Social Media Recommendation

Zhi Wang, Wenwu Zhu, Peng Cui, Lifeng Sun, and Shiqiang Yang

Abstract Social media recommendation is foreseen to be one of the most important services to recommend personalized contents to users in online social network. It imposes great challenge due to the dynamical behavior of users and the large-scale volumes of contents generated by the users. In this chapter, we first present the principal concept of social media recommendation. Then we present the framework of social media recommendation, with a focus on two important types of recommendations: interest-oriented social media recommendation and influence-oriented social media recommendation. For each case, we present the design of the recommendation that takes both social property and content property into account, such as user relations, content similarities, and propagation patterns. Furthermore, we present theoretical results and observations on the social media recommendation approaches.

1 Social Media Recommendation in Online Social Network

Online social network is attracting more and more people in today's Internet, where users can share and consume all kinds of multimedia contents. Social media sharing is based on online social network, where users are able to reach various contents shared by others. With the exponential growth in social media contents, such as images and videos generated by users, it is of great importance to study how to provide personalized contents in the social media service. Recommendation is foreseen to be one of the most important services that can provide such personalized multimedia contents to users. However, social media recommendation is different

Z. Wang (✉) • W. Zhu • P. Cui • L. Sun • S. Yang
Department of Computer Science and Technology, Tsinghua University, Beijing, 100084, China
e-mail: wangzhi04@mails.tsinghua.edu.cn; wwzhu@tsinghua.edu.cn; cuip@tsinghua.edu.cn; sunlf@tsinghua.edu.cn; yangshq@tsinghua.edu.cn

N. Ramzan et al. (eds.), *Social Media Retrieval*, Computer Communications and Networks, DOI 10.1007/978-1-4471-4555-4_2,

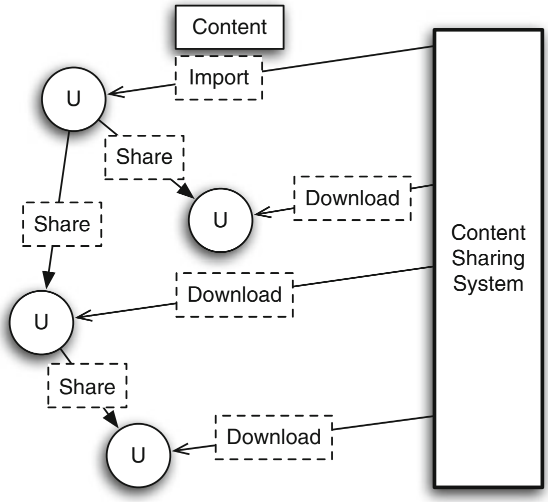


Fig. 1 Contents shared in online social network

from traditional content recommendation in that social media recommendation needs to take not only the content information but also users' social relationship and behavior into account.

1.1 Why Do People Need Recommendation in Social Media?

In online social network, many contents are generally "imported" by users from other systems, i.e., the *content sharing system* where contents are hosted to serve users. Online social network service and online content sharing service are two of the most popular applications in today's Internet. For example, online social network services like Facebook[1] and Twitter[2] have attracted hundreds of millions of users all over the world, and online content sharing system like YouTube[3] and Flickr[4] are also providing contents to billions of viewers per day.

Recent years have witnessed a rapid convergence of online social network and online content sharing network. An example of the convergence is illustrated in Fig. 1. We observe that contents are first generated by users and uploaded to the online content sharing system, e.g., more than 60 h worth of videos is uploaded every minute to YouTube.[5] Then the contents which are originally hosted by the

[1]2012: http://www.facebook.com

[2]2012: http://www.twitter.com

[3]2012: http://www.youtube.com

[4]2012: http://www.flickr.com

[5]http://www.telegraph.co.uk/technology/news/9033765/YouTube-uploads-hit-60-hours-per-minute.html

content sharing system are imported by users into the online social network and shared between users. Due to the dynamical behavior of users in the online social network and the massive number of contents generated by users, it imposes great challenge for traditional recommendations to provide personalized contents to users.

To effectively perform recommendation in the online social network, in this chapter, we present the conceptual design of social media recommendation, which takes both users and contents into account. We illustrate the advantages of social media recommendation by two types of social media recommendations in the online social network, namely, *interest-oriented social media recommendation* and *influence-oriented social media recommendation*, as follows:

- *Interest-oriented social media recommendation* answers the question "how to find the contents that can interest users in the online social network?" In this recommendation, the relevance between a user and a content is to be evaluated, so that contents that can mostly interest a user are recommended. The recommendation is based on the user relations and actions and the content similarity.
- *Influence-oriented social media recommendation* answers the question "what contents should one share to maximize one's influence?" which can be extended into two information retrieval scenarios: (1) ranking of users, given a content item, who should share it so that its diffusion range can be maximized in a social network; and (2) ranking of social media contents, given a user, what should one share to maximize one's influence among one's friends.

1.2 Traditional Approaches for Social Media Recommendation

Traditional recommendation relies on the content similarities and the collaborative references from users. Collaborative filtering-based [6, 28] and content-based [24, 26] approaches have been widely used in the existing recommendation systems, where users' rating scorings can be used to predict others' interests. Since both approaches have their shortcomings when being used individually, some designs are proposed to combine their advantages. Melville et al. [23] have proposed to use the content-based predictor to enhance the existing user data so as to provide better personal suggestions through collaborative filtering. Debnath et al. [11] have proposed to improve the content-based recommendations by some weights which determine the content attributes' importance to users. The weight values are estimated from a set of linear regression equations obtained from a social network graph which captures human judgment about similarity of items.

