The Wisdom of Networks: Matching Recommender Systems and Social Network Theories

R. Dandi[1]

Abstract. This paper aims to analyzing the match between social network theories and recommender systems. Several social network theories provide explanations on why nodes link to each others. At the same time, recommender systems recommend users to connect to some items according to different internal algorithms. The study identifies the theoretical mechanisms behind the main types of recommender algorithms, and specifically behind network-based ones. Main design implications for recommender algorithms are derived.

Introduction

We can define a recommendation as the communication of a prediction: the prediction of the utility of an item to somebody. Recommendations are a key part of knowledge exchange, where an expert communicates to a non-expert the utility to access an information or a piece of knowledge, to perform a task, or to connect to another expert.

Recommender systems automatically generate information and knowledge by mimicking the recommendation process among humans. Recommender systems available online may suggest different "items" such as books (Amazon.com, Barnes&Noble.com), movies (Movielens), music (Pandora.com), jokes (Jester.com), people to date (Meetic.it, YahooPersonals), friends (Facebook.com), restaurants (2spaghi.it), etc. Users instruct the system about their preferences. The focus is not just on retrieving information but on filtering it. Only the relevant information must be recommended in order to avoid information overload.

Recommender systems are an essential part of Web 2.0 websites [1] where they put in practice the concept of the so-called "wisdom of crowds", from the title of the book by James Suriowecki [2]. In Web 2.0 websites, users are also content generators and their behaviour and relationships produce recommendations to

1 Post-doc researcher at LUISS Guido Carli University, Rome, Italy, rdandi@luiss.it

other users. On Facebook, for example, the system recommends Facebook pages that friends have voted positively with the rating system "ILike". Explicit or implicit "votes" (e.g. number of downloads) are also used in other Web 2.0 websites (Amazon, eBay, Youtube, MySpace, etc.). In a sense, the democratic voting process is a way to identify items that can be useful to others. These types of recommender systems are called *collaborative filtering systems* because they rely on the fictional collaboration of the entire set of users in providing recommendations.

In addition to reducing overload and collaboratively identifying useful items, recommender systems are one of the most powerful tools for commercial websites to build the ideal of "mass customization". In fact they: (i) convert "browsing" users into buyers by suggesting appropriate items to users who otherwise would not have purchased an item [3]; (ii) increase cross-selling by recommending additional and complementary items to those already chosen by a user [3], [4]; (iii) build customer loyalty, as "the more a customer uses a recommender system – teaching it what he wants – the more loyal he is to the site" [5]; (iv) better the understanding of customer needs and of market segments [6].

As company data are more easily integrated, organized, and analyzed thanks to enterprise data warehouses, recommender systems are increasingly adopted in knowledge-intensive organizations for the support of knowledge exchange. Recommender systems are used for example to support knowledge exchange in the research process [7] or to improve knowledge exchange and expert recognition within business organizations [8] and within communities [9], [10]. Recommender systems in fact permit both (i) a "repository view" of knowledge management [11] emphasizing the gathering, providing, and filtering of explicit knowledge, and (ii) an "expertise sharing" model, emphasizing the identification and access of experts and the connection between people [12].

Finally, latest recommender systems not only are based on collaborative voting and item features but also on social networks among users. Items are ranked according to the usefulness to people with a connection to the focal user [8], [13], [14], [15], [16]. In this sense, it is the "wisdom of networks" that produces recommendations.

The aim of this paper is to highlight the theoretical mechanisms behind the major types of recommender systems and to find the intersections between network research and recommender algorithm design.

Types of Recommender Systems Algorithms

There are two basic entities that are included in any recommender system: items (books, people to contact, songs, articles, etc.) and users. Recommender systems need some input information on these entities such as: demographic data of the users, item features, and users' ratings of the items (ratings can be explicit or implicit, binary or on different scales). The recommendation problem is to make a prediction of the user ratings of unrated items, and to recommend one or more items that maximize the utility of the user [17].

