Inferring Emotions from Social Images Leveraging Influence Analysis

Boya Wu[1,2], Jia Jia[1,2], Xiaohui Wang[1,2], Yang Yang[1], and Lianhong Cai[1,2]

[1] Department of Computer Science and Technology, Tsinghua University, Beijing 100084, China
[2] Key Laboratory of Pervasive Computing, Ministry of Education,
Tsinghua National Laboratory for Information Science and Technology (TNList)
stella.1991@163.com, jjia@mail.tsinghua.edu.cn,
wangxh09@mails.tsinghua.edu.cn, sherlockbourne@gmail.com,
clh-dcs@tsinghua.edu.cn

Abstract. Nowadays thriving image-based social networks such as Flickr and Instagram are attracting more and more people's attention. When it comes to inferring emotions from images, previous researches mainly focus on the extraction of effective image features. However, in the context of social networks, the user's emotional state is no longer isolated, but influenced by her friends. In this paper, we aim to infer emotions from social images leveraging influence analysis. We first explore several interesting psychological phenomena on the world's largest image-sharing website Flickr[1]. Then we summarize these pattern into formalized factor functions. Introducing these factors into modeling, we propose a partially-labeled factor graph model to infer the emotions of social images. The experimental results shows a 23.71% promotion compared with Naïve Bayesian method and a 21.83% promotion compared with Support Vector Machine (SVM) method under the evaluation of F1-Measure, which validates the effectiveness of our method.

Keywords: Images, emotion, social influence.

1 Introduction

Emotion plays a vital role in human life. It is said that emotion stimulates the mind 3,000 times faster than rational thoughts[1]. Understanding the internal dynamics and external manifestation of human emotion can not only pushes forward the frontier of sociology and psychology, but also benefits the promotion of product design and user experience.

The factors that determine human emotional state varied. We may be delighted for a feast or a beautiful melody, and we may feel sad because of our friends' bitter experience. The way we express our feelings also varies. Besides texts, images provide us with a novel way to convey what we feel. Compared with texts, images are more vivid, freer, but more difficult and more subjective to understand as well.

[1] www.flickr.com

H.-Y. Huang et al. (Eds.): SMP 2014, CCIS 489, pp. 141–154, 2014.

Nowadays, with the thriving of many image-based social networks, such as Flickr and Instagram, we're used to upload images on the Internet and share them with our friends. The rapid development of these websites and applications provides us with large amounts of available data, thus offering us a great opportunity to study the problem of the inference of emotions embedded in images.

Previous researches mainly focus on the exploration of effective image features, such as colors, composition and shapes. J.Machajdik and A.Hanbury[2] pick up four categories of low-level features like wavelet textures and GLCM-features and proved the effectiveness of these features on three data sets. Moreover, from an aesthetic point of view, X.Wang[3] explored the interpretable image features, including the color combinations, saturation, brightness, the warm and cool color ratio, etc. and conduct experiments on two public data sets. Other works can be found in [14], [15]. However, in the context of social networks, the user's emotion state is no longer isolated. That is to say, the user's emotional state is not only determined by images features, but also affected by her friends' emotional states and her unique experience.

This point of view is also supported by sociological and psychological theories. S.Hareli and A.Rafaeli[4] discovered that one's emotion will influence other people's emotion, thoughts and behavior. In turn, other people's reaction will influence the future interaction between them as well as this person's future emotion and behavior. J.Fowler and N.Christakis[5] studied social networks and mentioned that people's happiness depends on the happiness of those who are connected to them.

In this paper, we aim to tackle the image emotion inference problem by leveraging the influence analysis. Figure 1 illustrates the procedure of our work. Our study is based on the world's largest image-sharing website Flickr and 2,060,353 images and 1,255,478 user profiles are constructed as the data set of our experiment. Based on these data, three types of correlations are considered and studied, namely, 1) Attributes correlation: the correlation between image emotion and image features. 2) Temporal correlation: the correlation between the emotion of current image and the images the user uploaded before. 3) Social correlation: the correlation between the image emotion and the user's interaction with her friends. We first make observations and discover several interesting psychological phenomena about the emotion influence on the social network and then summarize them into formalized factor functions. Next by introducing these factor functions we propose a partially-labeled factor graph model which fulfils the inference of images emotion. The experimental results show that the F1-Measure reaches 0.4251 on average, which increases by 23.71% compared to Naïve Bayes and 21.83% compared to SVM, thus validating the effectiveness of our method.

