Applied Geography 60 (2015) 217–223

Contents lists available at ScienceDirect

Applied Geography

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

GIS analysis of depression among Twitter users

Wei Yang, Lan Mu*

Department of Geography, University of Georgia, United States

ARTICLE INFO

Article history:
Available online 20 November 2014

Keywords:
Depression
Tweets
Clustering
GIS
Social media

ABSTRACT

Depression is a common chronic disorder. It often goes undetected due to limited diagnosis methods and brings serious results to public and personal health. Former research detected geographic pattern for depression using questionnaires or self-reported measures of mental health, this may induce same-source bias. Recent studies use social media for depression detection but none of them examines the geographic patterns. In this paper, we apply GIS methods to social media data to provide new perspectives for public health research. We design a procedure to automatically detect depressed users in Twitter and analyze their spatial patterns using GIS technology. This method can improve diagnosis techniques for depression. It is faster at collecting data and more promptly at analyzing and providing results. Also, this method can be expanded to detect other major events in real-time, such as disease outbreaks and earthquakes.

Introduction[1]

Depression is a common chronic disorder with adverse effects for well-being and daily functioning, and is associated with high suicide rates (Barlow & Durand, 2011). Major depressive disorder (MDD) (Barlow & Durand, 2005) is the most common type. The centers for disease control and prevention has reported that an estimated 3.4 percent U.S. adults report MDD (MMWR, 2010).

Depression often goes undetected due to the absence of reliable laboratory test, and therefore new methodology for its diagnosis is needed. Social media provides a platform for people to share their activities and feelings. Paul and Dredze (2011) highlighted a slogan "You are what your tweet" and suggested that Twitter had broad applicability for public health research. Twitter is a popular social media website that allows users to post tweets with a maximum length of 140 characters. It has a large user base and provides geo-tagged tweets to researchers, thus the collected data are not restricted to geographic, time, and mobility constraints. Recent work has shown that it is feasible to use Twitter for MDD detection (Park, Cha, & Cha, 2012) (De Choudhury, Counts, & Horvitz, 2013).

Former research showed that depression exhibited geographic clusters. Neighborhood racial composition (Mair, Diez Roux, Osypuk, et al., 2010), household income and education (Akhtar-Danesh & Landeen, 2007), and neighborhood family structures (Mair, Diez Roux, & Morenoff, 2010) are all related to depression. Their limitation lies in traditional data collection that conducts an interview or uses self-reported measures of mental health. Most people are unaware of the symptoms when they have depression; people may also conceal facts for privacy reasons. This may result in the same-source bias and spurious associations (Mair et al., 2009).

We contribute to the study of depression using social media in two aspects. First, we design a procedure to automatically detect MDD users before their estimated onset using Twitter. Because tweets are short and noisy texts and a term can have different meanings, we not only use the key word "depress" and its variations to filter tweets, but also leverage an advanced text mining algorithm, namely *non-negative matrix factorization* (NMF) (Lee & Seung, 1999). We evaluate our procedure on real tweets. This procedure offers a novel and automated way to diagnose depression. It reduces human labor for MDD screening and facilitates personalized interventions for particular groups of population. Second, we further detect the spatial distribution of these MDD users on Twitter and examine the association with socioeconomic status (SES) (Krieger, Williams & Moss, 1997) at county level.

* Corresponding author. 210 Field St., Rm103, GG Building, Department of Geography, University of Georgia, Athens, GA 30602, United States. Tel.: +1 706 542 1058; fax: +1 706 542 2388.

E-mail addresses: weiyang@uga.edu (W. Yang), mulan@uga.edu (L. Mu).

[1] Abbreviations in this paper: GIS, Geographic Information System; MDD, Major Depressive Disorder; NMF, Non-negative Matrix Factorization; SES, Socioeconomic Status; API, Application Programming Interface; DSM-IV, Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition; TP, True Positive; FP, False Positive; FN, False Negative; MSA, Metropolitan Statistical Areas.

http://dx.doi.org/10.1016/j.apgeog.2014.10.016

Using Twitter for depression study

There were some papers using Twitter to predict depression. Park et al. (2012) applied sentiment analysis on tweets and showed that Twitter provided meaningful data for clinical studies on depression. De Choudhury et al. (2013) from Microsoft Research compared tweet text of depressed Twitter users to those of the normal users and highlighted the potential of Twitter as a tool for predicting MDD. Harman, Coppersmith, and Dredze (2014) pointed out that although Twitter users are not a representative sample of the entire population suffering from MDD, individual level and population level analysis can still be made because of the diverse set of quantifiable signals related to MDD in Twitter. A recent survey revealed that 26 percent of the online U.S. adults discussed their health information online, and 42 percent of them use social media to post or seek information about health conditions (GE Healthcare, 2012). Rudd et al. (2006) showed that signs identified by the American Association of Suicidology were frequently included on websites. Among all the signs, the largest individual category for psychological symptoms is about depression and anxiety states. Additionally, they found that young adults like to text to share their feelings.

