한국빅데이터학회지
제1권 제1호(창간호), 2016, pp. 33-39

스마트폰 사용과 MBTI 사용자 특성간의 관계 평가*

Assessing the Relationship between MBTI User Personality and Smartphone Usage

라저스리 소카세인 · 김경백†

전남대학교 전자컴퓨터공학부

요 약

최근 스마트폰 사용 형태의 도움을 받아 사용자 특성을 예측하는 것은 매우 흥미롭고 주의를 사로잡는 연주 주제이다. 현재 몇몇 연구들은 사용자의 특성을 예측하기 위해 전화 사용 기록, 문자 메시지 사용 기록, 소셜 네트워크 서비스 사용 기록 등을 이용하고 있다. 이 논문에서, 우리는 MBTI 사용자 특성과 스마트폰 사용로그 간의 관계를 평가한다. 이를 위해, 스마트폰 사용 기록에서 부터 몇몇 특징들을 추출하고 이를 Naive Bayes와 SVM등의 분류기에 적용하여 사용자의 특성을 구분하였다. 사용자 특성 분석 결과의 분석을 통해 facebook사용 기록이 외향적인 사람과 내향적인 사람을 가장 잘 구분하는 것을 알 수 있었고, SVM 분류기가 Naive Bayes보다 사용자의 특성을 잘 예측하는 것을 확인하였다.

■ 중심어 : 사용자 특성, 스마트폰, MBTI, 사람 컴퓨터 인터액션, 서포트 벡터 머신

Abstract

Recently, predicting personality with the help of smartphone usage becomes very interesting and attention grabbing topic in the field of research. At present there are some approaches towards detecting a user’s personality which uses the smartphones usage data, such as call detail records (CDRs), the usage of short message services (SMSs) and the usage of social networking services application. In this paper, we focus on the assessing the correlation between MBTI based user personality and the smartphone usage data. We used Naïve Bayes and SVM classifier for classifying user personalities by extracting some features from smartphone usage data. From analysis it is observed that, among all extracted features facebook usage log working as the best feature for classification of introverts and extraverts; and SVM classifier works well as compared to Naïve Bayes.

■ Keyword : User Personality, Smart Phone, MBTI, HCI, SVM

Ⅰ. Introduction

Nowadays usage of smartphone is tremendously increased. The increased usage of smartphone has increased the need of user behavior study for providing good services based on their personality. Determining a user’s personality type by using their smartphone usage and providing services according to their personality preferences, is become more attention-grabbing topic in today’s research space. Knowing the users’ personality can be a strategic advantage for the design of adaptive and personalized user interfaces [4]. In recent years, the field of Human-computer interaction (HCI) has emphasized the importance of identifying the users’ personality traits and preferences in order to build adaptive and personalized systems with an improved user experience [7]. For example, Personality Psychology has provided evidence on the influence of different personality traits over leadership, performance and group interaction styles [1]. Assessment of personality preferences is based on MBTI (Myers-Briggs Type Indicator) theory [9]; the indicator is frequently used in the areas of pedagogy, career counseling, team building, group dynamics, professional development etc. [6].

Existing research work in the direction of user’s personality detection is based on smartphone usage data,

2015년 12월 14일 접수; 2016년 1월 7일 수정본 접수; 2016년 1월 14일 게재 확정.
* 이 논문은 2015년 정부(교육부 또는 미래창조과학부)의 재원으로 한국연구재단의 지원을 받아 수행된 연구임(NRF-2014R1A1A1007734).
† 교신저자 kyungbaekkim@jnu.ac.kr

such as information extracted from call detail records (CDRs) [9], the usage of short message services (SMSs) and the usage of web, music, video, maps, proximity information derived from bluetooth etc., in addition to the traditional call and SMS usage information [2]; GPS log data is used for clustering user locations with the possibility of detecting user's personality preferences [11]. This paper presents the correlation between MBTI user personality and the smartphone usage of a user towards detecting user's personality.