The rest of this paper is organized as follows: Section 2 presents the data observation. Section 3 formulates the problem. Section 4 explains the proposed model and algorithm. Section 5 shows the experiments we conduct and investigates the experimental results. Section 6 concludes the work of this paper.

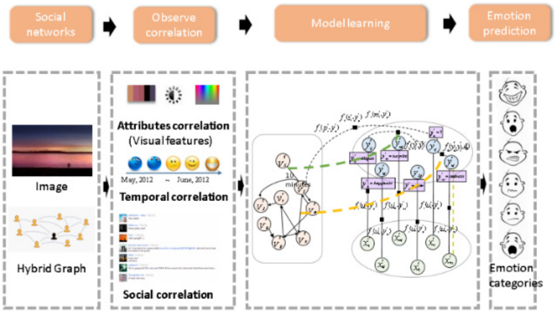

Fig. 1. The procedure of inferring emotions from social images leveraging influence analysis

2 Data Observation

In order to leverage the information of emotion influence, to begin with we need to explore the patterns emotion influence acts on the social networks. In this work, we direct our attention to the following three aspects: the intimacy between the user and her friends, the dominant emotion of friends and the composition of friends.

2.1 Data

We randomly downloaded 2,060,353 images and 1,255,478 users from Flickr and employ them as our data set. To evaluate the performance of inferring emotions, we first need to know the primitive emotions embedded in the images. Due to the large-scale of our data set, manually labeling the emotion for every image is not practical. Therefore we adopt a method used by Y.Yang[12] to automatically label the emotions. Based on Ekman's theory[6] that human emotion can be generally classified into six categories, namely, happiness, surprise, anger, disgust, fear and sadness, we first construct six word list in which words are closely related to the basic emotions from WordNet[7] and HowNet[8]. Then we manually verify every word list to make sure that the words are relative to the emotions. Next we compare every word in the tag written by the user who upload the image and find out which word list matches the tag the best. The image is labeled with that basic emotion if the best match exists. By this method, 354,192 images are qualified. To make sure that the data we used for observation is typical and representative, we further restrict the completeness of data, which means that the time stamp of the image and the related information is not absent. In this way, we pick out 10,448 images for data observation.

2.2 The Intimacy between Friends

The intimacy between friends may exert a profound influence on the degree of emotion influence. For example, the user may experience the same emotion as her best friend, but may not feel linked with a newly acquainted person. Because comments of

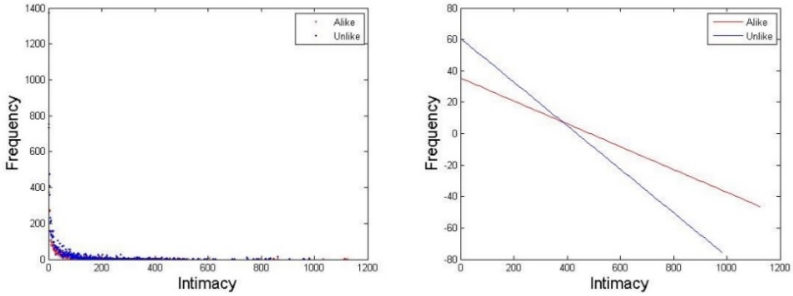

Fig. 2. The correlation between emotion influence and the intimacy between friends

images are public to every user, so herein we define friends as the other users whose images are ever commented by this user. Meanwhile, we carefully make an assumption that the user's emotional state can be conveyed through the image the user uploads. Given the definition and the assumption, we observe the emotional states of users and their friends on the day they upload an image and the day before. We also observe the interactions between users and their friends.

Figure 2 illustrates the observation results. It is noted that when the user is intimate with a friend, she probably feel the same with that friend, but when the user is not familiar with a friend, she may not likely to influenced by that friend.

2.3 The Dominant Emotion of Friends

The dominant emotion of the user's friends may affect the user's emotional state. Interestingly, will the user feel happy if all her friends feel happy? Will the answer remain the same when it comes to different emotion categories? To find answers of these questions, we conduct experiments and the observation results are shown in Figure 3.