Geographic analysis of health issues using Twitter

There were very few papers on geographic analysis of public health issues using social media data. Lee, Agrawal, and Choudhary (2013) examined influenza spread in the U.S. by measuring the weighted percentage of tweet volumes mentioning “flu” at state level. The limitations were that they only used the key word “flu” to extract tweets and only considered user profiles that have valid U.S. state information in their home location field.

Ghosh and Guha (2013) compared spatial clusters of obesity-related tweets with the distribution of McDonald's restaurants. They didn't consider that the frequency of posting tweets is different for different people. Using tweets as units instead of Twitter users to explain obesity distribution is not meaningful. For example, someone who posting many tweets with obesity theme at one location doesn't represent there is a higher incident there.

Morstatter, Pfeffer, Liu, and Carley (2013) showed that when geographic bounding boxes were used for Twitter data collection, the collected data were almost the complete set of geo-tagged tweets and thus can be trusted for analytical purposes. Also, text analysis was most accurate when data were downloaded from the Twitter streaming APIs (application programming interfaces).

Table 1
Categories A–E, labels and codes 1–23 for tweets. Categories A and B are expressing one's own depressed feelings or delivering depression information. Tweets belong to categories C, D, E are noise in our study.

A. about my depressed feelings	
1	depressed feeling
2	reason for depression
3	my own remedy for depression
4	treatment on one's depression
5	change of depression feelings
6	pattern of depression
7	meeting doctors
8	take medication
9	confess diagnosis
10	not depressed
B. delivery of depression info	
11	my symptoms
12	Info w/o URL for depression
13	seek depression info
C. sharing thoughts related to depression	
14	my perception of depression
15	attitude towards depression
16	comments to encourage others with depression
D. other topics	
17	various usage of the word depress
18	investigating depression
19	pet depression
20	will be depressed if …
21	meaning is not clear
E. about other's depressed feelings	
22	stories of others on TV, article..
23	tweets about friends' depressed feelings

Material and method

Pilot study

Because a term can have different meanings, we conduct a pilot study to explore the different expressions in tweets related to depression. We use two kinds of Twitter APIs for downloading data: Twitter Streaming APIs give low latency access to the global stream of Twitter data. Twitter REST APIs provide interfaces for most of Twitter's online functionality. We only keep tweets written in English and posted in the U.S. from 2013/09/05 to 2014/03/05, and remove all the others. First, we filter tweets by the key word “depress” or its variations to select tweets related to depression. Then, we write a computer program to randomly select a uniform sample with 0.1% rate (402 tweets) for content analysis.

We follow the classification scheme of tweets related to depression proposed by Park et al. (2012), and manually label the tweets according to their meaning and purpose. Similar labels are placed into the same category in a hierarchical way (Table 1). Each label has a code. For example, “*don't want to be in this hospital anymore, I'm depressed.*” is labeled as “reason for depression” in category A – about my depressed feeling.

Fig. 1 shows the number of tweets under different categories and labels. Among the 402 randomly selected tweets, the word “depress” is most frequently used to express one's own depressed feelings (70 percent of the sample). Among them, 146 tweets express depression directly, 119 tweets even give detail reasons for being depressed, and three tweets describe the change of depressed feelings. Some of them even post private information about their own remedy or diagnosis for depression, and pattern of depression. Eight tweets in category B deliver information regarding depression. Among them, seven tweets contain links leading to a song or a video for fighting against depression, or some medical articles and causes for depression. One tweet mentions the user's own depression symptoms.

Text mining for depression

In this section, we introduce a text mining technique we employed to select MDD candidates more accurately. Because a term can usually have different meanings, it is not sufficient to use the keyword “depress” or its variations to select tweets related to depression. Therefore, we apply a data mining technique, called *non-negative matrix factorization* (NMF), to differentiate the word context associated with depression from those not related to depression.

The basic idea of matrix factorization is that each tweet can be represented as a high-dimensional vector in the space of words, and such high-dimensional data can often be described approximately in a latent subspace with much lower dimensions (Deerwester, Dumais, Landauer, Furnas, & Harshman, 1990). We can interpret the dimensions of the latent subspace as a recurring pattern of word contexts. NMF is a special type of matrix factorization that applies to nonnegative data (Lee & Seung, 1999). It

