Design of Decision-making Process Based Personalized Recommendation System

Luo Ya[1]

Guizhou University of Finance and Economics, Guiyang, 550001

141756@qq.com

Key words: decision-making process, personalized recommendation, profitability, diversity, precision

Abstract. In-depth development of e-business requires the combination of information technology and principles des affaires. Most existing personalized recommendation systems emphasize the precision of contents, but neglect the profitability of results recommended and the diversity of recommended lists. Proceeding from the decision-making process of consumers and enterprises, this paper simplified the decision-making process of consumers to awareness and satisfaction, and the decision-making process of enterprises to response to recommendation and profit calculation, designed a three-tier personalized recommendation system with the profit maximization for merchandize recommended by enterprise as the objective, and provided logic algorithms for the data supporting tier, the classification tier and the recommendation result tier. Experimental results of the system indicated that good results have been achieved in terms of precision, diversity and profitability of recommendation.

Introduction

The personalized recommendation system is an important tool for enterprises to increase their market share and accrue operation revenue in the e-business market. For a long time, people have taken accurate matching as an objective for the design of personalized recommendation system, believing the higher the matching degree is the more the sales would be.

Logical Principles of Decision-making Based Personalized Recommendation System

In essence, a personalized recommendation system is an auxiliary decision-making means for companies. Companies provide a personalized recommendation system not only to improve the shopping experience of consumers, discover potential customers and increase the sales, but also in its ultimate objective to acquire profit. Hence, in product recommendation, companies should not just recommend the products customers are willing to purchase, but shall also persuade consumers to buy products of large profit margin (Ulrike & Daniel, 2006) [1]. Assuming consumer u has a shopping history H at a company, and after algorithm analysis, the probability of the consumer to purchase commodity G_1 during the shopping is $P(G_1|H) = 0.9$, and the probability to purchase commodity G_2 is $P(G_2|H) = 0.7$; as the selling profits r_1 and r_2 of G_1 and G_2 are 30 yuan and 50 yuan respectively, the company should recommend its commodities in the sequence of $R = \text{Rank}\{r_i * P(G_i|H)\} \quad i = 1,2\ldots n$. In a list of limited recommendation results, display optional

commodities not per the purchase probability of a consumer, but per the profit volume that the commodities may bring to the company. From the perspective of consumers, the existing recommendation system displays possible commodities to them through correlative rule, content filtering or collaborative filtering. Some attributes of the displayed commodities may attract the attention of consumers, enabling them to realize some features meet their demands, thus to motivate the consumer awareness. However, motivation of awareness is simply not enough. More strong stimulation is needed after the formation of awareness, to make consumers feel satisfied about the commodities. Consumers will purchase only when they are aware and satisfied. Therefore, decision-making process based personalized recommendation is the result of combination of company's recommendation and consumer's initiative purchase behavior. The logic process of recommendation system is:
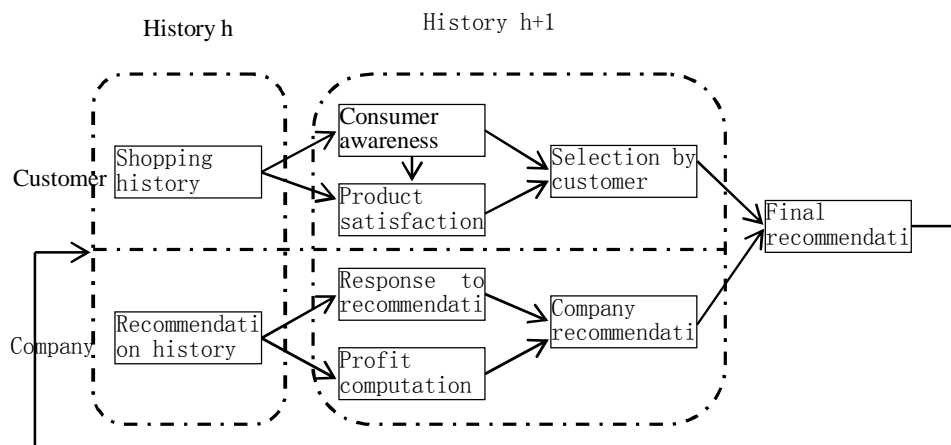

Figure 1 Selling profit based personalized recommendation logic

The final recommendation result shown in Figure 1 is the intersection of optional commodity set of customer and to-be-recommended product set of company. The shopping history of consumers contain two types of data, browsing data and purchase data, where the browsing data of consumer indicate the consumer awareness, the purchase data indicate the consumer satisfaction, and the intersection of these two types of data form an optional commodity set for the consumer; a company's recommendation history include products that have been successfully recommended and products of high profit, and their intersection constitute a to-be-recommended product set of the company.