As can be seen from Figure 3, whether we consider emotion influence or not, over half of the users on the image-based social networks are happy. What's more, negative emotions, namely, surprise, anger, disgust, fear and sadness are more capable of influencing others. In particular, when most of the user's friends feel surprised, angry or disgust, the possibility of the user feel the same doubled or even trebled. However, in terms of positive emotion, which refers to happy in the six basic emotion categories, the tendency is not obvious, indicating that negative emotions have a rather larger influence than positive emotions.

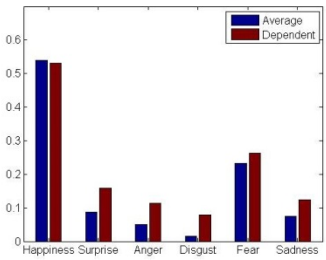

Fig. 3. The correlation between emotion influence and the dominant emotion of friends

2.4 The Composition of Friends

The composition of friends may have something to do with the degree of emotion influence. For instance, are those having friends of various locations or occupations more likely to be influenced or those having similar friends? Are those having more friend inclined to be influenced or those having less friends?

Excitingly, several interesting as well as important phenomena are discovered. The left figure in Figure 4 shows that users who have less friends are more likely to be influenced by their friends and users who have more friends are less likely to be influenced, while no obvious difference is found in terms of the location diversity and occupation diversity. The middle figure indicates that when users have more female friends, they are more probable to be affected. The right figure is a little bit different. It's not about the users' friends, but the users themselves. Interestingly, it can be seen that female users and taken users have a tendency to feel the same with their friends compared with male users and single users.

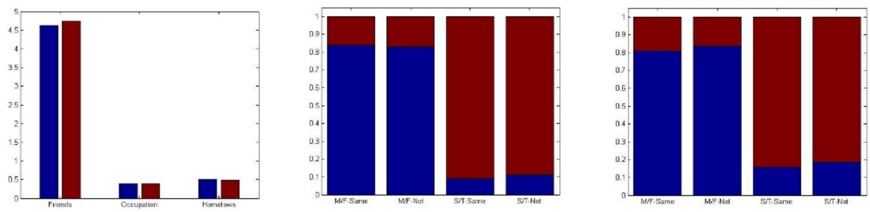

Fig. 4. The correlation between emotion influence and the composition of friends

3 Problem Definition

By exploring the patterns of emotion influence, we discovered many interesting psychological phenomena. In this section, we'll define the notations of this work and formalize the problem we study.

The goal of our work is to analyze and predict the emotion of images uploaded by users in the context of social networks leveraging influence analysis.

A static social network can be represented as $G = (V, E)$ where V is the set of $|V| = n$ users, $E \subset V \times V$ is the set of relationship among users. In our work, this relationship is can be regarded as the friendship. First, we should give friend a definition.

Definition 1. Friend: If user v_i ever comments any image uploaded by user v_j before time t, then v_j is called v_i's friend and v_i's friends at time t are denoted as $FR^t(v_i)$.

Based on the definition, it can be seen that user v_i's friends $FR^t(v_i)$ at time t may be empty, which means that v_i has never commented any image before. As time goes by, when v_i makes her first comment, $FR^t(v_i)$ won't be empty anymore. So $FR^t(v_i)$ is time-dependent, and more precisely E is better denoted as E^t.

Acquiring the concept of friend, we can further define the intimacy between friends.

Definition 2. Intimacy: After standardization, the frequency of user v_i making comments of v_j's images at time t denoted as μ_{ij}^t is called the intimacy between user v_i and v_j.

The standardization here is the process of discretizing the frequency into three degrees of intimacy, namely, $\mu_{ij}^t \in \{1,2,3\}$. The larger $\mu_{ij}^t$is, the higher intimacy between user v_i and v_j is.

On the base of these definitions, we define emotion next.

Definition 3. Emotion: The emotional state of user v_i at time t is denoted as y_i^t. The emotion of the image x_{ij}^t uploaded by user v_i at time t is denoted as y_{ij}^t. In our work, an assumption is carefully made as follows: According to Ekman's theory, human emotion can be classified into six basic categories, which are happiness, surprise, anger, disgust, fear and sadness. We adopted Ekman's theory and denote the emotional space as R, where $y_i^t \in R$, $y_{ij}^t \in R$.

