A Survey on Recommendation System Using Web Usage Mining

Harshit Patel[1], Prof. Pooja Jardosh[2]

[1] Student, M.E (CE), SOCET, GTU
Ahmedabad, Gujarat, India
Harshit06patel@gmail.com

[2] Assistant Professor, M.E (CE/IT), SOCET, GTU
Ahmedabad, Gujarat, India

ABSTRACT

Recommender systems are intelligent systems which make suggestions about user items. Many of the largest commerce Web sites are already using recommender systems to help their customers find products to purchase. Recommender system has become an important part of any entertainment or E-Commerce website. Various personal services in business play important roles in the success of current marketing field. The personalized recommendation technique in recommender systems, one of the most important tools of personal service in websites, makes great significance in Internet marketing activities of e-Commerce.

Keywords - E-Commerce, Recommender System, Web Personalization, Web Usage Mining

I. INTRODUCTION

More and more E-commerce Websites provide products with different prices which made it hard for consumers to find the products and services they want. Technological improvement has led to an explosive growth of recorded information, with the Web being a huge storehouse under no editorial control. Recommender systems are used by E-commerce sites to suggest products to their customers and to provide consumers with information to help them decide which products to purchase. Web personalization can be defined as any action that personalizes the Web experience to a particular user, or a set of users.

1. Recommendation System

The products can be recommended based on the top overall sellers on a site, or on an analysis of the past buying behavior of the consumer as a prediction for future buying behavior. The forms of recommendation include suggesting products to the consumer, providing personalized product

information, summarizing community opinion, and providing community critiques. Broadly, these recommendation techniques are part of personalization on a site because they help the site adapt itself to each customer. Text summarization is a chain of a compressing a given document into an abbreviated variant by extricating the most imperative information from it.

1.1 Collaborative filtering

These approaches building a model from a user's past behavior (items previously purchased or selected and/or numerical ratings given to those items) as well as similar decisions made by other users. This model is then used to predict items (or ratings for items) that the user may have an interest in. Collaborative filtering approaches often suffer from three problems: cold start, scalability, and sparsity.

• There are two major approaches for collaborative filtering algorithms:

- **Model-based approaches:** These approaches use training data to generate a model. These models have been used to predict the ratings for the items that a user has not been rated before. In this approach the raw data is usually processed offline. For example, decision trees, aspect models, latent factor models and clustering methods are model-based approaches for collaborative filtering.
- **Memory-based approaches:** These approaches look at similar users or items based on their previous rating and combine their ratings in order to make new predictions. In this approach, the raw data is kept and processed in memory. Examples of memory-based collaborative filtering algorithms are user-based and item based methods.

1.2 Content-based filtering

This approach uses a set of discrete characteristics of an item to recommend more items with similar properties. It is based on item description and user preference. For example, the music recommendation 'Pandora' uses the properties of a song or the singer in order to create a station which plays music with similar properties. In content based recommendation you compare items based on their features for movies things like title, genre, release date, director, producers, studio, etc.

1.3 Demographic Recommender System

Demographic recommender systems aim to categorize the user based on personal attributes and make recommendations based on demographic classes. A demographic recommender system uses the personal information of the user to make predictions regarding their likes. The benefit of a demographic approach is that it may not require a history of user ratings of the type

needed by collaborative and content-based techniques.

1.4 Knowledge-Based Recommender System

Knowledge-based recommender systems are a specific type of recommender_system that is based on explicit knowledge about the item assortment, user preferences, and ultimately, how the item is useful for the user. Knowledge-based systems tend to work better than others at the beginning of their deployment but if they are not equipped with learning components they may be surpassed by other shallow methods that can exploit the logs of the human/computer interaction.

1.5 Hybrid Recommendation

In some cases, combining the results of any two recommendation technique like collaborative and content-based filtering proves to be more efficient. It can be done by making collaborative and content based recommendations separately and then merging them or by adding content based to collaborative approach or vice versa. Netflix is an example of this method. They provide recommendations after analyzing the browsing habits of similar users (collaborative filtering) as well as by offering movies which have similar traits with the films that a user has rated highly (content filtering).

2. Web Mining

It is the application of data mining techniques to discover patterns from the World Wide Web. Web mining has mainly three types.

2.1 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. Usage data captures the identity of Web users along with their browsing behavior at a Web site.

2.2 Web Structure Mining: Web structure mining is the process of using graph theory to analyze the node and connection structure of a web site.

2.3 Web Content Mining: Web content mining is the mining, extraction and integration of useful data, information and knowledge from Web page content. The heterogeneity and the lack of structure that permits much of the ever-expanding information sources on the World Wide Web, such as hypertext documents, makes automated discovery, organization.

II. COMPARISONS OF RECOMMENDATION TECHNIQUES

All recommendation techniques have strengths and weaknesses discussed below and summarized in Table I.

Technique	Advantages	Disadvantages
Collaborative	• It can produce personalized	• New user ramp-up Problem

filtering	recommendations • Domain knowledge not needed. • Adaptive: quality improves over time. • It not require content information • Implicit feedback sufficient	• New item ramp-up problem • "Gray sheep" problem • Quality dependent on large historical data set. • Sparsity problem
Content-based	• Domain knowledge not needed. • Adaptive: quality improves over time. • Implicit feedback sufficient • Does not require any user input to recommend	• New user ramp-up problem • Quality dependent on large historical data set. • Stability vs. plasticity problem • Absence of personal recommendations
Demographic	• Can identify cross-genre slot. • Domain knowledge not needed. • Adaptive: quality improves over time.	• New user ramp-up problem • "Gray sheep" problem • Stability vs. plasticity problem • Must gather demographic information
Knowledge-based	• No ramp-up required • Sensitive to changes of preference • Can include non-product features • Can map from user needs to Products	• Knowledge engineering required. • Suggestion ability statics
Hybrid	• No Item ramp-up required • No user ramp-up required • Reduce the sparsity problem	• Implementation is very complex

Table I [4]

III. Related Works

Many authors presented different recommendation systems and techniques to improve the suggestions or recommendations for users. This section comprises a brief mention about some methods proposed by different authors.

