대한산업공학회/한국경영과학회 2002 춘계공동학술대회
한국과학기술원(KAIST) 2002년 5월 3일-4일

Collaborative Recommendations Using Adjusted Product Hierarchy : Methodology and Evaluation

Jae Kyeong Kim [1*], Su Kyung Park[2], Yoon Ho Cho [3], and Il Young Choi [4]
1,2,4 School of Business Administration, KyungHee University#1, Hoeki-Dong, Dongdaemoon, Seoul, 130-701, South Korea

3 Department of Internet Information Systems, Dongyang Technical College62-160 Kochuk, Kuro, Seoul 152-714, Korea

Abstract

Today many companies offer millions of products to customers. They are faced with a problem to choose particular products. In response to this problem a new marketing strategy, recommendation has emerged.

Among recommendation technologies collaborative filtering is most preferred. But the performance degrades with the number of customers and products. Namely, collaborative filtering has two major limitations, sparsity and scalability. To overcome these problems we introduced a new recommendation methodology using adjusted product hierarchy, grain.

This methodology focuses on dimensionality reduction to improve recommendation quality and uses a marketer's specific knowledge or experience. In addition, it uses a new measure in the neighborhood formation step which is the most important one in recommendation process.

1. Introduction

Recommender systems are a personalized information filtering technology, used to identify a set of top-N products that will be of interest to a certain user. One of the most successful recommender technologies is collaborative filtering(CF)[Hill, et al., 1997; Konstan, et al., 1997; Resnick, et al., 1994; Shardanand, et al., 1995]. However, CF-based recommender systems have two major limitations. The first is related to sparsity. The second is related to scalability.

This research intends to solve these two limitations by proposing adjusted product hierarchy based methodology. This adjusted product hierarchy is used to improve the quality of recommendations. And we applied to the real world.

2. Backgrounds

The underlying techniques used in today's recommendation systems fall into two distinct categories: content-based filtering and collaborative filtering(CF)[Billsus, et al., 1997].

However, in this paper, only collaborative filtering is considered in our research to make recommendations. And dimensionality reduction using a marketer's knowledge and experience is being stressed from researchers, so we use this technique, especially the product hierarchy representing her/his knowledge.

2.1 Collaborative filtering

Collaborative filtering is the most successful recommendation method to date and used in many of the recommender systems. This recommends products to a target customer based on the opinions of other customers.

This paper applies a collection of algorithms such as traditional data mining, nearest-neighbor collaborative filtering, and dimensionality reduction on two different data sets.

2.2 A marketer's specific knowledge in a recommendation procedure

There have been few research work done in the area of a marketer's specific knowledge. However, in data mining field many researchers have emphasized the usage of this knowledge [Lawrence, et all, 2001; Adomavicius & Tuzhilin, 2001; Berry & Linoff, 1997]. In particular, Adomavicius and Tuzhilin (2001) have shown that an explicit participation of the marketer is required to build more personalized profiles.

The formal use of hierarchies as the most important background knowledge in data mining is introduced by Han, Cai, and Cercone(1993). Lu(1997) also deals with hierarchies as one of the most important background knowledge in the context of data mining.

Brew(1991) and Mellish(1991) have shown that hierarchies are important in knowledge representation and reasoning.

3. Research Methodology

In general, recommendations based on collaborative filtering can be simply described as follows.

Based on this general collaborative

recommendation, we suggest CF-based recommendation methodology is experimented using actual data set of H department store.

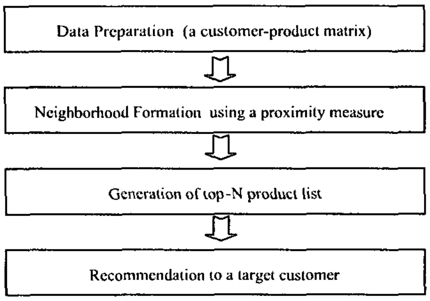


Figure 1 The procedure of collaborative recommendation

3.1 Overall Procedure

In this paper, we divide the entire procedure of CF-based recommendation into three sub-tasks; data representation, neighborhood formation, and recommendation generation.

