Dorin MILITARU, PhD
Department of MARKETING, ESC Amiens
Groupe Sup de Co Amiens Picardie;
18, place Saint-Michel - 80038 Amiens FRANCE
E-mail: dorin.militaru@supco-amiens.fr
Costin ZAHARIA, PhD
MCF Université du Mans (IUT-GEA)
Chercheur ENSAM Paris/ESTP Cachan/IAE de Paris I
5, rue Joseph Granier – 75007 Paris FRANCE
E-mail : Costin.Zaharia@univ-lemans.fr

ONLINE COLLABORATIVE FILTERING-BASED SYSTEMS: SEMANTICS AND EFFICIENCY

Abstract. *Internet services operate on a vastly larger scale and allow virtual interactions. Both Web and Internet have created vast new opportunities, providing an infrastructure that enables buyers and sellers to find each other online. Companies can now offer many products, services and information easily and with lower costs. It becomes more and more difficult for customers to find quickly what they are looking for. Nevertheless, recommendation systems are playing a major role. Collaborative filtering (CF), or recommender system based-CF, has appeared as one methodology designed to perform such a recommendation task. These systems allow people to use expressed preferences of thousands of other people in order to find the product they desire based on the level of similarity between tastes. The concept has appeared from convergent research on search browsers, intelligent agents and data mining, and it allows to avoid the difficult question of "why" consumers prefer this or that product or brand.*

Early studies of electronic markets tools and recommender systems took a simplistic view of consumers as economic agents whose behavior was guided by the search for the lowest cost transactions. Moreover, most studies take into account only technical aspects of these systems like algorithms' development and computational problems. No study had been interested in recommendation's efficiency of collaborative filtering-based systems.

This article explores the current state of research in recommender systems-based collaborative filtering, and proposes an experiment to find if such electronic recommendations are better than human recommendations.

Key-words: Electronic Commerce; Collaborative filtering; Recommender Systems; Marketing, Experimentation.

JEL Classifications: C11, C63, D11, D12, D83, M31.

1. EMERGENCE OF COLLABORATIVE FILTERING

Among the several new situations or opportunities raised by the development of Internet, one is allowing the achievement of more rapid and extensive communication, but in a virtual world without face-to-face communication. It is therefore hypothesized that such an environment should favor the development of recommendation systems on which consumers could rely to help them first to assess the quality of products or services, and then eventually to purchase those goods. Indeed, "word-of-web" is starting to replace the classical word-of-mouth that circulates in traditional markets [1]. However, the impact of recommendation systems on markets and marketing practices is now a major – and challenging, if not puzzling – area of research.

Collaborative filtering has its origins in the earlier system of information filtering. This system removes redundant or unwanted information from an information stream using automated methods prior to presentation to a human user. Its main goal is the management of the information overload. These characteristics may originate from the information item - content-based approach - or the user's social environment - the collaborative filtering approach. On the presentation level, information filtering takes the form of user preferences-based feeds aggregator[1].

Collaborative filtering (CF) (sometimes called "social filtering" [2], or recommender system, [3]) has appeared as one methodology designed to perform recommendation through Internet. According to Goldberg et al., (2000) [4], Rich (1979) [5] is considered as the first reference for CF, but it is only in 1992 that Goldberg et al. [6] coined the term *"collaborative filtering"* in the context of a system for filtering emails. The concept has appeared from convergent research on search browsers, intelligent agents and data mining. Among the first reported algorithms and results were GroupLens [7] and Ringo [2]. The GroupLens team initially implemented a neighborhood-based CF system for rating Usenet articles. They used a 1-5 integer rating scale and computed a distance using Pearson correlations. The Ringo system, designed for music recommendations, tested a number of measures of distance between users, including Pearson correlation, constrained Pearson correlation, and vector cosine.

Applications of collaborative filtering typically involve very large data sets. CF is defined as the process of filtering for information using techniques involving collaboration among multiple agents, viewpoints, data sources, etc. It use a database containing a large number of users' known preferences to predict other preferences for additional products, services, or topics a new user might want to experience [8]. It is

[1] Software that runs on a user's computer and fetches, filters, and organizes selected information from the Internet.

based on three types of inputs: a user *(u),* an object *(o)* and a rating of this object by the user *(r(u,o)).* It is important to point out that in our conceptualization the data are ratings and not just observations of behaviors (for some mishaps resulting from confusion between preferences and behaviors, see the example in [9]).

