Pervasive and Mobile Computing 69 (2020) 101269

Contents lists available at ScienceDirect

Pervasive and Mobile Computing

journal homepage: www.elsevier.com/locate/pmc

When phones get personal: Predicting Big Five personality traits from application usage

Ella Peltonen [a,*], Parsa Sharmila [a], Kennedy Opoku Asare [a], Aku Visuri [a], Eemil Lagerspetz [b], Denzil Ferreira [a]

[a] Center for Ubiquitous Computing, University of Oulu, P.O. Box 4500, FI 90014, Finland
[b] Department of Computer Science, University of Helsinki, P.O. Box 64, FI 00014, Finland

ARTICLE INFO

Article history:
Received 4 July 2020
Received in revised form 26 August 2020
Accepted 3 October 2020
Available online 10 October 2020

Keywords:
Smartphones
Personality
HCI

ABSTRACT

As smartphones are increasingly an integral part of daily life, recent literature suggests a deeper relationship between personality traits and smartphone usage. However, this relationship depends on many complex factors such as geographic location, demographics, or cultural influence, just to name a few. These factors provide crucial knowledge for e.g. usage support, recommendations, marketing, general usage improvements. We use six months of application usage data from 739 Android smartphone user together with the IPIP 50-item Big Five personality traits questionnaire. As our main contribution, we show that even category-level aggregated application usage can predict Big Five traits at up to 86%–96% prediction fit in our sample. Our results show the effect of personality traits on application usage (mean error improvement on random guess 17.0%). We also identify which application usage data best describe the Big Five personality traits. Our work enables future personality-driven research, and shows that when studying personality, application categories can provide sufficient predictions in general traits.

1. Introduction

Smartphones are increasingly intertwined into our communities and everyday life [1,2]. Consequently, there is rising interest to study how people use their smartphones and which factors affect application choice, interaction and communication routines. Smartphones are more than phones, including features such as navigational maps, cameras, information search, and games. It is possible to study smartphone usage in the wild [3] and collect large datasets on naturalistic human behaviour, radically different from traditional laboratory experiments [4,5].

After more than a decade of research to understand how smartphones are appropriated, the question remains unanswered on the relationship between ones' personality and how they use their smartphone to predict potential future smartphone use. There are several conceptualised physiological measures proposed for assessing one's personality, and here we follow the Big Five personality traits and investigate how they are manifested on smartphone usage.

Studying someone's personality traits using technology with smartphones is not novel. For example, understanding personality traits can increase user experience and engagement [6,7], helping people with disabilities to become part of the world-scale smartphone user community [8], and therefore make a smartphone use more useful, more enjoyable, and to explore new and innovative functionalities. Understanding the user, the actual stakeholder of all of these technological

* Corresponding author.
E-mail addresses: ella.peltonen@oulu.fi (E. Peltonen), parsa.sharmila@oulu.fi (P. Sharmila), kennedy.opokuasare@oulu.fi (K. Opoku Asare), aku.visuri@oulu.fi (A. Visuri), eemil.lagerspetz@helsinki.fi (E. Lagerspetz), denzil.ferreira@oulu.fi (D. Ferreira).

https://doi.org/10.1016/j.pmcj.2020.101269

designs, is key to creating enticing smartphone experiences. Personality traits are arguably as important as classical demographic factors (e.g., location, time of day) when predicting deeds and needs of the smartphone use [9]. On a larger scale, understanding the culture and one's society [2] expands personality traits from individuals to their communities.

We analyse smartphones' application usage data from 739 Android users to predict their personality traits determined by the IPIP 50-item Big Five questionnaire [10,11]. These participants are a subset of the whole participant sample from the Carat dataset which fulfilled the Big Five questionnaire. Our analysis and personality model can, in the best case, predict the personality traits in accuracy of 1–2 points (in a scale of maximum 50). We find evidence that communication-oriented apps and games have a direct relationship to one's personality traits. We make the following contributions to the state-of-art:

1. We analyse 739 Android users' app usage and their IPIP 50-item Big Five questionnaire, achieving an optimal prediction fit of 96% and 86% at worst;
2. We find app categories that best describe the Big Five personality traits, i.e., high Extraversion and Agreeableness are best described by use of communication apps, and Conscientiousness and Emotional Stability by use of casino games.
3. Our results suggest that, when investigating personality based on smartphone usage, app category-level aggregation is sufficient for accurate predictions.

