Information Processing and Management 000 (2016) 1–19

Contents lists available at ScienceDirect

Information Processing and Management

journal homepage: www.elsevier.com/locate/ipm

A rough set-based association rule approach for a recommendation system for online consumers

Shu-hsien Liao*, Hsiao-ko Chang

Department of Management Sciences, Tamkang University, No. 151, Yingjuan Road, Danshuei Dist, New Taipei City 251, Taiwan, ROC

ARTICLE INFO

Article history:
Received 27 November 2015
Revised 3 May 2016
Accepted 7 May 2016
Available online xxx

Keywords:
Data mining
Rough set
Association rule
Rough set association rule
Analytic hierarchy process
Recommendation systems

ABSTRACT

Increasing use of the Internet gives consumers an evolving medium for the purchase of products and services and this use means that the determinants for online consumers' purchasing behaviors are more important. Recommendation systems are decision aids that analyze a customer's prior online purchasing behavior and current product information to find matches for the customer's preferences. Some studies have also shown that sellers can use specifically designed techniques to alter consumer behavior. This study proposes a rough set based association rule approach for customer preference analysis that is developed from analytic hierarchy process (AHP) ordinal data scale processing. The proposed analysis approach generates rough set attribute functions, association rules and their modification mechanism. It also determines patterns and rules for e-commerce platforms and product category recommendations and it determines possible behavioral changes for online consumers.

1. Research background

Most online businesses that are involved in the sales of products/services, such as commercial websites, are aware of the need to acquire knowledge about their online consumers. However, knowledge about online consumers, though available, is not accessible, so it is critical to analyze all of the available knowledge if online users in search for information, products, or services and then highlight potential product promotions and marketing alternatives from online firms. In this regard, recommendation systems are increasingly used by online businesses to suggest options to online consumers (Miao, Yang, Fang, & Goh, 2007). A recommendation system supports users a search for information, products, or services (such as books, movies, music, digital products, Web sites and TV programs) by aggregating and analyzing suggestions from other users, reviews from various authorities and user attributes (Gao, Tang, & Liu, 2015).

As an information technology that supports a personalized service, recommendation systems are widely used by e-commerce practitioners and have become an important research topic in the field of information sciences and decision support systems (Liang et al., 2008; Wu, Kao, Wu, & Huang, 2015). Recommendation systems are decision aids that analyze a customer's prior online behavior and present information on products that match the customer's preferences. By analyzing the consumer's purchase history or communicating with the consumer, recommendation systems employ quantitative and qualitative methods to determine the products that best suit the customer. Most of the current recommendation systems recommend products that have a high probability of being purchased (Bodapati, 2008). They employ content-based filtering (CBF) (Shih, Chen, Chu, & Chen, 2012; Zenebe & Norcio, 2009), collaborative filtering (CF)

* Corresponding author. Fax: +886229453007.
E-mail address: michael@mail.tku.edu.tw (S.-h. Liao).

http://dx.doi.org/10.1016/j.ipm.2016.05.003
0306-4573/

(Herlocker, Konstan, Tervin, & Riedl, 2004; Ortega, Bobadilla, Hernando, & Gutiérrez, 2013), recommendations based on network structure and graph theory (NSGT) (Wang, Dai, & Yuan, 2008), hybrid recommendations (Yin & Peng, 2012), Radio-Frequency Identification (RFID) route recommendation systems (Tsai & Chung, 2012) and other data mining techniques, such as clustering (Kuo, Liao, & Tu, 2005), association rules (Jie, Yong, Wang, & Chu, 2009; Lee & Lee, 2011), rough sets (Su, Wang, Hsiao, & Tseng, 2010) and semantic approaches (Liang et al., 2008). Other studies have determined the effect of recommendation systems on customer's purchasing behavior (Bodapati, 2008; Mandl et al., 2011). These studies argue that the recommendation decision is based not on the probability of a purchase, but rather on the sensitivity of probability of a purchase, as affected by the recommendation action. General practice regards a recommendation system as successful if customers purchase the suggested products (Jiang, Shang, & Liu, 2010).

