

E-Commerce Recommendation over Big Data Based on Structural Balance Theory with the Prediction Rates

Pavithra R & Dr. K.Satyanarayan Reddy

¹Student, CSE department, Cambridge Institute of Technology.

²HOD, ISE, Cambridge Institute of Technology.

Abstract: Through E-commerce, users can browse and select the products that they like, in a more convenient manner which brings great facility to the E-commerce users. Generally, in an E-commerce company, there are varieties of products that are ready to be compared, selected and of products that are ready to be compared, selected and purchased by users. Therefore, from the E-commerce companies perspective, predicting users interests, preferences and recommending or suggesting the appropriate products to him/her, is becoming the key to ensure the success of E-commerce companies. With the well-known Collaborative Filtering (i.e.CF) recommendation, we can determine the similar friends of the user, or the similar product of the user's preferred products, and further put forward CF recommendation methods. The traditional CF-based recommendation approaches can work very well, when the user has one or more similar friends, or the user's purchased preferred products have one or more similar products. While due to the sparsity of rating data in E-commerce, similar friends and similar products may be both absent from the user-product purchase network, which lead to a big challenge to recommend appropriate products to the user. Considering the challenge, Structural Balance Theory based Recommendation (i.e., SBT-Rec) approach was proposed where user's "enemy" (i.e., the users having opposite preference with target user are searched for; then the user's "possible friends" are determined, according to "enemy's enemy is a friend" rule of Structural Balance Theory, and recommend the product preferred by "possible friends" of user to the user. In this paper, we enhance the SBT-Rec technique and extend it to help the users in selecting restaurants, airlines etc. We collect all the existing restaurants and users data from a big E-commerce company and recommend the list of restaurants to each user based on his

interests which are collected during the time of registration and also considering all the restaurants preferred by the friends with similar interests. Also we calculate the prediction rate for each user which represents how correctly we have recommended the list of restaurants to the user by analyzing how many times the user have selected that particular restaurant and based on his ratings. This prediction rate can be used for further research and analysis of the correctness of the proposed recommendation methods and helps in choosing the best method.

1. Introduction

Due to the null or insufficient feedback incentive in E-commerce applications, many online shopping users are not willing to give their ratings on product items after the purchase, which generates a big but sparse user-product rating matrix. In this situation, for the target user, his/her similar friends and similar product items are both absent from the user-product purchase network, which may lead to a failure of traditional CF-based recommendation approaches and a big challenge for accurate product item recommendation for the target user, has increased the users of E-commerce.

As the popularity and usage of internet has increased, the users of E-commerce have increased as well, all over the world. Using E-commerce websites, users can search, compare and select any kind of product items including electronic gadgets, house hold items, clothing booking hotel, flight, movie tickets etc., they are interested at their own convenience. Today, many E-commerce companies (e.g. Amazon, Flipkart) have provided various product items to their massive online users.

Generally, in each E-commerce company, that are ready to be compared, selected & purchased by target users. From the perspective of E-commerce companies, accurately predicting target users'

preference and further recommending appropriate product items to him/her, is becoming the key to ensure the continuous success of Ecommerce companies. In view of this, many recommendation approaches are brought forth, e.g. the well-known collaborative Filtering (i.e., CF)-based recommendation. Through observing the big rating data in user-product purchase network, we can determine the similar friends of target user, or the similar product items of target user's preferred product items, and further put forward CF recommendation methods such as item-based one user based one or hybrid one.

2. Existing system

2.1. A Practical Model to Predict The Repeat Purchasing Pattern of Consumers in The C2C E-commerce

In C2C e-commerce, the enterprises just provide the platform, and they are not involved in the trading. The seller and consumers can free trade on this platform. The seller can provide goods for online auction, and the buyer can choose goods online for bidding. C2Ce-commerce contains various groups of consumers irrespective of the age, gender, etc. It decides based on the demand recognition. When the consumers create psychological demand, they start gathering the information, selection, and evaluation and then finally make the decision of purchasing. Consumer demand confirmation will be generated at this time. After knowing consumer's requirements, C2C would begin to show the relevant information. Firstly, it shows the commodity information, then the third party evaluation. The consumers would collect all this collective information of the goods which is very important in the buying decision. Compared with business information, consumers show more concern about the objective evaluation of the third party. If consumers are satisfactory in the shopping experience, their evaluation is the propaganda for the industry which helps further analyzing and making it easy for the users to search and select the products of their interest.

2.1. Understanding Customers' Repeat Purchase Intentions in B2C E-commerce: The Roles of Utilitarian Value, Hedonic Value and Perceived Risk

In a Business-to-Consumer (B2C) E-commerce, companies sell their online goods to consumers who are the end users of their products or services. Any users have an open access to the B2C E-commerce web sites. A consumer opens the website, selects and orders the catalog and then the business organization

is intimated through the e-mail. After receiving the order, goods would be dispatched to the consumer. In B2C website, the products are traded directly between the manufacturers and consumers, the business layers responsible for intermediary functions are removed- Disintermediation. There is another process called Re-intermediation wherein the business layers responsible for intermediary functions are shifted from traditional to electronic mediums.

