# *Predicting Subjective Well-Being by Smartphone Usage Behaviors*

Yusong Gao[1], He Li[2] and Tingshao Zhu[1]
[1] *Institute of Psychology, Chinese Academy of Sciences, 16 Lincui Road, Beijing 100101, China*
[2]*The 6th Research Institute of China Electronics Corporation, Beijing 100083, China*
*gaoys36@163.com, lihessq@163.com, tszhu@psych.ac.cn*

Keywords: Smartphone Usage Behaviours, Subjective Well-Being, Mental Health

Abstract: Subjective Well-Being (SWB) refers to how people experience the quality of their lives, thus to acquire people's SWB levels timely and effectively is very important. Self-report and interviewing are mostly used techniques for assessing SWB, but cannot be done in time. This study aims to predict one's SWB levels by smartphone usage behaviours. We collect users' smartphone usage and self-reported subjective well-being, and found that several usage behaviours correlate with SWB, especially for females. For example, smartphone users with higher SWB scores tend to use more communicating apps, play more games and read more, but take fewer photos. Based on these findings, we trained a predicting model of user's SWB based on smartphone usage behaviours, and the accuracy is up to 62%. The result indicates that SWB can be identified based on smartphone usage fairly well.

## 1 INTRODUCTION

Subjective well-being (SWB) means how people experience the quality of their lives, comprising longer-term levels of pleasant affect, lack of unpleasant affect, and life satisfaction (Angner, 2010; Diener, 1994). SWB focus on how a person evaluates his/her own life, including emotional experiences of pleasure versus pain in response to specific events and cognitive evaluations of what a person considers a good life (Diener, 2000).

Positive psychologists have done much research on SWB (Lyubomirsky, 2001), and they claimed that the pursuit of happiness is regarded as one of the most valued goals in almost every culture. Some studies found that health and SWB influence each other, as good health leads to greater happiness (Okun, 1984), and a number of studies also found that positive emotions and optimism have a beneficial influence on health (Frey, 2011).

As SWB plays an important role in both mental and physical health of people, it is critical to acquire either individual or population SWB level. Currently, SWB is generally measured by using self-reported questionnaires, in which participants fill in the related scales to get the SWB levels. However, this method has three disadvantages.

1. It's not convenient to collect SWB on large population. It may take a lot of manpower and material resource, and it is prone to human errors in the process of questionnaires.
2. It is not easy to identify the driving force of SWB level, if only SWB is presented. We cannot identify any detailed personal conditions or daily behaviours that can explain his or her mental state.
3. Sometimes, it is difficult to acquire SWB in time. Self-reported method is always time-consuming, and conditions may have some changes after the data process (i.e., questionnaire results) has been totally completed.

In order to avoid these disadvantages, we propose to predict SWB based on users' phone usage behaviours. Since a smartphone is built on a mobile operating system with advanced computing capability, it becomes an indispensable part of our daily lives as an attractive tool for communication and interpersonal interaction. IDC (the Internet Data Center) statistics show that in the first quarter of 2013, total global mobile phone sales to 418.6 million, in which smartphones to 216.2 million, accounting for 51.6% of the total mobile phone sales volume.

There are various definitions of phone usage behaviour, some researchers take it as part of the functions of mobile phones, such as telephone,

calendar, SMS using frequency, etc. (Coen et al., 2003). Suh et al. took some specific phone features as phone usage behaviour, such as mobile e-commerce usage (Suh et al., 2003). Falaki et al. described it as users' interactions with the device and the applications used (Soikkeli, Karikoski and Hammainen, 2011). In this paper, we investigate the interactions between users and their devices to explore individual's phone usage behaviour, such as frequency of telephone, SMS, application use, phone switch habits, wallpaper changing, GPS using, etc.

Since smartphone is now a very powerful platform, it is possible to record various interactions between users and devices, ranging from the most basic communication functions to a variety of third-party applications use, such as the frequency of making calls, sending text messages, App using, GPS using. This data intensive platform provides a new opportunity, as it allows us to study users' various phone usage behaviours.

In order to record user's usage behaviour automatically, we have developed an Android application named MobileSens (Li et al., 2013; Guo et al., 2011). This application can record smart phone usage and upload data to the server. At the same time, the application also allows the user to fill a questionnaire and upload the answers to the server. In this paper, we recruited 106 participants using smartphones to record phone usage and SWB scores. We analysed relations between user's phone usage and SWB, and found that it is possible to predict SWB by phone usage behaviours with an accuracy of 62%.

