

COLLECTING DATA OF SNS USER BEHAVIOR TO DETECT THE SYMPTOM OF ITS EXCESSIVE USAGE – DESIGN OF DATA COLLECTION APPLICATION

Ploypailin Intapong¹, Tiranee Achalakul² and Michiko Ohkura³

Graduate School of Engineering and Science, Shibaura Institute of Technology¹

Department of Computer Engineering, King Mongkut's University of Technology Thonburi²

College of Engineering, Shibaura Institute of Technology³

nb15508@shibaura-it.ac.jp¹, tiranee@cpe.kmutt.ac.th², ohkura@sic.shibaura-it.ac.jp³

ABSTRACT

Social Networking Sites (SNSs) have become a popular mode of communication where groups of people virtually meet and interact with other people with similar interests. The use of SNSs dramatically increased over the past several years. People spend so much time on SNSs than expected, sometimes in uncontrollable way. Many studies suggest that the excessive use of SNSs can cause emotional relational, health-related, and performance problems including the potential of addictive behavior. For this reason, it is necessary to detect the symptom of the excessive SNSs usage. In order to detect the symptom, collecting data of SNSs usage behavior is essential. The article aims to design an application for collecting data of SNSs usage behavior. A questionnaire is used for gathering self-report data. It is designed to have three parts, which are personal information, SNSs usage information and SNSs addictive behavior information. Internet Addiction Test (IAT) and Bergen Facebook Addiction Scale (BFAS) are employed to reflect the six core components of addictive behavior (salience, mood modification, tolerance, withdrawal symptoms, conflict and relapse). We will modify IAT and BFAS to be used for SNS with retainment of the original concept. Then, we will conduct a preliminary experiment for content validity and usability of the questionnaire. From previous studies, the sample size of surveying is small and specific. Therefore, crowdsourcing technique will be employed to distribute the survey through an implemented web questionnaire system. However, the self-report data may be inaccurate. Thus, we will retrieve data from SNSs via API and network traffic between user and SNSs. The obtained data from application will be analyzed to clarify the relation between them and the characteristic of SNSs usage at present. These results will be applied to detect the symptom for preventing excessive use of SNSs.

1. INTRODUCTION

In 2014, the use of SNSs dramatically rose 12% from 1.9 billion to 2.1 billion active users, an astonishing total that equals almost 30% of the world's population (We Are Social, 2015). People use SNSs for many activities and purposes. However, some people spend so much time on SNSs and use them in ways that may become excessive and/or addictive.

The excessive use of SNSs can cause emotional, relational, health-related, and performance problems, especially the risk of addiction. For these reasons, we should understand how people interact with SNSs. Collecting data is an important first step for analyzing individual usage behaviors of SNSs.

This article designs an application for collecting SNS data. We employed many methods in designing. We will analyze the obtained data from applications to clarify the relationships and the characteristics of SNS usage. We will also use our analysis results for detecting symptoms and preventing the excessive use of SNSs.

2. LITERATURE REVIEW

2.1 Social Networking Sites

Social networking sites are virtual communities where groups of people with similar interests can create individual public profiles and interact with others (Kuss & Griffiths, 2011). However, researchers may use the different words that refer to SNSs. Ellison & Boyd (2013) suggested that *"Terminology varied widely with the interchangeable use of 'social networking sites', 'online social networks' or even simply, 'social network' to refer to a diffuse – and sometimes improbable – range of sites and services."* Boyd & Ellison (2007) defined social network sites as follows: *"web-based services that allow*

individual to (1) construct a public or semi-public profile within a bounded system, (2) articulate a list of other users with whom they share a connection, and (3) view and traverse their list of connections and those made by others within the system.” However, Ellison & Boyd (2013) argued that the term “Social Network Sites” is more accurate than other terminologies because it emphasizes the role of networks, unlike previous online interaction space.

2.2 Negative Consequences of Excessive SNSs Usage

Many studies have suggested that excessive SNS usage leads to various negative consequences. According to Kuss & Griffiths (2011), excessive SNS usage causes relationship problems and impacts academic and work performances. For instance, people who spend excessive time on SNSs appeared less involved in their real life communities and students who used the internet for studying had lower grades because they were easily distracted and exercised poor time management. J. Al-Menayes (2015) concluded that Kuwaiti university students who spent time on SNSs instead of studying had lower grades. In Thailand, many teenagers suffered from negative effects of excessive SNS usage, such as lack of sleep, academic performance, inappropriate manners, negative emotional expressions, impairment of family and social functions, and mental health problems (Phanasathit, 2015). Excessive SNS usage delays bedtimes and reduces sleep quality (Andreassen et al., 2012).

