Association for Information Systems
AIS Electronic Library (AISeL)

MCIS 2009 Proceedings

Mediterranean Conference on Information Systems (MCIS)

2009

Applying Recommendation Techniques In Conventional Grocery Retailing

Matina Charami
Department of Management Science and Technology, Athens University of Economics and Business, stcharami@aueb.gr

George Lekakos
Department of Management Science and Technology, Athens University of Economics and Business, glekakos@aueb.gr

APPLYING RECOMMENDATION TECHNIQUES IN CONVENTIONAL GROCERY RETAILING

Charami, Matina, Department of Management Science and Technology, Athens University of Economics and Business, 47A Evelpidon and Lefkados Str, 11362 Athens, Greece, stcharami@aueb.gr
Lekakos, George, Department of Management Science and Technology, Athens University of Economics and Business, 47A Evelpidon and Lefkados Str, 11362 Athens, Greece, glekakos@aueb.gr

Abstract

In grocery retailing, promotions and recommendations, derived from traditional data mining techniques, apply uniformly to all customers and not to individual ones, thus failing to meet each customer's personal needs. On the other hand, recommender systems have been widely explored in the field of e-commerce managing to provide targeted personalized recommendations for products and services. Despite the great success of recommender systems in internet retailing, their application in many other fields remains practically unexplored. RFID and pervasive networking technologies now offer the potentials to utilize recommender systems in physical environment. The scope of this paper is to examine the individual characteristics of the new domain along with the applicability of various recommendations techniques. The results indicate the superiority of the e-commerce recommendation techniques against the traditional approaches currently used in grocery retailing.

Keywords: *Recommendation, Recommender Systems, Personalization, Personalized Recommendations, One-To-One Marketing*

1 INTRODUCTION

The vast amount of products and services offered through the Internet, as well as the great diversity in users' tastes has driven the need for personalized systems (Prassas et al., 2001). Personalized systems are able to identify and differentiate the users in order to provide content, services and information that best suit user's individual needs.

Recommender systems form a special category of such personalized systems and aim to predict user's preferences based on her previous behavior. They have experienced a great success and still continue to efficiently apply on various popular websites. Amazon.com is the most familiar and successful example of using recommender systems in order to provide personalized promotions for books, movies, etc. In Pandora.com the user can create her own radio station which will play her favorite kind of music. Whereas, Baynote.com is specialized in websites and content search that best satisfy the user's requirements.

The key idea underlying recommender systems is based on the observation that people very often rely upon opinions and recommendations from friends, family or associates to make choice decisions (Lekakos et al., 2006). They imply preferences either implicitly or explicitly and their main goal is to identify and select among all available goods the new ones that the users will be interested in. As an input, they take product votes/ratings and/or characteristics, and as an output, they generate personalized recommendations or they manage the personalized guidance to items of interest among a maze of possible choices.

Two major approaches have been developed for formulating predictions: collaborative filtering (CF) and content-based filtering (CBF). In collaborative filtering (Breese eta al., 1998, Herlocker et al., 2002) the users provide ratings for the products denoting, in this way, their degree of satisfaction. The users that commonly like some items present similar behaviour and are considered to be neighbors. We call target user the user for whom we want to make recommendations. The target user may receive recommendations from her neighbors for new, previously inexperienced items with a great probability that she will like them, as the neighborhood is characterized by similar behavior and taste.

Content-based filtering makes the assumption that a user's previous preferences are reliable indicators for her future behavior. This approach requires that items are described by features and is usually applied upon text-based documents or in other domains with structured data where feature extraction is possible and applicable (Mooney et al., 2000). For example, in books recommender systems the features used can be the title, the author, the genre and perhaps even a brief summary. Thus, recommendations about new books for a particular user will include books of the same author, close genre and theme, etc.

In addition to the above methods, data mining refers to extraction of implicit but useful information from large volumes of data and can be used to support a wide range of business intelligence applications such as customer profiling, targeted marketing, cross-selling, up-selling, etc. Classification methods, clustering techniques and association analysis can be used to formulate markets according to tastes and to produce predictions inside these markets. All the previously mentioned approaches can either be applied autonomously or be combined to form a hybrid recommender system. Hybrids comprise a very challenging field of research as they aim to combine techniques in such a way that advantages are inherited but limitations are eliminated so as to globally improve the task of generating predictions (Burke, 2002).

