Artificial Intelligence Review manuscript No.
(will be inserted by the editor)

The State-of-the-Art in Personalized Recommender Systems for Social Networking

Xujuan Zhou, Yue Xu, Yuefeng Li, Audun Josang and Clive Cox

Received: date / Accepted: date

Abstract With the explosion of Web 2.0 application such as blogs, social and professional networks, and various other types of social media, the rich online information and various new sources of knowledge flood users and hence pose a great challenge in terms of information overload. It is critical to use intelligent agent software systems to assist users in finding the right information from an abundance of Web data. Recommender systems can help users deal with information overload problem efficiently by suggesting items (e.g., information and products) that match users' personal interests. The recommender technology has been successfully employed in many applications such as recommending films, music, books, etc. The purpose of this report is to give an overview of existing technologies for building personalized recommender systems in social networking environment, to propose a research direction for addressing user profiling and cold start problems by exploiting user-generated content newly available in Web 2.0.

Keywords Social Networking · Recommender Systems · Trust · User Profiles · User Generated Content

1 Introduction

With the explosion of Web 2.0 applications such as blogs, discussion forums, social and professional networks, and various other types of social media, the users online activities have been changed. Rosa *et.al* [2007] described the change in their report: "Online activities can no longer be characterized by just searching or browsing. Usage is evolving to interacting, and quickly to creating and sharing content."

Xujuan Zhou Yue Xu Yuefeng Li
Faculty of Science and Technology, Queensland University of Technology, Australia
Tel.: +61-07-31387758
E-mail: x.zhou, yue.xu, y2.li@qut.edu.au

Audun Josang
UNIK Graduate Center, University of Oslo, Norway
E-mail: josang@unik.no

Clive Cox
Rummble Limited, UK.

The new generation of Web applications are no longer read only. The Web users are no longer mere consumers of information, but the "producers of information". They actively participate in social networks, upload their personal photos, share their bookmarks, write blogs, and annotate and comment on the information provided by others. They create information, build content and establish online communities. They not only contribute information, but also contribute "themselves", creating detailed personal profiles on social sites and sharing that information to establish new relationships with hundreds of new virtual friends.

The millions of users spend hours daily in these sites, and generate rich information and various new sources of knowledge that has not been available before. The abundance and popularity of social networking sites flood users with huge volumes of information and hence pose a great challenge in terms of information overload and also create many new research issues such as social web storage, search and mining, social network construction, expertise oriented search and association search in social networks, etc. It also introduces many real-world applications. For examples, web community detection and search, hot-topic detection in a specific web community, accurate and timely recommendations in commercial applications. These research issues have been receiving growing attentions in recommender systems field, data mining and among others in the recent years. Many significant researches have been done on these research topics. The main purpose of this document is to provide a survey of the development on these research challenges.

The rest of the report is organized as follows. The Section 2 begins with a brief introduction to social networking to provide the foundations of the Web 2.0 concept. The Background to social networking is given first, then review some well-known social networking sites. In Section 3, we review the origins of recommender systems and explore the use of Web 2.0 technologies for improving the recommender systems in social networking environment. Section 4 presents overview of personalized social recommender systems and some live applications such as tagging systems, blog mining and trust-based recommender systems. Section 5 concludes this study.

2 Social Networking

Although the term social networking is being used in new ways since the availability of the digital medium the concepts behind it are not new. The modern digital medium technology makes sharing contents, collaborating with others, and connecting with each other to create a community faster, easier and more accessible to a wider population than ever before.

2.1 Definition of Social Networking

According to the free on-line dictionary of computing [1], a social network is " a Web site where one connects with those sharing personal or professional interests, place of origin, education at a particular school, etc.". In this report, a social network is defined as "a web site that facilitates meeting people, finding like minds, communicating and sharing content, and building community ". The terms of the social network and social networking are used interchangeably in this report.

Generally, social networks are used to allow or encourage various types of activity whether commercial, social or some combination of the two. Today, People engage in a

[1] http://dictionary.reference.com/browse/social+network

growing variety and number of Web activities on social websites, from buying on commercial sites, to blogging, to online dating, to post personal information profiles. A user's personal profiles are unique pages where one can express their thoughts and feelings, post photographs and show off their network of friends. The most popular social networking websites put a strong emphasis on the user's profiles. The users are encouraged to upload a profile photo, to enhance their profiles by adding multimedia content.

Social networks vary greatly in their features which include music sections, video uploads, groups, games and more. Some have photo sharing or video-sharing capabilities; Others have built-in blogging and instant messages technologies. On May 24, 2007, Facebook launched the Facebook Platform which provides a framework for software developers to create applications that interact with core Facebook features. Applications that have been created on the Platform include lots of games which allow users to play games with their friends.

In a nutshell, social networking sites are the epitome of Web 2.0, in which the network of users is the platform and the community drives the content. In these communities, like-minded individuals can share information, knowledge and interests and provide feedback and reviews. Such sites can act as collaborative platforms, allowing entire networks to grow in value as the user base increases. Furthermore, businesses can focus their marketing efforts on those who are truly interested.

