



Full length article

An analysis of sleep complaints on Sina Weibo

Xianyun Tian^{*}, Guang Yu, Fang He

School of Management, Harbin Institute of Technology, Harbin, PR China



ARTICLE INFO

Article history:

Received 24 February 2016

Accepted 4 April 2016

Keyword:

Social media

Social networking

Insomnia

Mental health

ABSTRACT

This study explores sleep complaints on Sina Weibo to gain insights into social networking about mental health. A random sample ($n = 1000$) of insomnia-related postings was coded and analyzed to investigate the themes and symptoms expressed in postings. The most common theme mentioned is the disclosure of insomnia. The difficulty with sleep initialization is the most common symptom revealed in postings. Besides, the prevalence of insomnia is higher in certain areas and the people who may suffer from insomnia tend to be active from midnight to noon. Our findings can be used to detect and provide help for those Weibo users who may suffer from insomnia.

© 2016 Elsevier Ltd. All rights reserved.

1. Introduction

Like Twitter, Weibo is a popular social media for communication and information sharing in China with 500 million users and 100 million messages posted per day. It is the most influential social media in China which directly influences a third of China's population. Weibo allows people to post updates that are less than 140 characters. These postings can then be reviewed by those individuals who follow them.

Weibo's openness enables people to connect with many acquaintances and strangers, and this makes it a very popular tool for people to express themselves and exchange with each other. Postings on Weibo cover many different themes, including news, emotional messages, and comments about events and so on. Weibo is a good data source that is rich enough to conduct various types of research. Moreover, because of the anonymity, people feel free and safe to talk about some private issues, such as physical illness and mental disorder (Guo & Goh, 2014; Han & Wang, 2015; Liang et al., 2015; Wang & Liu, 2015).

Insomnia is the most common sleep disorder, with high impact on individual and society. It is estimated that 10%–15% of the adult population experiences chronic insomnia (Roth, 2001). The third edition of the International Criteria for Sleep Disorders (ICSD-3), proposed by the American Academy of Sleep Medicine, defines insomnia as “a repeated difficulty with sleep initiation, duration, consolidation, or quality that occurs despite adequate opportunity

and circumstances for sleep, and results in some form of daytime impairment” (Sateia, 2014). It is known that poor socio-occupational functioning and cognitive impairments are closely associated with insomnia (Fortier-Brochu, Beaulieu-Bonneau, Ivers, & Morin, 2012). Individuals with insomnia tend to perform poorly on tasks regarding working memory, episodic memory and problem solving. Additionally, insomnia can lead to more motor vehicle accidents and a higher risk of psychiatric disorders (Breslau, Roth, Rosenthal, & Andreski, 1996; Léger, Guilleminault, Bader, Lévy, & Paillard, 2002; Ohayon & Smirne, 2002). Despite the high prevalence rates and severe consequences, insomnia is underrecognized, underdiagnosed, and undertreated (Roth, 2001). The key to these problems is the dissemination of information regarding sleep routine and sleep therapy. Social media is a perfect candidate for disseminating such information for its widespread use and low cost.

Emerging research has shown that social media are being used as tools to disclose various personal health issues. Specifically, several studies have been conducted to understand better the information discussing mental disorder or physical illness, such as depression, self-harm and suicidality. For instance, the association between tweets involving depression on Twitter and offline diagnosis had been established (Daine et al., 2013; Moreno et al., 2012). Their studies found that individuals who had been diagnosed with MDD (Major Depressive Disorder) tended to have lowered social activity, higher negative emotion and self-attentional focus. Additionally, the postings about feeling depressed echoed with the depression symptoms they self-reported on a depression screening tool. Previous research also suggested the feasibility of using social media data as a method for evaluating and detecting HIV risk

^{*} Corresponding author. 92 Xidazhi St, Nangang, Harbin, Heilongjiang, 150001, PR China.

E-mail address: uncertainorcertain@gmail.com (X. Tian).

behaviors and outcomes (Guo & Goh, 2014; Young, Rivers, & Lewis, 2014). For instance, in a study concerning PHA support group on Sina Weibo, researchers refined a classical theory by breaking down two broad categories of informational messages and socio-emotional messages into more specific components (Guo & Goh, 2014). However, the research focusing on establishing an association between insomnia and social media is still in its infancy. As far as we know, only a content analysis was conducted to analyze the Tweets that contained the word “insomnia” and it found that these tweets contained more negative health information, strongly suggesting that the users were disclosing their sleep disorder (Jamison-Powell, Linehan, Daley, Garbett, & Lawson, 2012).

There is still more to learn about the sleep complaints on social media. Keyword-based methods have been developed to retrieve tweets that contain “insomnia” (Broniatowski, Paul, & Dredze, 2013; Chew & Eysenbach, 2010; Denecke et al., 2013; Jamison-Powell et al., 2012); however, the tweets collected in this way are not all complaints. In other words, new algorithms should be developed to retrieve the postings that are complaining of sleep disorder. A thematic analysis of the tweet has been given (Jamison-Powell et al., 2012); however, the demographic characteristics of the individuals who complain of sleep disorder are not well studied. Besides, no attention has been paid to Sina Weibo, through which we can learn about Chinese’s levels of insomnia at very low cost.