The evolution of Internet-based CF has been very rapid, starting with applications in the area of Web pages (e.g. how to find useful and interesting web pages based on the behavior of similar users), and generalizing over products such as movies (SEPIA Video Guide [10]; Cinemax [11]; EachMovie; MovieFinder; MovieLens), books (barnes&noble.com [12]; Amazon.com), music (Amazon.com; Ringo), toys (Toys "R" Us [13]), restaurants (Entree [14]; Bostondine) and online newspaper (Los Angeles Times [15]; NewsDude). These systems have seemingly succeeded in low-risk content domains (products or services) but are still ignored in high-risk domains such as mutual funds or honeymoon destinations [16].

CF is said to be an "unprecedented system for the distribution of opinions and ideas and the facilitation of contacts between people with similar interests" [17] (p. 385). The same author also writes that it can be a significant tool in order to increase sales, cross-sells and up-sells, advertising revenues and the benefit of targeted promotions, to deepen customer loyalty (p. 363). But, in our opinion, it does offer more than a strategic advantage; it could help us understand better some important aspects of consumer behavior, such as the genesis of preferences with its consequences on decision making and evaluation, but also the role of culture, of social referents and other important aspects of consumer behavior.

In this paper, we firstly present the current state of research in recommender systems-based collaborative filtering; secondly, we propose an experiment to find whether such electronic recommendations are better than human recommendations.

2. CF SYSTEM ALGORITHM

The underlying assumption of CF approach is that those who agreed in the past tend to agree again in the future. People generally decide which movie to see, book to buy, or restaurant to eat at by talking to their friends. Then, they make a decision based on those opinions. The basic idea behind collaborative filtering concept is to automate word-of-mouth. The goal of collaborative filtering is to generate recommendations automatically, and on a large scale, using Internet as a mean to compile and diffuse others' evaluations.

Typically, in a CF web site, users are asked to tell about their preferences for a catalog of items, books, other web sites, or information. These opinions reveal how much a user liked an item. Then the system compares these profiles with other users and finds people with similar opinions, the neighbors. The system can predict how much the user will like what their neighbors said they were interested in but that the user has not yet experienced. Based on that, collaborative filtering can start recommending new material that the user is more likely to be interested in.

More formally, a recommender system based CF recommends objects (***o***) to the active user (***a***) based on the ratings of ***n*** others users. If we denote the set of all objects as ***O*** and the rating of user ***u*** for object ***o*** as ***r(u, o)***, the function ***r(u, o)*****:** ***O*** **→** ***R****,* maps objects to real numbers or to zero, which means "no rating". Usually ***R= {0, 1, 2, 3, 4, 5}*** or ***{0, 1,…, 7}***. We denote the vector of all user ***u's*** ratings for all objects as ***r(u, O)***, and the vector of all of subset of objects that the active user has rated is ***O-NR****,* where ***NR*** is all objects non rated. The vector ***r*****(u,** ***S*****)** is all of user ***u***'s ratings for any subset of objects ***S***. Finally, we denote the matrix of all users' ratings for all objects as ***r***. In general terms, a collaborative filter is a function ***f*** that takes as input all ratings for all users, and outputs the predicted ratings for the active user: ***r(a, NR) = f(r(1, O), r(2, O), ... , r(n, O)) = f(r).***

Collaborative filtering is a complex mechanism. The two main aspects of the collaborative filtering technology are *agents* and *algorithms.* Agents are entities that are capable of taking action on their own, and algorithms predefined sequences agents use to complete a task. Both work in background to evaluate all the users' preferences in order to make recommendations for a specific user. The complex system decides which algorithm is better to accomplish a specific task for a specific set of data. The entire process of CF based system is structured into three sub-tasks namely, representation of input data, neighborhood formation, and recommendation generation, as shown in the figure below.