2.2 Brief History of Social Networking

The social networking began with people wanting to reconnect with lost school friends. Classmates.com (created in 1995) is often considered the first social networking site. It focuses on ties with former school mates. It had approximately 40 million members in June 2007, according to the Classates.com site. Other social networking sites quickly followed. SixDegrees.com was launched in 1997. It promoted itself as a tool to help people connect with and send messages to others. While SixDegrees attracted millions of users, it failed to become a sustainable business. In 2000, the service closed. Epinions.com developed a trust-based social networking in 1999. The Epinions has a Web of Trust scheme. Members of Epinions can decide to either trust or block (distrust) another member. A member's list of trusted members represents that member's personal Web of Trust.

The next wave of social networking sites embark on when the business social networking site Ryze was launched by entrepreneur Adrian Scott in 2001.Business social networking can create a pool of contacts from which you can draw leads, referrals, ideas, and information for your job search and career progress. Later the more popular and successful business social networking site Friendster launched in 2002. Social networking began to flourish as a component of business internet strategy at around 2003 when LinkedIn and Xing launched in 2003.

Linkedin is a business-oriented social network dedicated to helping professionals maintain a list of connections. From its birth, LinkedIn was geared more to fostering network connections within the business community and it's common for entrepreneurs and corporate managers alike to seek out partnership opportunities for their business. Through the business social network, the job seekers can find out more about potential employers, and people in job recruitment or human resource positions can find out the potential employees to fill job vacancies.

XING is a social network that powers business professionals by giving them the tools to tap into the vast resources of their own personal network to open doors to thousands of

companies, find jobs and search for employees. With over six million business professionals using XING each day in any one of 16 different languages, XING is a worldwide leader in business social networking.

Nowadays, the social networking sites MySpace launched in 2003, Facebook in 2004, Bebo in 2005, and Twitter in 2006 emerged as the most popular sites in the world. MySpace is a social networking website targeted at a general audience. It became one of the most visited websites in the world within a few years. In 2006, Facebook opened to every one. Since then, it has experienced tremendous growth ranking only behind MySpace among social networks. However, April 2008 was the milestone: Facebook officially caught up to MySpace in terms of unique monthly worldwide visitors, according to data released by Comscore [2].

2.3 In Summary

Social networking sites (SNS) are a type of virtual community that has grown tremendously in popularity over the past few years. Through SNSs, users connect with each other, share and find content, and disseminate information.

When people join social networking sites, they begin by creating a profile, then make connections to existing friends as well as those they meet through the site. User profile is a backbone of SNS which is a list of identifying information. It can include your real name, or a pseudonym. It also can include photographs, birthday, hometown, religion, ethnicity, and personal interest. Members connect to others by sending a "friend" message, which must be accepted by the other party in order to establish a link. Users can adjust their privacy settings as well.

Given the success of item recommendation systems in commercial websites, such as Amazon.com and Netflix, it is considered worthwhile to revisit the recommendation problem through the novel perspective of social networking. In general, recommendation systems aim to provide personalised recommendations of items to users based on their previous behaviour as well as on other information gathered by item descriptions and user profiles. However, no emphasis has been placed yet on personalisation based explicitly on social networks.

3 User Profiling for Recommender Systems

Recommender systems have become an essential research area as an important response to the so-called information overload problem since the appearance of the first papers on collaborative filtering in the mid-1990s [Resnick et al 1994]. The essential recommendation making mechanism of current recommender systems is to firstly identify the target user's neighbours based on user profile similarity, and then suggest the target user items that the neighbours have liked in the past. User profiling is the process of acquiring, extracting and representing the features of users. The profile data may include users' selected items, ratings for specific items, and demographic data etc. User profiling is one of most challenging tasks. In current recommender systems, user profiles are usually generated based on data with limited relevance that are too simple to produce quality recommendations [Adomavicius and Tuzhilin 2005].

[2] http://www.techcrunch.com/2008/06/12/facebookno-longer-the-second-largest-social-network/, June 12 2008

Massive quantities of User Generated Content (UGC) on social networks are now available from blogs, tags, item reviews, knowledge-sharing sites, collaborative filtering systems, online gaming, newsgroups, chat rooms, etc. A warehouse of UGC can be mined and analyzed to expand user profiles based on which more reliable recommendations can be made to users. The richness of the online UGC challenges the current personalization techniques and also provides new possibilities for accurately profiling users. Thus, how to incorporate the new features and practices of Web 2.0 into personalized recommender applications becomes an important and urgent research topic.

In this section, we review traditional recommender systems first and then articulate some problems of current recommender systems, and also identify some new challenges to recommender systems in social network environment.

3.1 Traditional Recommender Systems

Recommender systems combine ideas from user profiling, information filtering and machine learning to deliver users a more intelligent and proactive information service by making concrete product or service recommendations that match their learned user preferences and needs. The recommender technology is superior to other information filtering applications because of its ability to provide personalized and meaningful information recommendations. For example, while standard search engines are very likely to generate the same results to different users entering identical search queries, recommender systems are able to generate results to each user that are personalized and more relevant because they take into account each user's personal interests.