Design of decision-making process based personalized recommendation system

A personalized recommendation system includes the data supporting tier, the data computation tier and the recommendation result tier.

Tier function of decision-making process based personalized recommendation system

The data supporting tier is the data source of personalized recommendation system. It provides the data computation tier with various types of data required, mainly including the basic information data, consumption preference data and consumption behavior data of consumers, the basic attribute data of products, and the basic resource data and operating condition data of company. These data are the foundation for matching by personalized commodity recommendation information.

The main function of the type analysis tier is to, based on the information of various participants, compute the similarity between participants of the same type, and provide the recommendation basis for the result of recommendation. The customer similarity is used to identify customers with similar consumption preference or consumption experience, to recommend commodities to a

customer via the purchase behavior of other customers; the product similarity is used to identify the similarity of products that have been searched, viewed and purchased by customers, to recommend appropriate commodities; the profit level is used to compute the profit margin in the price range of current recommendation, to provide a decision-making basis for company in determining the product(s) to be recommended.

The recommendation result tier, on the basis of the recommendation computation tier, makes intersection computations for the recommendation list for each participant, and conducts a weighting computation on the recommendation result to determine the recommendation sequence and generate a recommendation list. As the browsing paths of consumers are quite varied, the system will, according to the possible consumption path of consumers, display different recommendations at different positions; collect the response information of recommendations to generate the recommendation feedback and collect data for the next recommendation.

Decision-making process based personalized recommendation data set

Data in commodity purchase process present in structured form. For the sake of computation efficiency, metadata of participants of various types are extracted. Consumer database stores the personal information and the behavior information of consumers. The personal information includes gender, age, education degree, occupation, income, residential district, personal hobbies, etc. of consumers; the behavior information includes the browsing path information and the purchase information. When a consumer visits the homepage of an e-business website, upon login, his/her basic information and historic purchase information will be displayed based on his/her historic purchase information, and similar consumers will be predicted according to these information; if the consumer doesn't login, similar consumers will be searched according his/her browsing path, directory index or keyword search. Consumer information is represented with this 9-tuple.

$$Consumers_{n,i} = (Gend_n, Age_n, Edu_n, Ocup_n, Wage_n, Region_n, Prefer_n, Path_{n,i}, Purch_{n,i})$$

Where, n is the consumer number, i is the number of login times, and $Consumers_{n,i}$ is the information of the nth consumer at his/her i^{th} login.

The product database records product attributes, such as category, function, material, quality, model, purchase price, selling price, sales volume, evaluation, etc. It is represented with the following 8-tuple:

$$Products_{m,t,j} = (Classif_{m,j}, Funct_{m,j}, Mater_{m,j}, Quality_{m,j}, Size_{m,j}, PP_{m,t,j}, SP_{m,t,j}, Sale_{m,t,j}, Remark_{m,t,j})$$

Where, m is the mth product, t is the tth purchase, and j is the jth company; $Products_{m,t,j}$ means the mth product of the jth company sold for the tth time.

The company database records the basic information of the operating condition information of company, where the basic information includes company name, sector, nature of operation, registered capital, number of employees, annual sales volume, qualification certificate, etc.; the operating condition information includes term of e-business operation, annual fee, credit rating, marketing activity, etc. It is represented with the following 11-tuple:

$$Company_j = (Name_j, Indus_j, Type_j, Captial_j, Num\_Emp_j, Ave\_Sales_j, Num\_Quanti_j, Years_j, Fee_j, Reput_j, Mark_{j,t,m})$$

As

e-business operation platforms are characterized by multiple consumers, if a platform operator is the sales company, j=1, and the subscript j for counting the number of company can be ignored; if a platform operator is the integrator of sales companies, $Company_j$ means the information of the jth company. Because a sales company may participate different marketing activities of an e-business platform at different times, $Mark_{j,t,m}$ means the jth company participated the marketing activity with its mth product at the tth purchase of a consumer.

