Using Data Mining Technology to Explore Internet Addiction Behavioral Patterns

Mu-Jung Huang
Information Management
National Changhua University of Education
Changhua, Taiwan
mjhuang@cc.ncue.edu.tw

Mu-Yen Chen
Information Management
National Taichung University of Science and Technology
Taichung, Taiwan
mychen@nutc.edu.tw

Chin-Chun Cheng
Information Management
National Changhua University of Education
Changhua, Taiwan
initiald781207@gmail.com

The purposes of this study were to explore psychological satisfaction and emotional reaction of Internet users through emotional perspectives and to discuss whether Internet use behaviors would lead to addiction to the Internet. From previous literature and studies, it was found that most studies explored this topic by testing hypotheses. The study used data mining to identify association rules among affective ambivalence, Internet use behavior and Internet addiction.

Online and paper questionnaires were distributed for this study. Online questionnaires were put on BBS, Facebook and major forums; paper questionnaires were distributed via convenience sampling. A total of 565 questionnaires were recovered. Among these, 502 copies of the questionnaires were valid, making the effective response rate about 88%. It was found from the affective ambivalence that different use behaviors would result in different affective states. Different individuals would also show different behavior and creativity.

Keywords—Internet Addiction; Affective Ambivalence; Internet Use Behavior; Association Rules

I. INTRODUCTION

Internet was originated from the "ARPA NET" (Advanced Research Project Agency Network) developed by the U.S. military in 1969. However, it was not until 1990 that Berners-Lee created the World Wide Web project and made simplified use of the network so prevalent in this era. The Internet became an important source for modern people to access information. Whether in terms of communication, data transmission and exchange, online shopping, video conferencing or online entertainment, Internet has become an integral part of our life, making information exchange more convenient. In recent years, with the advancement of Internet technology, the mobile network technology also becomes more mature. Users can access the Internet via 3G wireless network regardless of their location through a variety of mobile devices, such as smart phones or tablet computers.

Although Internet has brought a lot of convenience to the public, it also causes many negative effects to personal physiology, psychology as well as the society, such as hackers, web spoofing, network crimes, cyber personal data crimes, and cyberbullying. One of the hottest issues is Internet addiction. The concept of "Internet addiction" was brought forth by an American doctor, Ivan Goldberg, in 1995. It refers to "a psychological and physiological dependence on pleasure brought by using the Internet." Being addicted to the Internet will bring many negative effects to people, such as less communication with their families and poor academic performance. In recent years, there were also a lot of cases of sudden death due to excessive Internet use. A new term " phubber" was coined because of mobile network. Phubbing all the time causes various harms to our bodies. Studies done by Ferris (1996) [1] found that the core symptoms and negative impacts of Internet addiction included: knowing that one should limit the time to surf the Internet, but failing to control the desire to surf the Internet; showing physical or psychological discomfort when not surfing the Internet; the desire to surf the Internet cannot be satisfied and the time spent online was becoming longer.

Many factors lead to Internet addiction nowadays. One of the reasons is accessibility. Since cheap or free Internet access via Wi-Fi or 3G mobile network is provided on the campus and in businesses, it is easy to access the Internet through various channels. Another possible reason is that people want to make friends and expand relationships through the Internet. The characteristics of anonymity on the Internet make people feel a sense of security, thus reducing the risk of developing virtual relationships. Some people with poor social skills can develop virtual friendship. As a result, incidences of Internet addiction are increasing every year. The study expected to explore users' Internet use behavior, and emotional patterns at that moment by researching the topic of Internet addiction. Most previous studies discussed Internet addiction from the perspective of affective ambivalence. This study integrated the two factors to

978-1-4799-5375-2/14/$31.00 ©2014 IEEE

explore Internet addiction and to find a different conclusion. Therefore, the purposes of this study included the following:

- To identify the relationship between affective ambivalence and Internet use behavior via data mining
- To identify the relationship between affective ambivalence and Internet addiction via data mining

II. Literature Reviews

A. Affective ambivalence

The term “affect” was first derived from the Latin word "affectus," meaning individuals' “mental state" or "mindset" [2]. The mood and emotion were used as basis for definition. Various changing moods and emotions can be divided into positive affectivity and negative affectivity. Affective characteristics decided how a person acknowledges or perceives external environment or stimuli [3,4,5], and further affected personal feeling and thinking as well as evaluation and feeling towards the external environment. Finally, people would adopt different behaviors and attitudes in the face of events [6,7].

