Summarization helps in minimizing the size of the documents by analyzing the semantics of the statements. Categorization deals with placing the documents in some predefined groups while Clustering helps to group similar items together on the basis of characteristics recognized. For creating a graphical representation on basis of indexing and feature extraction techniques Information Visualization is being employed. Taking the view of structured data, web crawlers are considered which are designed to travel across the hyperlinks in world wide web. The Wrapper Generation technique employs certain rules for extraction useful information from the web pages like meta information by using page ranking. The Page Content Mining similarly uses page ranking for extracting and displaying the content of the web pages. Another technique referred as Object Exchange Model (OEM) which also sometimes used for semi-structured data help to understand the information structure of the web pages. The semi-structured data uses Top Down Extraction technique which helps in transferring complex objects of rich resources into simpler ones. Another technique known as Web Data Extraction Language used for converting the data into structured formed which can then be used by end users. Considering Multimedia data, it uses SKICAT technique which is based on astronomical data analysis system. The Color Histogram matching technique removes unwanted artifacts by finding the correlation among color components employing smoothing methods. The Multimedia Miner used for extracting images and videos for feature extraction. The shot Boundary Detection technique help in finding automatic boundaries among images [5].

B. TOOLS EMPLOYED

Web Content Mining is equipped with different tools that help in extracting and managing web data in proper way. Different tools of Web Content Mining are categorized as-

• Mozenda

This tool help in managing, extracting and reframing the information previously stored in a system. The agents can help in storing, publishing and extracting information. It comprises of Mozenda Web Console which help in viewing and organizing result and exporting data. The Agent Builder help to build data extraction project which is easy to use and platform independence.

- **Web Info Extractor**

This data mining tool can extract tabular structured and unstructured data from the web pages without using complex template and help to reform into local file which can then be stored into web server. It is quite helpful in extracting new content and updating and monitoring web pages in any language. The tool is multitasking and facilitates recursive task definition

- **Screen Scrapper**

Used for searching a database and provide help in extracting information from websites. Various languages like Java, .Net, PHP, Visual Basic and Active Server Pages can be used to access Screen Scrapper. It provide scraping of information in cycle time by using the graphical interface and help in downloading the data to spreadsheet. Meta-search engine is a classical example which helps in executing query entered by the user by running it under multiple websites in real time.

- **Web Content Extractor**

It is a powerful, friendly which provide a wizard driven interface for data extraction. The interface help in forming data extraction pattern and useful for creating crawling rules which crawl different websites like online stores, shopping and business sites and transfer the extracted data to different formats like text, XML, CSV, SQL etc. This tool is quite helpful for job seeking user, auction of products, book information, news articles and extracting other online information about holiday spots.

- **Automation Anywhere**

It also helps in getting web data with its automation technology that automates complex tasks. With the help of wizard facility it record keyboard and mouse strokes to automate any task.

- **Rapid Miner**

It is open source software which helps in extracting information from web by using inbuilt algorithm and can also generate algorithm. It reduces time for accessing information and is easy to use [6, 7, 8].

C. ALGORITHM USED

Different types of algorithms are employed in Web Content Mining which helps in easy retrieval of the web content. Some of the algorithms are-

- **K-nearest neighbor**

This algorithm is based on training by analogy where training tuples are compared with a test tuple. The training tuples are defined by n attributes and tuple signify a point in n dimensional pattern space. Similarly all the training tuples are stored in n dimensional pattern space. The k-nearest-neighbor classifier searches for an unknown tuple in pattern space for the closest tuple. The Nearest neighbor classifiers can be used for prediction also for an unknown tuple by returning a real-valued prediction. It suffers from poor accuracy for noise and irrelevant attributes when used for distance-based comparisons which is done by assigning intrinsically equal weight for each attribute.

- **Decision Tree**

Decision tree is a tree induction classification method in form of flowchart. It is tree structure capable of handling high dimensional data. Each internal node (non leaf node) denotes a

test on an attribute, each branch represents an outcome of the test, and each leaf node (or terminal node) holds a class label. Algorithms used for building decision trees are Classification and Regression Trees (CART), ID3 and C4.5.