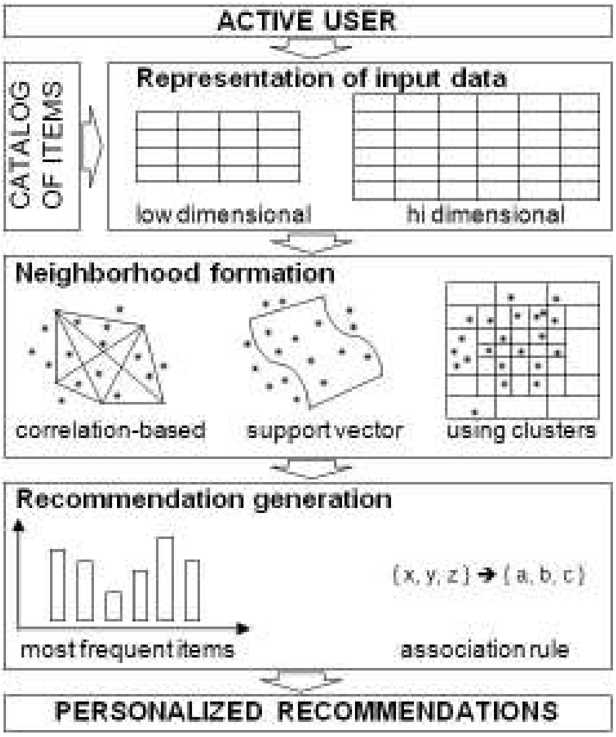

Figure 1: Example of item-based collaborative filtering

2.1 Representation of input data

The input data is a collection of expressed preferences, which is represented by a high- or low- dimensional matrix. The preferences are in fact ratings. A "rating" is a measure of a customer's opinion of a product, a service or a piece of information. There are explicit ratings (information that the user gives intentionally) and implicit ratings (an opinion inferred from watching user actions).

An important aspect of CF is the *critical mass.* How many ratings are necessary before a CF system can make a reliable recommendation? To start predicting the tastes, usually, ten to fifteen ratings from a user are enough. This is possible because systems-based CF also uses other data, such as the gender, zip code, or age of the end-user, gathered from the registration form each user must fill in to access a Web site. This representation of the preferences - as ratings - may potentially pose some problems such as scalability and synonymy.

2.2 Neighborhood formation

The *neighborhood selection sub-task* tries to find the neighbors (neighborhood) that are both strongly correlated with the active user and able to predict preference for many items not rated by the active user. There are many algorithms used to determine the neighborhood. Pearson correlation-based method, sometimes called “nearest neighbor algorithm”, is the most widely referenced in the literature. The support vector method and a scalable Pearson correlation-based method that uses clustering to improve scalability and accuracy are algorithms having already been used in the context of the EachMovie data set. Others algorithms used to determine the neighborhood include vector similarity-based methods and graphical models [8]. We describe here the first three algorithms, which are the most important.

Correlation-based prediction

The correlation-based method was originally used in the GroupLens project [7]. It uses Pearson correlation to compute a user’s predicted rating of an item. It weights users’ similarity, by using all available correlated neighbors, and computes a final prediction by performing a weighted average of derivations from the neighbor’s mean. The predicted rating $\boldsymbol{r_{predicted}(a, o)}$ of user $\boldsymbol{a}$ on item $\boldsymbol{o}$ is given by:

$$r_{predicted}(a,o) = \bar{r}_a + k\sum_{\substack{k=1 \\ k\neq a}}^{n} w(a,k)(r_{k,o} - \bar{r}_a)\,, \tag{2.1}$$

where $\bar{r}_a$ is the average rating for active-user, $\boldsymbol{k}$ is a normalizing factor ensuring that the absolute value of the weights sum to 1; $\boldsymbol{r_{k,o}}$ is the explicit rating given by user $\boldsymbol{k}$ for the item $\boldsymbol{o}$. The weights $\boldsymbol{w(a,\ k)\in [-1,\ 1]}$ can reflect distances, correlations, or similarities between user $\boldsymbol{i}$ and other users that have rated the same items. Most commonly, $\boldsymbol{w(a, k)}$ is the Pearson correlation coefficient between users $\boldsymbol{a}$ and $\boldsymbol{k}$:

