

Personalized Trip Information for E-Tourism Recommendation System Based on Bayes Theorem

Patcharee Srisuwan and Anongnart Srivihok

Department of Computer Science, Faculty of Science, Kasetsart University, Bangkok
10900, Thailand g4764233@ku.ac.th anongnart.s@ku.ac.th

Abstract. This paper presents the personalized recommendation system for e-tourism by using statistic technique base on Bayes Theorem to analyze user behaviors and recommend trips to specific users. The system is evaluated by using Recall, Precision and F-measure. Results demonstrate that it is possible to develop Personalization Recommendation System. Past and recent information of customer's behaviors are used for recommending products and services best fit to relevant customers in the e-Tourism.

Keywords: *E-Commerce, Business analytics, Personalize, Recommendation system, Bayes theorem, E-Tourism*

1. INTRODUCTION

Nowadays, recommendation systems are applied for recent business in e-commerce. The aim of this system is an ability to correctly filter and analyze the customer preferences and behaviors [1] to recommend the right information to specific customers in selecting their interesting products or services.

Many research used the personalization in the recommendation systems [1-10]. The techniques are grouped into 2 types: (1) Content – based filtering and (2) Collaborative filtering [1].

Content – based filtering [11] provides an advice depends on the essence of items more than the other user opinion this method makes the comparison between the features of products or services that system offers with the user profile. If they are matched, the products will be recommended to relevant users [2-4].

Collaborative filtering [11] is the popular technique, it works by considering and comparing the feature of active user and original user who would have the similarly user database. The systems which use this technique include Movie Lens [5].

Many systems [7-10] used Machine learning technique such as learning algorithms to analyze user behaviors such as reinforcement learning [7, 8] or K nearest neighbor [9, 10].

Bayes theorem is a statistical principle for combining prior knowledge of the classes with new evidences gathered from data. The probability principle is the basic of Bayes theorem that is able to measure accurately. According to the prior research [11-14], using Bayesian Theorem is simple in prediction the answers of questions that effect to the similarity of answers than other theories. Bayes theorem is also used for

Please use the following format when citing this chapter:

Srisuwan, P., Srivihok, A., 2007, in IFIP International Federation for Information Processing, Volume 255, Research and Practical Issues of Enterprise Information Systems II Volume 2, eds. L. Xu, Tjoa A., Chaudhry S. (Boston: Springer), pp. 1271-1275.

prediction of classification problem from statistical perspectives. No study has used Bayes theorem for the prediction of user behaviors in personalized recommendation system [12-14].

Currently many travel programs offer the trip information on the websites. It is very time consuming for users in searching the right trip for them. Actually, to develop a system which recommends the right trips to the right user is compulsory. In this system, transactions of user behaviors in selecting trips are stored in the database. The system can predict user interesting trips and recommend the relevant trip when user connects to the system next time.

The objective of this research is to propose the Personalized Recommendation System for e-Tourism (PRSET). This system uses Bayesian Learning in analyzing the historical information on trip selection and recommends relevant trips to the same user personally.

2. PRSET SYSTEM ARCHITECTURE

The basic feature would effect to the user decision. It consists of five trip features. The features include type, day, zone, price and season. The components of the system are the User Interface, User transaction database, Trip database, Bayesian learning and Rank module. Data collection were obtained from 110 internet users, their ages were ranged from 15 to 50 years old. They are about 7,000 transactions of system usages.

2.1 Two parts of System database

1. Trip database: there are 75 relevant trips, each trip have five features which include (1) type, (2) number of day, (3) zone, (4) price and (5) season of trips
2. User transaction database: it includes user behaviors on web usages which are data of user selection on relevant trips, the feature of relevant trips and the selected feature values.

2.2 Procedure

1. The user starts to select that relevant trips or feature of relevant trips from the provided website
2. The system keeps the data usages on Transaction database.
3. Using Bayesian learning to analyze and learn user behaviors. Figure 1 shows PRSET learning model. The probability of trip selection by users is depend on five factors include type, day, zone, price and season

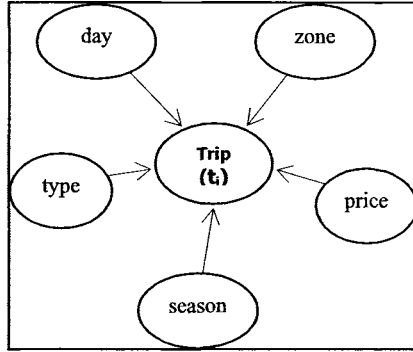


Figure 1. PRSET Model

4. From the 3rd procedure, the system learns important features of the relevant trips from user behaviours. The probability that user selected each features of relevant trips in database would be calculated. When user selected feature of any trips the system used trip feature comparing with another trips. For example, user select feature price is 5,000 bahts and select feature type is the seaside, so the trip which price is 5,000 bahts and type is the seaside must increase trip probability to be selected by the same user.