Authors [5] have improved the efficiency of the collaborative filtering approach based recommendation which overcomes the problem of sparsity and scalability by using web usage mining, decision tree induction method, association rule mining algorithms and data warehousing technologies. They have used web logs as a source to find the frequent patterns using Apriori algorithm and built product taxonomy. Decision tree induction method is used to classify the customers, finally a recommendation system

with five levels have introduced to recommend the items to the customer.

Authors [6] have done a recommendation using web usage mining with two major data mining algorithms such as clustering and association rule mining. They have used Hierarchical Bisecting Mediods for clustering the users with respect to time framed session. Association rules are applied to above formed groups to find the similar kind of students in future.

Authors [7] have proposed a recommendation system mainly based on applying association rule mining concept in the web log files for better recommendation. They have introduced one algorithm for finding the association rule from the web logs called Formal Concept Analysis (FCA) using lattice theory. They have proved that, the new algorithm is better than the Apriori algorithm.

Authors [8] have developed a novel recommendation system for the students who are all themselves learning technologies through e-learning environment using web usage mining. Their recommendation system will automatically suggest the educational resources for the students based on their browsing history. Learner and Content module is built in offline then recommendations are suggested based on the above built model.

Authors [9] have done the recommendation based on collaborative filtering technique only for the trustworthy customers. Entropy based similarity measure is used to identify the similarity between the users.

Authors [10] have presented a web page recommendation algorithm using weighted sequential patterns and markov model.

The researchers have [11] constructed a new model to understand the user personal behavior about webpage navigation using Latent Dirichlet Allocation (LDA). They have also proposed three types of recommendation models namely pure-LDA, LDA-knn, and LDA-tran.

Authors [12] have discussed web usage mining is considered as the major source for web recommendation in association with Collaborative filtering approach. Authors use the association rule mining and Markov model to recommend the web pages to the user. The traditional Apriori algorithm is improved by adding the time duration spent on each web page. Markov model is used for recommending the web pages based on user's past history.

IV. CONCLUSION

Web is growing rapidly, but on the other hand the user's capability to access Web content remains constant. Currently, Web personalization is the most promising approach to alleviate this problem and to provide users with tailored experiences. Web-based applications (e.g.,e-commerce sites, e-learning systems, etc.) improve their performance by addressing the individual needs and preferences of each user,

increasing satisfaction of user. In this paper, we discussed Web personalization as one of the solutions to this problem, which makes use of Web usage mining. Summarizing, in this paper we explored the different faces of personalization.

V. REFERENCES

[1]. Recommender Systems Handbook Francesco Ricci · Lior Rokach · Bracha Shapira · Paul B. Kantor .Springer.

[2]. Hybrid Recommender Systems: Survey and Experiments, Robin Burke, California State University, Fullerton Department of Information Systems and Decision Sciences.

[3]. Bridge, D., G¨oker, M., McGinty, L., Smyth, B.: Case-based recommender systems. The Knowledge Engineering review 20(3), 315–320 (2006).

[4]. Intelligent Decision and Policy Making Support Systems edited by Da Ruan, Frank Hardeman, Klaas van der Meer

[5]. Yoon Ho Cho, Jae Kyeong Kim, Soung Hie Kim. 2002.A personalized recommender system based on web usage mining and decision tree induction. Expert Systems with Applications 23, 329–342.

[6]. Feng Hsu Wanga, Hsiu-Mei Shao. 2004. Effective personalized recommendation based on time-framed navigation clustering and association mining. Expert Systems with Applications. 27, 365–377.

[7]. Baoyao Zhou, Siu Cheung Hui, Kuiyu Chang. 2005. A Formal Concept Analysis Approach for Web Usage Mining. Intelligent Information Processing II IFIP International Federation for Information Processing. 163, 437-441.

[8]. Mohamed Koutheaïr Khribi, Mohamed Jemni1 and Olfa Nasraoui. 2009. Automatic Recommendations for E-Learning Personalization Based on Web Usage Mining Techniques and Information Retrieval. Educational Technology and Society, 12 (4), 30–42.

[9]. Harita Mehta, Shveta Kundra Bhatia, Punam Bedi and Dixit, V., S. 2011. Collaborative Personalized Web Recommender System using Entropy based Similarity Measure. IJCSI International Journal of Computer Science. 8(6),3, 1694-0814.

[10]. Suneetha, K., and Usha Rani, M. 2012. Web Page Recommendation Approach Using Weighted Sequential Patterns and Markov Model. Global Journal of Computer Science and Technology. 12(9).

[11]. Qinjiao Mao, Boqin Feng, Shanliang Pan,. 2013. Modeling User Interests Using Topic Model. Journal of Theoretical and Applied Information Technology. 48(1).

[12]. R. Suguna. , D. Sharmila. 2013. An Efficient Web Recommendation System using. Collaborative Filtering and Pattern Discovery Algorithms. International Journal of Computer Applications (0975 – 8887). Volume 70– No.3, May 2013. 37.