$$w(a,k) = \frac{\sum_j (r_{a,j} - \bar{r}_a)(r_{k,j} - \bar{r}_k)}{\sqrt{\sum_j (r_{a,j} - \bar{r}_a)^2 \sum_j (r_{k,j} - \bar{r}_k)^2}}\,, \tag{2.2}$$

where the summations over $\boldsymbol{j}$ include items that both users $\boldsymbol{a}$ and $\boldsymbol{k}$ have rated in common ([8], [2]). This method has the advantage to be popular, intuitive, and relatively accurate [8], but it suffers from several limitations like the fact that the correlation between two users profiles can only be computed based on items that both users have rated.

Support vector method

Support vector method (SVM) is an alternative approach to the prediction of discrete ratings. It relies on existing expressed preferences to identify ratings classes used to make predictions. The SVM algorithm constructs an optimal margin hyperplane which maximizes the margin between the two classes; then, new "points" (predicted ratings) can be classified by identifying their position on this hyperplane. Support vector method is a statistical approach (see [19]). An advantage of this method is that it guarantees an optimal predicted rating because the solution is the mathematically optimal margin hyperplane. The disadvantage is that the accuracy of the prediction is affected by the missing data.

Correlation-based prediction using clusters of users
The idea behind this method is to assemble users into clusters. This reduces the number of users that are examined for similarity with the active user, and also address the scalability problems. To produce predicted ratings the algorithm treats the clusters as composite users and uses the clusters for determining likely neighbors.

2.3 Recommendation generation

The final step of a CF-based recommender system is to provide recommendations based on the determined neighborhood. There are two important methods with that respect [18]: the *most-frequent item recommendation,* and an *association rule-based recommendation.* The first method consists in scanning over all products rated in the neighborhood and performing a frequency count. The recommended product is chosen among the most frequent products that have not yet been rated by the active user. The second method is a traditional data-mining concept. It consists to find association rules between a set of co-rated products. More explicitly, the presence of some product in a particular transaction implies that products from other transactions are also present in the same transaction.

A number of applications combine the memory-based and the model-based CF algorithms. These overcome the limitations of native CF approaches. It improves the prediction performance. Importantly, it overcomes the CF problems such as sparsity and loss of information. However, they have increased complexity and are expensive to implement [19].

3. STUDY DESIGN

We conducted an empirical study to evaluate the efficiency of recommendations made by users' friends to those made by online CF systems. CF systems can typically take explicit or implicit input or a combination of the two [20]. In our study, we only examined a system that relied upon explicit input. The system studied was Movie Gift

Recommender (MovieLens[1]), a movie recommender system. This system is the most referenced in the literature on collaborative filtering-based systems.

Participants: 59 people participated in the experiment. Study participants were French Internet users. (*Participant Details*: Age range: 15 to 55 years; Gender Ratio: 33 males and 26 females; Study level: 3 undergraduate; 39 graduate; 17 postgraduate). Participants are asked to recommend and to evaluate movies. Each participant rates at least ten movies recommended by their friends, as well as at least ten recommendations made by CF system.

Procedure: The first part of the experiment involves the human recommendations. Each participant receives one type of questionnaire and completes the following tasks: (a) Answer some socio-demographics questions. (b) Tell us how many movies he watches by week, (considered his expertise level in movies). (c) Give us e-mail addresses for a friend familiar enough with their tastes to be able to recommend ten movies. The only constraint was that the friend must agree to participate to the experiment. (d) Tell how well he knows the person for which he provides recommendations (e) Recommend to his friend at least 10 chosen movies among a list of 100 suggested movies. For each movie recommended, he provides a rating supposed to express his friend's opinion on the movie. At the end of the list, he can add other movies.

The second part of experiment involved CF system recommendations. Participants receive another type of questionnaire and complete the following tasks: (a) Evaluate the friends' recommendations on the same scales as recommendations were made. (b) Complete online registration process and rate movies in order to get recommendations. (Systems may require users to complete a second step, where they will be asked for more ratings to refine recommendations.) (c) For each movie recommended by CF system, users provide ratings. They evaluate these recommendations on the same scales as recommendations made by their friends.