In this paper, we focus on understanding the relationship between MBTI user personality and the smartphone usage of a user for detection of user's personality. To determine the relationship between MBTI user personality and the smartphone usage of a user for detecting user personality, we gathered about 31 user's smartphone usage data which contains call logs, SMS logs, messenger logs and GPS log data; and also the personality type of a user based on MBTI theory [9]. By using the available dataset, we extracted some features such as call data, SMS data, kakao, facebook application usage data. We developed some assumptions based on MBTI theory. We used two well-known classifiers named Naïve Bayes and Support Vector Machine (SVM) to classify smartphone usage data based on personality type and assess the relationship between them. SVM gives better classification results than Naïve Bayes, when we use single feature classifying introverts and extraverts; and Naïve Bayes works better when we use all features for classifying introverts and extraverts. From analysis we can say that facebook usage log is the best feature for classifying introverts and extraverts.

The paper is organized as follows. Section 2 describes methodology which explains dataset used for classification, user personality based on MBTI theory and hypothesis we used to classify the user personality. Also section 2 describes the features we extracted from smartphone usage data for classification in detail as well as classification techniques we used to classify the user personalities. Results of evaluation study are described in section 3. Finally, discussion and conclusion is provided in last section.

Ⅱ. Methodology

2.1 Dataset

To assessing the relationship between user personality and smartphone usage of a user, we gathered a dataset for a month from 2013 October to 2013 November. Android application is designed and implemented on the smartphone of each participated user to collect user behavior data based on their usage of smartphone. We collected 31 users' smartphone usage data such as call logs (incoming and outgoing call log), SMS logs (SMS received and SMS sent log), kakao, facebook logs (application usage log) and GPS logs (GPS data for every 15 minute). And also we collected MBTI-based personality type data of each user. Out of 31 users, 18 users are extraverts and 13 users are introverts. During our experimentation, we mainly focused on I and E attributes of MBTI-based user personality, which describes attitude of a user.

2.2 User Personality

Knowing the users' personality can be a strategic advantage for the design of adaptive and personalized user interfaces. Based on the answers to the questions on the inventory, people are identified as having one of 16 personality types. The goal of the Myers-Briggs Type Indicator (MBTI) is to allow user to further explore and understand their own personalities including their likes, dislikes, strengths, weaknesses, possible career preferences, and compatibility with other people. The 16 distinctive personality types are generated by using the four pairs of preferences or dichotomies viz. 1. Extraversion (E)-Introversion (I), 2. Sensing (S)-iNtuition (N), 3. Thinking (T)-Feeling (F), and 4. Judgment (J)-Perception (P).

Extraversion (E)-Introversion (I): The extraversion-introversion dichotomy is used as a way to describe how people respond and interact with the world around them. Extraverts are "outward-turning" and tend to be action-oriented, enjoy more frequent social interaction, and feel energized after spending time with other people. Introverts are "inward-turning" and tend to be thought-oriented, enjoy deep and meaningful social interactions, and feel recharged after spending time alone.

Sensing (S)-iNtuition (N): This scale involves looking at how people gather information from the world around them. People who prefer sensing tend to pay a great deal of attention to reality, particularly to what they can learn from their own senses. Those who prefer intuition pay more attention to things like patterns and impressions.

Thinking (T)-Feeling (F): This scale focuses on how people make decisions based on the information that they gathered from their sensing or intuition functions. People who prefer thinking place a greater emphasis on facts and objective data. Those to prefer feeling are more likely to consider people and emotions when arriving at a conclusion.

Judging (J)-Perceiving (P): The final scale involves how people tend to deal with the outside world. Those who lean toward judging prefer structure and firm decisions. People who lean toward perceiving are more open, flexible and adaptable. The judging-perceiving scale helps describe whether a user is extravert when user is taking in new information (sensing and intuiting) or when user is making decisions (thinking and feeling).

