Information & Management 52 (2015) 859–869

Contents lists available at ScienceDirect

Information & Management

journal homepage: www.elsevier.com/locate/im

Subjective well-being measurement based on Chinese grassroots blog text sentiment analysis

Jiayin Qi [a,*], Xiangling Fu [b,1], Ge Zhu [a,1]

[a] School of Economic and Management, Beijing University of Posts and Telecommunications, No. 10 Xitucheng Road, Haidian District, Beijing 10086, China
[b] School of Software, Beijing University of Posts and Telecommunications, No. 10 Xitucheng Road, Haidian District, Beijing 10086, China

ARTICLE INFO

Article history:
Received 26 August 2014
Received in revised form 12 March 2015
Accepted 1 June 2015
Available online 10 June 2015

Keywords:
Subjective well-being
Social media
Chinese text
Sentiment analysis
Text mining

ABSTRACT

In this study, we propose a new method to measure the subjective well-being (SWB) of Chinese people. Based upon the classic framework in psychology, our model constructs a system of multiple weighted emotions in positive and negative affect by applying a text-sentiment analysis. To study SWB in the Chinese context, we also establish and supplement our model with a new lexicon, Ren-CECps-SWB 2.0. Tests on the data of 7 years of grassroots blogs on Sina.com demonstrate the validity of our model. Employing the same data, we find interesting patterns of the SWB of Chinese people on weekly and monthly bases.

1. Introduction

Though self-report scales are the most popular means in psychological studies to measure SWB, self-report scales have disadvantages, such as the limited samples they can assess, their high cost, and their sensitivity to participants' memory, which makes it hard to present the real-time status of respondents [10,15,16,18]. With the rapid development of online social network services (SNS) [25], increasing numbers of people are creating user-generated content (UGC) on the Internet to express their emotions [14]. Because of the rich information in UGC [1,2,22], scholars have attempted to measure SWB through UGC. For example, Dodds and Danforth used Affective Norms for English Words (ANEW) to measure SWB [10,11]. In 2009, Facebook released Facebook Gross National Happiness (FGNH) to measure the aggregate level of SWB [12,13]. Bollen and his colleagues [3] used a similar method with FGNH to calculate SWB. Dodds and Danforth's method is based on economic utility theory and uses valence value to predict SWB. FGNH applies the dualistic classification of negative and positive words to compute SWB. Despite the insights generated by prior studies, as far as we know, the SWB measurements from text sentiment analysis seldom follow the established methods in psychology. Without harvesting the accumulative insights in psychology for SWB research, we might not be able to stand on the shoulders of giants to gain new understandings of the phenomenon under investigation. Moreover, the extant studies did not offer sufficient reasons why some emotions are selected into the SWB measurement and others are not. In this research, we attempt to overcome these research limitations by following the established psychological measurement for SWB based on multiple weighted emotions through text sentimental analysis.

In addition, although the research that uses sentiment analysis to measure SWB in an English-speaking context is gaining momentum, no research as of today has focused on Chinese text to measure the SWB of Chinese people. However, Chinese semantic analysis and English semantic analysis have striking differences. Applying the results from a sentiment analysis in English contexts to measure Chinese SWB may blind the real SWB of Chinese people because of the huge differences between English and Chinese cultures and languages. Fourth, no present corpus in Chinese can be used to build an SWB measurement model directly. Therefore, we intend to fill the current research gaps in Chinese SWB measurement applying a sentiment analysis technique in the Chinese language. We selected the Positive Affect and Negative Affect Schedule (PANAS) from psychology to construct our SWB

* Corresponding author. Tel.: +86 10 62282039; fax: +86 10 62282039.
E-mail addresses: qijiayin@bupt.edu.cn (J. Qi), fuxiangling@bupt.edu.cn (X. Fu), wwzhuge@gmail.com (G. Zhu).
[1] Tel.: +86 10 62282039; fax: +86 10 62282039.

http://dx.doi.org/10.1016/j.im.2015.06.002
0378-7206/© 2015 Published by Elsevier B.V.

model. On the basis of Ren-CECps, we constructed a Ren-CECps-SWB 2.0 Chinese lexicon. Integrating localized Chinese PANAS and Ren-CECps-SWB 2.0, we built an SWB measurement model for the Chinese context. We validated our model using data collected from the blog Sina.com from 2008 to 2013.

We selected Sina.com for SWB measurement for three major reasons. First, Sina.com has the largest number of users among all Chinese blog platform services. Sina.com has more than 10 million registered users, and its daily page views exceed three million. Second, Sina.com has the longest operation history in China. For most of the major events that occurred in China in recent years, there are corresponding blogs that can be searched and accessed on the Sina.com blog platform. The data availability facilitated our analysis and tests. Third, bloggers at Sina.com have clear categories, such as entertainment star blogs, intellectual celebrity blogs, and grassroots blogs. Grassroots bloggers normally express their own feelings and emotions in their blogs. They care about the people and events around them. Therefore, grassroots blogs reflect the ups and downs of common people in China. Accordingly, in this research, we chose the blogs of grassroots bloggers as our data source.

This research makes three contributions to the SWB research. First, we build a novel SWB measurement model based on multiple weighted emotions from PANAS. Second, we construct a lexicon, Ren-CECps-SWB 2.0, used specifically to measure the SWB implied by Chinese texts based on the Ren-CECps corpus. Third, we provide an SWB measurement model including five basic emotions in Chinese.

The remainder of this paper is organized as follows. We first provide a literature review. Then, we develop an SWB measurement model based on PANAS. Third, we modify Ren-CECps corpus and construct the Chinese lexicon Ren-CECps-SWB 2.0. Fourth, we build a novel SWB measurement model for Chinese and validate our measurement model. Finally, we draw implications for research and practice.

2. Literature review

2.1. Definition of SWB

Though SWB has been well studied, there is no consensus definition for it. In this research, following Ed Diener, one of the leading researchers in the field of SWB, we define SWB as the way a person evaluates his/her own life, including emotional experiences of pleasure versus pain in response to specific events and cognitive evaluations of what a person considers a good life [7,6]. Based on this definition, SWB consists of cognitive well-being and affective well-being [8]. In this paper, we restrict our investigation on SWB to affective well-being merely due to data availability. (As of today, there is no valid method available for us to mine measurements for cognitive wellbeing from text sentiment analysis.) This study conceives of certain types of pleasant experiences (viz. pleasant moods) as more valuable than others, e.g., "transient pleasant sensations" [4]. In our study, therefore, SWB is a summarization of individuals' emotional experiences of the continuous events in their daily life and work, including both positive emotions, such as love and happiness, and negative emotions, such as sorrow and anxiety.