- **Naïve Bayes**

To predict class membership probabilities for the statistical classifiers Bayesian classifiers are used which tell that a given tuple belong to a particular class. Based on Bayes' theorem exhibited high accuracy and speed when applied to large databases. The attribute value of a particular class does not depend on values of other attribute in the Naive Bayesian classifiers. This is called class conditional independence and is based upon the principle of Maximum A Posteriori (MAP).

- **SVM**

Support vector machines (SVMs) are regarded as a set of related supervised learning methods which are used for classification which construct a separating hyperplane in the n-dimensional space for viewing input data as two sets of vectors. For calculating the margin two parallel hyperplanes are constructed which are pushed up against the two data sets. The larger the distance, better the generalization error for the classifier and this is achieved by extending the distance for neighboring datapoints that separates the hyperplane.

- **Neural Network**

A neural network is a set of connected input/output units in which each connection has a weight associated with it. A neural network learning algorithm is referred as Backpropagation. For correcting the class label for input tuples, the network learns by adjusting the weights during the learning phase. It is also referred as connectionist learning due to the connections between units. It require a number of parameters and long training times which are determined empirically, such as the network topology or "structure." The Backpropagation algorithm was designed for performing learning which predicts weights for class label and the task is performed iteratively. It consists of an input layer, one or more hidden layers, and an output layer. Each layer is made up of units. The outputs of the hidden layer units can be input to another hidden layer, and so on and the number of hidden layers is arbitrary. For input to units that make up the output layer the weighted output of last hidden layer is subjected as input that emits network prediction for tuple. It is fully connected in that each unit provides input to each unit in the next forward layer [9].

D. METRICS

In order to access the performance of the web content mining certain metrics is being used which help in accessing the execution time, service and ability of the system. Some of them are-

- **Scale up**

It tests the capability of the system in managing performance when more computers are integrated together.

Scale up = $\frac{\text{Throughput After}}{\text{Throughput Before}}$

- **Cost Performance**

It is the ratio of throughput and cost.

Cost Performance= Throughput / cost

- **Concurrency**

It helps in providing service to different users at a particular instance of time.

Concurrency= Successful Operations / Total Operations

- **Durability**

It is related to maintaining the information for a longer period of time.

Durability= Current Reads / Total Reads [7]

E. FILTERING METHODS

Different techniques are used which help in enhancing the process of finding relevant information. They are-

- **Information Filtering/ Categorization**

These techniques employ information retrieval methods for retrieving, filtering and categorizing information in web pages.

- **Intelligent Search agents**

These agents take the advantage of user profile and domain characteristics to interpret the discovered information among the web pages.

- **Personalized Web Agents**

These are related with user preferences and discover web information which is of same interest [6].

F. NLP TECHNIQUES

Some of the Natural Language Processing methods are used which help in handling the web content easily and smoothly. They are-

- **Structured Data Extraction**

The information which is retrieved from the databases is in proper structured format like list of products or services which provide value added services. Two techniques are used Wrapper Induction which is a supervised machine learning technique that is applied to learn extraction rules or patterns. It is useful in extracting target items from pages. Different wrapper induction systems are WIEN, Stalker, BWI, IDE etc. The problem with wrapper is that it is costly when sites change frequently. The other method is Automatic Extraction which is unsupervised technique help in generating extraction patterns from multiple data records. The problem arises when data is extracted from multiple sources.

- **Information Integration**

To design a consistent and reliable database for structured data, integration is performed through Schema Match or through Data Instance Match. The Schema Matching integrates two or more database schemas into a single global schema that are semantically same. In Data Instance Match value characteristics are used for matching which are from different fields.

- **Information Knowledge/ Synthesis**

It is the combining activity of different elements to form a coherent whole. It is being used in web search paradigm to rank

the list of pages according to the query given by the user. The technique is quite efficient for navigational search but not suited for open-ended research, learning and exploration as in case of information search.

- **Opinion Mining**

The search ranking strategy is not appropriate for opinion search. Opinion at review sites can contain valuable information which is quite helpful in business organizations. The documents are classified based on overall sentiments expressed by opinion holders [10].