Given the definition of friend and emotion, we are able to obtain the size of user's friends and the dominant emotion of friends. Herein we devote the size of user's friends as s_i^t and denote the dominant emotion of friends as m_i^t, where $m_i^t \in R$. It is clearly noted that $s_i^t = size(FR^t(v_i))$.

As for the user v_i herself, the user's personal profile, including v_i's gender and whether v_i is single or taken, can be defined as vector p_i. p_i is stable.

Besides, we use X to represent all images uploaded by all users from V. In detail, we use X^L to represent the subset of X where the images are labeled, and use X^U to represent the subset of X where the images are unlabeled. For an image $x_{ij}^t \in X$ uploaded by user v_i at time t, the image features are denoted as vector u_{ij}^t.

We can see that the relationship between users are varying with time. Therefore a time-varying social network is to be defined.

Definition 4 Partially-Labeled Time-Varying Social Network: A partially-labeled time-varying social network can be defined as $G = (V, E^t, X^L, X^U)$, where V is the set of users, E^t represents the friendship between users at time t, X^L represents the labeled images and X^U represents the unlabeled images.

Accomplishing the definition of these notations, the learning task of our model is put forward as follows.

Learning Task: Given a partially-labeled time-varying social network $G = (V, E^t, X^L, X^U)$, find a function f to predict the emotion of all unlabeled images:

$$f: G = (V, E^t, X^L, X^U) \to Y \tag{1}$$

where $Y = \{y_{ij}^t\}$, $y_{ij}^t \in R$, meaning the emotion of images.

4 Model

In this paper, we proposed a factor graph model to infer the emotion of social images. In the model, many types of correlations are defined and introduced into the model through factor functions. These correlations are: 1) Attributes correlation: The emotion of the image is basically induced by the image features. 2) Temporal correlation: The emotion of current image may be related to the emotion of images the user upload before. 3) Social correlation: The emotion embedded in the image may be influenced by the user's interaction with her friends. These correlations can better help us with the emotion inference problem.

4.1 The Predictive Model

The graphical representation of the model is illustrated in Figure 2. The model is constructed on a basic factor graph model. The input of the model is a partially-labeled time-varying network G, and after the learning and training process, the emotions of the unlabeled images are inferred effectively.

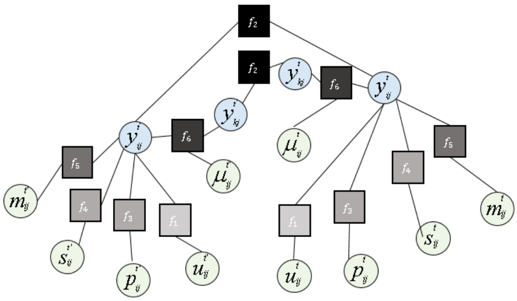

Fig. 5. The illustration of the model

There are two types of nodes in the factor graph model, one called variable node and the other called function node. Shown in Figure 2, the variable nodes are round and the function nodes are square. In our work, all images uploaded by users and their relative information can be regarded as variable nodes and the attributes correlation, temporal correlation and social correlation are leveraged in the form of formalized factor functions.

In the model, the three correlations mentioned above are defined as follows.

1) Attributes correlation: The attributes correlation refers to the correlation between image emotion and image features. It can be encoded as $f_1(u_{ij}^t, y_{ij}^t)$, where u_{ij}^t means the image features of image x_{ij}^t uploaded by user v_i at time t, y_{ij}^t means the emotion of image x_{ij}^t. By introducing parameter α, it can be instantiated as exponential-liner function:

$$f_1(u_{ij}^t, y_{ij}^t) = \frac{1}{z_\alpha} exp\,\{\alpha^T \cdot u_{ij}^t\} \tag{2}$$

2) Temporal correlation: The temporal correlation indicates the correlation between the emotion of current image and emotions of images the user uploaded before. It can be depicted as $f_2(y_i^{t'}, y_i^t)$, $t' < t$, where y_i^t and $y_i^{t'}$ means the emotion of user v_i at time t and t', conveyed by the images the user uploaded at that time. By introducing the function $g(y_i^{t'}, y_i^t)$ to depict the correlation, it can be instantiated as follows:

$$f_2(y_i^{t'}, y_i^t) = \frac{1}{z_\xi} exp\,\{\xi_i \cdot g(y_i^{t'}, y_i^t)\} \tag{3}$$

3) Social correlation: The social correlation means the correlation between the emotion embedded in the image and the user's interaction with her friends. Based on the data observation, the social correlation can be detailed into following four aspects.