In general, two recommendation techniques have come to dominate: content-based filtering (CBF), collaborative filtering (CF). The content-based approach [Mooney and Roy 2002] recommends a user to items whose content is similar to content that the user has previously viewed or selected. In a movie recommender application, for instance, a CBF system will typically rely on information such as genre, actors, director, producer etc. and match this against the learned preferences of the user in order to select a set of promising movie recommendations. CBF recommender systems need a technique to represent the features of the items. Feature representation can be created automatically for machine-parsable items (such as news or papers) but must be manually inserted by human editors for items that are not yet machine-parsable (such as movies and songs). Obviously this activity is expensive, time consuming, error-prone and highly subjective. Moreover, for some items such as jokes, it is almost impossible to define the right set of describing features and to "objectively" classify them [Massa and Bhattacharjee 2004].

Collaborative filtering (CF) collects information about a user by asking them to rate items and makes recommendations based on highly rated items by users with similar taste. CF approaches make recommendations based on the ratings of items by a set of users (neighbours) whose rating profiles are most similar to that of the target user [Breese et al 1998]. CF algorithms generally compute the overall similarity or correlation between users, and use that as a weight when making recommendations. In a book recommendation application, for example, the first step for the CF system is try to find the "neighbours" of the target user. The "neighbours" refer to other users who have similar tastes in books (rate the same books similarly). In the second step, only the books that are highly rated by the "neighbours" would be recommended.

In contrast with the content-base approaches, the CF techniques rely on the availability of user profiles that capture the past ratings histories of users [Breese et al 1998] and

don't require any human intervention for tagging content because item knowledge is not required. Therefore, the CF techniques can be applied to virtually any kind of items: papers, news, web sites, movies, songs, books, jokes, locations of holidays, stocks and promise to scale well to large item bases [Massa and Bhattacharjee 2004]. Collaborative filtering is the most widely used approach to build online recommender systems. It has been successfully employed in many applications, such as recommending books, CDs, and other products at Amazon.com, Movies by MovieLens [Adomavicius and Tuzhilin 2005]. Some methods combine both content and collaborative filtering approaches to make recommendations [Schein et al 2002].

3.2 User Profiles Learning Issues in Recommender Systems

A crucial factor for the success of recommender systems is the availability and quality of the user profiles. The user profile information can be input explicitly by users or implicitly gathered by software agents that monitor user activity [Gauch et al 2007]. For explicit acquisition, users are required to rate or select items. For implicit acquisition, the users' behaviours will be passively observed as they interact with the system and then the users' interests will be inferred from these interactions. Currently the user profile information for online recommendation is mainly obtained by analyzing usage log data such as users' click streams and navigation patterns etc. Both the explicit and implicit methods have their respective strengths and weaknesses. The explicit acquisition is more accurate, because information comes directly from the users when a user rates the relevance of a set of items. However, it may place an increased cognitive burden on the users [Morita and Shinoda 1994]. The implicit acquisition places little or no burden on the users. However, inferences drawn from the user interaction are not always valid because of the indicators of the user interests are often erratic [Kelly and Teevan 2003].

The user profiles are often difficult to obtain and their quality is also hard to ensure. Current existing user profiling for recommender systems is mainly using user rating data. Usually, hundreds of thousands of users and items are involved in a recommender system, but only a few items are viewed, selected or rated by users. As Sarwar et.al reported in [Sarwar et al 2001], the density of the available ratings in commercial recommender systems is often less than 1%. Moreover, as for new users, they will start with a blank profile without selecting or rating any items at all. These situations are commonly referred to as the data **sparseness** and **cold start** problem [Schein et al 2002]. The current recommender algorithms are impeded by the sparsity and cold start problems.

With the increasing use of recommender systems in e-commerce and social networks, maliciously or unfairly influences to the outcomes of recommender systems by creating false user rating data are also intensified. For example: a simple but effective attack to recommender system is to deliberately create a bunch of fake users with pseudo ratings favour or disfavour to some particular products. With the fake information, user profile data becomes unreal and not reliable.

In summary, without sufficient knowledge about users, even the most sophisticated recommendation strategy will not be able to make satisfactory recommendations. The **cold start**, **sparseness**, **malicious rating** are formidable problems for user profiling. They cause user profiles to become the weakest link in the whole recommendation process.

3.3 Trust Issues in Recommender Systems

Traditional recommender systems purely mine the user-item rating matrix for making recommendations. However, recommendations are not made in rational isolation, which means that they are not evaluated merely by their information value [Perugini et al 2004]. The social embedding of a recommendation is crucial to understanding the decision making process of an individual; it is determined by factors such as experience, background, knowledge level, beliefs and personal preferences [Lueg 1997]. It was found in paper [Sinha et al 2001] that given a choice between recommendations from friends and recommender systems, in terms of quality and usefulness, friends' recommendations are preferred even though the recommendations given by the recommender systems have high novelty factor. Friends are seen as more qualified to make good and useful recommendations as compared to recommender systems. In many markets, people typically trust and act on recommendations from friends more than from the company selling the product. A recent study found that positive word of mouth [Shardanand and Maes 1995] among customers is by far the best predictor of a company's growth. Word-of-mouth marketing has the key advantage that a recommendation from a friend or other trusted source has the credibility that advertisements lack.

