European Journal of Mental Health 9 (2014) 220–230
DOI: 10.5708/EJMH.9.2014.2.5

György Csepeli[*] & Richárd Nagyfi

FACEBOOK DIAGNOSTICS

Detection of Mental Health Problems Based on Online Traces[**]

(Received: 6 May 2013; accepted: 19 February 2014)

Users of the social media network communicate with each other 24 hours a day. Their presence on Facebook leaves traces. Based on the analysis of the distribution patterns of the traces users had left, this paper aims to explore how mental states can be detected among Facebook users. Results of the analysis raise the possibility of the development of innovative methods for screening persons with mental problems among consumers and producers of online social media.

Keywords: social media, Facebook, scale-free distribution, mental health problems, HCSM (healthcare social media), depression, Big Data

Facebook-Diagnostik: Erschließung von mentalhygienischen Problemen mithilfe von Spuren im Internet: Die Benutzer der Plattformen sozialer Netzwerke können rund um die Uhr miteinander in Kontakt bleiben. Bei der Nutzung von Facebook hinterlässt man allerdings auch unwillkürlich Spuren. Die Studie berichtet über ein Experiment, bei dem aufgrund von Abweichungen dieser Aktivitätsspuren von der Potenzverteilung ein Einblick in den psychischen Zustand der Facebook-Benutzer gewonnen werden konnte. Diese Forschungsergebnisse könnten es ermöglichen, die mentalhygienischen Probleme der aktiven Benutzer sozialer Netzwerke automatisiert, online zu diagnostizieren.

Schlüsselbegriffe: soziale Netzwerke, Facebook, Potenzverteilung, mentalhygienische Störungen, HCSM (healthcare social media), Depression, Big Data

[*] Corresponding author: György Csepeli, Interdiscplinary Doctoral Program at the Faculty of Social Sciences of Eötvös Loránd University, Pázmány Péter sétány 1/a, H-1117 Budapest, Hungary; csepeli.gyorgy@gmail.com.

[**] This article has been prepared in the the framework of the grant titled 'PIAC_13-1-2013-0205 (Prediktív e-kereskedelmi megoldások kutatása online és offline tartalmak forgalmazása során keletkező nagy tömegű adattartalmak feldolgozását segítő fúziós módszertan fejlesztésére)'.

ISSN 1788-4934

1. Introduction

Humans have always been communicating with each other but the process of direct, mainly oral communication has remained traceless in everyday life up until the latest ages (BUDA 1988). The development of information technology in the 20th century has made the recording of direct human communication possible, but the recording of the acts of communication has generally been bound by the constraints of the situation, and the analysis of the data of the recorded video and audio materials has been limited by the objectives of recording. The emergence of broadband internet has resulted in a radical change in all aspects of the process of human communication. People have been able to widen their circles of communication, the number of partners has increased considerably, the line between 'private' and 'public' has become blurred and the traditional communication roles that had formerly divided participants of public social communication into 'producers' and 'consumers' has disappeared in the communication space created by broadband internet (CASTELLS 2013). The forms of social media have become the platform where people can express freely what they think, feel and wish to do. As a result of the current changes in information technology mass communication goes beyond the traditional division between producers and consumers of media content. The internet has made public communication possible for anybody, at any time and about anything.

People logged into Facebook and other social media websites can communicate with each other 24 hours a day. Their topics are varied. Most of them are personal while others reflect public interest. Anyone can take a topic selected from the ones discussed by traditional mass media or can pick up something that has been stored on the internet. Facebook is not just a platform of online communication, it is simultaneously a social network service. Users of Facebook must register before communicating on the site. The users create a profile of their own, label other users as friends. The number of Facebook users is well above one billion.