- The correlation between image emotion and the user's personal attributes: The correlation can be encoded as $f_3(p_i, y_{ij}^t)$, where p_i is the personal information of user v_i, including v_i's gender and whether v_i is single or taken, and y_{ij}^t is the emotion of image x_{ij}^t. By introducing parameter β, it can be instantiated as:

$$f_3(p_i, y_{ij}^t) = \frac{1}{z_\beta} exp\,\{\beta^T \cdot p_i\} \tag{4}$$

- The correlation between image emotion and the size of user's friends: The correlation can be encoded as $f_4(s_i^t, y_{ij}^t)$, where s_i^t represents the size of user v_i's friends and y_{ij}^t is the emotion of image x_{ij}^t. By introducing parameter γ, it can be instantiated as:

$$f_4(s_i^t, y_{ij}^t) = \frac{1}{z_\gamma} exp\,\{\gamma^T \cdot s_i^t\} \tag{5}$$

- The correlation between image emotion and the dominant emotion of the user's friends: The correlation can be encoded as $f_5(m_i^t, y_{ij}^t)$, where m_i^t represents the dominant emotion of user v_i's friends and y_{ij}^t is the emotion of image x_{ij}^t. By introducing parameter δ, it can be instantiated as:

$$f_5(m_i^t, y_{ij}^t) = \frac{1}{z_\delta} exp\,\{\delta^T \cdot m_i^t\} \tag{6}$$

- The correlation between image emotion and the user's intimacy with her friends: The correlation can be encoded as $f_6(y_i^t, y_j^t, \mu_{ij}^t)$, where y_i^t and y_j^t represents the emotional states of user v_i and v_j at time t and μ_{ij}^t measures the intimacy between user v_i and v_j at time t. By introducing the function $h(y_i^t, y_j^t, \mu_{ij}^t)$ to depict the correlation, it can be instantiated as follows:

$$f_6(y_i^t, y_j^t, \mu_{ij}^t) = \frac{1}{z_\eta} exp\,\{\eta_{ij} \cdot h(y_i^t, y_j^t, \mu_{ij}^t)\} \tag{7}$$

So far we have formalized all the correlations in the model. Thus we can define the joint distribution of the model:

$$P(Y|G) = \frac{1}{Z} \times \prod_{x_{ij}^t \in X, v_i \in V} f_1\left(u_{ij}^t, y_{ij}^t\right) \times \prod_{x_{ij}^t \in X, v_i \in V} \prod_{y_i^{t'} \in SU^t(v_i)} f_2(y_i^{t'}, y_i^t)$$
$$\times \prod_{x_{ij}^t \in X, v_i \in V} f_3\left(p_i, y_{ij}^t\right) \times \prod_{x_{ij}^t \in X, v_i \in V} f_4\left(s_i^t, y_{ij}^t\right) \times \prod_{x_{ij}^t \in X, v_i \in V} f_5\left(m_i^t, y_{ij}^t\right)$$
$$\times \prod_{x_{ij}^t \in X, v_i \in V} \prod_{v_j \in FR^t(v_i)} f_6\left(y_i^t, y_j^t, \mu_{ij}^t\right) = \frac{1}{Z} exp\, \{\theta^T S\} \tag{8}$$

where $Z = Z_\alpha Z_\xi Z_\beta Z_\gamma Z_\delta Z_\eta$ is the normalization term, S is the aggregation of factor functions over all nodes, θ denotes all the parameters, i.e., $\theta = \{\alpha, \beta, \gamma, \delta, \xi_i, \eta_{ij}\}$, and $Y = \{y_{ij}^t\}$, $y_{ij}^t \in R$, meaning the inferred emotion for images.

Therefore the target of the inference process is to maximize the log-likelihood objective function $O = log\, P(Y|G)$, and the keynote of the training process is to learn $\theta^* = \arg\max O(\theta)$.

