Two-Stage Friend Recommendation Based on Network Alignment and Series Expansion of Probabilistic Topic Model

Shangrong Huang, Jian Zhang, *Senior Member, IEEE*, Dan Schonfeld, *Fellow, IEEE*, Lei Wang, *Senior Member, IEEE*, and Xian-Sheng Hua

***Abstract*—Precise friend recommendation is an important problem in social media. Although most social websites provide some kinds of auto friend searching functions, their accuracies are not satisfactory. In this paper, we propose a more precise auto friend recommendation method with two stages. In the first stage, by utilizing the information of the relationship between texts and users, as well as the friendship information between users, we align different social networks and choose some "possible friends." In the second stage, with the relationship between image features and users, we build a topic model to further refine the recommendation results. Because some traditional methods, such as variational inference and Gibbs sampling, have their limitations in dealing with our problem, we develop a novel method to find out the solution of the topic model based on series expansion. We conduct experiments on the Flickr dataset to show that the proposed algorithm recommends friends more precisely and faster than traditional methods.**

***Index Terms*—Friend recommendation, series expansion, topic model.**

I. Introduction

FRIEND recommendation is a primary function in social network services and aims to recommend new social links for each user. Today when we lodge on the main social website such as Facebook, Twitter, and LinkedIn etc., we receive many recommendations of online friends. Seeing and hearing what the friends look at and listen to, or sharing our experience with our friends is an unparalleled experience. However, the decision of making friends is a complex human behaviour and can be affected by many different factors such as age, gender, location, interest [1], etc. As a consequence, similar to real life, finding a good on-line friend is not easy without the help of good recommendations. Traditional friend recommendations widely applied by Facebook and Twitter are often based on common friends and similar profiles such as having the same hobbies or studying in the same fields. These methods usually provide a long ranked possible friend list, but the recommendation precision is usually not satisfactory due to its complexity.

Nowadays people are commonly retained in a multi-resource environment, and usually do not seek friends based on only one kind of information anymore. Recently cross domain friend recommendation technologies have been extensively explored [2]–[4]. [2] applies a matrix factorisation method to combine the image and text information, [3] considers the proximity and homophily information for synthesised recommendation, and [4] specifies individuals' requirements from different domains. Most of these papers utilise information from different resources simultaneously for recommendation. In this paper, we approach this recommendation problem in a different way by utilizing the multi-domain information in different stages for a more precise recommendation.

The reason why we apply the multi-stage friend recommendation scenario lies in the complexity of multi-source information and the decision making behaviour of people. For example, an individual might make an on-line friend because they discuss a hard mathematical problem, or it is possible that he/she makes a friend because they both enjoy a film. The reason for friend making might be very diverse. It would be relatively difficult if we consider different factors together at the same time for recommendation. In our opinion, it is more convenient and clearer to analyse these factors step by step, rather than to deal with such cross-domain information as a whole. By untwisting the different factors in the recommendation procedure and analysing each factor in depth, a more precise recommendation performance is expected. As a consequence, we apply a two-stage framework to synthesise heterogeneous information from different domains.

In this paper, we concentrate on the widely-used image and image-related experience sharing website Flickr, where individuals can upload photos and tags for sharing as well as make online friends (Flickr Contact) and join communities (Flikr Group). Tag (text) information is quite useful for friend recommendation since it is simple and direct. For example, two individuals that both have interest in tags "travel" and "historical people" have higher probability to be friends with

Manuscript received April 11, 2016; revised August 31, 2016 and November 21, 2016; accepted December 28, 2016. The associate editor coordinating the review of this manuscript and approving it for publication was Dr. Zhen Wen.

S. Huang and J. Zhang are with the School of Computing and Communications, University of Technology Sydney, Ultimo, NSW 2007, Australia (e-mail: shangrong.huang@student.uts.edu.au; Jian.Zhang@uts.edu.au).

D. Schonfeld is with the Multimedia Communications Laboratory, University of Illinois at Chicago, Chicago, IL 60607, USA (e-mail: dans@uic.edu).

L. Wang is with the School of Computing and Information Technology, University of Wollongong, Wollongong, NSW 2522, Australia (e-mail: leiw@uow.edu.au).

X.-S. Hua is with the Alibaba Group, Hangzhou 311121, China (e-mail: xianshenghua@gmail.com).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TMM.2017.2652074

each other. Text-based friend recommendation have been successfully developed in [5], [6]. We apply the text information in our first stage.

Image information is also helpful for friend recommendation. On the other hand, image information is vaguer and more complex for this task. For example, it is hard to claim that two individuals who both enjoy some vivid and colourful photos, or some photos of beautiful women have the higher probability to be friends. As a consequence, in our algorithm, we utilise the image information as a supplementary source in the second stage of the algorithm, to refine the result of the first stage.

In the first stage, similar to [5], based on the correlation of different networks, we align the tag-similarity network to friend network to obtain a possible friend list. Specifically, we consider each user as one node in a graph, and we crawl the uploaded tags from each user and calculate the tag similarity between any two users as the edges to form a tag-similarity network. On the other hand, we also obtain the friendship information in Flickr, and if two users have friendship with each other, we add an edge between the two to form a contact network. In this way we build a tag-similarity network and a contact network that have the same nodes but different topologies. Because the tag-similarity network and contact network on Flickr are related to each other, we dig their correlation by choosing important tag features, to make the tag-similarity network more similar to the contact network. In this way, the chosen tag features provide a guideline for friend recommendation. This stage makes a mass election of possible friends.

In the second stage, to overcome the problem that the mass election considering only the tag information might not be precise, we build a topic model to illustrate the relationship between user's friend making behaviour and the image features they have uploaded. This stage refines the list obtained in the first stage. The main reason for applying a topic model in our second stage lies in the fact that the topic model has the ability to tell on what probability a user would prefer a photo/item/friends.

The probabilistic topic model discovers the abstract "topics" that occur in a collection of documents/datasets, and it has been widely used in recommendation systems [7]–[9]. By introducing some latent variables and applying the Bayesian rule, it is conceptually easy to combine information from different domains and make specific recommendations [7], [9]. Generally it assumes that people's various behaviours such as shopping, posting and friend making are controlled by some latent topics. Certain people have particular bias on different latent topics. For an individual that acts differently in different domains, his/her latent interest topic might be similar. For example, a user who posts many different photos about food on Flickr might have higher probability to be interested in the topic of cooking, and thus it is reasonable to recommend some kitchenware to him/her on Amazon. Furthermore, the topic model provides a relatively precise probability to show to what extent an individual is interested in a topic, and thus makes it easy for further recommendation.

In this paper, we propose a topic model to correlate the data about the Flickr image information and the contact information. Compared with some previously cross-domain topic models, our model is more compact with less parameters, which leads to some computational convenience. Briefly, we assume that the attractiveness of photos is controlled by a latent variable, and individuals' photo uploading behaviours and their friend making behaviours are controlled by some other latent variables. By determining the values of these latent variables we can predict individuals' friends.

However, it is often not easy to find the solution of a topic model when different domains are concerned, for it involves the integrals of several coupled random variables, which is a complicated mathematical problem in general [10]. Two methods are widely used to deal with this problem: Gibbs sampling [11] and variational inference [10], or the combination of the two [12]. Although applied successfully in many cases, both of them have some disadvantages: for Gibbs sampling, it is inefficient for large count values since it requires averaging over many samples to reduce variance; for variational inference, though it is efficient to deal with large scale data, the variational step makes it hard to control the precision when approximating the integrals when making the Bayesian inference. In this paper, with the help of Mellin and inverse Mellin transform, we propose a novel way based on series expansion to calculate the coupled integrals that are required in the Bayesian inference.

Matrix factorization (MF) method can be also applied to deal with the cross domain recommendation problems [13], [14]. It decomposes different social networks into latent vectors to find the important factors that influence individuals' social behaviours, and make recommendations based on these latent factors. However, it lacks a mechanism to draw the complete distributions of the whole social network, and thus might lead to some local optimum. Our proposed method provides a way to describe the whole distribution of the social network, to perform a better recommendation.

To sum up, we build a two-stage friend recommendation system based on text and image data: in the first stage, we apply tag-user information to get a possible friend list, and in the second stage we refine the list by utilizing the image-user information. Our main contributions are as follows: Firstly, we build a compact topic model to analyse the relationship of the data from different domains. Secondly, we propose a novel method based on the study of the distribution of algebra of random variables to find a solution of the model. The solution is given in a series expansion form, and can lead to more precise solutions of the model. As far as we know,this is the first time to solve a topic model from the aspect of integral series expansion. We also make comprehensive experiments to show the effectiveness of our method.

The rest of the paper is organised as follows: Section II outlines related work. Section III introduces our system framework. Section IV gives the detailed explanation of our series expansion method. Section V evaluates the performance of our method and some analysis is made according to the results. Lastly, Section VI concludes our work.

II. Related Work

Our work in this paper is mainly related to the following research fields: friend recommendation, topic model, and algebra of random variables.

A. Friend Recommendation

Friend recommendation is a relative challenging issue compared with item or group recommendations, for there might be various reasons for two persons to become friends, and online and offline friendships are quite different. Recently, [15] even provides some method to distinguish the online and offline friends.

[6] makes a survey of individuals' daily life, and then summarises the reports as their "life styles" using latent Dirichlet allocation algorithm (LDA). [8] collects individuals' posts in Micro-blogs and arranges them in a chronological order. By building a temporal-topic model it can recommend different friends to each user at different time, as the user's interest changes from time to time. [16] utilises the information from different platforms (Flickr, Twitter, Google+, etc) to alleviate the sparsity problem of social networks, the idea is that Google+ can provide a information bridge between these different social platforms. In this paper, We dig the friend recommendation problem deep by considering multimedia information one platform, and applying a two-step scenario to refine the result.

1) Cross-Domain Recommendation: As mentioned in Section I, individuals' decision of making friends are often multi-dimensional. As a result, recently many researchers consider friend recommendation based on cross-domain information. [17] considers the friend recommendation problem at working places and conferences, by utilising both users' temporal location as well as their common friend information. [2] combines three aspects of each user's information: the items one likes, the friends one has, and the groups one belongs to. Such information of different aspects is synthesised and integrated into one cost function. By optimizing the cost function, the heterogeneous data are fused for item, group and friend recommendations. In [16], individuals that have both accounts in Flickr, Twitter and Google+ are collected to build the relationship of the two social websites. The common behaviours of each user in Flickr and Twitter are analysed and the friend recommendation of the two domain is made based on these common behaviours.

[4] divides the different data in Flickr into two classes: interaction data(comments, making favorite photos) and similarity data(common friends, groups, tags, geo, visual), and applies these two classes of data comprehensively to estimate the strength of the ties between users. [18] utilises Flickr social relations for further multimedia recommendation. It builds a topic model to combine the image, text, and friendship information to discovery individuals' preferences. The topic model is solved via Gibbs sampling.

For the works listed above, the data from different domains are processed simultaneously or fused together to get the final recommendation result. On the one hand, the above methods take the advantage that data from different domains might be related to each other; On the other hand, these methods combine the cross-domain information in one step ([16]) or synthesise it in one cost function ([2]), thus usually can not give a good explanation of how the data from a specific domain contribute to the final recommendation result, and the twisted data from different domains often makes the problem more complex. To have a better understanding of the effectiveness of the data from each domain, in this paper, we design a two-step recommendation that in each step we utilise the data from one domain.

2) Multistage Recommendation: Existing multi-stage recommendations are usually applied to find some patterns of users or items. For example, in [19], a two-stage mobile recommendation is proposed to help users find the correct events. The first stage clusters people according to their profile similarity and the second stage discovers the event-participating pattern. [20] designs the first stage to find some related resources that one user requires, and the second stage is used to find some patterns that the user might prefer from the previous stage for further recommendation. Both [19] and [20] can handle the cold-start problem well but do not consider much about the cross-domain problem.

In this paper we apply a different strategy: in the first stage, some relatively good results are chosen by observing the text data; then we refine the results in our second stage, with the help of image data. In our previous paper [21], we provide a two-stage recommendation and each stage utilises data from different domains by alignment and co-clustering. However, co-clustering method lacks the ability to tell the intimacy distance between two individuals exactly but only to group people roughly with similar properties, and thus can not make precise recommendation. To overcome this problem, in this paper we propose a probabilistic topic model in the second stage for a better recommendation. We also provide a novel and more precise method to solve the topic model problem.

B. Probabilistic Topic Model

In the second stage of our model, a topic model is applied to get a more precise recommendation.

1) Topic Model in Recommendation: The probabilistic topic model is a successful approach solving the problem for information retrieval[10] and recommendation[7]–[9]. For example, [8] recommends temporary friends to users by building models that contain latent variables that illustrate users' interests change with time.

By assuming some latent factors it is conceptually easy to build the relationships among different domains. [7] designs a model that connects the Flickr and Foursquare data for image, topic and item recommendation. It assumes that both domains have some common latent factors and each domain also has its own latent factors, and the users' activities on these two platforms are the synergism of all these factors. Gibbs sampling is applied to find the value of the latent factors. [9] considers the friendships and the votings on the large Film rating website. To predict individual's flavour about films his/her social relationships and scores of films are combined with some latent factors. Variational methods are applied to solve the model.

To make the model to illustrate the situation of the real world more accurately and reasonably, both [7] and [9] make many assumptions of the latent topic and thus contain many unknown parameters to infer: [9] contains more than 10 unknown parameters and [7] has more than 30. The presence of so many unknown variables not only greatly increases the complexity of the algorithm, but also leads to other problems such as overfitting or redundancies. In this paper we build the model in a more compact manner.

2) Gibbs Sampling, Variational Inference and Matrix Factorization: Due to the coupling of latent variables, the direct inference is usually impossible for a specific topic model. Generally there are two methods to find a solution for topic model: Gibbs sampling [22] and variational inference [10]. For some complex multivariate probability distributions, to determine the parameters of the distribution, direct sampling is difficult. Gibbs sampling samples the marginal distribution of one variable each time, and iteratively samples all the marginal distributions. The variational method, on the other hand, approaches the solution by approximating the original complex distribution with a factorised one, which is easier to handle.

As stated in Section I, both of the two methods have some weaknesses: Gibbs sampling has difficulties in handling big data problems, and the variational method can not determine if the approximation is close to the original one. Some researchers consider combining the two in one problem: In [12], small counts of data are sampled and the variational method is applied to update large counts, which improves the performance on the large dataset. However, how accurate the approximation of variational method is not yet discussed in [12]. In this paper, we propose a new solution to a topic model by directly calculating the distribution of the latent variables.

Compared with the above two methods, MF-based method also assumes some latent variables but instead of determining the marginal distribution of the observed data, it factorizes the observed data into different latent factors, which leads to some computational convenience and efficiency. Both of [13] and [14] utilize user friendship network and user-item network and obtain some latent factors that show the preference of individuals. The recommendations based on these latent factors are relatively effective. On the other hand, they do not try to find the probabilistic distribution of the network and all of these methods apply some gradient descent methods, that are relatively easy to be trapped into a local optimum. Our method avoids this drawback by deducing the distribution of the whole probabilistic model.

C. Algebra of Random Variables

The essential problem of our approach in this paper is to get the exact mathematical expression of the coupling of different random variables, mainly the sum and product of random variables. These problems were extensively discussed in the 1950s to 1970s year, last century, during which time the random process was a hot research topic but the computer simulation technology was not well developed. In [23]–[25], the products of typical distributions such as Beta, Gamma and Rayleigh are discussed. Most of these works utilise the Mellin transform [26] as the essential tool for deducing. [27] gives a good summary of these works and also discusses the distribution of the sum of random variables. The algebra of random variables has also been studied recently in certain fields such as wireless communication in [28] and [29]. These works show that the product and quotient of random variables with certain distributions can be expressed analytically. We will mainly apply some of the results in [25], [27] later in our work. As Gaussian distribution has some good properties(its domain of definition is all the real values, and has a central point, etc.) we assume that our latent variables to be Gaussian distributed.

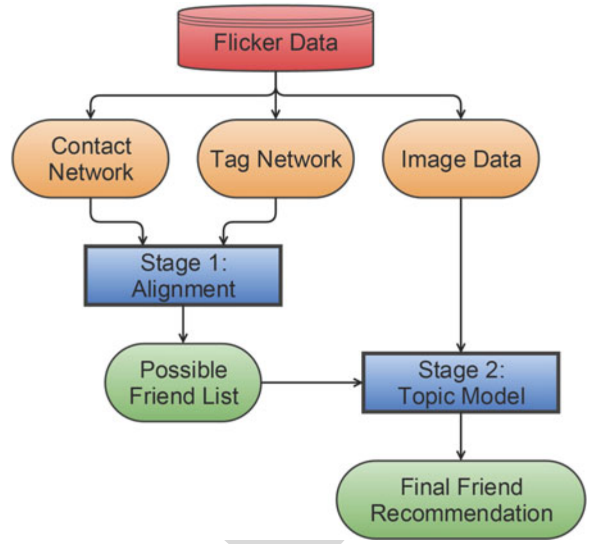


Fig. 1. Two-stage system illustration.

III. System Model

The main framework of our model is shown in Fig. 1, which contains two stages: In the first stage, network alignment is applied to generate a possible friend list, by correlating the tag and contact data in Flickr; In the second stage, the user-uploaded image features generate some topics by utilising a probabilistic topic model, and a new method is developed to solve the model for precise friend recommendation.

A. First Stage: Network Alignment

The detailed alignment method has been discussed in [5]. The following is an introduction of its basic idea. An individual may join different social networks for different purposes. For example, one may at the same time join a football fan network for physical practice and a restaurant information sharing network to look for the best food. He/she plays different social roles in these different networks, and might make different friends. However, these different social roles for one individual are not independent, but related to each other.(The man might look for some food that helps quickly recuperate after hard physical practice). The motivation for social network alignment lies on the fact that these different networks, though having different edges (relationships), are usually related to each other. Taking Flickr as an example, according to the uploaded-tag-similarity of each user and their contact list, a tag similarity network and a contact network are formed. Although the topologies of the two networks are not the same, they are related to each other, for users uploading similar tags on Flickr have higher probability to make friends with each other. By digging the correlation of the topologies of different networks we may make inference for the knowledge from one domain to another.

Specifically, we align the Flickr tag-similarity network with the contact network, so that after the alignment, one tight edge between two users in the tag similarity network would imply that these two users have higher probability to have contact with each other. We align the tag-similarity network with the contact network by selecting important tag features. The reason we apply feature selection here lies in the phenomenon that when we look for online friends, it is common that we do not take care of all the factors of a person but concentrate on certain points that would interest ourselves. As an example, a traveller might post his photos with the following tags: “Sydney”, “Blue Mountain”, “great view”, and “street”. Among these tags some people might contact him/her for some more details about the experiences in ‘’Sydney” and ‘’Blue Mountain”, but seldom would have interests about “great view” or “street” because they are too common. We can treat these two tags as redundancy for friend making. Based on this observation, we believe that some Flickr tags can be more indicative in the task of friend recommendation, because they are more important to reflect the connections on the contact network. We can treat these tags as important features for friend recommendation. Inspired by this phenomenon, we design a method to choose some important features that are more helpful for friend making decision.