In general, a user is much more likely to believe statements from a trusted acquaintance than from a stranger. The possibility of people taking recommendations made by trusted friends is higher than that of taking recommendations made by strangers. However, current recommendation techniques make recommendations to a target user mainly based on other users' item preference, these users have similar rating data with the target user, but the trust between users has not been well exploited. Therefore, another big challenge for traditional recommender systems is how to embed the social elements of decision-making and advice seeking or the trust relations among users.

4 Social Recommender Systems

There has been a tremendous increase in user-generated content (also referred to as user-created content) in the past a few years via the technologies of Web 2.0. It is now well recognized that the user-generated content (e.g., product reviews, tags, forum discussions and blogs) contains valuable user opinions that can be exploited for many applications. By exploiting the UGC more effectively via the use of the latest collaborative filtering, data mining techniques, and trust management technology, more accurate and sophisticate user profiles can be built which contain not only users' item preferences (i.e., item ratings) but also users' topic interests and trustworthiness between users. Based on the enhanced user profiles, high quality and reliable recommendations can be generated. Many significant researches have been done to investigate new strategies available in Web 2.0 framework. In this section, we review some new strategies for social recommender systems.

4.1 User Generated Content

Unlike the user rating data which is numeric data, the UGC comprises various forms of media and creative works such as written, audio, visual, and combined created by users explicitly and pro-actively. Therefore, it contains rich semantic information and provides a huge potential to obtain deeper knowledge about users, items, and the various relationships among users and items. It has become an important information resource in addition

to traditional website materials. From the UGC information, it is possible to acquire users' opinions, perspectives, or tastes towards items or other users. The growing and readily available user-generated content is rising the new opportunity to construct user profiles accurately compared with the existing personalized recommender techniques and to mitigate the cold start and malicious rating problems considerably.

The UGC expresses users' opinions or sentiments towards items and is transforming how people seek advice and consider recommendations. The opinion mining and sentiment analysis such as customer opinion summarization [Zhuang et al 2006] and sentiment analysis of user reviews [Ding et al 2008] are possibly as augmentations to recommendation systems [Tatemura 2000], since it might behoove such a system not to recommend items that receive a lot of negative feedback.

The individual users show their interest in online opinions about products or services. They share their brand experiences and opinions, positive or negative, regarding any product or service. The vendors of these items are increasingly coming to realize that these consumer voices can potentially wield enormous influence in shaping the opinions of other consumers and they are paying more and more attention to these issues [Hoffman 2008]. There are already many companies that provide opinion mining services and examples include Epinions.com [3], Amazon.com [4].

4.2 Tagging Systems

A tag is a keyword that is added to a digital object (e.g. a website, picture or video clip) to describe it, but not as part of a formal classification system. Tags are freely chosen keywords and they are a simple yet powerful tool for organizing, searching and exploring the resources. Web-based tagging systems such as Del.icio.us, Technorati, Flickr, Last.fm, or citeulike have become increasingly popular. These systems enable users to assign tags to Internet resources (e.g., web pages, images, videos,songs, blogs, urls, scientific papers) without relying on a controlled vocabulary and allow users to share their tags for particular resources. In addition, each tag serves as a link to additional resources tagged the same way by others.

One of the first large-scale applications of tagging was the del.icio.us website, which launched the 'social bookmarking' phenomenon. The services like Flickr (photos), YouTube (video) and Odeo (podcasts) allow a variety of digital artefacts to be socially tagged. CiteULike is a free service to help academics to store, organise and share the academic papers they are reading. When you see a paper on the Web that interests you, you click a button and add it to your personal library. CiteULike automatically extracts the citation details, so you do not have to type them in.

Heymann et al. [2008] conducted an analysis study of social tagging on the popular del.icio.us bookmarking system. The results have shown that searches on del.icio.us can be improved by a navigable hierarchical taxonomy of tags which is derived from tag usage. The taxonomy of tags is used to help users broadening/narrowing the set of tags that best describe their interests.

Halpin et al. [2007] study the distribution of tags in the social bookmarking site del.icio.us and propose a generative model of collaborative tagging in order to evaluate the dynamics that lie beneath the act of collaborative recommendation. Their findings prove that the dataset collected follows a power-law distribution.

[3] http://www.epinions.com/

[4] https://www.amazon.com/

Like other UGC information, the tag information is becoming an important information source to profile user's topic interests as well as to describe the content or classification of items. Compared with other traditional implicit user information such as click stream and web log, the tag information has some distinctive advantages. One important advantage is that tags are pieces of light weighted textural information but contain very rich and explicit topic information since they are given by users explicitly and proactively. Another important advantage is that it is independent with the content of the items, which makes it possible to do content filtering for any items such as videos, music files etc. Moreover, the tagging behavior forms a three dimensional relationship among users, items and tags such as the additional implied item-tag, user-tag besides the typical implicit user-item relationship.

