Designing Social Filtering Systems by Embedding Explicit Receiver-Source Relationships

Ofer Arazy
School of Business
University of Alberta
Edmonton, Alberta, T6G 2R6, Canada
ofer.arazy@ualberta.ca

Nanda Kumar
Baruch College, City University of New York
55 Lexington Avenue, Box B11-220
New York, NY 10011, USA
nanda_kumar@baruch.cuny.edu

Bracha Shapira[1]
Department of Information Systems Engineering
Faculty of Engineering Sciences, Ben-Gurion University of the Negev
Beer-Sheva, Israel 84105
bshapira@bgumail.bgu.ac.il

ABSTRACT
For over a decade now, the ad-hoc standard in social filtering (or Collaborative Filtering) has employed an approach, where recommendations were generated by computing "shared interests" based on users' preferences for items. The rapid growth in online social networks such as myspace.com presents an opportunity to identify other relevant relationship characteristics among participants who know each other and use these characteristics to improve the quality of the recommendations generated. This paper proposes a conceptual framework for determining how alternative measures of source quality could be obtained in an online setting, and develops a model that incorporates users' explicit perception of sources' trustworthiness to improve the quality of the recommendations.

Author Keywords
Recommender Systems, Collaborative Filtering, Social Networking, Trustworthiness

ACM Classification Keywords
I.4.3, H.5.4

INTRODUCTION
Information overload impedes organizations' performance. The rapid growth of available information online, fuelled by the rapid adoption of the internet and the web, is making access to relevant information analogous to that of finding a needle in a haystack. This problem is accentuated in today's markets, where information and knowledge play a critical role in firms' competitive positioning. Recommendation systems play a significant role in reducing information overload and provide users with information relevant to their specific interests. Recommendation systems are now an integral part of firms' information architecture, serving both customers (e.g. Amazon's book recommendations) and internal employees.

SOCIAL FILTERING
Recommendation literature suggests that people usually access information through social means, i.e. relying on their friend's and acquaintances advice. In the mid 1990s social (or collaborative) filtering techniques emerged, and since then the social approach to the design of recommendation systems has become the industry standard. Social Filtering works as follows: (1) identifying sources that the recipient trusts, (2) summarizing the recommendations (i.e. quality evaluations) of these trusted sources, (3) and generating a recommendation [1, 2].

How could trusted sources be identified in an online setting? One specific operationalization of the Social Filtering model, where a source's trustworthiness is estimated based on the degree to which receiver and source share interests [2], has gained dominance over the last decade. Most of the current recommendations systems employ the Shared-Interests approach (e.g. Amazon's recommendation system) as it is easy to automatically establish the relations between recommendation recipient and sources. Users' ratings of products (i.e. recommendations) could be gathered either implicitly (e.g. based on a purchase or other recorded transaction) or

[1] Authors' names are presented in alphabetical order

explicitly without requiring much effort. The recorded history of users' behaviour is then used to identify similar individuals, thus generating recipient-sources relationships automatically.

Another advantage of the Shared Interests approach was the large number of recommendations sources that were made available for each user, since shared-interests relationship are not restricted to source the recipient knows. Access to the ratings of these large number of sources is an important factor in ensuring recommendations quality. In recent years, the Shared Interests SF approach has attracted significant attention, and research in this area has provided enhancements along various dimensions such as automatic elicitation of accurate user feedback, algorithms for measuring users' similarities, and improving prediction methods, resulting in better system effectiveness over alternative filtering approaches [3].

EMBEDDING EXPLICIT TRUST RELATIONSHIPS

Knowledge Sharing literature suggest that several factors determine the likelihood that a recipient will accept a source's recommendation, e.g. source's trustworthiness and source's reputation [4, 5]. The limitations of the Shared-Interests operationalization stem from the fact that similarity of interests may not be the best proxy for sources' trustworthiness, and thus may not produce the best possible recommendations. How alternatively could trust-relations be extracted in an online setting? Two promising alternatives have emerged in recent years:

1. In online social networks, users establish relations with friends, and these relations are used for communicating and sharing information. To the extent that these relationships indicate common trust among friends, they could potentially be utilized by recommendation systems. The growing interest in this area in academia, as well as the emergence of social information access systems (such as Yahoo's MyWeb) illustrates the promise of this approach.
2. Online rating and reputation systems help to establish users' expertise in an online social setting. To the extent that automatically-calculated reputation represents experts' collective trust, reputation mechanisms could be incorporated into recommendation systems, such that an expert's trustworthiness is estimated based on his reputation.