2.2. Prior research on SWB measurement through text sentiment analysis

Several methods have been proposed to measure SWB through text sentiment analysis. One method is used by Facebook to build its Facebook Gross National Happiness (FGNH) index. When constructing the index, the company first uses the number of positive (negative) words in users' status updates to proxy positivity/negativity, and the FGNH index is the standardized difference between positivity and negativity. Another method is proposed by Dodds and Danforth [10]. They used Affective Norms for English Words (ANEW) to measure the implied SWB by estimating the overall valence score for a text. Though the two basic methods have their merits, they also have limitations. For example, in FGNH, only genetic positive emotions and negative emotions are involved. The dualistic classification and equal weights of general positive emotion and negative emotion in SWB are over simplified. Dodds and Danforth integrated good-bad (valence), but psychologists insist that good and bad emotions are independent scales for SWB [9]. Moreover, SWB involves multiple dimensions of emotion, and each emotion may make a different contribution to SWB [17]. To overcome the limitations of prior methods for measuring SWB in text sentiment analysis, as discussed in the next section, this paper constructs a new SWB method using automated UGC sentiment analysis. Our approach is based on a more delicate classification of emotions, and each emotion has a specified weight in the SWB analysis.

3. SWB measurement model and its specification in Chinese context

In this section, we propose a new SWB model specific to the Chinese context based on UGC. Our model extends the PANAS framework by constructing and measuring its key components, positive affect and negative affect, using the online UGC and text sentiment analysis techniques. To construct SWB based on Chinese UGC, we also introduce a new sentiment lexicon by extending the Ren-CECps lexicon.

3.1. SWB measurement model based on PANAS

The Positive and Negative Affect Schedule (PANAS) is one of the most widely used scales to measure mood or emotion [24]. This brief scale is comprised of 20 items, with 10 to measure positive affect (PA, e.g., excited, inspired) and the other 10 to measure negative affect (NA, e.g., upset, afraid). Each item is rated on a five-point Likert scale. PANAS provides a classification for both PA and NA. Schmukle et al. [21] provide solid evidence that the PA and the NA are unrelated, which suggests that they can be used as independent indicators for measuring SWB [21].

We propose an SWB measurement model based on PANAS. The key difference between our model and PANAS scale is that in PANAS,[1] PA and NA are measured by a self-report survey. In our model, however, we use the UGC of online grassroots blogs to conduct text sentiment analysis to measure PA, NA, and SWB. In our model, first we calculate the proportion of each sentiment word in a text. The more frequently a word appears in an online text conversation, the more representative the word is in the conversation, and the more weight will be assigned to this word in measuring SWB. Then, we calculate the value of each emotion vector from PA and NA in a text by summing up each emotion vector value from each sentiment word in a text. Finally, we obtain the weighted sum of all emotion vectors in a text to gauge SWB in this text. When we average all SWB from those blogs in a period of time, we obtain the aggregate SWB in this period of time. On

[1] Because we want to measure Chinese people's well-being, the PANAS framework from Watson, Clark and Tellegen cannot be used directly because of huge differences between English and Chinese. We have to establish a Chinese version of the PANAS framework to measure Chinese SWB. Based on the original PANAS from Watson, Clark and Tellegen, Qiu, Zheng and Wang developed a Chinese localized PANAS, which includes nine emotions in PA and NA, respectively [20]. Our model is based on this Chinese-localized PANAS.

Table 1
Comparison among the three Chinese corpuses.

Chinese corpus	Contents	Advantages	Disadvantages
NTUSD	11,086 words (2810 positive, 8276 negative)	Large volume of words	Only positive and negative are notated; too few are positive words
HowNet	11,000 words 271 information structures 49 syntax distributions 58 syntax structures	Large volume of words, flexibility, expansibility	Only positive and negative are notated
Ren-CECps	More than 11,000 words 53,422 records	Large volume of words, 8-dimension emotion notation	Not as flexible and expansible as HowNet

this basis, we calculate the basic emotions of the texts in online conversations. Our model specification can be written as follows:

$$p_{ik} = \frac{f_{ik}}{\sum_{n_i}^{k=1} f_{ik}} \tag{1}$$

$$d_{ij} = \sum_{k=1}^{n_i} p_{ik} e_{jk} \tag{2}$$

$$H_i = \sum_{j=1}^{m} \omega_j d_{ij} \tag{3}$$

$$SWB_T = \left(\sum_{t_i \in T} H_i \right) / n_T \tag{4}$$

where

p_{ik} is the ratio of the frequency of word k to the number of all sentiment words in text i;
f_{ik} is the frequency of word k in text i;
n is the total number of sentiment words in text i;
d_{ij} is the value of j the emotion vector in text i;
e_{jk} is the j the emotion vector for word k in the sentiment lexicon;
H_i is the affective well-being contained in text i;
m is the number of emotions used in SWB measurement[2];
ω_j is the weight of emotion j in SWB;
SWB_T is the aggregate SWB in the period of time T;
t_i is the release time of text i;
n_T is the total number of texts in the period of time T.

Eqs. (1)–(4) suggest that to measure SWB in a Chinese context through text sentiment analysis, we need to construct a Chinese lexicon and determine parameters ω_j and m.

3.2. Chinese lexicon Ren-CECps-SWB 2.0

Prior research suggests that to measure SWB, a sentiment lexicon suitable to describe emotion must be constructed [10]. For example, Dodds and Danforth [10] used the English sentiment lexicon (ANEW) to measure SWB. As such, to study SWB in the Chinese context, constructing a viable Chinese sentiment lexicon is a key to our study. In this study, we propose a new lexicon by extending a popular Chinese corpus, Ren-CECps.

[2] There are 10 emotions in PA and 10 emotions in NA, but more is not necessarily better. On one hand, if we include all emotions, the computation will be too complex, and the accuracy of the sentiment analysis for all emotions will decrease. On the other hand, if we include fewer emotional words in our model, we may lose some key information in SWB measurement by omitting important dimensions. Therefore, we need to balance our measurement results with computation complexity and measurement adequacy to determine parameter m.

3.2.1. Comparison between Chinese corpuses

Many Chinese corpuses have been proposed. The popular ones are NTUSD, HowNet, and Ren-CECps, each of which has its advantages and disadvantages. Table 1 provides a brief comparison of the three Chinese corpuses.

Because the three corpuses are widely applied corpuses for Chinese sentiment analysis, we decided to choose one of them as the basis of our lexicon construction. Ideally, the selected corpus should have a large volume of words. It should also contain multiple emotion categories, and those categories should perfectly overlap with the Chinese localized PANAS developed by Qiu et al. [19]. Although the three corpuses have no obvious distinction in terms of word volume, Ren-CECps includes four positive emotions and four negative emotions, which are highly consistent with the Chinese localized PANAS. In contrast, HowNet and NTUSD have only two emotion dimensions: positive and negative. Moreover, Ren-CECps uses online blog services such as Sina.com, Baidu.com, Tencent.com, and Qzone as data sources. Thus, it includes many Internet words and emerging new words. As such, Ren-CECps has an inherent advantage in processing blog texts. Accordingly, we chose Ren-CECps as our basic corpus.