However, since there is no restriction or boundary on selecting words for tagging items, the tags used by users are free-style and contain a lot of noise such as semantic ambiguity which means that the same tag name has different meaning for different users, tag synonym which means that different tags actually have the same meaning. Another serious situation of tags is that nearly 60% tags are personal tags that are only used by one user [Sen et al 2009]. All these disadvantages of tags bring challenges to make use of tags to profile users' topic preferences accurately or describe the topics of the items correctly. Thus, how to solve these problems caused by the free-style vocabulary of tags is a key issue to improve the accuracy of recommendation systems based on tag information.

The work of Tso-shuter et al. [2008] extended the user-item matrix to user-item-tag matrix to make collaborative filtering item recommendation. However, this work didn't consider the noise of tags. More recently, the noise of tags has become an important research question. In the recent work of [Sen et al 2009], a special tag rating function was used to find user's preferences for tags. Along with the tag preferences, the click streams, tag search history of each user were used to get user's preferences for items through the inferred tags preferences. However, Sen' work needs various kinds of extra information or special function, which makes the work incomparable and gives restrictions to the application of the work. Moreover, it is difficult to determine the influence of tag information when the click streams, search queries were combined together.

Different from Sen's work the approach proposed in [Liang et al 2009] makes use of the standard item taxonomy or ontology given by experts to represent each user's tag individually to remove the noise of tags. Item taxonomy is a set of controlled vocabulary terms or topics designed to describe or classify items, which is available for various domains, for example, the book classification taxonomy of Amazon.com, world knowledge ontology such as Library of Congress Subjects Headings. Because item taxonomy is usually designed and developed by experts, reflecting the common views to the description and classification of items, providing not only a standard vocabulary but also a hierarchical structure to represent the relationships among concepts or categories, it can be used to eliminate the inaccuracy caused by the users' free-style vocabulary in social tags.

4.3 Blogs Mining

The term web-log, or blog refers to a simple webpage consisting of brief paragraphs of opinion, information, personal diary entries, or links, called posts, arranged chronologically with the most recent first, in the style of an online journal. Most blogs also allow visitors to add a comment below a blog entry. People express their opinions, ideas, experiences, thoughts, wishes through these free-form writings. A typical blog post can combine text, images, and links to other blogs, web pages, and other media related to its topic. The individuals who

author the blog posts are referred as bloggers. The universe of all these blog sites is often referred as Blogosphere [Stewart et al 2007]. Linking is also an important aspect of blogging as it deepens the conversational nature of the blogosphere and its sense of immediacy. It also helps to facilitate retrieval and referencing of information on different blogs.

Blogs notoriously contain quite a bit of subjective content. General topics include personal diaries, experiences, opinions, information technology, and politics to name a few. Thus blogs are more relevant than shopping sites for queries that concern politics, people, or other non-products. However, the desired material within blogs can vary quite widely in content, style, presentation. Mining opinions and sentiments from bloggers poses several challenges as compared to the historic feedback and surveys.

State-of-the-art content analysis techniques could be used for basic clustering, classification of the blog posts/blog sites. For example, a prototype system called Pulse [Gamon et al 2005] uses a Naive Bayes classifier trained on manually annotated sentences with positive/negative sentiments and iterates until all unlabeled data is adequately classified.

The researchers Joshi and Belsare [2006] developed a blog mining program called BlogHarvest which searches for, and extracts, a blogger's interests in order to recommend blogs with similar topics. The program uses classification, links, topic similarity clustering and tagging based on opinion mining to provide these features. The program design is based on the knowledge that blogging communities are not formed randomly, but as a result of shared interests. It is also designed to provide a useful search facility to bloggers while generating large amounts of revenue for advertising services and providers.

4.4 Trust in Social Recommender Systems

The term "trust" is being used with a variety of meaning. In order to give the reader a reference point for understanding trust, some general definitions from existing research are given first.

4.4.1 Definition of Trust

Mayer et al. [1995] defined trust as "the willingness of a party to be vulnerable to the actions of another party based on the expectation that the other will perform a particular action important to the trustor, irrespective of the ability to monitor or control that other party". Another definition from Mui et al. [2002] is "trust is a subjective expectation an agent has about another's future behaviour based on the history of their encounters". The definition given by Olmedilla et al. [2005] is:"Trust of a party A to a party B for a service X is the measurable belief of A in that B behaves dependably for a specified period within a specified context (in relation to service X)".

Trust is the outcome of observations leading to the belief that the actions of another may be relied upon. Trust is subjective and personal. Trust is asymmetrical. This means for two people involved in a relationship, trust is not necessarily identical in both directions. Being trustworthy is one step, but that's not the same as being trusted. Trust is dynamic as well. Generally people's trust to others is gradually built up and keeps changing over time. A user's initial trust to another user only reflects his beliefs to the user at a static point in time. Therefore, techniques are needed to automatically construct or update users' trustworthiness based on users' online behaviour. Trust has the property of transitivity since a trusted acquaintance will also trust the beliefs of her friends. Trusts may propagate (with

appropriate discounting) through the relationship network [Guha et al 2004] or infer trust using a trust metric.

4.4.2 Trust-based Social Recommender Systems

There has been a lot of work that deals with trust in social networks and recommender systems. A vast literature on trust has grown in these area of research. Some preliminary work has shown the benefits of utilizing trust information in recommendation making [Avesani et al 2005]; [Golbeck 2006]; [Massa and Avesani 2007]; [Andersen et al 2008]; [Bedi et al 2007]; [Ma et al 2009].

