December 2001

Finding Information Just for You: Knowledge Reuse Using Collaborative Filtering Systems

Il Im
New Jersey Institute of Technology

Alexander Hars
University of Southern California

FINDING INFORMATION JUST FOR YOU: KNOWLEDGE REUSE USING COLLABORATIVE FILTERING SYSTEMS

Il Im
College of Computing Sciences
New Jersey Institute of Technology
ilim@homer.njit.edu

Alexander Hars
Information & Operations Management
University of Southern California
hars@usc.edu

Abstract

In today's networked business environment, with the endless increase in available information, relevant information is becoming more and more difficult to find. Collaborative filtering (CF) generates recommendations for users based on others' evaluations. CF has great potential to improve information search and knowledge reuse. Previous studies have mostly focused on the improvement of CF algorithms. Little research has been done on the effect of users and types of product domains on the performance of CF systems. In this study, four factors—product domain, user characteristics, user's search mode, and number of users—that are expected to affect the accuracy of CF systems were identified and investigated. The effects of the four factors were tested using data collected from two experiments in two different product domains: movies and research papers. It was shown that CF systems work better for knowledge-intensive domains than consumer product domains. The accuracy of CF systems is affected by users' search mode and knowledge in a domain. This study demonstrates that CF systems have great potential in information search and customization. It also shows that a successful CF system needs to be designed to handle multiple modes of search, even within a domain and user group.

Keywords: Collaborative filtering, e-commerce, information search, recommendation system.

INTRODUCTION

We are moving toward an economy where competitive advantage is determined by knowledge rather than by access to raw materials and cheap labor. In this environment, organizations are forced to create more knowledge in as short a time as possible. This requires that they improve the reuse of knowledge that already exists within their organization to facilitate knowledge creation among their employees. A problem these companies encounter is in identifying reusable knowledge. Even with the support of classification and search tools, it takes a long time for a user to locate needed knowledge within a large repository.

One of the solutions to this problem is collaborative filtering (CF), a technology that provides users with personalized recommendations based on their preferences or evaluations (Schafer et al. 1999). CF is a relatively young, but not immature, technology (Konstan et al. 1997). CF is used in an increasing number of online companies such as Amazon (www.amazon.com), Levi's (www.levis.com), and Moviecritic (www.moviecritic.com) (Schafer et al. 1999).

While CF has been successfully used by many companies to recommend CDs, books, and movies (Wingfield 1998), only a few companies have applied it to their internal knowledge repositories such as documents. CF has great potential in facilitating knowledge reuse by reducing the time needed to search and evaluate information.

This study investigates CF in knowledge-intensive domains and compares it with the CF in consumer products. Another goal of this study is to answer the question, what are the critical factors that determine the performance (accuracy) of a CF system in a knowledge-intensive domain? Important factors for CF performance are identified, and the effect of these factors on the performance of CF systems is examined.

COLLABORATIVE FILTERING TECHNOLOGY

The concept of collaborative filtering was enunciated by Goldberg et al. (1992), who applied the technology to information retrieval. Collaborative filtering generates recommendations for users based on others' evaluations (Miller et al. 1997). Without a CF system, evaluations on information items must be transferred by word-of-mouth, which is haphazard and inaccurate. By institutionalizing evaluations, CF makes it possible to transfer consistent and precise evaluations among larger numbers of people.

CF Algorithms

A CF system needs a certain number of initial evaluations from a user in order to be able to generate recommendations for her/him. The system stores and then compares these evaluations with those of other users to identify those who have similar evaluations (preferences). The degree of similarity—termed ***similarity index*** in this study—between the current user and each of the other users is calculated to identify people with similar preferences. After a set of people with similar preferences—termed a ***reference group*** in this study—is identified, the system looks up the items that have not been seen by the user who is to get recommendations. The system calculates the user's expected (predicted) evaluation score for each item based on the evaluations of the people in the reference group. Then, it recommends the items with the highest predicted evaluation scores. This overall recommendation process is similar for most CF applications although the detailed algorithms differ with each application.

Previous Research on CF Systems

One of the research streams of past CF studies is the algorithms used in CF systems. The objective of this stream of research is to improve the CF algorithms. Some of the studies compared different algorithms (Breese et al. 1998; Shardanand and Maes 1995), sought variations of algorithms to improve performance (Herlocker et al. 1999), and investigated combinations of CF with other methods (Ansari et al. 2000; Good et al. 1999).