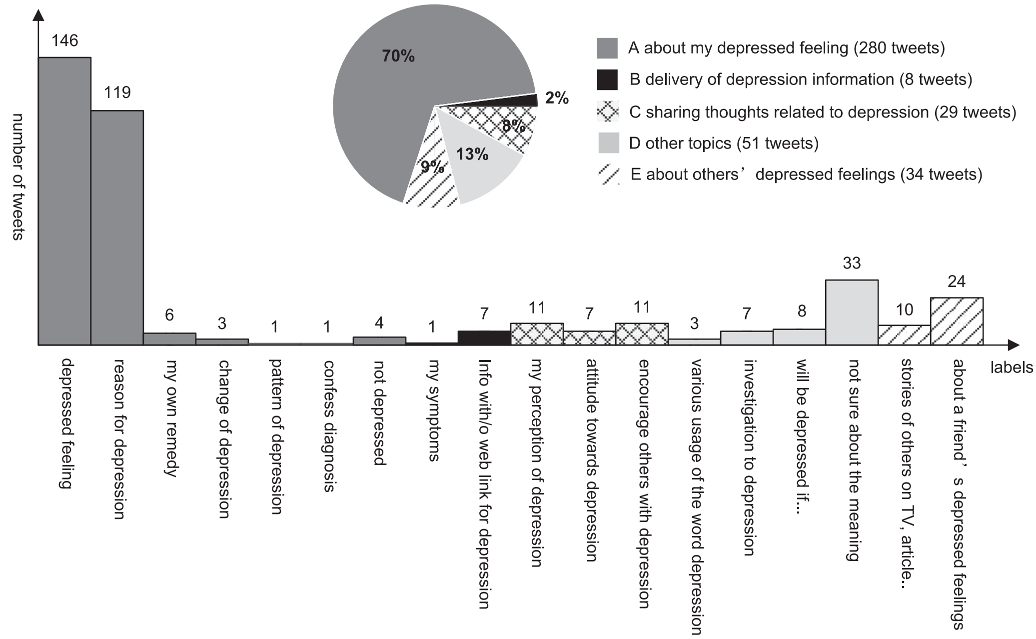


Fig. 1. Tweets counts for different categories and labels. We use different colors and textures to represent five different categories A–E. The pie chart shows the proportion of sample of tweets belonging to each category. The legend lists the number of sample tweets corresponding to each category. The bar chart shows the number of tweets under different labels.

often drives better interpretation of the latent dimensions and is first applied to document clustering by Xu, Liu and Gong (2003).

First, we transform the tweets into a word-document matrix (Manning, Raghavan, & Schütze, 2008). A common assumption is that each tweet (document) is composed of a bag of words and the sequence of words in a document can be ignored. Thus, we can represent a document as a vector of word counts for each word in a vocabulary V. We consider a fixed vocabulary containing m unique words, that is, $|V| = m$. We arrange n documents represented as vectors into an $m \times n$ matrix A, as shown in Fig. 2.

In NMF, given a word-document matrix A as above and a positive integer k, the goal is to approximate A by the product of two nonnegative matrices, W and H:

$$A \approx WH \tag{1}$$

Here W is an $m \times k$ matrix and H is a $k \times n$ matrix. We show an intuitive illustration of NMF in Fig. 3.

In Equation (1), each column of A, denoted as $\mathbf{a}_i$, is approximated by a weighted sum of the columns of W, that is,

$$\mathbf{a}_i \approx W \times \mathbf{h}_i \tag{2}$$

The weights are contained in the i-th column of H, denoted as $\mathbf{h}_i$.

We can interpret each column of W as a vector of frequencies for all the words in V, which we call a word context. The k word contexts define a k-dimensional latent subspace where we can describe the documents, and $\mathbf{h}_i$ contains the coordinates for the i-th document $\mathbf{a}_i$ in the latent subspace. We can interpret $\mathbf{h}_i$ as the proportions corresponding to the k word contexts that constitute $\mathbf{a}_i$. Hence, $\mathbf{a}_i$ can be assigned to the word context with the largest proportion value.

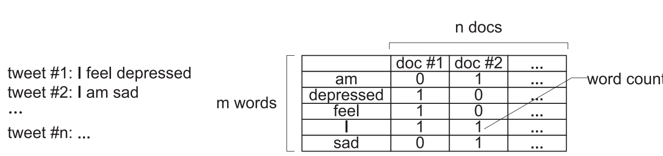


Fig. 2. Example of a word-document matrix. We use two tweets to illustrate. In the matrix, each column represents a document. Each row represents a word. The entry in the matrix represents word count.

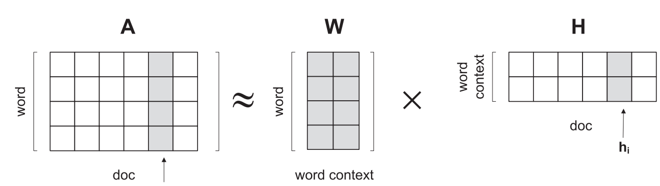


Fig. 3. Illustration of NMF. In matrix A, each column represents a document, each row represents a word. In matrix W, each column represents a word context, each row represents a word. In matrix H, each column represents a document, each row represents a word context. Matrix A can be approximated by the product of W and H.