Mathematically, assume that the feature selection matrix to be $\mathbf{W}$, the known tag-user matrix to be $\mathbf{X}$, the tag distance matrix to be $\mathbf{L}$, and the first d eigenvector-matrix of the contact network to be $\mathbf{V}$, the important feature can be obtained by solving the following problem:

$$\min_{\mathbf{W}} \|\mathbf{XW} - \mathbf{V}\| + \mu tr(\mathbf{W}^T\mathbf{X}^T\mathbf{LXW}) + \lambda\|\mathbf{W}\|_{2,1}. \quad (1)$$

The first term of (1) aligns the tag-similarity network to the contact network so that they become more similar to each other, and the second term preserves the local structure of the original tag-similarity network. The third term is for regularization. In this way the tag feature selection matrix $\mathbf{W}$ makes the topology of the tag-similarity network more similar to the contact network, while preserving the topology of the tag-similarity network as much as possible. In other words, we align the tag-similarity network to the contact network. By comparing the similarity of two users on the those important tags we can generate a possible friend list for each user. The solution of $\mathbf{W}$ in (1) is discussed more thoroughly in [5].

B. Second Stage: Topic Model

In the previous stage we get a possible friend list by considering the correlation between the tag and contact networks on Flickr. However, as the real world friend relationship is affected by many factors[1], one stage is usually not enough for a precise friend recommendation. In the following stage, we introduce the image data as auxiliary information to refine the recommendation list.

We apply the topic model to combine the image data and the friendships in Flickr. It is common sense that a person uploads a photo on Flickr because he/she likes the photo. Why does he/she like the photo? We assume that in one’s mind, some latent interest factors control his/her taste of image. For example, some people like colourful, vivid photos, while others prefer spectacular or imposing ones; children enjoy comic-style pictures while adults have more interests in realistic-style paintings; young women pay much attention to photos of beautiful clothes while young men to electrical devices. These latent factors are determined by various aspects such as age, gender, living experiences, etc. and can not be observed or simply summarised. We assume individuals’ interest latent factor to be $\mathbf{v}$. Each image contains the factors that attract people, such as colour, line, or history, which we assume to be $\mathbf{a}$. The correlation of $\mathbf{v}$ and $\mathbf{a}$ determines whether a user would upload an image.

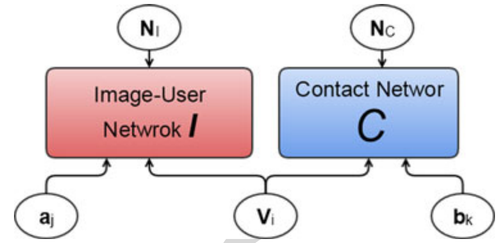


Fig. 2. Probabilistic topic model combining image-user network and contact network.

Similarly, we assume that each user exhibits some attractive factors during his/her activities in Flickr such as uploading photos, writing descriptions of photos and making comments, etc. We also summarise these attractive factors with the third latent variable $\mathbf{b}$. Notice that the same user’s interest latent factor $\mathbf{v}$ and attractive factor $\mathbf{b}$ are not the same. The combination of $\mathbf{b}$ and $\mathbf{v}$ determines whether two users should make friends with each other. For simplicity we view them as independent from each other. The topic model is shown in Fig. 2.

In Fig. 2, $\mathbf{C}$ and $\mathbf{I}$ stand for the $0-1$ contact network and image-user network, respectively. $\mathbf{C}$ is an $n \times n$ matrix where n is the number of users. $\mathbf{I}$ is an $n \times f$ matrix where f stands for the number of total features. For $\mathbf{C}$, if user k and user j are friends with other, then $\mathbf{C}_{\mathrm{kj}}$ equals one, and zero otherwise. For $\mathbf{I}$, if the uploaded photos of user i contain an image feature j, then $\mathbf{I}_{\mathrm{ij}}$ equals one, and zero otherwise. a stands for image factor, and b stands for individuals’ social interest factor, respectively. v stands for individuals’ common interest factor that has effect on both his choice of images and friends. $\mathbf{N_I}$ and $\mathbf{N_C}$ stand for zero-mean additive noises. The relationship can be mathematically expressed as follows:

$$\mathbf{I}_{\mathrm{ij}} = a_i \times v_j + \mathbf{N}_{\mathrm{Iij}}, \mathbf{C}_{\mathrm{kj}} = b_k \times v_j + \mathbf{N}_{\mathrm{Ckj}}. \quad (2)$$

We assume that all the latent random variables a_i, b_k and v_j are Gaussian distributed with the parameters of means and variances of μ_a, σ_a, μ_b, σ_b, μ_v, and σ_v, respectively. The reason we choose Gaussian distribution is as follows: Although some other distributions that are in the form of an H-function (such as Beta, Gamma or Rayleigh distributions) would lead to some calculation convenience [27], we assume Gaussian distribution here because it is defined on the whole real domain and contains negative values and has a central point, while other distributions such as Beta are only defined on the positive real domain.

The coupling between random variables a, b and v makes the integral of (2) often intractable. Traditional methods dealing with (2) contain Gibbs sampling [11] and variational inference [10]. Gibbs sampling meets with difficulties when the data scale is large, and the variational method applies some approximation that the precision is hard to control. In the following we develop a new approach to solve (2) that is based on Mellin transform and series expansion.

IV. Series Expansion

A. Product of Gaussian Random Variables

When dealing with the distribution of product of random variables, the Mellin transform is an essential tool [27]. We take the first equation in (2) to explain its basic idea. For simplicity we first neglect the noise term $\mathbf{N}_{\rm ij}$ (its effectiveness is to be discussed later) and we have $\mathbf{I}_{\rm ij} = a_i v_j$ for two random variables a_i and v_j with different probability distribution functions. One useful property for Mellin transofrm is: the Mellin transform of the product of two probability density functions (PDF) is equal to the product of the Mellin transforms of their PDFs.

Mathematically, we recall the following rule [27]: If a_i and v_j are two non-negative random variables with the PDFs $f_a(a_i)$ and $f_v(v_j)$, their product $\mathbf{I}_{\rm ij} = a_i v_j$ has a distribution $h(I_{\rm ij})$, and then the Mellin transform of $h(I_{\rm ij})$ is precisely the product of Mellin transform of $f_a(a_i)$ and $f_v(v_j)$, respectively. The expression is given as

$$\mathcal{M}(h(\mathbf{I}_{\rm ij})) = \mathcal{M}(f_a(a_i))\mathcal{M}(f_v(v_j)) \tag{3}$$

where the Mellin transform and its inverse of an analytical function $f(x)$ are defined as follows:

$$\mathcal{M}(s) = \int_0^{+\infty} x^{s-1} f(x)dx \tag{4}$$

$$\mathcal{M}^{-1}(x) = \frac{1}{2\pi i}\int_{c-i\infty}^{c+i\infty} x^{-s}\mathcal{M}(s)ds \tag{5}$$

where c in (5) stands for an arbitrary real number. With the help of (3)–(5) and the known distribution of a_i and v_j, we can give an exact mathematical expression for distribution of the coupling of the two random variables a_i and v_i.

In this way we can first deduce the Mellin transform of each of the distributions, then make product of the two, and finally inverse the Mellin transform to get the final product distribution. In this way, we first calculate the distribution of $\mathbf{I}$ in (2).

From the previous assumption we know that a_i, b_k and v_j follow the Gaussian distribution with mean $\mu_{\rm ai}$, $\mu_{\rm bk}$, $\mu_{\rm vj}$ and the variance $\sigma_{\rm ai}$, $\sigma_{\rm bk}$, $\sigma_{\rm vj}$. We further take the symbol of $f_{\rm ai}$, $f_{\rm bk}$ and $f_{\rm vj}$ as their PDFs. We first do the Mellin transform on a_i and v_j separately to get $\mathcal{M}(f_a(a_i))$ and $\mathcal{M}(f_v(v_j))$, and then we product them and do the inverse Mellin transform to finally get the distribution of product of two random variables, which is the distribution of the variables in image-user matrix $\mathbf{I}$. The details are given in [27] and [25], which provide two equivalent expressions for the distribution of two Gaussian random variables. We apply the expression from [25] and the details are briefly outlined in the following.

To calculate the distribution of $\mathbf{I}_{\rm ij} = a_i v_j$ with Gaussian random variables a_i and v_j, we take the Mellin tranform of $f_a(a_i)$ and $f_v(v_j)$. Notice that according to (4), the positive and negative parts of the distribution of a_i and v_j should be considered separately. We apply the property that the Mellin transform of the standard Gaussian distribution is Gamma function[30]: $\mathcal{M}\{\mathrm{e}^{-x^2/2}\} = 2^{s/2-1}\Gamma(s/2)$, and a non-central Gaussian distribution can be expressed as a standard Gaussian distribution multiplied by a series in the form: $\mathrm{e}^{-\frac{1}{2}(x-\mu)^2} = \mathrm{e}^{-\mu^2/2}\sum_{j=0}^{\infty}\frac{1}{j!}\mu^j x^j \mathrm{e}^{-x^2/2}$. If we define the following:

$$\begin{aligned}
a_{i1} &= max(a_i,0), v_{i1} = max(v_j,0)\\
a_{i2} &= min(a_i,0), v_{i2} = min(v_j,0)\\
\mathbf{I}_{ij-1} &= a_{i1}v_{i1}, \mathbf{I}_{ij-2} = a_{i1}v_{v2}\\
\mathbf{I}_{ij-3} &= a_{i2}v_{v1}, \mathbf{I}_{ij-4} = a_{i2}v_{j2}.
\end{aligned}$$

And we also define the probability distribution function of $\mathbf{I}_{ij-1}$, $\mathbf{I}_{ij-2}$, $\mathbf{I}_{ij-3}$, and $\mathbf{I}_{ij-4}$ to be $h_1(\mathbf{I}_{\rm ij})$, $h_2(\mathbf{I}_{\rm ij})$, $h_3(\mathbf{I}_{\rm ij})$ and $h_4(\mathbf{I}_{\rm ij})$, respectively. Following the methods of [25], and taking $\mathbf{I}_{ij-1}$ as an example, we have

$$\mathcal{M}_{\mathbf{I}_{ij-1}}(s) = \sum_{o=0}^{\infty}\frac{\mu_{a_i}^{2o}}{(2o)!}\frac{\mu_{v_j}^{2o}}{(2o)!}\Gamma^2(s). \tag{6}$$

To get the distribution of $\mathbf{I}_{ij-1}$, we do the inverse Mellin transform of (6) as:

$$h_1(\mathbf{I}_{\rm ij}) = \sum_{o=0}^{\infty}\left(\frac{1}{2\pi i}\right)\int_{c-i\infty}^{c+i\infty}(y^2)^{-s}\frac{\mu_{a_i}^{2o}}{2o!}\frac{\mu_{v_j}^{2o}}{2o!}\Gamma^2(s+o)ds. \tag{7}$$

Equation (7) is an integral on half of the complex plane. According to Residue Theorem [31], the solution is expressed with the infinite residues that are related to the poles on the real plane. By calculating the residues we get (8), shown at the bottom of the page, where $C1 = \frac{1}{\pi}\mathrm{e}^{-\frac{1}{2}(\frac{\mu_{\rm ai}^2}{\sigma_{\rm ai}}+\frac{\mu_{\rm vj}^2}{\sigma_{\rm vj}})}$, $C2 = ((\frac{1}{(2o)!})^2(2\frac{\mu_{\rm ai}^2}{\sigma_{\rm ai}}\frac{\mu_{\rm vj}^2}{\sigma_{\rm vj}})^o)$, and $\psi(1)$ is the Euler-Mascheroni constant.

Similarly we should also consider the case of $h_2(\mathbf{I}_{\rm ij})$ for $a > 0 \cap v < 0$, $h_3(\mathbf{I}_{\rm ij})$ for $a < 0 \cap v > 0$, and $h_4(\mathbf{I}_{\rm ij})$ for $a <$

$$h_1(\mathbf{I}_{\rm ij}) = C1\left[\sum_{o=0}^{\infty}C2\sum_{s=o}^{\infty}\left[\frac{(\mathbf{I}_{\rm ij})^{2s}}{\prod_{t=0}^{s-o-1}(-s+o+t)^2}\left(2\psi(1) - 2\sum_{w=0}^{s-o-1}\frac{1}{-s+o+w}\right) - \frac{(\mathbf{I}_{\rm ij})^{2s}\ln((\mathbf{I}_{\rm ij})^2)}{\prod_{w=0}^{s-o-1}(-s+o+w)^2}\right]\right] \tag{8}$$

$0 \cap v < 0$. To sum up, we have

$$\begin{aligned} h(\mathbf{I}_{\mathrm{ij}}) &= h_1(\mathbf{I}_{\mathrm{ij}}) + h_2(\mathbf{I}_{\mathrm{ij}}) \qquad (y > 0) \\ &= h_3(\mathbf{I}_{\mathrm{ij}}) + h_4(\mathbf{I}_{\mathrm{ij}}) \qquad (y < 0). \end{aligned} \tag{9}$$

In a similar manner we can give the expression for $h(\mathbf{C}_{\mathrm{kj}})$

Here we give a short discussion about this series. In the first place, this is basically an alternating and power series [32] with infinite terms, with some of the terms multiplied with a logarithm factor. This is a series that when the sequence number of the term increases, the absolute value of the term increases. Some of the terms are positive and some are negative, and the sum of the terms eventually becomes convergent, as discussed in [33]. However, similar to some of the convergent Taylor series, when the absolute value of the series terms is large, these series converge only when the term number of the series is also large. In order to make the series to converge rapidly with relatively a small number of terms, in practice, we may normalise the value of $\mathbf{I}_{\mathrm{ij}}$ to be relatively small (In the experiments, the ground truth of $\mathbf{I}_{\mathrm{ij}}$ and $\mathbf{C}_{\mathrm{ij}}$ are 0 or 1, which is small enough).

B. *Additive Noise*

From Fig. 2 we see that after the products of a, v and b, v, the results should also add a bias value or noise to get the value of $\mathbf{I}_{\mathrm{ij}}$ and $\mathbf{C}_{\mathrm{kj}}$. In practice it can be interpreted as all the outer environmental influences other than the users and the items. For example, the change of seasons for the favour of clothing, or the change of temperature for the preference of food, etc. Mathematically the PDF of two independent random variables are the convolution of their PDFs of the two [27]. In our case, we can simply consider the environmental influences $\mathbf{N}_I$ and $\mathbf{N}_C$ to be independent from the image factor a, social attractive factor b and individual's latent factor v. For simplicity we assume the additive noise of $\mathbf{N}_I$ and $\mathbf{N}_C$ to be Gaussian distributed with zero mean and variance of σ_{Ni} and σ_{Nc}, respectively. Taking $\mathbf{I}_{\mathrm{ij}}$ for example, from (8) we see that the most important calculation is the convolution of the Gaussian function from additive noise $\mathrm{e}^{-\mathbf{I}_{\mathrm{ij}}^2/\sigma_{\mathrm{Ni}}^2}$ and the term ${\mathbf{I}_{\mathrm{ij}}}^{2s}\log({\mathbf{I}_{\mathrm{ij}}}^2)$ from (8), which is formally written as follows:

$$d_2(\mathbf{I}_{\mathrm{ij}}) = \mathrm{e}^{-\mathbf{I}_{\mathrm{ij}}^2/\sigma_{\mathrm{Ni}}} * \mathbf{I}_{\mathrm{ij}}^{2s}\ln(\mathbf{I}_{\mathrm{ij}}^2). \tag{10}$$

By calculating the convolution we see (10) can be expressed as follows:

$$d_2(\mathbf{I}_{\mathrm{ij}}) = {\mathbf{I}_{\mathrm{ij}}}^{2s+2}\left(\frac{\ln {\mathbf{I}_{\mathrm{ij}}}^2}{2s+2} - \frac{1}{(2s+2)^2}\right)\mathrm{e}^{\frac{(-{\mathbf{I}_{\mathrm{ij}}}^2)}{\sigma_{\mathrm{Ni}}}}. \tag{11}$$

In this way we can get a series expression of (9).

$$\begin{aligned}
h\left(\mathbf{I}_{\mathrm{ij}}\right) = {} & \frac{1}{\pi} e^{-\frac{1}{2}\left(\frac{\mu_{\mathrm{ai}}^2}{\sigma_{\mathrm{ai}}} + \frac{\mu_{\mathrm{vj}}^2}{\sigma_{\mathrm{vj}}}\right)} \Bigg[\Bigg[\Bigg[\sum_{t=0}^{\infty}\left(\frac{1}{(2t)!}\frac{1}{(2t)!}\left(2\frac{\mu_{\mathrm{ai}}^2}{\sigma_{\mathrm{ai}}}\right)^t\left(2\frac{\mu_{\mathrm{vj}}^2}{\sigma_{\mathrm{vj}}}\right)^t\right)\sum_{s=t}^{\infty}\Bigg[\frac{\mathbf{I}_{\mathrm{ij}}^{2s}\mathrm{e}^{\frac{(-{\mathbf{I}_{\mathrm{ij}}}^2)}{\sigma_{N}{}^i}}}{\prod_{m=0}^{s-t-1}(-s+t+m)^2}\Bigg(2\psi(1) - 2\sum_{m=0}^{s-j-1}\frac{1}{-s+t+m} \\
& - \frac{\mathbf{I}_{\mathrm{ij}}^{2s}\ln(\mathbf{I}_{\mathrm{ij}}^2)}{\prod_{i=0}^{s-t-1}(-s+t+m)^2}\Bigg] + \sum_{r=0}^{\infty}\sum_{t=0}^{\infty}\left(\frac{1}{(2t)!}\left(2\frac{\mu_{\mathrm{ai}}^2}{\sigma_{\mathrm{ai}}}\right)^t\frac{1}{(2r)!}\left(2\frac{\mu_{\mathrm{vj}}^2}{\sigma_{\mathrm{vj}}}\right)^r + \frac{1}{(2t)!}\left(2\frac{\mu_{\mathrm{vj}}^2}{\sigma_{\mathrm{vj}}}\right)^t\frac{1}{(2r)!}\left(2\frac{\mu_{\mathrm{ai}}^2}{\sigma_{\mathrm{ai}}}\right)^r\right)\sum_{s=t}^{r-i} \\
& \times \left[\mathbf{I}_{\mathrm{ij}}^{2s}\frac{\prod_{m=1}^{r-s-1} m}{\prod_{q=0}^{s-j-1} -q-1}\ln(\mathbf{I}_{\mathrm{ij}}^2)\right] + \sum_{t=0}^{\infty}\sum_{r=t+1}^{\infty}\left(\frac{1}{(2t)!}\left(2\frac{\mu_{\mathrm{ai}}^2}{\sigma_{\mathrm{ai}}}\right)^t\frac{1}{(2r)!}\left(2\frac{\mu_{\mathrm{vj}}^2}{\sigma_{\mathrm{vj}}}\right)^r + \frac{1}{(2t)!}\left(2\frac{\mu_{\mathrm{vj}}^2}{\sigma_{\mathrm{vj}}}\right)^j\frac{1}{(2r)!}\left(2\frac{\mu_{\mathrm{ai}}^2}{\sigma_{\mathrm{ai}}}\right)^{2r}\right) \\
& \times \sum_{s=r}^{\infty}\left[\frac{\mathbf{I}_{\mathrm{ij}}^{2s}(2\psi(1) - \sum_{m=0}^{r-t-1}\frac{1}{-s+t+m} - \sum_{q=r-t}^{s-t-1}\frac{2}{-s+t+q})}{\prod_{m=0}^{r-t-1}(-s+t+m)\prod_{q=r-t}^{s-t-1}\prod_{q=r-t}^{s-t-1}(-s+t+q)^2} - \frac{\mathbf{I}_{\mathrm{ij}}^{2s}\ln(\mathbf{I}_{\mathrm{ij}}^2)\mathrm{e}^{\frac{(-{\mathbf{I}_{\mathrm{ij}}}^2)}{\sigma_{N}{}^i}}}{\prod_{i=0}^{r-j-1}(-s+j+i)\prod_{k=r-1}^{s-j-1}(-s+j+k)^2}\right]\Bigg] \\
& \pm\Bigg[\sum_{k=0}^{\infty}\left(\frac{1}{(2k+1)!}\frac{1}{(2k+1)!}\right)\sum_{s=k}^{\infty}\Bigg[\frac{(\mathbf{I}_{\mathrm{ij}}^2)^{s+1/2}}{\prod_{m=0}^{s-q-1}(-s+q+m)^2}\left(2\psi(1) - s\sum_{m=0}^{s-q-1}\frac{1}{-s+q+m}\right) - \frac{(\mathbf{I}_{\mathrm{ij}}^2)^{s+1/2}\ln(\mathbf{I}_{\mathrm{ij}}^2)\mathrm{e}^{\frac{(-{\mathbf{I}_{\mathrm{ij}}}^2)}{\sigma_{N}{}^i}}}{\prod_{i=0}^{s-q-1}(-s+q+i)^2}\Bigg] \\
& + \sum_{p=1}^{\infty}\sum_{q=0}^{p-1}\left(\frac{1}{(2q+1)!}\left(2\frac{\mu_{\mathrm{ai}}^2}{\sigma_{\mathrm{ai}}}\right)^{(p+0.5)}\left(2\frac{\mu_{\mathrm{vj}}^2}{\sigma_{\mathrm{vj}}}\right)^{p+0.5} + \frac{1}{(2q+1)!}\left(2\frac{\mu_{\mathrm{vj}}^2}{\sigma_{\mathrm{vj}}}\right)^{q+0.5}\frac{1}{(2p+1)}\left(2\frac{\mu_{\mathrm{ai}}^2}{\sigma_{\mathrm{ai}}}\right)^{p+0.5}\right)\sum_{s=q}^{p-1} \\
& \times \left[\frac{(\mathbf{I}_{\mathrm{ij}}^2)^{(}s+1/2)\prod_{m=1}^{p-s-m}(m)}{\prod_{n=0}^{s-q-1}(-n-1)}\ln(\mathbf{I}_{\mathrm{ij}}^2)\right] + \frac{1}{(2q+1)!}\left(2\frac{\mu_{\mathrm{vj}}^2}{\sigma_{\mathrm{vj}}}\right)^{q+0.5}\frac{1}{(2p+1)!}\left(2\frac{\mu_{\mathrm{ai}}^2}{\sigma_{\mathrm{ai}}}\right)^{p+0.5}\Bigg)\sum_{s=p}^{\infty} \\
& \times \left[\frac{(\mathbf{I}_{\mathrm{ij}}^2)^{s+1/2}(2\psi(1) - \sum_{m=0}^{p-q-1}\frac{1}{-s+q+1} - \sum_{l=p-q}^{s-q-1}\frac{2}{-s+q+l})}{\prod_{i=0}^{p-q-1}(-s+q+m)\prod_{l=p-q}^{s-q-1}(-s+q+l)^2} - \frac{(\mathbf{I}_{\mathrm{ij}}^2)^{s+1/2}\ln(\mathbf{I}_{\mathrm{ij}}^2)\mathrm{e}^{\frac{(-{\mathbf{I}_{\mathrm{ij}}}^2)}{\sigma_{N}{}^i}}}{\prod_{i=0}^{p-q-1}(-s+q+m)\prod_{l=p-q}^{s-q-1}(-s+q+l)^2}\right]\Bigg]\Bigg]
\end{aligned} \tag{12}$$