2.3 Data Collection Methods

Since SNSs have become so popular, many researchers have sought to understand user behavior in them by collecting the data of SNS usage behavior. Many types of data and collection methods exist. Abdesslem, et al. (2012) concluded that SNS data can be collected from three sources: SNSs, surveys, and monitoring applications.

The most common way to retrieve data directly from SNSs is to use the Application Programming Interface (API) provided by each SNS. However, some data available on SNSs cannot be collected through APIs. Some studies employ automated script that automatically retrieves data using HTTP requests/responses. Additionally, to understand browsing activities, some researchers collect data by tracing the network traffic from ISP providers or gathering data by social network aggregator websites. Shahrak & Mohammadi (2014) collected data from social network aggregator websites to study the effects of SNSs on e-commerce interactions in middle-east countries.

Another method to collect data is to directly ask users about their experiences through questionnaires. Most researchers have employed this method in their studies for various purposes, even though research in human behavior areas argued that self-report measures are inaccurate when compared to actual behavior (Abdesslem et al., 2012) and (Junco, 2013).

Furthermore, deploying a monitoring application is a way to obtain SNS user behavior. Abdesslem et al. (2012)

argued that this method provided more flexibility and privacy. Nevertheless, it has some limitations. Researchers need to install applications in user devices to get data manually from user devices. Even though some cloud-based applications exist, most are commercial and limit the number of monitoring devices.

3. THE DESIGN OF DATA COLLECTION APPLICATION

3.1 Conceptual Design

From a literature review, many studies identify the negative consequences of excessive SNS usage (Kuss & Griffiths, 2011), (J. Al-Menayes, 2015), (Phanasathit, 2015) and (Andreassen et al., 2012). To detect the symptoms of excessive SNS usage, we must understand the SNS behaviors of users as the first step. Various methods have been employed for collecting SNS data to precisely understand how people interact with them. For this research, we will collect SNS data from self-report questionnaires, SNSs, and ISP logs. However, different methods might be employed in future research. Figure 1 shows our conceptual design for collecting SNS data.

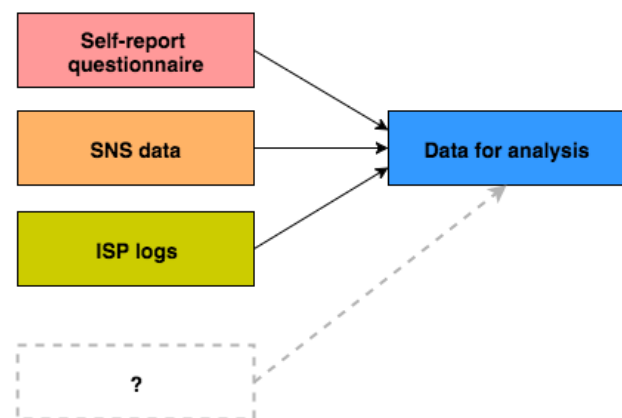


Fig. 1 Conceptual design for collecting SNSs data

We will use self-report questionnaires for gathering the user experiences of SNSs. SNS data will be retrieved from each SNS by APIs. In the beginning, we will focus on the most popular SNSs: Facebook and Twitter. We will also use logs from ISP providers and will trace the network traffic data from the computer center of the Shibaura Institute of Technology (SIT) in Japan. We will analyze the obtained data to understand the current user behavior.

In previous studies, the sample size of the self-report questionnaires was small and specific (Kuss & Griffiths, 2011). Therefore, in order to get a sufficient amount of data, we will employ an effective crowdsourcing technique that reduces time and cost (Brabham, 2008) and distribute our survey through an implemented data collection application.

A data collection application is a web-based application designed for aggregating SNS data from questionnaires and SNSs. Figure 2 shows an overview of

this application that can also be referred to as client-server architecture. The application has two parts: questionnaire and quizzes. Users can answer the questionnaires and/or quizzes online, and then the data are collected on a database. Quizzes are employed to engage and motivate users to retrieve data from SNSs. Each quiz queries data through APIs provided by each SNS. The application allows users to share the results of quizzes on SNSs to spread the contributions by crowdsourcing.