A second stream of research has focused on the application and use of CF. The studies in this stream investigated the applications of CF in e-mail messages (Goldberg et al. 1992), music (Shardanand and Maes 1995), movies (Greening ????), Web pages (Rucker and Polanco 1997), Usenet messages (Konstan et al. 1997; Miller et al. 1997), Internet resources (Terveen et al. 1997), and TV programs (Podberezniak 1998).

Most of the current CF tools are used for consumer products such as CDs, movies, and books, but rarely for knowledge-intensive items. Here, knowledge-related items are defined as codified knowledge objects that are results of or inputs to intellectual activities that yield economic or technological value for organizations. Documents, engineering drawings, and news are examples of knowledge-related items.

Since CF systems generate recommendations based on other users' evaluations, the accuracy of the recommendation is affected by the characteristics of the evaluations: mean, variance, how they are distributed or how the users are clustered. The characteristics of the evaluations would differ from product (domain) to product. Previous research has not sufficiently explored the effect of product type on the performance of CF systems. Another area that has not been adequately investigated by previous studies is the user's behavioral aspect of CF—the effect of the user's behavior on the performance of CF systems.

RESEARCH ISSUES AND HYPOTHESES DEVELOPMENT

One of the most important issues in CF research is how do we improve the effectiveness (or performance) of the system? In the case of CF, the effectiveness of the system is determined by the accuracy of the recommendation. The system characteristics—type of similarity index, reference group selection methods, and other methods for generating recommendations—are important factors that influence the accuracy of recommendations.

Many factors, other than system characteristics, affect the accuracy of CF systems. The number of users is an important factor because as the number of users increases, it will be easier to find people who have similar preferences. The characteristics of users will also be an important factor that determines the accuracy of recommendations. For example, evaluations by users with a lot of experience with the given domain will differ from those by novice users in terms of accuracy and consistency. Even for users with the same level of experience, the evaluations will vary depending on the situation and purpose of use.

Critical Mass of Users and Evaluations

Research on human cognitive processes implies that people have different mental models or cognitive maps (Montazemi and Conrath 1986; Santhanam and Sein 1994), so an increase in the number of users increases the probability of finding people with similar cognitive maps. Shardanand and Maes (1995) found that users of a CF system commented more and more positively about the generated recommendations as the number of users increased, which implies the accuracy of CF systems increases as the number of users increases. The minimum number of users required for a certain level of recommendation accuracy is termed ***critical mass***. The relationship between the number of users and performance will differ depending on the given circumstances. Miller et al. (1997) showed that the characteristics of evaluations (e.g., correlations of the evaluations across users) were dependent on the type of products. This implies that the critical mass in different domains would be different.

Two hypotheses can be formulated regarding the pattern of accuracy increase. First, the accuracy will increase as the number of users increases. Adding more users will improve the chance to find people with similar preferences unless all users have exactly same preferences. Second, the pattern of accuracy increase will vary depending on the given circumstances including user characteristics, product characteristics, and the purpose of information search.

H1a. The performance of a CF system increases as the total number of users increases.

H1b. The pattern of increase will differ depending on the type of product, user characteristics, and the search mode of the users.

Mode of Search

Even for the same CF system, performance will vary according to usage behavior and the current tasks of users. Studies have categorized search behavior in various ways. Vandenbosch and Huff (1997) used two modes of search: *scanning* and *focused.* If the underlying need driving the information search is, "What is the answer to this question?," the person is performing a focused search. Scanning is the behavior of people when they are browsing through data in order to understand trends or sharpen their general understanding of the business (Vandenbosch and Huff 1997). El Sawy (1985) also categorized managers' information retrieval behaviors into *scanning* and *problemistic search*. A problemistic search is a search stimulated by a problem and directed toward finding a solution to the problem while scanning is not directed toward any particular problem. In this study, the labels **scanning** and **problemistic search** are used to highlight the difference between the two searches—with and without specific problems.

Studies indicate that users have different mindsets when they are in different search modes (El Sawy 1985; Vandenbosch and Huff 1997), which implies that users may have different criteria for information evaluation. In problemistic search mode, users already have very specific ideas on what they are looking for while users in scanning mode do not. There is less chance in a problemistic search than in a scanning to find enough people with similar problems. Therefore, CF would be more effective for scanning than for a problemistic search, especially with small number of users. However, an earlier study (Miller et al. 1997) has shown that a CF system is more effective for items that have more highly correlated users. In problemistic search mode, users with common problems will have very high correlations. This implies that CF systems would be more effective for a problemistic search mode, where there are highly correlated users.