2. Related work

Smartphones are inherently personal, full of sensors (e.g., accelerometer, camera, microphone, GPS), and have a large number of available apps. Instrumenting the sensors e.g., to capture app usage and user interactions, with research-oriented data logging apps [3–5] has enabled real-life data collection from the study participants. Such data provide insight into various aspects of the human behaviour and to develop interventions, e.g., smartphone app usage [2,12,13], technological issues such as smartphone battery consumption [14,15], health-related conditions such as mental health [16–18], Parkinson's Disease progression [19,20], and environmental and urban sensing [21–24].

When focusing on understanding how and why people use smartphones, demographics such as age [25,26], countries and geographic locations, cultural factors, gender, income, and education [2] are influential. Zhao et al. [27] use demographic factors and application usage to define user clusters, such as "night communicators", "screen checkers", "financial users", and "young parents". Hiniker et al. [28] define instrumental and ritualistic use of different smartphone applications; users seeking ritualistic gratifications are more likely to play games, use social media, and read news. Böhmer et al. [12] show the daily patterns in the mobile app usage highlighting late-night gamers of the then popular "Angry Birds".

Based on these wide aspects of studies committed in understanding smartphone users, it is somehow surprising that personality traits have been understudied in our field. In psychology, they are too often linked to only harmful or disruptive technology use [29], which seems rather contradictory to how HCI perceives technology, i.e., as a facilitator or enabler to productivity and quality of life [30]. Lee et al. [31] investigate the dark patterns in smartphone usage and covers socially anxious behaviour, materialism, and the compulsive use of smartphones. Takao et al. [32] link smartphone usage to addictive behaviour, and several other authors link smartphone usage to mental health, especially depression [17,33]. However, there are also support to understanding personality of smartphone users, e.g., by providing them with more comprehensive experiences, better targeting of user designs, and aiming at more accurate recommendations. Ehrenberg et al. [34] link willingness of texting instead of call to different personality traits, and Montag et al. [35] link WhatsApp usage in personality traits. We question how personality plays a role in device usage. Especially communication and types of communication seems to be an important aspect to be studied and considered together with personality analysis.

3. The Big Five personality traits

The Big Five personality traits model consists on five factors, commonly called as Agreeableness, Conscientiousness, Extraversion, Openness, and Neuroticism [36], but the names have varied over time [37]. We borrowed the IPIP — International Personality Item Pool that states for a scientific collaboratory for the development of advanced measures of personality and other individual differences [10,11], with traits categorised as: (1) Extraversion, (2) Agreeableness, (3) Conscientiousness, (4) Emotional Stability (instead of Neuroticism), and (5) Intellect/Imagination (instead of Openness). The factors and their meanings are as follows: *Agreeableness* is a personality trait associated with kindness, concern for getting along with others i.e social cohesion and friendliness, trustworthiness, and optimism [13]. People with high *Conscientiousness* exhibit self discipline and focuses on achievements, are methodical and thorough. *Extraversion* is a personality trait associated with energetic behaviour, sociability and assertiveness. *Emotional Stability* describes how person is predisposed to experiencing sadness, embarrassment, distrust, anxiety and anger. *Intellect/Imagination* describes a person's willingness and openness to new intellectual experiences, and curious attitude for new things.

Despite Big Five personality model's popularity, it is not without criticism: personality is mutable; "Big" is meant to describe broadness of the traits instead of their greatness of perfection; and the model is not meant to be comprehensive presentation of the personality [36]. The Big Five traits may not have a clear relation to the real-life performance, such as leadership skills, work motivation or attitudes, coping stress, or team effectiveness [38]. Nonetheless, the Big Five is the most popular model to describe personality for behavioural and sociological sciences, which we appropriate as a lens to understand mobile app use.

Previous work has shown that the Big Five personality traits of individuals can affect individual's usage behaviour of certain applications and social media. Klobas et al. [13] study certain personality traits and their strong association with compulsive use of YouTube. Hierarchical regression models analysed self-reported surveys administered to 807 (528 female and 279 male) students. The students were classified into Compulsive user and Non-Compulsive user based on the responses from a compulsive use survey [39]. The personality trait of students were also measured using the 44-item Big Five scale [36]. The study concludes that there is no effect of Intellect/Imagination and Extraversion on YouTube compulsive use. Conscientious ($R^2 = 0.07$) and Agreeable ($R^2 = 0.02$) students are better placed to prevent or resist the temptation for compulsive use of YouTube, while neurotic ($R^2 = 0.10$) students are more inclined to become compulsive users. In addition, Quercia et al. [40] compare Twitter usage to the personality traits and present correlations especially to Emotional Stability and Extraversion, and Ortigosa et al. [41] show that social media interactions in Facebook can predict personality traits with accuracy of 60%–80%.