Facebook as a social service provider allows the formation of networks of users who mark their relationships as friendships. Friendships between Facebook users can be characterised typically as weak ties showing low values on measures of the amount of time, the emotional intensity, the intimacy and the reciprocal services which are generally seen as major dimensions of any type of interpersonal relationship (GRANOVETTER 1973). The nature of online friendships, however, differs from the nature of offline friendships. One can argue that online friendships by definition are weak because of the lack of physical contact. Nevertheless, online friendships might involve commitments of time and emotion and reciprocal services just as offline friendships do. According to the results of NEWCOMB's classical research, friends tend to be more similar to each other than mere acquaintances (1961). Newcomb found that the beliefs and attitudes of initial strangers tend to converge as a function of the progression of the acquaintance process. Consequently we might argue that the users who bond by online relationship will be more similar to each other than the users who are not connected. Thus, if online ties connect P to O1 and

P to O2, both O1 and O2 will likely be similar to P and probably will be similar to each other. Similarity of the network members can be predicted by Fritz HEIDER's theory of cognitive balance too (1958). Moreover, if online ties between P and O1 and P and O2 exist, then the emergence of dissimilarity between them would introduce strain into the network since O1 will want his own attitudes and beliefs to be congruent with the attitudes and beliefs of his friend P, and naturally with his attitude of his friend O2. On the basis of the congruency principle we can expect not only the similarity of attitudes and beliefs among the members of the network, but the similarity of their moods and mental states as well.

The relationship between social capital and computer-mediated communication has been investigated recently by a study of ELLISON and her colleagues (2007). The researchers asked college students about the ways they use social media and measured their social support stemming from strong ties and weak ties. The results of the study showed a positive correlation between Facebook participation and both varieties of social ties. According to these results Facebook usage can be considered as a predictor of bonding social capital which comes equally from offline and online relationships. Consequently, we can safely assume that there is an overlap between the two sorts of relationships (ELLISON et al. 2007).

Many people visit Facebook just to see what is going on. Other people use Facebook to promote social networking. Most of the friends are people with whom the user has never had an actual relationship. Social media, including Facebook, leads to less genuine empathy (MCCLINTOCK GREENBERG 2013). Most of the psychological studies on Facebook users focus on the negative psychological effects of exposure to Facebook demonstrating that the more people have used Facebook during a period of time, the more they have become depressed (KROSS et al. 2013).

The concern of this paper, however, is not to investigate the link between depression and Facebook usage. Our concern is to explore mental disorders including depression by analysing the behaviour of Facebook users. The research on the nature of the relationship between social media usage and depression usually focuses on the comparison of users who classify themselves as 'depressed' or 'non-depressed'. In the course of investigating the level of social activity, the presence or frequency of negative emotional outbursts, the level of self-attentional focus, and the share of weak and strong ties in their network is measured. Researchers claim that by leverage of the multiple types of measures obtained they can build a Major Depression Disorder (MDD) classifier that can predict whether an individual is vulnerable or not to depression.

There is extensive literature on the problem of predicting MDD on the basis of psychological, social psychological and sociolinguistic data emerging as a consequence of presence on the Internet. All these data, however, come from depression screening tests, self-reported information. There have been data collected from social media sites such as Twitter or Facebook. The total number of posts, the mean number of posts per user in a given period of time, and the variance of the number of posts per user were measured (CHOUDHURY et al. 2013).

The research that we are reporting on focused also on the prediction of mental disorders including MDD. Online communication however – just like offline communication – cannot be conceived without taking into account the social network within which the communication takes place. Consequently, we have had to deal with the behavioural patterns of the members in the individual user's social network as well.

2. Our research questions

Can we make inferences about the mental state of persons on the basis of the traces they leave on Facebook? This question can obviously be asked only about those who are regularly present on Facebook, react to the content they are exposed to by clicking the 'Like' button, share content with others, produce content themselves and share it with others, send messages, receive messages and participate in the life of communities organised on Facebook. However, we were not concerned about the total number of activities of the individual user, we have not calculated the mean number of likes per user over a period of time, and we were not interested in the variance of number of Facebook activities per day per user. It was not the quantity of Facebook activities but the pattern of the intermittence of active and inactive periods over time that we were interested in.

A Facebook user can choose from many activities. In our research, only the activities that are visible on the user's feed were taken into account.

The tracking of activities in everyday life, as we have already mentioned in the introduction, is an unsolved issue. Our words and acts tend to disappear without a trace. However, all activity produced on Facebook (and on any ICT-mediated communication platform) will be recorded. Consequently, presence on Facebook does not remain traceless. The users who log in to Facebook leave traces and create data as long as they are there. The question is: what can we do with the huge amount of data generated by the users? Is there any thread that can help us get out of the labyrinth created by Big Data?

