Logging Off: Motivations for Taking a Break from Twitter

Giorgia Fujita, Bob Guan, Yu Guo

Abstract

While social media has become increasingly widespread, some users choose to take a break or leave social media altogether. This paper presents results on a collection of 5000 tweets filtered to 346 tweets which express users' desire to take a break from Twitter. Results show the lack of a clear relationship between prior frequency of posts and the duration of break afterwards based on the limitation of 150 posts per week as a measure of frequency of Twitter usage. Furthermore, most users did not express any motivations for taking a break in their break tweets. However, the effect of taking a break did positively affect users' emotions as expressed in their return tweets. These results build on previous work to help advance how we account for non-use of social media.

1 Introduction

Social media has become an almost indispensable part of daily life with 90% of 18-29 year olds using at least one social networking site (Perrin, 2015). However, some have chosen to take breaks from social media or cease using social media sites altogether.

Despite the abundance of research on those who use social networking sites, relatively little research has been conducted on those who choose not to use such sites. This paper explores the relationship between prior frequency of posts and the duration of break afterwards, the motivations for taking a break, and the effect of taking a break on users' emotions.

2 Previous Research

2.1 Motivations for Leaving Social Media

Recent research shows that users choose to leave social media sites for a number of reasons. Using a questionnaire distributed to over 400 Internet users, Baumer et al. (2013) examined motivations for intentionally leaving social media. The questionnaire presented two types of questions, one type being straightforward, factual questions about Facebook usage and the second type exploring the experience of deactivating or deleting Facebook through a set of open-ended, free response questions. The second portion of the questionnaire which focused on the latter type of questions revealed that Internet users were motivated to leave Facebook based on privacy, data misuse, productivity, banality, addiction, and external pressures. Despite the extensive reasons for which those users deactivated or deleted Facebook, nearly half the respondents who left Facebook subsequently returned to the site. Based on this insight we explore the following:

RQ1: Is there a relationship between prior frequency of posts and the duration of break afterwards?

Hypothesis 1: Users who post frequently prior to taking a break are more likely to take a shorter break.

Another study focused on Grindr suggested departure from social media as a gradual and reversible process based on similar motivations. Brubaker et al. (2016) conducted 16 semi-structured interviews during a 2-month period in mid-2012 to discover why participants left Grindr. Some participants were motivated to leave Grindr due to concerns around productivity and banality, as they found the application to be time-consuming and distracting. It is worth noting that the participants ranged in age from 27 to 43 years.

Focusing on users primarily under the age of 34, Maier et al. (2012) surveyed 571 participants in developing a model on social media usage. Individuals completed all items of the survey on a 7-point Likert scale (1 meaning totally disagree; 7 meaning totally agree). Certain motivations for leaving social media were identified from the survey including privacy, banality, addiction, and external pressures. In regards to privacy, most users expressed that they feel Facebook invades their personal life with a correlation of 0.933. Participants also indicated sentiments of banality, feeling "tired from [their] Facebook activities" with a correlation of 0.859. Additionally, addiction and external pressures played a role in participants' sentiments towards social media usage. Many users felt they pay "too much attention" to posts of their friends on Facebook (correlation of 0.866) and that they "perceive pressure" from their friends to check their news on Facebook regularly (correlation of 0.842).

Based on research by Baumer et al. (2013), Brubaker et al. (2016), and Maier et al. (2012), we explore the following research question and hypothesis:

RQ2: What reasons do people give in their posts when they decide to take a break?

Hypothesis 2: People give a variety of reasons for taking a break. Potential ones are:

-- Privacy concerns
-- Data misuse
-- Unproductivity
-- Banality
-- Addiction
-- External pressures

While numerous studies have confirmed the motivations categorized by Baumer et al. (2013), there may be motivations which may not fit into those categories. Schoenebeck et al. (2014) observed Twitter users tweeting that they were "giving up Twitter for Lent" on Ash Wednesday in 2011, 2012 and 2013. Using the Twitter API and Perl scripts, tweets were captured from the 40 days prior to Lent throughout the Lenten period until Easter. Three concerns surfaced among interview participants with respect to social media use: spending too much time on it (addiction), tradeoffs of not spending time elsewhere (productivity), and a concern about social media not being "real life." While addiction and productivity are categorized motivations, the concern about social media not being "real life" may require further research in order to fully understand.