Similarly, Mehrotra et al. [42] investigate the relationship between personality traits and people's perceived receptivity and interruptibility *i.e* the tendency of a person to react or interact with notifications, alert dialog, and Experience Sampling Method (ESM) questionnaires. In their study, a smartphone app that collects contextual sensor data (Wi-Fi, GPS, ambient noise, light intensity), ESM questionnaire, and the 50-item Big Five [43] personality test was deployed on Google's Play Store. The study's findings suggested that persons who exhibited more Extraversion are more likely to perceive notifications as disruptive($R^2 = 0.737$). Some previous results [13,42] demonstrate the potential to develop personalised systems that takes the personality individuals into consideration. For instance, consider a use-case in customer service management where a chat-bot or conversational agent is able to automatically detect or predict the personality trait based on the individuals' voice [44], or used text [45], to better respond accordingly to the personality of the individual to improve customer relationships and retention.

Chittaranjan et al.'s [46] 2011 study of 83 Nokia N95 users shows supporting evidence: office-related apps are linked to users being conscientious but not emotionally stable and exhibit low openness; internet is mostly used by introverts; multimedia is less used by conscientious; and mail users as conscientious and neurotic. In 2013, Chittaranjan et al. [47] continue with 117 Nokia N95 smartphone users and found extroverts preferring less use of internet, games, and camera applications, agreeable users using hardly any apps, conscientiousness correlating negatively to the music apps, emotional stability correlating negatively to the office and calendar apps, and openness correlating widely negatively to multiple app categories. Today, even with a larger variety of applications at our disposal, and their functionalities are largely social and different in comparison, we can identify similar patterns in our work. Noteworthy, Oliveira et al. [48] call for action to predict personality traits from smartphone usage, and Montjoye et al. [49] show analysing phone call logs can predict personality traits 42% better than random. Recently, Stachl et al. [9] study 137 individuals and 2835 apps and show that Extraversion, Conscientiousness and Agreeableness predict well smartphone app usage, and outperform basic demographic variables. Continuation study of Stachl et al. [50] with 624 individuals highlights there is indeed a strong connection between smartphone usage and personality traits. When they look smartphone usage as a whole, our study focuses on applications and especially application categories, showing that categories as an aggregation model provide sufficient prediction fit to study personality.

Predictive analysis to determine the Big Five personality traits in supervised machine learning methods are promising. In a study that predicts Big Five personality traits from voice [44], Support Vector Machine (SVM) and Hidden Markovs Model (HMM) classifiers are trained with 640 (322 unique speakers) speech corpus that has been labelled with the scores from the 10-item Big Five Inventory (BFI-10), a substantially shorter version of 50-item Big Five personality test. The prediction models are then tested with 15 new voice recordings from three journalists (*2 females, 1 male*), that have been also labels with BFI-10 scores from twelve assessors. The classifier results in prediction accuracy; Intellect/Imagination 78.98%, Conscientiousness 90.78%, Extraversion 70.15%, Agreeableness 66.72%, and Emotional Stability 77.66%. Staiano et al. [51] use the Random Forest approach when studying personality traits from social network structures. In this paper, we use the 50-item validated IPIP questionnaire from 739 participants, thus a significantly larger sample when compared to previous work. Similarly [51], we implement a Random Forest Regression-based prediction model that is easily exploitable and is computationally lightweight for future personality-aware smartphone apps. To position our work, we revisited the literature that utilised BigFive questionnaires and smartphone-based sensor data to predict users' personality traits (Table 1).

4. Investigating personality traits using smartphones

4.1. Mobile application dataset

Our dataset originates from Carat [4], an open source mobile data gathering platform. The Carat platform users are anonymous worldwide volunteers interested in partaking in citizen science. We study a subset of the openly licenced dataset collected between 14 March 2018 and 25 August 2018, approximately a six-month long window, from an initial count of 843 Android users worldwide [2]. We filtered smartphone use data from users who had at least 10 days of app

Table 1
Summary of the previous contributions in comparison to this paper.