G. APPLICATION

Web Content Mining found its usage in wide area spanning from grouping to classifying information present on the world wide web. It is quite helpful in online marketing by tracking the behavior of the users which help in market analysis. Customer's feedback and reviews help in improving online shopping sites which can be reframed according to the needs of the user. Online content from social sites can be clustered together which help in decision making tasks and personalized data can be maintained. Digital libraries help to maintain automated indexing using web mining techniques. Opining mining and web wide tracking can be performed using different mining techniques [5,11, 12].

V. WEB STRUCTURE MINING

Web Structure Mining deals with the link structure of the web pages. It focuses on the link structure of the hyperlinks and tries to find pages that are source of information. It helps in generating structural summary about web sites and web pages. The main focus in web structure mining is on links [13]. By using the links among the web pages it establish relationship among each other and generate information like similarity between web pages. It is quite useful for information retrieval tasks by performing intra-page and inter-page mining [14]. Web page comprises of nodes and hyperlinks. The links allow to access to the desired information from the web pages and is divided into hyperlink and document structure. The hyperlink helps to connect different location of a web page while document structure comprises of the content within the page which can be structured in tree form based on HTML and XML [5]. Web Structure Mining uses the hyperlink structure which is useful for social network analysis [15]. The emphasis lies in finding the structural summary about the web sites and web pages.

The objects found in world wide web comprises of web pages, links (inlinks and outlinks) for connecting web pages and co-citation where two pages are linked by same page. Link based mining techniques help to perform different task like-

- On basis of words on page and links between pages category prediction is feasible.
- Same types of links are clustered to discover hidden patterns.
- Purpose, number of links and weight of the links can be assigned on basis of the task performed by it[16].

A. ALGORITHM USED

The links in web structure mining are used to infer knowledge which relates to web documents from which content related to the query is extracted. Analyzing the graph is the main goal of web structure mining and for ranking the web pages based on

links different algorithms are proposed like-

- **Page Rank Algorithm** - This algorithm is used by Google which rank the pages by using web structure and ranking is calculated based on number of pages which point to backlinks. It is rooted in social network analysis and considers a page is of higher rank if the sum of all the backlinks is larger than the page rank [17]. It is used to manage technological drift and large data sets [18].
- **HITS Algorithm** -calculates rank at query time by using weight of authority and hub pages [19].
- **Weighted Page Rank Algorithm** -work on rank score by taking into account the popularity of web pages [20].
- **Weighted Link Rank Algorithm** - rank score is calculated by weight value assigned to links and help in improving precision [21].
- **EigenRumor Algorithm** -focus on blogging sites by calculating hub and authority score of bloggers [22].
- **Distance Rank Algorithm** - find the shortest logarithmic distance for ranking [23].

VI. WEB USAGE MINING

It plays its role when there is a requirement to find out the behavior of the user on basis of how he interacts with the web [15]. Web Usage Mining is automatic discovery of user access patterns which also include referrer logs containing referring pages from web servers. It focuses on predicting the behavior of users by analyzing the navigational patterns which help to discover useful information about the users. WUM is divided into two categories -general Access Pattern where history pages are checked to track the information of the user while in Customized Usage Tracking it is targeted on specific users. Four data sources for collecting information in Web Usage Mining is client level, browser level, server level and proxy level [1]. On the basis of type of usage it can be further classified into Web Server Data- In this user logs are collected by the web server which normally includes IP address, page reference and access time. Application Server Data- Various commercial application servers like Weblogic provide the ability for tracking various kinds of business events and log them in application server logs. Application Level Data-History for some new events can be generated which are defined for an application and logging facility can be turned on for them.

A. TECHNIQUES

The Web Usage Mining comprises of different steps which help in overall cleaning and finding relevant pattern from the data. Different steps are-

- **Preprocessing**

It is the first step used for data which is stored in web logs. It comprises of data cleaning, session reconstruction, content and structure information retrieval and data abstraction. These stages deals with cleaning of data, finding user session and reconstruction, enriching web log data and development of abstraction inform of session and page views. This step is further categorized into usage preprocessing which is difficult task as it has incomplete server log data, content preprocessing which deal with converting unstructured and semi-structured documents into suitable forms by using vector space model and structure

preprocessing which handle hyperlinks.