2 PROBLEM STATEMENT

The techniques currently used to attract customers and increase purchases in grocery retailing inside physical stores include massive products promotions, special offers and occasional discount coupons. Data mining techniques and especially association rules are widely applied in grocery retailing. The main problem is that these approaches cover only general and by no means personalized trends, thus failing to satisfy the diverse needs of millions of customers. To overcome this major problem, our work aims to examine and incorporate recommender systems inside physical environments. In literature, there are a few references about recommender systems in physical stores, which in fact abstractly address the issue without providing any implementation and experimental evaluation (see Buser, 2007, Raman, et al., 2005).

Our objective is to apply personalized recommendations in the domain of grocery retailing, according to the following scenario: "While entering the supermarket, the customer dispatches a trolley with an embedded display on it or picks-up a PDA provided by the retailer. Either of them will serve as the mean through which the customer will receive recommendations while in the supermarket. Recommendations produced are based on customer's purchasing history, her so-called basket. When the customer completes her transactions, the purchased products are stored in a database updating her personal basket."

Identifying the customer is a primary prerequisite for the scenario described above. This task can be accomplished utilizing the RFID technology, where readers are placed inside the store and customers acquire tags. Alternatively, scanners and barcodes can be also used. The implementation of our system is independent of the technology used for customer identification. Though, we should examine the advantages and disadvantages of each one. On the one hand, barcodes can be applied with a minimum cost using the already provided by supermarkets loyalty cards. The limitation about this approach is that the customer should manually identify himself to the system by scanning her card. On the other hand, RFID technology provides a fully automated system with the cost of acquiring, installing and maintaining a more expensive investment. But eventually, as far as RFID technology is amortized, it definitely outclasses the barcodes.

As previously mentioned a customer's basket is composed by all the products she has ever purchased. So, in this domain the products are binary denoting that a product is or is not bought (Mild et al., 2001). Alternatively, we could use each customer's purchasing frequency for a product or even provide a rating scale as used in e-commerce. Still these are some issues that need to be further examined. Last but not least, features for products of that kind are not yet available and the task of extracting them can be very cumbersome with uncertain outcome. As a result, the luck of these features currently makes the use of content-based filtering extremely complex.

Taking into consideration the needs stated above, this paper has four key contributions:

1. Statement of the need for personalized recommendations in grocery retailing.
2. Implementation of two pure recommendation techniques, collaborative filtering and association rules, plus a hybrid combination of them.
3. Experimental evaluation with real dataset.
4. Proposals for new kinds of applications for the RFID technology.

3 DESIGN AND IMPLEMENTATION

Collaborative filtering is a very successful technique in the field of e-commerce, which has not yet been used in grocery retailing and physical stores. Association analysis is the leading approach for estimating general trends in grocery retailing and it would be very interesting to also apply and test it in the generation of personalized recommendations. Based on Sarwar et al. (2000), we have implemented these two pure algorithms and also combined them in one hybrid. In the remainder of this section we describe the algorithms of collaborative filtering, association rules and the hybrid.

3.1 Collaborative filtering

For the target user α for whom we would like to produce recommendations, collaborative filtering proceeds as follows:

- Compute the similarity between the target user's basket c_α and the baskets c_i of the remaining customers *i* using the Jaccard or Tanimoto coefficient $w(a,i)=\frac{n(c_a \cap c_i)}{n(c_a \cup c_i)}=\frac{n(c_a \cap c_i)}{n(c_a)+n(c_i)-n(c_a \cap c_i)}$,
  where $n(x)$ is the number of products in basket *x*.
- Select the nearest neighbors for the target user α that will best serve in recommendations generation. The selection task can be performed either by obtaining neighbors that are above a predefined threshold (neighborhood-based technique) or by keeping the *k* nearest neighbors (*k* nearest neighbors or *k–NN* technique).
- Generate prediction for the target user α and the product *j* as the sum of the weighted preferences of her neighbors for this product. c_{ij} denotes neighbor's *i* preference for the product *j*. In literature, c_{ij} generally stands for the rating score that user *i* has provided for the product *j.* In shopping basket data, as is our case, c_{ij} is binary. $c_{ij}=1$ means that neighbor *i* has experienced (bought) the product *j,* while $c_{ij}=0$ means that neighbor *i* has not experienced (bought) the product *j* The denominator serves as a normalizing factor for similarities to sum up to unit.

$$p_{a,j} = \frac{\sum_{i=1}^{n} w(a,i) c_{ij}}{\sum_{i=1}^{n} w(a,i)}$$

3.2 Overall association rules (OAR)

As far as the association analysis is concerned, the notion of transactions is equal to the customer's basket. In other words, a customer's basket is considered to be a single transaction.