2.2 Motivations for Returning to Social Media

A recent study showed that 61% of Facebook users have taken a voluntary break from the site (Pew Research), but less than half of Facebook users ages 18-40 (between 34-42%) self-reported that their time spent on Facebook on a typical day decreased since the prior year. Thus, while taking breaks is common, users usually do not leave social media sites altogether. Baumer et al. (2015) explored social media reversion, or when a user intentionally ceases using a social media site but then later resumes use of the site. By analyzing data from people who volunteered to stay off Facebook for 99 days with some returning before that time, Baumer et al. (2015) identified factors predicting the likelihood of reversion to Facebook, including perceived addiction and friends' reaction to non-use.

Stieger et al. (2013) compared Facebook quitters (n=310) against Facebook users (n=321) across three measures: privacy concern scale (PCS), internet addiction test (IA-T), and mini international personality item pool (Mini-IPIP) personality measure. The PCS is a short six-item questionnaire which asks participants to answer on a scale how frequently (1 meaning Never; 5 meaning Always) they engage in certain behaviors related to privacy in everyday situations to gauge general privacy concerns. The IA-T is a 20-item questionnaire measuring Internet addiction by asking how personal Internet use affects daily routine, social life, productivity, sleeping pattern, and feelings (five-point scales; 1 meaning Never; 5 meaning Always). The Mini-IPIP is a 20-item measure of the Big Five personality dimensions (five-point scales; 1 meaning Very accurate; 5 meaning Very inaccurate). This study focused on four items of the Mini-IPIP which refer to extraversion, agreeableness, conscientiousness, neuroticism, and openness to experience. Stieger et al. (2013) found that Facebook quitters had a significantly higher general concern about privacy, higher Internet addiction scores, and were more conscientious than Facebook users. Thus, Facebook users may feel more inclined to stay on the platform if their Internet addiction is not as high and they express less concern over privacy and conscientiousness.

2.3 The Effect of Breaks from Social Media

While motivations for returning to social media after a break have been identified, the effect of such breaks on a user's emotional state has not been as thoroughly researched. Thus, we explore:

RQ3: What effect does a break have on the sentiment expressed in users' first post upon return?

Hypothesis 3: Users' return posts show more positive sentiments.

We hypothesize that users returning to social media show more positive sentiments based on insights from Stieger et al. (2013) in which Facebook users felt more inclined to stay on the platform if their Internet addiction is not as high and they express less concern over privacy and conscientiousness.

3 Method

The main goal initially was to maintain efficiency and quality in data extraction. For Twitter in particular, tweet data can be loosely coupled and hard to examine using normal data extraction tools such as the Twitter API. The most optimal way in dealing with social media is by using custom-made libraries that can provide a layer of abstraction on top of raw information. We decided to use Tweepy as our data extraction tool due to its simplicity and efficiency. The framework is built under Python and the Twitter API. The Python environment is known for having many data manipulation APIs such as Matplotlib by which further consolidated the decision.

Three main functions of Tweepy were used in the data extraction.

For break tweet extraction, API.Search was used for such requests. The Search function is a powerful indexing function that allows the request service to query Twitter's backend servers for corresponding tweets in bulk. Additional parameters such as count and since_id are available as constraints for designated searches. However, due to the non-exhaustiveness of Twitter's search service and the extensive properties of the Tweepy API, tweets returned were within a weekly time period in which the preliminary data samples were unsatisfactory. Considering these downsides, the initial planned extraction size was further increased and no geolocation nor language restrictions was put into place. This approach expanded the scope of the response which in turn, enhanced the randomness of the raw data. The number of retrieved results in each response was controlled by the Cursor function that will be further illustrated in later sections.