We can identify 6 types of computer-supported recommender systems:

1) **Content-based systems** recommend items that are similar to the items the focal user liked in the past. For example, the internet radio Pandora.com provides song recommendations based on the structural features (types of vocals, music genre, etc.) of the first song selected by a user.

2) **Collaborative filtering systems** recommend items to a particular user based on the ratings of users who show similar tastes for the same items [4]: starting from a user-by-item matrix, the systems groups users whose ratings are similar; the system then recommends the items selected by the groups to each member of the groups. Ringo [18] is an early example.

3) **Item-based collaborative filtering** [19] or **item-by-item collaborative filtering** [20] is the method adopted by Amazon.com for its famous functionality "customers who bought this book also bought these books". Starting from a user-by-item matrix, the recommender creates an item-by-item matrix where the generic cell represents the similarity between items. This method does not need to compute a neighbourhood of similar users; therefore it is much faster than "user-based" collaborative filtering methods.

4) **Demographic filtering** categorizes users according to demographic data and generates recommendations based on demographic classes [21].

5) **Average ratings** are non-personalized recommendations based on what other users have said about the items on average**.** An example is the average rating of videos on Youtube.com.

6) **Recommendation support systems** [22] are systems that do not automate the recommendation process but support people in sharing recommendations. For example, users of IMBD.com may leave text comments about pros and cons of movies on each movie page. These are also non-personalized recommendations (every user can see the same texts).

Main problems of these recommender systems are [23] [24]: (i) *cold start / long tail problem*: the systems can not recommend items not yet rated. Many items usually have few ratings; (ii) *lack of novelty/serendipity*: content-based filtering recommends similar items (not likely a radical change) while collaborative filtering recommends very popular items to anyone; (iii) *start-up / new user problem*: new users of a system have few or no ratings of items. Therefore both content-based systems and collaborative filtering systems cannot find similar items or similar users for the production of recommendations.

Social Network-based Algorithms

The main idea of network-based recommender systems [8], [13], [14], [15], [16] is that items are recommended not because of their features or because people similar to the focal user liked them but because people in the user social network (i.e. friends or colleagues) voted them positively. In other words, if person A trusts person B and B trusts person C, it is likely that items liked by C can be recommended to A, even if A and C are not directly connected. We can call this logic, *social network filtering*.

[15] tested a social network filtering algorithm using data from Epinions.com, the only available dataset that combines explicit user ratings and social networks. On this website, consumers can rate different types of goods (cars, movies, books, music, computers, software, ect.). Reviewers can also rate positively other reviewers if they found their ratings valuable. In this model, a focal user trusts both the people he/she rated positively and the users that were rated positively by his/her trusted users. Trust propagates with a limitation: the further away the user is from current user, the less reliable is the inferred trust value. This measure of trust was used to create a trust network for the production of trust-based recommendations.

In the same way, *TrustWebRank* [16] is a metric inspired by Google's algorithm PageRank which has been tested on the same Epinions data. The metric computes user trust based on the centrality of the user's connections in the trust network. A model simulates also the dynamics of trust in the network.

Finally, C-IKNOW [14] provides recommendations based on similarity, heterogeneity and Exponential Random Graph Models (ERGM). ERGM models compute the probability of a link between each pair of nodes based on the structural tendencies, such as transitivity or reciprocity [25]. The idea of C-IKNOW is that if there is a probability of a link between user A and item B, then B is recommended to A.. Also, C-IKNOW includes a recommender system for the creation of teams called Team Assembly. Recommendations are the grouping options and may be based on potential team member similarity or on existing social relationships.

This new logic potentially improves recommendations by reducing some of the computational limitations of traditional recommender systems:

(i) *Cold start / long tail*: if an item has not been rated (explicitly or implicitly) it can not be recommended in any system. However, network-based recommender systems, rather than other systems, may include more items with fewer ratings in their recommendations if these items have been rated by trusted people. This hypothesis needs to be tested in the future.