2.3 Hypothesis

Extraverts usually open to and motivated by outside world of people and things; introverts are "inward-tuning", they like to be alone. Based on this theory, we developed some assumption with MBTI user personality which is generated by using four pairs of dichotomies (mainly focused on extraversion and introversion) and the smartphone usage data.

Assumption1: Extraverts like to be in contact with outside world of people (Stay connected with people), in contrast introverts like to stay alone. Based on this assumption, we used number of incoming and outgoing call logs to check user's connectivity with outside world.

Assumption 2: Extraverts get energized after communicating with outside world of people but introverts feel recharged after spending time alone. By considering this assumption, we used incoming and outgoing call duration to check user's talkative nature with outside world.

Assumption 3: Extraverts like to share more information with outside world and also like frequent social interaction; in contrast introverts enjoy occasional social interaction. With this assumption we used SMS logs and social networking services logs such as kakao and facebook.

2.4 Features

With the available dataset we decided to extract some features to classify user personality.

2.4.1 Call Logs

We consider call logs is one of the feature for classifying user personality. Call log contains number of incoming and number of outgoing call, duration of call information. From the dataset we collected it is observed that the number of incoming calls is less than number of outgoing calls in case of extraverts. From this observation we can say that, extraverts like to stay connected with outside world. Call duration in case of extraverts is greater than introverts; it shows that extraverts are more talkative as compared to introverts.

2.4.2 SMS Logs

SMS log woks as one of the feature for classification of user personality. SMS log contains number of SMS received and sent as well as the length of SMS. SMS sent by extraverts has greater length than to SMS sent by introverts. It shows that extraverts are interested in sharing more information as compare to introverts.

2.4.3 Social Networking Services Logs

In this category we gathered kakao and facebook social networking services log as a feature to classify user personality. Extraverts use these social networking services applications more frequently than introverts. From this we can observed that extraverts enjoy frequent social interaction in contrast introverts enjoy occasional social interaction.

2.5 Classifier

A classification task usually involves training and test sets which consist of data instances. Each instance in the training set contains one target class and several features. The goal of a classifier is to produce a model which is able to predict target class of data exist in the testing set, for which only the features are known. To assess the relationship between user personality and smartphone usage data, we applied two well-known classifiers, Naïve Bayes Classifier and SVM classifier.

2.5.1 Naive Bayes

Naive Bayes models are effective classification tools that are easy to interpret. Naive Bayes classifier requires a small amount of training data to estimate the parameters (means and standard deviation) necessary for classification. Naive Bayes is particularly appropriate when the dimensionality of the independent space (i.e., number of input features) is high.

2.5.2 Support Vector Machine (SVM)

As a classification method, SVM is a global classification model that generates non-overlapping partitions. SVMs are based on maximum margin linear discriminants, and are similar to probabilistic approaches, but do not consider the dependencies among attributes (features).

Ⅲ. Evaluation Results

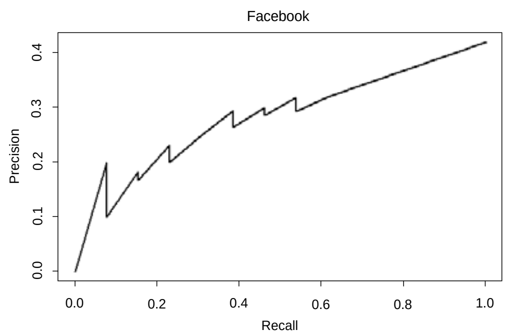
Based on the assumptions explained above in hypothesis section, we used smartphone usage data to extract some features for personality prediction/ classification. Recall and precision is used to get good features for classification. Recall measures how many of the related personalities in a collection have actually been judged as relevant. Precision measures how many of the personalities judged in analysis are actually relevant. Table I shows the percentage of precision and recall for four pairs of dichotomies of MBTI personality type by using Naïve Bayes and SVM classifiers in detail. We used features Call logs, SMS logs, kakao logs and facebook logs for classification of extraverts and introverts. Among all features we used for classification, facebook log works better than all others.