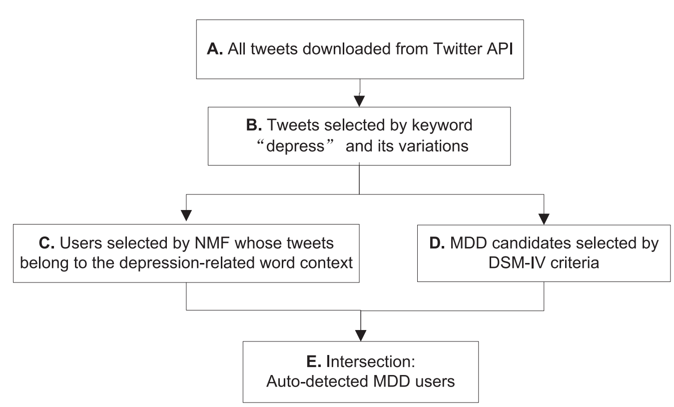


Fig. 4. Workflow of detecting MDD users. Follow the path A–B–C, we can get a group of MDD candidates whose tweets belong to depression word context. In part D, we add DSM-IV criteria for further MDD diagnosis. The final auto-detected MDD users are the intersection of part C and D.

Choosing an appropriate number of word contexts k is difficult but important. Because tweets are very noisy, we manually choose the number of word contexts, and examine the word contexts found by NMF under different choices. Then we pick the word context directly related to the depression mood, and collect all the documents assigned to that context.

For illustration purposes, for each word context (a column of W), we pick the words corresponding to the ten largest frequencies in order to display the word contexts.

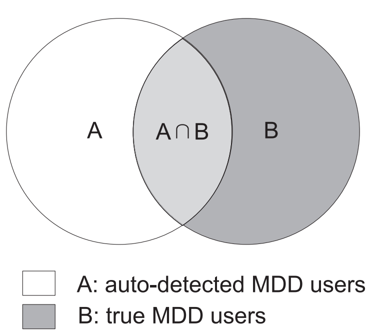
Auto-detected MDD users

Using text clustering to identify the word context for the depressed mood is not sufficient to detect the MDD candidates. Additionally, we use the diagnostic and statistical manual of mental disorders (DSM-IV) criteria (American Psychiatric Association, 2014) for MDD (Fig. 4).

The DSM-IV criteria require that MDD patients must suffer from at least five of the nine typical depression symptoms for more than two weeks (Skinner, Whiteley, & Ratner, 1990). These specific symptoms include depressed mood, decreased interest in daily activities, significant weight change or change in appetite, change in sleep or activity, fatigue, feelings of guilt or worthlessness, loss of concentration and having suicide plan. However, when analyzing Twitter text, we only focus on the depressed mood of the users. As a result, in order to apply DSM criteria to identify MDD candidates, we make the following change to the criteria: We require five or more tweets that are associated with depressed mood within a two-week period.

Procedure evaluation

We use the evaluation measures arising from information retrieval to assess the accuracy of our procedure (Rijsbergen, 1997): True positive (TP) is the number of times that our procedure correctly indicates that a user has MDD. False positive (FP) is the number of times that our procedure wrongly indicates that a user has MDD, while in fact the user does not have MDD. False negative (FN) is the number of times that our procedure indicates that a user does not have MDD, while the user actually has MDD.



$$\text{recall} = \frac{A \cap B}{B} = \frac{TP}{TP \cup FN}$$

$$\text{precision} = \frac{A \cap B}{A} = \frac{TP}{TP \cup FP}$$

$$\text{F-score} = \frac{2 \times \text{precision} \times \text{recall}}{\text{precision} + \text{recall}}$$

Fig. 5. Illustration of calculation for recall and precision. Set A represents auto-detected MDD users. Set B represents true MDD users. The intersection of A and B represents users who are suffering from MDD and also have been detected as MDD users by our procedure.

To find the true MDD users, we focused again on those MDD candidates identified by both key word and DSM-IV criteria. See part D in Fig. 4. We manually go through their tweets related to depression and exclude users whose tweets do not express depressed feelings, such as those regarding tropical depression.

Recall measures the fraction of auto-detected MDD users out of all the true MDD users, and precision measures the fraction of true MDD users out of all the auto-detected MDD users (Fig. 5). We use the F-score as a summary statistic to evaluate our procedure, which is the harmonic mean of precision and recall. A larger F-score measure corresponds to a more accurate procedure.

Spatial analysis for MDD users

In this research, we want to detect the spatial patterns for MDD users at county level in New York-Newark-Jersey City Metropolitan Statistical Area (abridged as "NY MSA" in the following) (Nussle, 2008). The time period is from 2013/09/05 to 2014/03/05. The null and alternative hypotheses are as follows:

H0. The MDD users are geographically randomly distributed.

H1. The distribution of MDD users are spatially clustered or dispersed.