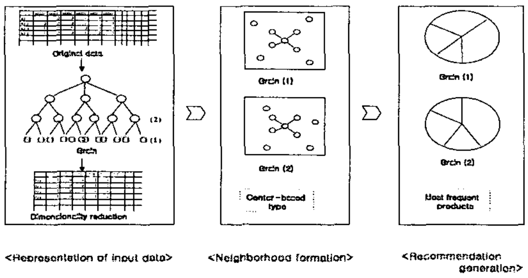


Figure 2 The procedure of collaborative recommendation

3.2 Grain generation

To solve these sparsity and scalability problems we introduce the method of dimensionality reduction using a product hierarchy that is an original product hierarchy in Figure 3.

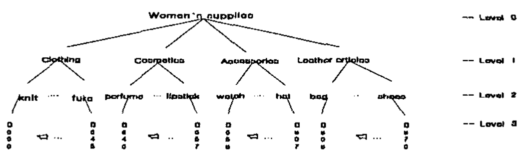


Figure 3 A product hierarchy

With the usage of a product hierarchy, we adopt a "grain", which is the bottom band of a newly adjusted product hierarchy. The concept of "cut" proposed by Adomavicius and Tuzhilin(2001) is a similar to our "grain" concept. However, we generate a grain in a different way by grouping leaf-nodes into an integrated leaf node using a marketer's knowledge or experience. In this paper, four types of grains are presented

3.2.1 Grains at flat-level

In this case, a grain is generated on a particular level of product hierarchy and consists of nodes belonged to that level.

Flat-level 1: In this case we consider non-leaf nodes at the level 2 as leaf-nodes. We term this 'Flat 1' in experiments and it is marked with the shaded box in Figure 4.

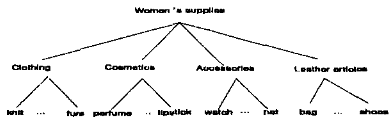


Figure 4 A grain at the level 2

Flat-level 2: In this case all the leaf-nodes of the original product hierarchy are considered as a grain since constructing a product hierarchy itself, leading to the dimensionality reduction. This grain is termed with 'Flat 2' in the experiments and marked with the shaded box in Figure 5.

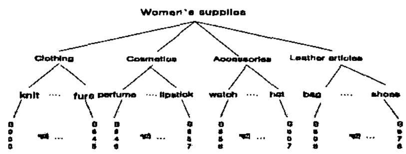


Figure 5 Flat -level 2

3.2.1 Grains at cross-level

Unlike grains at flat-level, this grain is formed using two kinds of levels of product hierarchy. This grain is presented in the shape of cross-leveled band.

Cross-level 3: Cross-leveled grains are generated at the level 3, the leaf-nodes of product hierarchy and at the level 2, the class level of the leaf-nodes. This process is divided into two steps. First, a marketer chooses nodes over a particular threshold value at the level 3 of product hierarchy. Second process is based on a marketer's knowledge or experience and so this criterion depends on her/his knowledge Various shape of each grain can be generated like Figure 6 in the product hierarchy.

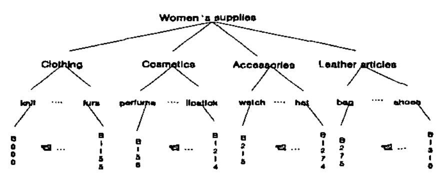


Figure 6 Cross-level 3

Cross-level 4: Whereas the Cross 3 is made the target of all customers, Cross-level 4 is made a particular target group. To generate this grain we first select a particular target customer group and then we analyze the propensity to costumer and generate a grain fitted with this inclination. This grain is also generated by using a marketer's knowledge or experience like Cross-level 3 and termed with 'Cross 4' in the experiments and marked with the shaded box in Figure 7.

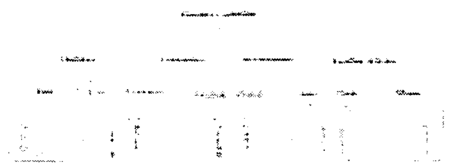


Figure 7 Cross-level 4

3.3 Neighborhood Formation

The goal of neighborhood formation is to find, for each customer u, and ordered list of l customers $N = \{N_1, N_2, \cdots, N_l\}$ such that u (N and sim(u, N_1) is maximum, sim(u, N_2) is the next maximum and so on [Sarwar et al., 2000].