2.2. Which Is More Important in Internet Shopping, Perceived Price Commerce Research and Applications or Trust?

In the context of online shopping, the effect of price is stronger as customers can easily compare prices from one store to another. Trust is considered as important factor too in customers' decision of purchasing. It tries to find the relative influence of perceived trust and perceived price on customers' intention to buy. In this method to measure perceived trust, respondents should have tried the online market, to measure perceived price, respondents should have seen the prices offered online. To obtain respondents who fit the criteria, respondents were instructed to visit the site. Respondents were also asked whether they had signed up online site or not. It was intended to group respondents into two distinct groups: potential customers and repeat customers. They were asked to enter an item that had been chosen along with the price to a form. Limitation of item value were made to ensure that respondents saw the price of the products offered, so their perceived price can be measured. After meeting the criteria, respondents were asked to fill out online questionnaires in Google Docs. Respondents were given questions to measure the variables perceived trust, perceived price, perceived value and purchase intentions. Questionnaires were randomly distributed using iMacros program via social media.

2.3. Predicting Future Purchases with The Poisson Lognormal Model

We propose an approach that creates the Poisson log normal distribution (PLN) for modeling purchase frequency counts and predicting future purchases based on past performance. The proposed PLN model has been shown that the model gives a better fit for count data compared to the negative binomial model. It is evident that the log-normal distribution's tails are heavier than that of the gamma distribution. Not only does the PLN model appeal in fitting to empirical data, but also the lognormal distribution has an attractive theoretical interpretation. The factors could include different needs and different exposures to marketing activity such as advertising,

promotion and word of mouth. If these many independent unobserved factors have proportional but different impacts on each consumer in a multiplicative process, then central limit theorem suggests the geometric mean purchase rate across consumers converges to a log-normal distribution. The previous neglect of the Poisson log normal model in the literature should be reconsidered in future applied work.

2.4.A Trustworthiness Evaluation Method for Software Architectures Based on the Principle of Maximum Entropy (POME) and the Grey Decision-Making Method (GDMM)

Trustworthiness of an entity means that when an entity reaches the target being set, its behaviors and consequences can always be trusted. Similarly, Trustworthy software means that the service provided by the software system always conforms to people's expectations and will be stable in case of any interference. It can be seen that the trustworthiness of software architecture is a subjective feeling which can be related to the trustworthiness attribute for the people, and the trustworthiness attribute can further describe the trustworthiness of software architecture.

The evaluation of software trustworthiness is a new direction in software studies. The trustworthiness attribute model of software architecture belongs to a multi-attribute model. Each trustworthiness attribute has both greyness and fuzziness, which contributes to the trustworthiness. Grey system theory has been successfully applied to many different areas. On the basis of the trustworthiness attribute model of software architecture, the maximum entropy principle and the grey decision-making method are used to evaluate the trustworthiness of software architecture. The weight problem is addressed using the maximum entropy principle. There is a famous theorem about the information entropy—the Principle of Maximum Entropy (POME) which is similar to the principle of entropy in the thermodynamics

2.5.A New Similarity Function for Selecting Neighbors for Each Target Item in Collaborative Filtering

The main idea of CF is to predict the user's interest for the corresponding items via the previous data of users. Traditional collaborative filtering consists of four parts: (1)Construct a matrix about users and items from rating information; (2)Calculate a target user's similarity with other users by similarity

function and select target's neighbors; (3)Predict the ratings of an item, on the basis of target users' neighbors; (4)Recommend top-rated n items.

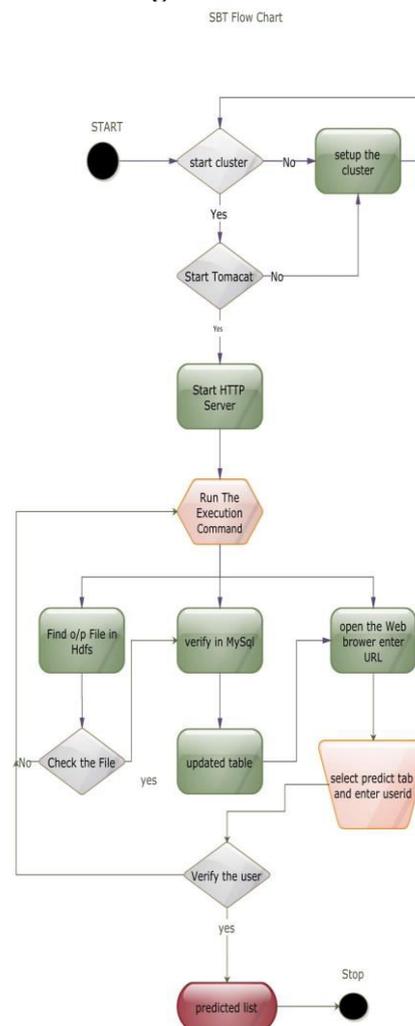
Build User-Rating Matrix- At first, set up the user-rating matrix. In the matrix, the column represents an item's set of scores rated by all users and the row represents a user's set of scores to items which he used to rate

Calculate a target user's similarity with other users - We choose the Pearson to calculate the similarity between two users. Pearson(a, b) shows the similarity between user a and user b. Rai, , Rbi, represent that the rate of user a or user b to item the m in formula (1) is the number of items. Pearson(a, b) is closer to 1, then the similarity between user a and user b is higher.