To infer the location for an MDD user from their tweets related to depression, we create a user-by-county matrix. Each row represents an MDD user and each column represents a county. Each entry is the number of tweets by a user geo-tagged within a county. The sum of each row is the total number of tweets posted by that user, and the probability of user #1 living in county #1 is estimated as the proportion of tweets posted in county #1 by that user. The sum of each column is the number of MDD users located in that corresponding county.

After determining the number of MDD users in each county, we calculate Moran's I and use hot spot analysis to examine the spatial pattern of MDD users.

We then examine the association between depression rate at county level and local SES. It is reasonable to assume that the number of Twitter users is positively correlated with the size of population. Since over 88 percent of Twitter users are between 15 and 35 years old, we normalize the number of MDD users in each county by the population between 15 and 35 years old. Population and socioeconomic data such as education, income, and race are downloaded from the U.S. Census Bureau.

Results and discussion

Word context and auto-detected MDD users

To find the word context related to depression, we use all the data downloaded from 2013/09/05 to 2014/03/05 in the U.S. When using the keyword "depress" and its variations to select tweets related to depression, we choose not to remove the tweets whose subject is not the user, since the pilot study shows that only eight percent of our sample of tweets are talking about other people's

Table 2
Top ten words for each of the five auto-detected word contexts. The fifth word context is identified as the word context related to depression.

1	feel, today, makes, make, sad, reason, bad, hate, idk, depresso
2	lol,*, tweets, im, music, ass, listening, twitter, songs, drake
3	depression, post, concert, deep, real,*, anxiety, great, seasonal, back
4	depressing, tweets, day, music, weather, today,*, news, song, movie
5	depressed, im, feeling, makes, stressed, dressed, making,*, happy, mood

Note: * represents cursed words or swear words.

Table 3
Actual number of users corresponding to the workflow (See Fig. 4). The auto-detected MDD users in procedure E are the intersection of users in procedure C and D.

Procedure referring to the workflow	Number of users
C. Users selected by NMF whose tweets belong to depression-related word context	1225
D. MDD candidates selected by DSM-IV criteria	448
E. Intersection: auto-detected MDD users	286

depressed feeling. Additionally, for detecting MDD users, we require the occurrence of five or more tweets related to depression within two weeks, which can exclude most of the tweets whose subjects are not the Twitter user.

For NMF, we experiment with different numbers of word contexts k ranging from two to six, and find that the word context related to depression emerges when $k = 5$. This word context remains when k gets larger. Table 2 shows the NMF results: The word contexts 2, 3 and 4 are related more to music or weather; the word context 5 contains the key words "depressed, mood, stressed" and thus is identified as the word context related to depression. A more objective way to choose the depression-related word context is to determine the set of tweets assigned to each word context via NMF, and find which set of tweets has the largest intersection with the users selected by the DSM criteria for MDD. In this way, the word context 5 is chosen as the depression-related word context again.

To focus on NY MSA, Table 3 lists the statistics of the group of users in our experiment corresponding to Fig. 4. In order to evaluate the accuracy of our method, we identify 404 out of the 448 users are true MDD users. We then calculate TP (242 users), FP (44 users), and FN (162 users). The precision is 0.85, recall is 0.60, and F-score is 0.70.

Spatial analysis of MDD users in NY MSA

There are 25 counties in NY MSA. 249 out of 286 MDD users have shared location information. Dutchess, Pike, Union, Richmond, Monmouth and Ocean counties have the highest rate of MDD users (Fig. 6). Moran's I results (z-score = 3.54, p-value = 0.0004) also shows the clusters are significant. Hot spot analysis reveals that the hot spots clustering are concentrated around Ocean and Dutchess Counties (Fig. 7).

Bringing in SES variables to understand the auto-detected MDD results, we can see that the proportion of White population by county has a positive correlation with the total number of MDD users by county (Fig. 8). Counties with higher proportion of population with college education have a lower MDD rate. Counties with 8–15 percent of population with less than high school education have a higher MDD rate (Fig. 9). Middle-class people with a median household income at around 60,000 dollars have a higher prevalence of depression (Fig. 10), and this echoes some of the findings in the literature (Akhtar-Danesh & Landeen, 2007).

Conclusion and limitation

In this paper, we present a new method for public health research combining GIS with social media. Compared with traditional data collection method, our automated method for detecting MDD users is faster and cheaper for analysis and diagnosis. The system can be applied to some online forum for detecting depression topic and forwarding related questions to psychiatrists. Our GIS results also provide novel knowledge about this disorder by examining the geographic clustering of MDD users and relationship with SES.