The rest of the paper is organized as following: Section 2 introduces some related works conducted by other researchers. We describe the details of our data in Section 3. Section 4 mainly describes the process of feature extraction, and the results will be thoroughly discussed in Section 5. Section 6 concludes our work in this paper and gives a brief discussion on future work.

## 2 RELATED WORKS

Much research has been conducted to investigate the users' psychological characteristics and their usage of Internet and social media, such as Youtube, blogs, Facebook. It is reported that Internet and social media usage can reflect psychological characteristics (Back et al., 2010; Biel et al., 2011; Counts and Stecher, 2009; Stecher, 2008; Yeo, 2010).

Since smartphones become one kind of ubiquitous communication and interpersonal interaction tools nowadays, smartphone usage may reflect our mood and mental health. Traditional researches mostly used questionnaires to acquire users' phone usage preference, and then analysed these data with the psychological scale scores. Butt pointed out that mobile phone supports interpersonal interaction, and concluded that psychological theory can explain patterns of mobile phone use (Butt and Phillips, 2008). Ehrenberg analysed mobile IM (instant messaging, IM) application use, and found agreeable and low self-esteem have a negative relationship with IM-related applications usage among teenagers (Ehrenberg and Juckes, 2008). With a self-designed questionnaire assessing problematic mobile phone use, Billieux found impulsivity played a specific role in mobile phones usage (Billieux and Van der Linden, 2008). Chittaranjan found some aggregated features obtained from smartphone usage data can indict the Big-Five traits (Chittaranjan and Blom, 2011). Gross et al. reported that time spent on-line was not associated with dispositional or daily well-being. However, the closeness of instant message communication partners was associated with daily social anxiety and loneliness in school, above and beyond the contribution of dispositional measures (Gross et al., 2002).

Some research have been done to investigate mobile phone usage and psychological characteristics. Chittaranjan installed a data collection procedure into Nokia N95, collected 117 participants' phone use logs during 17 months, including text messaging, applications, phone card use logging (Chittaranjan et al., 2011). They found several aggregated features obtained from smartphone usage can be indicators of the Big-Five traits. LiKamWa developed an iPhone software system to record phone use data and remained user of noting down their mood regularly (LiKamWa, 2013). He found smartphone use patterns fluctuate as mood changes, and he built a smartphone software system which can infer user's mood by how the smartphone is used.

In this study, we explore relationship between smartphone usage and SWB using data collected automatically. Differing from studies mentioned above, we intend to predict users' SWB based on smartphone usage, thus to acquire the user's SWB timely and accurately.

# 3 METHODS

After MobileSens has been implemented, this study was carried out by three steps. Firstly, we recruited 106 participants who owned Android smartphones and also use phones in daily life. During the experimental time (one month), all the participants were required to fill out the questionnaire at least once, and should keep the phone network connectivity when phone use records were uploading. Secondly, we extract 48 phone use behavioural features, including SMS, telephone use and wallpaper changed frequency, GPS use frequency, etc. We analysed the relationships between these features and user's well-being by calculating the correlation coefficients. Finally, we extract features and build a prediction model using data mining method.

## 3.1 Participants

We recruited participants by advertising on social networks. Finally, 98 participants between 18 and 32 years (mean 23.5, SD=2.48) were finally included in the data analysis; 61.22% were male, and 38.78% were female. Most of the participants are college students and postgraduates, and Fig. 1 shows detailed education levels of the participants.

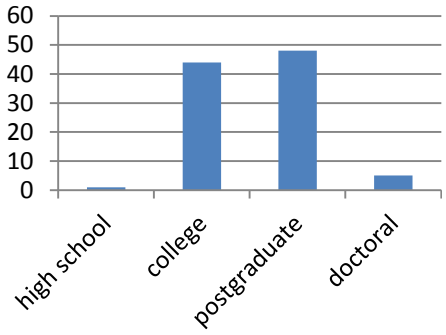


Figure1: Participants Education level distribution

## 3.2 Data collection

MobileSens consists of two modules, one is for collecting phone usage, and the other is for filling the questionnaire, which can get usage data and the corresponding questionnaire results at the same time. Its framework is shown in Fig. 2.

After one user installed MobileSens into their Android devices, most of their interactions with the device were recorded in the Android Sqlite, and these data would upload to our server later. MobileSens recorded 14 categories of information, as listed in Tab. 1.