Reference	N	Data	Analysis method	Findings
Chittaranjan et al. [46], 2011	83	10-item BigFive; 8 months of app usage, SMS, call, and Bluetooth data	Pearson Correlation Multiple Regression	Office-related apps are linked to users being conscientious but not emotionally stable and exhibit low openness; internet is mostly used by introverts; multimedia is less used by conscientious; and mail users as conscientious and neurotic.
Staiano et al. [51], 2012	53	44-item BigFive; 8 weeks call and Bluetooth proxity logs	Random Forest Regression	Call logs predict personality traits better than random. Mobile phones-based behavioural data can be superior to surveys for personality classification.
Montjoye et al. [49], 2013	69	44-item BigFive; 1 year phone calls, SMS and GPS locations	Support Vector Machine	Extraversion and Neuroticism are the best predicted by basic phone usage.
Mehrotra et al. [42], 2016	20	50-item BigFive; Experience Sampling Questionnaires; Notifications, WiFi, GPS, noise, light	Linear Regression	People with higher Extraversion are more likely to perceive notifications as disruptive
Wang et al. [52], 2018	646	BigFive; 14 days of Ambient Noise Intensity	Bivariate Linear Mixed Models Gradient Boasted Regression	Predicted BigFive scores significantly correlated with self-reported personality traits.
Gao et al. [53], 2019	52	52-item BigFive; Accelerometer, call and messaging logs	Support Vector Regression	Predicted BigFive scores highly correlate with the baseline, e.g. conscientiousness
Stachl et al. [54], 2019	137	300-item BigFive (German version); 60 days of smartphone app usage	Spearman Correlation and LASSO penalised Regression Modelling	Extraversion, Conscientiousness, and Agreeableness predict well smartphone app usage, and outperform basic demographic variables.
Stachl et al. [50], 2020	624	300-item BigFive (German version); 30 days of smartphone usage logs (GPS, Application usage, Bluetooth and Wifi)	Elastic Net Regularised Linear Regression, Random Forest	All personality traits was successfully predicted from behavioural patterns derived from smartphone data. Communication and app usage were among the most significant predictors of personality traits.
This work, 2020	739	50-item BigFive; Six months smartphone usage including applications and categories	Random Forest Regression, Deep Neural Network and Support Vector Regression	Optimal prediction fit for all traits. High Extraversion and Agreeableness are best described by communication apps, and Conscientiousness and Emotional Stability by casino games. Category-level aggregation is enough to give accurate results.

usage data, to remove those users who for example, installed the Carat only at the end of the data collection period, or who did not succeed to send their data to the server due to the technical difficulties. In order to omit less significant users from the prediction model, we discarded 104 number of users who have data less than 10 days resulting in our selected 739 users. From the self-reported demographics, out of those 739 users, 55% of them the country is not known, 17% are based on US, and the rest to 41 other countries, mainly the UK, Finland, Germany, and Canada. The age distribution is as follow: 11% of age 18–24 years, 31% of age 25%–34%, 27% of age 45–64, and 6% 65 years or older. In comparison to studies mainly focused on students [55], we have 38% of our users working as professionals, 13% working as associate professionals, 10% managers, and 13% students. Other occupational groups present, e.g., the retired, self-employed, and home parents. However, mainly due to the fact we cannot really control who answers a crowdsourced questionnaire, we have a strong bias towards males (88%). This may be due to the Carat app being marketed to the tech-savvy community.

The Carat data consists on measurements containing the current running apps, timestamp, and user identifier, and the app state, i.e., is it on the foreground or background. Originally designed for the battery consumption research, the Carat [4] platform collects snapshot data on 1% of the battery drainage. As we study usage as routines of app use over time, not granular app interactions, we focused on foreground apps, i.e., the ones' users had interaction with. We created two binary matrices, following the previous pre-processing approach [2]. For app categorisation, we consider Android's Google Play categories and how the apps are mapped accordingly at the time of analysis. The dataset has 7852 apps that are available in Google Play. They belong to 41 categories, such as *Tools*, *Travel and Local*, and *Games*. The percentage of user population statistics for the five most and least used app categories are given in Table 2. Among the most used categories, *Tools* are used by all the users — it includes apps like Antivirus, Battery Status Tool, File Manager, Performance Tuner, and App Organiser. Followed by second most used app category is Communication which includes apps like Messaging, App Contacts, Call Recorder, Google Voice, and so on. In the least used categories, we can see all the apps are under different *Game* categories. It is noteworthy that several misclassification can be seen for the apps under the game categories because these categories are described with similar words from word tokenization [56].