In this research, we didn't say that our method for MDD diagnosis can replace the work of a clinical psychologist. Our method can improve diagnosis techniques for depression. Further detailed

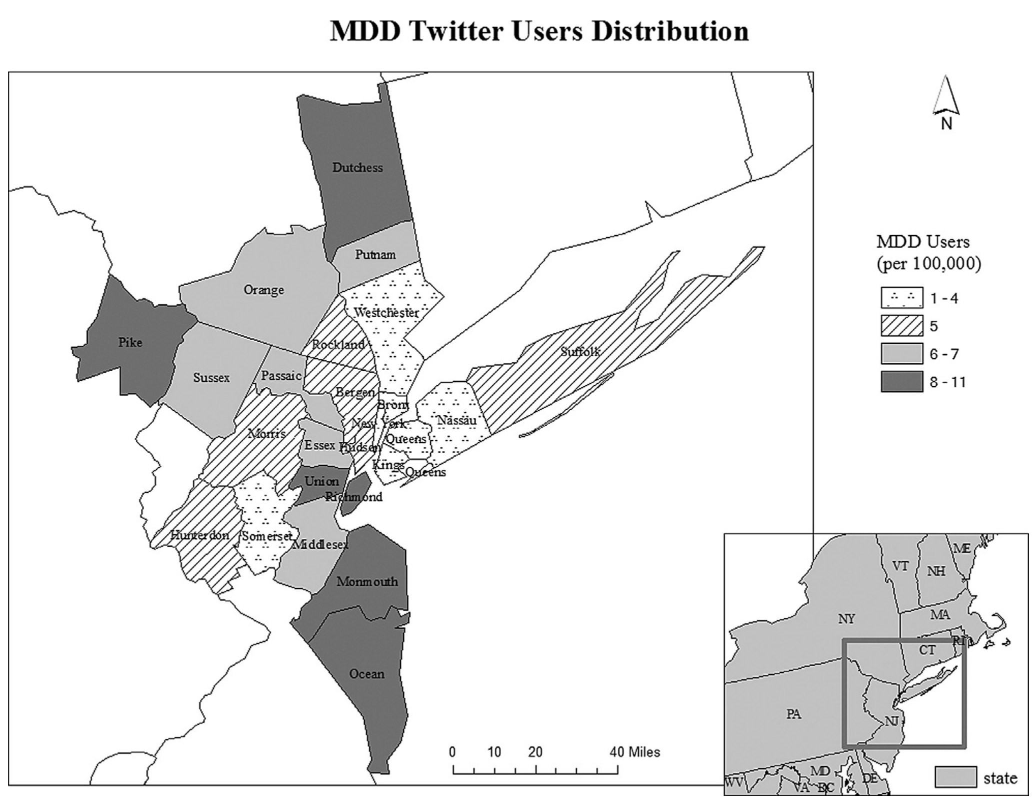


Fig. 6. Distribution of MDD Twitter users in NY MSA. We use different colors and textures to represent four levels of MDD rate.

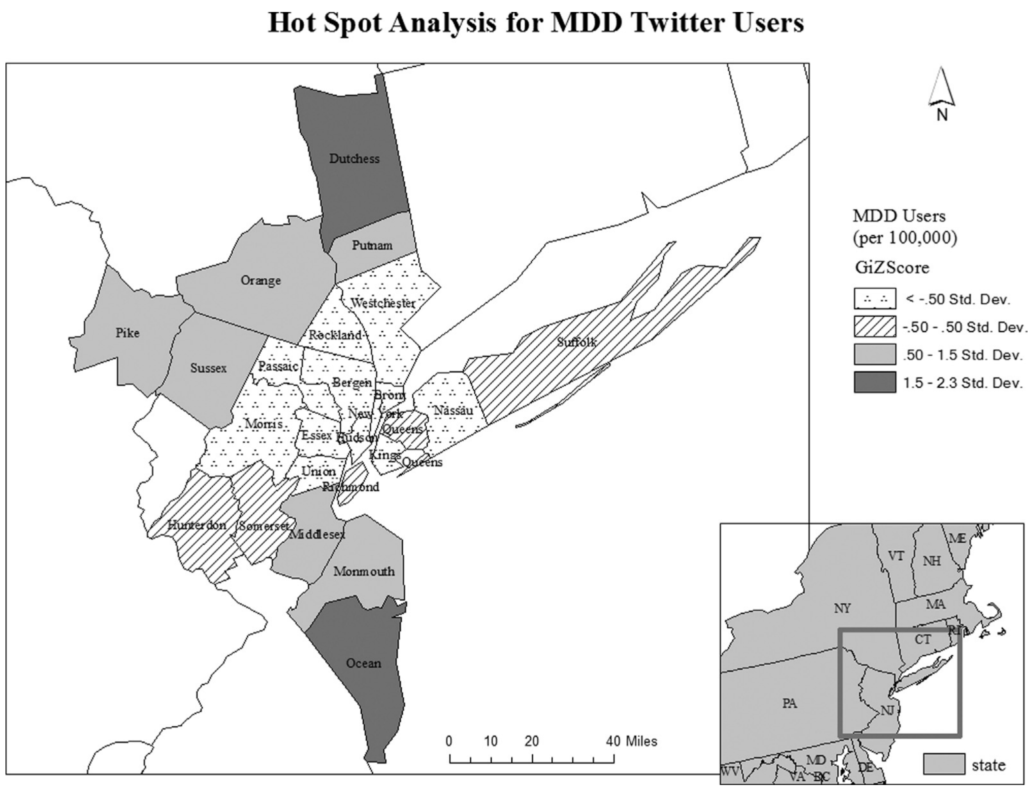


Fig. 7. Hot spot analysis for MDD Twitter users in NY MSA. We use different colors and textures to represent four levels of MDD spatial clustering.