Table 1: Categories of Phone usage information and details

| Type | Record Content |
|---|---|
| activity application log | creating, starting, resuming, stopping, and exiting of activity application |
| application package log | adding, changing, and removing package |
| calling log | state, number, contact, and direction of calling |
| configuration log | configuration change information (e.g., font, screen size, and keyboard type) |
| contact log | adding, changing, and deleting of contacts |
| date changed log | changing of system date and time |
| GPS log | user's locale, altitude, latitude, longitude and direction of movement |
| headset log | plugging in headset or not |
| power connected log | connecting or disconnecting the power |
| power log | powering on smartphone or not |
| screen log | state of the screen |
| service application log | creating, starting, and deleting service application |
| sms log | state, content, and contacts of SMS |
| wallpaper log | changing wallpaper |

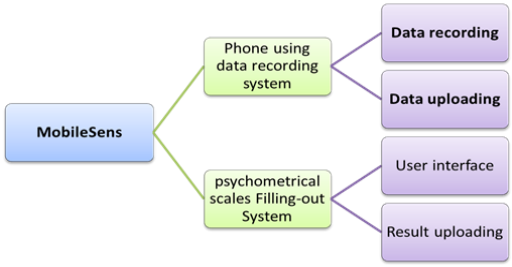


Figure 2: MobileSens Modules

### 3.3 Psychological Scale

In our study, SWB is assessed by Satisfaction With Life Scale (SWLS), which was developed to assess subjective feelings of well-being (Diener et al., 1985). SWLS is a single scale that has been used by UNESCO (The United Nations Educational, Scientific and Cultural Organization), the CIA (the Central Intelligence Agency), the New Economics Foundation, and the WHO (the World Health Organization) to measure how one views his or her self-esteem, well-being and overall happiness with life (Diener, 1999). Among the various components of subjective well-being, SWLS focuses on assess global life satisfaction and does not tap related constructs such as positive affect or loneliness. SWLS items are rated on a 7-point scale: 0 = strongly disagree, 1 = disagree, 2 = slightly disagree, 3 = neither agree nor disagree, 4 =slightly agree, 5 = agree, 6 = strongly agree. Five items are shown as follows.

1. In most ways my life is close to my ideal.
2. The conditions of my life are excellent.
3. I am satisfied with my life.
4. So far I have gotten the important things I want in life.
5. If I could live my life over, I would change almost nothing.

The distribution of participants' SWB scores in this study is shown in Tab. 2.

Table 2. The mean, standard deviation and range of SWLS scores in this study

| | Mean | StdDev | Range |
|---|---|---|---|
| Male | 14.600 | 6.239 | 4~30 |
| Female | 15.395 | 5.299 | 7~28 |
| Total | 14.908 | 5.877 | 3~30 |

## 4 FEATURES EXTRACTION

After data collection, we then extract several features from smartphone usage. There are three types of features extracted as shown in Fig. 3, such as the frequency of some basic smartphone functions usage. We classify Apps according to Wandoujia Android market (http://www.wandoujia.com/apps), which is very popular and comprehensive in China.

48 features are extracted from phone usage data by MobileSens running as a background service. Therefore, these features were objective and captured various aspects of communication and applications use on the phone.

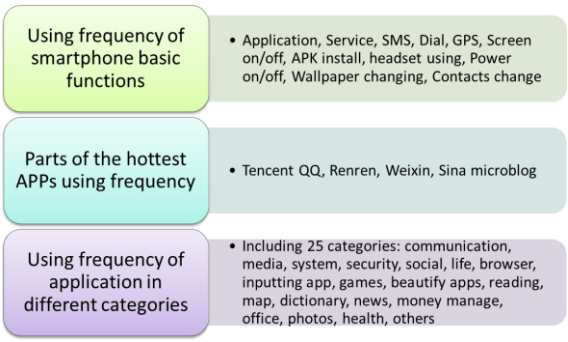


Figure 3: The extracted feature classification and details

## 5 DATA ANALYSIS

### 5.1 Correlation analysis

In this section, we run statistical analysis to examine the correlation between smartphone usage and SWB across genders.

In psychology literature, Pearson's correlation coefficient is commonly used as a bounded measure of correlation, or linear dependence between two variables. For two random variables X and Y, it is given by

$$r = \frac{\mathrm{cov}(X,Y)}{\sqrt{DX \cdot DY}} \quad (1)$$

where cov(X,Y) is the covariance between the random variables X and Y. r=1 denotes a positive sloped linear relationship, and r=-1 denotes a negative one. Values in-between -1 and 1 indicate sub-linear relationships between the variables.

We compute the Pearson's correlation coefficient between the Well-being and the smartphone features. We enlist parts of the correlation analysis results where $p<0.05$ in Table 3.