Table 2
Percentage of users with 5 most and 3 least used application category (739 participants).

Category	Users (%)	Category	Users (%)
Tools	100	Game music	0.0
Communication	99.6	Game casino	0.1
Game card	96.8	Game educational	0.3
Transportation	92.0	Game sports	0.4
Productivity	92.0	Game trivia	0.4

Table 3
Descriptive statistics of the Big Five values gathered by 739 participants.

Statistic	Extra-version	Agree-ableness	Conscientiousness	Emotional stability	Intellect/ Imagination
Mean	27.6	36.7	35.0	32.5	39.0
Standard Deviation	8.0	6.5	6.2	7.9	5.9
Minimum	11	10	17	10	17
25th Percentile	22	33	31	27	35
50th Percentile	28	37	35	33	39
75th Percentile	33	42	39	38	43
Maximum	49	50	50	50	50
Median	28	37	35	33	39

Ethical Considerations. We only consider aggregate-level data that contains no personally identifiable information, following the privacy protection mechanisms of the Carat platform [4]. Data collection is subject to the IRB process of the University of California, Berkeley. The mobile users are informed about the collected data and have given their consent from their devices. The 50-item IPIP user questionnaire performed for this work have been approved on 14 June 2016 by the IRB process of the University of Helsinki, Finland. Participation in the study has been voluntary and the users have been informed about the data collection and management procedures.

4.2. Big Five personality trait dataset

The Big Five personality questionnaire consisting on 50 standardised questions were given to the volunteering Carat platform users through to the user questionnaire tool implemented as a part of the Carat Android version. Through an anonymised user ID, we are capable of matching questionnaire results to the measured app usage in Carat. We obtained 739 answers for our study. We follow the Big Five personality trait questionnaire format as per the official website of IPIP 50-item scale [43]. Users indicate themselves with any of the five criteria, 1. Very Inaccurate, 2. Moderately Inaccurate, 3. Neither Accurate Nor Inaccurate, 4. Moderately Accurate, or 5. Very Accurate for each of the 50 statements. The statements describe a person to understand the five traits which are Extraversion, Agreeableness, Conscientiousness, Emotional Stability, and Intellect/Imagination. Table 3 summarises the descriptive statistics of the traits.

5. Personality from smartphone data

Our dataset has 7852 apps, sparsely distributed as few apps are popular among the majority of users. In order to identify the apps not contributing well to the generalisation capacity of our personality model, we use Principal Component Analysis (PCA) on the app data. Based on the PCA, 600 components preserve 99% of the total variance. The mobile usage data is multinominal and personality traits continuous, and this nature of the data guides us to use the following algorithms for the prediction: Random Forest Regression (RF), Deep Neural Network (DNN), and Support Vector Regression (SVR). The basic 10-fold cross-validation over users is applied for training the models.

As an evaluation metric for the analysis, we use Root Mean Square Error (RMSE). It is a standard deviation of prediction errors, in other words, RMSE is the difference between the predicted values of a model and the actual values observed for the data set. RMSE describes how data is condensed around the best fit. Formally,

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(p_i - d_i)^2} \tag{1}$$

where n total number of samples, p is the predicted value, and d is the actual value. For goodness of the fit measure of regression performance, we calculate $1 - RMSE$.

The prediction results for the Big Five with RF, DNN, and SVR, separately for app and category data, are given in Table 4. When comparing the prediction errors, we can see that the difference between RF and SVR is minimal. Between RF and DNN, the prediction fit is better by DNN for app categories but for apps only, the prediction fit of DNN is low compared to RF. Considering the comparatively better fit of RF for both apps and app categories, we choose RF over DNN and SVR and

Table 4
Goodness of fit comparison ($1 - RMSE$) for Random Forest Regression (RF), Support Vector Regression (SVR) and Deep Neural Network (DNN), for apps and app categories.

Big Five factor	Applications			App categories		
	RF	SVR	DNN	RF	SVR	DNN
Extraversion	91%	90%	87%	91%	90%	91%
Agreeableness	93%	93%	87%	93%	92%	94%
Conscientiousness	94%	93%	86%	93%	93%	93%
Emotional stability	91%	91%	84%	92%	91%	91%
Intellect/ Imagination	93%	93%	87%	93%	93%	94%

Table 5
Normalised RMSE values for each factor and based on whether apps or app categories.

