American Journal of Engineering Research (AJER) ***2018***
American Journal of Engineering Research (AJER)
e-ISSN: 2320-0847 p-ISSN : 2320-0936
Volume-7, Issue-8, pp-161-164
www.ajer.org
Research Paper Open Access

Importance of Cloud Computing In Web Mining

Dr. Amamer Khalil Masoud Ahmidat
Higher Institute of Medical Technology in Baniwaleed, Libya
Faculty of Sciences Baniwaleed University -Libya
Corresponding Author: Dr. Amamer Khalil Masoud

ABSTRACT:
This paper deals the web mining using Cloud Computing Technology. Web mining includes how to extract the useful information from the web and gain knowledge using data mining techniques. Here so many resources and techniques are available i.e. web content mining, web structure mining, web usage mining and access through the web servers. Web mining techniques (specially web usage mining techniques) and applications are much needed in cloud computing. The implementation of these techniques through cloud computing will allow users to retrieve relevant and meaningful data from virtually integrated data warehouse which reduces cost and infrastructure.

Date of Submission: 31-07-2018 Date of acceptance: 15-08-2018

I. INTRODUCTION

Web mining is the application of data mining techniques to discover patterns from the Web. According to analysis targets, web mining can be divided into three different types, which are Web usage mining, Web content mining and Web structure mining. Web usage mining is the process of extracting useful information from server logs e.g. use Web usage mining is the process of finding out what users are looking for on the Internet using cloud computing . Some users might be looking at only textual data, whereas some others might be interested in multimedia data. Web Usage Mining is the application of data mining techniques to discover interesting usage patterns from Web data in order to understand and better serve the needs of Web-based applications. Usage data captures the identity or origin of Web users along with their browsing behavior at a Web site. Web usage mining itself can be classified further depending on the kind of usage data considered. Several data mining methods are used to discover the hidden information in the Web.

However, Web mining does not only mean applying data mining techniques to the data stored in the Web. The algorithms have to be modified such that they better suit the demands of the Web.

II. METHODS OF WEB MINING USIGN CLOUD COMPUTING

Concept of web usage mining

The web usage mining generally includes the following several steps: data collection, data pretreatment, and knowledge discovery and pattern analysis.

➢ **Data collection**

Web usage mining is the process of extracting useful information from server logs e.g. use Web usage mining is the process of finding out what users are looking for on the Internet. Some users might be looking at only textual data, whereas some others might be interested in multimedia data. Web Usage Mining is the application of data mining techniques to discover interesting usage patterns from Web data in order to understand and better serve the needs of Web-based applications. Usage data captures the identity or origin of Web users along with their browsing behavior at a Web site. Web usage mining itself can be classified further depending on the kind of usage data considered:

Web Server Data: The user logs are collected by the Web server. Typical data includes IP address, page reference and access time.

Application Server Data: Commercial application servers have significant features to enable e-commerce applications to be built on top of them with little effort. A key feature is the ability to track various kinds of business events and log them in application server logs.

Application Level Data: New kinds of events can be defined in an application, and logging can be turned on for them thus generating histories of these specially defined events. It must be noted, however, that many end applications require a combination of one or more of the techniques applied in the categories.

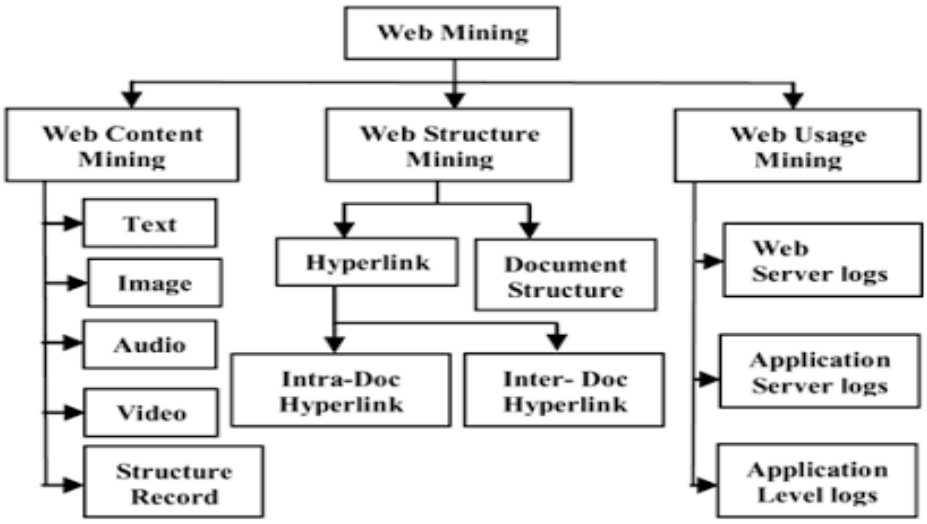
➢ **Data preprocessing**

Web Usage Mining in cloud computing is one of the categories of data mining technique that identifies usage patterns of the web data, so as to perceive and better serve the requirements of the web applications. The working of WUM involves three steps - preprocessing, pattern discovery and analysis. The first step in WUM - Preprocessing of data is an essential activity which will help to improve the quality of the data and successively the mining results. This research paper studies and presents several data preparation techniques of access stream even before the mining process can be started and these are used to improve the performance of the data preprocessing to identify the unique sessions and unique users in cloud computing . The methods proposed will help to discover meaningful pattern and relationships from the access stream of the user and these are proved to be valid and useful by various research tests. The paper is concluded by proposing the future research directions in this space.