Key algorithm for decision-making process based personalized recommendation

According to the logic frames shown in Figures 1 and 2, a decision-making process based personalized recommendation is the intersection of consumer oriented recommendation aggregate and enterprise oriented recommendation aggregate. Consumer oriented recommendations need to take consumers' awareness and satisfaction into account, and enterprise oriented recommendations need to take the sales profit of enterprises into account, where the former falls under subjective awareness, and is highly subject to external influence, to which, in addition to the purchase history of consumer, display of other consumers and products is also an important factor; the latter is objective data, and can be computed according to the operation ledger and the sales record. In consideration of the complexity of recommendation process, the correlative rule based collaborative filtering is used as the decision-making process based personalized recommendation method.

When a consumer logins an e-business platform, his/her basic information ($Gend_n, Age_n, Edu_n, Ocup_n, Wage_n, Region_n, Prefer_n$) will all be displayed, based on which, the similarity with other consumers can be computed, a consumer with a longer shopping history can be chosen based on the similarity, and commodities that have been purchased by this consumer can be recommended to the current consumer to enable him/her to realize the existence of a commodity; if the consumer does not login the e-business platform, the consumer awareness can be determined according to his/her behavior information ($Path_{n,i}$). The computation method of consumer similarity is to: after the login of a consumer, use the improved Pearson correlation coefficient method.

$$\text{sim}(Cx, Cy) = \frac{\sum_{i=1}^{7}\left[\left(Item_{Cx,i} - \overline{Item_{Cx}}\right) * \left(Item_{Cy,i} - \overline{Item_{Cy}}\right)\right]}{\sqrt{\sum_{i=1}^{7}\left(Item_{Cx,i} - \overline{Item_{Cx}}\right)^2}\sqrt{\sum_{i=1}^{7}\left(Item_{Cy,i} - \overline{Item_{Cy}}\right)^2}} \qquad (1)$$

Formula (1)

Where, Item means the first seven elements of consumer data aggregate. After the consumer similarity is computed, the possible commodity labelcan be computed based on the purchase history of similar consumer(s), and at this moment, the consumer aware commodity is

$$AP = \left\{Purch_{n,i}, Purch_{n+1,i}, Purch_{n+n,i} \ldots \ldots\right\}.$$

If a consumer does not login, the consumer awareness can be determined according to the consumer's behavior data. Consumer's behavior data at an e-business website record the skip and duration of webpage browsing, and the longer a consumer stays at a webpage, the stronger awareness of the consumer on the webpage. Hence, $Lable_{n,i,k}$ is the label for consumer n's i^{th} time browsing the kth webpage, $Time_{n,i,k}$ is the duration of consumer n's i^{th} time browsing the kth

webpage, and the consumer awareness is:

$$\text{Rank}\{Time_{n,i,k}\}; \quad \text{getnum(n,i,k)}; \quad AP = \{Lable_{n,i,k}\}$$

After a consumer is aware of the existence of a commodity, the recommendation system will enter the product database, and determine the consumer satisfaction based on the sales volume and evaluation of product. Set a threshold value, and assume the full mark of commodity evaluation rule is S and the consumer satisfaction is 95%*S if the confidence coefficient is 95%, hence the

threshold value $\frac{\sum_{m=1}^{M}\sum_{j=1}^{J}\sum_{t=1}^{T}Remark_{m,t,j}}{\sum_{m=1}^{M}\sum_{j=1}^{J}\sum_{t=1}^{T}Sale_{m,t,j}} \geq 95\% * S$. Filter product m of company j according to

the threshold value to generate a consumer recommendation aggregate.

The probabilities of the consumer to purchase different commodities in the recommendation aggregate are varied. The purchase probability of a target consumer can be determined with consumer similarity, and the purchase record of a product by the consumer is provided in his/her purchase history, assuming

$$P_{Cx,m} = \begin{cases} 1 & consumer \; Cx \; has \; purchased \; commodity \; m \\ 0 & consumer \; Cx \; hasn't \; purchased \; commodity \; m \end{cases}, \text{ the purchase probability of target}$$

consumer Cy is:

$$P_{Cy,m} = sim(C_x, C_y) * P_{Cx,m} \qquad （2）$$

Formula (2)

When a consumer hasn't logged in, determine the purchase probability with the browsing time based preference method, assume $\text{Max Time}_{n,i,k}$ is the maximum time consumer n has browsed

label k in the ith browsing, $\text{Min Time}_{n,i,k}$ is the minimum time, and $\overline{\sum_{j=1}^{J} Time_{n,i,k}}$ is the average

time consumer n has browsed label k, the purchase probability of consumer n for the product under label k is:

$$P_{n,k} = \frac{\text{Max Time}_{n,i,k} - \text{Min Time}_{n,i,k}}{\sum_{j=1}^{J} Time_{n,i,k}} \qquad （3）$$