〈Table 1〉 Detail Analysis of All Extracted Features

		Naïve Bayes		SVM	
	Type	Precision	Recall	Precision	Recall
Call log	**E**	**78**	**78**	**83**	**100**
	I	**69**	**69**	**100**	**70**
	N	60	43	-	0
	S	85	92	86	100
	F	54	78	100	28
	T	89	73	78	100
	J	77	81	69	93
	P	55	50	60	17
SMS log	**E**	**61**	**94**	**86**	**100**
	I	**67**	**15**	**100**	**75**
	N	41	100	100	58
	S	100	58	88	100
	F	40	89	100	73
	T	91	45	90	100
	J	100	48	90	100
	P	48	100	100	71
Kak-ao log	**E**	**58**	**39**	**91**	**100**
	I	**42**	**61**	**100**	**85**
	N	50	14	91	77
	S	79	96	92	97
	F	50	22	100	67
	T	74	91	87	100
	J	74	95	95	100
	P	75	30	100	83
Facebook log	**E**	**58**	**100**	**100**	**100**
	I	**-**	**0**	**100**	**100**
	N	-	0	91	100
	S	77	100	100	97
	F	-	0	87	100
	T	71	100	100	95
	J	67	100	94	100
	P	-	0	100	89
All feat-ures	**E**	**84**	**89**	**71**	**90**
	I	**83**	**77**	**75**	**45**
	N	60	86	-	0
	S	95	83	72	100
	F	75	100	0	0
	T	100	86	73	97
	J	82	86	60	100
	P	67	60	-	0

However our work is mainly focused on classification of I and E, but from the analysis we observed that we may able to classify well the other pairs of dichotomies such as S and N, F and T as well as J and P. From table it is observed that SVM classifier works well to classify all other dichotomies of MBTI theory. In some cases for example call log with S and N is not classified well enough because of bias or because of insufficient dataset.

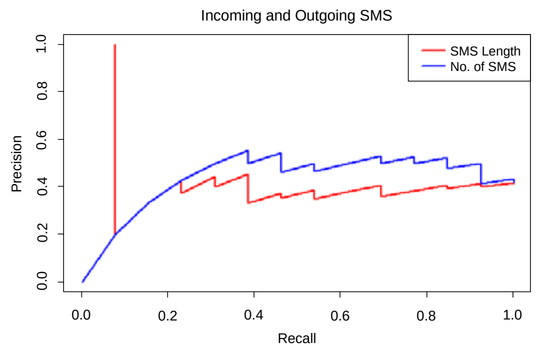
Precision and recall graph for facebook usage is shown in <Figure 1>. The graph shows that if we use facebook usage log as a feature for classification then it can classifies user personalities more appropriately. From table it is observed that SVM classifier is working better than Naïve Bayes classifier to classify with single feature and Naïve Bayes works better than SVM while we use multiple features for classification.



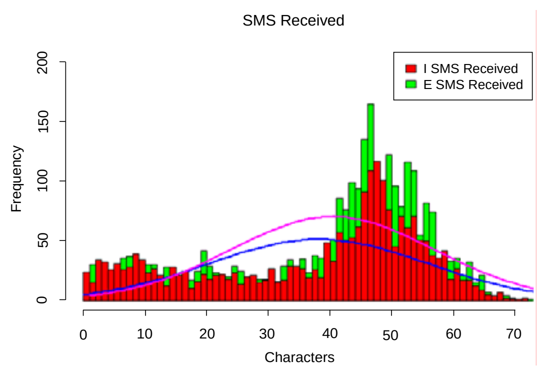
〈Figure 1〉 Precision and Recall for Facebook Logs

<Figure 2> shows the precision and recall graph for SMS log, specifically for number of SMS and SMS length features. From <Figure 2> it is observed that, if we use number of SMS as a feature then it classify the extraverts and introverts very well, but if we use SMS length as a feature to classify extraverts and introvert, it will not classify extraverts and introverts very well. Further we extracted more features from SMS log such as SMS received and SMS sent. <Figure 3> shows the frequency of SMS Vs the length of SMS received by extraverts and introverts. From analysis we observed that, this feature is not good for classifying the extraverts and introverts.