Indeed, the rapid growth in online social networks presents an opportunity for refining the existing social filtering approach, namely in exploiting the relationships embedded in a social network to detect trusted sources for generating recommendations. Epinions.com is one of the most notable example of a commercial social filtering system, where users' trust relationships form a "web of trust", which is then utilized to filter product recommendations. In a closely related field - web search - we are now witnessing the emergence of new social-based systems, such as del.icio.us and Yahoo's MyWeb (still in Beta version).

In academic research, since the early work of Goldberg et al. [1], very few works addressed the issue of explicit trust relations for establishing quality sources in filtering systems. In a few recent formulations of trust, it has been suggested that explicit trust relations could be utilized by search and filtering systems [6, 7]. Gnasa et al. [8] propose a search architecture that includes social relations (although these are restricted to shared interests), and have developed a prototype (Iskodor) to illustrate their ideas. Freyne and Smith [9] introduce the I-Spy search engine, which employs un-named explicit social relations (and the recorded browsing experiences of related individuals) to enhance search performance within defined communities that share interests. The studies reviewed above reflect works-in-progress, demonstrating the rising interest in SF systems with explicit receiver-source trust relationships.

Research on these alternative approaches for establishing source's trustworthiness in recommendation systems is in its infancy, and to date key questions remained unanswered, such as the comparative usefulness of the three approaches for estimating source's trustworthiness. Hence, our research seeks to develop a conceptual framework to answer the following questions: (a) What attributes of the source are indicators of source's trustworthiness and the recipient's likelihood of accepting the recommendation? and (b) Which of the three alternative approaches for operationalizing 'source's trustworthiness' - based on similarity of interests, explicit friendship relations that are embedded in online social networks, or experts that are identified through online reputation mechanisms - is most useful for recommendation systems"?

THEORETICAL FRAMEWORK

Based on the review of Marketing, Communication and Knowledge Management literature, we identified three complementary measures of source's quality that will lead to system effectiveness enhancements over the common existing model (i.e. employing only the shared-interests measure). Figure 1 below is a graphic illustration of the proposed conceptual social filtering model.

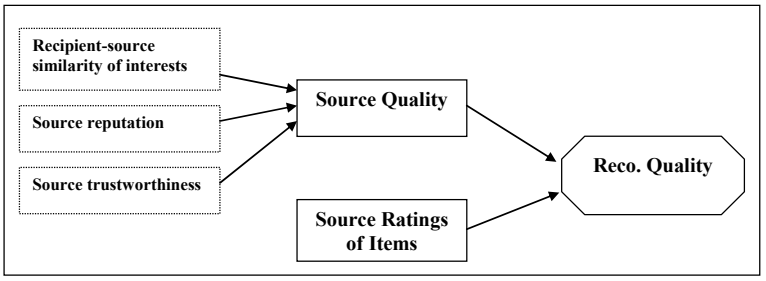

Figure 1: Proposed Conceptual Social Filtering model

As seen in the Figure, the recommendation to a recipient is based on *n* sources. The model identifies the quality of each of the sources, and aggregates their former ratings, considering each source's quality. The quality of the source is determined by three main factors: the first is the traditional recipient-source similarity of interests (the shared-interest approach); the second and the third are the additional social indicators that we propose in this paper: namely, source's reputation and trustworthiness. The final recommendation of an item to the recipient is then computed by measuring the quality of each of the sources, assigning weights to sources based on their quality, and aggregating sources' former ratings, factoring-in source's weights.

The usefulness of the proposed model depends to a large extent on the availability of the various source-quality indicators. In an online setting, similarity of interests is the most accessible measure and could be automatically calculated whenever history logs of users' behavior are maintained. Measures of users' reputation are only available in settings where rating and reputation mechanisms are in place. The challenge in implementing such mechanisms is not the technology involved in computation of recommendations, but rather in setting-up technological and social controls that would assure normative behavior of users, as realized by commercial web sites such as Amazon, eBay, and ePinions. Successful implementation of reputation mechanisms requires that users rate others' recommendations, and incentives for inducing such a behavior are necessary.

Trustworthiness measures – the third type of source-quality indicator - are least accessible, as they require users to explicitly define their trusted sources. Although trust relationships could be extracted from online social networks (i.e. in leveraging relationships that were already explicated in forming recipient's social network), the number of sources for each recipient is bound to be small. This limitation could be addressed by propagating the trust relationship in the network. Trust related literature suggests that trust between individuals could be propagated, such that if user A trusts user B and User B trusts user C, user A most probably trusts user C. Such trust propagation mechanisms are already utilized in various settings, e.g. in Slashdot.org users can filter information coming not only from friends, but also from friends of friends, and ePinions.com propagates trust in their network to establish indirect links between users. Trust propagation is a recent developed area of research and various propagation algorithms are being explored.

A practical advantage of the proposed model is that it associates the recipient with more sources, and thus is likely to provide more robust recommendations. One of the major limitations of existing realizations of the shared-interests model is data sparsity – i.e. having insufficient product evaluations (and thus unreliable recipient-source associations), especially during the start-up period. Utilizing additional source-quality indicators will increase the reliability of recipient-source associations.

CONCLUSIONS AND FUTURE WORK

In order to successfully incorporate our proposed framework into social filtering systems, it is required that we establish the relative weights of shared-interests, source's reputation, and trustworthiness in determining source quality. However, it is a challenge to determine the relative weights of the various source-quality indicators, as these weights may vary across settings

We plan to conduct a series of empirical studies in 2006 – across different settings – in which we will measure the social relations (e.g., 'Trust', 'Friendship', etc.) and individual's reputation in a group of subjects that are familiar to one another. In addition, we will ask subjects to evaluate products, and we would establish shared-interests similarity measures. We plan to feed these various source-quality indicators, along with subjects' product evaluations, into a standard collaborative filtering system, to generate recommendations. The performance of the system would be evaluated based on the extent to which the system's predictions (i.e. recommendations) match subjects' own evaluations of the products. We plan to study system performance when using each of the three distinct source-quality indicators, as well as to explore combinations of these indicators to produce optimal performance. If given a chance, we will be in a position to present the results of the study in the conference.

In conclusion, recommendation systems, and specifically Social Filtering (SF) systems, play a significant role in reducing information overload and providing users with information relevant to their specific interest. For over a decade now, the ad-hoc standard in social filtering employed an approach, where the shared-interests relation was used to generate recommendations, mainly due to difficulty in obtaining information about alternative types of receiver-source relations. The rapid growth in online social networks and reputation systems presents an opportunity for a new social filtering approach, where reputation measures and explicit relations from social networks could be utilized to indicate sources' quality. This approach could complement, or potentially even replace, the current practice of using similarity of interests to detect quality sources, and has the potential to enhance recommendation systems performance.

REFERENCES

[1] D. Goldberg, D. Nichols, B. Oki, and D. Terry, "Using Collaborative filtering to Weave an Information Tapestry," *Communicatioins of the ACM*, vol. 35, pp. 61-70, 1992.

[2] U. Shardanand and P. Maes, "Social Information Filtering: Algorithms for Automating Word of Mouth," presented at Proceedings of the SIGCHI conference on

Human Factors in Computing Systems, Denver, Colorado, USA, 1995.

[3] D. Kelly and J. Teevan, "Implicit Feedback for Inferring User Preferences," *SIGIR Forum*, vol. 37, pp. 18-28, 2003.

[4] D. Z. Levin and R. Cross, "The Strength of Weak Ties You Can Trust: The Mediating Role of Trust in Effective Knowledge Transfer," *Management Science*, vol. 50, pp. 1477-1490, 2004.

[5] H. McKnight, V. Choudhury, and C. Kacmar, "Developing and Validating Trust Measures for e-Commerce: An Integrative Typology," *Information Systems Journal*, vol. 13, pp. 334-359, 2002.

[6] A. Abdul-Rahman and S. Hailes, "Relying on Trust to Find Reliable Information," presented at Proceedings 1999 International Symposium on Database Web and Cooperative Systems, Baden-Baden, Germany, 1999.

[7] R. Guha, R. Kumar, R. Prabhakar, and A. Tomkins, "Propagation of Trust and Distrust," presented at Proceedings of the 13th Conference on World Wide Web, New York, NY, USA, 2004.

[8] M. Gnasa, M. Won, and A. Cremers, "Three Pillars for Congenial Web Searching – Continuous Evaluation for enhancing Web Search Effectiveness," *Journal of Web Engineering*, vol. 3, pp. 252-280, 2004.

[9] J. Freyne and B. Smyth, "An Experiment in Social Search," presented at Proceedings of the 3rd International Conference on Adaptive Hypermedia and Adaptive Web-Based Systems, Eindhoven, The Netherlands, 2004.