This is the authors’ final peer reviewed (post print) version of the item published as:


Available from Deakin Research Online:

http://hdl.handle.net/10536/DRO/DU:30072737

©2014 IEEE. Personal use of this material is permitted. However, permission to reprint/republish this material for advertising or promotional purposes or for creating new collective works for resale or redistribution to servers or lists, or to reuse any copyrighted component of this work in other works must be obtained from the IEEE.

Reproduced with the kind permission of the copyright owner

Copyright: 2014, IEEE
Analysis of Circadian Rhythms from Online Communities of Individuals with Affective Disorders

Bo Dao, Thin Nguyen, Svetla Venkatesh and Dinh Phung
Centre for Pattern Recognition and Data Analytics (PRaDA), Deakin University, Australia.
Email: {dbdao, thin.nguyen, svenkatesh, dinh.phung}@deakin.edu.au

Abstract—The circadian system regulates 24 hour rhythms in biological creatures. It impacts mood regulation. The disruptions of circadian rhythms cause destabilization in individuals with affective disorders, such as depression and bipolar disorders. Previous work has examined the role of the circadian system on effects of light interactions on mood-related systems, the effects of light manipulation on brain, the impact of chronic stress on rhythms. However, such studies have been conducted in small, preselected populations. The deluge of data is now changing the landscape of research practice. The unprecedented growth of social media data allows one to study individual behavior across large and diverse populations. In particular, individuals with affective disorders from online communities have not been examined rigorously. In this paper, we aim to use social media as a sensor to identify circadian patterns for individuals with affective disorders in online communities. We use a large scale study cohort of data collecting from online affective disorder communities. We analyze changes in hourly, daily, weekly and seasonal affect of these clinical groups in contrast with control groups of general communities. By comparing the behaviors between the clinical groups and the control groups, our findings show that individuals with affective disorders show a significant distinction in their circadian rhythms across the online activity. The results shed light on the potential of using social media for identifying diurnal individual variation in affective state, providing key indicators and risk factors for noninvasive wellbeing monitoring and prediction.

Index Terms—social media, circadian rhythms, psycholinguistics, affective disorders, online communities.

I. INTRODUCTION

Fundamental to many of our biological processes is the circadian rhythm, endogenously generated such as hormonal secretion, cardiovascular and respiratory metabolic functions; although external factors such as sunlight, temperature, sleeping habits and work patterns might have indirect influences [1]. It is characterized by daily repeating of 24-hour circle and is known to link directly to the patterns of brain wave activities, production of hormone, cell activities and other biological and physiological signals [2], [3]. It is also known that disruptions of chronobiological rhythms, such as circadian rhythms, cause destabilization in individuals with affective disorders, such as depression and bipolar disorders [4]. Hence it is of great interest to study circadian patterns for individuals with affective disorders.

Traditionally, these questions have been often examined at laboratory studies [2], [3], [5]. A key limitation of these traditional approaches, as highlighted by Golder and Macy [6], is that their conclusions are often drawn from data collected on a small homogeneous sub-population and they relied on the error-prone process of self-reporting, questionnaires to gather the ground-truth for data analysis. They further state that "...Researchers have acknowledged the limitations of this methodology [...] but have had no practical means for in situ real-time hourly observation of individual behavior in large and culturally diverse populations over many weeks". Echo on this sentiment, this paper uses social media data from thousands of public posts from LiveJournal weblog as a form of sensor data to understand this problem. In particular, we identify individual circadian rhythms of users in online communities who declared to have, or being affected by affective disorders.

This paper uses a large scale study cohort of data from nearly 10,000 people in 24 online communities declared to be related to affective disorder which include bipolar, depression, self-harm, suicide, and separation groups – for which we shall call them clinical groups. We analyze changes in hourly, daily, weekly and seasonal affect of these clinical groups in contrast with control groups of 23 generic online communities. We also measure affective rhythms including positive affect (PA) and negative affect (NA) of the individuals within both groups. By comparing the behaviors of the clinical groups and the control groups through circadian rhythms, our findings demonstrate that individuals with affective disorders show a significant difference in their circadian rhythms across online activities. It also demonstrates the potential use of social media as a sensor in identifying diurnal individual variation in affective state, providing a novel mechanism to extract key indicators and risk factors for understanding individualized wellbeing from online data for individuals with affective disorder such as depression or autism.