This personalization of product information is one of the most important factors that effect a customer's product selection and satisfaction in today's competitive and challenging market. A personalized service requires that firms understand their customers and offer goods or services that meet their needs (Aksoy et al., 2011). However, most of the traditional recommendation systems focus on extracting and recommending the common preferences, according to historical data for users. Although common preferences for general users may be a relevant consideration, each individual user also has his/her own personal preferences. He/she may also rely on the domain expert's knowledge to some extent when making decisions (Su et al., 2010). When using traditional recommendation systems, it is often difficult for online consumers to determine whether the items that are presented on a page are actual recommendations or simply the contents of the page displayed indiscriminately to all users.

Currently, personalized recommendation agents are addressing the impersonal nature of integrated recommendations by using technology to assist customers in decision-making and by treating each customer individually (Jiang et al., 2010). However, the key characteristic of e-commerce/business applications is that there is an increasing move towards a customer-centric paradigm, in order to increase competitiveness over other online firms (Liu, Lai, & Lee, 2009). Therefore, consumer preferences and profile building is critical to market segmentation for both online consumers and providers that seek to make personalized recommendations. This study considers that integrate both rough set theory and association rule as a new approach in terms of processing AHP ordinal scale data for developing a recommendation system on electronic commerce.

2. Related research

2.1. Recommendation systems

Since the development of the first recommendation system by Goldberg, Nichols, Oki, and Terry (1992), various recommendation systems and related technologies, such as CBF and CF, have been reported (Herlocker et al., 2004; Zenebe & Norcio, 2009). Of these, user-based CF is considered to be the most successful recommendation technique and is successfully used by many e-commerce systems, such as Amazon.com and Dell.com (Konstan et al., 1997). CF elicits superior preference information from the collaborative transaction logs to assist active users in making a choice from potential items. In other words, the goal of a CF-based recommendation approach is to predict the target item values for active users by learning from a set of users' rating behaviors. It finds a user group that is similar to the target buyer and recommends products that have been rated by users in the reference group that have not yet been viewed by the target buyer.

However, user-based CF has some limitations. It has difficulty in measuring the similarities between users and there is a scalability issue. As the number of customers and products increases, the computation time for algorithms increases exponentially (Hung, 2005). Item-based CF was proposed to overcome the scalability problem by calculating item similarities offline. It is assumed that a user is more likely to purchase items that are similar or related to the items that he/she has already purchased (Deshpande & Karypis, 2004). However, in terms of personalized recommendations, the existing recommender systems suffer from cold-start, first-rater limitations, sparsity and scalability problems. The fusion of rough sets and average-category-rating (FRSA) integrates multiple contents and collaborative information to predict user's preferences according to a FRSA, in order to reduce the gap between the user's preferences and the automated recommendations (Su et al., 2010).

Content-based filtering (CBF) uses content analysis to target items. Target items are described in terms of their attributes, such as color, shape, or material. A user profile is constructed by analyzing responses to questionnaires, ratings for products and navigation history. The recommendation system proposes items that have a correlation with a user's profile. However, a pure CBF system also has limitations, in that users can only receive recommendations that are similar to their earlier experiences and some items, such as music, photographs and multimedia, are difficult to analyze (Cheung, Kwok, Law, & Tsui, 2003). To increase the efficiency of CF and CBF, data-mining techniques that use decision trees, association rules, regression models and Markov chains have been developed to recommend movies and books (Ansari et al., 2000) and to support one-to-one online marketing (Huang, 2012), associative classification (Jiang et al., 2010) and real-time content recommendation in online social communities (Li et al., 2012).

In contrast to the recommendation of products according to the likelihood of purchase, Bodapati (2008) argued that the recommendation decision should also determine a customer's sensitivity to such a recommendation. A model was constructed to measure the role that recommendation systems play in modifying customers' purchase behavior, relative to what the customers would have done without such a recommendation intervention. Although the existing recommendation sys-

Please cite this article as: S.-h. Liao, H.-k. Chang, A rough set-based association rule approach for a recommendation system for online consumers, Information Processing and Management (2016), http://dx.doi.org/10.1016/j.ipm.2016.05.003