The conflicting predictions may be resolved by critical mass. In scanning mode, users evaluate items based on potential usefulness. In problemistic search mode, users rate the items based on usefulness for the current task (hit or miss). Therefore, potential usefulness will have less variability across users than usefulness for current task. Less variability in evaluations means fewer required users (critical mass) for the same performance level. With a small number of users, therefore, a CF system will perform better when the users are in scanning mode. If there are a sufficient number of users, even the users in problemistic search mode can be classified into groups that are large enough but with small variations within groups. The above discussion can be summarized in the following hypotheses:

> **H2a**. With a small number of users, the performance of a CF system is better for the users in scanning mode than in problemistic search mode. (Scanning mode needs less critical mass.)
>
> **H2b**. The performance of a CF system for the problemistic search mode increases faster than for the scanning mode as the number of users increases.

User Characteristics

The performance of CF systems is also be affected by the characteristics of users. If CF is applied to a knowledge intensive domain, the knowledge of users will be a critical factor that affects the accuracy of the system. A number of studies have investigated the human mental structure and applied it to information system design. A ***cognitive map*** is a representation of the relationships that are perceived to exist among the elements of a given environment (Daniels et al. 1995; Montazemi and Conrath 1986). Studies show that the cognitive mapping technique is useful for information requirement analysis (Montazemi and Conrath 1986). It was also shown that a person's mental model affects the success of learning (Santhanam and Sein 1994). Another important theory on the human cognitive process is ***categorization theory*** (Dutton and Jackson 1987). A critical assertion of this theory is that people form cognitive categories based on their observations of the features or attributes of objects. Cognitive categories comprise objects with similar perceived attributes and they reflect the perceived structure of the objects in the environment.

Cognitive mapping and categorization theories imply that each person forms his/her own mental model of artifacts, and that this mental model affects his/her categorization and evaluation of new objects. The people categorized by CF into a same group may have similar mental maps or concept schemas, at least in the given problem domain, because people's evaluations of information items partly reflect their evaluations of features and attributes of the objects.

Studies show that experts are more selective in the information they acquire and more flexible in the manner in which they search for information (Johnson 1988). Spence and Brucks (1997) find that experts make more accurate and tightly clustered judgments than do novices. Shanteau (1988) shows that more knowledgeable people agree more than novices on what information is important. Experts appear to possess comparatively richer schemata for ascribing meaning to given information (Carter et al. 1988; Schenk et al. 1998).

In summary, previous studies show that experts (experienced people) are better at judging the value of information and have higher cohesiveness across people and a more consistent evaluation scheme or mental model across items. These characteristics of experienced people leads to better performance of CF systems for them because of their precise and consistent evaluations.

> **H3**. The system with more experienced users performs better.

EMPIRICAL STUDY

In order to test the hypotheses developed, data were collected from two different domains—movies and research papers—and analyzed. The accuracy of CF systems was calculated using simulation, which is the most widely used method in CF research (Ansari et al. 2000; Breese et al. 1998; Herlocker et al. 1999).

Selection of Domains and Similarity Measures

Among consumer products, movies form one of the best domains for this study because the evaluation of an individual toward a specific movie item is mostly determined by his/her preferences. Movies have also been widely used in CF research (Ansari et al. 2000; Breese et al. 1998; Herlocker et al. 1999). A research paper is a typical knowledge-intensive domain, used mainly for solving a problem or carrying out a certain task. An individual's evaluation of a specific item (paper) depends on the task or job of the person evaluating the item. In this study, movies and research papers were selected as a consumer product domain and a knowledge-intensive product domain., respectively.

Even though comparing different recommendation methods is not a goal of this study, selecting the type of similarity index to be used is a very important issue, because similarity index is one of the critical factors that determine the accuracy of CF systems. Previous studies (Breese et al. 1998; Herlocker et al. 1999; Shardanand and Maes 1995) have shown that correlation is one of the best similarity indices for CF. Therefore, a correlation coefficient was used as the similarity index in this study.

Experiments

Two experiments were conducted in this study—one for movies and the other for research papers. The systems for the two experiments were developed on top of an existing web application using Microsoft Access, Oracle, and Inprise's Delphi. The system for the first experiment contained about 150 movies from various genres. The subjects were recruited from undergraduate classes at two major universities in California. The system for the second experiment contained abstracts of about 2,000 academic articles from the recent issues (1991–2000) of five leading IS journals: *Communications of the ACM, Information Systems Research, Journal of MIS, MIS Quarterly,* and *Management Science*.