Ren-CECps was constructed based on a relative fine-grained annotation scheme at three levels: sentence, document, and paragraph [20]. At the sentence level, annotations include emotion categories (expectance, joy, love, surprise, anxiety, sorrow, anger and hate), emotion intensity, emotional keywords/phrases, degree words, negative words, conjunction, rhetoric, punctuation, objective/subjective, and emotion polarity. At the document and paragraph levels, emotion category, emotion intensity, topic words and topic sentences are annotated. The main purpose for constructing this emotion corpus is to support the development and evaluation of emotion analysis systems in Chinese. In Ren-CECps, the emotions of each word are represented by an emotion vector:

$$d\rightarrow = \langle e_1, e_2, \ldots, e_8 \rangle \tag{5}$$

where $d\rightarrow = \langle e_1, e_2, \ldots, e_8 \rangle$ is a basic emotion set contained in a word. The value of e_i (i = 1, 2, ..., 8) ranges from 0.0 to 1.0 (discrete), indicating the intensity of one of the eight basic emotions (expectance, joy, love, surprise, anxiety, sorrow, anger and hate). For example, for the word "like," $d\rightarrow = \langle 0.0, 0.3, 0.9, 0.0, 0.0, 0.0, 0.0, 0.0 \rangle$. That is, the word "like" indicates weak joy and strong love.

3.2.2. Chinese Lexicon for measuring SWB

The Ren-CECps corpus cannot be used directly in SWB measurement. There are two major reasons why. First, Ren-CECps is merely a corpus, as shown in Fig. 1. It is not a lexicon, but we need a lexicon to construct our SWB measurement model. Second, if we extract all the emotion words from the Ren-CECps, we cannot

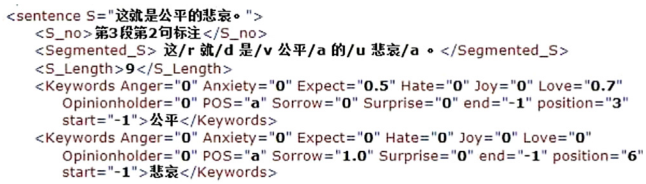
```
<sentence S="这就是公平的悲哀。">
    <S_no>第3段第2句标注</S_no>
    <Segmented_S> 这/r 就/d 是/v 公平/a 的/u 悲哀/a 。</Segmented_S>
    <S_Length>9</S_Length>
    <Keywords Anger="0" Anxiety="0" Expect="0.5" Hate="0" Joy="0" Love="0.7"
        Opinionholder="0" POS="a" Sorrow="0" Surprise="0" end="-1" position="3"
        start="-1">公平</Keywords>
    <Keywords Anger="0" Anxiety="0" Expect="0" Hate="0" Joy="0" Love="0"
        Opinionholder="0" POS="a" Sorrow="1.0" Surprise="0" end="-1" position="6"
        start="-1">悲哀</Keywords>
```

Fig. 1. An example of notations in Ren-CECps.

use it immediately because we must select the most effective words to construct a special lexicon for measuring SWB. We therefore modify and extend it such that we can use it to measure Chinese SWB.

We extracted 95,612 records with emotion vectors and part-of-speech (POS) tags from the Ren-CECps corpus. After deleting duplicate ones, 52,631 records of 20,814 words were retained. In notating the corpus, Ren-CECps considered the usage context and grammar structure such that same words could have the same POS tag but different emotion vectors. For example, "爱国"s in Table 2 are all adjectives, but they indicate different sentiments.

To reserve all connotations of a word as best as we can, we calculate the average value of the emotion vectors of records that contain the same words with the same POS tags to be the emotion vector value of the word of this very POS tag. For the word "爱国" in Table 2, the average intensity of love is 0.643, and the average intensity of joy is 0.086; the other emotions are not indicated by this word (see Table 3).

On the one hand, by calculating average value, we can reserve all possible emotions indicated by a word under a specific POS tag. On the other hand, we measure SWB by analyzing a large amount of text, and this method will not lead to significant bias. Dodds and his colleagues [10] applied a similar method when extending ANEW. After processing, 20,814 records remained as the primary lexicon for measuring SWB. (We label it as Ren-CECps-SWB 1.0.)

After pretesting, however, we find the SWB measurement results based on Ren-CECps-SWB 1.0 do not have sufficient volatility. This is because some words are apparently not emotional, such as "是" (yes) and "的" (of). We thus further refine Ren-CECps-SWB 1.0 by deleting the low-emotional-intensity words from Ren-CECps-SWB 1.0. More specifically, we set one emotion intensity value as the threshold value. We delete the words for which all emotion vector values are below the threshold value, and we check the change of the coverage. If the change of the coverage is not obvious, we lift the threshold value by 0.1 and repeat the process. If the change of the coverage decreases sharply, we confirm that the threshold value is the optimal. The trial and error results are shown in Table 4.

We start with 20,814 words. If the threshold value is 0.1, the coverage is 0.999 (20,789/20,814), which has low volatility and

Table 2
Notations of "爱国" as adjective.

Word	POS	Surprise	Sorrow	Love	Joy	Hate	Expectance	Anxiety	Anger
爱国	a	0	0	0.7	0	0	0	0	0
爱国	a	0	0	0.6	0	0	0	0	0
爱国	a	0	0	0	0.6	0	0	0	0
爱国	a	0	0	1	0	0	0	0	0
爱国	a	0	0	0.9	0	0	0	0	0
爱国	a	0	0	0.8	0	0	0	0	0
爱国	a	0	0	0.5	0	0	0	0	0

Table 3
Average emotion intensity of "爱国" as adjective.

Word	POS	Surprise	Sorrow	Love	Joy	Hate	Expectance	Anxiety	Anger
爱国	a	0	0	0.643	0.086	0	0	0	0

high coverage. This suggests that we need to further tune down the threshold value. When the threshold value is 0.3, the coverage is 0.984 (20,485/20,814); when the threshold value is 0.4, the coverage is 0.863 (17,961/20,814), which is the first obvious decease of coverage. We can see that the optimal threshold value should be 0.4 after the trade-off between coverage and volatility. That is to say, if a word has an average emotion vector in which the value of each component is under 0.4, we deleted the word. This trial and error process resulted 17,961 words in our final lexicon. We label it as Ren-CECps-SWB 2.0. We use Ren-CECps-SWB 2.0 as the building blocks to measure SWB from Chinese texts in online blogs.

3.3. Weights of eight basic emotions

We used Ren-CECps-SWB 2.0 to measure SWB. There are eight emotion dimensions in Ren-CECps-SWB 2.0. The next question, therefore, is to determine the weight of emotion, where $\omega_j, j = 1, \ldots, 8$. The incorporation of ω_j differentiates this study from the work of Dodds and Danforth [10], who use only one emotion dimension valence from ANEW in their SWB measurement. This is also a difference between our study and the method used in FGNH, in which the SWB measurement considers positivity and negativity with equal weights. To determine parameter ω_j, we applied two different methods, a public questionnaire survey and the Delphi method.