- **Pattern discovery**

This phase focuses on some algorithms which are quite useful in extracting pattern and understandable knowledge by using statistics, machine learning and data mining techniques. Various algorithms are -Statistical Analysis which helps in improving system performance by focusing on data. Association Rules uses Apriori algorithm which help in finding correlations among web pages. Clustering which is used for grouping similar session and helps to find user having similar behavior. Classification deals with categorization of documents which help in detecting interesting patterns. Sequential Patterns help to find sessions that are placed in a particular order.

- **Pattern Analysis**

This step helps in system performance by separating interesting patterns from uninteresting ones. Various techniques like Knowledge Query Mechanism, Visualization tools and Intelligent Agents are used for analyzing discovered patterns [1]. This phase help in redesigning of the websites, improving access time of fetching the pages and helps in enhancing the browsing of desired pages [24].

B. STAGES

Web Usage Mining helps in recovering user access patterns from web servers. It comprises of four processing stages-

- **Data Collection**

This stage helps the web managers for improving the management, performance and controlling of web servers by collecting the hidden information and usage patterns.

- **Data Preprocessing**

This stage focuses on selecting useful data by eliminating irrelevant and noisy data. It tries to arrange most recently accessed data with higher index with respect to least used by providing lower index value. It is regarded as a critical step as it helps in the analysis to obtain precise result.

- **Data Clustering**

Clustering algorithm are most frequently used in mining web pages and forming cluster objects. This method is helpful in finding usage patterns and user profiles.

- **Pattern Discovery & Analysis**

With help of data analysis and graph relevant and useful information can be easily predicted [d].

C. TOOLS EMPLOYED

Number of sources which help in accessing data in Web Usage Mining includes cookies, web access log, login information, client or server side scripts etc [25]. Different tools are used which help in recovering and performing data preprocessing, pattern discovery and pattern analysis are Data Preparator, SumatraTT, Lisp Miner, SpeedTracer, SEWEBAR-CMS, i-Miner, Argonaut, Webalizer, WebViz and WebMiner. These tools performs the task of cleaning, preprocessing, reconstructing, pattern discovery, rules selection, clustering, analyzing patterns and forming visualization of graphical patterns [1].

D. APPLICATION

As web usage mining keep track of user profile, sessions, transactions, user generated queries, cookies, bookmark data which help in getting an overview about the user interest and behavior. This overall helps in redefining and designing the web structure. It also helps in personalization of web users on the basis of their behavior in tracking web pages which directly helps in designing and implementing websites. It reduces the time of web servers by employing retrieving and caching strategies [3].

VII. CONCLUSION

The paper basically outlines an insight into web mining which is very useful in different set of tasks especially in retrieving information. The paper portrays detailed analysis about web mining categories like web content mining, web structure mining and web usage mining. The Web Content Mining techniques and tools are quite useful in handling user queries, visualizing information as per need, help in maintaining relationship among customers and many more. Different tools are covered which help in achieving these tasks. At the same time it has been found out that Web Structure Mining by using the hyperlinks among pages helps in retrieving desired pages by using different ranking algorithms. These algorithms use different parameters which help them in accessing the desired information in lesser amount of time. Similarly Web Usage Mining by using web logs and web server data help in predicting the behavior of the users which can help in restructuring the websites according to the interest of the users. The process of mining information from the web is a tricky task and different mining techniques can be combined together so that extraction of information and finding hidden pattern can be smoothened.