TABLE I
SUMMARY OF PARAMETERS

μ_{ai}	mean of image factor a_i
μ_{bk}	mean of individuals' social attractive factor b_k
μ_{vj}	mean of individual's interest factor v_j
σ_{NI}	variance of image noise N_I
σ_{NC}	variance of social noise N_C

So the expression for the distribution of $\mathbf{I}_{\mathrm{ij}}$ considering the additive noise is given in (12), shown at the bottom of the page.

In a similar way we can also obtain the distribution of $\mathbf{C}_{\mathrm{jk}}$.

C. EM for Parameter Estimation

Applying the above, we obtain the exact infinite expansion expression of the PDF. of $\mathbf{I}$ in a series form given in (12). The expression of $\mathbf{C}_{\mathrm{kj}}$ can be obtained in a similar way. From (12) we can see that the exact value of μ_{ai} and σ_{ai} does not matter much, but the value of $\frac{\mu_{\mathrm{ai}}^2}{\sigma_{\mathrm{ai}}}$ matters. So we can assume that a_i has standard derivation of 1, and we only need to calculate the average value of a_i. Similarly, we also do not need to calculate σ_{bk} and σ_{vj} but only assume that v_j and b_k have standard derivation.

All the parameters $\mathcal{P}$ are summarised in Table I. As mentioned in Section IV-A, in the experiments, when we choose the starting point of the parameters not too large, we can make the series converge in a relatively small number of terms. Then we can apply the standard EM method to refine the parameters iteratively. Experimental result shows that the number of series terms can be no longer than 10 and after several EM iterations, the precision becomes stable.

The EM training process is introduced as follows. For E step, Consider Eq.(12), which is the Equation we want to maximize by knowing the value of $\mathbf{I}_{\mathrm{ij}}$, with respect to the parameters $\mathcal{P}$ as follow:

$$\max_{\mathcal{P}} h(\mathbf{I} \mid \mathcal{P})). \tag{13}$$

In the M step, we find the derivative of each parameter in $\mathcal{P}$ by fixing other parameters. Then we set the derivative to be zero to get the value for each parameter. The whole process goes until convergence.

One problem to solve (12) is that (12) contains not only polynomial terms but also exponential terms for the parameters. For simplicity we can make an assumption that the parameters are relatively small, and then we can use the first several terms, or following [34] to get a polynomial expression of the parameters, to make (12) solvable.

Another problem is that for some parameters such as μ_{ai}, it contains infinity high order terms that makes the solution intractable. Again we can make the assumption that these parameters to be smaller than one, and discard the high order terms. In practice we keep the terms whose orders are equal or lower than 4, and follow the method discussed in [35] to calculate the values of the parameters.

From (12) we can obtain the parameters that related to the image-user matrix $\mathbf{I}$, such as μ_{ai}, μ_{vj}, and σ_{NI}. In a similar manner we can also get the parameters related to the contact matrix $\mathbf{C}$, such as σ_{NC}, μ_{bk}, and also μ_{vj}. By iteratively updating these parameters relating to the two matrix we can finally determine the value of all the parameters.

After the EM iterations we fix all the parameters in Table I and according these parameters we can make the final friend recommendation.

D. Recommendation Method

When a new user i comes into the network, he/she may upload some favourite photos as well as some tags. The recommendation procedure is divided in two stages. In the first stage, a list of possible friends is generated according to the similarity of the selected important tags. In the experiments, we put the top 200 users into the list.

In the second stage, according to the features of the images uploaded by use i, we get the individuals' interest factor v_i of this user. For a user k in the possible friend list obtained from the first stage, we can also calculate his/her attractive and interest factors b_k. The similarity score of user i and k is obtained by $S_{\mathrm{ik}} = v_i b_k$. The higher the similarity score, the more likely that they are to be friends. So we can rank the 200 users in the list according to the similarity score with user i, and recommend the top ones as user i's friends

The whole procedure is given in Algorithm 1.

E. Complexity Analysis

The complexity analysis of our algorithm is also divided by the two stages as follows:

Considering the first stage, the complexity of the network alignment is mainly decided by two steps: the eigenvalue calculation and the inverse of the similarity matrix, which is given by $\max(\min\{n, e\}^3, dn^2)$ as discussed in [5]. e stands for the number of total tags. As previous defined, n stands for the number of users, and d stands for the first d eigenvectors.

To solve the topic model of the second stage, Assume together we need to make L time iterations. in each iteration of the EM step, assume that we calculate the first g terms of the series of (12) (In practice we make $g = 4$). And it takes e steps to solve a 4th order polynomial equation, as mentioned in Section IV-C. Then the complexity would be of $\mathcal{O}(L * e * g * (n * f + n * n))$, where f is the number of image features, as previously defined.

V. EXPERIMENTS

In this section, we make experiments to show the advantage of our proposed method. First, we introduce our social media dataset, and then we discuss the results of our algorithm by comparing it with reference methods. We utilise a cluster containing 16 cores and 128G memories to run our experiments.

A. Dataset and Feature Extraction

We crawled a social network from the big image sharing site Flickr. As the data set is quite large, a relatively unbiased dataset was obtained. In total we crawled the data of 30000 users, and for each user, we crawled all their photos, and tags of each photo. In this paper we tried the SIFT feature and the deep network

TABLE II
DATASET STATISTICS

Users	30000
Photos	1,356,293 photos from 30000 users
CNN features	4096
Contact	628,153 friend links among users
Tags	42,739 words after filtering

extracted features through an CNN autocoder realized by Caffe [36]. For the CNN features we follow the steps of the widely used AlexConvNet [37] and use the 4096 dimensional features vectors from the last full-connected layer. In most cases the CNN features performs better than the SIFT features, so we chose the CNN extracted features for the rest of our experiments. In the future we can also refine feature extraction method for better performance. We then crawled the user contact information to form the contact network. The contact information in Flickr was acquired by checking if a user added another user to his/her friend list, or vice versa. We crawled all the contacts between any two users in our dataset. A short summary of our dataset is given in Table II.

B. Settings and Metrics

Our task is to make precise contact information prediction. When a new user enters into the social network, we recommend new friends according to key words and photos that represent the user's interests.

In friend recommendation, assume we recommend T friends to each user. We use the existing contact information as the ground truth for training and testing. In the first stage, the parameter μ of (1) is determined on the training set by a four-fold cross validation to find the best. The range for the parameter is: $\mu \in 10^{[-2:1:3]}$.

We use the method summarised in Algorithm 1 to recommend friends to new users. We use the recommendation precision metrics to show the effectiveness of the proposed algorithm. In our experiment, precision is defined as the number of correctly recommended friends divided by all the recommended users. We also introduce the precision-recall curve to further show the advantage of our algorithm, where recall is defined as the number of the correctly recommended friends divided by the number of all friends.

During our experiments we divide the whole usere set randomly into two groups: 4/5 of all the users are in the training set and the rest are in the test set. The important features in the first stage are selected on the training set, where the parameters in the second stage are also trained. When a new user in the test set comes into the system with some uploaded tags and photos, T friends will be recommended to him/her from the training set. Assume that together we have recommended $RecAll$ real friends to the test users (totally 6000 users), then the overall precision is calculated by $RecAll/(6000 \times T)$. We adopt a five-fold cross validation to ensure that all the users are utilised as training and testing data once

Algorithm 1: Two Stage Friend Recommendation.

Input:
tag feature matrix $\mathbf{T}$, contact matrix $\mathbf{C}$, image-user matrix $\mathbf{I}$, tag and image feature of the new user $\mathbf{t}$ and $\mathbf{i}$, the numbers of possible friends in Stage 1 and final friends $k1$ and k, respectively

Output:
Friend recommendation list of the new friend

Training:
Stage I
1: Determine λ and μ in (1) via cross validation.
2: Solve (1) with the method in [5]
Stage II
3: Generate the expression of distribution of h[(12)] in the form of series.
4: Apply EM method determining the parameters in Table I

Testing:
5: Stage I: Use $\mathbf{W}$ calculated in Step 2 to obtain $k1$ possible friend list.
6: Stage II: Use the parameters in Step 4 to refine the final recommendation friend list,recommend top k users

C. Reference Methods

The performance analysis of our first stage: network alignment methods can be seen in some previous related papers such as [21], [5]. For the performance analysis of the second stage in which the topic model method is applied, we choose several widely-used methods for comparison.

The first is the variational method, which has been widely applied in this decade for solving the Bayesian network problem[10]. Basically we apply the methods in [9] with some slight modifications to our problem.

The second is the widely-used Gibbs sampling method, which is also very popular in dealing with topic model. Compared with the variational method, the idea of Gibbs sampling is simpler but usually it has difficulty in dealing with large scale problems. We apply the method based on [7] for comparison.

The third method is a co-clustering based method [21]. It is not a topic model-based method, but has a relatively simpler concept: In the second stage, we do co-clustering of image features, users and tags to get a . We apply a simple ranking method, similar to [21] for the final friend recommendation.

To further check the advantage of our method, we also compare our whole two-stage recommendation algorithm with several state-of-the-art recommendation systems. The first one is based on matrix factorization(MF). MF method decomposes the item-user or user-user matrix to infer the latent factors that catch individuals' interests and has been widely discussed for different kinds of recommendation prolems[13], [14]. In this paper we apply a recent method proposed in [14] for comparison, for it jointly considers the information from two different domains.

Another recent method is based on Bayesian collaborative filtering that takes the social connections into account, called SBPR [38]. As a widely-used recommendation method, collaborative filtering assumes that two users that choose the same

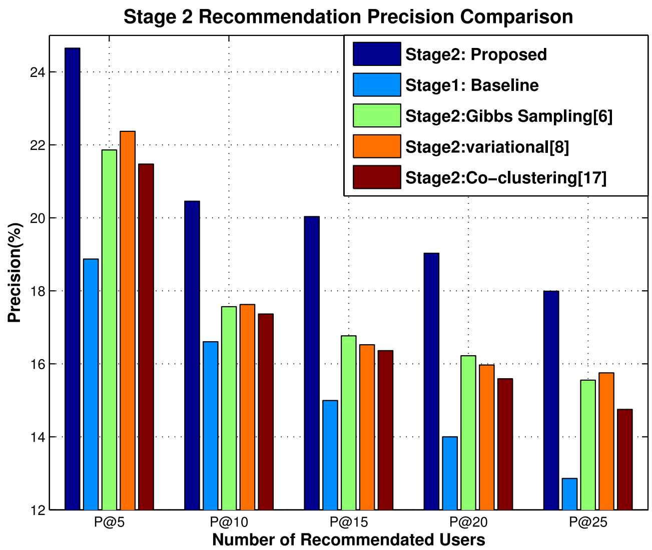


Fig. 3. Stage 2 recommendation precision comparison.

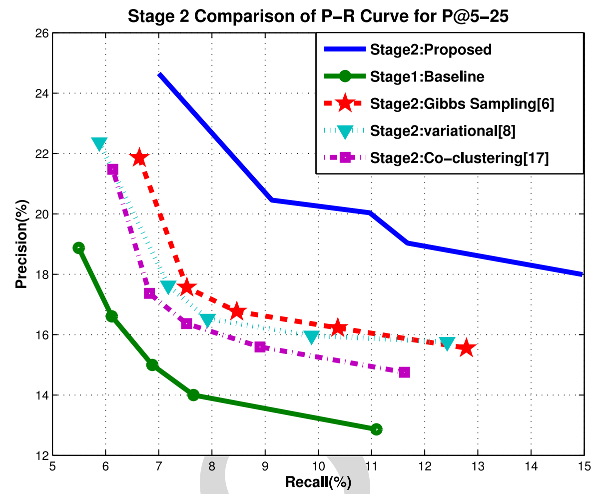


Fig. 4. Recommendation precision and recall for stage 2.

items behave similar on other items. Traditional collaborative filtering methods do not consider much about the social connections between users. SBPR removes this drawback by taking the social connections into account by assigning a social coefficient to each user.[1]

At last we consider a multi-network based algorithm for comparison. When considering social multiple network problems, transition probability propagation is a method that is frequently used [39], [40]. We choose [39] as a reference method for the following reasons: 1) It considers the relationships of different networks, which is similar to our idea; 2) It uses the information of other networks for recommendation, which again has some similarities with ours. [39] enhances the links in one network and between different networks using a random walk propagation method. After a sufficient number of walks, it obtains the modified link weights between each user pair. We use the weights for friend recommendation.

D. *Experimental Results*

Here we report the results of our method for friend recommendation as follows.

1) Performance of Series Expansion: In this experiment we compare the proposed series expansion method with the variational, Gibbs sampling, and co-clustering methods in the second stage. We treat the performance of the first stage as the baseline.

From Fig. 3 we can see that our method has the best performance for accurate recommendation. $P@X$ stands for that each time we recommend the top X friends to users. Generally, the second stage improves the recommendation precision from only applying the first stage, illustrating the effectiveness and necessity of applying the two staged methods. Our proposed method improves about 5–7% compared with the performance of the first stage, and also makes about 2–3% improvement compared with the Gibss sampling method and the variational method. The reason for the improvement mainly lies in that we apply an exact expression to approach the PDF of the data, rather than an approximation or sampling method. The co-clustering method lacks the ranking ability and thus the performance is not good.

Fig. 4 illustrates the precision-recall curve of the proposed and reference methods. Based on the result of the first stage, the series expansion method achieves the highest performance(The upper right line on the figure). We can see from Fig. 4 that when precision or recall is fixed, we can achieve a 3–4% improvement over the best reference methods. This means that the proposed method can achieve both the highest precision and recall. This experimental results shows that the series expansion method can best approximate the real distribution of the data, and thus makes the most precise recommendation.

On the other hand, the proposed method have also imposed Gaussian distribution assumption to the latent variables a, b, and v. This may also cause some negative effect although it can give an analytic expression. It is worthy to make a depth observation of the distribution of the latent variables in our future studies.

2) Performance of the Proposed Two-Stage Method: Now we compare our two-stage method with some recently-proposed recommendation systems as mentioned in V-C. The main results for precision and precision-recall curve are shown in Figs. 5 and 6.

From Figs. 5 and 6 we can see that our system achieves the best performance, compared with other state-of-the-art recommendation systems. In average, our system improves the recommendation accuracy by about 3-4%, compared with the second best one. MF based method [14] has the best performance among all the reference methods, for it decomposes the item-user and user-user matrix into different social factors in a proper way. The reason that the proposed method performs better than MF might lies in that the MF method does not consider the whole distribution of the network and is trapped into some local optimum. Collaborative filtering based method [38] has slight lower performance than [14], the reason might be that its assumptions about the users' positive and negative feedback are not very proper for the Flickr dataset. Finally, the random-walk based method [39] has the lowest performance, since the

[1]The realization of [14] and [38] is based on the existing open-source Java package LibRec at http://www.librec.net/.

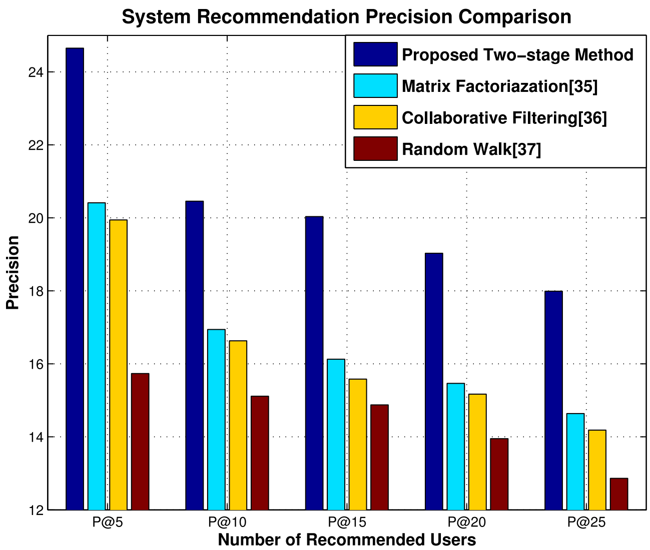


Fig. 5. Two-stage recommendation precision compared with state-of-the-art systems.