3.3.1. Public questionnaire

The purpose of the survey was to determine the weights of eight emotions for the common Chinese public. We started our questionnaire with an online survey, which was conducted using the free service of SOJUMP.com, a national, popular, professional online survey platform. Most participants in our online survey were young people. The respondents did not represent middle-aged white-collar workers or scholars. To overcome this limitation, we conducted an offline survey using the same measurement instrument. More specifically, we hand-delivered our questionnaires to Central Business District (CBD) white-collar workers, and university professors and college students in Beijing. We collected a total of 427 responses, 286 online and 141 offline. After deleting 25 incomplete and inconsistent responses, we had 402 valid responses (an effective response rate of 94.1%).

In our questionnaire, we adopted a nine-point Likert-scale questionnaire to measure SWB, with −4 indicating "extremely unhappy" and 4 indicating "extremely happy". The questionnaire we employed is presented in Appendix 1. We calculated the average degree of SWB of each basic emotion denoted. We then normalized the scores of positive and negative emotions. Table 5 presents the results.

In Table 5, the first four emotions have positive scores, and the scores of next four emotions are negative. This is finding consistent with our expectation and suggests that the first four emotions can improve psychological well-being, whereas the other four can decrease psychological well-being. Accordingly, we measure positive sentiment using the first four emotions and negative sentiment using the last four emotions.

One could argue that our respondents are not randomly selected because the 286 online survey participants were users at SOJUMP.com, who participated in the survey voluntarily. The

Table 4
Number of words that remained under different thresholds.

POS tag	Intensity: 0.0	Intensity: 0.1	Intensity: 0.2	Intensity: 0.3	Intensity: 0.4	Intensity: 0.5	Intensity: 0.6	Intensity: 0.7	Intensity: 0.8	Intensity: 0.9
a	1622	1620	1611	1577	1528	1391	1150	741	293	99
ad	116	115	113	109	101	82	59	31	7	1
Ag	30	30	30	30	29	27	22	10	1	0
an	15	15	15	14	14	13	11	8	5	0
b	99	99	96	90	83	62	40	22	4	3
c	42	42	41	36	30	22	12	6	2	0
d	504	504	494	458	409	325	234	123	35	8
Dg	4	4	4	4	3	3	3	2	1	0
e	4	4	4	4	4	3	1	0	0	0
f	17	17	16	15	14	13	11	2	0	0
h	6	6	6	6	5	3	2	1	0	0
i	804	803	794	781	745	657	534	320	113	31
j	18	18	18	17	14	13	9	7	3	0
l	170	170	169	159	148	129	94	58	17	3
m	91	91	91	86	81	60	47	16	3	2
n	7692	7685	7560	7150	6527	5345	3764	2051	595	164
Ng	51	51	50	47	41	31	22	13	2	0
nr	80	80	79	72	63	51	32	14	6	4
ns	19	19	19	19	16	16	9	5	1	0
nz	18	18	17	17	16	14	9	6	4	3
o	2	2	2	1	1	1	0	0	0	0
p	47	47	46	42	40	32	24	10	4	0
q	46	46	44	42	39	32	24	14	7	0
r	90	90	85	79	73	57	44	26	6	0
s	10	10	9	7	6	5	3	1	0	0
t	38	38	37	35	32	23	10	5	3	0
Tg	3	3	3	3	3	1	1	0	0	0
u	120	120	119	110	100	80	41	23	4	1
Ug	1	1	0	0	0	0	0	0	0	0
v	8113	8100	7991	7617	6985	5825	4131	2340	671	171
vd	4	4	4	3	3	3	2	0	0	0
Vg	45	45	43	42	40	35	24	12	4	0
vn	748	747	732	701	645	533	396	219	85	21
w	5	5	5	5	5	4	3	2	0	0
y	18	18	17	15	13	12	10	7	5	2
z	122	122	121	114	105	88	59	30	5	2
Total	20,814	20,789	20,485	19,507	17,961	14,991	10,837	6125	1886	515

141 offline surveys were white-collar workers, university professors or students in Beijing. To address this concern, we applied the Delphi method to make sure that our measurement instrument had both internal and external validity.

3.3.2. Delphi method

We invited 20 experts from Beijing Normal University and Chinese Academy of Sciences (CAS) to compare the intensity of positive (negative) emotions in pairs and rate the relative intensity of the emotions. For example, if an expert judges that the SWB contained in "love" is four times stronger than that of "expectance", then he should rate love/expectance as 4. All basic emotions are required to be compared in pairs at the same strength level. For example, the comparison between "strong love" and "low expectance" is not allowed, whereas that between "strong love" and "strong expectance" is accepted. The questionnaire used to solicit the experts' responses is reported in Appendix 2.

The Delhi method was strictly employed in accordance with the guideline suggested by Dalkey and Helmer [5]. The 20 experts completed questionnaires in three rounds. Because the results from the first round and the second round differed significantly, we presented the results of the second assessment to these experts and invited them to make a third assessment. The second and third assessments were basically consistent. We used the third assessment for further processing. Table 6 presents the results of the Delphi analysis.

3.3.3. Final weights of eight basic emotions

Comparing the weights of the questionnaire and the Delphi method, except for the weight of "hate", which has a difference slightly greater than 0.1 (0.136), the differences between the other pairs of emotions from the questionnaire survey and the Delphi method are smaller than 0.1. The responses from common people and experts were highly consistent. To assign weights to different emotion dimensions, we used the average weight of the results of the questionnaire and Delphi method, each with a weight of 50%. The final weight of for each emotion is reported in Table 7.

Table 5
Weights of emotions for the SWB of the public.

Basic emotion	Positive sentiment				Negative sentiment			
	Expectance	Love	Joy	Surprise	Anxiety	Sorrow	Anger	Hate
Weight	0.268	0.339	0.338	0.055	−0.205	−0.255	−0.237	−0.303

Table 6
Weights of emotions: Delphi method.

Basic emotion	Positive sentiment				Negative sentiment			
	Expectance	Love	Joy	Surprise	Anxiety	Sorrow	Anger	Hate
Weight	0.181	0.356	0.339	0.124	−0.105	−0.248	−0.208	−0.439

Table 7
Final weights of emotions.

Basic emotion	Positive sentiment				Negative sentiment			
	Expectance	Love	Joy	Surprise	Anxiety	Sorrow	Anger	Hate
Questionnaire	0.268	0.339	0.338	0.055	−0.205	−0.255	−0.237	−0.303
Delphi method	0.181	0.356	0.339	0.124	−0.105	−0.248	−0.208	−0.439
Final weights	0.2245	0.3475	0.3385	0.0895	−0.155	−0.2515	−0.2225	−0.371

3.4. Number and weights of emotions used in the model

In consideration of computation efficiency, we also tried to reduce the number of emotions used in our model to measure SWB. Ren-CECps-SWB 2.0 has eight emotions. If a subset of the eight emotions could guarantee a reliable measurement result, we could use the subset instead of the entire set of eight emotions. To find out if there is such a subset, we tested all the combinations of the eight basic emotions. The results are shown in Table 8.