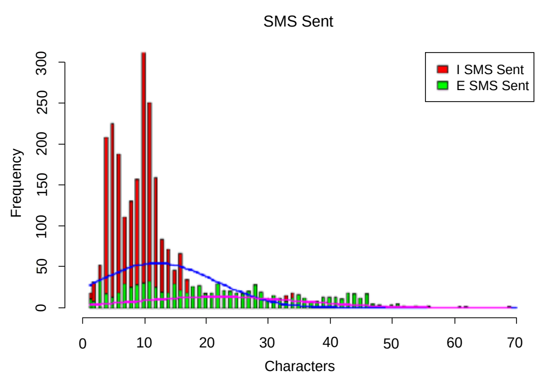
<Figure 4> shows the frequency of SMS Vs the length of SMS sent by extraverts and introverts. We assumed that extraverts like to share more data, so we consider that length of SMS can be feature to classify extraverts and introverts. But according to our assumption 3 this feature is not working well to classify extraverts and introverts, as observed in <Figure 2> (precision and recall curve for length of SMS). Distribution curve with blue and magenta color indicates average length of SMS sent by introverts and extraverts respectively. <Figure 4> shows that frequency of SMS sent is higher in introverts than extraverts.



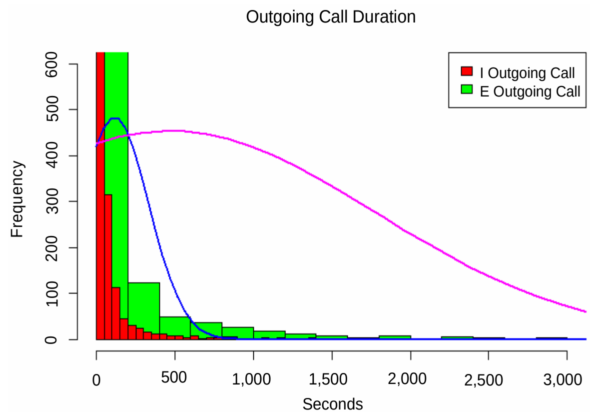
〈Figure 2〉 Precision and Recall for SMS logs



〈Figure 3〉 Frequency and Length of SMS Received



〈Figure 4〉 Frequency and Length of SMS Sent



〈Figure 5〉 Frequency and Length of Outgoing Call

In <Figure 5> we use outgoing call duration of extraverts and introverts as a feature and distribution curve to show average of given data. In <Figure 5> the blue color distribution curve shows average outgoing call duration for introverts and the magenta color distribution curve shows average outgoing call duration for extraverts. From <Figure 5> it is observed that if the duration of call is small then possibly the user is classified as introverts, and if the duration of increased probability of classifying the user as extraverts is increased. From this result we can say that our assumption 2 to classify the user personality with smartphone usage data works well with given data.

According to our assumption 3 “extraverts like to share more data and also like frequent social interaction than introverts”, facebook log can be the best feature for classifying introverts and extraverts; but if we can observe <Figure 2>, SMS log while considering length of SMS as a feature will not give proper classification between introverts and extraverts. From these results we can say that our assumption 3 works well with frequent social interaction but not with sharing data.

The call log data with number of calls as a feature gives good results. It is observed that assumption 1 “extraverts like to stay connected with people and introverts like to stay alone”, also works well. With these analysis results we can say that some of extracted features and considered assumptions we may able to assess the relationship between MBTI based user personality and smartphone usage data.

Ⅳ. Conclusion

In this paper, we focus on understanding the relationship between MBTI user personality and the smartphone usage of a user for detection of user’s personality. We applied two well-known classifiers Naïve Bayes Classifier and SVM classifier for predicting user’s personality. We developed some assumptions for classifying user personality by extracting features from smartphone usage data. Precision and recall is used to get good feature for classification. Based on the features extracted from smartphone usage data, we classified user personality mainly extraverts and introverts. From analysis we observed that by using SVM, we may able to classify not only E and I but also S and N, T and F as well as J and P up to some extent.