Formula (3)

Response to a company's recommendations can be judged from sales volume and commodity evaluation. To avoid the improvement in sales volume and evaluation due to low-price promotion, all commodities shall be recommended except for those with sales volume and evaluation both low. Hence, the algorithm for recommendation response is by ranking,

$\text{Rank}(Sale_{m,t,j}) \cap \text{Rank}(Remark_{m,t,j})$, and by applying the quarter bit method, the front three

fourths commodities shall be recommended. The profit of commodities recommended by a company is the gross revenue minus the operating cost, and the formula is:

$$\text{Profit}_{m,t,j} = (PP_{m,t,j} - SP_{m,t,j}) * Sale_{m,t,j} - Fee_{m,j} - Mark_{j,t,m} \qquad （4）$$

The final commodity recommended by the company is:

$$R = Profit_{j,m} * P_{n,m} \qquad （5）$$

Formula (5)

To sum up the above design thoughts, the formalized algorithm of the decision-making process based personalized recommendation system is:

Input: Input consumer information, product information and company information;

Output: A consumer decides to purchase, and a company sales a product of high profit margin.

Step 1: Compute the consumer awareness, and obtain the commodity aggregate that the consumer has become aware of.

Step 2: Compute the consumer satisfaction, and obtain the consumer satisfied commodity aggregate.

Step 3: Obtain the consumer aware and satisfied commodity aggregate.

Step 4: Compute the purchase probability of the consumer for commodities in the commodity aggregate in Step 3.

Step 5: Compute the commodity aggregate of the company with the highest recommendation response rate.

Step 6: Compute the commodity profit of the company, and obtain a commodity aggregate with a high profit level.

Step 7: Obtain the company's commodity aggregate of high profit under the purchase probability of the consumer.

Output of decision-making process based personalized recommendation result

According to the above design, recommendation results from computation with consumer similarity, rather than product similarity, are different in category, i.e. the commodities recommended to target customers are of different categories, solved the commodity homogenization problem existing in a common recommendation system. These commodities of different categories are ranked by the product of the company's selling profit and the consumer's purchase probability, so those commodities just with high selling profit or purchase probability are filtered out, to ensure the precision and effectiveness of the recommendation list.

Different display positions of recommended commodities on a webpage have different influence levels on consumers' attention. The common recommendation positions are as shown in Figure 3. Suspension points in the diagrams indicate prompts or operation contents to be displayed to consumers per the internet shopping procedure, and the boxes indicate the recommended products. Among the four options, the recommended commodities in (d) occupy the core position of the webpage, and are more likely to attract the attention of consumers, however, suspicion of advertisement should be avoided in design, and pictures shall provide quality and valuable information to the greatest extent.
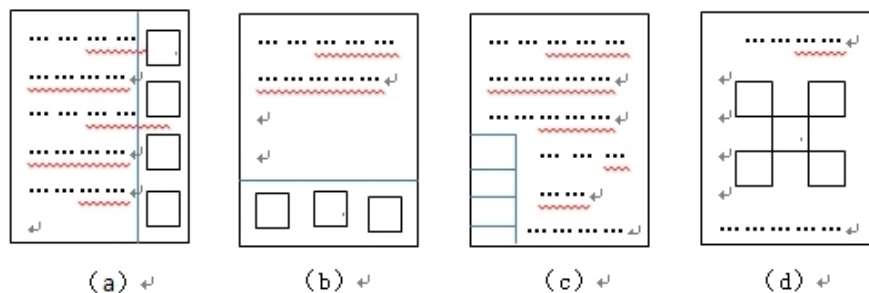
Figure 3 Display position of recommended position

Evaluation information of recommended commodity is an important basis for consumption decision-making, and negative evaluation information has more intensive impact on the judgment of consumers than positive evaluation information. To prevent the amplification effect of negative evaluation information, and enable consumers to objectively understand the negative evaluation information, the possible strategies include: (1) displaying the positive feedback rate, and prompting the reliability analysis result in statistics, e.g. if the positive feedback rate is 98%, the higher than the positive feedback reliability of the commodity is higher than statistical acceptance rate, 95%; (2) displaying the negative feedback rate, prompting the validity analysis result in statistics, e.g. if the negative feedback rate is 2%, the validity of negative feedback contents of the commodity is less than the criterion, 5%. Showing the positive and the negative feedback rates at the same time will produce comparison, improving the effect of positive feedbacks; in the meantime, consumers will be educated to raise their tolerance for negative feedbacks.