In 2006, Fong further explained positive and negative affectivities and brought forth the concept of "affective ambivalence.” The study concluded that when individuals feel intense positive or negative affectivity within a short period of time, their sensitivity and feelings to the surroundings would be improved, producing different effects on individual’s behavior and creativity. Most people experienced various affective states in everyday life, and there were also a lot of situations that would make people experience positive and negative affectivities in a short period of time [8].

"Positive affectivity" reflected the extent of happiness and pleasure of an individual in an environment [3]. Positive affectivity was more like feelings that were conducive to positive emotion, making an individual feel happy and healthy towards all things and interpret things around from a positive and optimistic viewpoint. Additionally, behaviors reflecting positive affectivity were more enthusiastic, energetic, active, alert and self-efficacious. In the process of social interaction and achievement of goals, people would be in a more pleasant mood [4, 9]. Previous studies found that people with higher positive affectivity reacted in a more positive way towards all stimulations. Their feelings, thinking and behaviors would be more positive. On the contrary, people who were not as positive would show characteristics of lethargy and fatigue. And in the process of getting things done, they would approach things with a cool emotion, and even showed tendency towards and feelings of frustration sometimes [4].

"Negative affectivity" referred to negative emotion experienced at certain time or in certain context, particularly fretfulness, anxiety and irritability. They would see themselves with negative emotion [3]. Liaw et al. defined it as individual’s general subjective distress and unpleasant emotion, which would further affect the individual's cognition, self-concept and values [3]. Such emotion would make people feel unhappy and irritating, such as anxiety and neuroticism. People who were very negative would experience subjective distress, depression, nervousness, discomfort and anxiety when they were in an environment with non-obvious pressure sources [10].

In 1988, Watson et al. used the 60 words used by Zevon & Tellegen (1982) to describe positive and negative emotions to develop the "Positive and Negative Affect Scales (PANAS) [11]." The purpose was to measure the emotional experience of an individual at different time (including now, today, in recent days, in the past few weeks, in the past year and under general condition). Through factor analysis, two dimensions - positive affectivity and negative affectivity were generated. Types of positive and negative affectivities were sorted out by using the scale proposed by Watson and the research results found by Fleur & Steenkamp (2005). From which, topics suitable for the study were identified and incorporated into the scale [12].

B. Internet Use Behavior

For modern people, the Internet was something they will come into contact with every day. It provides many virtual activities such as sending and receiving E-mails, shopping online, chatting online, searching the latest information, and so on. In the online world, people can communicate more freely due to the anonymous feature of the Internet [13]. As for the investigation of Internet behavior, BCG (Boston Consulting Group) conducted a survey on Internet behavior in 2009 and classified a total of 14 Internet use behaviors: instant messaging, online music, news reading, online video, search engine, online games, e-mail, blogs, social networking sites, e-commerce, online banking, bulletin boards/forums, job hunting and others.

Tosun & Lajunen (2009) explained from psychological viewpoint that motivation to use social network would affect one’s Internet use behavior [14]. The study explored social motivation from the concept of "real me" and found the following motivations: to establish a new relationship, to have friends on the Internet, to keep in touch with friends far away, and to assist face to face interpersonal relation. It was also proven that the Internet can replace face-to-face social support; therefore, from this perspective, if the purpose of using the Internet was to satisfy social needs, it was possible that these Internet users might lack it in real life.

C. Internet Addiction

The term "Internet Addiction" was developed in modern time. Researchers began to believe that people would be addicted to the Internet and therefore designed studies to prove it [15]. Young (1999) published related research and definition for Internet addiction, he concluded that individuals would be easily attracted by characteristics of the Internet such as virtual, real-time, and anonymous in order to escape life stress and frustration [16]. They choose to indulge in the Internet world in order to give vent to their real-life discomfort and to seek satisfaction.

Young (1999) classified the types of Internet addiction according to the content of the Internet users used and found five categories [16]:

- Internet Sex Addiction: To be attracted by sex-related Web contents or indulged in erotic activities derived from the Internet.
- Internet Relationship Addiction: To be indulged in interpersonal activities on the Internet.
- Internet Compulsion: Include getting indulged in online games, online gambling, online shopping and trading activities.
- Internet Information Overload: To indulge in searching and collecting information on the Internet.
- Internet Computer Addiction: To indulge in activities and needs of computer operations related to the Internet.