After the retrieval of the break tweets, users are labeled by the screen name (account name) and tweets are identified using its unique identification number. Tweets from before and after the 'break' were required for the analysis of the trail of the user's emotions. The User information was retrieved via the API.user_timeline function and timeline was separated into prior and post by appending the unique identifier as max_id and since_id respectively. Taking into the account of computation and space complexity, limits were added to mitigate the time span of the service requests. This approach also augmented the precision in emotion prediction by confining emotion fluctuation in a shorter duration.

In addition to the retrieval functions, the Cursor function served as a utility method that was used not only in data extraction but also in data cleaning. It acted mainly as a wrapper function for other native Tweepy API functions that efficiently cutdown development duration by integrating pagination loops and provided a way to append supplementary parameters for better control over the content request. In order to retrieve the complete tweet information, Cursor function was used to enable extended mode for full-text retrieval. This was important in ensuring information completeness in tweet text filtering.

Since the definition of 'taking a break' on social media platforms can be loosely defined, we needed to come up with a way to clearly pinpoint such a topic in tweets as well as ensuring the subjective-objective relationship between the user and the break. The Twitter Search engine played a major role in filtering out unnecessary and unrelated tweets that are similar in meaning but differ in context. The search engine is able to break apart a single plain string query into key sections and apply regular expression search on to each key section. It was time-saving and simplified the filtering process. The preliminary tweet data was then sent to a custom

Python regular expression filter for an additional round of processing.

After obtaining an accurate list of users who are on a break or have already taken a break from Twitter, for the prior tweet data that was associated with these users a datetime constraint was placed on the response from the server and the quantity was gauged. Furthermore, the request for break tweets was also presumably time sensitive, however the Twitter API and hence the Tweepy API intentionally omitted such an option so no time restrictions or expansion in this case was feasible.

As mentioned previously, the Cursor function provided additional features for the query request. In the filtering process, using the native Python items function, the quantity of the user tweets was confined, which reduced the number of requests and the cap for tweets immediately followed the break laid the foundation for emotion prediction.

The final step was applying emotion RNN models on collected user tweets to generate the corresponding emotion vectors. The models were trained on character attributes hence no preprocessing of tweets in the semantic level was mandatory. The setting of the emotion predictor was configured to Multi-Class as the alternative mode Multi-Label setting would not be able to produce a probability table for each emotion indicators.

Although we originally applied three emotion RNN models, we decided on using one emotion RNN model in order to evaluate the tweets collected. We chose to use Plutchik's model which suggests eight primary bipolar emotions: joy versus sadness; anger versus fear; trust versus disgust; and surprise versus anticipation (Plutchik 1980). By categorizing emotions as opposites (Figure 1), Plutchik's model allowed us to categorize the sentiment of users' tweets as either negative or positive. Positive emotions were identified based on the following categories: joy, surprise, anticipation, and trust. Negative emotions were identified as either sadness, fear, disgust, or anger. We applied Plutchik's model to break tweets and then return tweets in order to compare the emotions of users and identify if there is an emotional effect caused by taking a break from Twitter.

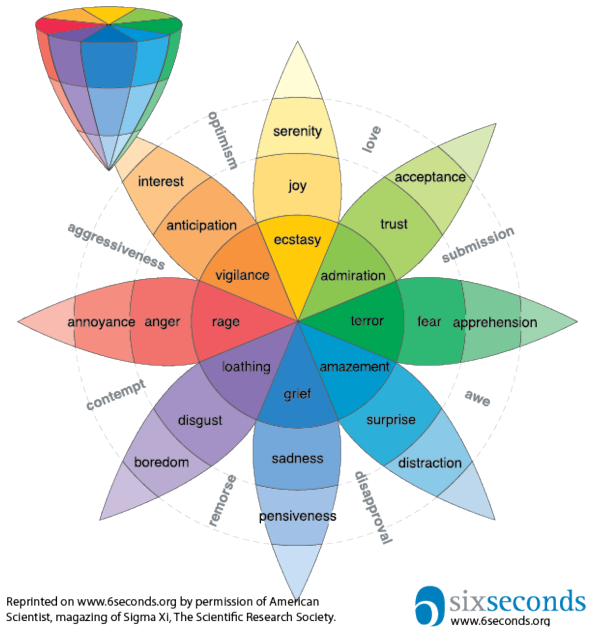

Figure 1: Plutchik's Model

4 Results and Evaluation

With the methods and techniques that we elaborated on in the previous section, we conducted experiments on Twitter data to find answers to the research questions that we posed. In this section, we will introduce the detailed settings and steps of our experiment.

