ABSTRACT: The amount of information over internet has been growing last few years. And it has caused risk of information problem of accessing related data to the users. The information demand of the online users can be figured out by evaluating user’s web navigation behavior. Web Usage Mining (WUM) is used to extract knowledge from Web user’s access logs by using Data Mining Techniques. One of the applications of WUM is Web Sites Recommendation system which is personalized information filtering technique used to either determine whether a certain user will approve a given item or to identify a list of items which can be of significant importance to the user. In this paper the modified architecture that integrates item information with user’s access log data and then find pattern and make pattern clustering. There after generates a set of recommendations for the user. So execution time and fetching time is reduced.

Keywords: Web Usage Mining, Online Web Recommendation System, Pattern Matching, K-Means.

I. INTRODUCTION

Web mining is the application of data mining techniques to extract knowledge from web data, including web documents, hyperlinks between documents, usage logs of web sites, etc. Web mining can be broadly divided into two main categories, according to the kinds of data to be mined.

Web Content Mining

Web content mining is the process of extracting useful information from the contents of web documents. Content data is the collection of facts a web page is designed to contain. It may consist of text, images, audio, video, or structured records such as lists and tables. Application of text mining to web content has been the most widely researched. Issues addressed in text mining include topic discovery and tracking, extracting association patterns, clustering of web documents and classification of web pages. Research activities on this topic have drawn heavily on techniques developed in other disciplines such as Information Retrieval (IR) and Natural Language Processing (NLP). While there exists a significant body of work in extracting knowledge from images in the fields of image processing and computer vision, the application of these techniques to Web content mining has been limited.

Web Structure Mining

The structure of a typical Web graph consists of Web pages as nodes, and hyperlinks as edges connecting related pages. Web Structure Mining is the process of discovering structure information from the Web. This can be further divided into two kinds based on the kind of structure information used. Hyperlinks

Document Structure

Web Usage Mining

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 [4]. 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

Application Server Data

Application Level Data

II. THEORETICAL REVIEW

Various recommendation systems use different approaches based on the sources of information they utilize. The accessible sources are user information (demographics), the product information (keywords, genres) and the user-item ratings [8]. Recommendation system exhibits certain limitations such as intelligence, adaptability, flexibility, Limited accuracy. This paper aims to provide an architecture that integrates product information with users access log data and then generates a set of recommendations for that particular user is presented [8]. These disadvantages can be overcome by implementing a hybrid architecture that integrates product information with user’s access log data and then generates a set of recommendations for that particular user using Boyer-Moore Pattern Matching Algorithm and K-Means Clustering Algorithm.

III. ONLINE RECOMMENDATION RELATED ALGORITHMS

3.1 An Online Recommendation System based on web usage mining and semantic web using LCS Algorithm [1]

In this paper they used two tier architecture for integrating semantic information with web usage mining. The architecture for recommending the list of products to the users can be partitioned into two main phases; offline phase and online phase [14]. These two phases work together strongly. In the offline phase there are two main modules, data pre-processing and semantic knowledge base. The main module of the online phase is Recommendation engine. They have used LCS algorithm to generate a list of recommended products to the user and resolved it using Huffman code and
This paper have put forth a customer end based system where at customer end web action of the customer framework is assessed. This assessment helps in developing the transactional patterns that helps in personalization and behavioral analysis. A frequent access pattern algorithm is used to divide the data into time slices. This discovering the pattern in data which aids data is formulated using Hidden Markov Model (HMM) for in generating much accurate and efficient recommendations.

In this paper system implemented the significant elements of web recommendation and personalization like Modelling of web pages or web products and customers, Mapping between appropriate customers and products and determining the set of recommendations. This system made use of Ontology and OWL (Web Ontology Language) for web product classification. This system avoided false positive recommendations i.e. the product will be recommended even though it would not be relevant to the customer. The system does not recommend to a customer if the number of earlier recommendations to his/her has not exceeded a dedicated amount of threshold.

3.8 Enhancement of Online web Recommendation System Using a Hybrid Clustering and Pattern Matching Approach [8]
In this paper author used K-Means clustering algorithm and boyar more pattern matching algorithm. This Online Web Recommendation System displays a list of recommended products based on the user’s recent history. It involves clustering of user profiles which leads to searching of patterns in clusters rather than searching whole user logs, thereby reducing time consumption and thus increasing performance of the overall system. It Provides Recommendation based on previous history. The recommendation will be generated based on feature.

IV. CONCLUSION
This Advance System displays a list of recommended items. It involves clustering of pattern items which leads to searching of patterns in clusters rather than searching whole items from databases. It reduces execution time and thus increasing performance of the overall system. The system doesn’t suffer from searching items in whole database ,it just find from pattern cluster which is to be recommended. Future scope includes using K-Means clustering the performance values obtained as a new parameter which is going to improving the performance of this advance recommendation system.
<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Summary</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.1</td>
<td><strong>An Online Recommendation System based on web usage mining and semantic web using LCS Algorithm.</strong></td>
<td>The system comprised of online and offline phase. RDF model is used for Incorporating Semantic information. System doesn’t suffer from new item problem. If the user is visiting the server for the first time there will be no list of recommended products available until the user purchases any product. Does not involve clustering of user profiles.</td>
</tr>
<tr>
<td>2.2</td>
<td><strong>A Personalized Product Based Recommendation System Using Web Usage Mining and Semantic Web.</strong></td>
<td>It solve the new item problem. System does not involve clustering of user profiles Reduced the performance of system.</td>
</tr>
<tr>
<td>2.3</td>
<td><strong>A Web Usage Mining Approach Based on LCS Algorithm in Online Predicting Recommendation Systems</strong></td>
<td>The system involved two phases. Data pre-treatment and navigation pattern mining is carried out in offline. Increase the accuracy up to 73%. It use a fixed-size sliding window over current active session to capture the current user’s activities.</td>
</tr>
<tr>
<td>2.4</td>
<td><strong>Using Semantic Information for web usage mining based recommendation</strong></td>
<td>Sequential association rule mining is used. Time consuming process.</td>
</tr>
<tr>
<td>2.5</td>
<td><strong>Research on the Model of Integration with Semantic Web and Agent Personalized Recommendation System</strong></td>
<td>It use AI-Multi Agent technique. The system exhibits limited degree of intelligence, autonomy and flexibility.</td>
</tr>
<tr>
<td>2.6</td>
<td><strong>Analysis on Recommend ed System for Web Information Retrieval Using HMM</strong></td>
<td>A frequent access pattern algorithm is used to divide the data into time slices. Data is formulated using Hidden Markov Model (HMM). Generating much accurate and efficient recommendations. Only work with client end system. Use rating to generate the recommendation</td>
</tr>
<tr>
<td>2.7</td>
<td><strong>A Semantic Recommendation Procedure for Electronic Product Catalogue.</strong></td>
<td>This system made use of Ontology and OWL (Web Ontology Language) for web product classification. This system avoided false positive recommendation. The system does not recommend to a customer if the number of earlier recommendation s to his/her has not exceeded a dedicated amount of threshold.</td>
</tr>
<tr>
<td>2.8</td>
<td><strong>Enhancement of Online web Recommendation System Using A Hybrid Clustering And Pattern Matching approach</strong></td>
<td>It involves clustering of user profiles. Searching of patterns is done in clusters rather than searching whole user logs. It reduce the time. Increase the performance. Current recommendation system exhibits certain limitations such as intelligence, adaptability, flexibility, limited accuracy</td>
</tr>
</tbody>
</table>

**REFERENCES**


International Conference IEEE 2009.


[13] www.w3c.org