Table 3: Correlations between Well-being and parts of Smartphone use features

| | Female | Male | Total |
|---|---|---|---|
| Communication App using | 0.371* | 0.234* | 0.251* |
| Strategy Game | 0.320* | 0.004 | -0.002 |
| Reading App using | 0.299* | -0.099 | 0.063 |
| Camera using | -0.311* | -0.242* | -0.251* |
| Competitive Game | -0.050 | 0.265* | 0.197* |
| Browser using | 0.368* | -0.027 | 0.124 |

*p < 0.05

Table 3 shows that users with high SWB tend to use communication Apps (MSN, weixin, feixin) more frequently, but less camera Apps usage, especially for female participants. Besides, all the users with higher SWB are more likely to play games, with females prefer strategy games while males prefer competitive games. In addition, females with higher Well-being may use reading Apps and browsers more frequently, which is not significant on male participants.

## 5.2 Prediction Model

To identify SWB based on smartphone usage behaviours, we train prediction models by machine learning.

To achieve better performance, we need to get more meaningful features at first. In this study, we use StepWise regression as a feature extraction method. StepWise Regression algorithm is a regression-based algorithm for automatic filtering features, and this selection method is applicable to a variety of other models (Kwak and Choi, 2002). It is a forward selection algorithm, which involves starting with no variables in the model, testing the addition of each variable using a chosen model comparison criterion, adding the variable (if any) that improves the model the most, and repeating this process until none improves the model. In this study, we have adopted StepWise regression algorithm for feature selection. The selection result shows that communicate apps usage, certain types of game (strategy games are for female and competitive games are for male), and photo-taken frequency are the most important variables for predicting SWB.

After feature extraction, we built prediction models on WEKA. WEKA contains a collection of visualization tools and algorithms for data analysis and prediction model, together with graphical user interfaces (Witten and Frank, 2005).

In this study, we choose pace regression to build the prediction model in WEKA. Pace regression is one kind of linear regression algorithm, to create parametric linear relationship between independent variables and dependent variable (Wang, 1999). Compared with other algorithms, pace regression performs much better in this study. To make the trained model more general, we use four criterions to measure the quality of prediction model.

1. Correlation Coefficient (CORR)

$$\rho_{X,Y} = \frac{\mathrm{cov}(X, Y)}{\sigma_X \sigma_Y} = \frac{E\left[(X-\bar{X})(Y-\bar{Y})\right]}{\sigma_X \sigma_Y} \quad (2)$$

2. Mean Absolute Error (MAE):

$$MAE = \frac{1}{n}\sum_{i=1}^{n}\left|\hat{f}(x_i) - y_i\right| \quad (3)$$

3. Root Mean Squared Error (RMSE):

$$RMSE = \frac{1}{n}\sqrt{\sum_{i=1}^{n}(\hat{f}(x_i) - y_i)^2} \quad (4)$$

The results of CORR, MAE and RMSE are shown in Table 4.

Table 4: Results of the prediction model

| | CORR | MAE | RMAE |
|---|---|---|---|
| Female | 0.62 | 3.49 | 4.23 |
| Male | 0.34 | 4.82 | 5.89 |
| Total | 0.39 | 4.76 | 5.92 |

We can see that the prediction model for females obviously get a much higher accuracy than models for males or total users. The reason for the result hasn't been found up to now, but we have supposed that females interact with their smartphones more by using more kinds of Apps and spending more time on phones than male users. Therefore, we can predict females' SWB more easily based on females' rich datasets.

# 6 CONCLUSION AND FUTURE WORK

This paper mainly investigates how smartphone usage correlates with SWB, and further develops a SWB prediction model. We intend to provide an

alternative method to measure subjective well-being based on smartphone usage behaviours.

The result demonstrates that users with different SWB level behave differently on their smartphones. Users with higher SWB would like to use more communication Apps but less camera Apps, especially for females users. Besides, all the users with higher well-being are more likely to play games, female users prefer to strategy games and males play competitive games more. In addition, females with higher Well-being use reading Apps and browsers more frequently, which does not appear on male users.

In the near future, we plan to tune the performance of prediction model by developing new behavioural features, and take into account short message and phone voice. We also intend to implement SWB predicting App, to get more feedback from real users.

## ACKNOWLEDGEMENTS

The authors gratefully acknowledge the generous support from National High-tech R&D Program of China (2013AA01A606), National Basic Research Program of China(973 Program， 2014CB744600), Key Research Program of CAS (KJZD-EW-L04), Strategic Priority Research Program (XDA06030800) and 100-Talent Project (Y2CX093006) from Chinese Academy of Sciences.