First, we called Twitter API's search function using keyword query "take a break" and collected an initial set of 5000 datapoints. We realized that those raw data contained too much noise, and since we have no exact knowledge on how the API handles the query, we decided that we should create filters to ensure we get valid data.

Thus, we then applied our handcrafted regular expression filter to the 5000 data points to pluck out the tweets that we are confident are expressing the users' desire of taking a break from Twitter. We get 346 data points that pass the rigorous filter. Even though this filtering process might leave out false negative samples, we argue that precision is of more importance, and that 346 samples are enough for us to analyze our problems.

Next, we call Twitter API again to check if those break-takers have returned and collect their return tweets. In our experiment, 294 of the 346 users posted a return post. Those 294 users are consequently our main subject of attention. Subsequent steps will be mentioned in light of their related research questions.

RQ1: Is there a relationship between prior frequency of posts and the duration of break afterwards?

To get an answer to this, we first calculated the duration of a break for the 294 users that returned. Since the data points have a field "created-on," we were able plot a histogram (Figure 2) to look at how the distribution of break duration looks like. Unlike uniform distribution or normal distribution, our plot shows higher concentration on the left-hand side, which means that a major proportion of people claiming to take a break tend to return pretty fast. Yet we also notice certain populations that do keep away from Twitter for several days.

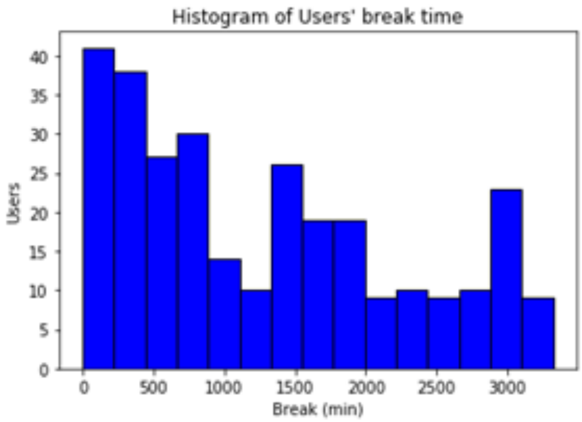

Figure 2: Histogram of Users' Break Time

Next, we decided on a criterion for users' prior frequency: the number of posts a week prior to the break tweet, capped at a maximum of 150. This data is collected by calling Twitter API to search each users' timeline and filter by time of creation. With this and the break duration calculated as above, we create a scatter plot (Figure 3) to gain intuition of their correlation or lack thereof.

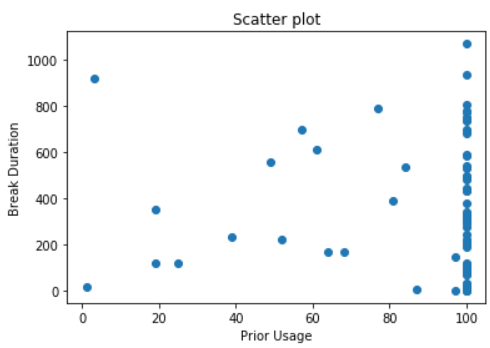

Figure 3: Break Duration vs Prior Usage

It is worth noting that the plot has a series of concentrated dots scattered on the x=150 axis, which is because we capped the maximum count of prior posts to 150. Despite this flaw, we agree that the dots show no clear and simple pattern: unlike our hypothesis, there is no clear indication of either positive relationship or negative relationship. So, we pronounce for our first research question that prior frequency of posts and the duration of break afterwards share no evident correlation.

Nonetheless, we arrived at a conclusion in a different perspective. By plotting the histogram of users' prior usage, as shown in Figure 4, we find that the majority of break-takers are heavy users of Twitter, which can be studied further later.

