

Does Social Media Behaviors Reflect Users' Anxiety

A Case Study of Twitter Activities

Tian Wang

Illinois Informatics
University of Illinois, Urbana-Champaign
Champaign, USA
Email: tianw7@illinois.edu

Masooda Bashir

School of Information Sciences
University of Illinois, Urbana-Champaign
Champaign, USA
Email: mnb@illinois.edu

Abstract—Anxiety has been a common mental health disorder that affects many people, especially young adults, but it is often undetected or untreated. Meanwhile, mental health professionals and services available are insufficient for the assessment and treatment of mental illnesses, including anxiety, due to cost, availability, and other factors. Therefore, there is an urgent need to develop new approaches for mental healthcare services. Considering that the number of young adults using social media has increased significantly in the last decade, data from social media activities could possibly be a new approach for individual's mental health assessment. In this study, we recruited 95 participants and examined their Twitter activities to learn if young adults' social media activities are relevant to their anxiety level. Twitter activities were assessed between two groups of users, namely, those participants who reported symptoms, and participants reporting few or no anxiety symptoms. Our preliminary results show that certain social media behaviors differ for individuals that report anxiety symptoms and those with less or no symptoms of anxiety. Results from this study could be beneficial for both researchers and mental health professionals as a supplemental source of information and could potentially provide a new approach for mental health assessment and treatment in the future.

Keywords – Social media behaviors; Mental health; Data analytics.

I. INTRODUCTION

While mental health is a growing concern around the globe in recent years, anxiety disorders have been one of the most commonly diagnosed mental illnesses. According to National Institute of Mental Health, about 18% of U.S. adults aged 18 or older are affected by anxiety in any given year [1], and over 20% of those adults with anxiety disorders have severe symptoms. Moreover, 31.1% of U.S. adults are estimated to experience anxiety disorders at some time in their lives [2]. In particular, anxiety disorders are very common among children, adolescents, and young adults. Statistics show that about 30% of young adults aged 18-29 in U.S. have some type of anxiety disorders [3]. However, assessments and treatments for mental illness have been considered to be insufficient in the recent past. A previous study in 2017 [4] pointed out that although effective treatment may be theoretically available, healthcare services

are still lacking for many mental illnesses including anxiety disorders. According to a survey study by National Comorbidity Survey (NCS), about 75% of nearly 10,000 people with generalized anxiety disorder or social anxiety disorder did not receive any treatment [2]. Therefore, it is necessary for researchers and mental health professionals to explore new approaches to provide more effective and efficient mental healthcare services besides the current applied methods.

Nowadays, interactive communication technologies such as social media have become ever more popular all over the world. Research studies have found that 72% of the US population uses some type of social media [5]. Especially, social media continues to be popular among young adults. Reports found that usage of social media for young adults has increased from 9% in 2005 to 89% in 2013 [6]. Social media is mostly used by individuals to share information, express feelings and emotions, catch up with latest news, and connect with other people. Hence, the information from an individuals' social media activities which provides a great reflection of their daily life could be helpful in determining their mental wellbeing.

This research study aims to examine the relationship between young adults' social media behaviors and their mental health status, especially general anxiety level. We selected Twitter as the social media platform to observe and analyze in this study based on its large number of users and content. As of 2019, Twitter has 139 million active users per day [7] with a total of 500 million Tweets being posted [8]. To understand the relationship between individuals' Twitter activities and anxiety level, two research questions will be addressed when conducting the study:

- Is there a relationship between Twitter activities and one's anxiety level?
- Are there any differences in Twitter activities among people with symptoms of anxiety and people without such symptoms?

The subsequent parts of this paper are organized in the following way. In Section II, previous literature will be reviewed and summarized to provide an overview of the current practices related to social media analytics and mental health predictions. Section III will explain the methodology used in this research study, including the process of data collection, and elements measured in the study. Results will

be displayed in Section IV, and Section V will discuss the characteristics of the preliminary results. We conclude the work in Section VI.