## REFERENCES

Angner, E., 2010. Subjective well-being. *The Journal of Socio-Economics*, *39*(3), 361-368.

Diener, E., 1994. Assessing subjective well-being: Progress and opportunities. *Social indicators research*, *31*, 103--157.

Diener, E., 2000. Subjective well-being: The science of happiness and a proposal for a national index. *American psychologist*, *55*(1), 34.

Lyubomirsky, S., 2001. Why are some people happier than others? The role of cognitive and motivational processes in well-being. *American Psychologist*,*56*(3), 239.

Billieux, J., Van der Linden, M., & Rochat, L., 2008. The role of impulsivity in actual and problematic use of the mobile phone. *Applied Cognitive Psychology*, *22*(9), 1195-1210.

Chittaranjan, G., Blom, J., & Gatica-Perez, D., 2011. Mining large-scale smartphone data for personality studies. *Personal and Ubiquitous Computing*, *17*(3), 433-450.

Ehrenberg, A., Juckes, S., White, K. M., & Walsh, S. P., 2008. Personality and self-esteem as predictors of young people's technology use. *Cyberpsychol Behav*, *11*(6), 739-741.

Soikkeli, T., Karikoski, J., & Hammainen, H., 2011. Diversity and End User Context in Smartphone Usage Sessions. 7-12.

Frey, B. S., 2011. Happy people live longer. *Science*, *331*(6017), 542-543.

Coen A, Dai L, Herzig S, et al., 2002. The Analysis: Investigation of the cellular phone industry [J]. Retrieved October, 2002, 21: 2006.

Suh, B., & Han, I., 2003. Effect of trust on customer acceptance of Internet banking. *Electronic Commerce research and applications*, *1*(3), 247-263.

Back MD, Stopfer JM, Vazire S, Gaddis S, Schmukle SC, Egloff B, Gosling SD., 2010. Facebook profiles reflect actual personality, not self-idealization. *Psychol Sci* 21:372–374

Biel, J. I., Aran, O., & Gatica-Perez, D., 2011. You Are Known by How You Vlog: Personality Impressions and Nonverbal Behavior in YouTube. In *ICWSM*.

Counts, S., & Stecher, K. B., 2009. Self-Presentation of Personality During Online Profile Creation. In *ICWSM*.

Stecher, K. B., & Counts, S., 2008. Spontaneous Inference of Personality Traits and Effects on Memory for Online Profiles. In *ICWSM*.

Yeo, T. D., 2010. Modeling Personality Influences on YouTube Usage. In *ICWSM*.

Butt, S., & Phillips, J. G., 2008. Personality and self reported mobile phone use. *Computers in Human Behavior*, *24*(2), 346-360.

Gross, E. F., Juvonen, J., & Gable, S. L., 2002. Internet use and well-being in adolescence. *Journal of Social Issues*, *58*(1), 75-90.

LiKamWa R., 2013. MoodScope: Building a Mood Sensor from Smartphone Use Patterns. In *ACM*.

Diener, E. D., Emmons, R. A., Larsen, R. J., & Griffin, S., 1985. The satisfaction with life scale. *Journal of personality assessment*, *49*(1), 71-75.

Diener, E., Suh, E. M., Lucas, R. E., & Smith, H. L., 1999. Subjective well-being: Three decades of progress. *Psychological bulletin*, *125*(2), 276.

Kwak, N., & Choi, C. H., 2002. Input feature selection for classification problems. *Neural Networks, IEEE Transactions on*, *13*(1), 143-159.

Witten, I. H., & Frank, E., 2005. Data Mining: Practical machine learning tools and techniques. *Morgan Kaufmann.*

Wang, Y., & Witten, I. H., 1999. Pace regression.

Li, A., Li, H., Guo, R., & Zhu, T., 2013. MobileSens: A Ubiquitous Psychological Laboratory based on Mobile Device. *International Journal of Cyber Behavior, Psychology and Learning (IJCBPL)*, *3*(2), 47-55.

Guo, R., Zhu, T., Wang, Y., & Xu, X., 2011. MobileSens: A framework of behavior logger on Andriod mobile device. In *Pervasive Computing and Applications (ICPCA), 2011 6th International Conference on*. *IEEE*, 2011: 281-286.

Workshop on New Computationally-Enabled Theoretical Models to Support Health Behavior Change and Maintenance. (2012). Realizing effective behavioural management of health. Available: http://www.behaviorchange.be/