

# A Survey on Online Shopping Recommendation Based on Customer Transactions

**Priya Kathirvel, Information Technology, IFET College of Engineering, Villupuram, India.**

**Jeba Moses T, Assistant Professor, Information Technology, IFET College of engineering, Villupuram, India.**

**Abstract—** This paper provides a survey of different shopping recommendation methods and techniques. Customer based recommendation play an essential role in many data mining tasks that try to find interesting patterns from databases, such as association rules, sequences, classifiers, clusters and, many more of which the mining of association rules is one of the most popular techniques. Association rules describe how often items are purchased together. The aspects of Web mining, like clustering, association rule mining, information retrieval, text and image mining are considered under the existing taxonomy. The role of fuzzy logic is to identifying the promising transactions based on the customers.

**Index Terms—** Fuzzy logic, Information retrieval, Clustering, Association rule.

## I. INTRODUCTION

Web mining refers to the use of data mining techniques to automatically retrieve, extract and evaluate (generalize/analyze) information for knowledge discovery from Web documents and services. Web data is typically unlabeled, distributed, heterogeneous, semi-structured, time varying, and high dimensional. Ordinary end-users often face difficulties in formulating a precise representation of their information needs in a Boolean query. This affects the efficiency of the information retrieval process.

Data mining or Knowledge Discovery is the process of analyzing data from different perspectives and summarizing it into useful information. This information can then be used to increase revenue, cuts costs, or both. A software created with Data mining as its basic theme should allows users to analyze data from many different dimensions or angles, categorize it, and summarize the relationships identified. Technically, data mining is the process of finding correlations or patterns among dozens of fields in large relational databases. As a marketing tool it is employed to mine out the frequent item sets in a large no of transactions. Thus it is also called “Frequent Item-set Mining”. Frequent

itemset mining: Itemset mining focused on discovering frequent itemset. Itemset is frequent if its support satisfies given minimum support threshold. Infrequent itemset mining: Patterns that are rarely found in database are considered to be unexciting and are eliminated using the support measure. Item set is infrequent if its support is less than or equal to predefined support threshold. This method has a great interest as they deal with rare but crucial cases. Now a days Infrequent itemset mining can be achieved through many applications. Applications in infrequent itemset mining include identifying rare diseases, predicting equipment failure, and finding association between infrequently purchased items.

## II. LITERATURE SURVEY

In this section we are focus on different techniques used in recommendation systems. [1] Utility Itemsets Mining Algorithm: Association rule mining (ARM) identifies frequent itemsets from databases and generates association rules by considering each item in equal value. However, items are actually different in many aspects in a number of real applications, such as retail marketing, network log, etc. The difference between items makes a strong impact on the decision making in these applications. The advantage of this paper is that it performs very efficiently in terms of speed and memory cost. The main disadvantage of this paper is “Downward closure property” doesn’t apply to utility mining.

[2] Mining Frequent Itemsets from Data Streams with a Time-Sensitive Sliding Window: An effective bit-sequence based, one-pass algorithm, called MFI-TransSW (Mining Frequent Itemsets within a Transaction-sensitive Sliding Window). Used to mine the set of frequent itemsets from data streams within a transaction-sensitive sliding window which consists of a fixed number of transactions. A mechanism that is self-adjusting under the memory limitation is presented. All these methods have to rescan the original database because non-frequent itemsets can be frequent after the database is updated. This is one main drawback of this method.

[3] Mining of Temporal Maximal Utility Itemsets from Data Streams: Temporal High Utility Item (THUI) mining has become an emerging research topic in the data mining field, and finding frequent item sets is an important task in data mining with wide applications. A novel approach, TOUI-tree (Temporal Optimal Utility Item set

tree), is also used for efficiently capturing the utility of each item set with one-time scanning. The advantage of this technique is Knowledge discovery from online e-business or transaction flows, analysis of network flows, monitoring of sensor data, and so on. Disadvantage is, infinite transactions could not be stored, multi-scan algorithms are no more allowed.

[4] Incremental Mining of Frequent Patterns in Evolving Time Series Databases: Time series databases are characterized by two features: The continuous arrival of data and the time dimension. These features raise new challenges for data mining such as the need for online processing and incremental evaluation of the mining results. This approach updates the mining results with the arrival of every new data item by considering only the items and patterns that may be created by the newly arrived item. Advantage of this paper is Address the problem of discovering frequent patterns in databases with multiple time series. It is crucial to provide for these applications algorithms that are more efficient than general sequence mining algorithms is a drawback of this paper.