II. LITERATURE REVIEW

Although typical diagnosis on mental health status is based on self-reported information and mental status examinations [9], social media provides reflections of an individual's personal life given its massive amount of information on people's personal experiences, feelings, and emotions that is being shared. Individuals use social media for various reasons: social interaction, information seeking, entertainment, relaxation, or expression of opinions [10], and each of these activities involve significant personal data collection, sharing, and processing. Previous research suggested that social media activities could be used for personal identification, including age, gender, and personality [11]. For example, one research study reported that users' personality traits could be predicted by their information shared publicly on Facebook, and the results of the predictive models were more accurate than the predictions made by close friends and family [12]. The statistical models also provided accuracy rates on predicting other personal information, for example, sexual orientations (if someone was homosexual or heterosexual), or political opinions (if the user was Democrat or Republican) [13]. In addition, expressed emotions on social media websites have been used by commercial companies to improve their business decision making and tailor their marketing strategies [23].

Previous studies on social media found that individuals' social media activities could be used to assess, identify, and predict their health status. For instance, a study on Twitter behaviors by Choudhury et al. [9] found that people with major depression may behave differently on social media compared with people without any symptoms. Another study also suggested that people with depression symptoms may have different perspective on sharing personal information on social media compared with people without such symptoms [21]. Besides predicting depression, similar work by Choudhury et al. [14] also applied predictive models to forecast if new mothers could possibly experience significant postpartum changes by analyzing their social media activities. Social media behaviors are not only used for personal health assessment, but also could be used for overall community healthcare. Social media analytics have been widely used in the field of medical sciences, examples including biomedical outcome prediction [15], infectious disease risk prediction [16], predicting and tracking the trends of disease outbreaks [17]. Using social media for health predication have also been shown to have accuracy rates in previous studies. One such study found that Facebook postings on two health issues were very similar to the actual results on official reports about the same topics [18].

III. METHODS

An online survey was designed and used to conduct this research study. Participants were recruited randomly via

Amazon Mechanic Turk (MTurk), which is a crowdsourcing platform. The online survey used General Anxiety Disorder-7 (GAD-7), which is a seven-item, self-report questionnaire, to assess participant's health status based on the scores. The GAD-7 has been used in both research and clinical settings to detect individual's anxiety level [19]. The survey also includes questions on Twitter usage, as well as general demographic information. Participants were asked to provide their Twitter links and give permission to researchers for observing their Twitter activities. No personal identifiable information was collected in this study. The recruitment process and the online survey were reviewed and approved by the Institutional Review Board. In addition, relevant resources were provided to participants at each page of the survey in case they felt upset or needed any help during the study.

By using filters on MTurk, only adults aged 18-25 in U.S. holding a Twitter account were recruited in this study since the aim of this study is to analyze the relationship between young adults' social media activities and their mental health status. A total of 200 participants were randomly recruited from MTurk, and 95 responses were recorded after excluding noisy data (i.e., incomplete survey or filling out the survey within extremely short time period), with 52 out of 95 (54.74%) participants being female. The average age of participants was 25.35, and 65 participants (68.42%) held a bachelor's degree or higher. Over half of the participants (56 out of 95) indicated that they use Twitter every day. Participants' responses on their health status were converted to numerical scores based on GAD-7 guideline. Total score for each participant was calculated, and participants were divided into two groups, participants with symptoms of anxiety and participants with less or no symptoms of anxiety, by using the cut-off point (10 or greater) recommended by GAD-7. There were 34 participants who received a total score of 10 or greater, while 61 participants received less than 10 in the GAD-7 questionnaire.

By using Twitter links provided by each participant, data on all the publicly available information was collected, including information on user's profile, followings and followers, Twitter postings (Tweet), liked Tweet, and photos. Twitter's developer API was used to retrieve content of Tweets from user's timeline. One notable exclusion during the retrieving process was that only historical Tweets posted before the participant took the survey were collected, since the measurements on GAD-7 are based on previous experience. In addition, this step was taken to minimize social desirability on Twitter based on the study participation. A total number of 353514 Tweets were retrieved from the 95 participants' public Twitter links.

To analyze participants' Twitter activities, three elements were measured: overall engagement, negative emotions, and level of personal information revealed publicly. In a previous study, Choudhury et al. have defined and explained the measurements of engagement [9], and similar methods were applied in this study. Besides that, since posting personal content on social media is on the rise [10], we were also interested in assessing negative emotions and the level of

personal information revealed on participant's Twitter links to see if those two elements are also related to their mental health status.