(ii) *Lack of novelty/serendipity*: as demonstrated by [16] through simulations, social-network based systems give users the possibility to get recommendations from unknown users (trusted by trusted users) which may bring novel items in the recommendation list, "e.g. recommendations on travel books for people usually interested in tools for gardening" [16].

(iii) *Start-up/new user problem*: [15] demonstrated through experiments that a trust network-based recommender system has more coverage (number of predictable ratings) and a decreased error than a collaborative filtering system. They argue that the network-based system is more efficient because collecting few trust statements from new users is more useful than collecting an equivalent amount of item ratings.

Social Network Theories and Recommender Systems

The rise of "collaborative filters" and social-network based collaborative filters constitute a strong call for the involvement of social scientists in the field of recommender systems.

Table 1

Social network-based theories [26] [27]	**Recommender systems**
Homophily theory: Individuals establish relations with people who are similar to them (e.g. same age, gender, education, etc.)	Demographic filtering: similar users on MySpace.com, Meetic.it, etc. Collaborative filtering: similar users in terms of tastes influence each others.
Balance theory: people prefer to build balanced relationships to avoid discomfort (reciprocity: if A is friend to B, B is friend to A; transitivity: if A is friend to B, and B to C, then A to C)	Social network filtering: Facebook "People you may know" transitivity algorithm.
Proximity theory: people physically close to each other are more likely to connect	Demographic and content-based filtering: geography of users (LinkedIn.com,) or items (restaurants in 2spaghi.it)
Heterophily: or the love of the different. People who differ on certain features tend to group together	Demographic filtering: recommendations are based on demographic data on MySpace.com, Meetic.it, etc.
Collective action theory: people group together to achieve results otherwise unachievable	Social network filtering: no available examples
Social contagion theory: people choose items chosen by people in their social / trust network	Social network filtering: Facebook's "ILike" ("your friends like this"), TrustWebRank metric
Self interest: people choose items on the basis of the balance between benefits and costs associated to those items	Recommendation support systems: text comments on pros and cons of items
Transactive Memory Theory: people connect to those whom they recognize experts or to those items they think may be informative.	Social network filtering: LinkedIn Best Answer, C-IKNOW [14]
Structural hole theory: people connect to non-connected others in order to enhance their structural autonomy	Social network filtering: no available examples
Structural equivalence: people connect to people connecting to the same people	Item-by-item collaborative filtering: Amazon's People who bought this book bought also this [20]

Source: our own elaboration

The literature [26] [27] identifies many theoretical mechanisms explaining why people create social networks. These theories can be interpreted also as explanations of why a certain user may need to connect to a certain "item" (be it a human being or a thing). In table 1 we try to match those theories, briefly summarized, with the types of recommender algorithms. The purpose of this table is to understand whether or not main social theories explaining the emergence of links between people are taken into account by current recommender algorithms. For each algorithm type we provide an example among existing systems.

Some theoretical mechanisms remain unexplored: *collective action theory* [26] could inspire algorithms recommending complementary items or people needed to accomplish a task; algorithms based on *structural holes theory* [28] would recommend to connect with items and people disconnected among each others.

Some other theory, like *self-interest theory* and *transactive memory theory* (TMT) [29] could be better exploited. The first one usually is not automated in recommender systems. The balance between benefits and costs is provided by users through text comments. The second one, TMT, has already some applications among recommender systems but of limited extent.

TMT argues that knowledge exchange occurs because individuals serve as external memory aids to each other [29]. People benefit from each other's knowledge and expertise when they develop a good, shared understanding of who knows what in a group of people. The difficulty in including this mechanism in a recommender system actually relies on the fact that it is challenging to obtain a cognitive measure of other's expertise.

One application of TMT is the LinkedIn's service "Best Answer": people ask questions on categorized topics and rate as "best answer" one of the answers. Users who produced many "best answers" on a certain topic are recognized as experts on that topic. However, LinkedIn does not include these data in the "people search" engine, thus making this feature substantially ineffective. IKNOW [13] and C-IKNOW [14] can be used to include self-reported data about the cognitive measures of other's expertise, thus providing the necessary data for the recommendation of the experts. However, a system including an explicit rating of others' expertise is unlikely to be scalable. Therefore a future TMT-based recommender system may benefit from an implicit expertise rating system that could be obtained by crossing behavioral data such as citation of people, referrals, or other types of endorsements.