4.2 Model Learning

Given the model's input and output, we make clear the parameters and the objective function of the model. Next we'll detail the learning process of the model and the algorithm is summarized as follows.

Algorithm: the learning and inference algorithm of image emotion
Input: a partially-labeled time-varying network $G = (V, E^t, X^L, X^U)$, learning ratio λ
Output: parameter group $\theta = \{\alpha, \beta, \gamma, \delta, \xi_i, \eta_{ij}\}$, inference result $Y = \{y_{ij}^t\}$, $y_{ij}^t \in R$

Read in network G. Construct factor graph. Set up variable nodes and function nodes.
Initiate parameters $\theta = \{\alpha, \beta, \gamma, \delta, \xi_i, \eta_{ij}\}$
Repeat
    Calculate $E_{p_\theta(Y|Y^U,G)} S$ using LBP
    Calculate $E_{p_\theta(Y|G)} S$ using LBP
    Calculate the gradient of θ: $E_{p_\theta(Y|Y^U,G)} S - E_{p_\theta(Y|G)} S$
    Update θ with the learning ratio λ: $\theta = \theta_0 + \frac{\partial O}{\partial \theta} \cdot \lambda$
Until convergence
Get the inference result $Y = \{y_{ij}^t\}$, $y_{ij}^t \in R$ and the trained parameters $\theta = \{\alpha, \beta, \gamma, \delta, \xi_i, \eta_{ij}\}$

To observe which part each parameter is linked to in the objective function, $O = log\, P(Y|G)$ can be rewritten as follows:

$$O = log\, P(Y|G) = log \sum\nolimits_{Y|Y^U} exp\, \{\theta^T S\} - logZ$$
$$= log \sum\nolimits_{Y|Y^U} exp\, \{\theta^T S\} - log \sum\nolimits_{Y} exp\, \{\theta^T S\} \tag{9}$$

Thus the gradient of the parameter can be obtained. The gradient of θ can be represented as:

$$\frac{\partial O}{\partial \theta} = \frac{\partial (log \sum_{Y|Y^U} exp\{\theta^T S\} - log \sum_{Y} exp\{\theta^T S\})}{\partial \theta} = E_{p_\theta(Y|Y^U,G)} S - E_{p_\theta(Y|G)} S \qquad (10)$$

Then we can update θ by $\theta = \theta_0 + \frac{\partial O}{\partial \theta} \cdot \lambda$ according to the learning ratio λ. λ can be manually defined and in our work $\lambda = 0.1$.

Given the current parameters, we can obtain the inferred emotion of unlabeled images which fulfills the target of modeling.

5 Experiments

In this section, we validate the effectiveness of the proposed method through experiments. First we give an illustration of the data set we employed. Then we compare the experimental results with other baseline methods. Finally to find out which factor contributes the most, we conduct factor contribution analysis and present some interesting sociological and psychological phenomena.

5.1 Experimental Setup

The data set we employed is randomly downloaded from world's largest image-sharing website Flickr which comprised of 2,060,353 images and 1,255,478 users. As explained in Section 2, by extracting emotion from image tags and checking the completeness of relative information, 218,816 qualified images are left. In order to test the performance of every emotion category, we evenly and randomly pick out 6,000 images from every emotion category and 36,000 images are chosen in total.

As for image features, considering the interpretability of image features and the fine performance when used for emotion inference, we adopt the method proposed by X.Wang[3] to extract the features, including the five major color types, saturation, brightness and so on.

We conduct performance comparison experiments to demonstrate the effectiveness of our method. The two baseline methods we employed for comparison are the Naïve Bayesian method and the SVM method.

Naïve Bayesian: Naïve Bayesian is a widely used classifier. The image features are directly inputted into the classifier and the classifier outputs the inferred image emotions.

SVM: SVM is frequently used in many classification problems. Here it directly uses the image features as the input and outputs the inferred image emotion. In this work, we use LIBSVM design by C.Chang and C.Lin[9].

The data set we employed to the baseline methods and our proposed methods is the same, 60% for training and 40% for testing as well. The evaluation measure we used contains precision, recall and F1-Measure.

5.2 Experimental Results

Table 1 exhibits the experimental results.