- **Engagement:** Number of Tweets, Retweets, replies, liked Tweets, followers, and followings. We considered those elements to be a reflection of how active an individual is on Twitter. For example, if an individual posts a huge number of Tweets in a short time period, it may imply that he or she spends a lot of time on social media every day. We also considered Retweets and replies to be related to the interaction with other people, and the number of followings and followers to show how close the individual connects to the social media community [20].
- **Negative emotions:** Sentiment analysis has been widely applied in Twitter data analytics. For example, models of sentiment analysis such as polynomial regression, classification modeling, and lexicon-based sentiment analysis models, all have been helpful in categorizing data and hence enabling fast decision making [22]. To analyze the negative emotions expressed in participants' Twitter activities, we considered and examined two specific elements: frequency of words with negative sentiment used in original Tweets, and if any negative attitude was exhibited in Twitter activities other than original Tweets, such as liking or re-posting negative Tweets, or including negative words in profile description. Content and sentiment analysis were applied to assess the negative emotions.
- **Level of personal information shared publicly:** In a previous study on depression and social media activities, results showed a relationship between the level of personal information shared publicly and the person's depression [20]. More specifically, we observed activities such as using selfie as a profile photo, posting photos or videos related to personal life, or adding personal identifiable information on profile description. Therefore, examining publicly shared personal information might potentially indicate if individuals are open and willing to disclose and share their personal life publicly with the online community, including both acquaintances and strangers.

IV. PRELIMINARY RESULTS

As described in Section III, three elements are analyzed based on data retrieved from participants' Twitter activities: engagement, negative emotions, and level of personal information shared publicly. For each of the two groups, participants with anxiety symptoms and participants with few or no symptom of anxiety, the results were recorded accordingly and then compared. The results and comparisons for all of the three elements are shown in the following sub-sections:

A. Engagement

To analyze the engagement element, the average number of Tweets, followings, and followers are recorded. The statistics on the engagement for both groups are shown in Table 1.

TABLE I. STATISTICS ON ENGAGEMENT

Participants	With anxiety symptoms	Without anxiety symptoms
	<i>N=34</i>	<i>N=61</i>
Average # of Tweets	4970.30	4138.04
Average # of followings	449.96	373.13
Average # of followers	190.18	229.42
Ratio of Followings/Follower	2.37	1.63

From the table above, statistics show that participants with anxiety symptoms have a slightly larger number of Tweets posted. One reason could be that the sample size is small and outliers exist when analyzing the data (for example, there is one particular participant who posted over 30,000 Tweets, and another participant only posted 1 Tweet although the Twitter was registered in 2011). Note that the number of Tweets recorded in Table 1 includes both original posts, re-posts (Retweets), and replies, and most of the participants with larger number of Tweets have a lot of Retweet, which increased the total number. Although outliers may influence the results, those outliers were not excluded from our analysis because the number of Tweets could be a determining factor when evaluating how active a participant is on a given social media. Only Tweets posted before the participant took this study's survey were collected. One challenge that we faced was related to collecting all of the user's historical data and relating that to their mental health condition because, for this study, we did not assess or know the onset of anxiety. Therefore, our study results may be limited in explaining the relationship between participants' Twitter activities and their anxiety level. Further research studies with more comprehensive survey questions that assess the onset of the mental health condition would be recommended to have more precise results.

Based on observations of the data collected, participants with anxiety symptoms had significantly higher percentage of Retweets in their total number of Tweets, for example, sharing an external link from other websites, or re-posting a celebrity's Tweet. However, people with less or no symptoms of anxiety were more likely to post original content or interact with others by replies.

Also, participants with anxiety symptoms have more followings but less followers than participants with less or no symptoms. Ratios of number of followings to number of followers for these two groups are 2.37 and 1.63, respectively. Further observations found that participants with anxiety symptoms are more likely to follow verified Twitter accounts (celebrities, organizations' official accounts, website accounts), while participants with less or no

symptoms tend to follow personal accounts like their family and friends that are close to them in real life.