D. *Data mining*

Fayyad et al. (1996) suggested that the knowledge discovery process could be divided into five steps, namely (1) data selection, (2) data processing, (3) data transpose, (4) data mining, (5) data evaluation and interpretation [17]. Finally, useful knowledge could be produced. Data mining is one of the tools in the process of knowledge discovery. Data mining is a popular issue and research tool in recent years in the field of database. It is to use statistics and machine learning algorithm to identify commercially valuable knowledge and rules from the large amount of data. It is also applied in the automation business strategy.

- Classification: Classification is the most common model in data mining. Classifying things is the most original pattern of forming rules of knowledge by humans. Classification models can be further divided into decision trees, clustering algorithms and Naive Bayes methods.
- Cluster: Based on similarity, similar information will be put in the same group. Complex and large amount of data will be significantly simplified by clustering. This process is called clustering.
- Association Rules(AR): It is to identify correlation between valuable items from a large number of data. It is mainly used to process and analyze a large number of disorganized data, and to simplify data to become statistic data that is easy to be observed, understood and explained.
- Sequential: Also known as time series. It is to find out the order of things in a homogeneous grouping, and to identify the life cycle of things. Different responses in different point of time are used to assist decision analysis.
- Estimation: It is to predict the relevant property of continuous values. Representative calculus methods include regression analysis and neural network analysis.

III. RESEARCH METHODOLOGY

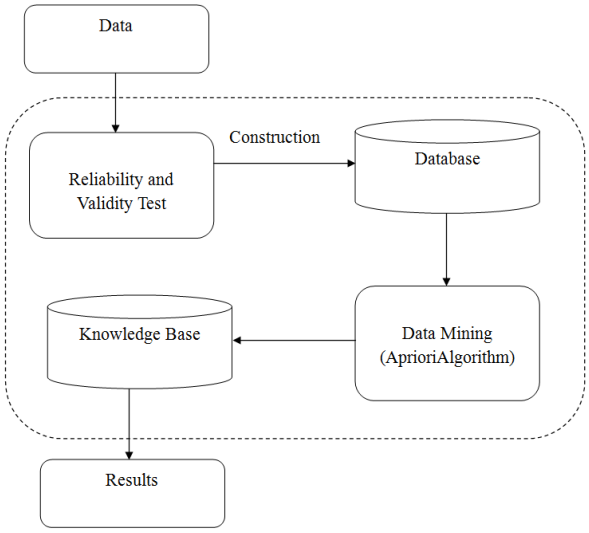

Figure 1. Research Architecture

A. *Research Architecture*

This study used affective ambivalence, Internet use behavior, and Internet addiction as the four variables to further discuss the correlation between Internet addicts and their Internet use behaviors. SPSS Clementine 14 was used to analyze association rules; SPSS 18 statistics software was used to analyze reliability and validity analysis of data processing. The main research framework was shown in Figure 1.

The study samples and data sources were mainly Internet users. Online and paper questionnaires were distributed. Online questionnaires were placed on BBS, Facebook as well as the major forums; paper questionnaires were distributed via convenience sampling.

For data processing, the recovered questionnaires first went through data preprocessing, and then were converted into formats that could be used by statistical methods and association rules. And the data of single variable was presented by statistical methods of frequency and percentage.

B. *Research Scale*

The study used four scales, namely affective ambivalence scale, Internet use behavior scale and Internet addiction scale. Likert five-point scale was used in the questionnaire. The response “Always” got five points; “often” got four points; “usually” got three points; “occasionally” got two points and “seldom” got one point.

- Affective ambivalence: Positive and Negative Affect Scales (PANAS) developed by Watson, Clark & Tellegen (1988) was used [5]. The study used the most representative questionnaire from PANAS as the basis for positive and negative affect questions to establish the scale of this variable in the study.
- Internet use behavior: Includes Internet use habits, referring to users’ most common Internet use habits, such as Internet browsing, blogs, online shopping, P2P

downloading, online games and so on; This aspect cited definition of Internet use behavior developed by Boston Consulting Group in 2009 as basis to create a question in the questionnaire.

- Internet addiction: The Internet Addiction Test Scale was translated from the "Internet Addiction Test (IAT)" developed by psychologist Young (1998) to determine whether the subjects reached the standard of Internet addiction [16]. Studies done by Widyanto & McMurran (2004) indicated that IAT had good internal consistency [18]. Cornbach's α value was 0.82.