Proximity Measure

There are two types of proximity between two users a and b. One is the Pearson correlation and the other is Cosine measure.

The Pearson correlation: Proximity between two users a and b is measured by computing the Pearson

$$\mathrm{corr}_{ab} = \frac{\sum_i (r_{ai} - \bar{r}_a)(r_{bi} - \bar{r}_b)}{\sqrt{\sum_i (r_{ai} - \bar{r}_a)^2 (r_{bi} - \bar{r}_b)^2}}$$

correlation corr_{ab}, which is given by

Cosine: In this case two customers a and b are thought of as two vectors in the m dimensional product space. The proximity between them is measured by computing the cosine of the angle between the two vectors, which is given by

However, there is a problem in these measures.

$$\cos(\vec{a}, \vec{b}) = \frac{\vec{a} \cdot \vec{b}}{\|\vec{a}\|_2 * \|\vec{b}\|_2}$$

So we introduce a new measure. In this measure proximity between a target customer, a and other customer, b $prox_{ab}$ is measured as follows.

$$prox_{ab} = \frac{P_a \cap P_b}{P_a} \times \frac{P_b - (P_a \cap P_b)}{T - (P_a \cap P_b)}$$

Here P denotes the number of products a customer purchased and T the number of total products.

Neighborhood Types: we use the center-based neighborhood formation. The center-based type forms a neighborhood of size k, for a target customer c, by simply selecting the l nearest other customers.

3.4 Generation of Recommendation list

In this step we have two sub-tasks; choosing top-L classes and producing top-N recommendation list. First, we look into the neighborhood N and then choose top-L classes that are most frequently purchased by neighborhood. Second, we check their purchase data and then perform a frequency count of the products in each class.

4. Performance Evaluation

In this chapter, a case application to the method proposed in chapter 3 is illustrated. This case is based on the real H department store.

4.1 Data Set

We used data from the H department store. We restrict our experiments to the customers who purchased only women's supplies in May 2000 and April 2001 and only consider women's supplies to construct a product hierarchy.

This data set contains purchase information of 50,000 customers on 876 products. We randomly selected enough users to obtain 11,414 purchase-records from the database.

Purchase-record in this context is defined to be a triplet *<customer, product, purchase amount>*. We divided the purchase records into training set and a test set by using the same 60%/40% train/test ratio. This value indicates that we divide the 11,414 purchase data set into 6,848 train cases and 4,566 test cases. The training data was converted into a customer-product matrix R that had 1,833 rows (i.e., 1.833 customers) and 557 columns (i.e., 557 products that were purchased by at least one of the customers). This matrix is used for the grains Flat 1, Flat 2, and Cross 3. However, Cross 4 using this matrix has a different matrix that had 616 rows and 503 columns because of its generation characteristics. The sparsity level of this data set is 0.9888.

4.2 Evaluation metrics

Among several measures there are two metrics used in the information retrieval (IR) community, *recall* and *precision* to evaluate top-N recommendation [Kowalski, G., 1997].

These two metrics are computed as follows.

Recall: For recommendation experiments, we define recall as the ratio of hit set size to the test size

$$\mathrm{Recall} = \frac{\text{size of hit set}}{\text{size of test set}} = \frac{|test \cap top-N|}{|test|}$$

Precision: In the context of the recommendation precision is defined as the ration of hit set size to the top-N set size.

$$Precision = \frac{\text{size of hit set}}{\text{size of top N set}} = \frac{|test \cap top-N|}{|N|}$$

But these two measures are often conflicting in nature [Sarwar, et al., 1999]. Therefore we introduce the standard F1 metric [Yang et al., 1999] that gives equal weight to both of them. This formula is computed as the following.