People in academia are good subject candidates for the second experiment. Academics have in-depth knowledge in a subject area, they often share common research interests, and most of them have specific research questions. In this study, therefore, academics in the IS field were selected as the subject pool. Soliciting e-mails were sent to the IS faculties whose e-mail addresses are listed in the IS Faculty Directory of ISWorld (www.isworld.org). Subjects were asked to evaluate two things per paper: ***overall usefulness/relevance of the paper for general IS research*** and ***usefulness/relevance of the paper for the user's specific research project***. The former is the evaluation of scanning mode and the latter is that of problemistic search. Once the subject finishes evaluating 10 papers, five recommended papers for each search mode (scanning and problemistic search) were given to the subject. To get better recommendations, the subjects could evaluate more than 10 papers. Actually, 47 subjects evaluated more than 10 papers. At the end of the session, each subject was asked to fill out a questionnaire designed to elicit the subject's background information and his/her perceived accuracy and ease of use. The second experiment lasted for about six weeks and, on average, 10 to 15 people visited the experimental web page every day.

Performance Measures

When the performance of CF systems are investigated, the most commonly used accuracy measures are mean absolute deviation (MAD) (Ansari et al. 1998; Good et al. 1999; Sarwar et al. 1998; Shardanand and Maes 1995), mean squared error (MSE))Miller et al. 1997), root mean squared error (Sarwar et al. 1998), and correlation between actual and predicted evaluations (Hill et al. 1995). These measures have some limitations when used as a performance measures for CF. First, they cannot be used to compare domains with different evaluations scales; for example, 1-5 scale or 1-7 scale. Second, for more accurate measure, the measures should be adjusted by average and variance.

The use of a rank-based performance measure can eliminate the limitations of these measures. Rank-based measure is obtained by calculating the sum of rank differences between actual and predicted evaluations.

Performance (accuracy) = Estimation error by average evaluations of everybody – Estimation error by CF

Estimation error by average evaluations of everybody and CF = $$\frac{\sum_{i=1}^{n}\sum_{j=1}^{m}\left| r(a_{ij}) - r(e_{ij}) \right|}{nm}$$

where $r(a_{ij})$ = rank of actual evaluation of user *i* on item *j*
$r(e_{ij})$ = rank of estimated evaluation (by CF or average of everybody) of user *i* on item *j*
n = number of users in the holdout set
m = number of items in the holdout set

If the system did not have the CF feature, the average ratings would be used to generate recommendations. In order for a CF system to work, the estimated evaluation by CF should be more accurate than by average ratings. Therefore, the average of everybody becomes a sort of baseline when the accuracy of CF is measured. A positive value of rank-based measure means that CF is more accurate than average of everybody. For example, a rank-based measure of 0.5 means that CF outperformed the average of everybody by 0.5 in ranks.

Simulation

For accuracy calculations, a simulation method is used in this study. The simulation in this study simulates recommendation generations with different sets of users and items. The evaluation data are divided into two sets: ***holdout*** sets and ***calibration*** sets as shown in Figure 1. The estimated evaluations for the items in the holdout set are calculated using only the evaluations in the calibration set. Since we know the actual evaluations in the holdout set, the estimation error (= actual evaluation – predicted evaluation) for the holdout set can be calculated. It is a common method used in time series analysis (Griffiths et al. 1993) and other CF studies (Ansari et al. 2000; Breese et al. 1998; Herlocker et al. 1999) to calculate the accuracy of estimations.