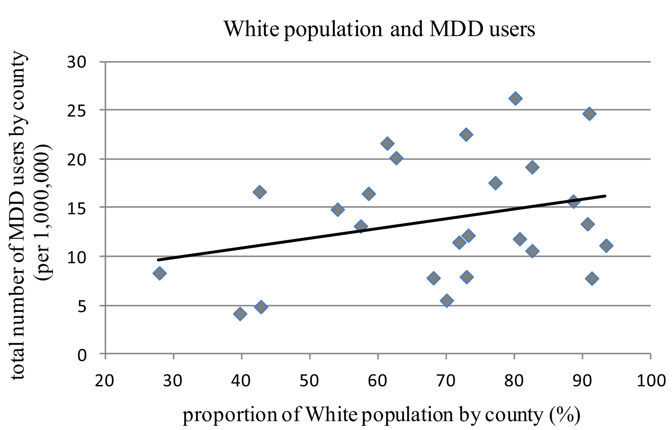


Fig. 8. Relationship between the proportion of White population and the rate of MDD users at county level.

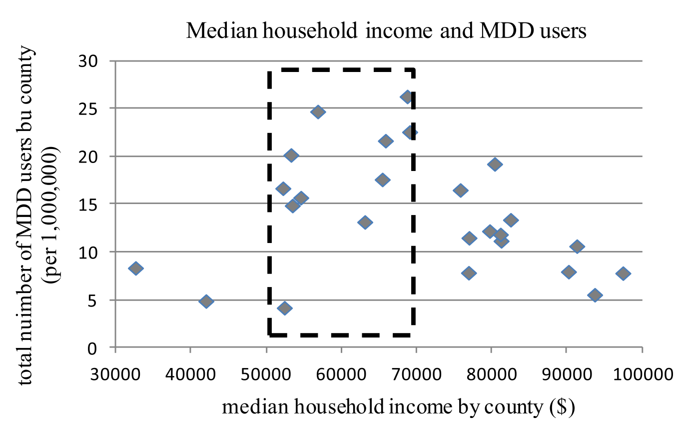


Fig. 10. Relationship between median household income in dollars and the rate of MDD users at county level.

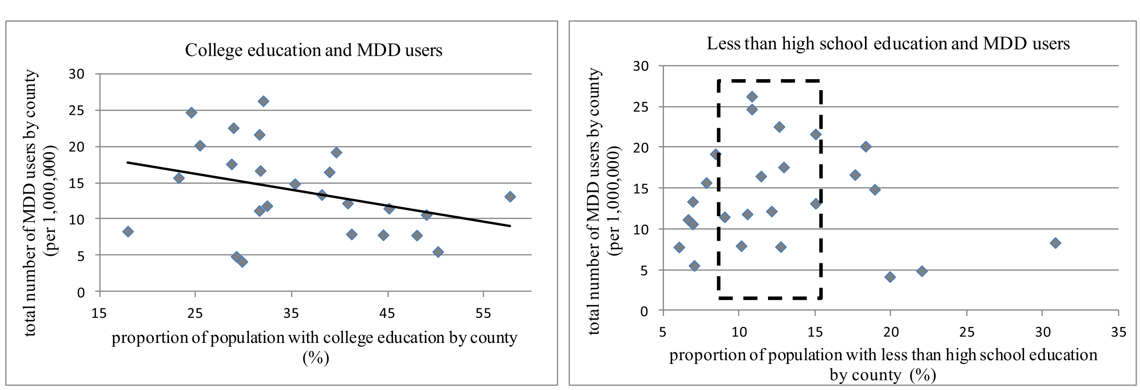


Fig. 9. Relationship between the proportion of population with college education, and those with less than high school education and the rate of MDD users at county level.

clinical contexts are needed to make a formal diagnosis. Future study should probe into the difference between depression detected online and self-reported depression reported by a professional clinical scale table.

Secondly, the Twitter APIs only allows free access to a one percent convenience sampling of tweets. Data acquired are restricted to users with public profiles. These may bring some bias to our result.

Additionally, we only include tweets written in English and users who identified themselves living in the U.S. Changing any of those may affect our results.