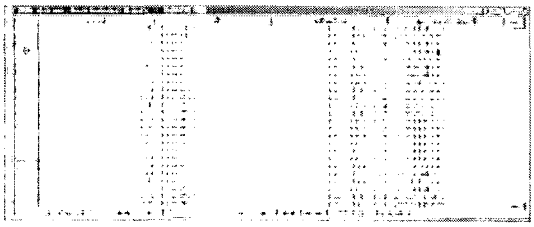
$$F1 = \frac{2 * \text{Recall} * \text{Precision}}{\text{Recall} + \text{Precision}}$$

4.3 Experimental methodology

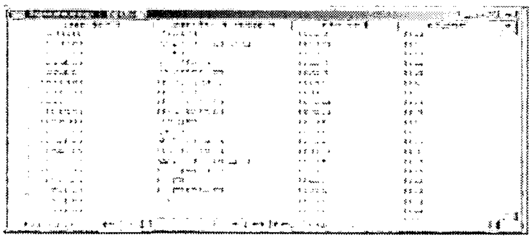
4.3.1 Experimental platform

Data set was converted to be used in the program as follows. Four tables among many ones are presented; a 'transaction' table of all customers, 'taxonomy' table to construct a product hierarchy, a 'rating' table showing purchase amount, and 'a marketer's grain' table to construct 'Cross 3', adjusted product hierarchy. Using these kinds of tables, we performed all the experiments. These tables are presented in Table 2 below.

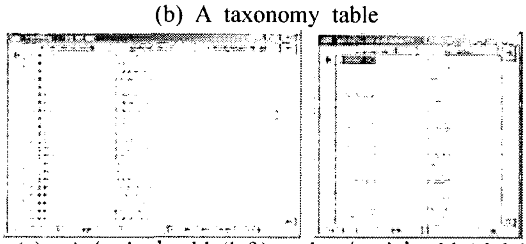
Table 2 Tables of data set



(a)A transaction table



(b) A taxonomy table



(c) A 'rating' table(left) and a 'grain' table(right)

4.3.2 Experimental steps

The ratio of training set is fixed at 60% and running counts of each experiment is 20 times. The evaluation metric, F1, is computed for each customer and then we use the average value as our metric. $Prox_{ab}$ is used to form neighborhoods.

We first evaluate the performance of each grain generated by the process of 'Cross 3' and fix the number of neighbors at 250 and choose the best grain among them.

Here the threshold value of each grain termed with 'marketer' is 20, 30, 50, 100 and 150 respectively for marketer1, 2, 3, 4, and 5. After performing the generation process of 'Cross 3', the leaf-nodes of each grain were 44, 44, 43, 40, and 35 respectively. As we can see from Figure 8, 'marketer4' having the threshold value of 100 produces the highest quality in recommendations.

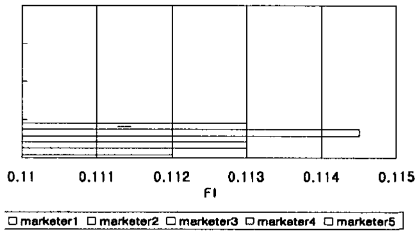
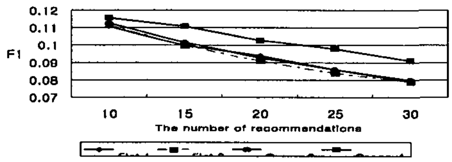


Figure 8 Grains of Cross 3

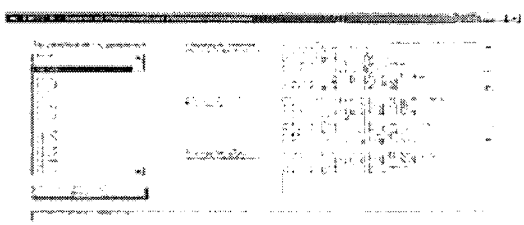
Second, we are interested in determining the number of recommendations which enhances the quality of recommendations. We fixed the size of neighborhood at 250. As we can see from Figure 9(a), the peak of the number of recommendations in each grain is 10. Unlike our expectation, the quality of recommendations becomes worse after 10.

Here we can observe similar shapes of each grain from Figure 9(a). However, the grain, Cross 4 representing a marker's knowledge and experience is better than others in F1 value.

We can also know that what products a customer bought from the recommendation list in Figure 9(b).