Based on the ideas of Albert-László Barabási (2010) and Péter Csermely (2008) we found that the 'scale-free' phenomenon will be the Ariadne's thread that will help us escape the labyrinth created by the mass of data preserving the 'behavioural fingerprints' of users. We have learned from Barabási and Csermely that complex networks of any sort are characteristically scale-free. Such networks have a relatively large number of high-degree nodes and the degree distribution follows a power law function (Yule–Simon distribution). The degree distribution ratio of nodes is independent from size, and high-degree nodes are also typically connected to each other. Barabási's tests on a variety of networks have shown that the majority of self-organising networks are characteristically scale-free. Society is typically full of self-organising networks, and the human nervous system is no different. Péter Csermely quotes the catchy definition of Benoît Mandelbrot, the founder of fractal

theory who named scale-freeness the 'Matthew effect' – even though Matthew only quoted Christ who stated that 'whoever has will be given more' (Mt 25:29).

In self-organising networks those who have will always be given more. This is the law of the scale-free distribution of goods. The law works reversely as well: everyone who has not, will not be given. According to Csermely, the scale-freeness occurs not only in space but also in time. Benoît Mandelbrot refers to the Bible again when he calls the temporal effect of scale-freeness the 'Joseph effect'. The story of Joseph refers to the seven lean and seven fat years in Egypt (Gen 41).

Against this background, we thought that by analysing the mass of data stemming from the 'behavioural fingerprints' of Facebook users, we could begin solving this task using the phenomenon of scale-free distribution. We assumed that the distribution of the data produced by the activity of Facebook users can be described by a scale-free distribution curve.

However, a question arises: what to do with those who are active on Facebook but the pattern of their activities cannot be described by the curve of scale-free distribution? Looking for the answer to this question, we found important hints in Albert-László Barabási's work titled *Bursts* (2010), where he describes the results of a Japanese research team and argues that based on the scale-free nature of human motion, depression is detectable purely by monitoring the individuals' wrist movements. The Japanese researchers analysed locomotor activity data, capturing even slight bodily acceleration counts in a continuous fashion. Motion sensors were affixed to the wrists of twenty-six people. The sensors recorded the smallest movements. As Barabási states, the researchers have shown that the length of resting periods, when the subject's hand does not move, generally tends to take a power law distribution (Nakamura et al. 2008). The majority of the resting periods lasted from seconds up to a few minutes. Between these, intervals of several hours without any movement occurred as well, which corresponded to periods of sleeping, relaxation or meditation. Among the 26 subjects, however, there were some whose movements were characterised as intermittent. These individuals were depressed in the clinical sense. Barabási states that the resting periods of healthy individuals lasted approximately seven minutes on the average, while those of depressed patients were over fifteen minutes. Moreover, the scaling exponents unique to all power distributions were higher in the case of healthy individuals than in the case of the depressed ones (2010).

Alteration in locomotor activity was found as one of the major signs of psychiatric disorders including depression. One of the better known symptoms of depression is *psychomotor retardation*, involving a slowing down of movements. The results of the study showed that the locomotion activity of persons with depression in contrast with that of healthy control subjects exhibited bursts in the counts and the cumulative distribution of resting period durations deviated from the universal power law observed on healthy control persons. The visually clear disorders of the sequence of movements indicated mental problems, shock or stress.

In our experiment no motion sensor was installed on the subjects' hands. The individuals' hand movements were analysed from the data collected when they se-

lected the 'like' option while logged into Facebook. Beginning with our acquaintances, using the snowball method, we reached a total of 195 persons for whom we developed a questionnaire application on Facebook. This application, before start-up, requested access to the sociologically relevant personal data available on Facebook as well. The data were collected in aggregated form, using one-way and two-way encryption algorithms. In the questionnaire application we asked questions accepted by the profession, also found in the *Depression Inventory* developed by Beck (BECK et al. 1996). This questionnaire measured depression, anxiety and stress level based on self-reporting. In addition, other questions were asked that served a better understanding of Facebook presence.