In online social network, users' rating scores on contents are usually modeled as a user-content matrix with each entry indicating the score a user rates a content item. The user-content matrix is highly sparse, as there are a large number of missing entries in the matrix. The recommendation system is to estimate the values of the missing entries. Matrix factorization is proposed to perform the recommendation in such model, where users and contents can be represented by vectors that are indicated in the factor matrices [18]. To finally suggest interesting contents to users, the product of the vector representing a user and the vector representing a content can be used to evaluate the relevance between them. The performance of such matrix-based model has been verified by the Netflix prize.[6]

In the context of recommendation for users' interests, Davidson et al. [10] have studied the challenges in the recommendation for such user-generated contents by taking YouTube as an example. Videos on YouTube are mostly short form, and user interactions are thus relatively short and noisy, making the traditional recommendations less effective. Baluja et al. [4] have proposed to use random walk through a co-view graph concluded from the viewing links to give video suggestions to users. Zhou et al. [40] have investigated the impact of such recommendation systems on the diversity of videos on YouTube. They observe that recommendation is the main source of views of videos on YouTube, and the number of views and the rank of recommendation are highly correlated. Walter et al. [32] have proposed a model to use the users' social connections to reach contents and filter the contents by their trust relationship. Golbeck et al. [13] have considered the social network as a recommendation network since the cascade of information is phenomenal in online social network. DuBois et al. [12] have proposed to improve the collaborative filtering recommendations by using the trust information as the weights between users.

In the context of influence maximization, there are studies on structure-level analysis [25, 29] and topic-level analysis [14, 30]. The goal of influence-oriented recommendation is to predict users' social influence for unobserved data [9] or to analyze the influence patterns from observed data [2, 3]. The existing social influence analysis research can be summarized into a diagram: *Who (A) influences whom (B) given what (C)*. A is often regarded as a single user (or node). For B, previous works can be categorized as *macroscale*, where B is the whole network [2, 37]; *microscale*, where B is a single user [8, 30]; and *mesoscale*, where B is the community or A's friends (neighborhoods) [1, 22]. From the side of C, contents can be analyzed at structure level [25, 29] and at topic level [14, 30].

1.3 What Information Can Be Utilized in Social Media Recommendation?

Traditional approaches have not taken users' social relations and their social actions into account, which is less effective in social media recommendation. In online social network, *users* and *contents* are the two core dimensions. On one hand, users can import and re-share contents in the online social network; their behavior not

[6] http://www.netflixprize.com

only indicates their interests but also reflects how contents can influence the users. On the other hand, the rapid convergence of online social network service and online content sharing service makes it possible to perform social media recommendation using information from the contents.

Our social media recommendation is based on the users and the contents generated by the users. To perform the interest-oriented recommendation and influence-oriented recommendation, we refer to the social relations, the user actions, and the content similarities:

- **Social connections**. The unique information that is available in the online social network is the set of social connections between users. On one hand, social connections reflect the social relations between people in the real world; on the other hand, social connections reflect users' interests and determine how contents can propagate among users. In different types of social network services, social connections are different, e.g., the "friending" connections in Facebook-like systems indicate users' general social interests, the "following" connections in a Twitter-like systems reflect users' information preference, and the "professional" connections in LinkedIn-like[7] systems usually show people's career interests.
- **Social actions**. The online social network has recorded valuable information on how contents are utilized by users, including which contents are *imported* by users, how these contents are *re-shared* by users, how users view these contents, and how users *comment* on these contents. On one hand, such actions make the contents propagate through the social connections and reach other users in the online social network; on the other hand, they reflect users' interests in these contents and determine the influences of users and contents, e.g., when a content item is being shared by a user, the user probably likes the content (interested) and wants his/her friends to see the content (influenced).
- **Social media analysis**. Based on the content analysis, we are able to investigate the similarities between content items. In the content analysis, multiple similarity analysis approaches can be utilized. For example, in Twitter-like systems where texts are the main form of media contents, the similarity can be evaluated using the following intuitive approaches: (1) Contents are more similar to each other if they are published in the same category, which can be inferred from the text tags in the content items, and (2) contents are similar to each other if they are published at the same location, which can be retrieved from the geographic tags. Such content similarity is not limited to text-based contents, e.g., images and videos can be analyzed using their visual features. In social media recommendation, on one hand, when performing interest-oriented recommendation, it is possible that a user likes contents that are similar to the ones he/she likes before; on the other hand, for the influence-oriented recommendation, similar contents are supposed to have similar influence properties.

[7]2012: http://www.linkedin.com/

1.4 Basic Formulation in the Social Media Recommendation

We formulate the basic models for the social media recommendation in the online social network.

1.4.1 Interest-Oriented Recommendation

In the interest-oriented recommendation, we formulate the social connections, social actions, and content analysis used in the social media recommendation as follows:

- The *social graph* is used to represent the social connections [37], e.g., how users follow each other in an online microblogging system. In the social graph, users are represented by the nodes and the social connections are represented by the edges. We use a user-user matrix **A** to denote the social graph. In an online microblogging system, **A** can be defined as follows:

$$\mathbf{A}_{ij} = \begin{cases} 1, & i \text{ follows } j, \text{ or } i = j \\ 0, & \text{otherwise} \end{cases}. \tag{1}$$

- The *user-content graph* is used to represent how users import and share contents, which are motivated by traditional matrix completion-based recommendation approaches [7]. Users and contents are represented by the nodes, and the importing/sharing actions are represented by the edges. We use a user-content matrix **B** to represent the user-content graph, which records the importing/sharing history of users as follows:

$$\mathbf{B}_{ij} = \begin{cases} 1, & \text{user } i \text{ has imported/shared content } j \\ 0, & \text{otherwise} \end{cases}. \tag{2}$$

- The *content graph* is used to represent the similarities between the contents according to content analysis [19]. Contents are represented by the nodes, and the similarities between each other are represented by the weights of the edges. We use a content-content matrix **C** to represent the content graph, which records the "similarities" between the contents as follows:

$$\mathbf{C}_{ij} \begin{cases} = 1, & i = j \\ \in [0, 1), & \text{otherwise} \end{cases}, \tag{3}$$

  where larger $\mathbf{C}_{ij}$ indicates that content i is more similar to content j.

The three matrices are then used to describe users' interests in media contents: (1) Based on the social graph **A** and user action **B**, we can infer which content items a user may like according to people one follows (friends) by using $\mathbf{A} \times \mathbf{B}$;

(2) based on the user action **B** and content similarity **C**, we can also infer users' interests, since a user may like content items that are similar to the ones he/she has already imported or shared before, by using $\mathbf{B} \times \mathbf{C}$. The operation "×" varies for different applications and different sparsities of the matrices. Using (1) and (2), user actions can be updated to a new user-content matrix $\mathbf{B}'$, where some missing entries can be estimated. Based on $\mathbf{B}'$, we are able to factorize user actions, i.e., the relevance between a user and a content can be calculated to perform the interest-oriented recommendation.

1.4.2 Influence-Oriented Recommendation

To formally define the recommendation for influence maximization problem, suppose we have M users with the ith user denoted as u_i and N content items with the jth item denoted as p_j. Let $\mathcal{N}(u_i)$ denote the collection of u_i's friends (in Facebook-like systems) or followers (in Twitter-like systems). There are two key factors involved in the influence-oriented recommendation:

- **Item-level social influence**: A straightforward way to define the strength of u_i's influence on $\mathcal{N}(u_i)$ given the content item p_j, denoted as f_{ij}, is the number of u_i's friends/followers who have viewed content p_j.
- **Social influence prediction**: There are $M \times N$ potential social influences in total. However, in practice, only a tiny fraction of them can be observed. The social influence prediction is to predict the unobserved social influences $\hat{f}_{ij}$ based on the observed f_{ij}'s and those predictive factors.

With the above terminologies, let us formally define the task of item-level social influence prediction. We denote the user-content influence matrix as $\tilde{\mathbf{X}} \in \mathbb{R}^{M \times N}$, with its (i, j)th entry

$$\tilde{X}_{ij} = \begin{cases} f_{ij} & \text{if } u_i \text{ shared } p_j \\ 0 & \text{otherwise} \end{cases}. \tag{4}$$

If we use g_i to denote the number of u_i's friends (i.e., $g_i = |\mathcal{N}(u_i)|$, where $|\cdot|$ is the cardinality of a collection), then $f_{ij} \leq g_i$. Also, it should be noted that in the matrix $\tilde{\mathbf{X}}$, there are two cases where an entry $\tilde{X}_{ij} = 0$. First, user u_i has not shared the content item p_j, and second, user u_i has shared content item p_j, but none of u_i's friends have viewed it.

Since different users have different numbers of friends/followers, the strength of social influence (if measured by f_{ij}) for each user-content pair should be evaluated in different scales. To alleviate its effect on the final performance, we use the following *percentile* influence matrix:

$$X_{ij} = \begin{cases} \frac{f_{ij}}{g_i} & \text{if } u_i \text{ shared } p_j \\ 0 & \text{otherwise} \end{cases}, \tag{5}$$

so that X_{ij}'s are normalized into the range of [0, 1].

The user-content influence matrix $\tilde{X}$ can be reconstructed by

$$\tilde{\mathbf{X}} = Diag(\mathbf{g}) \cdot \mathbf{X}, \tag{6}$$

where $\mathbf{g} = [g_1, g_2, \cdots, g_N]^\top \in \mathbb{R}^N$ and $Diag(\mathbf{g})$ is the diagonal matrix with $\mathbf{g}$ on the diagonal line.

In this way, the recommendation of influential content items is converted to the problem of predicting the unobserved entries in $\mathbf{X}$.

Next, we will discuss the interest-oriented and influence-oriented social media recommendations, respectively.

2 Interest-Oriented Social Media Recommendation

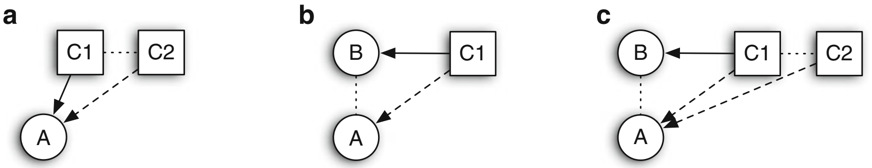
With the large number of user-generated contents in online social network, it is crucially important for social media providers to recommend users the ones that can interest them. In this section, we first show how users' interests are represented by their social behavior in online social network. Then, we present a joint social-content recommendation framework [34].

2.1 Representation of Users' Interests in Online Social Network

Explicit scores can be provided by users to indicate their interests in multimedia contents, which are commonly used in traditional recommendation frameworks. However, such rating mechanism is not suitable in the social media recommendation as follows. (1) Ratings are only a small fraction of users' social actions in online social network—many users do not rate a content even after having imported or shared it. (2) The scores cannot always accurately reflect users' interests, e.g., users can provide different scores to the same content at different times. To address the first problem, in the social media recommendation, we use the dominant social actions to study users' interests instead of the rating scores. To address the second problem, we separate the actions so that the recommendation is performed for each action, where the binary actions can be referred to in the recommendation.

We present social media recommendation for two important social actions: importing and sharing. Importing is used by users to generate new contents in the social network, while sharing is used by users to distribute contents that are already imported by others. Correspondingly, contents are recommended for users to import or share as follows:

- *Importing recommendation* which recommends users the contents to import to their profiles in online social network. Since in popular online social network



Content-based recommendation. Social-based recommendation. Joint social-content recommendation.

Fig. 2 Content-based, social-based, and joint social-content recommendations

systems such as Facebook and Twitter, contents (e.g., videos) are not hosted by the systems directly, instead, they are imported from other content sharing systems, the importing recommendation helps users in online social network to discover contents that they want to import to online social network, among all the contents from the external content sharing systems.

- *Sharing recommendation* which recommends users the contents to share in online social network. After users have imported contents to online social network, such contents will be distributed through the social connections. In online social network, users who obtain the contents shared by their friends or people they follow can further share the contents to their (other) friends or users following them, making the contents propagate in a cascade way [17]. The sharing recommendation helps a user discover the contents that he/she wants to share, among all the contents shared from his/her friends.

As important social actions in online social network [35], importing and sharing are implicit in representing users' interests, e.g., if a user imports a content, it only indicates that the user is interested in the content but cannot evaluate how much the user likes the content. Since rating scores are not available to connect the users and the contents, in the interest-oriented social media recommendation, social actions are divided into different groups so that the recommendation can be performed separately. The rationale is that the actions in the same group can be used to evaluate each other.

2.2 Recommendation Based on Users and Contents

Traditional interest-oriented recommendation for social media includes both *content-based approach* [26] and *social-based approach* [39]. In the content-based recommendation, content filtering and collaborative filtering [28] have been widely used. They make use of either the similarities based on content analysis or the similarities based on the historical users' ratings. Such recommendation can provide a user with contents that are similar to the ones he/she has viewed before, as illustrated in Fig. 2a. On the other hand, in the social-based recommendation, social network is used to filter the information distributed through the social connections, so that content items that one likes can be recommended to their friends [32]. Such recommendation is able to provide users with the contents that have previously interested their friends or people they follow, as illustrated in Fig. 2b. In existing social media recommendations, users' social connections and contents' similarities are used separately.

Figure 2c illustrates the new conceptual design of a *joint social-content recommendation*, where users and the contents are used jointly to perform the recommendation. The design has the following advantages:

- Using social graph and social actions in the recommendation makes it possible to take the propagation patterns into consideration [41]. Propagation determines how contents reach different people in online social network, which can be very important to the recommendation, e.g., if a content item repeatedly appear to a user since it is repeatedly shared by his/her friends, he/she may get interested in that content item as well [33]. Given the social relations and social actions, such propagation of content items can be simulated and the strength of the propagation can be evaluated for the recommendation.
- Since importing and sharing contents in online social network are implicit, i.e., users do not rate the contents they have imported or shared, which are required in many existing recommendation approaches, we need to further refer to other information from the social network and the content sharing network, e.g., comments of users in online social network can be used to further analyze users' emotions [31].
- Cold start is even more challenging in social media recommendation for two reasons as follows. (a) Users who have just joined the system have hardly imported or shared any content in the system. It is difficult to recommend any content for them, since existing recommendation systems rely on users' historical preferences. (b) A large number of user-generated contents have no or fewer viewers. It is difficult to decide which users these contents should be recommended to. In the joint social-content recommendation, (a) and (b) can be solved by updating the user-content matrix, based on the original matrices **A**, **B**, and **C**, where both users' social connections and the content similarities can be used to assist the updating of the missing entries in **B**.

In the joint social-content recommendation, users' social connections, user actions, and content similarities are utilized to derive an updated *user-content graph* to perform joint recommendation. More specifically, first, the joint social-content recommendation utilizes a user-content matrix completion to predict which contents are to be imported/shared by which users. In the completion, entries for cold users and cold contents are updated so as to perform recommendation for such users and contents. Second, based on the updated user-content matrix and the historical importing and sharing records of users, a joint *user-content space* can be built to measure the relevance between a user and a content for the recommendation. Next, we discuss the details of the joint social-content recommendation.

2.3 *The Joint Social and Content Recommendation*

The joint social-content recommendation is proposed to recommend interesting contents to users by using their social relations, social actions, and the content similarities. The joint social-content recommendation framework is facing the following challenging problems: (1) How to make use of the propagation patterns of contents to update the user-content matrix? (2) How to make use of the implicit actions of users' importing and sharing contents to perform the recommendation? (3) How to improve the recommendation for newly joint users who have little historical importing/sharing information and newly published contents which are viewed by fewer users? The answer to these questions is the design of the joint social-content recommendation framework.

First, a *completion* scheme based on the social propagation and the content similarity is used to update the user-content matrix. In the user-content matrix completion, the missing entries corresponding to some cold users/contents can be estimated. On one hand, the more a user's friends have imported or shared a content, the more possible it is for the user to import or share the content as well. On the other hand, if a user has imported or shared a content, the user will probably import or share contents similar to that content as well. The matrix update procedure can be executed based on new entries that are updated, simulating the propagation of contents in online social network. The user-content matrix completion solves the aforementioned problems (1) and (3).

Second, to connect the users and the contents by their implicit social actions, we use the importing and sharing records as input to construct a dynamical user-content space to measure the relevance between a user and a content. In the user-content space, a user can be represented by a vector, each entry in which is calculated according to the contents the user has imported/shared. Similarly, a content can be represented by a vector according to the user actions as well. Thus, the construction of the user-content space only relies on the binary entries in $\mathbf{B}'$. Each import or share will be regarded as a small increase of the user's interest in a particular dimension in the user-content space. The user-content space construction solves the aforementioned problem (2).

2.3.1 Completion of the User-Content Matrix

The user-content matrix which indicates users' social actions on the contents can be very sparse [15]. It is difficult for the traditional recommendation algorithms to recommend contents to users who have imported/shared no or fewer contents, which are based on users' historical preferences. In a joint social-content recommendation, the user-content matrix ($\mathbf{B}$) can be updated by the collaboration of social graph ($\mathbf{A}$), social actions (the original $\mathbf{B}$), and content analysis ($\mathbf{C}$). The completion makes use of the original three matrices simultaneously. On one hand, the social

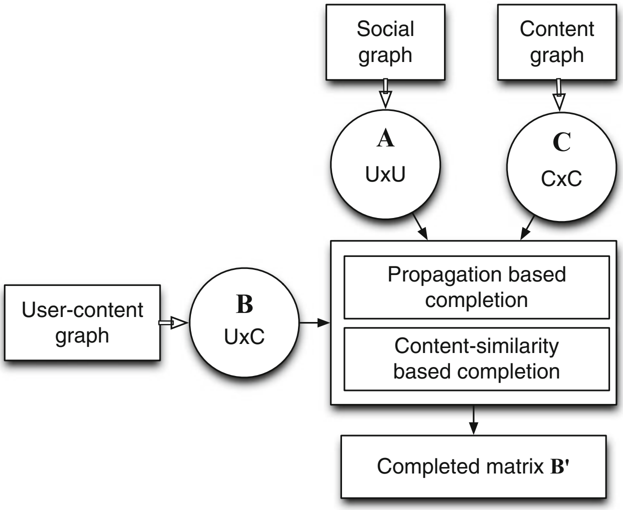


Fig. 3 User-content matrix completion based on propagation and content similarity

propagation model can be used to connect the contents to the users, i.e., contents imported/shared by a user's friends are likely to be imported/shared by the user as well. On the other hand, contents can be connected to users according to the content similarity analysis. The reason is that a user is likely to enjoy contents that are similar to the ones he/she has imported or shared before. Figure 3 illustrates the user-content completion framework. When updating the user-content matrix **B**, besides the original user-content matrix **B**, the user-user matrix **A** and the content-content matrix **C** are all used. In particular, **A** and **B** will be used for the social propagation-based completion, and **C** and **B** will be used for the content similarity-based completion.

2.3.2 Construction of the User-Content Space

The recommendation is based on constructing a joint user-content space to measure the relevances between users and content items in the space. Figure 4 illustrates how the relevance between a user and a content item is measured. The joint user-content space is constructed by combining a user space and a content space. The user space depends on the user-user matrix, and the content space depends on the content-content matrix. A user or a content item can be mapped to two *description vectors* [38] in both spaces. Since both the user space and the content space are constructed using a similar procedure, we take the user space as an example to illustrate how the space is constructed. First, several representative users are selected, which are the top-followed users in the online social network. Since these users can be followed by many other users, it is effective to use them to represent other users' interests.

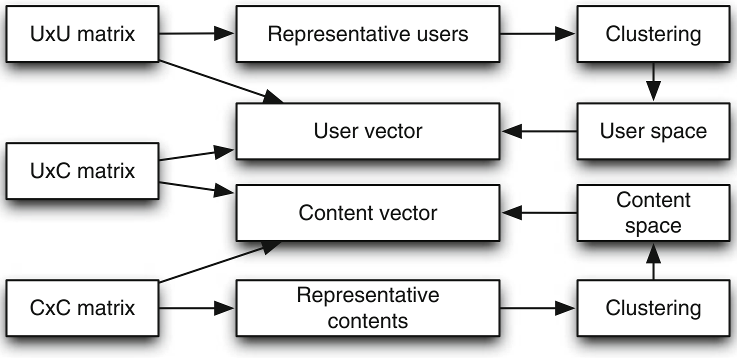


Fig. 4 Construction of the joint user-content space based on user actions

Second, these representative users are clustered into several groups, according to a similarity defined between any two representative users (two representative users are more similar to each other if they are followed by more common fans). Third, based on the clustered groups and the user-user matrix, we can define a user's description vector in the user space by counting the number of representative users he/she follows in each group, and based on the user-content matrix, we can define a content's description vector in the user space by combining the description vectors of users who have imported/shared the content.

The *relevance* between a user and a content is then measured by combining their dot products in both spaces, where a recommendation parameter is utilized to control their impacts for different social actions, i.e., importing and sharing. Finally, contents can be recommended according to their relevances to the user.

3 Influence-Oriented Social Media Recommendation

With the rapid proliferation of social applications, more and more user profiles, interactions, and collective intelligence (such as social tags and comments) are available online, which opens a new perspective for recommendation applications where more focus should be put on user collaborative information. At the same time, new recommendation scenarios, such as friend recommendation [16, 21], have also emerged. These scenarios propose a new challenge to traditional recommendation: how to effectively handle it in the social media scenario?

One key concept related to this challenge is *recommendation for influence maximization*, which has been becoming a prevalent and complex force governing the dynamics of social relationship or social network [37]. It is also a key dimension for modern recommendation in multiple aspects. To mention a few, (1) each user acts as an information source in social network, and the influence of a user is meaningful for the authority of the generated information; (2) in the influence-oriented recommendation, the social influence is the key indicator for content recommendation; (3) as different contents vary with the power to affect users to change their actions, they can be recommended by influence ranking for social purpose. Therefore, there is a clear need for techniques to analyze social influence and, more importantly, in recommendation field.

In the influence-oriented social media recommendation, we need to answer the question "who should share what?" which can be extended into two scenarios as follows. (1) Users ranking: given an item, who should share it so that its diffusion range can be maximized in a social network. (2) Content ranking: given a user, what should one share to maximize one's influence among one's neighbors.

Item-level social influence is not a general measure on users but on the interactions of users and contents. That is, we need to discriminate a user's social influences with respect to different contents. Different from most of the existing research works focusing on users' overall social influence analysis [25, 29] and topical social influence mining [14, 30], the social influence in this section is a finely grained measure of influence.

3.1 Motivating Recommendation Scenario

In social media, users and contents are two core dimensions, and users' sharing of contents (such as blog, news, and album) is the basic behavior. Actually, the spreading out of contents is because of the user sharing in social network. The owner of the contents, e.g., the advertisers, hopes to maximize the diffusion range of the contents [5]. This goal makes them desire to target the influencers, who are able to let many friends to click the information they share or even share further to extend the sharing cascades. Psychologically, users share contents with their friends mainly because they want to build their reputations and help others, in which *to influence others* is the important motivation for sharing [35].

According to the definition of social influence on Wikipedia,[8] social influence occurs when "an individual's thoughts, feelings or actions are affected by other users." In the context of online social network systems like Facebook and Twitter, when a user shares a content item, a portion of his/her friends (or neighbors) will click, comment, or even re-share the content, which are three levels of influence [36]. The resulted predictive model can be used in two angles. On one side, given a content item, we can find out the influencers for the diffusion. On the other, given a user, we can recommend a list of content items to share, which can improve the interactions between the user and his/her friends.

In predicting the item-level social influence, we mainly face the following challenges:

---

[8] 2012: http://en.wikipedia.org/wiki/Social_influence

- **User-content specific**. Item-level social influence is not a general measure on users but on the interactions of users and contents. That is, we need to discriminate a user's social influences with respect to different content items. Different from most of the existing research works focusing on users' overall social influence analysis and topical social influence mining, the social influence in this chapter is a finely grained measure of influence.
- **Sparsity**. The interactions between users and content items are extremely sparse compared with the total number of user-content pairs. According to our statistics of 34 K users in Renren,[9] which is a Facebook style social network site in China, each user only shares six web contents in average during a month, compared with a total of 43 K content items, and each item is only shared by four users, compared with a total of 34 K users. Thus, it is clear that we need subtle and effective prior knowledge for user and content grouping to alleviate the sparsity problem.
- **Complex factors**. There are a volume of factors that affect how many friends will click a shared content item and provide potential clues for user and content grouping, for example, the total number of friends, the tie strength between the user and his/her friends and the semantics of content items, which are often in different scales. How to select the effective factors and integrate these complex factors in one predictive model is also one of the focus in the influence-oriented recommendation.

We formulate the recommendation for influence maximization problem as the estimation of a user-content matrix, in which each element (i, j) represents the number of clicks by friends of user i on his/her jth shared content item. A *hybrid factor nonnegative matrix factorization (HF-NMF)* algorithm is designed for item-level social influence modeling [9]. The algorithm tries to find out the common hidden vector space for both the users and the contents, where their multiplication can well approximate the observed training interaction matrix. Meanwhile, in order to deal with the sparsity problem, we construct the priors on users and contents by incorporating the user-user similarity matrix and content topic distribution matrix. Also, in order to alleviate the over-fitting problem, we introduce the $L2$-norm as regulations for the hidden vector space to improve the generalization ability. We apply *projected gradient* [20] to solve the HF-NMF problem and carry out intensive experiments to demonstrate the effectiveness of the proposed method. To summarize, the algorithm for recommendation for influence maximization includes:

- The formulation of the item-level recommendation for influence maximization problem with HF-NMF and an efficient projected gradient method to solve it.
- The predicted item-level social influence from HF-NMF can support the applications, such as influencer ranking and contents recommendation by user-content matrix ranking in two directions.

[9] http://www.renren.com

- The strength of social influence in this method is well interpreted, which makes it easy to understand and extendable to higher-order social influence, e.g., the influence on all the friends and the friends of friends.

3.2 Recommendation for Influence Maximization

Let us demonstrate the necessity of the item-level social influence and validate the rationality of predictive factors by preliminary statistical analysis.

The dataset is acquired from the real social network Renren. Till now, the web site already owns more than 150 million active users. In this web site, a user can generate a content or share a web page as a content, and the user's friends will be informed through the news-feed mechanism. Then some of the friends will click, comment, or share the content. In this section, we only consider the click action as the manifestation of influence, and the number of clicks corresponds to the strength of influence. As the number of friends is different for each user, the upper bound of users' social influence strengths is also different. In order to make the strength of influence be measured in a unified scale for different users for the sake of observational and modeling study, we use the proportion of friends (of the user who publish a content) who click the shared content as the measure.

In the influence-oriented recommendation, we first assume that the influence should be specific on each user-content pairs. In order to validate the hypothesis, we randomly select three active users. Given a user, we calculate the proportion of his/her influenced friends (who clicked the shared content) for each of the shared contents. Then, we randomly select three popular contents. Given a content, we calculate the proportion of influenced friends when the content is shared by different users. In our observations, the social influence notably varies with different users and contents, which implies that (1) different users have different influence power to their friends, (2) different contents have different influence power (more intuitively, attraction) to users who are interested in, and (3) users' influences manifest differently for different contents. Therefore, only item-level social influence can reveal the users' real influence on friends, and the strength of influence should definitely be user-content specific.

3.3 Predictive Factors for Influence-Oriented Social Media Recommendation

The factors that affect the strength of social influence include the following three aspects:

- **User-specific factors**. Although users' social influence varies with the shared contents, the average of the social influences over contents determines the overall social influence of a user. We regard the factors that affect users' overall social influence (excluding the contents) as the user-specific factors.

- **Content-specific factors**. Similar as user-specific factors, we regard the factors affecting contents' overall social influence (excluding the users) as the content-specific factors.
- **User-content-specific factors**. As mentioned above, the social influence is user-content specific. The social influence of a user given a content cannot be well estimated only by the user and content-specific factors. The factors indicating the interactions between users and contents are also important for social influence-oriented recommendation.

One issue that is worthy of emphasizing here is that both the users and contents are essential for the predictive modeling. On one hand, the user-content interactions are very sparse. We need to find effective factors to "group" those users and contents to alleviate the sparsity problem. On the other hand, the user and content-specific factors also provide some effective prior knowledge to complement the inference from pure user-content interactions.

In order to find out the effective predictive factors for the influence-oriented recommendation, we first prepare a factor pool, which includes the available potential predictive factors including user profiles, number of users' friends, visiting frequency between users, and contents' topic distributions. Given each factor, we measure the correlation between the strength of social influence and the factor value. Finally, we select two user-oriented factors: the percentage of active friends, the average social tie strength (the interaction frequency) between a user and his/her friends, and one content-specific factor—the topic distribution of a content's content.

3.4 Influence Maximization: A Matrix Factorization Approach

Based on the symbols and basic model we have illustrated in Sect. 1, the influence maximization can be formulated as a matrix factorization problem. Let $\mathbf{U} \in \mathbb{R}^{M \times k}$ be the latent user feature matrix and $\mathbf{V} \in \mathbb{R}^{N \times k}$ be the latent content feature matrix, where k is the number of latent features. Then given the observed user-content-specific social influence matrix $\mathbf{X}$, the objective is to find the optimal latent user matrix $\mathbf{U}$ and latent content matrix $\mathbf{V}$ by minimizing the following objective:

$$\mathscr{J}_1 = \left\| \mathbf{X} - \mathbf{U}\mathbf{V}^\top \right\|_F^2 , \tag{7}$$

where $||.||_F$ denotes the matrix Frobenius norm.

The objective function $\mathscr{J}_1$ can be regarded as the quality of approximating $\mathbf{X}$ by the inner product of $\mathbf{U}$ and $\mathbf{V}$. However, in real cases, most of the elements in $\mathbf{X}$ are zero because of the sparse interactions between users and contents. Thus, in order to focus more on the valid elements, we propose to only measure the approximation loss on observed elements on $\mathbf{X}$. To formulate this, we introduce the *sharing matrix* $\mathbf{Y} \in \mathbb{R}^{M \times N}$ with its (i, j)th entry defined as

$$Y_{ij} = \begin{cases} 1 & \text{if } u_i \text{ shared } p_j \\ 0 & \text{otherwise} \end{cases} . \tag{8}$$

The objective function is converted to

$$\mathscr{J}_2 = \left\| \mathbf{Y} \odot (\mathbf{X} - \mathbf{U}\mathbf{V}^\top) \right\|_F^2 , \tag{9}$$

where $\odot$ is the Hadamard product.

As mentioned above, the severe sparsity of $\mathbf{X}$ makes it very challenging to directly learn the latent spaces for users and contents from only observed user-content interaction entries. We need to make full use of the user-specific and content-specific factors to compress the degrees of freedom, so that the correlation within users and contents can be exploited to alleviate the sparsity problem. The solution to the optimization is similar to the *maximum margin matrix factorization* approach in [27] and the *joint matrix factorization* approach in [18].

4 Conclusions

In this chapter, we introduce the principal concept and the framework of social media recommendation. We first present why recommendation is demanded in social media and what unique information is needed for effective social media recommendation from both users and contents. We then provide the theoretical formulation and key insights on the social media recommendation. In particular, we have presented two types of recommendations in online social network: interest-oriented social media recommendation and influence-oriented social media recommendation. To deepen the understanding of each social media recommendation, we further provide the following results:

- In the interest-oriented social media recommendation, we present the joint social-content recommendation. We have shown that user relations, user actions, and content similarities can be simultaneously utilized in the recommendation, so that two problems can be addressed as follows. (1) The completion of the user-content matrix based on the social propagation and the content similarity makes it possible to recommend newly published contents to newly joint users. (2) The construction of the joint user-content space enables recommendation for implicit social actions including importing and sharing.
- In the influence-oriented social media recommendation, we have presented the effective user-specific and content-specific predictive factors and used a matrix factorization method to incorporate these predictive factors for user-content-specific recommendation for social influence maximization at content-item level.

We see such an exploration will shed light on future applications in social media recommendation.

References

1. Agarwal, N., Liu, H., Tang, L., Yu, P.: Identifying the influential bloggers in a community. In: Proceedings of the ACM International Conference on Web Search and Web Data Mining, Palo Alto (2008)
2. Anagnostopoulos, A., Kumar, R., Mahdian, M.: Influence and correlation in social networks. In: Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, Las Vegas (2008)
3. Bakshy, E., Karrer, B., Adamic, L.: Social influence and the diffusion of user-created content. In: Proceedings of the ACM Conference on Electronic Commerce, Stanford (2009)
4. Baluja, S., Seth, R., Sivakumar, D., Jing, Y., Yagnik, J., Kumar, S., Ravichandran, D., Aly, M.: Video suggestion and discovery for Youtube: taking random walks through the view graph. In: Proceedings of the ACM WWW, Beijing (2008)
5. Bao, H., Chang, E.: Adheat: an influence-based diffusion model for propagating hints to match ads. In: Proceedings of the ACM WWW, Raleigh (2010)
6. Breese, J., Heckerman, D., Kadie, C.: Empirical analysis of predictive algorithms for collaborative filtering. In: Proceedings of the Fourteenth Conference on Uncertainty in Artificial Intelligence. Morgan Kaufmann, San Francisco (1998)
7. Cai, J.-F., Candès, E.J., Shen, Z.: A singular value thresholding algorithm for matrix completion. SIAM J. Optim. **20**(4), 1956–1982 (2010)
8. Crandall, D., Cosley, D., Huttenlocher, D., Kleinberg, J., Suri, S.: Feedback effects between similarity and social influence in online communities. In: Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, Las Vegas (2008)
9. Cui, P., Wang, F., Liu, S., Ou, M., Yang, S., Sun, L.: Who should share what? Item-level social influence prediction for users and posts ranking. In: Proceedings of the ACM SIGIR International Conference on Research and Development in Information. ACM, New York (2011)
10. Davidson, J., Liebald, B., Liu, J., Nandy, P., Van Vleet, T., Gargi, U., Gupta, S., He, Y., Lambert, M., Livingston, B., et al.: The Youtube video recommendation system. In: Proceedings of the ACM Conference on Recommender Systems, Barcelona (2010)
11. Debnath, S., Ganguly, N., Mitra, P.: Feature weighting in content based recommendation system using social network analysis. In: Proceedings of the ACM WWW, Beijing (2008)
12. Golbeck, J., Kleint, J., Srinivasan, A.: Improving recommendation accuracy by clustering social networks with trust. ACM RecSys'09 Workshop on Recommender Systems & the Social Web, New York (2009)
13. Golbeck, J., Hendler, J.: FilmTrust: movie recommendations using trust in web-based social networks. In: Proceedings of the IEEE Consumer Communications and Networking Conference, Las Vegas (2006)
14. Goyal, A., Bonchi, F., Lakshmanan, L.: Learning influence probabilities in social networks. In: Proceedings of the ACM International Conference on Web Search and Data Mining, New York (2010)
15. Johnson, C.: Matrix completion problems: a survey. In: Matrix Theory and Applications, vol. 40, pp. 171–198. American Mathematical Society, Providence (1990)
16. Kalashnikov, D., Chen, Z., Mehrotra, S., Nuray-Turan, R.: Web people search via connection analysis. IEEE Trans. Knowl. Data Eng. **20**(11), 1550–1565 (2008)
17. Kempe, D., Kleinberg, J., Tardos, É.: Maximizing the spread of influence through a social network. In: Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, Washington (2003)

18. Koren, Y., Bell, R., Volinsky, C.: Matrix factorization techniques for recommender systems. *Computer* **42**(8), 30–37 (2009)
19. Krippendorff, K.: Content Analysis: An Introduction to Its Methodology. Sage, Thousand Oaks (2004)
20. Lin, C.: Projected gradient methods for nonnegative matrix factorization. Neural Comput. **19**(10), 2756–2779 (2007)
21. Macdonald, C., Ounis, I.: Voting for candidates: adapting data fusion techniques for an expert search task. In: Proceedings of the ACM International Conference on Information and Knowledge Management. ACM, New York (2006)
22. McMillan, D., Chavis, D.: Sense of community: a definition and theory. J. Community Psychol. **14**(1), 6–23 (1986)
23. Melville, P., Mooney, R., Nagarajan, R.: Content-boosted collaborative filtering for improved recommendations. In: Proceedings of the National Conference on Artificial Intelligence. AAAI, Menlo Park (2002)
24. Mooney, R., Roy, L.: Content-based book recommending using learning for text categorization. In: Proceedings of the ACM Conference on Digital Libraries, San Antonio (2000)
25. Newman, M.: The structure and function of complex networks. SIAM Rev. **45**, 167–256 (2003)
26. Pazzani, M., Billsus, D.: Content-based recommendation systems. In: Brusilovsky, P., Kobsa, A., Nejdl, W. (eds.) The Adaptive Web, pp. 325–341. Springer, Berlin/New York (2007)
27. Rennie, J., Srebro, N.: Fast maximum margin matrix factorization for collaborative prediction. In: Proceedings of the ACM International Conference on Machine Learning. ACM, New York (2005)
28. Schafer, J., Frankowski, D., Herlocker, J., Sen, S.: Collaborative filtering recommender systems. In: Brusilovsky, P., Kobsa, A., Nejdl, W. (eds.) The Adaptive Web, pp. 291–324. Springer, Berlin/New York (2007)
29. Strogatz, S.: Exploring complex networks. *Nature* **410**(6825), 268–276 (2001)
30. Tang, J., Sun, J., Wang, C., Yang, Z.: Social influence analysis in large-scale networks. In: Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, Paris (2009)
31. Thelwall, M., Wilkinson, D., Uppal, S.: Data mining emotion in social network communication: gender differences in myspace. J. Am. Soc. Inf. Sci. Technol. **61**(1), 190–199 (2010)
32. Walter, F., Battiston, S., Schweitzer, F.: A model of a trust-based recommendation system on a social network. Auton. Agents Multi-Agent Syst. **16**(1), 57–74 (2008)
33. Wang, Z., Sun, L., Yang, S., Zhu, W.: Prefetching strategy in peer-assisted social video streaming. In: Proceedings of the ACM Multimedia, Scottsdale (2011)
34. Wang, Z., Sun, L., Zhu, W., Yang, S., Li, H., Wu, D.: Joint social and content recommendation for user generated videos in online social network. Technical report
35. Wasko, M., Faraj, S.: Why should i share? Examining social capital and knowledge contribution in electronic networks of practice. Mis Q. **29**, 35–57 (2005)
36. Weng, J., Lim, E., Jiang, J., He, Q.: Twitterrank: finding topic-sensitive influential twitterers. In: Proceedings of the ACM International Conference on Web Search and Data Mining, New York (2010)
37. Wolfe, A.: Social network analysis: methods and applications. Am. Ethnol. **24**(1), 219–220 (1997)
38. Xu, R., Wunsch, D., et al.: Survey of clustering algorithms. IEEE Trans. Neural Netw. **16**(3), 645–678 (2005)
39. Yang, W., Dia, J., Cheng, H., Lin, H.: Mining social networks for targeted advertising. In: Proceedings of the IEEE Annual Hawaii International Conference on System Sciences. IEEE Computer Society Press, Los Alamitos (2006)
40. Zhou, R., Khemmarat, S., Gao, L.: The impact of YouTube recommendation system on video views. In: Proceedings of the ACM IMC, Melbourne (2010)
41. Ziegler, C., Lausen, G.: Propagation models for trust and distrust in social networks. Inf. Syst. Front. **7**(4), 337–358 (2005)