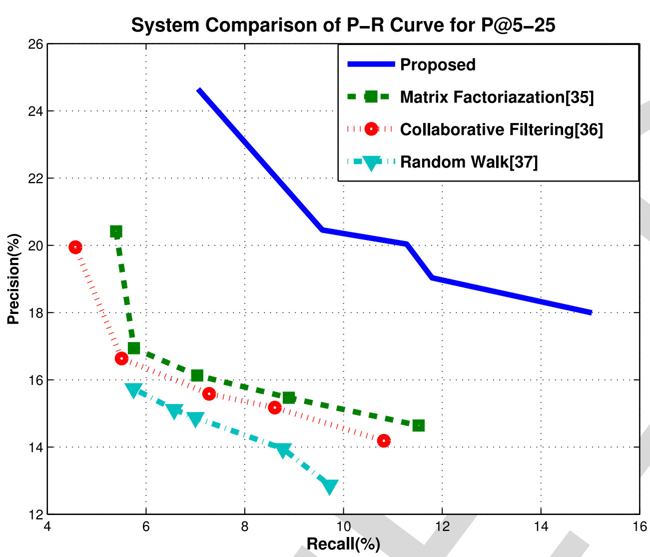


Fig. 6. Recommendation precision and recall compared with state-of-the-art systems.

random walk algorithm is not accurate enough for precise friend recommendation.

E. The Influence of Several Settings

1) The Influence of Additional Noise: The introduction of the additive noise, as shown in Section IV-B, makes the model more precise. However, it also leads to complicated inferences and calculations. In the following experiment we study the influence of the additive noise. In Table III, we compare the recommendation accuracy of the model that contains the additive noise and the model that does not contain the noise.

From Table III we see that by considering the additive noise we get a precision gain of about 1–2%, which is useful in the case where a more precise result is required.

TABLE III
INFLUENCE OF ADDITIVE NOISE

Precision(%)	P@5	P@10	P@15	P@20	P@25
Model With Noise	24.6	21.0	19.8	18.1	17.5
Model Without Noise	22.7	19.3	18.2	16.8	15.9

TABLE IV
INFLUENCE OF VALUES OF $\mathbf{C}$ AND $\mathbf{I}$

y	0.3	1	5	10
Precision(%)	19.6	24.6	13.7	11.0

2) The Influence of the Value of $\mathbf{C}_{\mathrm{kj}}$ *and* $\mathbf{I}_{\mathrm{ij}}$*:* As shortly discussed in Section IV, the convergence speed of the series is largely determined by the level of values of $\mathbf{C}$ and $\mathbf{I}$. If it is too large, then the convergence speed will decrease, leading to either the inaccuracy of the model, or larger number of terms. On the other hand, if the level is too small, the logarithmic terms in (12) will drop quickly and make the system unstable. In our experiments, contact network $\mathbf{C}$ stands for the intimacy of two individuals and in the image-user network $\mathbf{I}$, it stands for to what extent an individual favours an image feature. The values of each entry of $\mathbf{C}$ and $\mathbf{I}$ can be set according to our requirements. For example, we can set $\mathbf{C}_{\mathrm{jk}}$ to be 1 if two individuals are friends with each other and 0 otherwise; for image-user network we can also set $\mathbf{I}_{\mathrm{ij}} = 1$ if an individual has a certain image feature in his/photos, and 0 otherwise. On the other hand, we can also raise the level of the elements in $\mathbf{C}$ and $\mathbf{I}$ to be 5 or 10, or reduce it to be smaller than 1. The relationship between any two nodes would not change in the networks by varying the element value of $\mathbf{C}$ and $\mathbf{I}$, but the value does have an influence on the accuracy in our algorithm. We set the value of $\mathbf{C}$ and $\mathbf{I}$ on four levels to be 0.3, 1, 5 and 10 to check its influence on the performance.

In the following we compare the recommendation precision of these four levels.

From Table IV we see that the recommendation precision decreases rapidly as we increase the value of $\mathbf{C}$ and $\mathbf{I}$. On the other hand, if it is too small, the performance also goes down as the system becomes unstable around the poles of the logarithmic terms in (12). This indicates that we should choose the value of $\mathbf{I}$ and $\mathbf{C}$ around 1 for precise calculation.

VI. CONCLUSION AND FUTURE WORKS

In this paper, we develop a two-stage friend recommendation scenario utilizing multimedia information. In the first stage, tag information is utilised to build a tag-similarity network and is aligned to a contact network by a number of important features to generate a "possible friend list". In the second stage, a topic model is proposed and a new method based on series expansion is developed to combine image features and contact information to make more precise recommendations.

The experimental results show that the proposed method outperforms other methods in friend recommendation in that our method achieves the highest precision and recall in friend prediction. The network alignment of Stage One is effective. The topic model in Stage Two refines the result of stage one and the new series expansion method has better performance than the traditional variational and Gibbs sampling methods.

We will further develop our algorithm. For the series expansion method, it is a novel and effective method but not perfect. It is still to some extent mathematically complicated and has difficulties to apply on different models. We plan to refine the idea to make it more manoeuvrable and can be applied on general topic models. There are two directions to dig further. Firstly, for more complicated topic models, it might be viewed as a combination of some simpler models and thus are solvable based on our method. Secondly, our method is specially developed for Gaussian distributed random variables. For some other simple distributions, their algebra has been discussed in [23], [24], [27], etc. It is our future work to develop some general frameworks to combine all these distributions together.

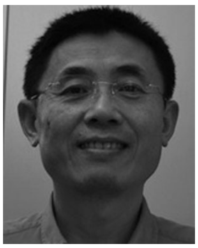
For our staged recommendation framework, we will extend our ideas to further applications such as product recommendation, media retrieval, etc. One problem of the current method is that in the first stage, some real friend might be omitted. We will further study how to increase the recalls in the first stage. We will develop other algorithms in each of our two stages, and to utilise the information form different domains. We will also make some studies about the ranks of the information from different domains. That is, which data should be applied in the first stage to achieve better performance. In the last, we can also introduce the concept of deep learning in our scenario for more efficient feature learning.

References

[1] W. Alex, "Why is it hard to make friends over 30?," *New York Times*, Jun. 2012.
[2] W. Zeng and L. Chen, "Heterogeneous data fusion via matrix factorization for augmenting item, group and friend recommendations," in *Proc. 28th Annu. ACM Symp. Appl. Comput.*, 2013, pp. 237–244.
[3] A. Chin, B. Xu, and H. Wang, "Who should I add as a "friend"?: A study of friend recommendations using proximity and homophily," in *Proc. 4th Int. Workshop Model. Soc. Media*, 2013, pp. 7:1–7:7.
[4] C. Guo, X. Tian, and T. Mei, "User specific friend recommendation in social media community," in *Proc. IEEE Int. Conf. Multimedia Expo*, Jul. 2014, pp. 1–6.
[5] S. Huang, J. Zhang, L. Wang, and X. S. Hua, "Social friend recommendation based on multiple network correlation," *IEEE Trans. Multimedia*, vol. 18, pp. 287–299, Feb. 2016.
[6] Z. Wang, J. Liao, Q. Cao, H. Qi, and Z. Wang, "Friendbook: A semantic-based friend recommendation system for social networks," *IEEE Trans. Mobile Comput.*, vol. 14, no. 3, pp. 538–551, Mar. 2015.
[7] W. Min, B.-K. Bao, C. Xu, and M. Hossain, "Cross-platform multi-modal topic modeling for personalized inter-platform recommendation," *IEEE Trans. Multimedia*, vol. 17, no. 10, pp. 1787–1801, Oct. 2015.
[8] N. Zheng, S. Song, and H. Bao, "A temporal-topic model for friend recommendations in chinese microblogging systems," *IEEE Trans. Syst., Man, Cybern.: Syst.*, vol. 45, no. 45, pp. 1245–1253, Sep. 2015.
Q1 [9] Y. Song, J. Zhang, and L. Cao, "A joint interest-social latent factor model for social recommendation," Tech. Rep., 2013.
[10] D. M. Blei, A. Y. Ng, and M. I. Jordan, "Latent dirichlet allocation," *J. Mach. Learn. Res.*, vol. 3, pp. 993–1022, Mar. 2003.
[11] T. L. Griffiths and M. Steyvers, "A probabilistic approach to semantic representation," in *Proc. 24th Ann. Conf. Cogn. Sci. Soc.*, 2002, pp. 381–386.
[12] M. Welling and Y. W. Teh, "Hybrid variational/Gibbs collapsed inference in topic models," in *Proc. Conf. Uncertain. Artif. Intell.*, 2008, pp. 587–594.
[13] H. Ma, H. Yang, M. R. Lyu, and I. King, "SoRec: Social recommendation using probabilistic matrix factorization," in *Proc. 17th ACM Conf. Inf. Knowl. Manage.*, 2008, pp. 931–940.
[14] G. Guo, J. Zhang, and N. Yorke-Smith, "TrustSVD: Collaborative filtering with both the explicit and implicit influence of user trust and of item ratings," in *Proc. 29th AAAI Conf. Artif. Intell.*, 2015, pp. 123–129.
[15] D. Lu, J. Sang, Z. Chen, M. Xu, and T. Mei, "Who are your real friend: Analyzing and distinguishing between offline and online friendships from social multimedia data," *IEEE Trans. Multimedia*,2017, to be published.
[16] M. Yan, J. Sang, T. Mei, and C. Xu, "Friend transfer: Cold-start friend recommendation with cross-platform transfer learning of social knowledge," in *Proc. IEEE Int. Conf. Multimedia Expo*, Jul. 2013, pp. 1–6.
[17] A. Chin, B. Xu, and H. Wang, "Who should i add as a "friend"?: A study of friend recommendations using proximity and homophily," in *Proc. 4th Int. Workshop Model. Soc. Media*, 2013, pp. 7:1–7:7.
[18] J. Sang and C. Xu, "Right buddy makes the difference: An early exploration of social relation analysis in multimedia applications," in *Proc. 20th ACM Int. Conf. Multimedia*, 2012, pp. 19–28.
[19] L.-H. Li, F.-M. Lee, Y.-C. Chen, and C.-Y. Cheng, "A multi-stage collaborative filtering approach for mobile recommendation," in *Proc. 3rd Int. Conf. Ubiquitous Inf. Manage. Commun.*, 2009, pp. 88–97.
[20] X. Zhao and J. Wang, "A theoretical analysis of two-stage recommendation for cold-start collaborative filtering," in *Proc. Int. Conf. Theory Inf. Retrieval*, 2015, pp. 71–80.
[21] S. Huang, J. Zhang, S. Lu, and X.-S. Hua, "Social friend recommendation based on network correlation and feature co-clustering," in *Proc. 5th ACM Int. Conf. Multimedia Retrieval*, 2015, pp. 315–322.
[22] S. Geman and D. Geman, "Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 6, no. 6, pp. 721–741, Nov. 1984.
[23] B. D. Carter and M. D. Springer, "The distribution of products, quotients and powers of independent H-fucntion variates," *SIAM J. Appl. Math.*, vol. 33, pp. 542–558, Dec. 1977.
[24] M. D. Springer and W. Thompson, "The distribution of products of beta, gamma and Gaussian random variables," *SIAM J Appl. Math.*, vol. 18, pp. 721–737, Jun. 1977.
[25] J. M. Pruett, "The distribution of products of some independent, non-standardized random variables," Ph.D. dissertation, Univ. Arkansas, Fayetteville, AR, USA, 1972.
[26] H. Michiel *et al.*, "Mellin transform," in *Encyclopedia of Mathematics*. Berlin, Germany: Springer, 2001.
[27] M. D. Springer, *The Algebra of Random Variables*. New York, NY, USA: Wiley, 1979.
[28] S. Ahmed, L.-L. Yang, and L. Hanzo, "Probability distributions of products of Rayleigh and Nakagami-m variables using Mellin transform," in *Proc. IEEE Int. Conf. Commun.*, Jun. 2011, pp. 1–5.
[29] R. Mallik and N. Sagias, "Distribution of inner product of complex gaussian random vectors and its applications," *IEEE Trans. Commun.*, vol. 59, no. 12, pp. 3353–3362, Dec. 2011.
[30] H. Bateman, *Tables of Integral Transforms*, vol. 1, New York, NY, USA: McGraw-Hill, 1954.
[31] L. Ahlfors, *Complex Analysis*. New York, NY, USA: McGraw Hill, 1979.
[32] E. L. Lady, "Convergence of infinite series in general and taylor series in particular," Oct. 1998. Q2
[33] W. Balser, *From Divergent Power Series to Analytic Functions*. Berlin, Germany: Springer, 1994.
[34] W. A. Richards and M. R. Acharya, "An efficient polynomial approximation to the normal distribution function and its inverse function," *J. Math. Res.*, vol. 2, 2010. Q3
[35] "Ferrari's solution of a quartic equation," *Maths notes*.
[36] Y. Jia et al., "Caffe: Convolutional architecture for fast feature embedding," *CoRR*, 2014. [Online]. Available: http://arxiv.org/abs/1408.5093
[37] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet classification with deep convolutional neural networks," in *Advances in Neural Information Processing Systems 25*, F. Pereira, C. J. C. Burges, L. Bottou, and K. Q. Weinberger, Eds., Red Hook, NY, USA: Curran Associates, Inc., 2012, pp. 1097–1105.
[38] T. Zhao, J. McAuley, and I. King, "Leveraging social connections to improve personalized ranking for collaborative filtering," in *Proc. 23rd ACM Int. Conf. Inf. Knowl. Manage.*, 2014, pp. 261–270.
[39] M. Jiang *et al.*, "Social recommendation across multiple relational domains," in *Proc. ACM Conf. Inf. Knowl. Manage.*, Oct. 2012, pp. 1422–1431.
[40] D. Liu, G. Ye, C.-T. Chen, S. Yan, and S.-F. Chang, "Hybrid social media network," in *Proc. ACM Multimedia Conf.*, Oct. 2012, pp. 659–668.

Shangrong Huang received the B.Sc. degree in telecommunication from Zhejiang University, Hangzhou, China, in 2005, and the M.Sc. degree in information and communication technology from the Darmstadt University of Technology, Darmstadt, Germany, in 2012, and is currently working toward the Ph.D. degree at the University of Technology Sydney, Sydney, NSW, Austraila.

His research interests include recommendation system, social medias, and big data analysis.



Jian Zhang (S'95–M'98–SM'04) received the B.Sc. degree from the East China Normal University, Shanghai, China, the M.Sc. degree in computer science from Flinders University, Adelaide, SA, Australia, and the Ph.D. degree in electrical engineering from the University of New South Wales (UNSW), Sydney, NSW, Australia, in 1999.

From 1997 to 2003, he was with the Visual Information Processing Laboratory, Motorola Labs, Sydney, NSW, Australia, as a Senior Research Engineer, and later became a Principal Research Engineer and a Foundation Manager in the Visual Communications Research Team. From 2004 to July 2011, he was a Principal Researcher and a Project Leader with National ICT Australia, Sydney, NSW, Australia, and a Conjoint Associate Professor with the School of Computer Science and Engineering, UNSW. He is currently an Associate Professor with the Global Big Data Technologies Centre and School of Computing and Communication, Faculty of Engineering and Information Technology, University of Technology Sydney, Sydney, NSW, Australia. He has authored or coauthored more than 100 paper publications and book chapters, and has six issued patents filed in the U.S. and China. His research interests include multimedia processing and communications, image and video processing, machine learning, pattern recognition, media and social media visual information retrieval and mining, human–computer interaction and intelligent video surveillance systems.

Dr. Zhang was the General Co-Chair of the International Conference on Multimedia and Expo in 2012 and Technical Program Co-Chair of the IEEE Visual Communications and Image Processing 2014. He was an Associate Editor of the IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS FOR VIDEO TECHNOLOGY (2006–2015) and the *EURASIP Journal on Image and Video Processing* (2016–present).

Dan Schonfeld (S'88–M'90–SM'05–F'10) received the B.S. degree in electrical engineering and computer science from the University of California at Berkeley, Berkeley, CA, USA, in 1986, and the M.S. and Ph.D. degrees in electrical and computer engineering from The Johns Hopkins University, in 1988 and 1990, respectively.

In 1990, he joined the University of Illinois at Chicago, Chicago, IL, USA, where he is currently a Professor with the Departments of Electrical and Computer Engineering, Computer Science, and Bioengineering. He has authored more than 200 technical papers in various journals and conferences. His research interests include signal processing, image and video analysis, video retrieval and communications, multimedia systems, computer vision, medical imaging, and genomic signal processing.

Dr. Schonfeld is Fellow of the SPIE for specific achievements in morphological image processing and video analysis. He has been elected as the University Scholar of the University of Illinois and was the recipient of the Graduate Mentoring Award of the University of Illinois at Chicago. He is currently serving as the Editor-in-Chief of the IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS FOR VIDEO TECHNOLOGY. He has served as the Deputy Editor-in-Chief of the IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS FOR VIDEO TECHNOLOGY and the Area Editor for Special Issues of the *IEEE Signal Processing Magazine*. He was the coauthor of a paper that was the recipient of the Best Paper Award at the ACM Multimedia Workshop on Advanced Video Streaming Techniques for Peer-to-Peer Networks and Social Networking 2010. He was also the coauthor of papers that were the recipient of the Best Student Paper Award at Visual Communication and Image Processing 2006 and the IEEE International Conference on Image Processing 2006 and 2007.

Lei Wang (SM'xx) received the Ph.D. degree from Nanyang Technological University, Singapore, in 2004.

He is currently with the School of Computing and Information Technology, University of Wollongong, Wollongong, NSW, Australia, as an Associate Professor. His research interests include machine learning, pattern recognition, and computer vision. For machine learning and pattern recognition, he is interested in feature selection, model selection, and kernel-based learning methods. For computer vision, he is interested in image categorisation and content-based image retrieval.

Xian-Sheng Hua received the B.S. and Ph.D. degrees in applied mathematics from Peking University, Beijing, China, in 1996 and 2001, respectively.

In April 2015, he was a Researcher and Senior Director of the Alibaba Group, leading the multimedia technology team in the search division. In 2013, he was a Senior Researcher of Microsoft Research Redmond, Redmond, WA, USA, where he worked on web-scale image and video understanding and search, and related applications. In 2011, he was a Principal Research and Development Lead of multimedia search for the Microsoft search engine, Bing, where he led a team that designed and delivered leading-edge media understanding, indexing, and searching features. In 2001, he joined Microsoft Research Asia, Beijing, China, as a Researcher. He has authored or coauthored more than 250 research papers in these areas and has filed more than 90 patents. His research interests include multimedia search, advertising, understanding, and mining, and recognition and machine learning.

Dr. Hua is an Associate Editor of the IEEE TRANSACTIONS ON MULTIMEDIA and an Associate Editor of the ACM TRANSACTIONS ON INTELLIGENT SYSTEMS AND TECHNOLOGY. He served as a Program Co-Chair for IEEE ICME 2013, ACM Multimedia 2012, and IEEE ICME 2012, as well as on the Technical Directions Board of the IEEE Signal Processing Society. He was the recipient of the prestigious 2008 MIT Technology Review TR35 Young Innovator Award for his outstanding contributions to video search. He was the recipient of the Best Paper and Best Demonstration Awards at ACM Multimedia 2007, the Best Poster Award at IEEE International Workshop on Multimedia Signal Processing 2008, the Best Student Paper Award at ACM Conference on Information and Knowledge Management 2009, the Best Paper Award at International Conference on MultiMedia Modeling 2010, the Best Demonstration Award at ICME 2014, and the Best Paper Award of the IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS FOR VIDEO TECHNOLOGY in 2014.