B. Negative Emotions

To understand the emotions expressed on participants' Tweets, only original posts (Tweets) and replies were retrieved and analyzed as measurement, since content on reposts (Retweets) were not created directly by participants themselves. Although those Retweets may also imply participants' attitude, in this study, we only focused on posts from the participants directly. The first step was to analyze word occurrence and frequency for Tweets and replies. After excluding common used words like pronouns and prepositions (i.e., "my", "you", "the"), or words without clear sentiment (i.e., "time", "video", "do"), Table 2 shows the top 10 most frequently used words (with sentiment) for participants with symptoms of anxiety and participants with few or no anxiety symptoms.

TABLE II. WORD FREQUENCY FOR ORIGINAL POSTS

Rank	With anxiety symptoms	Without anxiety symptoms
1	love (0.27%)	love (0.74%)
2	good (0.20%)	good (0.52%)
3	new (0.12%)	new (0.52%)
4	better (0.10%)	liked (0.41%)
5	great (0.09%)	best (0.40%)
6	best (0.09%)	chance (0.38%)
7	happy (0.08%)	right (0.37%)
8	over (0.07%)	never (0.37%)
9	stop (0.06%)	happy (0.33%)
10	bad (0.05%)	amazing (0.27%)

From Table 2, we could see that there are similarities for the most frequently used words with sentiments between the two groups. However, although participants with anxiety symptoms mentioned positive words in their original posts, the percentage of word occurrence (how many times the word occurred / total number of all the words) for those words is lower than the one for participants with less or no anxiety. For example, the percentage of word occurrence for the word "love" is 0.74% for participants with less or no anxiety, but only 0.27% for participants with reported anxiety symptoms. It is similar for the words with positive sentiment like 'good', 'new', and 'best'. Another observation from Table 2 is that the group with anxiety symptoms frequently used some words with negative sentiment, for example, the word "bad" was ranked 10th for participants with anxiety symptoms, but was not in the top 10 for the other group.

Since there were too many words in the list, we took a closer look at the words with negative sentiment. For participants with anxiety symptoms, the frequency of words

with negative sentiment in their Tweets and replies was higher than the one for participants with less or no symptoms. For example, the frequency for the word "sad" was about 0.07% for participants that reported anxiety, but was less than 0.01% for the other group.

Participants with anxiety symptoms were not only expressing negative emotions on Tweets, but also on their Twitter profile descriptions. One example was a participant who described herself as "an awkward girl trying to find herself" in her Twitter profile, and there was another participant from this group using the phrase "pseudo human" in his profile description.

C. Level of Personal Information Shared Publicly

In this study, the level of personal information shared publicly on Twitter is examined by three sub-categories:

- **Profile photo:** Is the user's selfie being used as profile photo? If so, is the photo clear enough to identify the person?
- **Profile description:** Is there any personal identifiable information disclosed in the profile description? For example, gender, age, education level/school, geolocation, occupation.
- **Media posts:** Is there any media post (photo or video) related to the user's personal life? For example, selfie, photos of personal experiences (hobbies, travels, daily life), or photos with family and friends.

The results showed that participants with less or no anxiety symptom (21 out of 61, 34.4%) are more likely to use their selfies as profile photos than participants with anxiety symptoms (7 out of 34, 20.6%). For participants who reported anxiety symptoms, 6 of them were using photos that were not clear enough to identify their faces in different ways: some of those pictures included blur effects, some pictures were taken from far away, and some selfies were taken from the right/left side of the face. Most of the participants with anxiety symptoms were using irrelevant pictures as profile photos, like animation picture, landscape, celebrities, and so on.

For profile description, 13 of the 61 participants with few or no anxiety symptoms had detailed information about themselves in the profile description. Most of them were indicating their job occupations, school and graduate year, place (state or city) of residency. However, only 2 of 34 participants who reported anxiety symptoms included some basic level of personal information in their profile description.

Participants with few or no anxiety symptoms were also more willing to post photos and videos about themselves, their family and friends, and photos related to personal experiences. On the other hand, participants who reported anxiety were more reserved to post any media that could potentially disclose their personal information. They were more cautious with media posts shared publicly, and they were more likely to only post default gif pictures by Twitter, or irrelevant pictures from the Internet.