Result of design experiments of decision-making process based personalized recommendation system

In this paper, data provided by our cooperative enterprise, www.gogbuy.com, were used as cold boot data, and were realized on JDK8.0 experimental platform in a lab environment of Windows Professional2010 and Microsoft Access 2010. Totally 162 students were convened in three times for experiments. The test parameters for the system designed per this paper include precision, diversity and profit margin of recommended result.

The common precision evaluation criteria include precision, recall rate and normalized discounted cumulative gain (NDCG). Precision was used as the evaluation indicator herein. Assuming the number of test sets of a consumer is N, the recommendation precision of the entire system can be represented by Formula (6), meaning the ratio of the number of non-empty intersection of system recommended commodity set and consumer finally purchased commodity set to the number of total recommended commodities[3].

$$\text{Probit(correct)} = \frac{Count\ (\sum_{n=1}^{N} Suggest_i(product) \cap Purchase_i(product) \neq \emptyset)}{N} \qquad (6)$$

Formula (6)

The common diversity evaluation criterion is coverage, i.e. the number of commodity categories included in a recommendation aggregate. A recommendation is effective only when the diversity is established based on precision. Therefore, the computation formula of diversity is:

$$\text{Probit(coverage)} = \frac{count(\sum_{n=1}^{N} count(suggest_i(product)) > 1)}{N} \times \text{Probit(correct)} \qquad (7)$$

Formula (7)

Different from profit margin, profit level refers to the profit margin rank[4] of a commodity, from commodities finally purchased by a consumer, in the profit margin chart of all commodities in a recommended commodity aggregate, and is represented with the reciprocal of the rank multiplied by the number of total recommended commodities. Its value varies between (0, 1), as shown in Formula (8).

$$\text{Degree(profit)} = \frac{\sum_{i=1}^{N} \frac{Rank(suggest_i(product))}{Count(suggest_i(product))}}{N} \qquad (8)$$

Formula (8)

Table 1 Overall end values of experiments

	Precision	Diversity	Profit level
Total number of samples	162	162	162
Effective sample	157	157	157
Max.	0.997	0.986	0.997
Min.	0.211	0.005	0.028
Average	0.783	0.625	0.752

According to the test results, the precision of the decision-making process based personalized recommendation system was 0.783, the diversity was 0.625, and the profit level was 0.752, therefore the effect of recommendation was good.

Conclusions

On the basis of decision-making process of consumers and enterprises, a recommended system consisting of data supporting tier, classification tier and recommendation result tier was designed in this paper, and the realization of each tier were described with algorithm. The design ideas of the system followed business rules, and were consistent with the shopping psychological feature of consumers. The recommendation list was designed with diversified display methods. Scientific prompts were added for response to recommendations, to neutralize the effect of negative feedbacks. These details facilitated the high precision of the system, with which, the purchase rate of consumers would be high through diversified options, and enterprises would realize the maximum profit. It is innovative in the design of recommendation system.

References

[1] Ulrike Gretzel, Daniel R. Fesenmaier. Persuasion in Recommender Systems[J].   International Journal of Electronic Commerce, 2006:11(2), 81-100

[2] Wang Jun, Zhang Zike. Social Label Information Entropy Based Personalized Recommendation Algorithm [J]. Library and Information Service, 2013, 23:31-35.

[3] Sun Yusheng, Dong Hui. Research on Semantic Network Based Personalized Recommendation System of Digital Library – System Architecture and General Framework [J]. Information Studies: Theory & Application, 2009, 06:63-67.

[4] Hu Xinming, Lui Ruizhi. Research on Metadata Based Library Recommendation System [J]. Journal of Intelligence, 2011, 04:152-156.

[5] He Guihe, Zeng Yitang. Research on Context-awareness Based Personalized Recommendation Model for E-business Platforms [J]. Information Studies: Theory & Application, 2013, 6:81-85.

[6] Zhao Yan, Wang Yamin. Research on Social Label Based Personalized Recommendation Model in P2P Environment [J]. New Technology of Library and Information Service, 2014, 5:50-57.

[7] Chen Zuqin, Ge Jike, Liu Xiwen. Personalized Recommendation Model Combined with Resource Semantics and User Access Path Analysis [J]. Information Studies: Theory & Application, 2014, 09:129-132

[8] Sun Hongfei, Wu Huijuan, Li Chengguang. Research Highlights and Frontal Analysis of Knowledge Diagram Based Personalized Recommendation. Intelligence Science, 2014, 1212-16: