# Interacting with Recommender Systems

Don Turnbull (donturn@fis.utoronto.ca)

## *Interests*

My primary interest in designing Recommendation Systems is from my research on designing a system to locate, evaluate, and recommend Web pages for intranets of Web users. I've chosen this problem domain mainly because of the ease of collecting Web use, the items used (Web pages), and the relatively open internet development standards.

In many ways, I'm approaching Recommender System design from an Information Retrieval perspective. I propose that some of the most applicable systems for recommendations are concerning documents, and also that the best way to help tune Recommender System interfaces is using statistics to measure their use. With multiple participating users, we can both match recommended resources and modify the view of the resources and the interface itself. I find a certain amount of similarity in ideas of precision and recall in relation to interface design. Maybe this is a move from *User Centered System Design* (Draper and Norman, 1986)(Draper & Norman, 1986) to *Use Centered Systems Design*.

## *Ideas*

I offered these comments on the COLLAB mailing list on what I thought the differences were between the different types of Recommender Systems:

**Social Filtering** seems to be when a group of users are helping each other anonymously or unaware of all that their preferences do- as in these "most-visited" web sites lists or to build an ad hoc kind of Yahoo for information objects (which is fairly obtuse and rigid, without much rapid change). For some reason words like "trends" and "memes" come to mind when thinking about how social filtering will try to point to what's hot or what's on the rise.

**Collaborative Filtering** seems to have a more team-oriented (a corporation, a group of people with a common purpose?) purpose and might involve some explicit ranking or rating of resources. This one seems to be the most complex as I tend to think about it having a more complex processing system- dealing with information objects and is constantly changing in that new information is added and relevancy of information is always shifting as well. A corporate intranet with filtered information on topics of interest or the underlying system to coordinate intranet (and internet) knowledge management might be two vague examples. We would have to measure the changing value of the information objects AND each user's value of the object (remembering that uniqueness, confidentiality or freshness of the information can influence the value).

**Recommendation Systems** seems to be the most simple of the three usually dealing with concrete objects (compact discs, movies, plumbing equipment) that are fairly static and have a definite, user-understandable set of characteristics (both for rating and comparing). The recommendation system can treat each characteristic as a value and simply run correlations for recommendation. I tend to think of these systems as being independent of the content they are recommending, changes in the recommended objects aren't likely to happen - however changes in the user's rating of the object might change over time. Obviously, Firefly would fit into this type of system. (Turnbull, 1998)

I still feel that Collaborative Filtering is a more accurate term for this genre of work, at least for the domain I'm interested in - the Web. I'm convinced that the best places to start building CF Systems are on the Web and for Web content. Here are a few reasons why the Web is a good domain for Recommender Systems

- URL (transmission protocol, domain name, document name, directory path)
- HTML-structured information
- Yahoo services to generally organize
- Date and time of information available
- May have been/can be indexed with available technologies.
- Log information (if available)

I've been working on some of these ideas to learn more about how people use the Web (Choo, Detlor, & Turnbull, 1998) which can also be used to recommend new Web resources. In many ways, this is simply new Information Retrieval work.

## *Information Retrieval and Recommendation Systems*

It is difficult to design an interface to do all of the information display for an application - sometimes it's easier to leverage the intelligence of its users. Typical interface issues are how to display a large amount of data at once and how to show (the perceived) most relevant information first. Even with the best-designed interface, users will gradually use some features more than others and read certain types of notes more often, filtering helps use this feedback to enhance the application's usability (or the information gathered via the application).

Belkin (Belkin & Croft, 1992) suggests that Collaborative Filtering is a new implementation of Information Retrieval. What makes it different from IR is that feedback is more important and aggregate feedback (how the information retrieved is used) among groups of users is measured as well. Some examples of reasonably filtering applications:

- delivering information to users, most likely unstructured data:
    - not from a controlled (managed or optimized) database
    - primarily textual information (is text still the most important) and some multimedia
    - large amounts
    - constantly incoming
- based on descriptions (types) and profiles (of users) as algorithms or tables (adaptively - profiles should change)
- can include text routing (email filters) and classification and categorization
- extract relevant information or organize it for seeking
- similar to IR Systems (more interactive and responsive)
- considers the source of the material (where, when, who the information originated)
- compares a history of queries (either locally or globally for users)

Is Collaborative Filtering, the next big thing in Information Retrieval? These are the key issues that take filtering beyond typical Information Retrieval:

- concerned with repeated use of the system (long-term goals)
- profiles can gradually be tuned to be more useful
- focuses on distributing information, not just finding it
- is done with dynamic amounts of information
- timeliness of information taken (more) into account
- much more user focused
- privacy concerns are being considered (again)

This next addition to IR is the idea of building and learning from user profiles. I think a lot of issues related to Data Mining and more specifically Knowledge Discover in Databases (Fayyad, Piatetsky-Shapiro, & Smyth, 1996; Fayyad, 1996) (now a new ACM SIG, SIGKDD) will prove useful in leveraging data for Recommender Systems.

## *Information Retrieval and the Web*

These systems comprise a variety of interaction techniques, interfaces, and types of recommendation information that have had success:

**Mosaic**- Berners-Lee, CERN

- Let users save bookmarks
- Users could publish or distribute bookmarks
- Comments added to Web pages

- Some searching of comments
- Links imply relations
- Graphical Interface

**Firefly** - (previously HOMR and Ringo) Metral, Shardanand, and Maes, MIT (Shardanand, 1995)

- "User modeling" (profile)
- "Word-of-mouth" recommendations
- Vector matching based on simple rating scale
- New user of algorithms (statistical models - mean squared difference, Pearson, correlation)
- Simple content (music and movies)
- AI language processing (Scheme and LISP)

Expanded from movies and music to community interests and other media.

**Yahoo** - Filo and Yang

- Expert classifiers (human and computer augmented)
- Dominant standard (their classifications are the Web's)
- User contributions

Point's Top 5% might also be in this category. These Web page classifier sites are de facto recommender systems as most of their links are sent in by users. It is not interactive or adaptive however.

**Lotus Notes** - Maltz and Ehrlich, Lotus, Corp. (Ehrlich, 1995; Krulwich, 1995; Maltz & Ehrlich, 1995)

- Built-in users (corporate mandate)
- Closed system of similar users (goals)
- Structured data (database triggers)
- Existing filtering and communication
- Search criteria assignable to documents (selection and classification)
- Public and Private data stores
- Active (Retrieval) and Passive (Seeking) filtering

InterNotes Navigator also adds browsing features to Notes and allows for sharing of bookmarks. It adds this feature on to an existing application, not a whole new application.

## *Interfaces for Collaborative Systems/Browsers*

The following systems provide some good design examples of how to use filtering and Collaborative Filtering to make finding (the right) things easier.

**GroupLens** - Resnick, et. al. (Univ. of Minnesota) (Resnick, Iacovou, Suschak, Bergstrom, & Riedl, 1994)
Explicit rating is required. It focuses on Usenet news and included the following features:

- Expands netnews characteristics (hierarchy and some moderation)
- Rating services (people agree over and over)
- Compares ratings
- Suggests information based on others' ratings
- Aggregate queries for each user
- Servers gather ratings and make (distribute) predictions - scalable
- Easy to enter ratings (buttons at top of window)

Of course, most of its benefits are derived mathematically, not with its interface (of the information space). However, its long success shows that its filtering capabilities are useful enough that users continue to use the system. I think explicitly rating something is the antithesis of a good interface however.

**SenseMaker** - Baldonado and Winograd, Stanford Digital Libraries Project (Baldonado & Winograd, 1996)
SenseMaker's real strengths are:

- Context-driven (examples)
- Cumulative use collected

- User-centered tasks
    1. Finite searching (one goal)
    2. Seeking/Browsing
    3. Reference searching (many goals or broad topic)
- Many dimensions and levels of detail (collaboration among services)
    - Expanding and Contracting Levels of information
    - Duplicate Deletion
    - Limits
    - Views (Bundling similar object of information)
- Iterative and Back-propagating

While many of these ideas are seen in many other examples, the right combination seems to be found here. SenseMaker provides a good interface and view into the data.