When completing the questionnaire, an automatic data-mining algorithm was running on the user's computer in the background. This script collected aggregated information about the user's past activity on Facebook. The analysis aimed to compare the results of the *Depression Inventory* and the curve representing the activity on Facebook in the previous three months. By comparing the results of the *Depression Inventory* and the pattern of the curve representing the activity of the user, we were able to see the correspondence between the mental state and the curve of activity in each user's case. The curve representing the activity of the mentally healthy user followed the pattern of scale-free distribution. In contrast, if the curve deviated from the scale-free distribution pattern, it was considered as a clear evidence of mental disorder (depression, anxiety or stress).

Our questionnaire was completed by a total of 195 people. In addition to gender and age, we asked for data on the place of residence, the language used on Facebook and the number of the user's Facebook friends.

Nearly half of the completed questionnaires were useless for different reasons. Many people just clicked over it, there were people who knew the questions and the 'right' answers and in some cases, Facebook refused to release data due to various reasons. Fortunately enough, in some cases after collecting some additional information, it was possible to find certain correlations from these questionnaires as well.

Finally, in the case of usable, successfully completed questionnaires we succeeded to carry out investigation into the relationship between the activity curve of presence on Facebook and mental state.

In *Figure 2* below we can see the curve of a 24-year-old, 'healthy' male. His curve roughly matches the ideal curve.

Figure 3 shows a 50-year-old, moderately depressed and stressed, slightly anxious female's curve. The curve is sunken.

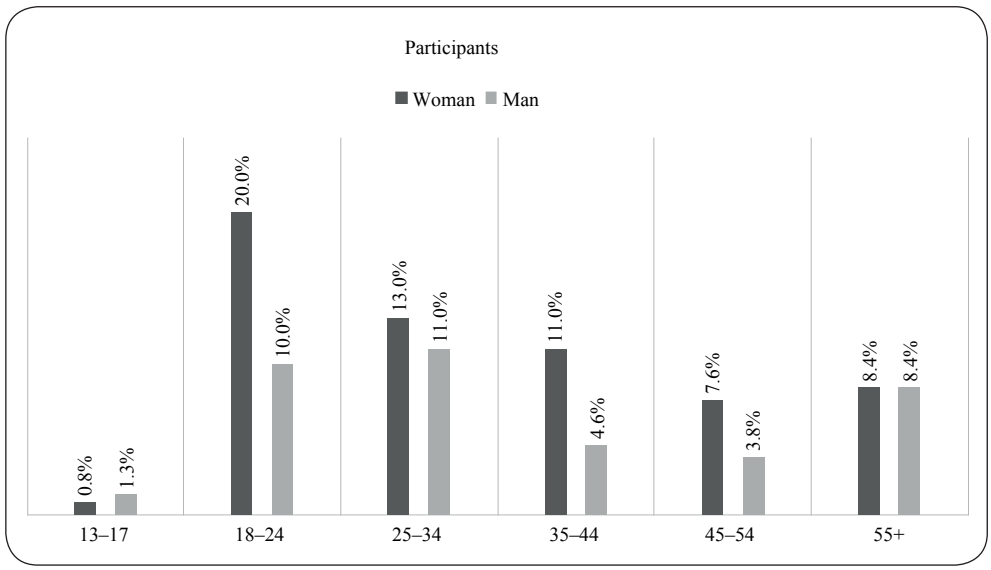

Figure 1
Breakdown of the experimental subjects' gender and age

Figure 2
Facebook diagnostics: detection of mental health problems
based on changes in individual Facebook usage frequency: 24-year-old male