We stepwise depleted emotions that had the least weight in the remaining emotions, and we tested the different dimensionality of the emotion vector using Sina.com grassroots blog texts in 2008. The results in Table 9 suggest that five dimensional vectors had the best performance.

Regardless of how many emotions we used, the trend of SWB in a week was always the same. (See the first seven lines in Table 9.) The weekly trend of SWB was stable as long as more than three emotions were used for calculation. The lowest values appear on Monday and Wednesday (with slight differences of less than 0.0003), followed by an increase from Wednesday to Sunday. On Sundays, the SWB of a week has the highest value. However, when we considered the effect of SWB measurement, we found that in applying the five-dimensional-vector, our model can generate the most significant performance. (See the two bottom rows “RANGE” and “SD” in Table 9.) Accordingly, we decided to make m be 5 and keep love, joy, sorrow, anger and hate as the components of the sentiment vector. Based on this result, Eq. (3) can thus be rewritten as:

$$H_i = \sum_{j=1}^{5} \omega_j d_{ij} \tag{6}$$

Based on Ren-CECps-SWB 2.0, Eqs. (1), (2), (4) and (6) form our final SWB measurement model for Chinese texts.

4. Analysis and results

4.1. Data collection

To test our model, we collected panel data and measured the SWB of each time period using our proposed model. We compared the results with historical events to find the patterns of Chinese people’s wellbeing.

As we explained earlier, we selected grassroots blog texts from Sina.com as our data source. Our data sample ranged from January 2008 to December 2013. During the sample period, 63,505 blogs were posted by 316 bloggers. Table 10 reports the descriptive statistics of our data for each year. As shown by the table, the standard deviation (SD) is fairly large, indicating that bloggers are substantively different in terms of activeness.

4.2. Model validation

We find strong evidence to support our SWB measurement model from the Sina.com data. Fig. 2 shows the monthly average SWB across the entire sample period (January 2008 to December 2013).[3] The zeniths and nadirs of the Chinese SWB identified by our model correspond to major national events during the time period, which are summarized in Table 11. This offers solid evidence of the validity of our model.

4.3. Patterns of Chinese public well-being

We applied our SWB measurement model to examine two patterns of Chinese public well-being. First, we explored the pattern of Chinese public SWB within a week. As illustrated by Fig. 3, we found that Wednesday is the unhappiest day in a week. As the first work day of the week, Monday is the second unhappiest day of the week. From Thursday to Friday, happiness gradually increases. As the weekend approaches, Chinese people’s SWB increases immediately and reaches its peak on Sunday. This pattern accords with the status quo of Chinese urban life.

Second, we explored the pattern of Chinese public SWB across the months of the year. Fig. 4 shows the Chinese public’s SWB for each month from 2008 to 2013. During each Spring Festival, the SWB of the Chinese public is relatively high. After the Spring Festival, it begins to fall. October brings another high well-being time period because of the Chinese National Day. We can see that the Spring Festival and National Day are the two most important holidays in influencing Chinese SWB, and March brings the lowest well-being for Chinese people each year. The curve of Chinese well-being also shows that emergent events can significantly affect Chinese people’s well-being (as shown in Fig. 2): Significant social crises decrease Chinese people’s well-being, whereas significant positive events increase Chinese people’s well-being.

Third, to see the Chinese SWB trend, we also crawled the grassroots’ blog texts in 2014, and we obtained 5033 texts from 130 grassroots bloggers. This enabled us to calculate Chinese SWB from 2008 to 2014, as shown in Table 12.

[3] The dates in Fig. 2 indicate that the corresponding events in Table 10 at time point.

Table 8
Weights of emotions in different combinations.

	Emotion	Means	Absolute value	Number of components						
				2	3	4	5	6	7	8
Weights	Hate	−2.753191	2.753191	−1	−1	−0.595984	−0.439053	−0.439053	−0.37	−0.371
	Love	2.574468	2.574468	1	0.50656	0.50656	0.506560	0.381658	0.381658	0.3475
	Joy	2.561702	2.561702		0.49344	0.49344	0.493440	0.371774	0.371774	0.3385
	Sorrow	−2.323404	2.323404			−0.404016	−0.297633	−0.297633	−0.2515	−0.2515
	Anger	−2.161702	2.161702				−0.263314	−0.263314	−0.2225	−0.2225
	Expectance	2.024043	2.024043					0.246568	0.246568	0.2245
	Anxiety	−1.868085	1.868085						−0.155	−0.155
	Surprise	0.417021	0.417021							0.0895

Table 9
Results of using emotion vectors of different dimensionalities.

Days	8-Vector	7-Vector	6-Vector	5-Vector	4-Vector	3-Vector	2-Vector
Monday	0.0498	0.0493	0.0459	0.0787	0.0920	0.1162	0.2037
Tuesday	0.0508	0.0503	0.0468	0.0798	0.0931	0.1169	0.2058
Wednesday	0.0496	0.0491	0.0457	0.0785	0.0921	0.1161	0.2036
Thursday	0.0502	0.0497	0.0462	0.0793	0.0923	0.1164	0.2050
Friday	0.0504	0.0499	0.0465	0.0795	0.0926	0.1170	0.2053
Saturday	0.0510	0.0505	0.0472	0.0805	0.0938	0.1162	0.2048
Sunday	0.0521	0.0515	0.0482	0.0813	0.0946	0.1174	0.2056
Range	0.0025	0.0024	0.0025	0.0028	0.0026	0.0013	0.0022
SD	0.000844	0.000814	0.000858	0.000986	0.000969	0.0005	0.000873

The trend of Chinese SWB in these seven years is shown in Fig. 5. We can see that the Chinese SWB in these 7 years generally increased, but in 2009, there was an obvious deterioration. In 2009, the world witnessed the global financial crisis. China was suffering. This financial crisis negatively affected the Chinese economy and slashed people's well-being. After 2009, however, Chinese SWB trended upward, reaching the highest point in 2014. From 2010 to 2011, the Chinese economy gradually recovered from the negative influence of the financial crisis. After 2012, in the political arena, one of the most serious anti-corruption storms in Chinese history increased the common public's SWB in general.[4]

5. Discussion

5.1. Theoretical implications

First, our research provides a new way to measure SWB using text sentiment analysis following the classic PANAS framework in psychology with multiple emotions in PA and NA. Research in psychology suggests that PA and NA are independent in PANAS. Therefore, to measure SWB, PA and NA should be considered simultaneously. Our study is one of first studies to integrate insights from psychology research to explore the measurement of SWB using sentiment- and text-analysis techniques. Moreover, emotions in PA and NA have different weights in one's SWB evaluation [19]. Thus, these multiple emotions should be assigned different weights to account for their effects on SWB. Based on these ideas, we provide a method to select the optimal emotion set by balancing computation efficiency, measurement adequacy and model parsimony. Our contribution is our use of text-sentiment analysis to extract all the selected emotions embedded in PA and NA. The solid psychological foundation of PANAS ensures that our SWB measurement method is both valid and reliable. Our work differs significantly from extant studies.