QUERIES

Q1. Author: Please provide the report number, as well as name and location of institution.
Q2. Author: Please provide complete bibliographic details in Refs. [32] and [35].
Q3. Author: Please provide page range in Ref. [34].
Q4. Author: Please provide the year in which the B.Sc. and M.Sc. degrees were earned.
Q5. Author: Please provide the year in which "Lei Wang" became the "Senior Member" of the IEEE.
Q6. Author: Please provide the location of Alibaba Group.

Two-Stage Friend Recommendation Based on Network Alignment and Series Expansion of Probabilistic Topic Model

Shangrong Huang, Jian Zhang, *Senior Member, IEEE*, Dan Schonfeld, *Fellow, IEEE*, Lei Wang, *Senior Member, IEEE*, and Xian-Sheng Hua

***Abstract*—Precise friend recommendation is an important problem in social media. Although most social websites provide some kinds of auto friend searching functions, their accuracies are not satisfactory. In this paper, we propose a more precise auto friend recommendation method with two stages. In the first stage, by utilizing the information of the relationship between texts and users, as well as the friendship information between users, we align different social networks and choose some "possible friends." In the second stage, with the relationship between image features and users, we build a topic model to further refine the recommendation results. Because some traditional methods, such as variational inference and Gibbs sampling, have their limitations in dealing with our problem, we develop a novel method to find out the solution of the topic model based on series expansion. We conduct experiments on the Flickr dataset to show that the proposed algorithm recommends friends more precisely and faster than traditional methods.**

***Index Terms*—Friend recommendation, series expansion, topic model.**

I. Introduction

FRIEND recommendation is a primary function in social network services and aims to recommend new social links for each user. Today when we lodge on the main social website such as Facebook, Twitter, and LinkedIn etc., we receive many recommendations of online friends. Seeing and hearing what the friends look at and listen to, or sharing our experience with our friends is an unparalleled experience. However, the decision of making friends is a complex human behaviour and can be affected by many different factors such as age, gender, location, interest [1], etc. As a consequence, similar to real life, finding a good on-line friend is not easy without the help of good recommendations. Traditional friend recommendations widely applied by Facebook and Twitter are often based on common friends and similar profiles such as having the same hobbies or studying in the same fields. These methods usually provide a long ranked possible friend list, but the recommendation precision is usually not satisfactory due to its complexity.

Nowadays people are commonly retained in a multi-resource environment, and usually do not seek friends based on only one kind of information anymore. Recently cross domain friend recommendation technologies have been extensively explored [2]–[4]. [2] applies a matrix factorisation method to combine the image and text information, [3] considers the proximity and homophily information for synthesised recommendation, and [4] specifies individuals' requirements from different domains. Most of these papers utilise information from different resources simultaneously for recommendation. In this paper, we approach this recommendation problem in a different way by utilizing the multi-domain information in different stages for a more precise recommendation.

The reason why we apply the multi-stage friend recommendation scenario lies in the complexity of multi-source information and the decision making behaviour of people. For example, an individual might make an on-line friend because they discuss a hard mathematical problem, or it is possible that he/she makes a friend because they both enjoy a film. The reason for friend making might be very diverse. It would be relatively difficult if we consider different factors together at the same time for recommendation. In our opinion, it is more convenient and clearer to analyse these factors step by step, rather than to deal with such cross-domain information as a whole. By untwisting the different factors in the recommendation procedure and analysing each factor in depth, a more precise recommendation performance is expected. As a consequence, we apply a two-stage framework to synthesise heterogeneous information from different domains.

In this paper, we concentrate on the widely-used image and image-related experience sharing website Flickr, where individuals can upload photos and tags for sharing as well as make online friends (Flickr Contact) and join communities (Flikr Group). Tag (text) information is quite useful for friend recommendation since it is simple and direct. For example, two individuals that both have interest in tags "travel" and "historical people" have higher probability to be friends with

Manuscript received April 11, 2016; revised August 31, 2016 and November 21, 2016; accepted December 28, 2016. The associate editor coordinating the review of this manuscript and approving it for publication was Dr. Zhen Wen.

S. Huang and J. Zhang are with the School of Computing and Communications, University of Technology Sydney, Ultimo, NSW 2007, Australia (e-mail: shangrong.huang@student.uts.edu.au; Jian.Zhang@uts.edu.au).

D. Schonfeld is with the Multimedia Communications Laboratory, University of Illinois at Chicago, Chicago, IL 60607, USA (e-mail: dans@uic.edu).

L. Wang is with the School of Computing and Information Technology, University of Wollongong, Wollongong, NSW 2522, Australia (e-mail: leiw@uow.edu.au).

X.-S. Hua is with the Alibaba Group, Hangzhou 311121, China (e-mail: xianshenghua@gmail.com).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TMM.2017.2652074

each other. Text-based friend recommendation have been successfully developed in [5], [6]. We apply the text information in our first stage.

Image information is also helpful for friend recommendation. On the other hand, image information is vaguer and more complex for this task. For example, it is hard to claim that two individuals who both enjoy some vivid and colourful photos, or some photos of beautiful women have the higher probability to be friends. As a consequence, in our algorithm, we utilise the image information as a supplementary source in the second stage of the algorithm, to refine the result of the first stage.

In the first stage, similar to [5], based on the correlation of different networks, we align the tag-similarity network to friend network to obtain a possible friend list. Specifically, we consider each user as one node in a graph, and we crawl the uploaded tags from each user and calculate the tag similarity between any two users as the edges to form a tag-similarity network. On the other hand, we also obtain the friendship information in Flickr, and if two users have friendship with each other, we add an edge between the two to form a contact network. In this way we build a tag-similarity network and a contact network that have the same nodes but different topologies. Because the tag-similarity network and contact network on Flickr are related to each other, we dig their correlation by choosing important tag features, to make the tag-similarity network more similar to the contact network. In this way, the chosen tag features provide a guideline for friend recommendation. This stage makes a mass election of possible friends.

In the second stage, to overcome the problem that the mass election considering only the tag information might not be precise, we build a topic model to illustrate the relationship between user's friend making behaviour and the image features they have uploaded. This stage refines the list obtained in the first stage. The main reason for applying a topic model in our second stage lies in the fact that the topic model has the ability to tell on what probability a user would prefer a photo/item/friends.

The probabilistic topic model discovers the abstract "topics" that occur in a collection of documents/datasets, and it has been widely used in recommendation systems [7]–[9]. By introducing some latent variables and applying the Bayesian rule, it is conceptually easy to combine information from different domains and make specific recommendations [7], [9]. Generally it assumes that people's various behaviours such as shopping, posting and friend making are controlled by some latent topics. Certain people have particular bias on different latent topics. For an individual that acts differently in different domains, his/her latent interest topic might be similar. For example, a user who posts many different photos about food on Flickr might have higher probability to be interested in the topic of cooking, and thus it is reasonable to recommend some kitchenware to him/her on Amazon. Furthermore, the topic model provides a relatively precise probability to show to what extent an individual is interested in a topic, and thus makes it easy for further recommendation.

In this paper, we propose a topic model to correlate the data about the Flickr image information and the contact information. Compared with some previously cross-domain topic models, our model is more compact with less parameters, which leads to some computational convenience. Briefly, we assume that the attractiveness of photos is controlled by a latent variable, and individuals' photo uploading behaviours and their friend making behaviours are controlled by some other latent variables. By determining the values of these latent variables we can predict individuals' friends.

However, it is often not easy to find the solution of a topic model when different domains are concerned, for it involves the integrals of several coupled random variables, which is a complicated mathematical problem in general [10]. Two methods are widely used to deal with this problem: Gibbs sampling [11] and variational inference [10], or the combination of the two [12]. Although applied successfully in many cases, both of them have some disadvantages: for Gibbs sampling, it is inefficient for large count values since it requires averaging over many samples to reduce variance; for variational inference, though it is efficient to deal with large scale data, the variational step makes it hard to control the precision when approximating the integrals when making the Bayesian inference. In this paper, with the help of Mellin and inverse Mellin transform, we propose a novel way based on series expansion to calculate the coupled integrals that are required in the Bayesian inference.

Matrix factorization (MF) method can be also applied to deal with the cross domain recommendation problems [13], [14]. It decomposes different social networks into latent vectors to find the important factors that influence individuals' social behaviours, and make recommendations based on these latent factors. However, it lacks a mechanism to draw the complete distributions of the whole social network, and thus might lead to some local optimum. Our proposed method provides a way to describe the whole distribution of the social network, to perform a better recommendation.

To sum up, we build a two-stage friend recommendation system based on text and image data: in the first stage, we apply tag-user information to get a possible friend list, and in the second stage we refine the list by utilizing the image-user information. Our main contributions are as follows: Firstly, we build a compact topic model to analyse the relationship of the data from different domains. Secondly, we propose a novel method based on the study of the distribution of algebra of random variables to find a solution of the model. The solution is given in a series expansion form, and can lead to more precise solutions of the model. As far as we know,this is the first time to solve a topic model from the aspect of integral series expansion. We also make comprehensive experiments to show the effectiveness of our method.

The rest of the paper is organised as follows: Section II outlines related work. Section III introduces our system framework. Section IV gives the detailed explanation of our series expansion method. Section V evaluates the performance of our method and some analysis is made according to the results. Lastly, Section VI concludes our work.

II. Related Work

Our work in this paper is mainly related to the following research fields: friend recommendation, topic model, and algebra of random variables.

A. *Friend Recommendation*

Friend recommendation is a relative challenging issue compared with item or group recommendations, for there might be various reasons for two persons to become friends, and online and offline friendships are quite different. Recently, [15] even provides some method to distinguish the online and offline friends.

[6] makes a survey of individuals' daily life, and then summarises the reports as their "life styles" using latent Dirichlet allocation algorithm (LDA). [8] collects individuals' posts in Micro-blogs and arranges them in a chronological order. By building a temporal-topic model it can recommend different friends to each user at different time, as the user's interest changes from time to time. [16] utilises the information from different platforms (Flickr, Twitter, Google+, etc) to alleviate the sparsity problem of social networks, the idea is that Google+ can provide a information bridge between these different social platforms. In this paper, We dig the friend recommendation problem deep by considering multimedia information one platform, and applying a two-step scenario to refine the result.

1) Cross-Domain Recommendation: As mentioned in Section I, individuals' decision of making friends are often multi-dimensional. As a result, recently many researchers consider friend recommendation based on cross-domain information. [17] considers the friend recommendation problem at working places and conferences, by utilising both users' temporal location as well as their common friend information. [2] combines three aspects of each user's information: the items one likes, the friends one has, and the groups one belongs to. Such information of different aspects is synthesised and integrated into one cost function. By optimizing the cost function, the heterogeneous data are fused for item, group and friend recommendations. In [16], individuals that have both accounts in Flickr, Twitter and Google+ are collected to build the relationship of the two social websites. The common behaviours of each user in Flickr and Twitter are analysed and the friend recommendation of the two domain is made based on these common behaviours.

[4] divides the different data in Flickr into two classes: interaction data(comments, making favorite photos) and similarity data(common friends, groups, tags, geo, visual), and applies these two classes of data comprehensively to estimate the strength of the ties between users. [18] utilises Flickr social relations for further multimedia recommendation. It builds a topic model to combine the image, text, and friendship information to discovery individuals' preferences. The topic model is solved via Gibbs sampling.

For the works listed above, the data from different domains are processed simultaneously or fused together to get the final recommendation result. On the one hand, the above methods take the advantage that data from different domains might be related to each other; On the other hand, these methods combine the cross-domain information in one step ([16]) or synthesise it in one cost function ([2]), thus usually can not give a good explanation of how the data from a specific domain contribute to the final recommendation result, and the twisted data from different domains often makes the problem more complex. To have a better understanding of the effectiveness of the data from each domain, in this paper, we design a two-step recommendation that in each step we utilise the data from one domain.

2) Multistage Recommendation: Existing multi-stage recommendations are usually applied to find some patterns of users or items. For example, in [19], a two-stage mobile recommendation is proposed to help users find the correct events. The first stage clusters people according to their profile similarity and the second stage discovers the event-participating pattern. [20] designs the first stage to find some related resources that one user requires, and the second stage is used to find some patterns that the user might prefer from the previous stage for further recommendation. Both [19] and [20] can handle the cold-start problem well but do not consider much about the cross-domain problem.

In this paper we apply a different strategy: in the first stage, some relatively good results are chosen by observing the text data; then we refine the results in our second stage, with the help of image data. In our previous paper [21], we provide a two-stage recommendation and each stage utilises data from different domains by alignment and co-clustering. However, co-clustering method lacks the ability to tell the intimacy distance between two individuals exactly but only to group people roughly with similar properties, and thus can not make precise recommendation. To overcome this problem, in this paper we propose a probabilistic topic model in the second stage for a better recommendation. We also provide a novel and more precise method to solve the topic model problem.

B. *Probabilistic Topic Model*

In the second stage of our model, a topic model is applied to get a more precise recommendation.

1) Topic Model in Recommendation: The probabilistic topic model is a successful approach solving the problem for information retrieval[10] and recommendation[7]–[9]. For example, [8] recommends temporary friends to users by building models that contain latent variables that illustrate users' interests change with time.

By assuming some latent factors it is conceptually easy to build the relationships among different domains. [7] designs a model that connects the Flickr and Foursquare data for image, topic and item recommendation. It assumes that both domains have some common latent factors and each domain also has its own latent factors, and the users' activities on these two platforms are the synergism of all these factors. Gibbs sampling is applied to find the value of the latent factors. [9] considers the friendships and the votings on the large Film rating website. To predict individual's flavour about films his/her social relationships and scores of films are combined with some latent factors. Variational methods are applied to solve the model.

To make the model to illustrate the situation of the real world more accurately and reasonably, both [7] and [9] make many assumptions of the latent topic and thus contain many unknown parameters to infer: [9] contains more than 10 unknown parameters and [7] has more than 30. The presence of so many unknown variables not only greatly increases the complexity of the algorithm, but also leads to other problems such as over-fitting or redundancies. In this paper we build the model in a more compact manner.

2) Gibbs Sampling, Variational Inference and Matrix Factorization: Due to the coupling of latent variables, the direct inference is usually impossible for a specific topic model. Generally there are two methods to find a solution for topic model: Gibbs sampling [22] and variational inference [10]. For some complex multivariate probability distributions, to determine the parameters of the distribution, direct sampling is difficult. Gibbs sampling samples the marginal distribution of one variable each time, and iteratively samples all the marginal distributions. The variational method, on the other hand, approaches the solution by approximating the original complex distribution with a factorised one, which is easier to handle.

As stated in Section I, both of the two methods have some weaknesses: Gibbs sampling has difficulties in handling big data problems, and the variational method can not determine if the approximation is close to the original one. Some researchers consider combining the two in one problem: In [12], small counts of data are sampled and the variational method is applied to update large counts, which improves the performance on the large dataset. However, how accurate the approximation of variational method is not yet discussed in [12]. In this paper, we propose a new solution to a topic model by directly calculating the distribution of the latent variables.

Compared with the above two methods, MF-based method also assumes some latent variables but instead of determining the marginal distribution of the observed data, it factorizes the observed data into different latent factors, which leads to some computational convenience and efficiency. Both of [13] and [14] utilize user friendship network and user-item network and obtain some latent factors that show the preference of individuals. The recommendations based on these latent factors are relatively effective. On the other hand, they do not try to find the probabilistic distribution of the network and all of these methods apply some gradient descent methods, that are relatively easy to be trapped into a local optimum. Our method avoids this drawback by deducing the distribution of the whole probabilistic model.

C. Algebra of Random Variables

The essential problem of our approach in this paper is to get the exact mathematical expression of the coupling of different random variables, mainly the sum and product of random variables. These problems were extensively discussed in the 1950s to 1970s year, last century, during which time the random process was a hot research topic but the computer simulation technology was not well developed. In [23]–[25], the products of typical distributions such as Beta, Gamma and Rayleigh are discussed. Most of these works utilise the Mellin transform [26] as the essential tool for deducing. [27] gives a good summary of these works and also discusses the distribution of the sum of random variables. The algebra of random variables has also been studied recently in certain fields such as wireless communication in [28] and [29]. These works show that the product and quotient of random variables with certain distributions can be expressed analytically. We will mainly apply some of the results in [25], [27] later in our work. As Gaussian distribution has some good properties(its domain of definition is all the real values, and has a central point, etc.) we assume that our latent variables to be Gaussian distributed.

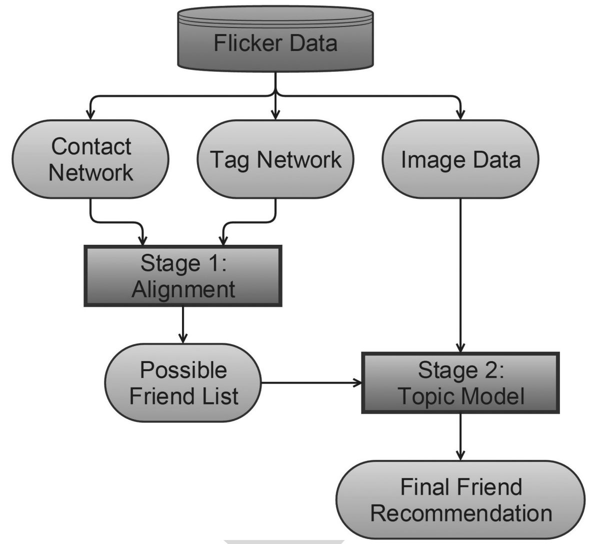


Fig. 1. Two-stage system illustration.

III. System Model

The main framework of our model is shown in Fig. 1, which contains two stages: In the first stage, network alignment is applied to generate a possible friend list, by correlating the tag and contact data in Flickr; In the second stage, the user-uploaded image features generate some topics by utilising a probabilistic topic model, and a new method is developed to solve the model for precise friend recommendation.

A. First Stage: Network Alignment

The detailed alignment method has been discussed in [5]. The following is an introduction of its basic idea. An individual may join different social networks for different purposes. For example, one may at the same time join a football fan network for physical practice and a restaurant information sharing network to look for the best food. He/she plays different social roles in these different networks, and might make different friends. However, these different social roles for one individual are not independent, but related to each other.(The man might look for some food that helps quickly recuperate after hard physical practice). The motivation for social network alignment lies on the fact that these different networks, though having different edges (relationships), are usually related to each other. Taking Flickr as an example, according to the uploaded-tag-similarity of each user and their contact list, a tag similarity network and a contact network are formed. Although the topologies of the two networks are not the same, they are related to each other, for users uploading similar tags on Flickr have higher probability to make friends with each other. By digging the correlation of the topologies of different networks we may make inference for the knowledge from one domain to another.

Specifically, we align the Flickr tag-similarity network with the contact network, so that after the alignment, one tight edge between two users in the tag similarity network would imply that these two users have higher probability to have contact with each other. We align the tag-similarity network with the contact network by selecting important tag features. The reason we apply feature selection here lies in the phenomenon that when we look for online friends, it is common that we do not take care of all the factors of a person but concentrate on certain points that would interest ourselves. As an example, a traveller might post his photos with the following tags: "Sydney", "Blue Mountain", "great view", and "street". Among these tags some people might contact him/her for some more details about the experiences in ''Sydney" and ''Blue Mountain", but seldom would have interests about "great view" or "street" because they are too common. We can treat these two tags as redundancy for friend making. Based on this observation, we believe that some Flickr tags can be more indicative in the task of friend recommendation, because they are more important to reflect the connections on the contact network. We can treat these tags as important features for friend recommendation. Inspired by this phenomenon, we design a method to choose some important features that are more helpful for friend making decision.