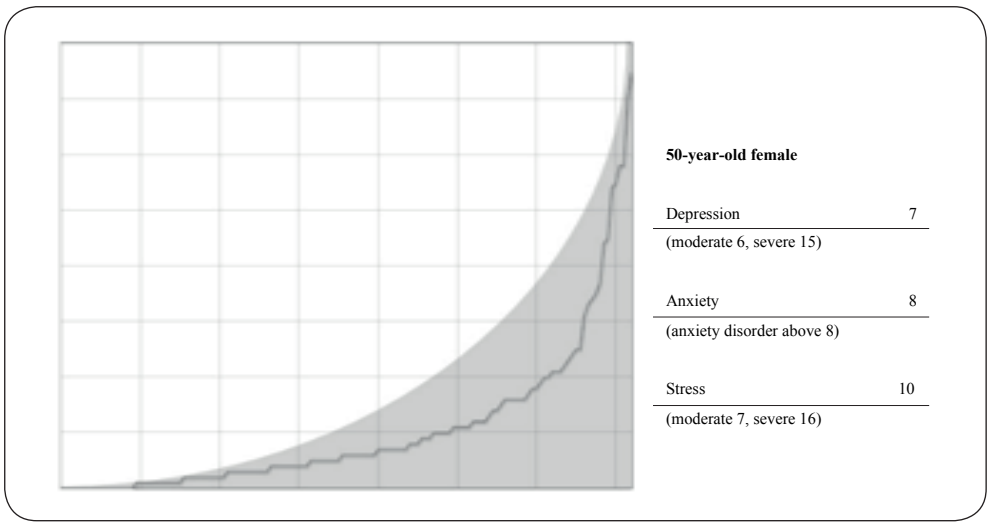

Figure 3
Facebook diagnostics: detection of mental health problems
based on changes in individual Facebook usage frequency: 50-year-old female

Figure 4 shows a 17-year-old, severely depressed, anxious and stressed male's curve. The theoretical, light blue area is not presented. His curve is nearly perpendicular to the axis. This means that after long periods of rest, every once in a while, very active periods occur in the person's life.

(The results, naturally, were not always so spectacular and easily discernible, and in order to understand the problems of the subject better, it was necessary to analyse other features of the graph as well). As we have seen, the analysis of curves mostly made the detection of depression, anxiety or stress possible.

The activity curve of the presence on Facebook, similarly to the curves obtained in the Japanese experiment mentioned by Barabási, has clearly shown the alternating carefree and 'low-flying' periods in the case of those struggling with mental problems. Our method enabled us to eliminate the effects of individually varying lengths of periods between activity and passivity. In all cases of deviation from the scale-free curve we managed to find a correlation between the shape of the 'like-curve' and mental problems.

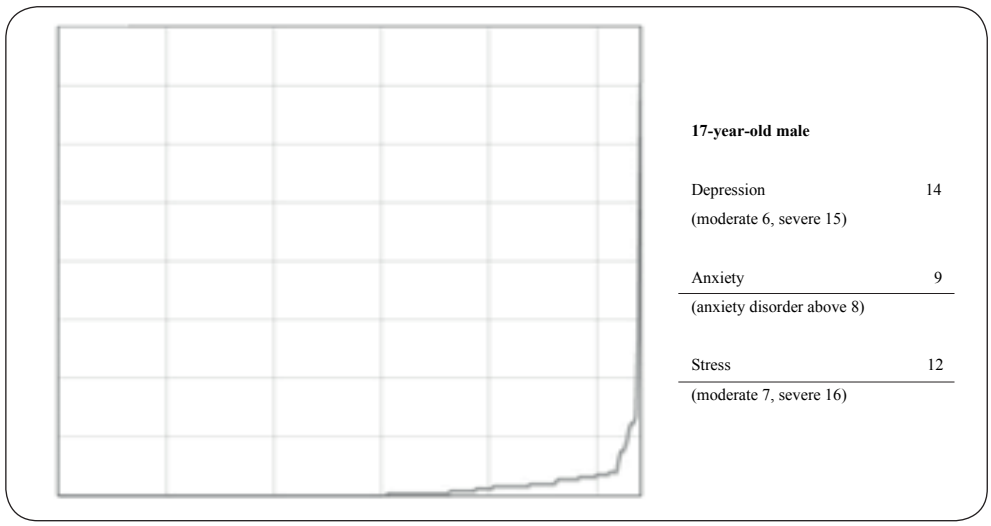

Figure 4
Facebook diagnostics: detection of mental health problems
based on changes in individual Facebook usage frequency: 17-year-old male

Naturally, we were not able to give accurate diagnosis without knowing other background variables. It seems to be particularly important that no curve indicating mental disorders was found isolated. The user's mental disorder was found not to be independent of the user's network. The friends of Facebook users showing mental problems struggled with disorders as well, while the friends of healthy individuals were found typically healthy.