- Generate rules $X \rightarrow Y$ using the entire dataset. We use the Apriori algorithm and obtain the rules that satisfy a minimum support $s(X \rightarrow Y) = \frac{\sigma(X \cup Y)}{N}$ and a minimum confidence $c(X \rightarrow Y) = \frac{\sigma(X \cup Y)}{\sigma(X)}$
- Apply the left side *X* of the rules to each customer. The left side of a rule can be applied to a customer *iff* it is entirely found in her basket.

When the left side of a rule is applied, recommend the right side *Y* of that rule. The right side *Y* of a rule generates a recommendation *iff* it contains products that are not yet in the customer's basket, for whom the recommendations are made.

3.3 Hybrid: Neighborhood-based association rules (NAR)

Combining the notion of neighborhood from collaborative filtering with association analysis, we implement the hybrid neighborhoodbased association rules. For the target user α for whom we aim to make predictions, we proceed as follows:

- Find target's *k* nearest neighbors.
- Generate rules $X \rightarrow Y$ using target's neighborhood.
- Apply the left side *X* of the rules to target.
- Recommend the right side *Y* of the applied rules to the target.

4 EXPERIMENTAL EVALUATION

The experimental evaluation of the above algorithms consists of data pre-processing, selection of the evaluation method and metrics, as well as the results driven from the experiments.

4.1 Data pre-processing

Our dataset derives from real purchasing transactions of a supermarket. The original dataset relates to 8126 customers, 17263 products and the density of the user-product matrix is just 0.23%. Using the user-product matrix shown in figure 1, density is calculated by the following formula:

$$density = \frac{filled\_in\_cells}{total\_number\_of\_cells}.$$

	Milk	Cereals	~~TV~~	Pasta	Coffee
George	1	1			1
~~John~~			1		
Mary	1			1	1
Tom		1	1	1	

1. $D_1 = 10/20 = 0.5$
   $D_{George} = 3/5 = 0.6$
   $D_{John} = 1/5 = 0.2$ X
   $D_{Mary} = 3/5 = 0.6$
   $D_{Tom} = 3/5 = 0.6$
2. $D_2 = 9/16 = 0.6$
   $D_{Milk} = 2/3 = 0.67$
   $D_{Cereals} = 2/3 = 0.67$
   $D_{TV} = 1/3 = 0.33$ X
   $D_{Pasta} = 2/3 = 0.67$
   $D_{Coffee} = 2/3 = 0.67$
3. $D_3 = 8/12 = 0.67$

Figure 1. Iterative deletion of customers and products.

We have reduced the original dataset in order to accelerate the experiments. Instead of randomly delete data, we introduce an iterative procedure, as shown in figure 1. We begin by calculating the overall density of the the user-product matrix. Similarly, we calculate the density of each row. We delete the rows that have density smaller than the overall density. Then we calculate the overall density again and the density of each column. We delete the columns that have density smaller than the new overall density. This procedure continues iteratively until we reach a desirable dataset size. The intuition behind this approach is that we delete customers with few purchases and products rarely bought. The final dataset consists of 1281 customers, 3993 products, while the density of the matrix is 2%.

4.2 Evaluation method and metrics

We have chosen 10-fold cross validation to be our evaluation method, as the most popular, robust and valid cross-validation technique. We divide our dataset in such a way that we keep 90% of it as training set and the rest 10% for testing. We repeat this procedure 10 times for non-overlapping test sets, so as all data have been used both for training and for testing.

We use precision, recall and F–1 measure to be our evaluation metrics.

- Precision stands for the percentage of recommendations generated that where actually found in the test set. That means the number of correct recommendations.
  $$PRE = \frac{testSet \cap recommendationSet}{recommendationSet}$$
- Recall stands for the percentage of the test set that the algorithm managed to recommend.
  $$REC = \frac{testSet \cap recommendationSet}{testSet}$$
- F–1 measure is a combination of equally weighed precision and recall.
  $$F-1 = \frac{2 \times PRE \times REC}{PRE + REC}$$

4.3 Results

The overall results of the experimental evaluation of all implemented techniques are shown on table 1.