Unlike SWB measurement in FGNH, our method involves detailed emotions in PA and NA. FGNH merely considers negativity and positivity. Moreover, FGNH assigns equal weight to negativity and positivity, which is contrary to the accumulative insights from several decades of psychology research that show that different emotions affect people's wellbeing differentially. Thus, the dualistic classification and equal weights of positive feelings and negative feelings for measuring subjective well-being oversimplified the measurement for SWB. It may produce biases and misleading results in wellbeing research and policy making. Based on solid evidence, Wang and his colleagues conclude that FGNH is not a valid measure for well-being [23]. Our study is based on the PANAS framework, whereas Dodds and Danforth's SWB measurement is based on economic utility theory. The most important difference is that Dodds and Danforth use only one emotion dimension valence from ANEW in their SWB measurement, although there are three semantic notations in ANEW, and the other two also affect SWB.

Second, our research provides a unique lexicon, labeled as Ren-CECps-SWB 2.0, for measuring SWB in Chinese text. Although there are many existing Chinese corpuses, such as NTUSD, HowNet, and Ren-CECps, not all of them aim to measure SWB. After comparing it with the current popular corpuses, we selected Ren-CECps as the

Table 10
Descriptive statistics of data.

Year	Total bloggers	Total posts	Mean	SD
2008	278	19,200	69.06475	88.39705
2009	249	14,315	57.48996	102.2461
2010	224	11,102	49.5625	77.12098
2011	197	8058	40.90355	68.70349
2012	179	5797	32.38547	60.9369
2013	141	5032	35.68794	84.29999

[4] New York Times commented, "At the moment, Mr. Xi has the public's support and political momentum". (Source: http://www.nytimes.com/2014/10/18/opinion/crony-communism-in-china.html?_r=0). People's Daily reported, "People's daily inventory Xi Jinping ruling achievements: Iron anti-corruption winning hearts and minds" (http://gd.people.com.cn/n/2014/0812/c123932-21961753.html).

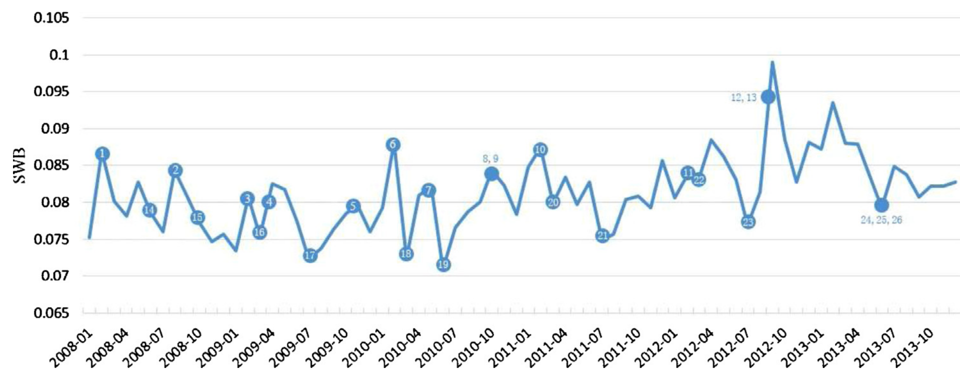

Fig. 2. The subjective well-being of the Chinese public from 2008 to 2013.

Table 11
Significant events in China from 2008 to 2013.

Time periods of low public SWB	Significant events in time periods of low public SWB	Time periods of high public SWB	Significant events in time periods of high public SWB
May–July, 2008	Sichuan Wenchuan earthquake[14]	February 2008	Spring Festival[1]
		August 2008	Beijing Olympic Games[2]
October, 2008	Chinese milk scandal[15]	February 2009	Spring Festival[3]
March, 2009	After Spring Festival[16]	April 2009	Boao Forum for Asia[4]
July, 2009	Xinjiang "7.5" terrorist event[17]	October 2009	The 60th national day of China[5]
March, 2010	After Spring Festival[18]	February 2010	Spring Festival[6]
June, 2010	CPI began to rise sharply[19]	May 2010	Expo 2010 Shanghai China[7]
March, 2011	After Spring Festival[20]	October 2010	The National Day holiday of China[8]; China launched Chang'e II successfully[9]
July, 2011	Red Cross scandal; Wenzhou train collision (40 deaths)[21]	February 2011	Spring Festival[10]
March, 2012	After Spring Festival[22]	February 2012	Spring Festival[11]
July, 2012	7.21 Beijing rainstorm (79 deaths)[23]	September 2012	China issued the "Statement of the Government of The People's Republic of China on the Territorial Sea Baselines for Diaoyu Dao and Its Affiliated Islands"[12]; Boxilai was expelled from the Party[13]
June, 2013	Xiamen bus fire (47 deaths)[24]; Edward Snowden revealed PRISM[25]; June 2013 Shanshan riots (24 deaths)[26]		

The number in Table 11 is the number of events in Fig. 2.

basic corpus for measuring subjective wellbeing. Making tradeoff between the volatility of results and the coverage of the corpus, we constructed Ren-CECps-SWB 2.0 with 17,961 words. We posted Ren-CECps-SWB 2.0 on the website http://wptrying.sinaapp.com/; anyone in the world interested in studying Chinese well-being can employ it to process information embedded in Chinese text.

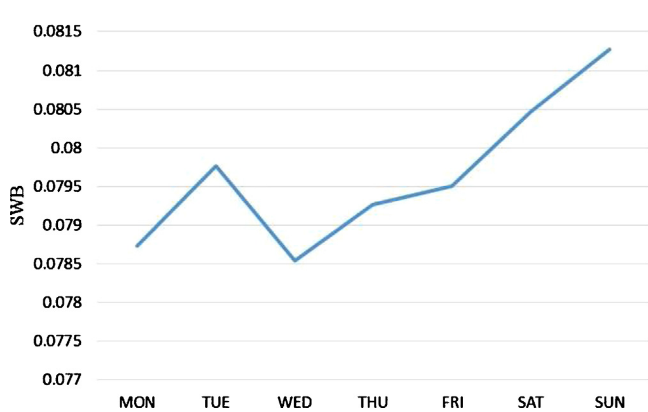

Fig. 3. Change in Chinese SWB in a week.

Third, based on Ren-CECps-SWB 2.0, we provide a unique automatic method that is different from the traditional self-reported method for measuring Chinese SWB based on UGC in Chinese text. Based on the SWB measurement model, we found the unique patterns of Chinese well-being on weekly and yearly bases. Specifically, we analyzed the weights of different emotions and which emotions are most appropriate to include in our measurement model. The robustness and reliability of our method was validated using grassroots blogs posted from 2008 to 2013 on Sina.com. As far as we know, no prior research as of today has studied Chinese well-being using sentiment- and text-analysis techniques based on data from UGC.