II. RELATED BACKGROUND

A substantial part of investigations on depression in psychology, psychiatry and sociolinguistics has been conducted to identify and understand interaction of affective disorders and social behaviors in individuals. For example, some core symptoms of depression showed circadian rhythm in their
clinical reflection such as diurnal mood variation [7]. In addition, diurnal mood variation of depressive symptoms is considered as a core symptom of major depressive disorder (MDD) [5]. It shows that depression is linked with mood swings in each individual. Therefore, laboratory studies have discovered that mood variability or diurnal mood variation appears to be the potential factor of diagnostic and treatment prediction for many depressed patients. Small changes in diurnal rhythms are found potentially to affect mood state of vulnerable individuals. Berk et al. [4] stated that the effect of daylight saving has impacted individuals with affective disorder, particularly bipolar disorder, increasing number of suicide in vulnerable people. Golder and Macy [6] emphasized that laboratory researches pointed out that the diurnal mood variation reflects the interaction of circadian rhythms with the duration of prior alertness or sleep of an individual. The authors also highlighted that experimental psychologists have repeatedly clarified both PA and NA show more psychometric measurement independently. Similarly, [5] found that the PA reflects a feeling of excitement, enthusiasm and activeness, whereas the NA shows distress including fear, shame and guilt.

Moving to studies in online setting, investigations have been conducted. Social media sites (e.g Facebook, Twitter and LiveJournal) provide platforms for people to express, exchange opinions and interact with others about their activities, their social interactions and the daily events around them. It gives opportunities for social/psychological scientists to study individual behaviors in the online context. For example, recently Ayers et al. [8], [9] found the circaseptan (weekly) rhythms in health considerations and the seasonal trends of seeking mental health terms via analyzing healthy behaviors on monitoring Google searching queries for healthy contemplations. In addition, current studies [6], [10], [11] investigated affective and sentiment analysis of online messages of individuals on social media. Particularly, these studies [10], [11] applied several methods to observe and extract diurnal rhythms of human behaviors based on their location or activities from unprecedented social media data scales. The study of Golder and Macy in [6] used Twitter messages as dataset to analyze changes in diurnal and seasonal mood rhythms in each individual around the world. The authors used a prominent lexicon for text analysis called Linguistic Inquiry and Word Count (LIWC) [12] to measure PA and NA from the message. They identified how PA and NA vary within individuals with daily and seasonal mood changes across diverse cultures.

Depression is currently a serious and widespread challenge in personal and public health. It is estimated that more than 350 million people of all ages suffering from depression [13]. Regarding to this prominent problem, many computational researches have paid more attention on using online social media platforms as a tool of mental illness measurement, surveillance and prediction. Indeed, recent studies [14], [15] examined the potential of using social media for identifying and predicting individuals with affective disorders, particularly focusing on a common mental disorder as MDD for individuals. The authors attempted to predict depressed users based on their tweets, including the diurnal rhythms of their posting. Moreover, postings of online people conveyed their thoughts and emotions of their daily activities or their happenings. These studies that the sentiment of emotion and language used in online messages may reveal feelings of individual on online communities. Furthermore, from web-blogs, inferred latent topics and linguistic features were used to determine if a blog post is made in an online depression community [16], [17]. Evidences of different mood swing patterns in low and high online social capital groups were studied in [18] using Bayesian nonparametric methods. Latent topics which are inferred from the corpus of posts were found to have stronger predictive power than linguistic styles when predicting if a post is written by a member from an online autism community [19]. In this investigation, we use data from LiveJournal weblog as a form of sensor to understand this problem. In particular, we identify circadian or circaseptan (weekly) or seasonal rhythms from online communities with bloggers who declared to have, or being affected by affective disorders.

III. Experiments

A. Data study cohort and analysis method

Data was collected from the LiveJournal (LJ) blogging sites for experiments in this paper. LJ was selected since it allows people in common of interest to form community of shared interests. It is suitable for observations of changes within the community context. In LJ, bloggers can either select a predefined mood of 132 common moods given by LJ or enter a free-text to label their posts at the time of writing, providing a chance for understanding affective aspects of communities. We construct two datasets for the experiments in the paper as follows.
Clinical dataset: We query all communities that contain the key word ‘depression’ in their interest description and have at least 200 posts. This process results in 24 affective disorder communities. These communities are further grouped into five categories according to the cohorts of bipolar, depression, self-harm, separation, and suicide, shown in Table Ia. These sub-groups are used as a clinical group in all experiments of the study in this paper. Table Ia lists statistics for these sub-categories.