In the data pretreatment work, mainly include data cleaning, user identification, session identification and path completion

✓ **Data Cleaning:**

The most important task of the Web Usage Mining in cloud computing process is data preparation. The success of the project is highly correlated to how well the data preparation task is executed. It is of utmost importance to ensure, every nuance of this task is taken care of. This process deals with logging of the data; performing accuracy check; putting the data together from disparate sources; transforming the data into a session file; and finally structuring the data as per the input requirements. The data used for this project is from the RIT Apache server logs, which is in the Common Log File format. This access log includes the agent and the referrer in the data as one of the attributes.



✓ **Path completion**

An implementation of data preprocessing system for Web usage mining and the details of algorithm for path completion are presented. After user session identification, the missing pages in user access paths are appended by using the referrer-based method which is an effective s

➢ **Knowledge Discovery**

olution to the problems introduced by using proxy servers and local caching. The reference length of pages in complete path is modified by considering the average reference length of auxiliary pages which is estimated in advance through the maximal forward references and the reference length algorithms. As verified by practical Web access log, the proposed path completion algorithm efficiently appends the lost information and improves the reliability of access data for further Web usage mining calculations.

In general, knowledge discovery can be defined as the process of identifying interesting new patterns in data. These patterns can be, e.g., relations, events or trends, and they can reveal both regularities and exceptions.

➢ **Pattern analysis**

Challenges of Pattern Analysis are to filter uninteresting information and to visualize and interpret the interesting patterns to the user. First delete the less significance rules or models from the interested model storehouse; Next use technology of OLAP and so on to carry on the comprehensive mining and analysis; Once more, let discovered data or knowledge be visible; Finally, provide the characteristic service to the electronic commerce website.

Survey on web mining

I found many paper related to web usage mining these are following:

A Research Area in Web Mining: This paper also discusses an application of WUM, an online Recommender system that dynamically generates links to pages that have not yet been visited by a user and might be of his potential interest. Differently from the recommender systems proposed so far, it does not make use of any off-line component, and is able to manage Web sites made up of pages dynamically generated.

Cloud Mining: This paper also discussed about Web usage mining and user behavior analysis using fuzzy C-means clustering: In this paper I got methodologies used for classifying the user using Web Usage data. This model analysis the users behaviors and depend on the interests of similar patterns provides appropriate recommendations for active user.

Discovery and Applications of Usage Patterns from Web Data: This paper provides an up-to-date survey of the rapidly growing area of Web Usage mining. With the growth of Web-based applications, specifically electronic commerce, there is significant interest in analyzing Web usage data to better understand Web usage, and apply the knowledge to better serve users.

Web Mining Using Cloud Computing: This paper provides present the technology of cloud computing using web mining .Web mining include how to extract the useful information from the web and gain knowledge using data mining techniques. Here so many online resources are available i.e. web content mining and access through the web servers.

➢ **Online Web Usages Mining in Cloud System**

Web based recommender systems are very helpful in directing the users to the target pages in particular web sites. Moreover, Web usage mining cloud model systems have been proposed to predict user's intention and their navigation behaviors. In the following, we review some of the most significant WUM systems and architecture that Cloud model for navigation pattern mining through Web usage mining to predict user future movements. The approach is based on the graph partitioning clustering algorithm to model user navigation patterns for the navigation patterns mining phase.

III. CONCLUSION

Cloud Computing is a broad term that describes a broad range of services. As with other many vendors have seized the term "Cloud" and are using it for products that sit outside of the common definition. In order to understand how the Cloud can be of value to an organization, it is first important different components. Since the Cloud is a broad collection of services, organizations can choose where, when, and how they Cloud Computing. In this paper it explained that the different types of Cloud Computing Service, Platform as a Service and Infrastructure as a Service with the help of web using mining in cloud Services.

REFERENCES

[1]. Yongjian Fu et al , "A Framework for Personal Web Usage Mining".
[2]. Rajni Pamnani, Pramila Chawan Department of computer technology, VJTI University, Mumbai "Web Usage Mining: A Research Area in Web Mining".
[3]. Jaideep Srivastava et al , "Discovery and Applications of Usage Patterns from Web Data" , Jan 2000 Volume 1, Issue 2 - page 12.
[4]. Kaikala Anjani Sravanthi1, "Web Mining Using Cloud Computing" , ISSN 2250-2459, ISO 9001:2008 Certified Journal, Volume 3, Issue 4, April 2013.
[5]. Vivek Raich, Pradeep Sharma, Shivlal Mewada and Makhan Kumbhkar, "Performance Improvement of Software as a Service and Platform as a Service in Cloud Computing Solution", ISROSET-International Journal of Scientific Research in Computer Science and Engineering, Volume-01, Issue-06, Page No (13-16), Nov -Dec 2013
[6]. V. Cadez, D. Heckerman, C. Meek, P. Smyth, and S. White. Visualization of navigation patterns on a web site using modelbased clustering. In In Proceedings of the sixth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining,pages 280{284, Boston, Massachusetts, 2000.

Amamer Khalil Masoud Ahmidat PhD. in the computer and information Faculaty of electrical Engand information technical University of Kosice Slovaki (2003-2008) . BSc electronic .Eng Computer Dept (1987-1991).MSc in Computer engineering and information Faculaty of electrical Eng.and information technical University of Kosice Slovaki (1995-1998) . **Head of Higher Institute of Medical Technology in Baniwaleed-Libya,** Assistant Professor from (01-10-2013.)

Dr. Amamer Khalil Masoud"Importance of Cloud Computing In Web Mining." American Journal of Engineering Research (AJER), vol. 7, no. 08, 2018, pp. 161-164