Mathematically, assume that the feature selection matrix to be $\mathbf{W}$, the known tag-user matrix to be $\mathbf{X}$, the tag distance matrix to be $\mathbf{L}$, and the first d eigenvector-matrix of the contact network to be $\mathbf{V}$, the important feature can be obtained by solving the following problem:

$$\min_{\mathbf{W}} \|\mathbf{XW} - \mathbf{V}\| + \mu tr(\mathbf{W}^T \mathbf{X}^T \mathbf{LXW}) + \lambda \|\mathbf{W}\|_{2,1}. \quad (1)$$

The first term of (1) aligns the tag-similarity network to the contact network so that they become more similar to each other, and the second term preserves the local structure of the original tag-similarity network. The third term is for regularization. In this way the tag feature selection matrix $\mathbf{W}$ makes the topology of the tag-similarity network more similar to the contact network, while preserving the topology of the tag-similarity network as much as possible. In other words, we align the tag-similarity network to the contact network. By comparing the similarity of two users on the those important tags we can generate a possible friend list for each user. The solution of $\mathbf{W}$ in (1) is discussed more thoroughly in [5].

B. Second Stage: Topic Model

In the previous stage we get a possible friend list by considering the correlation between the tag and contact networks on Flickr. However, as the real world friend relationship is affected by many factors[1], one stage is usually not enough for a precise friend recommendation. In the following stage, we introduce the image data as auxiliary information to refine the recommendation list.

We apply the topic model to combine the image data and the friendships in Flickr. It is common sense that a person uploads a photo on Flickr because he/she likes the photo. Why does

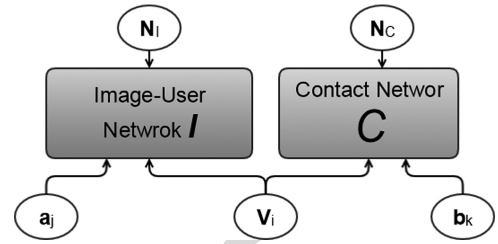


Fig. 2. Probabilistic topic model combining image-user network and contact network.

he/she like the photo? We assume that in one's mind, some latent interest factors control his/her taste of image. For example, some people like colourful, vivid photos, while others prefer spectacular or imposing ones; children enjoy comic-style pictures while adults have more interests in realistic-style paintings; young women pay much attention to photos of beautiful clothes while young men to electrical devices. These latent factors are determined by various aspects such as age, gender, living experiences, etc. and can not be observed or simply summarised. We assume individuals' interest latent factor to be $\mathbf{v}$. Each image contains the factors that attract people, such as colour, line, or history, which we assume to be $\mathbf{a}$. The correlation of $\mathbf{v}$ and $\mathbf{a}$ determines whether a user would upload an image.

Similarly, we assume that each user exhibits some attractive factors during his/her activities in Flickr such as uploading photos, writing descriptions of photos and making comments, etc. We also summarise these attractive factors with the third latent variable $\mathbf{b}$. Notice that the same user's interest latent factor $\mathbf{v}$ and attractive factor $\mathbf{b}$ are not the same. The combination of $\mathbf{b}$ and $\mathbf{v}$ determines whether two users should make friends with each other. For simplicity we view them as independent from each other. The topic model is shown in Fig. 2.

In Fig. 2, $\mathbf{C}$ and $\mathbf{I}$ stand for the $0-1$ contact network and image-user network, respectively. $\mathbf{C}$ is an $n \times n$ matrix where n is the number of users. $\mathbf{I}$ is an $n \times f$ matrix where f stands for the number of total features. For $\mathbf{C}$, if user k and user j are friends with other, then $\mathbf{C}_{\mathrm{kj}}$ equals one, and zero otherwise. For $\mathbf{I}$, if the uploaded photos of user i contain an image feature j, then $\mathbf{I}_{\mathrm{ij}}$ equals one, and zero otherwise. a stands for image factor, and b stands for individuals' social interest factor, respectively. v stands for individuals' common interest factor that has effect on both his choice of images and friends. $\mathbf{N}_{\mathbf{I}}$ and $\mathbf{N}_{\mathbf{C}}$ stand for zero-mean additive noises. The relationship can be mathematically expressed as follows:

$$\mathbf{I}_{\mathrm{ij}} = a_i \times v_j + \mathbf{N}_{\mathrm{Iij}}, \mathbf{C}_{\mathrm{kj}} = b_k \times v_j + \mathbf{N}_{\mathrm{Ckj}}. \quad (2)$$

We assume that all the latent random variables a_i, b_k and v_j are Gaussian distributed with the parameters of means and variances of $\mu_a, \sigma_a, \mu_b, \sigma_b, \mu_v$, and σ_v, respectively. The reason we choose Gaussian distribution is as follows: Although some other distributions that are in the form of an H-function (such as Beta, Gamma or Rayleigh distributions) would lead to some calculation convenience [27], we assume Gaussian distribution here because it is defined on the whole real domain and contains negative values and has a central point, while other distributions such as Beta are only defined on the positive real domain.

The coupling between random variables a, b and v makes the integral of (2) often intractable. Traditional methods dealing with (2) contain Gibbs sampling [11] and variational inference [10]. Gibbs sampling meets with difficulties when the data scale is large, and the variational method applies some approximation that the precision is hard to control. In the following we develop a new approach to solve (2) that is based on Mellin transform and series expansion.

IV. Series Expansion

A. Product of Gaussian Random Variables

When dealing with the distribution of product of random variables, the Mellin transform is an essential tool [27]. We take the first equation in (2) to explain its basic idea. For simplicity we first neglect the noise term $\mathbf{N}_{\mathrm{ij}}$ (its effectiveness is to be discussed later) and we have $\mathbf{I}_{\mathrm{ij}} = a_i v_j$ for two random variables a_i and v_j with different probability distribution functions. One useful property for Mellin transofrm is: the Mellin transform of the product of two probability density functions (PDF) is equal to the product of the Mellin transforms of their PDFs.

Mathematically, we recall the following rule [27]: If a_i and v_j are two non-negative random variables with the PDFs $f_a(a_i)$ and $f_v(v_j)$, their product $\mathbf{I}_{\mathrm{ij}} = a_i v_j$ has a distribution $h(I_{\mathrm{ij}})$, and then the Mellin transform of $h(I_{\mathrm{ij}})$ is precisely the product of Mellin transform of $f_a(a_i)$ and $f_v(v_j)$, respectively. The expression is given as

$$\mathcal{M}(h(\mathbf{I}_{\mathrm{ij}})) = \mathcal{M}(f_a(a_i))\mathcal{M}(f_v(v_j)) \tag{3}$$

where the Mellin transform and its inverse of an analytical function $f(x)$ are defined as follows:

$$\mathcal{M}(s) = \int_0^{+\infty} x^{s-1} f(x) dx \tag{4}$$

$$\mathcal{M}^{-1}(x) = \frac{1}{2\pi i} \int_{c-i\infty}^{c+i\infty} x^{-s} \mathcal{M}(s) ds \tag{5}$$

where c in (5) stands for an arbitrary real number. With the help of (3)–(5) and the known distribution of a_i and v_j, we can give an exact mathematical expression for distribution of the coupling of the two random variables a_i and v_i.

In this way we can first deduce the Mellin transform of each of the distributions, then make product of the two, and finally inverse the Mellin transform to get the final product distribution. In this way, we first calculate the distribution of $\mathbf{I}$ in (2).

From the previous assumption we know that a_i, b_k and v_j follow the Gaussian distribution with mean μ_{ai}, μ_{bk}, μ_{vj} and the variance $\sigma_{\mathrm{ai}}, \sigma_{\mathrm{bk}}, \sigma_{\mathrm{vj}}$. We further take the symbol of f_{ai}, f_{bk} and f_{vj} as their PDFs. We first do the Mellin transform on a_i and v_j separately to get $\mathcal{M}(f_a(a_i))$ and $\mathcal{M}(f_v(v_j))$, and then we product them and do the inverse Mellin transform to finally get the distribution of product of two random variables, which is the distribution of the variables in image-user matrix $\mathbf{I}$. The details are given in [27] and [25], which provide two equivalent expressions for the distribution of two Gaussian random variables. We apply the expression from [25] and the details are briefly outlined in the following.

To calculate the distribution of $\mathbf{I}_{\mathrm{ij}} = a_i v_j$ with Gaussian random variables a_i and v_j, we take the Mellin tranform of $f_a(a_i)$ and $f_v(v_j)$. Notice that according to (4), the positive and negative parts of the distribution of a_i and v_j should be considered separately. We apply the property that the Mellin transform of the standard Gaussian distribution is Gamma function[30]: $\mathcal{M}\{\mathrm{e}^{-x^2/2}\} = 2^{s/2-1}\Gamma(s/2)$, and a non-central Gaussian distribution can be expressed as a standard Gaussian distribution multiplied by a series in the form: $\mathrm{e}^{-\frac{1}{2}(x-\mu)^2} = \mathrm{e}^{-\mu^2/2} \sum_{j=0}^{\infty} \frac{1}{j!} \mu^j x^j \mathrm{e}^{-x^2/2}$. If we define the following:

$$\begin{aligned} a_{i1} &= max(a_i, 0), v_{i1} = max(v_j, 0) \\ a_{i2} &= min(a_i, 0), v_{i2} = min(v_j, 0) \\ \mathbf{I}_{ij-1} &= a_{i1} v_{i1}, \mathbf{I}_{ij-2} = a_{i1} v_{v2} \\ \mathbf{I}_{ij-3} &= a_{i2} v_{v1}, \mathbf{I}_{ij-4} = a_{i2} v_{j2}. \end{aligned}$$

And we also define the probability distribution function of $\mathbf{I}_{ij-1}$, $\mathbf{I}_{ij-2}$, $\mathbf{I}_{ij-3}$, and $\mathbf{I}_{ij-4}$ to be $h_1(\mathbf{I}_{\mathrm{ij}})$, $h_2(\mathbf{I}_{\mathrm{ij}})$, $h_3(\mathbf{I}_{\mathrm{ij}})$ and $h_4(\mathbf{I}_{\mathrm{ij}})$, respectively. Following the methods of [25], and taking $\mathbf{I}_{ij-1}$ as an example, we have

$$\mathcal{M}_{\mathbf{I}_{ij-1}}(s) = \sum_{o=0}^{\infty} \frac{\mu_{a_i}^{2o}}{(2o)!} \frac{\mu_{v_j}^{2o}}{(2o)!} \Gamma^2(s). \tag{6}$$

To get the distribution of $\mathbf{I}_{ij-1}$, we do the inverse Mellin transform of (6) as:

$$h_1(\mathbf{I}_{\mathrm{ij}}) = \sum_{o=0}^{\infty} \left(\frac{1}{2\pi i}\right) \int_{c-i\infty}^{c+i\infty} (y^2)^{-s} \frac{\mu_{a_i}^{2o}}{2o!} \frac{\mu_{v_j}^{2o}}{2o!} \Gamma^2(s+o) ds. \tag{7}$$

Equation (7) is an integral on half of the complex plane. According to Residue Theorem [31], the solution is expressed with the infinite residues that are related to the poles on the real plane. By calculating the residues we get (8), shown at the bottom of the page, where $C1 = \frac{1}{\pi} \mathrm{e}^{-\frac{1}{2}(\frac{\mu_{\mathrm{ai}}^2}{\sigma_{\mathrm{ai}}} + \frac{\mu_{\mathrm{vj}}^2}{\sigma_{\mathrm{vj}}})}$, $C2 = ((\frac{1}{(2o)!})^2 (2\frac{\mu_{\mathrm{ai}}^2}{\sigma_{\mathrm{ai}}} \frac{\mu_{\mathrm{vj}}^2}{\sigma_{\mathrm{vj}}})^o)$, and $\psi(1)$ is the Euler-Mascheroni constant.

Similarly we should also consider the case of $h_2(\mathbf{I}_{\mathrm{ij}})$ for $a > 0 \cap v < 0$, $h_3(\mathbf{I}_{\mathrm{ij}})$ for $a < 0 \cap v > 0$, and $h_4(\mathbf{I}_{\mathrm{ij}})$ for $a <$

$$h_1(\mathbf{I}_{\mathrm{ij}}) = C1 \left[\sum_{o=0}^{\infty} C2 \sum_{s=o}^{\infty} \left[\frac{(\mathbf{I}_{\mathrm{ij}})^{2s}}{\prod_{t=0}^{s-o-1} (-s+o+t)^2} \left(2\psi(1) - 2 \sum_{w=0}^{s-o-1} \frac{1}{-s+o+w} \right) - \frac{(\mathbf{I}_{\mathrm{ij}})^{2s} \ln((\mathbf{I}_{\mathrm{ij}})^2)}{\prod_{w=0}^{s-o-1} (-s+o+w)^2} \right] \right] \tag{8}$$

$0 \cap v < 0$. To sum up, we have

$$\begin{aligned} h(\mathbf{I}_{\rm ij}) &= h_1(\mathbf{I}_{\rm ij}) + h_2(\mathbf{I}_{\rm ij}) \qquad (y > 0) \\ &= h_3(\mathbf{I}_{\rm ij}) + h_4(\mathbf{I}_{\rm ij}) \qquad (y < 0). \end{aligned} \tag{9}$$

In a similar manner we can give the expression for $h(\mathbf{C}_{\rm kj})$

Here we give a short discussion about this series. In the first place, this is basically an alternating and power series [32] with infinite terms, with some of the terms multiplied with a logarithm factor. This is a series that when the sequence number of the term increases, the absolute value of the term increases. Some of the terms are positive and some are negative, and the sum of the terms eventually becomes convergent, as discussed in [33]. However, similar to some of the convergent Taylor series, when the absolute value of the series terms is large, these series converge only when the term number of the series is also large. In order to make the series to converge rapidly with relatively a small number of terms, in practice, we may normalise the value of $\mathbf{I}_{\rm ij}$ to be relatively small (In the experiments, the ground truth of $\mathbf{I}_{\rm ij}$ and $\mathbf{C}_{\rm ij}$ are 0 or 1, which is small enough).

B. *Additive Noise*

From Fig. 2 we see that after the products of a, v and b, v, the results should also add a bias value or noise to get the value of $\mathbf{I}_{\rm ij}$ and $\mathbf{C}_{\rm kj}$. In practice it can be interpreted as all the outer environmental influences other than the users and the items. For example, the change of seasons for the favour of clothing, or the change of temperature for the preference of food, etc. Mathematically the PDF of two independent random variables are the convolution of their PDFs of the two [27]. In our case, we can simply consider the environmental influences $\mathbf{N}_I$ and $\mathbf{N}_C$ to be independent from the image factor a, social attractive factor b and individual's latent factor v. For simplicity we assume the additive noise of $\mathbf{N}_I$ and $\mathbf{N}_C$ to be Gaussian distributed with zero mean and variance of $\sigma_{\rm Ni}$ and $\sigma_{\rm Nc}$, respectively. Taking $\mathbf{I}_{\rm ij}$ for example, from (8) we see that the most important calculation is the convolution of the Gaussian function from additive noise $\mathrm{e}^{-\mathbf{I}_{\rm ij}^2/\sigma_{\rm Ni}^2}$ and the term ${\mathbf{I}_{\rm ij}}^{2s}\log({\mathbf{I}_{\rm ij}}^2)$ from (8), which is formally written as follows:

$$d_2(\mathbf{I}_{\rm ij}) = \mathrm{e}^{-\mathbf{I}_{\rm ij}^2/\sigma_{\rm Ni}} * \mathbf{I}_{\rm ij}^{2s}\ln(\mathbf{I}_{\rm ij}^2). \tag{10}$$

By calculating the convolution we see (10) can be expressed as follows:

$$d_2(\mathbf{I}_{\rm ij}) = {\mathbf{I}_{\rm ij}}^{2s+2}\left(\frac{\ln {\mathbf{I}_{\rm ij}}^2}{2s+2} - \frac{1}{(2s+2)^2}\right)\mathrm{e}^{\frac{(-{\mathbf{I}_{\rm ij}}^2)}{\sigma_{\rm Ni}}}. \tag{11}$$

In this way we can get a series expression of (9).

$$\begin{aligned}
h(\mathbf{I}_{\rm ij}) = \frac{1}{\pi}e^{-\frac{1}{2}\left(\frac{\mu_{\rm ai}^2}{\sigma_{\rm ai}}+\frac{\mu_{\rm vj}^2}{\sigma_{\rm vj}}\right)} &\left[\left[\sum_{t=0}^{\infty}\left(\frac{1}{(2t)!}\frac{1}{(2t)!}\left(2\frac{\mu_{\rm ai}^2}{\sigma_{\rm ai}}\right)^t\left(2\frac{\mu_{\rm vj}^2}{\sigma_{\rm vj}}\right)^t\right)\sum_{s=t}^{\infty}\left[\frac{\mathbf{I}_{\rm ij}^{2s}\mathrm{e}^{\frac{(-{\mathbf{I}_{\rm ij}}^2)}{\sigma_N i}}}{\prod_{m=0}^{s-t-1}(-s+t+m)^2}\left(2\psi(1) - 2\sum_{m=0}^{s-j-1}\frac{1}{-s+t+m}\right.\right.\right.\right. \\
&- \left.\frac{\mathbf{I}_{\rm ij}^{2s}\ln(\mathbf{I}_{\rm ij}^2)}{\prod_{i=0}^{s-t-1}(-s+t+m)^2}\right] + \sum_{r=0}^{\infty}\sum_{t=0}^{\infty}\left(\frac{1}{(2t)!}\left(2\frac{\mu_{\rm ai}^2}{\sigma_{\rm ai}}\right)^t\frac{1}{(2r)!}\left(2\frac{\mu_{\rm vj}^2}{\sigma_{\rm vj}}\right)^r + \frac{1}{(2t)!}\left(2\frac{\mu_{\rm vj}^2}{\sigma_{\rm vj}}\right)^t\frac{1}{(2r)!}\left(2\frac{\mu_{\rm ai}^2}{\sigma_{\rm ai}}\right)^r\right)\sum_{s=t}^{r-i} \\
&\times \left[\mathbf{I}_{\rm ij}^{2s}\frac{\prod_{m=1}^{r-s-1}m}{\prod_{q=0}^{s-j-1}-q-1}\ln(\mathbf{I}_{\rm ij}^2)\right] + \sum_{t=0}^{\infty}\sum_{r=t+1}^{\infty}\left(\frac{1}{(2t)!}\left(2\frac{\mu_{\rm ai}^2}{\sigma_{\rm ai}}\right)^t\frac{1}{(2r)!}\left(2\frac{\mu_{\rm vj}^2}{\sigma_{\rm vj}}\right)^r + \frac{1}{(2t)!}\left(2\frac{\mu_{\rm vj}^2}{\sigma_{\rm vj}}\right)^j\frac{1}{(2r)!}\left(2\frac{\mu_{\rm ai}^2}{\sigma_{\rm ai}}\right)^{2r}\right) \\
&\times \sum_{s=r}^{\infty}\left[\frac{\mathbf{I}_{\rm ij}^{2s}(2\psi(1) - \sum_{m=0}^{r-t-1}\frac{1}{-s+t+m} - \sum_{q=r-t}^{s-t-1}\frac{2}{-s+t+q})}{\prod_{m=0}^{r-t-1}(-s+t+m)\prod_{q=r-t}^{s-t-1}\prod_{q=r-t}^{s-t-1}(-s+t+q)^2} - \frac{\mathbf{I}_{\rm ij}^{2s}\ln(\mathbf{I}_{\rm ij}^2)\mathrm{e}^{\frac{(-{\mathbf{I}_{\rm ij}}^2)}{\sigma_N i}}}{\prod_{i=0}^{r-j-1}(-s+j+i)\prod_{k=r-1}^{s-j-1}(-s+j+k)^2}\right]\Bigg] \\
&\pm \left[\sum_{k=0}^{\infty}\left(\frac{1}{(2k+1)!}\frac{1}{(2k+1)!}\right)\sum_{s=k}^{\infty}\left[\frac{(\mathbf{I}_{\rm ij}^2)^{s+1/2}}{\prod_{m=0}^{s-q-1}(-s+q+m)^2}\left(2\psi(1) - s\sum_{m=0}^{s-q-1}\frac{1}{-s+q+m}\right) - \frac{(\mathbf{I}_{\rm ij}^2)^{s+1/2}\ln(\mathbf{I}_{\rm ij}^2)\mathrm{e}^{\frac{(-{\mathbf{I}_{\rm ij}}^2)}{\sigma_N i}}}{\prod_{i=0}^{s-q-1}(-s+q+i)^2}\right]\right. \\
&+ \sum_{p=1}^{\infty}\sum_{q=0}^{p-1}\left(\frac{1}{(2q+1)!}\left(2\frac{\mu_{\rm ai}^2}{\sigma_{\rm ai}}\right)^{(p+0.5)}\left(2\frac{\mu_{\rm vj}^2}{\sigma_{\rm vj}}\right)^{p+0.5} + \frac{1}{(2q+1)!}\left(2\frac{\mu_{\rm vj}^2}{\sigma_{\rm vj}}\right)^{q+0.5}\frac{1}{(2p+1)}\left(2\frac{\mu_{\rm ai}^2}{\sigma_{\rm ai}}\right)^{p+0.5}\right)\sum_{s=q}^{p-1} \\
&\times \left[\frac{(\mathbf{I}_{\rm ij}^2)^(s+1/2)\prod_{m=1}^{p-s-m}(m)}{\prod_{n=0}^{s-q-1}(-n-1)}\ln(\mathbf{I}_{\rm ij}^2)\right] + \frac{1}{(2q+1)!}\left(2\frac{\mu_{\rm vj}^2}{\sigma_{\rm vj}}\right)^{q+0.5}\frac{1}{(2p+1)!}\left(2\frac{\mu_{\rm ai}^2}{\sigma_{\rm ai}}\right)^{p+0.5}\Bigg)\sum_{s=p}^{\infty} \\
&\times \left.\left.\left[\frac{(\mathbf{I}_{\rm ij}^2)^{s+1/2}(2\psi(1) - \sum_{m=0}^{p-q-1}\frac{1}{-s+q+1} - \sum_{l=p-q}^{s-q-1}\frac{2}{-s+q+l})}{\prod_{i=0}^{p-q-1}(-s+q+m)\prod_{l=p-q}^{s-q-1}(-s+q+l)^2} - \frac{(\mathbf{I}_{\rm ij}^2)^{s+1/2}\ln(\mathbf{I}_{\rm ij}^2)\mathrm{e}^{\frac{(-{\mathbf{I}_{\rm ij}}^2)}{\sigma_N i}}}{\prod_{i=0}^{p-q-1}(-s+q+m)\prod_{l=p-q}^{s-q-1}(-s+q+l)^2}\right]\right]\right]
\end{aligned} \tag{12}$$