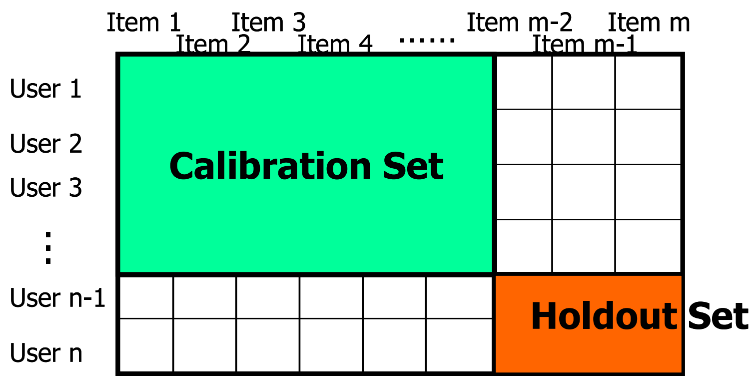


Figure 1. Holdout Set and Calibration Set in the Simulation

One issue here is that the accuracy of the recommendation may depend on who is in the holdout set (and calibration set) and which item is in the holdout set (and the calibration set). The ideal case would be trying all possible combinations of users and items. However, the number of possible combinations is virtually infinite. Therefore, a certain number of iterations with randomly selected combinations of users/items are carried out in the simulation for this study. From pilot simulations, the number of iterations was set at the point where the variance of output is stabilized.

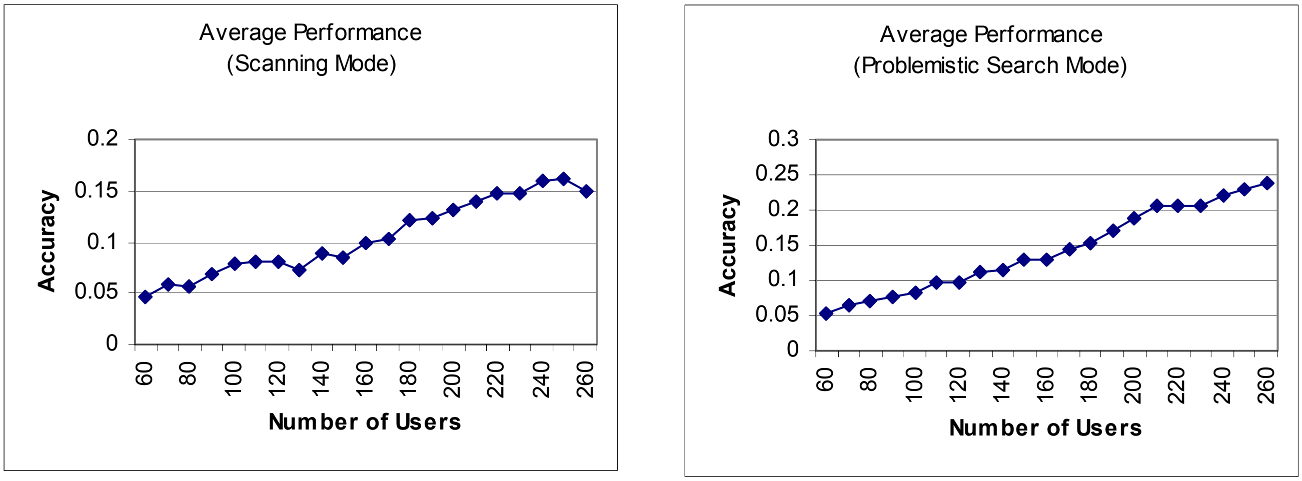
RESULTS AND DISCUSSIONS

The total number of subjects participating in the first experiment was 159. The total number of evaluations was 1,809. The total number of soliciting e-mails sent out for the second experiment was approximately 4,200. About 480 people visited the experimental site and 259 people participated in the experiment. The total number of evaluations was 3,634 or 1,817 for each search mode (scanning and problemistic search).

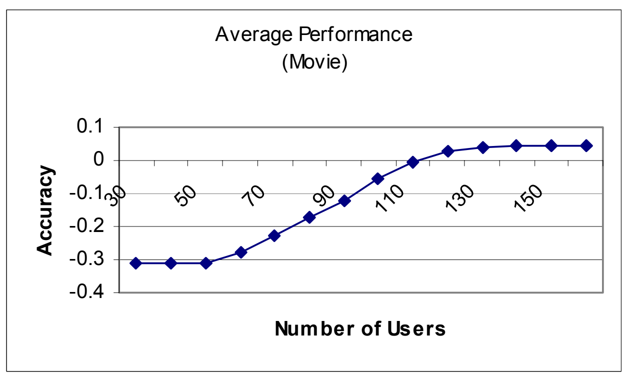
Number of Users and the Performance of CF Systems

In order to test Hypothesis 1, simulations for different numbers of users were conducted and the recommendation accuracies were collected. Figure 2 shows the simulation results of research papers and movies with holdout sample size 4. It is obvious from the figure that the performance of a CF system, as hypothesized, increases as the number of users increases.