A. Human recommenders will consistently outperform CF-based system recommendations.
B. CF-based system provides the greatest number of useful recommendations to individuals.
C. Human recommenders provide the greatest number of trust-recommendations to individuals.
D. Expertise level in movies is directly correlated with the number of good recommendations provided by individuals.

---

[1] MovieLens is a free service provided by GroupLens Research at the University of Minnesota.

E. How well he knows the advised person is directly correlated with the number of good recommendations provided by individuals.

Table1. Correlations: Metrics of Human Recommendations with Expertise level and How well he know the advised person

	Human Recommendations		
	No. of good recommend.	No. of useful recommend.	No. of trust recommend.
Expertise level	-0.341*	-0.067	-0.188
How well he knows the advised person	0.746**	0.122	0.312

** significant at the 0.05 level; ** significant at the 0.01 level*

4. RESULTS

Our analysis focused mainly on comparing the quality of the recommendations made by online CF system and recommendations made by friends on three metrics (good, useful, and trust recommendations). One of our hypotheses was that friends would make superior good recommendations since they know the user well, and have intimate knowledge of his / her tastes in a number of domains. On the other hand, CF system only has limited, domain-specific knowledge about the users. In addition, recommender system does not match yet the sophistication of human judgment processes.

Good Recommendations: Next, we examined the differences in the quality of the recommendations provided by both friends and CF system. As the Figure 2 shows, for good recommendations, friends performed at significantly higher level than CF system (Friends=49,78%; CF system=21,62%; in a total number of 624 recommendations provided by friends and 590 recommendations provided by CF system). This finding supports our first hypothesis that human recommenders outperform CF-based recommender system.

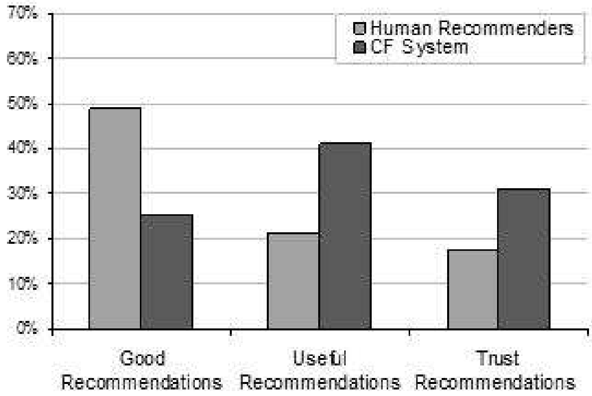

Figure 2. Comparing Human recommenders to CF

How are the recommendations provided by online systems different from those provided by humans on derived metrics?

For Useful recommendations metric experiment also supports our second hypothesis. CF system provides the greatest number of useful recommendations to individuals over entire number of good recommendations (CF system=42,12%, Friends=21,37%; in a total number of 310 good recommendations provided by friends and 127 good recommendations provided by CF system).

Trust-Recommendations: Recommended movies that the users had previously liked play a unique role in establishing the credibility of the human or online recommender system. Such recommendations are not useful in the traditional sense (i.e., recommendations which the user could use in the future), but help build trust in the provider of recommendation. Our experiment show that the number of trust recommendations provided by CF system is greater than the number of trust recommendations provided by individuals (CF system=31,11%; Friends=19,24%). This finding does not support our hypothesis that human recommenders provide the greatest number of trust-recommendations to individual. We propose possible explanations in the discussion section.

We also asked users to tell us how well they know the advised persons and how many films they watch by week. The table above shows the correlations between these two dependent metrics and the three metrics we created for human recommendations. We find that how well he knows the advised person is directly correlated with numbers of good recommendations. This finding supports our hypothesis. Expertise level does not correlate directly with the number of good recommendations. This finding is not consistent with our hypothesis.