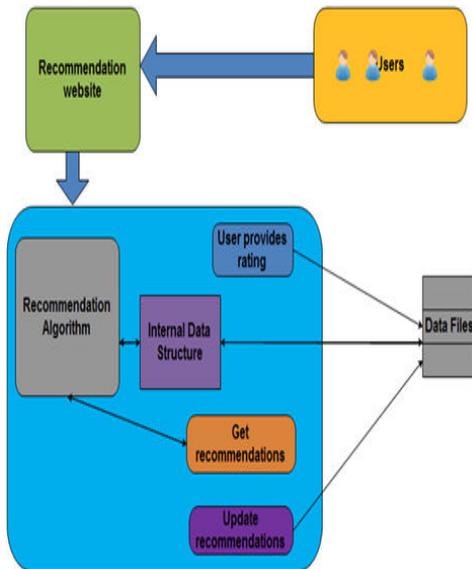
Predict Rates about Target User - Now predict a user's rates to items by selecting a set of users with a high similarity calculated by Pearson with the target user

3. Implementation Details

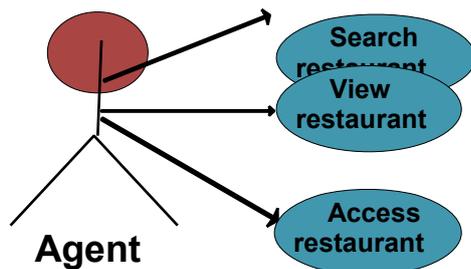
3.1. FlowDiagram



3.2. Architecture Overview



3.3 Use Case Diagram



Step1: User-based CF recommendation considering Structural Balance Theory. Determine user target's "possible friends" through "enemy's enemy is a friend" and "enemy's friend is an enemy" rules, and then select the product items liked by "possible friends" of user target as the recommendation candidates for user target.

Step2: Item-based CF recommendation considering Structural Balance Theory. For the product items liked by target user, determine their "possibly similar product items" based on "enemy's enemy is a friend" rule, and then consider them as recommendation candidates for target user.

Step3: Integration and recommendation. For the recommended product items derived in Step1 and Step2, calculate their integrated recommendation credibility and rank them in descending order for further recommendation.

Step4: For each target product item pro_item_{target} preferred by target user $user_{target}$, we can calculate

its similarity value with other product items pro_item_j ($\in PRO_ITEM$), i.e., $Sim(pro_item_{target}, pro_item_j)$.

4. Conclusion

According to, a novel product item recommendation approach named SBT-Rec for dealing with the specific recommendation situations, when the target user has no similar friends and the product items preferred by target user have no similar product items. On one hand, SBTRec makes full use of the valuable structural balance information hidden in user-product purchase network for precise recommendation, by considering "enemy's enemy is a friend" rule and "enemy's friend is an enemy" rule in Structural Balance Theory; integrates both user-based CF recommendation and item based CF recommendation, so as to improve the recommendation recall. Then we use the prediction rate to check the correctness of SBTRec and use it for further analysis.

5. References

- [1] Y.Tian, Z.Ye, Y.Yan, and M.Sun, "A Practical Model to Predict The Repeat Purchasing Pattern of Consumers in the C2C E-commerce", *Electronic Commerce Research*, vol.15,no.4, pp. 571-583, 2015.
- [2] C.Chui, E.Wang, Y.Fang, and H.Huang, "Understanding Customers' Repeat Purchase Intentions in B2C E-commerce: The Roles of Utilitarian Value, Hedonic Value and Perceived Risk", *Information Systems Journal*, vol.24, no.1, pp. 85-114, 2014.
- [3] K. Choi and Y. Suh, "A New Similarity Function for Selecting Neighbors for Each Target Item in Collaborative Filtering", *Knowledge-Based Systems*, vol.37, pp.146-153, 2013.
- [4] Z.Zheng, H. Ma, M.Lyu, and I. King, "QoS-aware Web Service Recommendation by Collaborative Filtering", *IEEE Transactions on Services Computing*, vol.4, no. 2, pp. 140-152, 2011.
- [5] V.Kavinkumar, R.Reddy, R.Balasubramanian, M.Sridhar, K.Sridharan, and D.Venkataraman, "A Hybrid Approach for Recommendation System with Added Feedback Component", *Proc. Int'l Conf. on Advances in Computing, Communications and Informatics (ICACCI'15)*, pp.745-752, August 2015.
- [6] N. Hummon and P. Doreian, "Some Dynamics of Social Balance Processes: Bringing Heider back into Balance Theory", *Social Networks*, vol. 25, no. 1, pp. 17-49, 2003.
- [7] B. Sarwar, G. Karypis, J. Konstan, et al, "Item-based Collaborative Filtering Recommendation Algorithms", *Proc. 10th International Conference on World Wide Web (WWW'01)*, pp. 285-295, May 2001.