Big Five factor	Source	25th	50th	75th
Extraversion	Applications	5%	14%	20%
	App Categories	5%	14%	20%
Agreeableness	Applications	4%	11%	14%
	App Categories	4%	10%	14%
Conscientiousness	Applications	5%	9%	14%
	App Categories	5%	10%	15%
Emotional Stability	Applications	6%	14%	20%
	App Categories	6%	13%	20%
Intellectual/Imagination	Applications	4%	10%	14%
	App Categories	4%	10%	15%

later discuss its results in Section 6. *Hyperparameter tuning* is done before training the RF model. The dataset is divided into training, testing and validation sets by 60%, 20% and 20% respectively. The parameter estimators of RF which are considered are: number of estimators, maximum depth of a tree, minimum samples for each split, and minimum samples for each leaf. We created a parameter grid of these estimators (n_estimators: ranged between 200 to 2000 as 200, 400, ...; min_samples_split: ranged between 2 to 5; min_samples_leaf: between 1 to 4; max_depth: ranged between 10 to 110 as 10, 20, 30...) and randomly searched for the best parameters using Python RandomizedSearchCV where three-fold cross-validation has been used with 100 iterations, fitting three folds for each of 100 candidates make it a total 300 fits to choose the best parameters.

6. Predicting the Big Five: The results

Based on the baselines discussed above, we perform Random Forest Regression (RF) to predict the Big Five personality traits by using two feature sets: app usage and category usage. Table 5 presents the normalised error as percentage, for 25th, 50th and 75th percentiles of the prediction error and thus the population. In the other words, there the each quartile includes 25% of the users when organised by optimal prediction fit, the first quartile including the most optimal fits and the fourth quartile the least optimal ones. Keeping in mind the distributions of the Big Five traits in Table 3, we can see that the median error is close to 4–7 "points" on the trait's scale (0–50), which is systematically 1–2 points better than the standard deviation of the data. When looking the 25th percentile of the prediction errors, we see that for every fourth user our method is accurate enough to predict the their personality traits by only 1–2 points away from the original Big Five value. As normalised results (Table 5), that means 4%–6% error. For the worst quarter of the user (over 75th percentile) we get error estimates between 7–9 points on the scale, which is close to the standard deviation and too much to give an accurate prediction to the personality. As normalised results (Table 5), this means 14%–20% error.

For approximately half of the population (50th in Table 5), we can predict personality traits with 9%–14% error. In terms of goodness of fit, that means around 86%–91%. However, even if the results are quite distributed, in the best cases we are very close to the actual Big Five traits, keeping in mind the personality being probably more versatile phenomenon than the user's smartphone usage can ever predict. If compared to work of Ortigosa et al. [41] based on 20988 Facebook users, social media interactions can predict personality traits with goodness of fit 60%–80% depending on the prediction model and definition of social interactions (e.g. number of active friends and number of posts in a month). Based on this comparison, we can see that smartphone usage has better predictive power for personality traits than social media interactions.

Interestingly, there seems to be no large difference if using the app data or app category data for prediction. App usage-based predictions are around one "point" more accurate than category-based. This is, indeed, an important indication that, at least when studying personality based on the smartphone usage, category-level aggregation would be enough to give accurate results. This may help the future studies to protect the user privacy by fading the actual app combinations.

Table 6
The highest pseudo-ρ^2 measures between the Big Five factors and the following categories: (A) Communication, (B) Game Action, (C) Game Board, (D) Game Casino, (E) Game Educational, (F) Game Simulation, (G) Game Trivia.

Big Five	A	B	C	D	E	F	G
Extraversion	0.246	0.182	0.211	0.311	0.102	0.165	0.223
Agreeableness	0.218	0.104	0.164	0.284	0.165	0.122	0.162
Conscientiousness	0.119	0.043	0.107	0.448	0.039	0.012	0.119
Emotional Stab.	0.046	0.047	0.125	0.515	0.272	0.179	0.214
Intellect	0.118	0.056	0.079	0.342	0.027	0.104	0.026

Table 7
The smallest pseudo-ρ^2 measures between the Big Five factors and the following categories: (A) Entertainment, (B) Finance, (C) Health and Fitness, (D) Media and Video, (E) Music and Audio, (F) News and Magazines, (G) Personalisation, (H) Travel and Local, (I) Weather.