Conclusions

The paper had the objective to link social network theories with recommender system research. The review identified which mechanisms are utilized and under-utilized for the design of recommender systems. The author believes that under-utilized mechanisms could increase the value of future recommender systems, especially in organizational settings. Collective action [26] may be a valuable criterion for the recommendation of people and items because it provides information about interdependent resources. Structural hole theory [28] may be useful to identify sources of potential lack of coordination. Finally, transactive memory theory [29] would keep track of cognitive networks about "who knows what" and "who knows who knows what" which are important for the development of knowledge exchange and expert recognition [26].

In general social network theories may provide great value to all the types of recommender systems as the different algorithms still rely on networks of users and items. Thus the "wisdom of networks" can be applied even in relation to non-human nodes.

Future research may be devoted to identifying the accuracy measures of these under-utilized mechanisms in order to predict their ability in identifying useful and relevant items, their learning capacity (what is the minimum density of data they need to produce recommendations), and the novelty of their predictions (whether or not the recommendations would have been discovered by the users without the system).

References

1. O'Reilly, T. (2005) What is Web 2.0? Retrieved May 18, 2008, from http://www.oreillynet.com/pub/a/oreilly/tim/news/2005/09/30/what-is-web-20.html
2. Surowiecki, J. (2004) *The wisdom of crowds: why the many are smarter than the few and how collective wisdom shapes business, economies, societies, and nations*, New York: Doubleday
3. Schafer, J. B., Konstan, J. A., & Riedl, J. (2001) E-Commerce Recommandation Applications. *Data Mining and Knowledge Discovery*, 5(1 - 2): 115-153.
4. Mirza, B. J. (2001) Jumping Connections: A Graph-Theoretic Model for Recommender Systems. *Unpublished Master of Science Thesis*, Virginia Polytechnic Institute and State University.
5. Schafer, J. B., Konstan, J. A., & Riedl, J. (1999) Recommender Systems in E-Commerce. In *Proceedings of the 1st Acm Conference on Electronic Commerce* (pp. 158-166). Denver, Colorado, USA: ACM Press.
6. Kangas, S. (2002) Collaborative Filtering and Recommendation Systems. *Research Report No. TTE4-2001-35*, Espoo, Finland: VTT Information Technology.
7. Resnick, P., Iacovou, N., Suchak, M., Bergstrom, P., & Riedl, J. (1994). Grouplens: An Open Architecture for Collaborative Filtering of Netnews. In *Proceedings of the ACM 1994 Conference on Computer Supported Cooperative Work*, Oct 22-26 1994 (pp. 175-186). New York, NY, USA: ACM Press.
8. McDonald, D. W. (2001). Evaluating Expertise Recommendations. In *Proceedings of the 2001 International ACM Siggroup Conference on Supporting Group Work* (pp. 214-223). Boulder, Colorado, USA ACM Press.
9. McNee, S.M., Albert, I, Cosley, D., Gopalkrishnan, P., Lam, S.K., Rashid, A.M., Konstan, J.A., Riedl, J. (2002). On the Recommending of Citations for Research Papers. In *Proceedings of ACM 2002 Conference on Computer Supported Cooperative Work* (Vol. 116-125). New Orleans, USA: ACM Press.
10. Terry, D. B. (1993). A Tour through Tapestry. In *Proceedings of the Conference on Organizational Computing Systems* (pp. 21-30). Cambridge, MA, USA: ACM Press.
11. Ackerman, M. S., Pipek, V., Wulf, V. (Eds.) (2003) *Expertise Sharing: Beyond Knowledge Management,* Cambridge MA, MIT Press.