From PRSET model, probability of trip feature with price is 5,000 bahts and the type is the seaside can be calculated from $P(\text{price} = 5,000 | \text{type} = \text{seaside})$ following Bayes theorem in equation (1). [15]

$$P(h|D) = \frac{P(D|h)P(h)}{P(D)} \quad (1)$$

5. Calculating each important relevant trips sequence value by considering the probability in the 4th procedure together with the weight (in equation 2) of that user select for ranking (in equation 3) the suitable relevant trips mostly as 5 programs offer to user.

$$w_j = \frac{TF_j}{TF} \quad (2)$$

When w_j = the weight of feature j

TF_j = the frequency of user selecting feature j

TF = the frequency of user selecting all features

$$RT_i = \sum_{j=1}^5 P(t_{ij} | f_j) \times w_j \quad (3)$$

When RT_i = the important program sequence value at item i

$P(t_{ij}|f_j)$ = The probability of relevant trips that "i" in feature is "j"

$\sum_{j=1}^5 P(t_{ij}|f_j)$ = the summation of probability of trip that would be

selected when feature is j

The trips from database are ranked according to their RT values calculated by the system. Then the top five trips which have maximum RT values are recommended to users for the next web usages.

3. THE EFFICIENCY MEASUREMENT

The evaluation of system learning and recommendation includes 3 methods (1) recall, (2) precision and (3) F-measure. Recall is the number of trips that system recommend and being selected divided by total number of trips that system recommended. Precision is the number of trips that system recommend and being selected divided by total number of trips being selected. F-measure is the value that show relationship between precision and recall value for recommending. The formula

$$F_1 = \frac{2 \times p \times r}{p + r}$$

4. CONCLUSIONS

From the measurement evaluation, it is shown that PRSET could offer the relevant trips alike with the user requirement. The system precision is 0.70, recall is 0.59 and F-measure is 0.63 which is good. The results demonstrate that it is possible to develop Personalization Recommendation System. By using Bayes theorem, past and recent information of customer's behaviors can be used to recommend products and services best fit to customers in e-Tourism and related area such as e-Commerce.

REFERENCE

1. S.H. Choi, S. Kang, and Y.J. Jeon, Personalized recommendation system based on product specification values, *Expert Systems with Applications*. Volume 31, Number 3, pp.607-616, (2004).
2. M. Balabanovic and Y. Shoham, Fab: Content-Based, Collaborative Recommendation, *Communications of the ACM*. Volume 40, pp.66-72, (1997).
3. M. Pazzani, J. Muramatsu, and D. Billsus, Syskill Webert: Identifying Interesting Web Sites, in *Proc. of The National Conference on Artificial Intelligence* (Portland, 1996), pp.54-61.
4. Q. Li and B.M. Kim, Clustering Approach for Hybrid Recommender System, in *Proc. of The IEEE /WIC International conference on Web Intelligent* (WI'03, 2003), pp.33-39.

5. D. Goldberg, D. Nichols, B.M. Oki, and D. Terry, Using Collaborative Filtering to Weave an Information Tapestry, *Communications of the ACM*. Volume 35, Number 12, pp.61-70, (1992).
6. Y.Y. Shih and D.R. Liu, Hybrid recommendation approaches: Collaborative filtering via valuable content information, in *Proc. of The 38th Hawaii International Conference on System Science, Track 8* (2005), pp.217-.
7. T. Joachims, D. Freitag, and T.M. Mitchell, WebWatcher: A tour guide for the World Wide Web, in *Proc. of The International Joint Conference on Artificial Intelligence* (1997), pp.770-775.
8. A. Srivihok and P. Sukonmanee, Intelligent Agent for e-Tourism: Personalization Travel, in *Proc. of The 14th International World Wide Web Conference Workshop (WWW)* (Chiba: Keio University, Japan, 2005), pp.10-14.
9. W.P. Lee, C.H. Liu, and C.C. Lu, Intelligence agent-based systems for personalized recommendation in internet commerce, *Expert Systems with Applications*. Volume 22, Number 4, pp.275-284, (2002).
10. D. Godoy, S. Schiaffino, and A. Amandi, Interface agents personalizing Web-based tasks, *Cognitive Systems Research*. Volume 5, Number 3, pp.207-222, (2004).
11. P. Srisuwan and A. Srivihok, Using Bayesian Learning for Personalization Recommendation System for e-Tourism, in *Proc. of The 4th International Joint Conference on Computer Science and Software Engineering (JCSSE)* (Thailand, 2007), pp.515-521.
12. A.D. Martin and K.M. Quinn, *Bayesian Learning about Ideal Points of U.S. Supreme Court Justices, 1953-1999*, Center for Statistics and the Social Sciences University of Washington (2001).
13. T-H. Lin, Bayesian Learning in Dual-Wing Harmoniums Applied to Information Retrieval and Genomics, *International Journal of Computer Science & Applications Technomathematics Research Foundation*. Volume 3, Number 3, pp.12-26, (2006).
14. Y. Tsuruoka and T. Chikayama, Estimating Reliability of Contextual Evidences in Decision-List Classifiers under Bayesian Learning, in *Proc. of The 6th NLPRS* (2001), pp.701-707.
15. T.M. Mitchell, *Machine Learning* (McGraw-Hill: New York, 1997).