More rigorous tests were conducted to investigate whether the increase is statistically significant. A parametric method cannot be used because it is not certain if the probabilistic distribution of the data is normal (Griffiths et al. 1993). One way of testing whether the increase is statistically significant would be to check whether the data are significantly deviated from the non-increasing line (horizontal line). The logic is similar to what is used in the t-test on the coefficients of independent variables in regression analysis.



a. Research Papers (Holdout Sample Size = 4)



b. Movies (Holdout Sample Size = 4)

Figure 2. Number of Users and CF Performance

A Wilcoxon Signed Rank Test (Conover 1980) was conducted to test whether the accuracy increase is statistically significant. The results are summarized in Table 1. From the tests, it can be concluded that the accuracy of CF, regardless of the domain, increases significantly as the number of users increases. The indirect accuracy measures—perceived accuracy of recommendation—also improved as the number of users increased, although it was not statistically significant.

Table 1. Test Results of Pattern of Performance Increase (Research Papers)

Domain	Willcoxon Signed Rank Test Z value (Horizontal Line vs. Straight Line)
Research Papers (Scanning Mode)	-3.84**
Research Papers (Problemistic Search Mode)	-3.88**
Movies	-3.24**

**Significant at α = 0.01 level

It is interesting that the pattern of increase in research papers differs from that in movies. The accuracy of CF for research papers increases almost linearly while the accuracy of CF for movies increases following an S-shape curve. It implies that for consumer products such as movies, there is a "take-off point" at which accuracy increases dramatically, and then a saturation state is reached with a small number of users. On the contrary, knowledge-intensive products such as research papers have a relatively steadily increasing accuracy.

Research papers have higher levels of accuracy, both rank-based measure and MSE, than do movies with the same number of users. This means that there is a greater opportunity for CF to provide recommendations for knowledge-intensive products than for consumer products.

Mode of search and the Performance of CF Systems

In order to test hypothesis 2, recommendation accuracies of the two search modes—scanning and problemistic search—were compared. The accuracies of scanning and problemistic search modes with a small number of users are similar. However, as the number of users increases, the accuracy of the problemistic search mode improves faster than in the scanning mode. This implies that CF systems would be more effective for those users who have a specific purpose for the search.

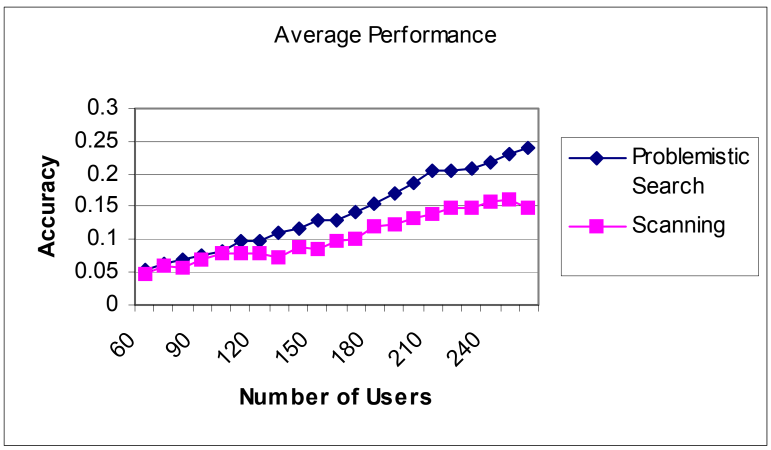


Figure 3. Search Mode and Recommendation Accuracy (Holdout Sample Size = 4)

Characteristics of Users and the Performance of CF Systems

Among the 259 respondents of the second experiment, a total of 68 people completed the post-experiment survey. "Years in the IS field" was used as the measure for experience of users in a general IS field, and "years in a specific project" was also used as a surrogate of the experience of users in the specific area. The subjects were divided into three sub-groups depending on the number of years in the IS field and in a specific project. The performance of the CF system for each group was simulated by iteratively calculating the accuracy of the system for the people in the group.

Table 2 shows the summary of simulation results for the sub-groups. In order to examine whether the differences are statistically significant, Kendall's W-test was conducted. Kendall's W-test is a non-parametric method to test mean difference across multiple related groups (Conover 1980). The results show that the accuracies of CF in both scanning mode and problemistic search mode for the three groups are statistically significant.