Golbeck and Hendler [2006] considered those social networking sites where users explicitly provide trust ratings to other members. However, for large social networks it is infeasible to assign trust ratings to each and every member so they propose an inferring mechanism which would assign binary trust ratings (trustworthy/non-trustworthy) to those who have not been assigned one. They demonstrate the use of these trust values in an email filtering application and report encouraging results. They also assume three crucial properties of trust for their approach to work: transitivity, asymmetry, and personalization. This is contrary to what was proposed in [Yu and Singh 2003], who assume symmetric trust values in the social network between two members. Also, consolidating the trust scores for a member and computing a global trust score for each member might not give a reasonable estimation. Trust of a member is absolutely a personal opinion. Therefore, authors propose personalization of trust which means that a member could have different trust values with respect to different members.

In paper [Ma et al 2009], it was assumed that every user's decisions on the Web should include both the user's characteristics and the user's trusted friends' recommendations. Under this assumption, the authors proposed a probabilistic matrix factorization framework that employs both the user-item matrix and the users' social trust network for the recommendations. Their experimental results on the Epinions dataset show that their method outperforms the state-of-the-art collaborative filtering and social trust-based recommendation algorithms, especially when the users have very few ratings. The complexity analysis indicates that their approach can be applied to very large datasets. They claimed that their method can help to alleviate the data sparseness problem and enhance the system scalability.

Walter et al. [2008] propose the use of social network information in recommendation systems and analyze the impact of trust dynamics on the performance of such a system. They study the effect of preference heterogeneity of agents and network density on usefulness of trust in the system. The authors take a random directed graph for the underlying social network structure and have considered only 100 agents with a limited number of items in which agents seek recommendations, taking it far from a real-world scenario. The algorithm would not scale well for large networks and large number of items. Moreover, trust according to them is based on past experience of recommendations. Also the paper makes some simplifying assumptions like the social network is static and there are not any malicious entities.

Guha [Guha et al 2004] proposes a propagation model for trust and distrusting social friendship networks based on a series of matrix operations, but works with an overly simplistic trust scale of 1, 0 and -1, which is unable to capture the gradation of interpersonal trust. Massa and Avesani [Massa and Avesani 2007] studied the trust-aware recommender systems. Their work replaces the similarity finding process with the use of a trust metric, which is able to propagate trust over the trust network and to estimate a trust weight. The experiments on a large real dataset shows that this work increases the coverage (number of ratings that are predictable) while not reducing the accuracy (the error of predictions).

Jøsang et al. [Jøsang et al 2006] described a method for trust network analysis using subjective logic (TNA-SL). Their method takes directed trust edges between pairs as input, and can derive a level of trust between arbitrary parties that are interconnected through the network. Even in case of no explicit trust paths between two parties exist, subjective logic allows a level of trust to be derived through the default vacuous opinions. TNA-SL therefore has a general applicability and is suitable for many types of trust networks. However, this method includes the same trust edges multiple times and will produce an inconsistent result. Theoretically they proposed an optimal model to describe how TNA-SL can preserve consistency without removing information in paper [Jøsang and Bhuiyan 2008]. The optimal TNA-SL avoids this problem by allowing the trust measure of a given trust edge to be split into several independent parts, so that each part is taken into account by separate trust paths.

Some researchers have found that given some predefined domain and context, people's interest similarity is a strong predictor of interpersonal trust [Jensen et al 2002]. The relationships between people's interest similarities and trust have been investigated by Ziegler and Golbeck [2005]. Their empirical analysis showed positive mutual interactions between interpersonal trust and interest similarity. That means, people who have similar interests tend to be more trustful towards each other.

In light of these studies, it can be said that the computational trust models can act as appropriate means to supplement current collaborative filtering approaches used by the recommender systems [O'Donovan and Smyth 2005].

5 Conclusion

The purpose of this survey has been to describe and analyse the state-of-the-art of personalized recommender systems in social networking environment, and to convey to the reader a sense of our excitement about the intellectual richness and breadth of the area.

Social Networking Sites contains a warehouse of information that can be mined and analyzed to expand user profiles and to build complex diagrams and maps of user-to-user and user-to-interest relationships. Data mining is an emerging research direction fulfilling various knowledge discovery tasks. Many data mining and opinion mining techniques have been used to underpin an effective recommender system. The applications of Tagging systems, Blogs mining systems, and trust techniques and approaches have been covered in this study.

In the recent years, many research groups have invested much effort on Web personalization, trust and reputation, and recommendations, and have made many great achievements. The rich literature growing around these topics. However, challenging problems still exist in these areas. In particular, the research issues on how to make breakthrough on the current recommender system for social networking environment, how to build the trust-based Web personalized recommender systems have attracted lots of the research attentions. We very much hope we have provided some helpful information to the readers who are encouraged to take up the many challenges that remain in the area.