This study is designed to evaluate the feasibility of using social media to understand whether and how sleep complaints are discussed online and the feasibility of using their data to inform insomnia detection and prevention. Specifically, this study seek to determine 1) whether sleep complaints can be extracted from social media data, 2) the content of these extracted data, 3) the demographic characters of the individuals who complain of the poor sleep, 4) the diurnal activity pattern of the users who complain of insomnia.

2. Method

2.1. Data collection

From May 17 to August 21, 2015, approximately 394,000,000 posts were collected from our Sina Weibo users of interest through its API. These users have been randomly sampled from Sina Weibo, and the size of this sample is 1 million. For every user, his/her 1000 latest posts were obtained, and if the number of a user's all posts was fewer than 1,000, then we crawled all of his/her postings. Varieties of metadata are available along with posting text including the demographic characteristics, the post's source and time the posting was posted.

Since it is hard to obtain online consent from everyone, many measures are adopted to protect the privacy of individuals studied in this research. For example, any information that can be used to trace whom the information belongs to is not published.

2.2. Selection of postings

The flow chart of posting selection is presented in Fig. 1. Of the 394,025,295 postings initially collected through the API, about 160,000 postings were included in the further analysis.

In this paper, we consider the sleep complaints detection as a supervised classification task and a text classifier was developed to complete this task. A labeling data set was required to train the classifier so that it can identify the complaint. However, there are no publicly available datasets that can meet our requirements. Thus, 20,000 postings were randomly selected and labeled by our researchers. Data labeling was undertaken independently by two

PhDs with extended mental health experiences, and any disagreement was resolved by consensus. There are only positive postings which account for 0.37% of the labeled data. The percent of the positive postings means that this is an imbalanced classification problem, and the AUC (Hanley & McNeil, 1982) should be selected as criteria to measure the performances of our classifiers. Ten-folded cross-validation was also adopted to evaluate the performances of our classification algorithms, including Logistic Regression (Kleinbaum & Klein, 2010), Support Vector Machine (Suykens & Vandewalle, 1999) and Random Forests (Breiman, 2001).

The experiment results show that among the classifiers, SVM algorithm is the best choice for sleep complaints detection, and the value of its AUC is 0.99. Consequently, the SVM classifier was used by us to classify the postings of users'.

The prediction of the SVM classifier is the probability of a posting belonging to the positive category. We can select a value as the threshold to decide the number of postings falling into positive. In this study, the labeled dataset shows that there are about 0.37% of postings that are sleep complaints. To gain balance between precision and recall and maximize the precision of our results, we selected top 0.1% percent of postings ranked according to the outcome of SVM in a descending order. In the end, approximately 160,000 postings are included in further analysis.

2.3. Themes

1000 postings were randomly sampled to identify themes contained in postings that complain of sleep disorder. Two members of our research 1000 postings were randomly sampled to identify themes contained in postings that complain of sleep disorder. Two members of our research team, each with experiences in mental health, were trained to scan the 1000 postings to determine the most common themes of the postings and create a codebook. Themes were defined as topics that occur or reoccur (Ryan & Bernard, 2003). Postings were coded as follows: 1) User just discloses sleep disorder; 2) User disclose sleep disorder and express negative emotions; 3) User discloses sleep disorder and gives the causes of poor sleep by himself/herself; 4) User discloses sleep disorder and seeks help; 5) User complains of poor sleep and goes on something to deal with it.

Our two trained researchers coded the 1000 postings using the codebook. They discussed each of the postings and came to an agreement on the final assigned code. A third member with extensive mental health experiences was introduced to code a sample of 200 postings randomly selected from the 1000 postings. Inter-coder reliability for each theme was as follows: 1) User just discloses sleep disorder: percent agreement 89.5%, kappa 0.769; 2) User disclose sleep disorder and express negative emotions: percent agreement 90%, kappa 0.781; 3) User discloses sleep disorder and gives the causes of poor sleep by himself/herself: percent agreement 96%, kappa 0.858; 4) User discloses sleep disorder and seeks help: percent agreement 97.5%, kappa 0.801; 5) User complains of poor sleep and goes on something to deal with it: percent agreement 98.5%, kappa 0.834.

2.4. Presence of insomnia symptoms

The randomly selected 1000 postings where users complaint of sleep disorder were further coded to learn what kinds of symptoms of insomnia were expressed. According to the third edition of the International Criteria for Sleep Disorders (ICSD-3), the symptoms of insomnia include a repeated difficulty with sleep initiation, duration, consolidation, or quality that occurs despite adequate opportunity and circumstances for sleep, and results in some form of

Download English Version:

<https://daneshyari.com/en/article/6836812>

Download Persian Version:

<https://daneshyari.com/article/6836812>

[Daneshyari.com](https://daneshyari.com)