General dataset: We also construct a general dataset that includes five sub-groups of general online communities from LJ community directory. We select a cut off to 23 communities as shown in Table Ib. Similarly, we show statistics for these communities in Table IIb. The earliest community in both datasets was created in 1999, we crawled the data until 12/2012, thus our datasets span nearly 13 years.

In addition to the emotion conveyed in the text of blog posts, a tagged mood allows us to reveal the user sentiment. Regarding to mood tags, we investigate the histogram of these predefined moods tagged in the posts of both two datasets to extract a set of 24 dominant moods, termed as the primary mood set. Figure 1 shows a cloud visualization of 24 primary moods tagged to blog posts made by both two groups. It can be seen that tagged moods are diverse reasonably, including moods tagged to blog posts made by both two groups. It can be seen that tagged moods are diverse reasonably, including moods tagged to blog posts made by both two groups. It can be seen that tagged moods are diverse reasonably, including moods tagged to blog posts made by both two groups.

To do this, we divide time axis into the regular interval of one hour. For example, for community \( j \) with \( m_j \) users, we extract the number of messages which have been posted during a fixed hour, say \( t \) where \( t = 0, 1, 2, \ldots, 23 \). The distribution of posts at this time \( t \) is computed as:

\[
post_j(t) = \frac{v_j(t)}{m_j}
\]

where \( v_j(t) \) is the mean number of messages posted at time interval \( t \) within community \( j \). To describe this more precisely, let \( x_{ji}(t, d) \) be the number of posts user \( i \) within community \( j \) has made at \( t \)-hour interval for \( d \)-th day where \( d = 1, \ldots, D_j \) with \( D_j \) denotes the number of days this community has been active.

\[
v_j(t) = \frac{\sum_{d=1}^{D_j} \sum_{i=1}^{m_j} x_{ji}(t, d)}{D_j}
\]

For each community \( j \), we further construct the data matrix \( C(j) = [c_{d,t}(j)] \) of \( D_j \times 24 \) to represent the data recorded where \( c_{d,t}(j) = \sum_{i=1}^{m_j} x_{ji}(t, d) \). These posting behavior data matrices are visualized in Figure 2 & Figure 3, where we can
D. Seasonal Posting Patterns

When the importance of sleep and the biological clock for affective rhythms on diurnal rhythms might extend to seasonal patterns, the authors in [8], [6] found that there are no any effect of day length on either PA or NA within individuals with seasonal changes in the length of the day at a given location over the year. Therefore, for our seasonal analysis, we group postings made per a quarter over a year (Jan-Mar, Apr-Jun, Jul-Sep, Oct-Dec). Similar to the diurnal analysis, we observe the change of seasonal variation through affective expression on both groups.

IV. RESULTS

A. Analysis of negative and positive affective

In the general group, all sub-groups show dominant positive emotion across all days of week and all hours. Figure 6b shows

\[ post_j (d) = \frac{v_j (d)}{m_j} \]  

where \( v_j (d) \) is the mean number of blogs posted at day \( d \) of the week within community \( j \). To describe this more precisely, let \( y_{ji} (t, d) \) be the number of posts user \( i \) within community \( j \) has made at \( t \)-hour interval for \( d \) day where \( d = 1, \ldots, 7 \) and \( t = 0, 1, 2, \ldots, 23 \).

Generally, all subgroups of both clinical and general groups posted almost their messages during afternoon and evening periods.

We describe daily patterns in user’s posting to uncover weekly (circaseptan) rhythms, where evaluating individual posts reveals both the user’s contemplations and when they are committing actions toward behavior changes in the online communities.
that three groups such as Fashion, Food and Pets seem more positive than two other groups, Parenting and Technology. Diurnal rhythms across these groups are not indistinguishable when the patterns are mirrored in the figure.

In contrary, posting patterns within individuals in general communities, dominant negative affective are shown in all clinical groups. The negative emotion in posting of people in three groups (Depression, Selfharm and Bipolar) rises in the early morning, declines during the afternoon, and rebounds peaks at late of evening and in the night period. Strangely, Figure 6a shows that individuals in both Separation and Suicide groups expressed their complicated emotion. These individuals reveal that they have just got only few hours with positive emotion at the morning and at afternoon period while they feel negative feelings at almost other times of the day.