12. McDonald, D. W., & Ackerman, M. S. (2000). Expertise Recommender: A Flexible Recommendation System and Architecture. In *Proceedings of the 2000 ACM Conference on Computer Supported Cooperative Work* (pp. 231-240). Philadelphia, Pennsylvania, USA ACM Press.
13. Contractor, N. S., Zink, D., & Chan, M. (1998). Iknow: A Tool to Assist and Study the Creation, Maintenance, and Dissolution of Knowledge Networks. In T. Ishida (Ed.), *Community Computing and Support Systems* (pp. 201-217). Berlin: Springer-Verlag.
14. Johnson, Z. (2010) C-IKNOW Complete Documentation & Workflow, *SONIC Laboratory at Northwestern University, accessed on June 1st 2010 on* http://sonic.northwestern.edu/ciknowdocs/ciknow-guide-all.pdf
15. Massa, P. and Avesani, P. (2004). Trust-Aware Collaborative Filtering for Recommender Systems, Lecture Notes in Computer Science, 3290, 492-508.
16. Walter, F.E., Battiston, S., Schweitzer, F. (2009) Personalised and Dynamic Trust in Social Networks, *RecSys '09: Proceedings of the third ACM Conference on Recommender Systems*, 2009, 197-204
17. Adomavicius, G., & Tuzhilin, A. (2005). Toward the Next Generation of Recommender Systems: A Survey of the State-of-the-Art and Possible Extensions. *IEEE Transactions on Knowledge and Data Engineering*, 17(6), 734- 749.
18. Shardanand, U., & Maes, P. (1995). Social Information Filtering: Algorithms for Automating “Word of Mouth”. *In Proceedings of the Sigchi Conference on Human Factors in Computing Systems* (pp. 210-217). Denver, Colorado, USA: ACM Press/Addison-Wesley Publishing Co.
19. Sarwar, B., Karypis, G., Konstan, J., & Reidl, J. (2001). Item-Based Collaborative Filtering Recommendation Algorithms. In Proceedings of the 10th International Conference on World Wide Web (pp. 285-295). Hong Kong, Hong Kong ACM Press.
20. Linden, G., Smith, B., & York, J. (2003). Amazon.Com Recommendations: Item-to-Item Collaborative Filtering. IEEE Internet Computing, 7(1), 76-80.
21. Montaner, M., Lopez, B., & de la Rosa, J. L. (2003) A Taxonomy of Recommender Agents on the Internet. *Artificial Intelligence Review*, 19(4), 285-330
22. Terveen, L., & Hill, W. (2001) Beyond Recommender Systems: Helping People Help Each Other, in J. Carroll (Ed.), *HCI in the New Millennium*. Reading, Massachusetts: Addison-Wesley.
23. Sarwar, B. M., Konstan, J. A., Borchers, A., Herlocker, J., Miller, B., & Riedl, J. (1998) *Using Filtering Agents to Improve Prediction Quality in the Grouplens Research Collaborative Filtering System*, Seattle, Washington, United States: ACM Press.
24. Breese, J., Heckerman, D., & Kadie, C. (1998) Empirical Analysis of Predictive Algorithms for Collaborative Filtering, in *Proceedings of the Fourteenth Conference on Uncertainty in Artificial Intelligence* (pp. 43-52). San Francisco, CA, USA: Morgan Kaufman.
25. Snijders, T.A.B., Pattison, P.E., Robins, G.L., Handcock, M.S. (2004) New specifications for exponential random graph models, *Working Paper no. 42,* Center for Statistics and the Social Sciences University of Washington, accessed on June 1st 2010 on http://www.csss.washington.edu/Papers/wp42.pdf
26. Monge, P. R., & Contractor, N. S. (2003) *Theories of Communication Networks*: Oxford University Press.
27. Kilduff, M. and Tsai, W. (2003) *Social networks and organizations*. London: Sage
28. Burt, R. S. (1992) *Structural Holes: The Social Structure of Competition*, Cambridge, MA, Harvard University Press.
29. Wegner, D. M. (1995). A Computer Network Model of Human Transactive Memory. *Social Cognition*, 13(3), 319-339.