(a)



(b)

Figure 9 Results with the number of recommendations

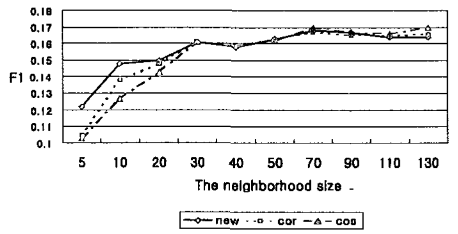
Using the number of recommendations, 10, we perform all the experiments in the following.

4.4 Results

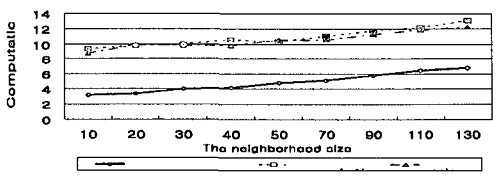
We present several experimental results using our adjusted product hierarchy to compare these performances with that of original product hierarchy at the level 3 termed with 'Flat 2'.

4.4.1 Results with the performance of each measure

we look into the efficiency of $prox_{ab}$ among them. We experimented to investigate the performance of each measure and the computation time taken to find neighbors. We used a grain, Cross 3 and fixed the number of recommendations at 10. The results are below.



(a)



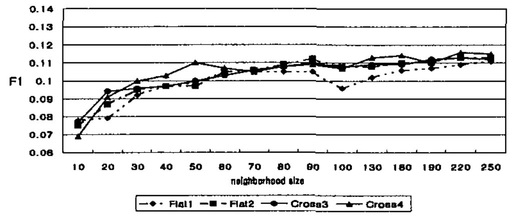
(b)

Figure 10 Results with the performance of each measure

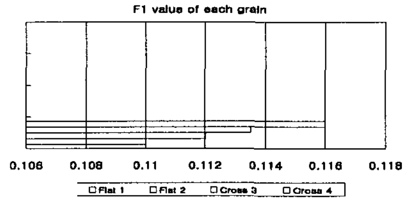
A new measure is better in the performance than other measures and reaches its peak earlier than others for the neighborhood size up to 30. Let's see the computation time of the Figure 10(b). The unit of computation time is a minute and the value of each measure in computation time indicates the average minute it takes to recommend products to all the customers. This shows the efficiency of our new measure and this efficiency is shown in all grains. The computation time of this measure grows with an increasing number of neighborhood size like others but it is two or two and a half times faster than those of others. This also leads to inexpensive experiments and can deal with numerous data sets under the condition of this measure.

4.4.2 Results with the neighborhood size

As the size of neighborhood affects the performance of recommendations, we conducted an experiment with varying neighborhood size. To determine the effectiveness of each grain, we computed the F1 metric. These results are shown in Figure 11.



(a)



(b)

Figure 11 Results with the neighborhood size

Figure 11(a) indicates that the size of neighborhood has impact on the recommendation quality. The F1 value continually increases as we increase the number of neighborhood. In particular, it rapidly increases in the range of 10 ~ 50. However, the increment in F1 value becomes dull and maintains a little constant value after a certain point. In general, each grain's peaks are reached in the range of 220 ~ 250. But each grain shows a little different aspect in reaching its peak. Consequently this experiment indicates that the size of neighborhood affects the quality of recommendations.

As we can see from Figure 11(b), there is a difference among grains. Grains of Cross 3 and 4 have a higher value. In particular, the grain of 'Cross 4' is better than others in F1 value. This tells us that the usage of a marketer's knowledge can improve the quality of recommendations and represents customers' preferences well. It also leads to very reduced dimensional matrix so that it takes less time to perform this experiment. In this case we fixed the size of neighborhood at 220.

4.2.3 Results with the level of sparsity

To overcome sparsity problem, we varied the minimum number of purchased products, that is, frequency. We fixed the number of frequency at 10, 20, and 30 respectively. Here the frequency 10 means that it considers persons who purchased above 10 products customers and this is the same with the frequency 20 and 30. In this experiment we used a grain, Cross 3 as this grain was higher in F1 value and required less time to compute F1 value than other grains. And we fixed the number of recommendations at 10 and used our new measure.