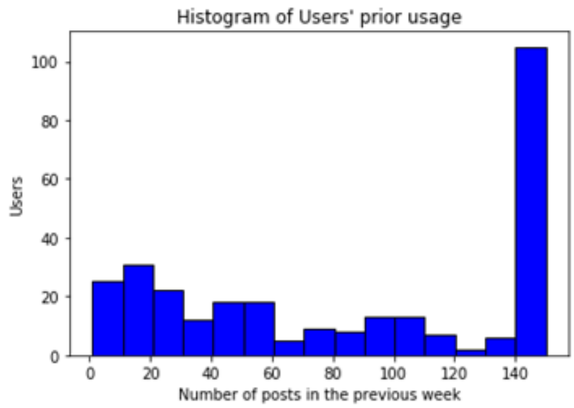

Figure 4: Histogram of Users' Prior Usage

RQ2: What reasons do people give in their posts when they decide to take a break?

With our collected break tweets at hand, we designed a classifier based on regular expression matching to identify the reasons break takers give in their farewell post. The regular expression filters are handcrafted and looked for key words that imply users' concern for privacy, data misuse, unproductivity, banality, addiction or external pressures.

The results show that only 12 users included a reason in their break tweets. Thus, we believe that most users do not explicitly say in their tweets the reason for their withdrawal from Twitter.

However, by looking at the posts that did include reasons, we were also able to confirm that the concerns we mentioned in our hypothesis existed. For instance, we have "Vivi7BTSArmyEGO" worried about addiction posting "I'm going to take a break from Twitter for a couple of days. I'll be back if there's something important and of course for the collab.

I just feel that I'm a bit addicted I want to calm down a little bit," and "B1uestk" expressing banality concerns in "I'm gonna take a break from twitter for a little while since I've just gotten tired of it recently. (I'll still be on discord though if anyone want to contact me)."

RQ3: What effect does a break have on the sentiment expressed in users' first post upon return?

For this question, we run the emotion predictor model on both the break tweets and the return tweets to capture a change in sentiment. For each tweet, we label it with the sentiment with highest probability, and construct a pie chart illustrating the distribution of emotions. Figure 5 corresponds to the break tweet, while Figure 6 corresponds to the return tweet, both of which are using Plutchik's model.

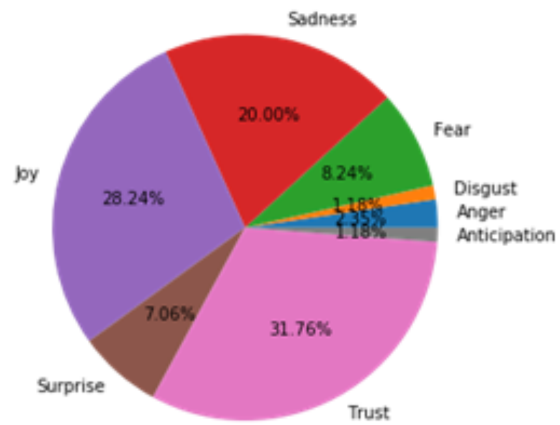

Figure 5: Plutchik's Model on Break Tweets

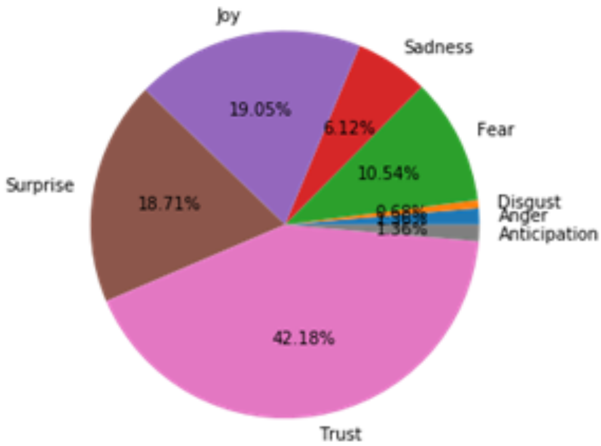

Figure 6: Plutchik's Model on Return Tweets

The overall percentage of negative emotions (sadness, fear, disgust, and anger) exhibited in tweets decreased from 32.95% across break tweets to 20.08% across return tweets. Further the overall percentage of positive emotions exhibited in tweets increased. This confirms our hypothesis that users' return posts would show more positive sentiments. Notably, trust increased from 31.76% to 41.18%.