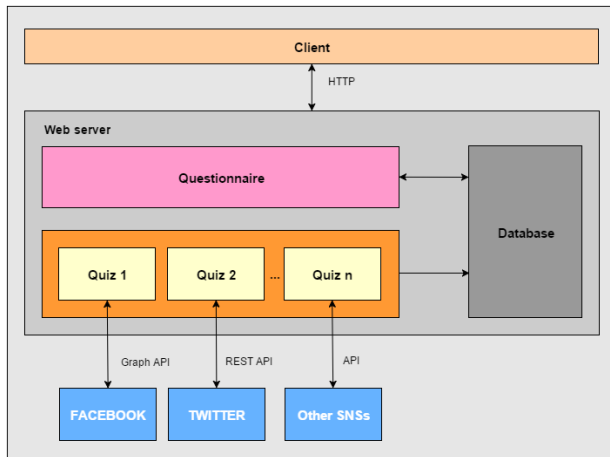


Fig. 2 Overview of data collection application

3.2 Questionnaire Design

Our social network usage questionnaire has three parts: personal information, SNSs usage information, and SNSs addictive behavior information in English language.

According to previous studies, SNS usage differs by age, gender, and country. Thus, the first part collected demographic information. To understand how people interact with SNSs, the second part asked about their SNSs, for instance, purpose, time spent, length, frequency, place, and device. In the third part, we employed the Internet Addiction Test (IAT) (Young, 1998) and the Bergen Facebook Addiction Scale (BFAS) (Andreassen et al., 2012) to reflect the six core components of addictive behavior. We modified IAT and BFAS for SNSs by retaining the original concept.

4. PRELIMINARY EXPERIMENT

To evaluate the design of our social network usage questionnaire, we conducted a preliminary experiment of its content validity and usability.

4.1 Method

Seventeen Shibaura Institute of Technology (SIT) students participated in this experiment: eight males and nine females, between 20-30 years of age. Their nationalities were Brazilian, Malaysian, Thai, Vietnamese, Indonesian, and Japanese. They completed the social network usage questionnaire and circled any words that confused them. We also allowed them to skip any question that they could not answer for content validity. Finally, participants completed an evaluation questionnaire that

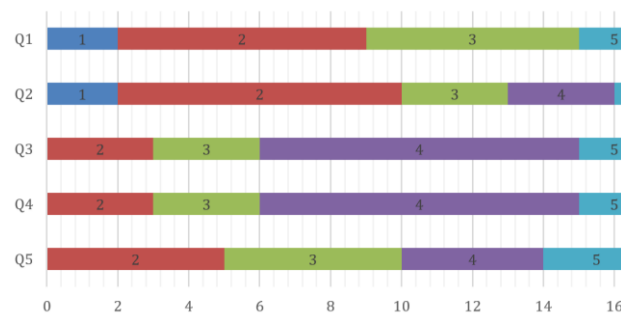
used a 5-point Likert scale that ranged from “strongly disagree” to “strongly agree” as feedback for questionnaire evaluations. Additionally, we observed and recorded the start to finish times for each part of the questionnaire while participants answered them.

4.2 Result

The average time for answering the questionnaires was 9.42 minutes. Most participants spent more time in the third part (SNS addictive behavior information) than on the first and second parts. Moreover, participants only circled confusing words and skipped questions in the third part. The most frequently skipped question was “How often do you prefer the excitement of SNSs to intimacy with your partner?” from IAT. The word “intimacy” was most frequently chosen as being confusing; that word is also in the question that had the highest skipping frequency.

The feedback for the social network usage questionnaires from the participants is shown in Figure 3. Most participants did not think that the questionnaire or the time spent answering was too long. Most agreed that the questions were clear and easy to understand without complicated syntax. 7 of the 17 participants understood each question clearly without any confusion.

Questionnaire Evaluation



1 = Strongly disagree, 2 = Disagree, 3 = Undecided, 4 = Agree, 5= Strongly agree

Note

Q1: Length of questionnaire is too long

Q2: Time spend on answer questionnaire is too long

Q3: Questions are clearly and easy to understand

Q4: Questions are grammatical and not contains complicated syntax

Q5: I can understand each question clearly without any confusion

Fig. 3 Evaluation of questionnaire

5. DISCUSSIONS AND CONCLUSION

In this article, we proposed an application design for collecting SNS data to understand how people interact with SNSs. We employed survey-based method for gathering user experiences. We designed self-report questionnaires and experimentally validated the content and the usability of them. Our results showed that most participants were satisfied with them. The time required for answering them and their length was appropriate. Also,

they clearly understood the questions. However, some confusing words might cause skipping answering in the third part where we employed IAT and BFAS. Probably these words were technical or uncommon. Also, the participants were not native English speakers. Unfortunately, we cannot rewrite or add descriptions to each question because that action might distort the original meaning. Thus, we will allow the participants to skip questions to avoid imprecise data. Then we will use an imputation technique to replace the missing data with mean values.