We now present the results of these experiments as follows.

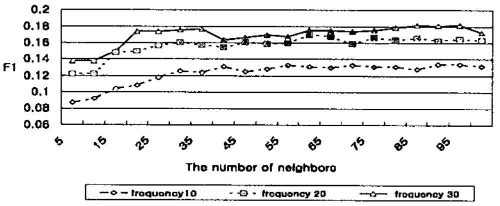


Figure 12 The different level of sparsity

This graph indicates that F1 value increases as we increase the number of frequency and the graph of frequency 30 is better in the quality of recommendations than others. Figure 12 shows that the lower the level of sparsity is, the higher the F1 value is.

Namely, in the lower level of sparsity a recommendation procedure can make better recommendations for a particular customer.

5. Conclusions

In this paper we identified the problems of collaborative filtering techniques and showed how these problems can be addressed through our methodology. Our results show that the performance of adjusted product hierarchy was better than the original hierarchy, especially in a marketer's grain, Cross 4. And also the usage of a new measure $prox_{ab}$ leads to very fast computation time. This measure may process numerous data sets and save much computation time and cost in experiments.

As we can see the above, dimensionality reduction using adjusted product hierarchy is proven to be very effective technique, which leads to improvement of recommendation quality. In particular, a marketer's participation in this procedure affects the recommendation result and produces much better performance in recommendations.

References

Adomavicius, G. & Tuzhilin, A.(2001). Expert-driven validation of rule-based user models in personalization applications . *Data Mining and Knowledge Discovery*, 5(1-2), 33-58.

Billsus, D., & Pazzani, M.(2000). Learning collaborative information filters . In Proceedings of Recommender Systems Workshop. Tech. Report WS-98-08. *AAAI Press.*

Brew, C. Systemic classification and its efficiency. *Computational Linguistics*, 17(4):375-408, 1991.

Cho, Y. H., Kim, J, K., &Kim, S. H.(2001). A methodology for personalized product recommendations based on web usage mining . Working paper, *KAIST.*

Good, N., Schafer, B., Konstan, J., Borchers, A., Sarwar, B., Herlocker, J., and Riedl, J. (1999). Combining Collaborative Filtering With Personal Agents for Better Recommendations . *In Proceedings of the AAAI-99 conference*, pp. 439-446.

Han, J., and Fu, Y. (1994). Dynamic Generation and Refinement of Concept Hierarchies for Knowledge Discovery in Databases . *AAAI 94 Workshop on Knowledge Discovery in Databases*, pp. 157-168.

Hill, W., Stead, L., Rosenstein, M., and Furnas, G. (1995). Recommending and Evaluating Choices in a Virtual Community of Use . *In Proceedings of CHI 95.*

Karypis, G.(2000). Evaluation of item-based top-n recommendation algorithms .
Communications of the ACM

Konstan, J., Miller, B., Maltz, D., Herlocker, J., Gordon, L, & Riedl, J. GroupLens: Applying collaborative filtering to usenet news . *Communications of the ACM.*

Mellish, C. The description identification problem. *Artificial Intelligence*, 52(2):151-167, 1991.

Resnick, P., Iacovou, N., Suchak, M., Bergstrom, P., and Riedl, J. (1994). GroupLens: An Open Architecture for Collaborative Filtering of Netnews . *In Preceedings of CSCW 94*, Chapel Hill, NC.

Resnick, P., and Varian, H. R. (1997). Recommender Systems . *Special issue of Communications of the ACM. 40(3).*

Sarwar, B., Karypis, G., Konstan, J., & Riedl, J.(2000). Analysis of recommendation algorithms for e-commerce . *In Proceeding of ACM E-Commerce 2000 Conference.*

Sarwar, B., Karypis, G., Konstan, J., & Riedl, J.(2001). Item-based collaborative filtering recommendation algorithm . *In Proceeding of the Tenth International World Wide Web Conference.*

Sarwar, B., Karypis, G., Konstan, J., & Riedl, J.(2001). Applicatin of Dimensionality Reduction in Recommender System A Case Study . *In Proceeding of the Tenth International World Wide Web Conference.*