5 Conclusion

While a major proportion of users who claimed to take a break on Twitter returned within hours or a few days at most, prior frequency of posts and the duration of break afterwards share no evident correlation. Most users did not explicitly state their reason for taking a break from Twitter and thus we were unable to determine a distribution of motivations for taking a break from the sample collected. We were able to identify that breaks positively affect users' emotions based on their return tweets.

6 Future Work

A larger sample size of Twitter users across a longer period of time may improve our results for future work. Although we were unable to identify a relationship between prior frequency of posts and the duration of a break, Twitter API's limit of 150 posts per week limited the measure of users' frequency of posts. By collecting tweets over the course of a month, we may be able to more clearly define users' frequency of posts, which could allow us to identify a relationship between prior frequency of posts and the duration of break afterwards.

Since most users did not explicitly state their motivations for taking a break from Twitter, we were unable to find a distribution of possible motivations suggested by Baumer et al. (2013). By collecting a larger sample size of tweets across future months, we may be able to collect more tweets in which users express their motivations for taking breaks.

In Stieger et al. (2013), Facebook users felt more inclined to stay on the platform if their Internet addiction is not as high and they express less concern over privacy and conscientiousness. Therefore, the increase in trust expressed in return posts found in our work may be based on users having less concern over privacy and conscientiousness. However, identification of motivations for users returning to Twitter would again require a larger sample size.

7 References

Baumer, E. P., Adams, P., Khovanskaya, V. D., Liao, T. C., Smith, M. E., Schwanda Sosik, V., & Williams, K. (2013, April). Limiting, leaving, and (re) lapsing: an exploration of Facebook non-use practices and experiences. In Proceedings of the SIGCHI conference on human factors in computing systems (pp. 3257-3266). https://doi.org/10.1145/2470654.2466446

Baumer, E. P., Guha, S., Quan, E., Mimno, D., & Gay, G. K. (2015). Missing photos, suffering withdrawal, or finding freedom. How Experiences of Social Media Non-Use Influence the Likelihood of Reversion. Social Media+ Society, 1(2), 2056305115614851. https://doi.org/10.1177/2056305115614851

Brubaker, J. R., Ananny, M., & Crawford, K. (2016). Departing glances: A sociotechnical account of 'leaving' Grindr. New Media & Society, 18(3), 373-390. https://doi.org/10.1177/1461444814542311

Maier, C., Laumer, S., Eckhardt, A., & Weitzel, T. (2012). Online social networks as a source and symbol of stress: an empirical analysis. https://pdfs.semanticscholar.org/d8b0/9272169acb2e026a61b73561d1bc0c0704f5.pdf

Perrin, A. (2015). Social media usage. Pew research center, 52-68. https://www.pewresearch.org/internet/2015/10/08/social-networking-usage-2005-201American Psychological Association. 1983. *Publica- tions Manual*. American Psychological Association, Washington, DC.

Plutchik, R., & Kellerman, H. (1980). *Theories of emotion, vol. 1 of emotion: Theory, research, and experience*. Academic Press.

Schoenebeck, S. Y. (2014, April). Giving up Twitter for Lent: how and why we take breaks from social media. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (pp. 773-782). https://doi.org/10.1145/2556288.2556983

Stieger, S., Burger, C., Bohn, M., & Voracek, M. (2013). Who commits virtual identity suicide? Differences in privacy concerns, internet addiction, and personality between Facebook users and quitters. Cyberpsychology, Behavior, and Social Networking, 16(9), 629-634. http://doi.org/10.1089/cyer.2012.0323

8 Contributions

Each team member contributed equally to all sections of the project including coding, data collection and results evaluation.