5. CONCLUSION

Since the goal of most recommender systems is to automate and replace what is essentially a social process, we decided to directly compare the two ways of providing recommendations (humans and online CF systems). The quantitative results of our experiment indicate that the user's friends consistently provided better recommendations than CF system. Results showed that the percentage of good recommendations provided by friends is two times greater than the percentage of good recommendation provided by CF system. However, on both derived metrics, useful- and trust- recommendations, online system perform better. This is an unexpected result. It is normal for CF system to provide many useful recommendations because it uses a big movie database and for this reason recommendations can easily be "new" and sometimes, "unexpected". Nevertheless, as the movies recommended by friends were chosen from a list of 100 suggested movies and from the movies they already know, it would have been more likely to find more trust-recommendations provided by friends. A possibly explanation is that because our subjects are interested in movies they have already seen many movies. Or perhaps considering the fact that the suggested movies are mostly American, the French subjects might have found them inadequate. Correlation test result shows that better they know the advised person, more objective the recommendations are.

7. REFERENCES

[1]**Rosen, E. (2000),** ***The Anatomy of Buzz.*** New York, NY: Doubleday;

[2] **Shardanand, U. and Patti, M. (1995),** ***Social Information Filtering: Algorithms for Automating.*** 'Word of Mouth', MIT Media-Lab; www.acm.org/sigchi/chi95/Electronic /documnts/papers/us_bdy.html;

[3]**Resnick, P. and Hal, R. V. (1997),** ***Recommender Systems***. Communications of the ACM, 40, 3, 56-58;

[4]**Goldberg, K., Theresa, R., Dhruv, G. and Chris, P. (2000),** ***Eigentaste: A Constant Time Collaborative Filtering Algorithm***. UCB Electronics Research Laboratory Technical Report M00/41, (August);

[5]Rich, E., 1979, User Modeling via Stereotypes, Cognitive Science, 3, 335-366.

[6]**Goldberg, D., Nichols,d., Oki, B. M. and Terry, D. (1992),** ***Using Collaborative Filtering to Weave an Information Tapestry***. Communications of the ACM, 35 (12), 61-70;

[7]**Resnick, P., Iacovou, N., Suchak, M., Bergstrom, P. and Riedl. J. (1994),** ***GroupLens: An Open Architecture for Collaborative Filtering of Netnews***. Proceedings of ACM 1994 Conference on Computer Supported Cooperative Work, Chapel Hill, NC, 175-186;

[8]**Breese, J. S., Heckerman, D. and Kadie, C. (1998),** ***Empirical Analysis of Predictive Algorithms for Collaborative Filtering*** . Technical Report MSR-TR-98-12, Microsoft Research, (October);

[9]**Ansari, A., Essegaier, S. and Kohli, R. (2000),** ***Internet Recommendation Systems***. Journal of Marketing Research, 37, (August), 363-375;

[10] vguide.sepia.com/

[11] www.cinemax.com/

[12] www.barnesandnoble.com/

[13] www.toysrus.com/

[14] infolab.ils.nwu.edu/entree/

[15] www.latimes.com/

[16**] Herlocker, J. L., Konstan, J. A. and Riedl, J. (2000),** ***Explaining Collaborative Filtering Recommendations***. CSCW '00, Philadelphia, PA., December 2-6;

[17**] McKim, R. (2000),** ***Collaborative Information Filtering Systems.*** Journal of Database Marketing, 7, 4, 385-389;

[18] **Sarwar, B., Karypis, G., Konstan, J. and Riedl, J. (2000),** ***Analysis of Recommendation Algorithms for Ecommerce***. EC'00, Minneapolis, Minnesota: ACM, October 17-20;

[19] **Das, A. S., Datar, M., Garg, A., Rajaram, S. (2007),** ***Google News Personalization: Scalable Online Collaborative Filtering***. *Proceedings of the 16th international conference on World Wide Web,* ACM New York, NY, USA;

[19**] Vapnik, V. (1998),** ***Statistical Learning Theory***; New York, NY: John Wiley, p.732;

[20] **Schafer, J. B., Konstan, J. and Riedl, J. (1999),** ***Recommender Systems in E-Commerce***. ACM Conference on Electronic Commerce. http://www.cs.umn.edu/Research/GroupLens/ec-99.pdf.