TABLE I
SUMMARY OF PARAMETERS

$\mu_{\rm ai}$	mean of image factor a_i
$\mu_{\rm bk}$	mean of individuals' social attractive factor b_k
$\mu_{\rm vj}$	mean of individual's interest factor v_j
$\sigma_{\rm NI}$	variance of image noise N_I
$\sigma_{\rm NC}$	variance of social noise N_C

So the expression for the distribution of $\mathbf{I}_{\rm ij}$ considering the additive noise is given in (12), shown at the bottom of the page.

In a similar way we can also obtain the distribution of $\mathbf{C}_{\rm jk}$.

C. EM for Parameter Estimation

Applying the above, we obtain the exact infinite expansion expression of the PDF. of $\mathbf{I}$ in a series form given in (12). The expression of $\mathbf{C}_{\rm kj}$ can be obtained in a similar way. From (12) we can see that the exact value of $\mu_{\rm ai}$ and $\sigma_{\rm ai}$ does not matter much, but the value of $\frac{\mu_{\rm ai}^2}{\sigma_{\rm ai}}$ matters. So we can assume that a_i has standard derivation of 1, and we only need to calculate the average value of a_i. Similarly, we also do not need to calculate $\sigma_{\rm bk}$ and $\sigma_{\rm vj}$ but only assume that v_j and b_k have standard derivation.

All the parameters $\mathcal{P}$ are summarised in Table I. As mentioned in Section IV-A, in the experiments, when we choose the starting point of the parameters not too large, we can make the series converge in a relatively small number of terms. Then we can apply the standard EM method to refine the parameters iteratively. Experimental result shows that the number of series terms can be no longer than 10 and after several EM iterations, the precision becomes stable.

The EM training process is introduced as follows. For E step, Consider Eq.(12), which is the Equation we want to maximize by knowing the value of $\mathbf{I}_{\rm ij}$, with respect to the parameters $\mathcal{P}$ as follow:

$$\max_{\mathcal{P}} h(\mathbf{I} \mid \mathcal{P})). \tag{13}$$

In the M step, we find the derivative of each parameter in $\mathcal{P}$ by fixing other parameters. Then we set the derivative to be zero to get the value for each parameter. The whole process goes until convergence.

One problem to solve (12) is that (12) contains not only polynomial terms but also exponential terms for the parameters. For simplicity we can make an assumption that the parameters are relatively small, and then we can use the first several terms, or following [34] to get a polynomial expression of the parameters, to make (12) solvable.

Another problem is that for some parameters such as $\mu_{\rm ai}$, it contains infinity high order terms that makes the solution intractable. Again we can make the assumption that these parameters to be smaller than one, and discard the high order terms. In practice we keep the terms whose orders are equal or lower than 4, and follow the method discussed in [35] to calculate the values of the parameters.

From (12) we can obtain the parameters that related to the image-user matrix $\mathbf{I}$, such as $\mu_{\rm ai}$, $\mu_{\rm vj}$, and $\sigma_{\rm NI}$. In a similar manner we can also get the parameters related to the contact matrix $\mathbf{C}$, such as $\sigma_{\rm NC}$, $\mu_{\rm bk}$, and also $\mu_{\rm vj}$. By iteratively updating these parameters relating to the two matrix we can finally determine the value of all the parameters.

After the EM iterations we fix all the parameters in Table I and according these parameters we can make the final friend recommendation.

D. Recommendation Method

When a new user i comes into the network, he/she may upload some favourite photos as well as some tags. The recommendation procedure is divided in two stages. In the first stage, a list of possible friends is generated according to the similarity of the selected important tags. In the experiments, we put the top 200 users into the list.

In the second stage, according to the features of the images uploaded by use i, we get the individuals' interest factor v_i of this user. For a user k in the possible friend list obtained from the first stage, we can also calculate his/her attractive and interest factors b_k. The similarity score of user i and k is obtained by $S_{\rm ik} = v_i b_k$. The higher the similarity score, the more likely that they are to be friends. So we can rank the 200 users in the list according to the similarity score with user i, and recommend the top ones as user i's friends

The whole procedure is given in Algorithm 1.

E. Complexity Analysis

The complexity analysis of our algorithm is also divided by the two stages as follows:

Considering the first stage, the complexity of the network alignment is mainly decided by two steps: the eigenvalue calculation and the inverse of the similarity matrix, which is given by $\max(\min\{n, e\}^3, dn^2)$ as discussed in [5]. e stands for the number of total tags. As previous defined, n stands for the number of users, and d stands for the first d eigenvectors.

To solve the topic model of the second stage, Assume together we need to make L time iterations. in each iteration of the EM step, assume that we calculate the first g terms of the series of (12) (In practice we make $g = 4$). And it takes e steps to solve a 4th order polynomial equation, as mentioned in Section IV-C. Then the complexity would be of $\mathcal{O}(L * e * g * (n * f + n * n))$, where f is the number of image features, as previously defined.

V. EXPERIMENTS

In this section, we make experiments to show the advantage of our proposed method. First, we introduce our social media dataset, and then we discuss the results of our algorithm by comparing it with reference methods. We utilise a cluster containing 16 cores and 128G memories to run our experiments.

A. Dataset and Feature Extraction

We crawled a social network from the big image sharing site Flickr. As the data set is quite large, a relatively unbiased dataset was obtained. In total we crawled the data of 30000 users, and for each user, we crawled all their photos, and tags of each photo. In this paper we tried the SIFT feature and the deep network

TABLE II
DATASET STATISTICS

Users	30000
Photos	1,356,293 photos from 30000 users
CNN features	4096
Contact	628,153 friend links among users
Tags	42,739 words after filtering

extracted features through an CNN autocoder realized by Caffe [36]. For the CNN features we follow the steps of the widely used AlexConvNet [37] and use the 4096 dimensional features vectors from the last full-connected layer. In most cases the CNN features performs better than the SIFT features, so we chose the CNN extracted features for the rest of our experiments. In the future we can also refine feature extraction method for better performance. We then crawled the user contact information to form the contact network. The contact information in Flickr was acquired by checking if a user added another user to his/her friend list, or vice versa. We crawled all the contacts between any two users in our dataset. A short summary of our dataset is given in Table II.

B. Settings and Metrics

Our task is to make precise contact information prediction. When a new user enters into the social network, we recommend new friends according to key words and photos that represent the user's interests.

In friend recommendation, assume we recommend T friends to each user. We use the existing contact information as the ground truth for training and testing. In the first stage, the parameter μ of (1) is determined on the training set by a four-fold cross validation to find the best. The range for the parameter is: $\mu \in 10^{[-2:1:3]}$.

We use the method summarised in Algorithm 1 to recommend friends to new users. We use the recommendation precision metrics to show the effectiveness of the proposed algorithm. In our experiment, precision is defined as the number of correctly recommended friends divided by all the recommended users. We also introduce the precision-recall curve to further show the advantage of our algorithm, where recall is defined as the number of the correctly recommended friends divided by the number of all friends.

During our experiments we divide the whole usere set randomly into two groups: 4/5 of all the users are in the training set and the rest are in the test set. The important features in the first stage are selected on the training set, where the parameters in the second stage are also trained. When a new user in the test set comes into the system with some uploaded tags and photos, T friends will be recommended to him/her from the training set. Assume that together we have recommended $RecAll$ real friends to the test users (totally 6000 users), then the overall precision is calculated by $RecAll/(6000 \times T)$. We adopt a five-fold cross validation to ensure that all the users are utilised as training and testing data once

Algorithm 1: Two Stage Friend Recommendation.

Input:
tag feature matrix $\mathbf{T}$, contact matrix $\mathbf{C}$, image-user matrix $\mathbf{I}$, tag and image feature of the new user $\mathbf{t}$ and $\mathbf{i}$, the numbers of possible friends in Stage 1 and final friends $k1$ and k, respectively

Output:
Friend recommendation list of the new friend

Training:
Stage I
1: Determine λ and μ in (1) via cross validation.
2: Solve (1) with the method in [5]
Stage II
3: Generate the expression of distribution of h[(12)] in the form of series.
4: Apply EM method determining the parameters in Table I

Testing:
5: Stage I: Use $\mathbf{W}$ calculated in Step 2 to obtain $k1$ possible friend list.
6: Stage II: Use the parameters in Step 4 to refine the final recommendation friend list,recommend top k users

C. Reference Methods

The performance analysis of our first stage: network alignment methods can be seen in some previous related papers such as [21], [5]. For the performance analysis of the second stage in which the topic model method is applied, we choose several widely-used methods for comparison.

The first is the variational method, which has been widely applied in this decade for solving the Bayesian network problem[10]. Basically we apply the methods in [9] with some slight modifications to our problem.

The second is the widely-used Gibbs sampling method, which is also very popular in dealing with topic model. Compared with the variational method, the idea of Gibbs sampling is simpler but usually it has difficulty in dealing with large scale problems. We apply the method based on [7] for comparison.

The third method is a co-clustering based method [21]. It is not a topic model-based method, but has a relatively simpler concept: In the second stage, we do co-clustering of image features, users and tags to get a . We apply a simple ranking method, similar to [21] for the final friend recommendation.

To further check the advantage of our method, we also compare our whole two-stage recommendation algorithm with several state-of-the-art recommendation systems. The first one is based on matrix factorization(MF). MF method decomposes the item-user or user-user matrix to infer the latent factors that catch individuals' interests and has been widely discussed for different kinds of recommendation prolems[13], [14]. In this paper we apply a recent method proposed in [14] for comparison, for it jointly considers the information from two different domains.

Another recent method is based on Bayesian collaborative filtering that takes the social connections into account, called SBPR [38]. As a widely-used recommendation method, collaborative filtering assumes that two users that choose the same

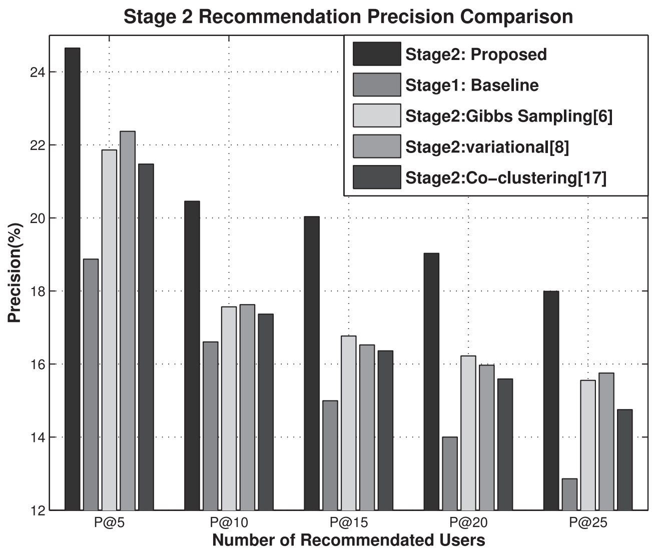


Fig. 3. Stage 2 recommendation precision comparison.

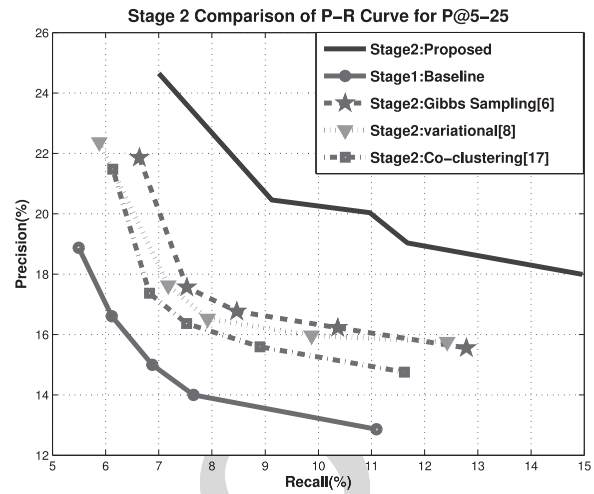


Fig. 4. Recommendation precision and recall for stage 2.

items behave similar on other items. Traditional collaborative filtering methods do not consider much about the social connections between users. SBPR removes this drawback by taking the social connections into account by assigning a social coefficient to each user.[1]

At last we consider a multi-network based algorithm for comparison. When considering social multiple network problems, transition probability propagation is a method that is frequently used [39], [40]. We choose [39] as a reference method for the following reasons: 1) It considers the relationships of different networks, which is similar to our idea; 2) It uses the information of other networks for recommendation, which again has some similarities with ours. [39] enhances the links in one network and between different networks using a random walk propagation method. After a sufficient number of walks, it obtains the modified link weights between each user pair. We use the weights for friend recommendation.

D. Experimental Results

Here we report the results of our method for friend recommendation as follows.

1) Performance of Series Expansion: In this experiment we compare the proposed series expansion method with the variational, Gibbs sampling, and co-clustering methods in the second stage. We treat the performance of the first stage as the baseline.

From Fig. 3 we can see that our method has the best performance for accurate recommendation. $P@X$ stands for that each time we recommend the top X friends to users. Generally, the second stage improves the recommendation precision from only applying the first stage, illustrating the effectiveness and necessity of applying the two staged methods. Our proposed method improves about 5–7% compared with the performance of the first stage, and also makes about 2–3% improvement compared with the Gibss sampling method and the variational method. The reason for the improvement mainly lies in that we apply an exact expression to approach the PDF of the data, rather than an approximation or sampling method. The co-clustering method lacks the ranking ability and thus the performance is not good.

Fig. 4 illustrates the precision-recall curve of the proposed and reference methods. Based on the result of the first stage, the series expansion method achieves the highest performance(The upper right line on the figure). We can see from Fig. 4 that when precision or recall is fixed, we can achieve a 3–4% improvement over the best reference methods. This means that the proposed method can achieve both the highest precision and recall. This experimental results shows that the series expansion method can best approximate the real distribution of the data, and thus makes the most precise recommendation.

On the other hand, the proposed method have also imposed Gaussian distribution assumption to the latent variables a, b, and v. This may also cause some negative effect although it can give an analytic expression. It is worthy to make a depth observation of the distribution of the latent variables in our future studies.

2) Performance of the Proposed Two-Stage Method: Now we compare our two-stage method with some recently-proposed recommendation systems as mentioned in V-C. The main results for precision and precision-recall curve are shown in Figs. 5 and 6.

From Figs. 5 and 6 we can see that our system achieves the best performance, compared with other state-of-the-art recommendation systems. In average, our system improves the recommendation accuracy by about 3-4%, compared with the second best one. MF based method [14] has the best performance among all the reference methods, for it decomposes the item-user and user-user matrix into different social factors in a proper way. The reason that the proposed method performs better than MF might lies in that the MF method does not consider the whole distribution of the network and is trapped into some local optimum. Collaborative filtering based method [38] has slight lower performance than [14], the reason might be that its assumptions about the users' positive and negative feedback are not very proper for the Flickr dataset. Finally, the random-walk based method [39] has the lowest performance, since the

[1]The realization of [14] and [38] is based on the existing open-source Java package LibRec at http://www.librec.net/.

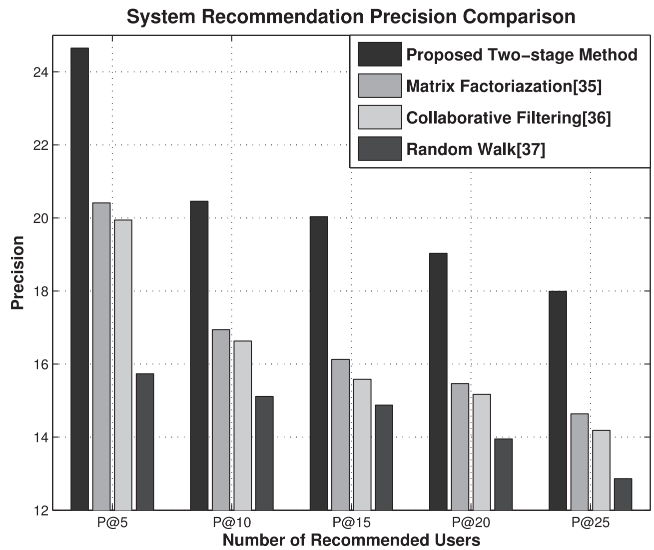


Fig. 5. Two-stage recommendation precision compared with state-of-the-art systems.