Big Five	A	B	C	D	E	F	G	H	I
Extraversion	0.001	0.003	0.000	0.001	0.001	0.001	0.004	0.009	0.003
Agreeableness	0.000	0.003	0.001	0.000	0.002	0.002	0.001	0.004	0.003
Conscientiousness	0.000	0.005	0.001	0.000	0.002	0.002	0.001	0.001	0.003
Emotional Stab.	0.002	0.030	0.001	0.000	0.005	0.003	0.008	0.004	0.007
Intellect	0.000	0.006	0.002	0.000	0.000	0.001	0.004	0.002	0.004

7. Deeper into the application categories

Given the prediction models done with the overall picture of the app usage, it is also interesting underlying question which categories affect individually to the Big Five factors. Thus, we compare each Big Five factor separately to each of the categories. We measure impact between category usage and each personality trait by using logistic regression. We report the McFadden's pseudo R-squared, ρ^2, which is a substitute of more familiar R^2 for linear regression analysis. Formally,

$$\rho^2 = 1 - \frac{logL(m_1)}{logL(m_0)}, \tag{2}$$

where $logL(m_1)$ is the log-likelihood of the logistic regression model and $logL(m_0)$ is the log-likelihood of the null hypothesis.

The categories with the highest ρ^2 i.e. most impact are presented in Table 6. We can see that the category *Communication* has one of the strongest impact Extraversion ($\rho^2 = 0.25$) and Agreeableness ($\rho^2 = 0.22$). Slightly higher use of communication apps can also be associated to both Conscientiousness and Intellect/ Imagination ($\rho^2 = 0.12$), whereas dependency between communication apps and Emotional Stability is not found. In the newest related studies, Stachl et al. [9] present communication apps being well predicted by Extraversion, which is also clearly in line with our study. In addition, Azucar et al. [57] define in their social media-based meta-analysis correlations 0.29 for Agreeableness and 0.40 for Extraversion, which is also in line with our study.

In addition, clear impact can be associated to variety of different games and the personality traits including *Action Games* ($\rho^2 = 0.18$) as well as *Board Games* ($\rho^2 = 0.21$) impacting to Extraversion. Also *Casino Games* ($\rho^2 = 0.31$) and *Trivia Games* ($\rho^2 = 0.22$) has similar effect. This fits to the study of Chittaranjan et al. [46] stating that the introverts use more internet, keeping in mind that back in 2011 variety of mobile games was clearly smaller but internet usage less common. *Casino Games* gained comparably high impact to all personality traits, including especially Conscientiousness ($\rho^2 = 0.45$) and Emotional Stability ($\rho^2 = 0.52$). This is also the case with Agreeableness ($\rho^2 = 0.28$), a result that is in line with Phillips et al. [58] report in 2006 where people with low agreeableness used more smartphone games. This is a trend visible in our study with high ρ^2 towards games.

Interestingly, *Educational Games* have impact on Extraversion ($\rho^2 = 0.10$), Agreeableness ($\rho^2 = 0.17$), and Emotional Stability ($R^2 = 0.27$), but not to Conscientiousness ($\rho^2 = 0.04$) or Intellect/Imagination ($\rho^2 = 0.03$). Similar results can be seen with *Trivia Games* affecting comparably to all the other personality traits expect Intellect/Imagination ($\rho^2 = 0.03$). This indicates towards not all games being equal in their impacts to the personality traits. In their later study in 2013, Chittaranjan et al. [47] found extroverts preferring less use of any sort of games, which we can agree on the basis of our study.

Table 7 presents categories having the least impact (the smallest ρ^2) as individuals to the Big Five factors, noting that, together with all the categories, they may still have stronger combined impact. These categories are surprising in a way they being also some of the most popular and widely used categories: *Entertainment*, *Finance*, *Lifestyle*, *Music and Audio*, and so on. In comparison, also highly used category *Communications* has alone notably impact to the personality traits. Chittaranjan et al. [46] state their categories Maps, Camera, Chat and Games have the least impact on application usage (again, in 2011), whereas we see Chat (Communication) and Games as the most impacting categories in our data, Maps and Camera operations staying more neutral or having only a minor impact.