CF					
threshold based			*k–NN*		
all recommendations			all recommendations		
threshold = 0.07			*k* = 10		
PRE	REC	F–1	PRE	REC	F–1
2.38%	2.94%	2.63%	2.57%	28.24%	4.71%
top-*N* recommendations			top-*N* recommendations		
threshold = 0.05			*k* = 60		
PRE	REC	F–1	PRE	REC	F–1
78.24%	13.04%	22.35%	**98.33%**	17.35%	29.50%

OAR		
all recommendations		
support = 0.03 confidence = 0.10		
PRE	REC	F–1
5.11%	3.33%	4.04%

Hybrid NAR		vs.		OAR	
all recommendations					
support = 0.10 confidence = 0.10 *NN*s = 60			support = 0.10 confidence $\geq$ 0.10		
PRE	REC	F–1	PRE	REC	F–1
9.19%	2.55%	3.99%	0.00%	0.00%	0.00%

Table 1. Experimental evaluation results.

For collaborative filtering, we observe that obtaining just the 10 best recommendations produces much better results instead of keeping all the predictions that the algorithm generates. The intuition for this is that in a physical store like a supermarket, customers will be unwilling to process an endless list of all possible recommendations. Instead, they expect to get a smaller number of recommendations that they will like the most. Fewer but more accurate and interesting recommendations are of vital importance as they help customers to build loyalty to the recommender system. In addition, we note that the selection of nearest neighbors is a much better technique than the threshold-based as it achieves better results. Table 1 also shows the results for the overall association rules, where it can be noticed that they are inferior to collaborative filtering. Finally, in the same table we may see a comparison between overall association rules and the hybrid neighborhood-based association rules. It can be easily observed that the later is superior to the former, but still hybrid performs much worse comparing to collaborative filtering.

5 CONCLUSION AND FUTURE WORK

This paper aims to describe a research in-progress, and hence reports the first results of such a work. The results we obtained indicate that collaborative filtering operates efficiently in the field of traditional retailing, in contrast to association analysis. Collaborative filtering's remarkable high precision may indicate an increase in purchases from the new, unseen products that this approach recommends to customers. As far as the association analysis is concerned, we may have to reconsider the approach of rules generation and move to less strong rules. The rationale behind this is based on our observation that very strong rules lead to products already bought and thus produce no recommendations. We aim to further extend our work by examining the behavior of association analysis obtaining only the *N* best predictions, while for the hybrid we aim to use new intelligent methods for combining recommendations at the output. The adoption of other data mining techniques, like clustering, may yield useful results in better combining techniques. Last but not least, it would be very interesting to expand the experiments in the original larger dataset, as well as integrate our system with an RFID reading system.

References

Breese, J. S., Heckerman, D., & Kadie, C. (1998). Empirical analysis of predictive algorithms for collaborative filtering. *Proceedings of the 14th Annual Conference on Uncertainty in Artificial Intelligence*, 43–52.

Burke, R. (2002). Hybrid recommender systems: Survey and experiments. *User Modeling and User-Adapted Interaction*, *12*(4), 331–370.

Buser, D. Ch. (2007). Context-based recommender systems in conventional grocery–an economic analysis. *IEEE Computer Society: Proceedings of the 40th Annual Hawaii International Conference on System Sciences*, 168.

Herlocker J., Konstan J. A., & Riedl J. (2002). An empirical analysis of design choices in neighborhood-based collaborative filtering algorithms. *Information Retrieval*, 5(4), 287–310.

Lekakos G., & Giaglis G. M. (2006). Improving the prediction accuracy of recommendation algorithms: Approaches anchored on human factors. *Interact. Comput*, 18(3), 410–431.

Mild A., & Reutterer T. (2001). Collaborative filtering methods for binary market basket data analysis. *Proceedings of the 6th International Computer Science Conference on Active Media Technology*, pp. 302–313.

J. Mooney R. J., & Roy L. (2000). Content-based book recommending using learning for text categorization. *Proceedings of the 5th ACM conference on Digital libraries*, pp. 195–204.

Prassas G., Pramataris K. C., & Papaemmanouil O. (2001). Dynamic recommendations in internet retailing. *ECIS*.

Sarwar B, Karypis G., Konstan J., & Riedl J. (2000). Analysis of recommendation algorithms for e-commerce, Proceedings of the 2nd ACM conference on Electronic commerce*, pp. 158–167*