Table 2. Simulation Results for User Groups with Different Experience

Years in IS Field	Accuracy	Kendall's W
Group1 (0-8 year)	0.158	0.418*
Group2 (9-15 years)	0.145	
Group3 (16 years and over)	0.149	
Years in the Project	**Accuracy**	**Kendall's W**
Group1 (0-1.5 year)	0.213	0.510**
Group2 (2-3 years)	0.229	
Group3 (4 years and over)	0.225	

*Significant at $\alpha = 0.1$ level ** Significant at $\alpha = 0.05$ level

Interestingly, the accuracies of the two search modes have reverse patterns. The accuracy is lowest in the group with moderate experience in the scanning mode. However, the accuracy of the same group is highest in the problemistic search mode. The reversed patterns of the two modes may be due to the differences of the evaluations in the two modes.

As shown in an earlier study (Carter et al. 1980), users may develop more dimensions (criteria) of evaluation as they gain more experience in the area. It seems that this addition of dimension may affect the accuracies of the two modes differently. More research is needed to investigate what happens as users gain experience in the area.

The hypothesis tests in this study are summarized in Table 3.

Table 3. Summary of Hypothesis Test Results

	Hypotheses	Test Result
H1a.	The performance of a CF system increases as the total number of users increases.	✓
H1b.	The pattern of increase will differ depending on the type of product and search mode of users.	✓
H2a.	With small number of users, the performance of a CF system is better for the users in scanning mode than problemistic search mode. (Scanning mode has less critical mass.)	✓
H2b.	The performance of a CF system for problemistic search mode increases faster than for scanning mode as the number of users increases.	✓
H3.	The system with more experienced users performs better.	✕

✓ - Supported ✕- Not Supported

CONCLUSIONS

This study investigated behavioral aspects of CF systems, which are critical factors in determining the accuracy of CF systems but have been neglected in previous studies. CF systems for a consumer product domain (movie) and a knowledge-intensive domain (research paper) were investigated. Hypotheses were developed from previous studies on CF and related areas and two experiments were conducted to test the hypotheses.

This study has some limitations due to the sample sizes and research methods. First, only two different types of products— movies and research papers—were examined in the study. Second, the small sample size is a limitation. Finally, the evaluation scale of this research was the Likert scale (1 to 13 for movies and 0 to 6 for research papers). Other types of measures and indirect measures—purchase, page visits, and duration of visit—need to be investigated.

Despite of the limitations, this study provides a better understanding about the behavioral aspects of CF systems and valuable implications in applying CF to knowledge-intensive domains. The analyses of the data shows that the accuracy of CF is different for different domains. The accuracy of CF systems increases as the number of users increases. However, the patterns of the accuracy increase differ for different domains. The accuracy of CF for knowledge-intensive domain was much better than it was for consumer product domain. When CF is applied to a knowledge intensive domain, the accuracy of the CF system is also affected by the search mode of users. CF seems to work better when users are looking for specific information than when they are searching for information of general use. Users' experience and knowledge in the domain appear to be the most critical factors influencing the accuracy of the system. In this study the accuracy was higher for users with moderate experience than for both those with little or those with a lot of experience.

References

Ansari, A., Essengaier, S., and Kohli, R. "Internet Recommender System," *Journal of Marketing Research*, August 2000.

Breese, J. S., Heckerman, D., and Kadie, C. "Empirical Analysis of Predictive Algorithms for Collaborative Filtering," in *Proceedings of the Fourteenth Conference on Uncertainty in Artificial Intelligence*, Madison, WI, July 1998.

Carter, K., Cushing, K., Sabers, D., Stein, P., and Berliner, D. "Expert-Novice Differences in Perceiving and Processing Visual Classroom Information," *Journal of Teacher Education*, May-June, 1988.

Conover, W. J. *Practical Nonparametric Statistics*, Wiley, New York, 1980.

Daniels, K., de Chernatony, L., and Johnson, G. "Validating a Method for Mapping Managers' Mental Models of Competitive Industry Structures," *Human Relations* (48:9), September 1995, pp. 975-991.

Dutton, J. E., and Jackson, S. E. "Categorizing Strategic Issues: Links to Organizational Action," *Academy of Management Review* (12:1), 1987.

El Sawy, O. A. "Personal Information Systems for Strategic Scanning in Turbulent Environments: Can the CEO Go On-line?," *MIS Quarterly*, March 1985.

Goldberg, D., Nicolas, D., Oki, B. M., and Terry, D. "Using Collaborative Filtering to Weave an Information Tapestry," *Communications of the ACM* (35:12), December 1992, pp. 61-70.