Regarding to weekly (circaseptan) rhythms, bloggers from online affective disorder communities show their negative affective all the time of the week. Specifically, Separation and Suicide groups still show their negative emotion all day of the week. Particularly, Depression has a peak of negative emotion on Monday, then declining and rebounding a peak on Thursday. Selfharm obtains their maximum negative emotion on Tuesday. Interestingly, three groups have reached a lower in their negative emotion on Wednesday, Thursday and Friday.

Likewise, bloggers from general online communities show their dominant positive emotion during the week as seen in Figure 7b with a lower level of positive affective at the middle of the week. People in both Parenting and Technology has a
lower positive emotion than others, indicating their weekly life is rough enough with difficulties.

For seasonal patterns, we examine the difference in affective processes over four quarters of the year of both two study cohorts. As expected, all 5 general online groups show a dominant positive emotion over all seasons of the year with a highest state on the fourth quarter of the year as shown in Figure 5b. It also shows that Parenting and Technology categories still have the lower positive emotion than others in the general cohort, indicating the difficulties of parenting duties or of technology troubles in Parenting and Technology sub-groups, respectively. Meanwhile, for clinical group, Figure 5a indicates that with the exception of the negative affective on the first quarter (Jan-Mar) of the year, only Suicide group shows their positive emotion all remaining quarters of the year with a peak of positive emotion during the second quarter (Apr-Jun). Other clinical groups have a variety of negative emotion season by season over the year.

**B. Analysis of affective valence rhythms on mood tags**

To inspect the affective rhythms on the usage of 24 primary moods, the valence of mood usage each hour can be examined. Figure 8 shows the difference in the proportion of valence of the moods tagged during daily posting between clinical and general groups. From 1:00 to 7:00, the proportion of valence of moods tagged in posts of the clinical group is higher than those in the general group. However, since the beginning of morning (7:00) the valence of the general group increases until 23:00 (midnight), excepting at 11:00 and 16:00 with a lower valence than those in the clinical group.

We investigate the usage of mood tags in both clinical and general groups. For each mood in 24 primary moods, we show the numbers of post tagged with that mood. Figure 9 shows the difference in using 24 primary moods between two groups. Negative valence moods (in left) are marked in red (or lighter color in gray-scale) and positive valence moods are in blue (or darker). During active time, people in clinical group tagged more negative valence moods than those in general group, while bloggers in the general group selected almost positive valence moods to tag to their posts. In addition, clinical groups used more two positive valence moods (e.g. contemplative and crazy) for their posts than others in general groups.

Furthermore, we investigate circadian rhythms of mood tagged by bloggers with and without affective disorder from clinical communities and general online communities. As we expected, the people in the clinical groups show their negative
valence moods across almost 24-hours of the day with the same dominant mood ‘depressed’ as seen in Figure 10a, excepting at 6:00 with ‘sad’ mood. This is an interesting sign for more analysis on affective disorder online communities. Similarly, for the general online group, their users post their messages with a variety of positive valence moods during the day. However, at the late of the night period (4am and 5am) people show their dominant negative valence moods (e.g. confused and tired moods).

V. CONCLUSION

The circadian system regulates 24 hour rhythms in biological creatures, impacting mood regulation. The disruptions of circadian rhythms in individuals with affective disorders cause their destabilization. Hence it is of great interest to examine circadian patterns for individuals with affective disorders. With the advent of social media, we now have large communities of bloggers with depressive conditions expressing their thoughts. In this paper, we use the data from thousands of public posts from LiveJournal weblog as a form of sensor data to understand this problem. In particular, we investigate a large scale cohort of data study from 24 LiveJournal online affective disorder communities. Then, we analyze changes of these clinical groups in contrast with control groups of 23 general communities in hourly, daily, weekly and seasonal patterns. We also measure affective rhythms of the individuals within both communities. The findings present that individuals with affective disorder from online communities show their distinguished circadian rhythms across the online activity. The investigation also gives us confidence that the potential of using social media in detecting diurnal individual variation in affective state, providing a novel mechanism of extracting key indicators and risk factor for online early warning systems for affective disorder such as depression, bipolar, suicide or autism.

REFERENCES