References

[Adomavicius and Tuzhilin 2005] 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

[Andersen et al 2008] Andersen R, Borgs C, Chayes J, Feige U, Flaxman A, Kalai A, Mirrokni V, Tennenholtz M (2008) Trust-based recommendation systems: an axiomatic approach. In: Proceeding of the 17th international conference on World Wide Web, ACM, New York, NY, USA, WWW '08, pp 199–208

[Avesani et al 2005] Avesani P, Paolo M, Roberto T (2005) A trust-enhanced recommender system application: Moleskiing. In: SAC '05: Proceedings of the 2005 ACM symposium on Applied computing, Santa Fe, New Mexico, ACM, New York, NY, USA, pp 1589–1593

[Bedi et al 2007] Bedi P, Kaur H, Marwaha S (2007) Trust based recommender system for the semantic web. In: IJCAI'07: Proceedings of the 20th international joint conference on Artifical intelligence, Morgan Kaufmann Publishers Inc., San Francisco, CA, USA, pp 2677–2682

[Breese et al 1998] Breese JS, Heckerman D, Kadie C (1998) Empirical analysis of predictive algorithms for collaborative filtering. Morgan Kaufmann, pp 43–52

[Ding et al 2008] Ding X, Liu B, Yu PS (2008) A holistic lexicon-based approach to opinion mining. In: Proceedings of the Conference on Web Search and Web Data Mining (WSDM'08), Palo Alto, California, USA, ACM, New York, NY, USA, pp 231- 240

[Gamon et al 2005] Gamon M, Aue A, Corston-Oliver S, Ringger E (2005) Pulse: Mining customer opinions from free text. pp 121–132

[Gauch et al 2007] Gauch S, Speretta M, Chandramouli A, Micarelli A (2007) User profiles for personalized information access. The Adaptive Web. Volume 4321/2007, pp54-89

[Golbeck 2006] Golbeck J (2006) Generating predictive movie recommendations from trust in social networks. In: Proceccdings of the 4th international conference on trust management, Pisa, Italy, Springer, pp 93-104

[2006a] Golbeck J, Hendler J (2006) Inferring binary trust relationships in web-based social networks. ACM Trans Internet Technol 6(4):497–529

[Guha et al 2004] Guha R, Kumar R, Raghavan P, Tomkins A (2004) Propagation of trust and distrust. In: Proceedings of the 13th international conference on World Wide Web, ACM Press, New York, NY, USA, pp.403–412.

[Hoffman 2008] Hoffman T (2008) Online reputation management is hot – but is it ethical? ComputerWorld, February, 2008.

[2008] Heymann P, Koutrika G, Garcia-Molina H (2008) Can social bookmarking improve web search? In: Proceedings of the international conference on Web search and web data mining. New York, NY, USA: ACM, 2008, pp. 195–206.

[2007] Halpin H, Robu V, Shepherd H (2007) The complex dynamics of collaborative tagging, in : Proceedings of the 16th international conference on World Wide Web. New York, NY, USA. ACM, 2007, pp. 211–220.

[Jensen et al 2002] Jensen C, Davis J, Farnham S (2002) Finding others online: reputation systems for social online spaces. In: CHI '02: Proceedings of the SIGCHI conference on Human factors in computing systems, ACM Press, New York, NY, USA, pp 447–454

[Jøsang and Bhuiyan 2008] Jøsang A, Bhuiyan T (2008) Optimal trust network analysis with subjective logic. In: SECURWARE '08: Proceedings of the 2008 Second International Conference on Emerging Security Information, Systems and Technologies, IEEE Computer Society, Washington, DC, USA, pp 179–184

[Jøsang et al 2006] Jøsang A, Hayward R, Pope S (2006) Trust network analysis with subjective logic. In: ACSC '06: Proceedings of the 29th Australasian Computer Science Conference, Australian Computer Society, Inc., Darlinghurst, Australia, Australia, pp 85–94

[2006] Joshi M, Belsare N (2006) Blogharvest: Blog mining and search framework. In: International Conference on Management of Data, Computer Society of India, Delhi, India, pp 226-229

[Liang et al 2009] Liang H, Xu Y, Li Y, Nayak R, Weng L (2009) Personalized recommender systems integrating social tags and item taxonomy. In: Proceedings of the 2009 IEEE/WIC/ACM International Joint Conference on Web Intelligence and Intelligent Agent Technology, Washington, DC, USA, IEEE Computer Society, pp 540-547

[Kelly and Teevan 2003] Kelly D, Teevan J (2003) Implicit feedback for inferring user preference: a bibliography. SIGIR Forum 37(2):18–28

[Lueg 1997] Lueg C (1997) Social filtering and social reality. In: In Proceedings of the 5th DELOS Workshop on Filtering and Collaborative Filtering, ERCIM Press, pp 10–12

[Ma et al 2009] Ma H, King I, Lyu MR (2009) Learning to recommend with social trust ensemble. In: SIGIR '09: Proceedings of the 32nd international ACM SIGIR conference on Research and development in information retrieval, ACM, New York, NY, USA, pp 203–210

[Massa and Avesani 2007] Massa P, Avesani P (2007) Trust-aware recommender systems. In: RecSys '07: Proceedings of the 2007 ACM conference on Recommender systems, ACM, New York, NY, USA, pp 17–24

[Massa and Bhattacharjee 2004] Massa P, Bhattacharjee B (2004) Using trust in recommender systems: An experimental analysis. In: Proceedings of iTrust2004 International Conference, pp 221–235

[1995] Mayer RC, Davis JH, Schoorman DF (1995) An integrative model of organizational trust. The Academy of Management Review, Vol. 20, No. 3, pp. 709-734

[Mooney and Roy 2002] Mooney RJ, Roy L (2002) Content-based book recommending using learning for text categorization. In: Proceedings of 5th ACM Conference on Digital Libraries. San Antonio, US. ACM Press, New York, US, 2000, pp. 195–204.