References

Akhtar-Danesh, N., & Landeen, J. (2007). Relation between depression and sociodemographic factors. *International Journal of Mental Health Systems, 1*, 4.

American Psychological Assocation, 2014, DSM http://www.psychiatry.org/practice/dsm

Barlow, D. H., & Durand, V. M. (2005). *Abnormal psychology: An integrative approach* (5th ed.). Belmont, CA, USA: Thomson Wadsworth. ISBN 0-534-63356-0.

Barlow, D. H., & Durand, V. M. (2011). *Abnormal psychology: An integrative approach*. Cengage Learning.

De Choudhury, M., Counts, S., & Horvitz, E. (2013). Predicting postpartum changes in emotion and behavior via social media. In *Proceedings of the 2013 ACM annual conference on human factors in computing systems* (pp. 3267–3276). ACM.

Deerwester, S. C., Dumais, S. T., Landauer, T. K., Furnas, G. W., & Harshman, R. A. (1990). Indexing by latent semantic analysis. *JASIS, 41*(6), 391–407.

GE Healthcare. (2012). Twenty six percent of online adults discuss health information online; privacy cited as the biggest barrier to entry. http://www.businesswire.com/news/home/20121120005872/en.

Ghosh, D., & Guha, R. (2013). What are we 'tweeting'about obesity? Mapping tweets with topic modeling and geographic information system. *Cartography and Geographic Information Science, 40*(2), 90–102.

Harman, C., Coppersmith, G., & Dredze, M. (2014). *Quantifying mental health signals in Twitter*.

Krieger, N., Williams, D. R., & Moss, N. E. (1997). Measuring social class in US public health research: concepts, methodologies, and guidelines. *Annual Review of Public Health, 18*(1), 341–378.

Lee, K., Agrawal, A., & Choudhary, A. (2013). Real-time disease surveillance using Twitter data: demonstration on flu and cancer. In *Proceedings of the 19th ACM SIGKDD international conference on knowledge discovery and data mining* (pp. 1474–1477). ACM.

Lee, D. D., & Seung, H. S. (1999). Learning the parts of objects by non-negative matrix factorization. *Nature, 401*(6755), 788–791.

Mair, C., Diez Roux, A. V., & Morenoff, J. D. (2010). Neighborhood stressors and social support as predictors of depressive symptoms in the Chicago community adult health study. *Health & Place, 16*(5), 811–819.

Mair, C., Diez Roux, A. V., Osypuk, T. L., Rapp, S. R., Seeman, T., & Watson, K. E. (2010). Is neighborhood racial/ethnic composition associated with depressive symptoms? The multi-ethnic study of atherosclerosis. *Social Science & Medicine, 71*(3), 541–550.

Mair, C., Diez Roux, A. V., Shen, M., Shea, S., Seeman, T., Echeverria, S., et al. (2009). Cross-sectional and longitudinal associations of neighborhood cohesion and stressors with depressive symptoms in the multiethnic study of atherosclerosis. *Annals of Epidemiology, 19*(1), 49–57.

Manning, C. D., Raghavan, P., & Schütze, H. (2008). *Introduction to information retrieval* (Vol. 1). Cambridge: Cambridge University Press.

Morbidity and Mortality Weekly Report (MMWR). (2010). Current depression among adults — United States, 2006 and 2008. http://www.cdc.gov/mmwr/preview/mmwrhtml/mm5938a2.htm.

Morstatter, F., Pfeffer, J., Liu, H., & Carley, K. M. (2013). Is the sample good enough? Comparing data from Twitter's streaming API with Twitter's firehose. In *Proceedings of ICWSM*.

Nussle, J. (2008). *Update of statistical area definitions and guidance on their uses. Office of Management and Budget*. http://www.whitehouse.gov/sites/default/files/omb/assets/omb/bulletins/fy2009/09-01.pdf.

Park, M., Cha, C., & Cha, M. (2012). Depressive moods of users portrayed in twitter. In *Proc. of the ACM SIGKDD Workshop on Healthcare Informatics, HI-KDD*.

Paul, M. J., & Dredze, M. (2011). You are what you Tweet: analyzing Twitter for public health. In *ICWSM*.

Rijsbergen, C. V. (1997). *Information retrieval. 1979*. Btterworths, London.

Rudd, M. D., Berman, A. L., Joiner, T. E., Nock, M. K., Silverman, M. M., Mandrusiak, M., et al. (2006). Warning signs for suicide: theory, research, and clinical applications. *Suicide and Life-Threatening Behavior, 36*(3), 255–262.

Skinner, B. F., Whiteley, J. M., & Ratner, H. (1990). *BF skinner*. American Psychological Association.

Xu, W., Liu, X., & Gong, Y. (2003). Document clustering based on non-negative matrix factorization. In *Proceedings of the 26th annual international ACM SIGIR conference on research and development in information retrieval* (pp. 267–273). ACM.