Good, N., Schafer, B., Konstan, J. A., Borchers, A., Sarwar, B., Herlocker, J., and Riedl, J. "Combining Collaborative Filtering with Personal Agents for Better Recommendations," in *Proceedings of the 1999 Conference of the American Association of Artificial Intelligence*, Orlando, FL, 1999.

Greening, D. R. "Building Consumer Trust with Accurate Product Recommendations: A White Paper on LikeMinds Personalization Server," Macromedia, (http://www.macromedia.com/software/likeminds/whitepapers).

Griffiths, W. E., Hill, R. C., and Judge, G. G. *Learning and Practicing Econometrics*, John Wiley & Sons, Inc., 1993.

Herlocker, J. L., Konstan, J. A., Borchers, A., and Riedl, J. "An Algorithmic Framework for Performing Collaborative Filtering," in *Proceedings of the 1999 Conference on Research and Development in Information Retrieval*, ACM Press, New York, 1999.

Hill, W., Stead, L., Rosenstein, M., and Furnas, G. "Recommending and Evaluating Choices in a Virtual Community of Use," in *Proceedings of ACM CHI'95 (Conference on Human Factors in Computing Systems)*, Denver, CO, 1995, pp. 194-201.

Johnson, E. J. "Expertise and Decision Under Uncertainty: Performance and Process," *The Nature of Expertise*, M. T. H. Chi, R. Glaser, and M. J. Farr (eds.), London, 1988.

Konstan, J. A., Miller, B. N., Maltz, D., Herlocker, J. L., Good, N., and Riedl, J. "GroupLens: Applying Collaborative Filtering to Usenet News," *Communications of the ACM* (40:3), March 1997, pp. 77-87.

Miller, B., Riedl, J., and Konstan, J. "Experiences with GroupLens: Making Usenet Useful Again," in *Proceedings of the 1997 Usenix Winter Technical Conference*, Anaheim, CA, January 1997.

Montazemi, A. R., and Conrath, D. W. "The Use of Cognitive Mapping for Information Requirements Analysis," *MIS Quarterly*, March 1986.

Podberezniak, E. *Collaborative Filtering in TV Recommender*, Unpublished Master's Thesis, New Jersey Institute of Technology, May 1998.

Rucker, J., and Polanco, M. J. "Siteseer: Personalized Navigation for the Web," *Communications of the ACM* (40:3), March 1997, pp. 73-75.

Santhanam, R., and Sein, M. K. "Improving End-User Proficiency: Effects of Conceptual Training and Nature of Interaction," *Information Systems Research*, December 1994.

Sarwar, B. M., Konstan, J. A., Borchers, A., Herlocker, J. L., Miller, B. N., and Riedl, J. "Using Filtering Agents to Improve Prediction Quality in the Grouplens Research Collaborative Filtering System," in *Proceedings of the Conference on Computer Supported Cooperative Work'98*, Seattle, WA, 1998.

Schafer, J. B., Konstan, J., and Riedl, J. "Recommender Systems in E-commerce," in *Proceedings of the ACM Conference on Electronic Commerce*, ACM Press, New York, November 1999.

Schenk, K. D., Vitalari, N. P., and Davis, S. "Differences between Novice and Expert Systems Analysts: hat Do We Know and What Do We Do?," *Journal of Management Information Systems* (15:1), Summer 1998, pp. 9-50.

Shardanand, U., and Maes, P. "Social Information Filtering: Algorithms for Automating 'Word of Mouth'," in *Proceedings of CHI (Conference on Human Factors in Computing Systems)*, Denver, CO, 1995.

Shanteau, J. "Psychological Characteristics and Strategies of Expert Decision Makers," *Acta Psychologica*, 1988, pp. 203-215.

Spence, M. T., and Brucks, M. "The Moderating Effects of Problem Characteristics on Experts' and Novices' Judgments," *Journal of Marketing Research*, May 1997, pp. 233-247.

Terveen, L., Hill, W., Amento, B., McDonald, D., and Creter, J. "PHOAKS: A System for Sharing Recommendations," *Communications of the ACM* (14:3), March 1997, pp. 59-62.

Vandenbosch, B., and Huff, S. L. "Searching and Scanning: How Executives Obtain Information from Executive Information Systems," *MIS Quarterly*, March 1997.

Wingfield, N. "Unraveling the Mysteries Inside Web Shoppers' Minds," *Wall Street Journal*, June 18, 1998.