[Morita and Shinoda 1994] Morita M, Shinoda Y (1994) Information filtering based on user behavior analysis and best match text retrieval. Proceedings of SIGIR '94 ACM pp 272–281

[2002] Mui L, Mohtashemi M (2002) A computational model of trust and reputation. In: Proceedings of the 35th Hawaii International Conference on System Science (HICSS)

[O'Donovan and Smyth 2005] O'Donovan J, Smyth B (2005) Trust in recommender systems. In: IUI '05: Proceedings of the 10th international conference on Intelligent user interfaces, ACM, New York, NY, USA, pp 167–174

[2005] Olmedilla D, Rana OF, Matthews B, Nejdl W (2005) Security and trust issues in semantic grids. In: Proceedings of the Dagsthul Seminar, Semantic Grid: The Convergence of Technologies, p 05271

[Perugini et al 2004] Perugini S, Gonçalves MA, Fox E (2004) Recommender Systems Research: A Connection-Centric Survey. J. Intell. Inf. Syst., 23(2), Kluwer Academic Publishers, pp107–143

[Resnick et al 1994] Resnick P, Iacovou N, Suchak M, Bergstrom P, Riedl J (1994) Grouplens: An open architecture for collaborative filtering of netnews. In: CSCW, pp 175–186

[2007] Rosa CD, Havens JCA, Hawk J, Jenkins L (2007) Sharing, privacy and trust in our networked world. A report to the oclc membership, OCLC, Dublin, Ohio USA

[Sarwar et al 2001] Sarwar B, Karypis G, Konstan J, Riedl J (2001) Item-based collaborative filtering recommendation algorithms, In: Proceedings of the 10th international conference on World Wide Web, Hong Kong, ACM Press, pp. 285–295.

[Schein et al 2002] Schein AI, Popescul A, Ungar LH, Pennock DM (2002) Methods and metrics for cold-start recommendations. In: SIGIR '02: Proceedings of the 25th annual international ACM SIGIR conference on Research and development in information retrieval, ACM, New York, NY, USA, pp 253–260

[Schein et al 2002] Schein AI, Popescul A, Ungar LH, Pennock DM (2002) Methods and metrics for cold-start recommendations. In: SIGIR '02: Proceedings of the 25th annual international ACM SIGIR conference on Research and development in information retrieval, ACM, New York, NY, USA, pp 253–260

[Shardanand and Maes 1995] Shardanand U, Maes P (1995) Social information filtering: Algorithms for automating "word of mouth". In: CHI, pp 210–217

[Sen et al 2009] Sen S, Vig J, Riedl J (2009) Tagommenders: Connecting Users to Items through Tags. In: Proceedings of the 18th international conference on World wide web, Madrid, Spain, ACM Press, pp 671-680

[Sinha et al 2001] Sinha R, , Sinha R, Swearingen K (2001) Comparing recommendations made by online systems and friends. In: Proceedings of the DELOS-NSF Workshop on Personalization and Recommender Systems in Digital Libraries

[Stewart et al 2007] Stewart A, Chen L, Paiu R, Nejdl W (2007) Discovering information diffusion paths from blogosphere for online advertising. In: ADKDD '07: Proceedings of the 1st international workshop on Data mining and audience intelligence for advertising, ACM, New York, NY, USA, pp 46–54

[Tatemura 2000] Tatemura J (2000) Virtual reviewers for collaborative exploration of movie reviews, In: Proceedings of Intelligent User Interfaces (IUI), 2000, pp. 272–275.

[2008] Tso-Sutter, KHL, Marinho L, Schmidt-Thieme L (2008) Tag-aware recommender systems by fusion of collaborative filtering algorithms. In: Proc. of Applied Computing, pp 1995–1999

[2008] Walter FE, Battiston S, Schweitzer F (2008) A model of a trust-based recommendation system on a social network, Autonomous Agents and Multi-Agent Systems, 16(1), pp 57–74

[Yu and Singh 2003] Yu B, Singh M (2003) Detecting deception in reputation management. In: Proceedings of the second international joint conference on Autonomous agents and multiagent systems. Melbourne, Australia, ACM, pp73–80

[Zhuang et al 2006] Zhuang L, Jing F, Zhu X, Zhang L (2006) Movie review mining and summarization. In: Proceedings of the 15th ACM international conference on Information and knowledge management, New York, NY, USA, ACM, pp43–50

[2005] Ziegler, C.-N., Golbeck, J. (2005). Investigating Correlations of Trust and Interest Similarity - Do Birds of a Feather Really Flock Together? Decision Support Systems, 1-34.