[5] Mining of User Behaviors by Temporal Mobile Access Patterns: Efficient Mining of User Behaviors by Temporal Mobile Access Patterns is aimed at modeling how to build mobile user's behavior pattern with a temporal association rule. In mobile agent systems, the mobile agent system consists of semi-structure data like XML data. Temporal association rules can be used to decide the next likely user's request services based on significant dynamic correlations. The advantage of this paper provides a data structure, that compactly stores the user's behavior pattern according to location and service information in memory. The disadvantage is that, these models were not designed for the user's temporal movement patterns in mobile agent systems.

### III. PERFORMANCE COMPARISON

Recommender systems are one of the most important proposed solutions for the exploration and management of users' favorites. This graph shows the comparison between different mining techniques. Content-based filtering has been considered in the web page recommender systems. The collaborative filtering (CF) increases the performance. In THUI mining the performance is high but it provides very poor user interface and accessing time. Based on the outcome of the following graph, it is clear that the collaborative filtering algorithm provides better performance than the other techniques. Temporal High Utility Item (THUI) mining has become an emerging research topic in the data mining field, and finding frequent item sets is an important task in data mining with wide applications. Data mining is the process of analyzing data from different perspectives and summarizing these data into useful information. This information can be used to increase revenue, cuts costs, or both of these. A software created with Data mining as its basic theme should allow users to analyze data from many different dimensions.

Factors	Performance	Time	User Interface
THUI Mining	4.2	2.5	1
MFI-TransSw	2.5	4.3	2
ARM	3.5	1.8	3
Collaborative Filtering	4.5	3	5

Table 1. Performance Comparison

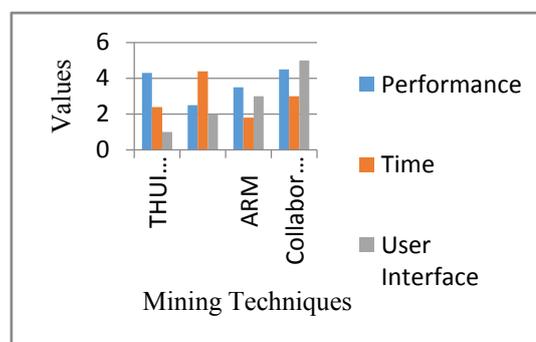


Fig.1. Performance Comparison Chart

### IV. CONCLUSION

With the help of Incremental Association Rule Mining and Transaction Clustering, It introduced a method to design an improved and well structured website design for an E-shop in the design phase. Assuming that the two thresholds, minimum support and confidence, do not change, the promising frequent algorithm can guarantee to discover frequent itemsets. It have used an efficient clustering algorithm for data items to minimize the SL ratio in each group. The algorithm is able to cluster the data items very efficiently. The fuzzy logic used in this algorithm not only incurs an execution time but also leads to the clustering results of very good quality.

### REFERENCES

- [1] Ying Liu, Wei-keng Liao, Alok Choudhary on "A Fast High Utility Itemsets Mining Algorithm".
- [2] Chih-Hsiang Lin, Ding-Ying Chiu, Yi-Hung Wu, Arbee L. P. Chen on "Mining Frequent Itemsets from Data Streams with a Time-Sensitive Sliding Window".
- [3] Bai-En Shie, Vincent S. Tseng, Taiwan, Philip S. Yu on "Online Mining of Temporal Maximal Utility Itemsets from Data Streams".

- [4]Mohamed Y. Eltabakh, MouradOuzzani, Mohamed A. Khalil on “Incremental Mining for Frequent Patterns in Evolving Time Series Databases”.
- [5]Seung-Cheol Lee, Juryon Paik, Jeewoong Ok, Insang Song and Ung Mo Kim on “Efficient Mining of User Behaviors by Temporal Mobile Access Patterns”.
- [6]W. Fan, K. Zhang, J. Gao, X. Yan, J. Han, P. Yu, and O. Verscheure, “Direct mining of discriminative and essential graphical and itemset features via model-based search tree,” In Proc. KDD, 2008, pp. 230–238.
- [7]A. Jorge, “Hierarchical clustering for thematic browsing and summarization of large sets of association rules,” in Proc. SDM, 2004, pp. 178–187.
- [8]S. Chen, B. Mulgrew, and P. M. Grant, “A clustering technique for digital communications channel equalization using radial basis function networks,” IEEE Trans. on Neural Networks, vol. 4, pp. 570-578, July 1993.

## AUTHORS BIOGRAPHY

**Ms. Priya Kathirvel** Currently pursuing B.Tech, Information Technology at IFET College of Engineering, Villupuram, India. Her area of interests includes OOPS concepts, Web Designing and Data Mining.

**Mr. T. Jebamoses** , M.Tech, MISTE. He is currently working as Assistant Professor in Department of Information Technology, IFET College of Engineering, Villupuram, India. He has published six papers in international journals. His area of interests includes Mobile Computing, Computer Networks and Object Oriented Programming.