From the table, it is apparent that our method significantly enhanced the performance of the image-based emotion inference problem. The average F1-Measure reaches 0.4254, increased by 23.71% compared with Naïve Bayesian and 21.83% compared with SVM. The Naïve Bayesian method and the SVM method barely use the image features for training and let go of the emotion influence on the social network. Under this circumstance, the classifier can only try to analyze the image emotion from the image's saturation, brightness and so on. Though these methods are effective to a certain extent, in the complex context of social networks, the user's emotional state is no longer isolated and influenced by other users' emotional states instead. So taking only the image features into account cannot measure the user's emotional state precisely. In contrast, because our method concerns about the emotion influence on the social network and formalize these patterns into the factor graph model, it better depicts the real situation of the social networks and examines the user's emotional state in its entirety, thus remarkably promote the performance of the emotion inference.

Table 1. Performance of emotion inference

Emotion	Method	Precision	Recall	F1-Measure
Happiness	Naïve Bayesian	0.2026	0.2287	0.2148
	SVM	0.2319	0.2251	0.2284
	Our method	**0.3880**	**0.4652**	**0.4232**
Surprise	Naïve Bayesian	0.1765	0.0061	0.0119
	SVM	0.2038	0.0179	0.0329
	Our method	**0.3911**	**0.3261**	**0.3556**
Anger	Naïve Bayesian	0.1962	0.1071	0.1385
	SVM	0.2025	0.1673	0.1832
	Our method	**0.3854**	**0.3752**	**0.3802**
Disgust	Naïve Bayesian	0.2047	0.2664	0.2315
	SVM	0.2216	0.3079	0.2577
	Our method	**0.4543**	**0.5990**	**0.5167**
Fear	Naïve Bayesian	0.2009	0.2707	0.2306
	SVM	0.1972	0.2253	0.2103
	Our method	**0.4778**	**0.3454**	**0.4010**
Sadness	Naïve Bayesian	0.2424	0.4008	0.3021
	SVM	0.2711	0.4223	0.3302
	Our method	**0.4845**	**0.4674**	**0.4758**
Average	Naïve Bayesian	0.2039	0.2133	0.1883
	SVM	0.2213	0.2276	0.2071
	Our method	**0.4318**	**0.4297**	**0.4254**

Wondered which factors play a vital role in the promotion, we conduct factor contribution analysis and several interesting sociological and psychological phenomena are discovered.

5.3 Factor Contribution Analysis

Herein, we test the contribution for each factor function used in the model. On the basis of the primitive model, we take out one of the factor and then examine the performance of the model while the other factors remain unchanged. The experimental results evaluated by F1-Measure are shown in Table 2.

As is shown in Table 3, the model which involves all factors achieves the best performance no matter it is evaluated by precision, recall or F1-Measure. The experimental results confirms that the $f3$, $f4$, $f5$ and $f6$ factor all make contribution to the promotion of the model. That is to say, on the image-based image social networks, the emotion influence indeed exist. It also proved that the correlations we observed in Section 3 is correct and effective.

More excitingly, it is clearly shown in the result that the $f3$ factor contributes the most in the model, which implies that whether the user is male or female and whether the user is single or taken did makes a great difference in the user's perception of images. It corresponds to the sociological and psychological theory. A.Fischer[10] pointed out that males tend to express more powerful emotions while females tend to express less powerful emotions, which indicates there is actually a gender difference in human emotion perception.

Besides, the $f6$ and $f4$ factor, which represents the intimacy between user and her friends and the size of her friends exert a profound influence on the user's emotional state. J.Pennebaker[11] discovered that people are more likely to share their emotion with those who are close to them, such as the family members, spouse and close friends.

J.Whitfield[13] found that happy people are more likely to be surrounded by happy people, while sad people tend to make friends with those who usually feel sad. People having similar emotions have a tendency to gather together, which illustrates the contribution of $f5$ factor.