Overall, from the comparison between the two groups in the three sub-categories described above, participants with

few or no anxiety symptoms are more open to reveal their personal information or daily life in their profile and posts. The level of personal information shared publicly is much lower for participants with anxiety symptoms.

V. DISCUSSION

Based on the preliminary results described in the above section, there are some initial findings on the relationship between social media behaviors and individuals' mental health status. In our study, we found that participants with reported anxiety symptoms behave differently on social media like Twitter compared with participants with few or no anxiety symptoms. For the social media engagement, study participants with high anxiety level are more likely to re-post content from others instead of creating original content. Considering they have much more followings than followers, this may be an indication that they are less active in connecting with others or with a community since they tend to follow verified accounts like official organizations or celebrities instead of interacting with others. Our sentiment analysis from Twitter content showed that participants with anxiety symptoms tend to be more passive and negative by using more words with negative sentiment and fewer positive words being mentioned. In contrast to participants with less or no symptoms of anxiety, who are more likely to express their negative attitude (for example, loneliness, sadness, anxiety, or frustration), not only on Twitter posts, but also on their Twitter profile descriptions. The results also reveal that participants with anxiety symptoms are more cautious with sharing their personal information publicly via Twitter, while participants with few or no anxiety symptoms being more open to share their photos, information, and personal life with the Twitter online social media community.

Considering the preliminary results found in this study, as well as previous research studies on social media and mental health, social media behaviors may be a useful source of information for individuals' general mental health status. Thus, in our quest of wanting to provide more timely, cost effective, and broader coverage of mental health services, perhaps social media data analysis may provide a way forward. Also, such information could be potentially used as a novel and effective approach to identify, detect, and predict one's mental health status other than traditional methods like self-reported information and screening tests, with two additional benefits. First, this approach is easy to implement and it includes rich content coverage of an individual's online activities. Second, compared to traditional methods, this approach could provide more efficient, objective, and unbiased information sources since it is entirely based on individuals' behaviors in their daily life without their awareness of being assessed or tested when applying traditional methods.

However, it is important to address privacy violations that may occur given the fact that people with reported anxiety symptoms could be identified according to their social media activities. There are already some commercial companies or third-party services using social media users' data to categorize them into certain groups and send targeted

advertisements [21]. While data sharing and personal identification might be unavoidable given current circumstances for the online environment, it is vital to recognize the importance of user privacy protections since social media users with mental health disorders may be more vulnerable and sensitive. Thus, such social media analytics should only be considered if and when privacy protections are of the highest priority.

Still, more research studies are needed to pave the way forward as our study is the first one and it includes some limitations. First, although our preliminary results suggest relationships between individuals' social media activities and their mental health status, especially anxiety level, more research studies with larger sample of participants and more accurate mental health assessment are necessary to confirm the results since the sample size is small and the mental health assessment is based on self-reported information. The crowdsourcing platform (MTurk) used for recruitment in this study also has its own limitations such as not being a representative of all populations and only including online participants. Second, only U.S. participants were selected to be involved in this study, but young adults from other countries may have different social media behaviors because of cultural differences. Studies on other social media platforms, such as Facebook and Instagram are also suggested to analyze individual's behaviors on various social media platforms, since this study only focused on one specific platform (Twitter). In addition, while social media activities could provide supplemental information for clinicians and researchers when assessing individual's mental health status, it is also important to address privacy protections when such level of personal information is involved.

VI. CONCLUSION

This research study shows initial findings that individuals' social media behaviors may be relevant to their mental health status, and people with symptoms of anxiety may behave differently than people without such symptoms. Information from social media activities could potentially be considered as a new method for mental health assessment and treatment along with traditional methods given how much it reveals an individual's personality and daily life. Results from this study may provide the initial steps towards a new direction in mental healthcare services that can be timely and reach broad and diverse populations.