In conclusion, we employed many methods to design a data collection application. Self-report data may be inaccurate. Thus, we employed a retrieved data from SNSs by APIs and a tracing network traffic method. We designed self-report questionnaire and experimentally evaluated them. Our results showed that questionnaire has validity and usability as an instrument for gathering data. Next we will modify our questionnaire and develop the data collection application. We will also analyze the obtained data from the application to clarify the relation between them and the characteristics of current SNS usage. These results will be applied for detecting the symptoms and preventing the excessive use of SNSs.

REFERENCES

- Abdesslem, F., Parris, I., and Henderson, T., Reliable online social network data collection, *Computational Social Networks*, Springer London, pp. 183–210, 2012.
- Andreassen, C. S., Torsheim, T., Brunborg, G. S., and Pallesen, S., Development of a Facebook addiction scale, *Psychological Reports*, vol. 110, no.2, pp. 501–517, 2012.
- Boyd, D. M. and Ellison, N. B., Social network sites: definition, history, and scholarship, *Journal of Computer-Mediated Communication*, vol. 13, no. 1, pp. 210–230, 2007.
- Brabham, D. C., Crowdsourcing as a model for problem solving: An introduction and cases, *Convergence: The International Journal of Research into New Media Technologies*, vol. 14, no. 1, pp. 75–90, 2008.
- Ellison, N. B., and Boyd, D. M., Sociality through Social Network Sites, *The Oxford Handbook of Internet Studies*, pp. 151–172, 2013.
- Jamal, J. A., Dimensions of social media addiction among university students in Kuwait, *Psychology and Behavioral Sciences*, vol. 4, no. 1, pp. 23–28, 2015.
- Junco, R., Comparing actual and self-reported measures of Facebook use, *Computers in Human Behavior*, vol. 29, no. 3, pp. 626–631, 2013.
- Kuss, D. J., and Griffiths, M. D., Online social networking and addiction-A review of the psychological literature, *International Journal of Environmental Research and Public Health*, vol. 8, no. 9, pp. 3528–3552, 2011.
- Phanasathit, M., Validation of the Thai version of Bergen Facebook Addiction Scale (Thai-BFAS), *J Med Assoc Thai*, vol. 98, no. 2, pp. S108–S117, 2015.
- Shahrak, M. Z., and Mohammadi, S., Middle East User Navigation in Online Social Networks and Interactions in E-Commerce, an Analogy, *Advances in Computer Science: an International Journal*, vol. 3, no. 2, pp. 32–36, 2014.
- We Are Social. Retrieved from <http://wearesocial.net/>, 2015.
- Young, K., Internet addiction: The emergence of a new clinical disorder, *CyberPsychology & Behavior*, vol. 1, no. 3, pp. 237–244, 1998.



Ploypailin Intapong received her B.Eng. (2009) and M.Eng. (2011) degree in computer engineering from King Mongkut's University of Technology Thonburi (KMUTT), Thailand. She is currently a Ph.D. candidate at Shibaura Institute of Technology, Japan. Her research interests include user behavior and social networking sites.



Tiranee Achalakul received her B.S. degree in computer engineering from King Mongkut's Institute of Technology Ladkrabang, Thailand, in 1994 and completed Ph.D. in computer engineering from Syracuse University, New York, in 2000. She is currently an associate dean in education development for the faculty of engineering at King Mongkut's University of Technology Thonburi. Her research interests include parallel and distributed computing, software engineering, Cloud and GPU technology, and data mining applications.



Michiko Ohkura received her B.S. (1976) and M.S. (1978) degrees in Mathematical Engineering and Ph. D. (1995) in Advances Interdisciplinary Studies from the University of Tokyo, Japan. She worked for some companies including Central Research Laboratory, Hitachi Ltd., and is currently a professor, Department of Information Science and Engineering, the Shibaura Institute of Technology, Japan. Her current research interests include human-friendly interactive systems and creation of Kansei values.