VIII. REFERENCES

- [1] Kamika Chaudhary and Santosh Kumar Gupta, "Web Usage Mining Tools & Techniques: A Survey", International Journal of Scientific & Engineering Research, Volume 4, Issue 6, June-2013.
- [2] Shadab Irfan, Subhajit Ghosh, " Optimization of Information Retrieval Using Evolutionary Computation: A Survey", International Conference on Computing, Communication and Automation (ICCCA2017).
- [3] Nasrin JOKAR, Ali Reza HONARVAR, Shima AgHAMIRZADEH, Khadijeh ESFANDIARI, "Web mining and Web usage mining techniques", Bulletin de la Société des Sciences de Liège, Vol. 85, 2016.
- [4] Shadab Irfan, Dr. B V Babu, "Information Retrieval in Big Data Using Evolutionary Computation: A Survey", International Conference on Computing, Communication and Automation (ICCCA2016).
- [5] Gadamsetty Vasavi and Dr. T. Sudha, "Information Extraction from Online Shopping Sites using Web Content Mining Methods and Techniques", International Conference on Innovative Applications in Engineering and Information Technology (ICIAEIT-2017).
- [6] Anurag kumar and Ravi Kumar Singh, "A Study on Web Content Mining", International Journal Of Engineering And Computer Science , Volume 6 Issue 1 Jan. 2017.
- [7] Andemariam Mebrahtu and Balu Srinivasulu, "Web Content Mining Techniques and Tools", International Journal of Computer Science and Mobile Computing, Vol.6 Issue.4, April-2017.
- [8] R.Malarvizhi and K.Saraswathi, "Web Content Mining Techniques Tools & Algorithms – A Comprehensive Study", International Journal of Computer Trends and Technology (IJCTT) – volume 4 Issue 8– August 2013.
- [9] Han, J., Kamber, M. Kamber. "Data mining: concepts and techniques". Morgan Kaufmann Publishers, 2000.
- [10] Karan Sukhija, "Web Content Mining equipped Natural Language Processing for handling web data", International Journal of Computer Applications Technology and Research Volume 4– Issue 3, 209 - 213, 2015.
- [11] N. R. Satish, "A Study on Applications, Approaches and Issues of Web Content Mining", International Journal of Trend in Research and Development, Volume 4(6), Nov-Dec 2017.
- [12] Ananthi.J, "A Survey Web Content Mining Methods and Applications for Information Extraction from Online Shopping Sites", International Journal of Computer Science and Information Technologies, Vol. 5 (3) , 2014.
- [13] Brijendra Singh and Hemant Kumar Singh, " WEB DATA MINING RESEARCH: A SURVEY", 2010 IEEE International Conference on Computational Intelligence and Computing Research.
- [14] Mohamed-K Hussein and Mohamed-H Mousa, "An Effective Web Mining Algorithm using Link Analysis", International Journal of Computer Science and Information Technologies, Vol. 1 (3) , 2010.
- [15] Raymond Kosala and Hendrik Blockeel, "Web Mining Research: A Survey", SIGKDD Explorations, ACM July 2000.
- [16] Anurag Kumar, Ravi Kumar Singh, "A Study on Web Structure Mining", International Research Journal of Engineering and Technology (IRJET), volume: 04 Issue: 1 , Jan -2017.
- [17] Ahmad Tasnim Siddiqui and Sultan Aljahdali, "Web Mining Techniques in E-Commerce Applications", International Journal of Computer Applications Volume 69– No.8, May 2013.
- [18] Sergey Brin and Larry Page, "The anatomy of a Large-scale Hypertextual Web Search Engine", In Proceedings of the Seventh International World Wide Web Conference, 1998.
- [19] Jon M. Kleinberg, "Authoritative Sources in a Hyperlinked Environment".
- [20] Wenpu Xing and Ali Ghorbani, "Weighted PageRank Algorithm", Proceedings of the Second Annual Conference on Communication Networks and Services Research , 2004 IEEE.

[21] Ricardo BaezaYates and Emilio Davis, "Web Page Ranking using Link Attributes", May 17–22, 2004,ACM.

[22] Ko Fujimura,Takafumi Inoue and Masayuki Sugisaki, "The EigenRumor Algorithm for Ranking Blogs", WWW 2005.

[23] Ali Mohammad Zareh Bidoki, Nasser Yazdani, "DistanceRank: An intelligent ranking algorithm for web pages', Information Processing and Management,2007 Elsevier.

[24] Yew-Kwong Woon Wee-Keong Ng Ee-Peng Lim, "Web Usage Mining: Algorithms and Results", Centre for Advanced Information Systems, Singapore.

[25] Sasa Bosnjak,Mirjana Maric, Zita Bosnjak, "The Role of Web Usage Mining in Web Applications Evaluation", Management Information Systems, Vol. 5 (2010).