REFERENCES

- [1] National Institute of Mental Health, Any Anxiety Disorder, Retrieved: August, 2020 from <https://www.nimh.nih.gov/health/statistics/any-anxiety-disorder.shtml>.
- [2] Harvard Medical School, National Comorbidity Survey (NCS), Retrieved: August, 2020 from <https://www.hcp.med.harvard.edu/ncs/index.php>, 2007.
- [3] Behavioral Health in the United States, Substance Abuse and Mental Health Services Administration, n.d., 7 February 2016.
- [4] J. Weissman et al, "Disparities in health care utilization and functional limitations among adults with serious psychological

- distress, 2006–2014," *Psychiatric Services*, vol. 68, no.7, pp. 653-659, 2017.
- [5] Social Media Fact Sheet, Pew Research Center, Retrieved: August, 2020 from <https://www.pewresearch.org/internet/fact-sheet/social-media/>, 2019.
- [6] D. C. Brooks, and J. Pomerantz, "ECAR Study of Undergraduate Students and Information Technology, 2017," EDUCAUSE, 2017.
- [7] I. Lunden, Twitter Q2 beats on sales of \$841M and EPS of #0.20, new metric of mDAUs up to 139M, Retrieved: August, 2020 from <https://techcrunch.com/2019/07/26/twitter-q2-earnings/>, 2019.
- [8] Twitter, The power of Twitter, Retrieved: August, 2020 from <https://business.twitter.com>, 2019.
- [9] M. D. Choudhury et al, "Predicting depression via social media," *lcwsm* 13, pp. 1-10, 2013.
- [10] A. Whiting, and D. Williams, "Why people use social media: a uses and gratifications approach," *Qualitative Market Research*, vol. 16, no. 4, pp. 362-369, 2013.
- [11] H. A. Schwartz et al, "Personality, gender, and age in the language of social media: The open-vocabulary approach," *PloS one*, vol. 8, no. 9, e73791, 2013.
- [12] J. Golbeck, C. Robles, and K. Turner, "Predicting personality with social media," CHI'11 extended abstracts on human factors in computing systems, pp. 253-262, 2011.
- [13] M. Kosinski, D. Stillwell, and T. Graepel, "Private traits and attributes are predictable from digital records of human behavior," *Proceedings of the national academy of sciences*, vol. 110, no. 15, pp. 5802-5805, 2013.
- [14] M. D. Choudhury, S. Counts, and E. Horvitz, "Predicting postpartum changes in emotion and behavior via social media," *Proceedings of the SIGCHI conference on human factors in computing systems*, pp. 3267-3276, 2013.
- [15] S. D. Young, "Behavioral insights on big data: using social media for predicting biomedical outcomes," *Trends in microbiology*, vol. 22, no. 11, pp. 601-602, 2014.
- [16] J. Rushmore et al. "Social network analysis of wild chimpanzees provides insights for predicting infectious disease risk," *Journal of Animal Ecology*, vol. 82, no. 5, pp. 976-986, 2013.
- [17] C. W. Schmidt, "Trending now: using social media to predict and track disease outbreaks," *Environmental Health Perspectives*, vol. 120, no. 1, pp. a30-a33, 2012.
- [18] N. Vincent, "Social Media and Environmental Health Crises: An Examination of Public Response to Imported Drywall and Perchlorate Health Risks," *American Public Health Association Annual Meeting and Exposition*, Washington, DC, vol. 31, 2011.
- [19] N. Williams, "The GAD-7 questionnaire," *Occupational Medicine*, vol. 64, no. 3, pp. 224-224, 2014.
- [20] T. Wang, and M. Bashir, "A Study of Social Media Behaviors and Mental Health Wellbeing from a Privacy Perspective," *International Conference on Applied Human Factors and Ergonomics*, vol. 1213, pp. 137-144, 2020.
- [21] K. Leetaru, What Does It Mean For Social Media Platforms To "Sell" Our Data? *Forbes*, Retrieved: August, 2020 from <https://www.forbes.com/sites/kalevleetaru/2018/12/15/what-does-it-mean-for-social-media-platforms-to-sell-our-data/#1d8005892d6c>, 2018.
- [22] X. Guo, and J. Li, "A Novel Twitter Sentiment Analysis Model with Baseline Correlation for Financial Market Prediction with Improved Efficiency," *2019 Sixth International Conference on Social Networks Analysis, Management and Security (SNAMS)*. IEEE, pp. 472-477, 2019.
- [23] P. Madhala, J. Jussila, H. Aramo-Immonen, and A. Suominen, "Systematic literature review on customer emotions in social media," *ECSM 2018 5th European Conference on Social Media*, Academic Conferences and publishing limited, pp. 154, 2018.