IV. Experimental Results

A. Demographics

This study distributed a total of 565 questionnaires. After deleting those with missing data, extreme values, and ineffective questionnaires, 502 valid questionnaires were left. The effective questionnaire recovery rate was about 88%. As for gender in the collected questionnaires, there were 284 men and 218 women. The proportion did not vary greatly. This study used convenience sampling and random sampling to collect questionnaires, so most subjects belong to the age groups of 16 to 20 years old, 21 to 25 years old, and 26 to 30 years old.

B. Reliability and Validity Analysis

The following showed Cronbach's α test results of each dimension of the question items in the study. The first one was affective ambivalence (Table 1) with a total of six questions. Q1-1 ~ Q1-3 were about positive affect; Q1-4 ~ Q1-6 were about negative affect.

Tabel 1. Affective ambivalence

Items	Cronbach's α
Q1-1	0.748
Q1-2	0.769
Q1-3	0.778
Q1-4	0.790
Q1-5	0.815
Q1-6	0.807

Questions in this part of the survey were to investigate Internet activities that users were often engage in (Table 2). Q2-1 was about Internet browsing; Q2-2 was about blogs and Facebook; Q2-3 was about communication software; Q2-4 was about online games; Q2-5 was about e-mail, Q2-6 was about online shopping.

Tabel 2. Internet activities

Items	Cronbach's α
Q2-1	0.866
Q2-2	0.865
Q2-3	0.883
Q2-4	0.901
Q2-5	0.873
Q2-6	0.866

A total of 14 questions asked about Internet use behavior (Table 3). Based on the literature, Internet use behaviors were classified into the following three categories: information gathering behavior, recreational use behavior, social networking behavior; questions about information gathering behaviors were Q3-1 ~ Q3-4; Q3-5 ~ Q3-10 were about recreational use behaviors; Q3-11 ~ Q3-14 were about social networking behaviors.

Tabel 3. Internet use behavior

Items	Cronbach's α
Q3-1	0.936
Q3-2	0.936
Q3-3	0.937
Q3-4	0.944
Q3-5	0.942
Q3-6	0.941
Q3-7	0.936
Q3-8	0.942
Q3-9	0.936
Q3-10	0.941
Q3-11	0.943
Q3-12	0.942
Q3-13	0.937
Q3-14	0.939

There were 20 questions about Internet addiction (Table 4). The degree of Internet addiction caused by users' Internet use behaviors were measured.

Tabel 4. Internet addiction

Items	Cronbach's α
Q4-1	0.939
Q4-2	0.938
Q4-3	0.940
Q4-4	0.940
Q4-5	0.938
Q4-6	0.939
Q4-7	0.943
Q4-8	0.939
Q4-9	0.940
Q4-10	0.939
Q4-11	0.938
Q4-12	0.939
Q4-13	0.939
Q4-14	0.940
Q4-15	0.941
Q4-16	0.938
Q4-17	0.939
Q4-18	0.940
Q4-19	0.941
Q4-20	0.941

Test results from Table 1, 2, 3, and 4 showed that all Cronbach's α values were greater than the highest reliability ($\alpha > 0.7$). Thus, it proved that samples collected by the study had consistency and stability.

C. Apriori Algorithm Mining

a) Internet Use Behavior and Affective Ambivalence

Take the two rules with the highest confidence as an example. "Information gathering = high and Internet social networking = high; highly positive affectivity will be created." "Information gathering = high and recreation = high; highly positive affectivity will be produced." It means that when users receive new information, or engage in social or entertaining activities to release emotion, they will have highly positive emotions (happiness, comfort and satisfaction). Therefore, frequency of use is pretty high.

As for negative affectivity, "Internet Socialization = high, and recreation = high, highly negative affectivity will be created." "Recreation = medium, and social networking = medium; highly negative affectivity will be produced." What was different from positive affectivity was that data collection about Internet use behavior does not show any rule. From this, it was found that when users were in a highly negative affectivity (stress, fear or shyness), they would want to release their negative affectivity through social networking, entertainment and other ways, and did not want to get new messages, which was not helpful to their mood. As a result, information gathering did not appear in the negative affectivity rule.

b) Internet Addiction and Affective Ambivalence

Observing users' affective status from Internet addiction, the benefits of Internet behavior were as shown in the previous section. Users would indulge in the online world because of the Internet's virtual, real-time, and anonymous characteristics.