**WebBook and WebForager**- Card, Robertson, and York, Xerox PARC (Card, Robertson, & York, 1996) InXight, WebBook/WebForager basically groups information into virtual books. These books facilitate moving among various books and pages of information (mostly web sites and wb pages) in a more native manner to beginning users. It's main strengths are:

- Interactive Seeking
- Levels of aggregation for Information
- Extra dimensions of data (3D)
- Relationships indicate meaning (distance = similarities)

Unfortunately, its weakness is trying to use a bad metaphor for a large amount of information. Who would want to be constrained by physical-oriented models when dealing with such a vast amount of information? Possibly useful as training wheels for new Web users.

## *Design Problems*

In looking at several of these more useful examples, we can see some obvious areas where more design work is needed:

- Make modifying viewing preferences (and there should be some) easy (and quick).
- Provide more interactive aids in using/seeking/selecting/reading recommendations.
- Making navigation through the recommendations more natural, keeping the user from getting lost or overwhelmed. What should be the metaphor?
- Notifications for information
- Reminders for information
- How to "age" information (appearance and automatic (re)moval)
- Simple feedback if recommendations are coming in real time.

## *My Background*

Don Turnbull (donturn@fis.utoronto.ca) is a Doctoral Candidate in the Faculty of Information Studies at the University of Toronto. He was a Lead Technical Architect at IBM Interactive Multimedia - Atlanta, working on the World Book/IBM Multimedia Encyclopedia and other Internet-related products. Don received his MS in Information, Design, and Technology from the Georgia Institute of Technology (Georgia Tech) in 1995. He also worked as a Methodologist at KnowledgeWare, Inc. designing CASE tools for client-server applications as well as working on Usability and Online Information Development issues. His research interests include intelligent agents, collaborative filtering, information seeking theory, and World Wide Web software architectures. He is a co-author of "Knowledge Management using the World Wide Web: Augmenting Organizational Intelligence through Intranets" due to be published in 1999.

## *References*

Baldonado, M. Q. W., & Winograd, T. (1996). *SenseMaker: An Information-Exploration Interface Supporting the Contextual Evolution of a User's Interests* (SIDL-WP-1996-0048). Palo Alto, CA: Standord University.

Belkin, N. J., & Croft, W. B. (1992). Information Filtering and Information Retrieval: Two Sides of the Same Coin? *Communications of the ACM, 35*(12), 29-38.

Card, S. K., Robertson, G. G., & York, W. (1996). *The WebBook and the Web Forager an information workspace for the World-Wide Web*. Paper presented at the Conference proceedings on Human factors in computing systems, Atlanta, GA.

Choo, C. W., Detlor, B., & Turnbull, D. (1998). *A Behavioral Model of Information Seeking on the Web - Preliminary Results of a Study of How Managers and IT Specialists Use the Web.* Paper presented at the Proceedings of the 61st Annual Meeting of the American Society of Information Science, Pittsburgh, PA.

Draper, S. W., & Norman, D. A. (1986). *User centered system design : new perspectives on human-computer interaction*. Hillsdale, N.J.: L. Erlbaum Associates.

Ehrlich, K. L., Alex; Ritari, Philio. (1995). *Hypertext Links In Lotus Notes And The World Wide Web* . Cambridge: Lotus Development Corporation.

Fayyad, U., Piatetsky-Shapiro, G., & Smyth, P. (1996). From Data Mining to Knowledge Discovery in Databases. *AI Magazine, Fall 1996*, 37-54.

Fayyad, U. M. (1996). *Advances in knowledge discovery and data mining*. Menlo Park, Calif.: AAAI Press : MIT Press.

Krulwich, B. (1995). *Learning user interests across hetergeneous document databases* (krulwich@andersen.com). 100 South Wacker Drive, Chicago, IL 60606: Andersen Consulting LLP.

Maltz, a., & Ehrlich, K. (1995). *Pointing the way: active collaborative filtering.* Paper presented at the CHI '95.

Resnick, P., Iacovou, N., Suschak, M., Bergstrom, P., & Riedl, J. (1994). *GroupLens: An Open Architecture for Collaborative Filtering of Netnews.* Paper presented at the CSCW 94, Chapel Hill, NC.

Shardanand, U. a. M., Pattie. (1995). *Social Information Filtering: Algorithms for Automating "Word of Mouth".* Paper presented at the ACM SIGCHI '95, Vancouver.

Turnbull, Don. RE: Social filtering or Collaborative filtering? From COLLAB list. May, 7 1998. http://www.sims.berkeley.edu/resources/mailing-lists/collab/0112.html.