Table 2. The contribution of different factor

Emotion	Model	Model-$f3$	Model-$f4$	Model-$f5$	Model-$f6$
Happiness	0.4232	0.4169	0.4136	0.4183	0.4149
Surprise	0.3556	0.3272	0.3208	0.3298	0.3303
Anger	0.3802	0.3686	0.3839	0.3835	0.3858
Disgust	0.5167	0.4459	0.4995	0.5022	0.5012
Fear	0.4010	0.3425	0.3519	0.3652	0.3642
Sadness	0.4758	0.4738	0.4643	0.4701	0.4692
Average	**0.4254**	**0.3958**	**0.4057**	**0.4115**	**0.4109**

6 Conclusion

In this paper, we study the problem of inferring emotions of social images leveraging influence analysis and proposed a novel method to solve the problem.

We first explored the patterns of emotion influence on the world's largest image-sharing website Flickr and observed several interesting psychological phenomena. Then we summarized these patterns into three types of correlations and introduced them in the factor graph model in the form of factor functions, which fulfills the modeling and inference of emotions of social images. By conducting experiments, we validate the effectiveness of our method and present a noteworthy promotion of the problem.

Image-based social networks such as Flickr and Instagram are thriving, making the problem of inferring emotions of social images of great significance. The research advances can provide a useful back-up for sociology and psychology and help Internet companies offer better services for customers.

Acknowledgement. This work is supported by the National Basic Research Program (973 Program) of China (2012CB316401), National Natural, and Science Foundation of China (61370023, 61003094). This work is partially supported by the major project of the National Social Science Foundation of China (13&ZD189) and funded by Microsoft Research Asia-Tsinghua University Joint Laboratory: FY14-RES-SPONSOR-111.

References

1. Tang, J., Zhang, Y., Sun, J., Rao, J., Yu, W., Chen, Y., Fong, A.: Quantitative study of individual emotional states in social networks. IEEE Transactions on Affective Computing 3(2), 132–144 (2012)
2. Machajdik, J., Hanbury, A.: Affective image classification using features inspired by psychology and art theory. In: Multimedia, pp. 83–92. ACM (2010)
3. Wang, X., Jia, J., Yin, J., Cai, L.: Interpretable aesthetic features for affective image classification. In: Proc. of ICIP 2013, Melbourne, pp. 3230–3234 (2013)
4. Hareli, S., Rafaeli, A.: Emotion cycles: On the social influence of emotion in organizations. In: Research in Organizational Behavior, vol. 28, pp. 35–59 (2008)
5. Fowler, J., Christakis, N., Steptoe, Roux, D.: Dynamic Spread of Happiness in a Large Social Network: Longitudinal Analysis of the Framingham Heart Study Social Network. British Medical Journal 338(7685), 23–27 (2009)
6. Ekman, P.: An argument for basic emotions. Cognition and Emotion 6(3-4), 169–200 (1992)
7. Fellbaum, C.: Wordnet. In: Theory and Applications of Ontology: Computer Applications, pp. 231–243 (2010)
8. Dong, Z., Dong, Q.: HowNet and the Computation of Meaning. World Scientific (2006)
9. Chang, C., Lin, C.: LIBSVM: A library for support vector machines. ACM TIST 2, 27:1–27:27 (2011)
10. Fischer, A., Rodriguez Mosquera, P., van Vianen, A., Manstead, A.: Gender and culture differences in emotion. Emotion 4(1), 87–94 (2004)

11. Pennebaker, J., Zech, E., Rimé, B.: Disclosing and Sharing Emotion: Psychological, Social and Health Consequences. In: Stroebe, M.S., Stroebe, W., Hansson, R.O., Schut, H. (eds.) Handbook of Bereavement Research: Consequences, Coping, and Care, pp. 517–539. American Psychological Association, Washington, DC (2001)
12. Yang, Y., Jia, J., Zhang, S., Wu, B., Chen, Q., Li, J., Xing, C., Tang, J.: How Do Your Friends on Social Media Disclose Your Emotions? In: Proceedings of the 28th AAAI Conference on Artificial Intelligence
13. Whitfield, J.: The secret of happiness: Grinning on the Internet. Nature (June 26, 2008)
14. Borth, D., Ji, R., Chen, T., Breuel, T., Chang, S.: Large-scale visual sentiment ontology and detectors using adjective noun pairs. In: Multimedia, pp. 223–232. ACM (2013)
15. Shin, Y., Kim, E.: Affective prediction in photographic images using probabilistic affective model. In: ICIVR, pp. 390–397. ACM (2010)