V. Conclusions

This study expected to find the relations and impacts among affective ambivalence, Internet use behavior and Internet addiction. It was hoped that the research results might be helpful for future studies. From previous literature and studies, it was found that most past studies adopted the research method of regression testing by assuming that these dimensions had already causal relationship among them. Therefore, this study assumed that there was no causal relationship among these dimensions and used correlation rule in data mining to explore the four dimensions. It was hoped that undiscovered relation may be found by a different way.

This study analyzed Internet use behavior and affectivity. From the results of analysis, it was found that under highly positive affectivity, most behaviors were entertainment-oriented, while under highly negative affectivity, behaviors were mostly about social activities. The result was identical to what Fong (2006) pointed out that in different affective states, individual's behaviors and creativity will be affected in different ways [19].

References

[1] Ferris, J. R. (1996). Internet addiction disorder: Causes, symptoms, and consequences. Retrieved June 20, 2003, from http://www.chem.vt.edu/chem-dept/dessy/honors/papers/ferris.html

[2] Adler, B., & Adler, H. (2001). *Information Affect Scale* (AFS). Retrieved October 06, 2011.

[3] Liaw, G. F., Cheng, M. J., & Wang, Y. S. (2005). A Study of Negative Affectivity Influencing the Relationships among Social Support, Stressor of Role and Job Burnout. *Journal of Human Resource Management, 5*(4), 155-180.

[4] Watson, D., Clark, L. A., & Carey, G. (1988). Positive and Negative Affectivity and Their Relation to Anxiety and Depressive Disorders. *Journal of Abnormal Psychology, 97*(3), 346-353.

[5] Watson, D., Clark, L. A., & Tellegen, A. (1988). Development and validation of brief measures of positive and negative affect: the PANAS scales. *Journal of Personality and Social Psychology, 54*(6), 1063-1070.

[6] Russell, J. A., & Barrett, L. F. (1999). Core affect, prototypical emotional episodes, and other things called emotion: Dissecting the elephant. *Journal of Personality and Social Psychology, 76*(5), 805-819.

[7] Naswall, K., Sverke, M., & Hellgren, J. (2005). The moderating role of personality characteristics on the relationship between job insecurity and strain. *Work & Stress, 19*(1), 37-49.

[8] George, J. M.(2011).Dual tuning: A minimum condition for understanding affect in organizations?. *Organizational Psychology Review,*1(2),147-164.

[9] Williams, S., Zainuba, M., & Jackson, R. (2003). Affective influences on risk perceptions and risk intention. *Journal of Managerial Psychology, 18*(1/2), 126-137.

[10] Watson, D., & Clark, L. A. (1984). Negative affectivity: The disposition to experience aversive emotional states. *Psychological Bulletin, 96*(3), 465-490.

[11] Zevon, Michael A.; Tellegen, Auke(1982) The structure of mood change: An idiographic nomothetic analysis. *Journal of Personality and Social Psychology, 43*(1), 111-122.

[12] Fleur J.M., Steenkamp, E.M. (2005),"Emotions in consumer behavior: a hierarchical approach," Journal of Business Research, 58(10), 1437-1445.

[13] T. Buchanan, A.N. Joinson, C. Paine and U.-D. Reips(2007). Looking for medical information on the internet: Self-disclosure, privacy and trust. *Health Information on the internet*, 8-9.

[14] Tosun, Lajunen (2009). Why Do Young Adults Develop a Passion for Internet Activities? The Associations among Personality, Revealing "True Self" on the Internet, and Passion for the Internet. *Cyber Psychology & Behavior, 12*(4), 401-406.

[15] Federwisch, A. (1997). Internet addiction? Web site for children (http://www.nurseweek.com/features/97-8/iadct.html).

[16] Young, S. K. (1999). Internet addiction: Symptoms, evaluation, and treatment. Web site for children(http://www.netaddiction.com/articles/symptoms.pdf).

[17] Usama M. Fayyad, Gregory Piatetsky-Shapiro, Padhraic Smyth: From Data Mining to Knowledge Discovery in Databases. AI Magazine 17(3): 37-54 (1996).

[18] Laura Widyanto, Mary McMurran, (2004). The Psychometric Properties of the Internet Addiction Test. CyberPsychology & Behavior, 7(4): 443-450.

[19] Fong, C. T.(2006).The effects of emotional ambivalence on creativity. *Academy of Management Journal, 49*, 1016-1030.