As discussed in results section it is observed that, SVM works well with an individual features. In contrast Naïve Bayes works well with all features rather than individual features. Our dataset has limited information, sometimes it is not properly predict the user personalities; to overcome such issue we will collect sufficient dataset to work further in this direction.

References

[1] Balthazard, P., R.E. Potter, and J. Warren, “Expertise, extraversion and group interaction styles as performance indicators in virtual teams: how do perceptions of its performance get formed?”, *SIGMIS DB*, Vol.35, No.1, pp.41-64, 2004.

[2] Chittaranjan, G., J. Blom, and D. Gatica-Perez, “Who’s who with Big-Five: Analyzing and Classifying personality traits with smartphones”, In Wearable Computers (ISWC), *2011 15th Annual International Symposium on*, *IEEE*, pp.29-36, 2011.

[3] de Montjoye, Y.A., J. Quoidbach, F. Robic, and A.S. Pentland, “Predicting personality using novel mobile phone-based metrics”, *In Social Computing, Behavioral-Cultural Modeling and Prediction*, pp.48-55. Springer Berlin Heidelberg, 2013.

[4] de Oliveira, R., A. Karatzoglou, P.C. Cerezo, A.A.L. de Vicuña, and N. Oliver, "Towards a psychographic user model from mobile phone usage", *In CHI'11 Extended Abstracts on Human Factors in Computing Systems*, pp.2191-2196, ACM, 2011.

[5] Ester, M., H.-P. Kriegel, J. Sander, and X. Xu, "A Density-Based Algorithm for Discovering Clusters in Large Spatial Databases with Noise", *In Proc. 2nd Int'l Conf. on Knowledge Discovery and Data Mining, Portland, Oregon*, pp.226-231, 1996.

[6] http://en.wikipedia.org/wiki/Myers-Briggs_Typ e_Indicator.

[7] Lee, K.M. and C. Nass, "Designing social presence of social factors in human computer interaction", *In CHI '03: Proc. of the SIGCHI conf. on Human factors in computing systems*, pp.289-296, New York, NY, USA, 2003.

[8] MacQueen, J., "Some methods of classification and analysis of multivariate observations", *In Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability*, pp.281-297 by J.B. Macqueen, 1967.

[9] Myers, I.B. and Mary H. McCaulley, "Myers-Briggs Type Indicator: MBTI", *Consulting Psychologists Press*, 1988.

[10] Nov, O., O. Arazy, C. López, and P. Brusilovsky, "Exploring personality-targe ted UI design in online social participation systems", *In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pp. 361-370. ACM, 2013.

[11] Sokasane, R. and K Kim, "Predicting User Attitude by Using GPS Location Clustering", *IEICE Transactions on Information and Systems*, Vol.E98-D, No.8, pp. 1600-1603, 2015.

저 자 소 개

라저스리 소카세인(Sokasane S. Rajashree)

- 2005년 : Shivaji University, Department of Electronics (학사)
- 2007년 : Shivaji University, Department of Computer Science (석사)
- 2013년~현재 : 전남대학교 전자컴퓨터공학부 (박사과정)
- 관심분야 : 빅데이터 활용, 위치정보 응용시스템

김 경 백(Kyungbaek Kim)

- 1999년 : 한국과학기술원 전기및전자공학과 학사
- 2001년 : 한국과학기술원 전기및전자공학과 석사
- 2007년 : 한국과학기술원 전기 및 전자공학과 (박사)
- 2007년~2011년 : University of California Irvine (박사 후 연구원)
- 2012년~현재 : 전남대학교 전자컴퓨터공학부 (교수)
- 관심분야 : 분산시스템, 미들웨어, 빅데이터 분석 및 활용, 오버레이 네트워크, 소셜 네트워크, 소프트웨어 정의 네트워크