'Offline' friends and 'online' friends, naturally, are different by nature. As we have already mentioned, 'online' friends do not necessarily know each other personally. In online social life we get in contact with people who are far from us in terms of physical distance but they are close to us in terms of values, attitudes and mentality. People with similar mental problems find each other on the internet without even being aware of the underlying reason that brings them close. The principle of congruence, proven by HEIDER (1958) and NEWCOMB (1961) in offline social life, prevails in online social life as well.

The investigation of Facebook activity (assuming voluntary participation) combined with answers to accompanying questions could make it cheaper and more effective to establish an early warning system of mental disorders. Provided that the mental disease has already been developed, the system could orient the patient in finding professional help and support.

3. Future perspectives

Based on the results of this research we find it worthwhile to carry out a comprehensive study aimed at collecting data created by the regular usage of social media. Our method envisages the development of a self-monitoring early warning system of mental disorders based on the recording of real-time real behavioural data stemming from the use of social media. In contrast with similar efforts, our method does not need measurement of symptoms of MDD. The method is focusing simply on recording and evaluating the 'behavioural fingerprints' of the individual users who can be alerted immediately if some risk of MDD has been detected.

As a result of a series of experiments carried out in a representative panel of volunteers, we would be able to identify patterns of mental health and mental illness on a mass scale that would revolutionise the approach of public healthcare organisations dealing with mental problems. Enormous progress could be achieved through a comprehensive research of the communication activity of the members of social media networks. Not only mental health problems but any other illnesses could be detected by a network-based medicine. Moreover, we believe that analysis of the ICT-mediated communication behaviour can result in the development of methods of automated mental health tracking. We are seeking to exploit the potentials of social media in identifying the stable and the unstable causes of mental disorders. As a result of the development of these new methods for public healthcare, the risks of self-aggression will likely be diminished.

References

Barabási, A.-L. (2010) *Bursts: The Hidden Patterns Behind Everything We Do, from Your E-mail to Bloody Crusades* (New York: Penguin).

Beck, A.T., R.A. Steer & G.K. Brown (1996) *Manual for the Beck Depression Inventory-II* (San Antonio: Psychological Corporation).

Buda, B. (1988) *A közvetlen emberi kommunikáció szabályszerűségei* (Budapest: Tömegkommunikációs Kutatóközpont).

Castells, M. (2013) *Communication Power* (Oxford: UP).

Choudhory, M. de, M. Gamon, S. Counts & E. Horvitz (2013) *Predicting Depression via Social Media* (Redmond: Association for the Advancement of Artificial Intelligence) retrieved 5 May 2013 from http://research.microsoft.com/pubs/192721/icwsm_13.pdf.

Csermely, P. (2008) *A rejtett hálózatok ereje* (Budapest: Vince).

Ellison, N.B., C. Steinfeld & C. Lampe (2007) 'The Benefits of Facebook "Friends": Social Capital and College Students' Use of Online Social Network Sites', *Journal of Computer Mediated Communication* 12, 1143–68.

Heider, F. (1958) *The Psychology of Interpersonal Relations* (Hillsdale: Erlbaum).

Granovetter, M.S. (1973) 'The Strenght of Weak Ties', *Americal Journal of Sociology* 78, 1360–80.

Kross, E., Ph. Verduyn, E. Demiralp, J. Park, D.S. Lee, N. Lin, H. Shablack, J. Jonides

& O. YBARRA (2013) 'Facebook Use Predicts Declines in Subjective Well-Being in Young Adults', *PLoS ONE* 8, No. 8 (DOI: 10.1371/journal.pone.0069841).

MCCLINTOCK GREENBERG, T. (2013) 'Facebook Depression: Real Friends and Social Media Friends', *Psychology Today,* published 26 Aug 2013, retrieved 28 Jan 2014 from http://www.psychologytoday.com/blog/21st-century-aging/201308/facebook-depression.

NAKAMURA, T., T. TAKUMI, A. TAKANO, N. AOYAGI, K. YOSHIUCHI, Z.R. STRUZIK & Y. YAMOMOTO (2008) 'Of Mice and Men: Universality and Breakdown of Behavioral Organization', *PLoS ONE* 3, No. 4 (DOI: 10.1371/journal.pone.0002050).

NEWCOMB, Th.M. (1961) *The Acquaintance Process* (New York: Holt, Rinehart & Winston).