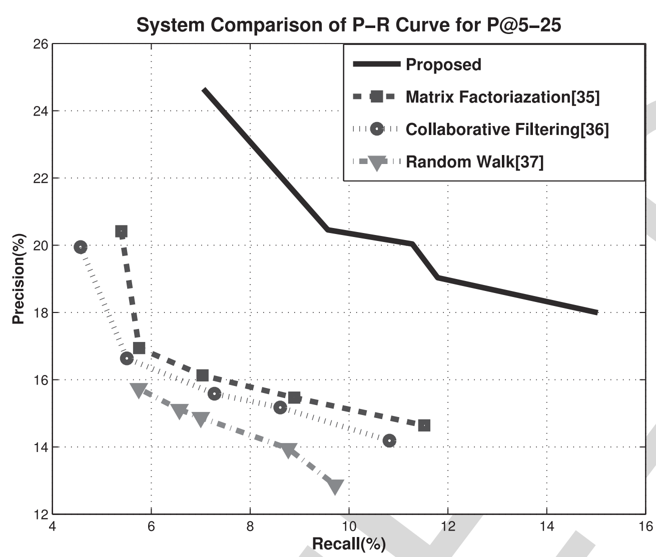


Fig. 6. Recommendation precision and recall compared with state-of-the-art systems.

random walk algorithm is not accurate enough for precise friend recommendation.

E. The Influence of Several Settings

1) The Influence of Additional Noise: The introduction of the additive noise, as shown in Section IV-B, makes the model more precise. However, it also leads to complicated inferences and calculations. In the following experiment we study the influence of the additive noise. In Table III, we compare the recommendation accuracy of the model that contains the additive noise and the model that does not contain the noise.

From Table III we see that by considering the additive noise we get a precision gain of about 1–2%, which is useful in the case where a more precise result is required.

TABLE III
INFLUENCE OF ADDITIVE NOISE

Precision(%)	P@5	P@10	P@15	P@20	P@25
Model With Noise	24.6	21.0	19.8	18.1	17.5
Model Without Noise	22.7	19.3	18.2	16.8	15.9

TABLE IV
INFLUENCE OF VALUES OF **C** AND **I**

y	0.3	1	5	10
Precision(%)	19.6	24.6	13.7	11.0

2) The Influence of the Value of $\mathbf{C}_{\mathrm{kj}}$ *and* $\mathbf{I}_{\mathrm{ij}}$*:* As shortly discussed in Section IV, the convergence speed of the series is largely determined by the level of values of **C** and **I**. If it is too large, then the convergence speed will decrease, leading to either the inaccuracy of the model, or larger number of terms. On the other hand, if the level is too small, the logarithmic terms in (12) will drop quickly and make the system unstable. In our experiments, contact network **C** stands for the intimacy of two individuals and in the image-user network **I**, it stands for to what extent an individual favours an image feature. The values of each entry of **C** and **I** can be set according to our requirements. For example, we can set $\mathbf{C}_{\mathrm{jk}}$ to be 1 if two individuals are friends with each other and 0 otherwise; for image-user network we can also set $\mathbf{I}_{\mathrm{ij}} = 1$ if an individual has a certain image feature in his/photos, and 0 otherwise. On the other hand, we can also raise the level of the elements in **C** and **I** to be 5 or 10, or reduce it to be smaller than 1. The relationship between any two nodes would not change in the networks by varying the element value of **C** and **I**, but the value does have an influence on the accuracy in our algorithm. We set the value of **C** and **I** on four levels to be 0.3, 1, 5 and 10 to check its influence on the performance.

In the following we compare the recommendation precision of these four levels.

From Table IV we see that the recommendation precision decreases rapidly as we increase the value of **C** and **I**. On the other hand, if it is too small, the performance also goes down as the system becomes unstable around the poles of the logarithmic terms in (12). This indicates that we should choose the value of **I** and **C** around 1 for precise calculation.

VI. CONCLUSION AND FUTURE WORKS

In this paper, we develop a two-stage friend recommendation scenario utilizing multimedia information. In the first stage, tag information is utilised to build a tag-similarity network and is aligned to a contact network by a number of important features to generate a "possible friend list". In the second stage, a topic model is proposed and a new method based on series expansion is developed to combine image features and contact information to make more precise recommendations.

The experimental results show that the proposed method outperforms other methods in friend recommendation in that our method achieves the highest precision and recall in friend prediction. The network alignment of Stage One is effective. The topic model in Stage Two refines the result of stage one and the new series expansion method has better performance than the traditional variational and Gibbs sampling methods.

We will further develop our algorithm. For the series expansion method, it is a novel and effective method but not perfect. It is still to some extent mathematically complicated and has difficulties to apply on different models. We plan to refine the idea to make it more manoeuvrable and can be applied on general topic models. There are two directions to dig further. Firstly, for more complicated topic models, it might be viewed as a combination of some simpler models and thus are solvable based on our method. Secondly, our method is specially developed for Gaussian distributed random variables. For some other simple distributions, their algebra has been discussed in [23], [24], [27], etc. It is our future work to develop some general frameworks to combine all these distributions together.

For our staged recommendation framework, we will extend our ideas to further applications such as product recommendation, media retrieval, etc. One problem of the current method is that in the first stage, some real friend might be omitted. We will further study how to increase the recalls in the first stage. We will develop other algorithms in each of our two stages, and to utilise the information form different domains. We will also make some studies about the ranks of the information from different domains. That is, which data should be applied in the first stage to achieve better performance. In the last, we can also introduce the concept of deep learning in our scenario for more efficient feature learning.

REFERENCES

[1] W. Alex, "Why is it hard to make friends over 30?," *New York Times*, Jun. 2012.

[2] W. Zeng and L. Chen, "Heterogeneous data fusion via matrix factorization for augmenting item, group and friend recommendations," in *Proc. 28th Annu. ACM Symp. Appl. Comput.*, 2013, pp. 237–244.

[3] A. Chin, B. Xu, and H. Wang, "Who should I add as a "friend"?: A study of friend recommendations using proximity and homophily," in *Proc. 4th Int. Workshop Model. Soc. Media*, 2013, pp. 7:1–7:7.

[4] C. Guo, X. Tian, and T. Mei, "User specific friend recommendation in social media community," in *Proc. IEEE Int. Conf. Multimedia Expo*, Jul. 2014, pp. 1–6.

[5] S. Huang, J. Zhang, L. Wang, and X. S. Hua, "Social friend recommendation based on multiple network correlation," *IEEE Trans. Multimedia*, vol. 18, pp. 287–299, Feb. 2016.

[6] Z. Wang, J. Liao, Q. Cao, H. Qi, and Z. Wang, "Friendbook: A semantic-based friend recommendation system for social networks," *IEEE Trans. Mobile Comput.*, vol. 14, no. 3, pp. 538–551, Mar. 2015.

[7] W. Min, B.-K. Bao, C. Xu, and M. Hossain, "Cross-platform multi-modal topic modeling for personalized inter-platform recommendation," *IEEE Trans. Multimedia*, vol. 17, no. 10, pp. 1787–1801, Oct. 2015.

[8] N. Zheng, S. Song, and H. Bao, "A temporal-topic model for friend recommendations in chinese microblogging systems," *IEEE Trans. Syst., Man, Cybern.: Syst.*, vol. 45, no. 45, pp. 1245–1253, Sep. 2015.

Q1 [9] Y. Song, J. Zhang, and L. Cao, "A joint interest-social latent factor model for social recommendation," Tech. Rep., 2013.

[10] D. M. Blei, A. Y. Ng, and M. I. Jordan, "Latent dirichlet allocation," *J. Mach. Learn. Res.*, vol. 3, pp. 993–1022, Mar. 2003.

[11] T. L. Griffiths and M. Steyvers, "A probabilistic approach to semantic representation," in *Proc. 24th Ann. Conf. Cogn. Sci. Soc.*, 2002, pp. 381–386.

[12] M. Welling and Y. W. Teh, "Hybrid variational/Gibbs collapsed inference in topic models," in *Proc. Conf. Uncertain. Artif. Intell.*, 2008, pp. 587–594.

[13] H. Ma, H. Yang, M. R. Lyu, and I. King, "SoRec: Social recommendation using probabilistic matrix factorization," in *Proc. 17th ACM Conf. Inf. Knowl. Manage.*, 2008, pp. 931–940.

[14] G. Guo, J. Zhang, and N. Yorke-Smith, "TrustSVD: Collaborative filtering with both the explicit and implicit influence of user trust and of item ratings," in *Proc. 29th AAAI Conf. Artif. Intell.*, 2015, pp. 123–129.

[15] D. Lu, J. Sang, Z. Chen, M. Xu, and T. Mei, "Who are your real friend: Analyzing and distinguishing between offline and online friendships from social multimedia data," *IEEE Trans. Multimedia*,2017, to be published.

[16] M. Yan, J. Sang, T. Mei, and C. Xu, "Friend transfer: Cold-start friend recommendation with cross-platform transfer learning of social knowledge," in *Proc. IEEE Int. Conf. Multimedia Expo*, Jul. 2013, pp. 1–6.

[17] A. Chin, B. Xu, and H. Wang, "Who should i add as a "friend"?: A study of friend recommendations using proximity and homophily," in *Proc. 4th Int. Workshop Model. Soc. Media*, 2013, pp. 7:1–7:7.

[18] J. Sang and C. Xu, "Right buddy makes the difference: An early exploration of social relation analysis in multimedia applications," in *Proc. 20th ACM Int. Conf. Multimedia*, 2012, pp. 19–28.

[19] L.-H. Li, F.-M. Lee, Y.-C. Chen, and C.-Y. Cheng, "A multi-stage collaborative filtering approach for mobile recommendation," in *Proc. 3rd Int. Conf. Ubiquitous Inf. Manage. Commun.*, 2009, pp. 88–97.

[20] X. Zhao and J. Wang, "A theoretical analysis of two-stage recommendation for cold-start collaborative filtering," in *Proc. Int. Conf. Theory Inf. Retrieval*, 2015, pp. 71–80.

[21] S. Huang, J. Zhang, S. Lu, and X.-S. Hua, "Social friend recommendation based on network correlation and feature co-clustering," in *Proc. 5th ACM Int. Conf. Multimedia Retrieval*, 2015, pp. 315–322.

[22] S. Geman and D. Geman, "Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 6, no. 6, pp. 721–741, Nov. 1984.

[23] B. D. Carter and M. D. Springer, "The distribution of products, quotients and powers of independent H-fucntion variates," *SIAM J. Appl. Math.*, vol. 33, pp. 542–558, Dec. 1977.

[24] M. D. Springer and W. Thompson, "The distribution of products of beta, gamma and Gaussian random variables," *SIAM J Appl. Math.*, vol. 18, pp. 721–737, Jun. 1977.

[25] J. M. Pruett, "The distribution of products of some independent, nonstandardized random variables," Ph.D. dissertation, Univ. Arkansas, Fayetteville, AR, USA, 1972.

[26] H. Michiel *et al.*, "Mellin transform," in *Encyclopedia of Mathematics*. Berlin, Germany: Springer, 2001.

[27] M. D. Springer, *The Algebra of Random Variables*. New York, NY, USA: Wiley, 1979.

[28] S. Ahmed, L.-L. Yang, and L. Hanzo, "Probability distributions of products of Rayleigh and Nakagami-m variables using Mellin transform," in *Proc. IEEE Int. Conf. Commun.*, Jun. 2011, pp. 1–5.

[29] R. Mallik and N. Sagias, "Distribution of inner product of complex gaussian random vectors and its applications," *IEEE Trans. Commun.*, vol. 59, no. 12, pp. 3353–3362, Dec. 2011.

[30] H. Bateman, *Tables of Integral Transforms*, vol. 1, New York, NY, USA: McGraw-Hill, 1954.

[31] L. Ahlfors, *Complex Analysis*. New York, NY, USA: McGraw Hill, 1979.

[32] E. L. Lady, "Convergence of infinite series in general and taylor series in particular," Oct. 1998. Q2

[33] W. Balser, *From Divergent Power Series to Analytic Functions*. Berlin, Germany: Springer, 1994.

[34] W. A. Richards and M. R. Acharya, "An efficient polynomial approximation to the normal distribution function and its inverse function," *J. Math. Res.*, vol. 2, 2010. Q3

[35] "Ferrari's solution of a quartic equation," *Maths notes*.

[36] Y. Jia et al., "Caffe: Convolutional architecture for fast feature embedding," *CoRR*, 2014. [Online]. Available: http://arxiv.org/abs/1408.5093

[37] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet classification with deep convolutional neural networks," in *Advances in Neural Information Processing Systems 25*, F. Pereira, C. J. C. Burges, L. Bottou, and K. Q. Weinberger, Eds., Red Hook, NY, USA: Curran Associates, Inc., 2012, pp. 1097–1105.

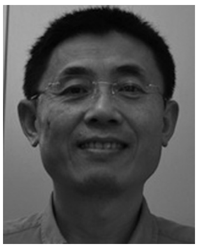
[38] T. Zhao, J. McAuley, and I. King, "Leveraging social connections to improve personalized ranking for collaborative filtering," in *Proc. 23rd ACM Int. Conf. Inf. Knowl. Manage.*, 2014, pp. 261–270.

[39] M. Jiang *et al.*, "Social recommendation across multiple relational domains," in *Proc. ACM Conf. Inf. Knowl. Manage.*, Oct. 2012, pp. 1422–1431.

[40] D. Liu, G. Ye, C.-T. Chen, S. Yan, and S.-F. Chang, "Hybrid social media network," in *Proc. ACM Multimedia Conf.*, Oct. 2012, pp. 659–668.

Shangrong Huang received the B.Sc. degree in telecommunication from Zhejiang University, Hangzhou, China, in 2005, and the M.Sc. degree in information and communication technology from the Darmstadt University of Technology, Darmstadt, Germany, in 2012, and is currently working toward the Ph.D. degree at the University of Technology Sydney, Sydney, NSW, Austraila.

His research interests include recommendation system, social medias, and big data analysis.



Q4 **Jian Zhang** (S'95–M'98–SM'04) received the B.Sc. degree from the East China Normal University, Shanghai, China, the M.Sc. degree in computer science from Flinders University, Adelaide, SA, Australia, and the Ph.D. degree in electrical engineering from the University of New South Wales (UNSW), Sydney, NSW, Australia, in 1999.

From 1997 to 2003, he was with the Visual Information Processing Laboratory, Motorola Labs, Sydney, NSW, Australia, as a Senior Research Engineer, and later became a Principal Research Engineer and a Foundation Manager in the Visual Communications Research Team. From 2004 to July 2011, he was a Principal Researcher and a Project Leader with National ICT Australia, Sydney, NSW, Australia, and a Conjoint Associate Professor with the School of Computer Science and Engineering, UNSW. He is currently an Associate Professor with the Global Big Data Technologies Centre and School of Computing and Communication, Faculty of Engineering and Information Technology, University of Technology Sydney, Sydney, NSW, Australia. He has authored or coauthored more than 100 paper publications and book chapters, and has six issued patents filed in the U.S. and China. His research interests include multimedia processing and communications, image and video processing, machine learning, pattern recognition, media and social media visual information retrieval and mining, human–computer interaction and intelligent video surveillance systems.

Dr. Zhang was the General Co-Chair of the International Conference on Multimedia and Expo in 2012 and Technical Program Co-Chair of the IEEE Visual Communications and Image Processing 2014. He was an Associate Editor of the IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS FOR VIDEO TECHNOLOGY (2006–2015) and the *EURASIP Journal on Image and Video Processing* (2016–present).

Dan Schonfeld (S'88–M'90–SM'05–F'10) received the B.S. degree in electrical engineering and computer science from the University of California at Berkeley, Berkeley, CA, USA, in 1986, and the M.S. and Ph.D. degrees in electrical and computer engineering from The Johns Hopkins University, in 1988 and 1990, respectively.

In 1990, he joined the University of Illinois at Chicago, Chicago, IL, USA, where he is currently a Professor with the Departments of Electrical and Computer Engineering, Computer Science, and Bioengineering. He has authored more than 200 technical papers in various journals and conferences. His research interests include signal processing, image and video analysis, video retrieval and communications, multimedia systems, computer vision, medical imaging, and genomic signal processing.

Dr. Schonfeld is Fellow of the SPIE for specific achievements in morphological image processing and video analysis. He has been elected as the University Scholar of the University of Illinois and was the recipient of the Graduate Mentoring Award of the University of Illinois at Chicago. He is currently serving as the Editor-in-Chief of the IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS FOR VIDEO TECHNOLOGY. He has served as the Deputy Editor-in-Chief of the IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS FOR VIDEO TECHNOLOGY and the Area Editor for Special Issues of the *IEEE Signal Processing Magazine*. He was the coauthor of a paper that was the recipient of the Best Paper Award at the ACM Multimedia Workshop on Advanced Video Streaming Techniques for Peer-to-Peer Networks and Social Networking 2010. He was also the coauthor of papers that were the recipient of the Best Student Paper Award at Visual Communication and Image Processing 2006 and the IEEE International Conference on Image Processing 2006 and 2007.

Lei Wang (SM'xx) received the Ph.D. degree from Nanyang Technological University, Singapore, in 2004. Q5

He is currently with the School of Computing and Information Technology, University of Wollongong, Wollongong, NSW, Australia, as an Associate Professor. His research interests include machine learning, pattern recognition, and computer vision. For machine learning and pattern recognition, he is interested in feature selection, model selection, and kernel-based learning methods. For computer vision, he is interested in image categorisation and content-based image retrieval.

Xian-Sheng Hua received the B.S. and Ph.D. degrees in applied mathematics from Peking University, Beijing, China, in 1996 and 2001, respectively.

In April 2015, he was a Researcher and Senior Director of the Alibaba Group, leading the multimedia technology team in the search division. In 2013, he was a Senior Researcher of Microsoft Research Redmond, Redmond, WA, USA, where he worked on web-scale image and video understanding and search, and related applications. In 2011, he was a Principal Research and Development Lead of multimedia search for the Microsoft search engine, Bing, where he led a team that designed and delivered leading-edge media understanding, indexing, and searching features. In 2001, he joined Microsoft Research Asia, Beijing, China, as a Researcher. He has authored or coauthored more than 250 research papers in these areas and has filed more than 90 patents. His research interests include multimedia search, advertising, understanding, and mining, and recognition and machine learning. Q6

Dr. Hua is an Associate Editor of the IEEE TRANSACTIONS ON MULTIMEDIA and an Associate Editor of the ACM TRANSACTIONS ON INTELLIGENT SYSTEMS AND TECHNOLOGY. He served as a Program Co-Chair for IEEE ICME 2013, ACM Multimedia 2012, and IEEE ICME 2012, as well as on the Technical Directions Board of the IEEE Signal Processing Society. He was the recipient of the prestigious 2008 MIT Technology Review TR35 Young Innovator Award for his outstanding contributions to video search. He was the recipient of the Best Paper and Best Demonstration Awards at ACM Multimedia 2007, the Best Poster Award at IEEE International Workshop on Multimedia Signal Processing 2008, the Best Student Paper Award at ACM Conference on Information and Knowledge Management 2009, the Best Paper Award at International Conference on MultiMedia Modeling 2010, the Best Demonstration Award at ICME 2014, and the Best Paper Award of the IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS FOR VIDEO TECHNOLOGY in 2014.

QUERIES

Q1. Author: Please provide the report number, as well as name and location of institution.
Q2. Author: Please provide complete bibliographic details in Refs. [32] and [35].
Q3. Author: Please provide page range in Ref. [34].
Q4. Author: Please provide the year in which the B.Sc. and M.Sc. degrees were earned.
Q5. Author: Please provide the year in which “Lei Wang” became the “Senior Member” of the IEEE.
Q6. Author: Please provide the location of Alibaba Group.