Table 12
Descriptive statistics of SWB per year.

Year	Max value	Min value	Mean	SD
2008	0.086864941	0.074677207	0.079213585	0.003869665
2009	0.082499316	0.072279045	0.077316445	0.003355965
2010	0.088088528	0.07095652	0.079522059	0.004754174
2011	0.087478833	0.074941794	0.081225975	0.003776581
2012	0.098974049	0.076916831	0.085182771	0.00555909
2013	0.093466147	0.079318709	0.084637331	0.003908459
2014	0.100627066	0.070316726	0.089311444	0.011287174

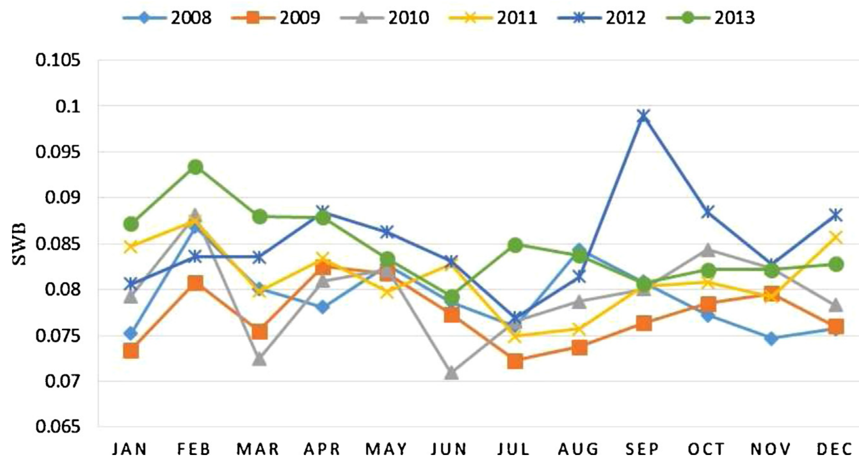

Fig. 4. Change in Chinese SWB monthly.

5.2. Practical implications

First, to make the Chinese public happy, the Chinese government should take actions to decrease the social emotion of hate. Although there are four positive emotions (love, joy, expectance, and surprise) and four negative emotions (hate, sorrow, anger, anxiety) in our model, the relative weights of these eight emotions for Chinese suggest that "hate" has the greatest influence of all the eight emotions to affect Chinese people's SWB, and the effect is negative. The Chinese government should check which factors may cause the social emotion of "hate". In the National People's Congress (NPC) and Chinese People's Political Consultative Conference (CPPCC), according to the hot public concern survey hosted by the official Chinese People.com website from 2002 to 2014 (http://npc.people.com.cn/GB/28320/374787/index.html), there were 12 times that anti-corruption was among the top three greatest concerns; six times, it ranked number 1. These surveys show that corruption is the most important influencing factor that causes the Chinese social emotion of "hate". Therefore, we suggest the Chinese government do more to reduce and prevent corruption to improve the public's SWB.

Second, to make Chinese public happy, the Chinese government should take actions to decrease the occurrence of shocking emergent events. Table 11 suggests that the Chinese public SWB is consistent with significant events. Positive events can lift the Chinese public SWB, whereas negative emergent events depress Chinese people's wellbeing quickly. Now, China is in its economic and social transformation period, and there are many complicated comprehensive social issues. These social issues could cause a relatively high risk of shocking emergent events and hidden risks for Chinese people's SWB. Based on the results of this study, the Chinese government should constantly improve its management qualification to resolve these potential risks.

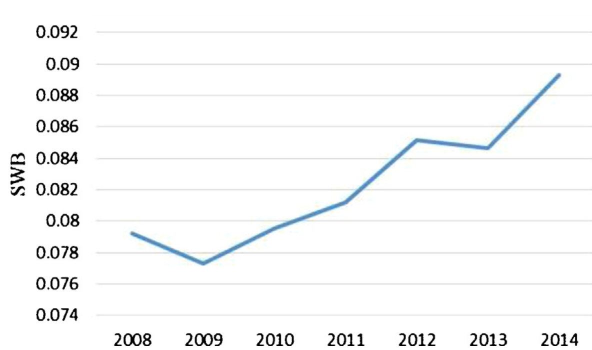

Fig. 5. The trend of Chinese SWB from 2008 to 2014.

Third, we find the patterns of Chinese people's SWB changes on weekly and monthly bases. These results will be helpful not only for the government but also for firms to adjust and formulate new strategies to improve employees' work efficiency and productivity. For example, Wednesday and Monday are the two unhappiest days in the week. To boost employees' mood, firms can arrange special activities on these two days to decrease the negative emotions of employees and thus improve workers' productivity. We found that March is the unhappiest month in a year after a long holiday in the Chinese Lunar New Year. Some employees want to seek new job opportunities at this time. Firms can provide appropriate incentives for their employees at this time to keep and motivate them.

6. Conclusion

Integrating PANAS with psychology and text sentiment-analysis techniques, we constructed a measurement model for SWB with multiple weighted emotions in PA and NA. We built a specific lexicon, Ren-CECps-SWB 2.0, to measure Chinese people's SWB. We included five emotions (hate, love, joy, sorrow and anger) in our model. We validated our method using data collected from Sina.com blog texts. However, our research has a few shortcomings. The contribution of this research should thus be read in light of these limitations. First, we did not analyze the influence of grammar, context and figures of speech in our analysis, which may lead to some bias in the results. Moreover, factors other than volatility and coverage may also influence the threshold of word filtering. We did not include these other potential factors in our analysis. Third, our method was not applied to short text. Further research on short text is therefore needed to validate our model and the results of this study.

The limitations of our research indicate a future research direction. First, Simple Chinese Word Segmentation (SCWS), which we used in this study, cannot be used to analyze grammar structures, different contexts, or varied figures of speech in Chinese. A better text analyzer could improve the accuracy of the results. Relatedly, our measurement model was built on the basis of vernacular Chinese. Future research can focus on, e.g., classic Chinese and ancient poetry to draw additional insights. Third, future studies may pay more attention to individual and public sentiment implied by short texts, such as Chinese Microblog.

Acknowledgements

This research was supported by Major State Basic Research Development Program of China (973 Program; Nos. 2012CB315805 and 2013CB329604), the National Natural Science Foundation of China (Project No. 71231002), the Research Fund for the Doctoral Program of Higher Education of China (Program No. 20120005110015), the Co-Research Fund of the Ministry of Education of P.R. China, and China Mobile (Program No. MCM20123021). In addition, the authors are very thankful to Hang Ye, Xu Liu and Ning Liu at Beijing University of Posts and Telecommunications for their initial data processing and analysis.

Appendix 1. Questionnaire survey for the public

Please make judgments on the relationships between eight basic emotions and affective well-being according to your daily life experiences.

Related concepts:

- **Affective well-being:** An overall individual emotional experience of events in daily life and at work.
- **Expectance:** The thought or belief that something will happen or that somebody will come.
- **Joy:** Great satisfaction, enjoyment, or joy.
- **Love:** Feelings of love, desire, or strong pleasure for something or somebody.
- **Surprise:** A brief emotional state experienced as the result of an unexpected significant event.
- **Anxiety:** An uncomfortable feeling in the mind usually caused by the fear or expectation that something bad will happen.
- **Sorrow:** A feeling of sadness or distress caused by, e.g., loss or regret.
- **Anger:** An emotional response related to one's psychological interpretation of having been threatened.
- **Hate:** Strong dislike; hatred.

A.1. Affective well-being contributions scale

Affective well-being									
	Extremely happy	Very happy	Happy	Slightly happy	Does not matter	Slightly unhappy	Unhappy	Very unhappy	Extremely unhappy
Expectance									
Love									
Joy									
Surprise									
Anxiety									
Sorrow									
Anger									
Hate									

Appendix 2. Questionnaire survey for experts

B.1. Related concepts: the same those in Appendix 1

- Guidelines:

The above eight basic emotions reflect positive sentiment and negative sentiment to some extent, but they indicate sentiment that varies in intensity. Please use your professional expertise to compare the intensity of the positive (negative) sentiment indicated by each pair of basic emotions. For example, if you decide the positive sentiment of "love" is four times stronger than that of "expectance", then you should rate love/expectance = 4. The levels of emotions themselves are not considered. That is, all basic emotions are compared at the same level. Therefore, the comparison between "love very much" and "a little expectance" is not allowed.

B.2. Positive sentiment judgment

Expectance/Joy=
Expectance/Love=
Expectance/Surprise=
Joy/Love=
Joy/Surprise=
Love/Surprise=

B.3. Negative sentiment judgment

Anxiety/Sorrow=
Anxiety/Anger=
Anxiety/Hate=
Sorrow/Anger=
Sorrow/Hate=
Anger/Hate=

References

[1] A.S. Abrahams, J. Jiao, G.A. Wang, W. Fan, Vehicle defect discovery from social media, Decis. Support Syst. 54 (1), 2012, pp. 87–97.
[2] A.S. Abrahams, J. Jiao, W. Fan, G.A. Wang, Z. Zhang, What is buzzing in the blizzard of buzz: automotive component isolation in social media postings, Decis. Support Syst. 55 (4), 2013, pp. 871–882.
[3] J. Bollen, B. Gonçalves, G. Ruan, et al., Happiness is assortative in online social networks, Artif. Life 17 (3), 2011, pp. 237–251.
[4] B. Brülde, Happiness and Suffering: Concept, Method, Explanation,(Lyckaochlidande: begrepp, metod, förklaring), Studentlitteratur, Lund, 2007.
[5] N. Dalkey, O. Helmer, An experimental application of the Delphi method to the use of experts, Manag. Sci. 9 (3), 1963, pp. 458–467.
[6] E. Diener, Subjective well-being, Psychol. Bull. 95, 1984, pp. 542–575.
[7] E. Diener, Subjective well-being: the science of happiness and a proposal for a national index, Am. Psychol. 55 (1), 2000, p. 34.
[8] E. Diener, Assessing Well-being: The Collected Works, Social Indicators Research Series, (vol. 39), Springer, The Netherlands, 2009.
[9] E.D. Diener, R.A. Emmons, R.J. Larsen, et al., The satisfaction with life scale, J. Pers. Assess. 49 (1), 1985, pp. 71–75.
[10] P.S. Dodds, C.M. Danforth, Measuring the happiness of large-scale written expression: songs, blogs, and presidents, J. Happiness Stud. 11 (4), 2010, pp. 441–456.
[11] P.S. Dodds, K.D. Harris, I.M. Kloumann, et al., Temporal patterns of happiness and information in a global social network: Hedonometrics and Twitter, PLoS ONE 6 (12), 2011, p. e26752.
[12] Facebook Data Team, Continuing Our Study of Happiness, 2010 http://www.facebook.com/notes/facebook-data-team/continuing-our-study-of-happiness/375901788858.
[13] Facebook Data Team, Gross National Happiness, 2011 http://apps.facebook.com/gnh_index/.
[14] W. Fan, M.D. Gordon, Unveiling the power of social media analytics, Commun. ACM 2014, forthcoming, Preprint available at: https://www.researchgate.net/publication/259148570_Unveiling_the_Power_of_Social_Media_Analytics.
[15] P.D. Killworth, H.R. Bernard, Informant accuracy in social network data, Hum. Organ. 35 (3), 1976, pp. 269–286.
[16] C. Martinelli, S.W. Parker, Deception and misreporting in a social program, J. Eur. Econ. Assoc. 7 (4), 2009, pp. 886–908.
[17] M. McGillivray, Human Well-being: Concept and Measurement, Palgrave Macmillan, 2007.
[18] W. Pavot, E. Diener, Review of the satisfaction with life scale, Psychol. Assess. 5 (2), 1993, p. 164.
[19] L. Qiu, X. Zheng, Y.F. Wang, Revision of the positive affect and negative affect scale, Chin. J. Appl. Psychol. 14 (3), 2008, pp. 249–254.
[20] C. Quan, F. Ren, Automatic annotation of word emotion in sentences based on Ren-CECps, in: Proceedings of LREC, 2010.
[21] S.C. Schmukle, B. Egloff, L.R. Burns, The relationship between positive and negative affect in the positive and negative affect schedule, J. Res. Pers. 36 (5), 2002, pp. 463–475.
[22] G.A. Wang, J. Jiao, A.S. Abrahams, W. Fan, Z. Zhang, ExpertRank: a topic-aware expert finding algorithm for online knowledge communities, Decis. Support Syst. 54 (3), 2013, pp. 1442–1451.
[23] N. Wang, M. Kosinski, D.J. Stillwell, et al., Can well-being be measured using Facebook status updates? Validation of Facebook's Gross National Happiness Index Soc. Indic. Res. 115 (1), 2014, pp. 483–491.
[24] D. Watson, L.A. Clark, A. Tellegen, Development and validation of brief measures of positive and negative affect: the PANAS scales, J. Pers. Soc. Psychol. 54 (6), 1988, p. 1063.
[25] D. Zeng, H. Chen, R. Lusch, S.-H. Li, Social media analytics and intelligence, IEEE Intell. Syst. 25, 2010, pp. 13–16.

Jiayin Qi is a professor of School of Economics and Management, Beijing University of Posts and Telecommunications (BUPT). She is also the director of Center of Network Information Management and Service, BUPT. Her research interests are social network analysis, user generated contents management and emergent event management, etc.

Xiangling Fu is an associate professor in School of Software Engineering of Beijing University of Posts and Telecommunications. She got Ph.D. degree from Peking University in 2004. Her research is focused on social network analysis, data processing in Chinese text and so on.

Ge Zhu is a graduate student of School of Economics and Management, Beijing University of Posts and Telecommunications (BUPT). His research interests are social network analysis, user generated contents management and customer relationship management, etc.