Especially interesting here is that those very same categories are closely linked in the cultural values in previous work [2], where cultures were presented by the Hofstede's Cultural Value Survey Model (VSM) providing six dimensions meant to describe populations: *Power distribution*, *Individualism vs collectivism*, *Masculinity vs femininity*, *Uncertainty avoidance*, *Long vs short-term orientation*, and *Indulgence vs restraint*. In that study, *Entertainment* applications and other leisure related categories, such as *Travel and Local*, *Health and Fitness*, and *Music and Audio*, linking closely to the individualistic cultures with low hierarchies. These same categories have almost zero effect ($\rho^2 \leq 0.03$) to all the Big Five personality factors, even if "individualism" could easily be considered also a feature of a individual, not only populations. Thus, we can conclude that, indeed, we are measuring two different aspects of people's lives here: personality and the culture in the other hand. It is interesting to observe that different features in people's personalities and cultures lead them to use different apps.

8. Discussion

Challenges of collecting personality traits. Personality traits are considered somewhat stable throughout someone's life, yet able to change if someone's motivated to do so over a long period of time. However, the 50-item survey for catching the Big Five personality traits, even if done only once, is very long. To collect such data it might be considerable to look also the shorter surveys and whether their accuracy can be compared to the 50-item version. We asked users of the Carat platform to perform the long Big Five questionnaire through their smartphones, and this may have also decreased the number of end participants. We estimated that it can take at least 20 min to answer all the questions through the app, which was informed to the user before they started the survey. This might have limited people's willingness to consider participating to the survey. Nonetheless, someone's personality trait, being slowly changing, means we do not need to check the Big Five that often.

Data representation. Based on the sample population discussed in the beginning of this paper, our gathered Big Five answers were well in distributed in terms of standard deviations of the traits. However, a cross-comparison with multiple datasets would validate whether our data is truly representative of smartphone users. Even so, it is important to remember that smartphone users – in our case, Android users only – do not describe the full human populations. Not everyone can afford a smartphone even if the prices have decreased in the past years, and people with some disabilities and elderly people are easily left out of the data samples collected only through the smartphones. Thus, our study represents only the Android smartphone users who are tech savvy, who have installed the Carat app and have been willing to participate in our study.

Privacy concerns of application usage data. Access to the list of installed applications allows unique identification of a device or a user in a population: 99.4% of all users have unique usage patterns among the top 60 globally used apps [59]. As shown by our findings, personality traits are tightly coupled with app usage, and we must therefore consider the privacy implications of the application usage data. For example, the advertising industry routinely collects data and classifies individuals into groups for targeted advertising. Knowledge of personality traits allows more precise targeting, as it can be used as an explaining factor behind user decisions, leading to more accurate interest predictions. On the other hand, we have shown that at least in the case of personality, analysing app categories instead of the app data itself leads to similarly accurate results. This means that user privacy can even be protected if data analysis apps transform collecting only the category-level aggregated information.

Smartphone data and their utility in personality studies. Smartphones are promising for improving our understanding of individuals personality traits. Some other previous studies show that, social media interactions in Facebook [41], YouTube usage [13], and Twitter activity [40] are linked to the Big Five personality traits. Previously, also cultural values [2] and demographic factors [27] have been shown to have a crucial impact how people chose apps and what smartphone functionalities they utilise. We argue that understanding users through smartphone data is already well-established and important field, however, more research is needed to fully cover all the aspects of personality and technology usage preferences. In this work, we further found that the used application categories are accurate up to 96% in predicting Big Five personality traits, establishing background for later studies of how specific demographic or cultural groups compare to personal preferences in the case of smartphone and technology usage.

Design implications. Understanding user's personality traits have various different implications into the smartphone apps design instantly. Distinct group of smartphone users can be identified by analysing app usage and demographics such as age, gender, and others, play significant role on app usage behaviour [27]. Application recommendation system can be boosted by more personal details. A study has been done to predict the popularity of different apps based on their usage to estimate the location-wise point of interest for these applications [60]. New apps and smartphone functionalities can be better marketed and targeted directly to the right audience. In addition, sociological research studying technological adaptation and describing smartphone users in the wild may benefit similar crowdsourced approach for delivering personality traits through smartphone usage behaviour without heavy questionnaires. In general, understanding that the underlying factors such as personality clearly have an impact on how we utilise technology, can lead technology and app designers into the better products and services.

Limitations. Our sample (N = 739) is fairly large, however the results presented here may not generalise to a wider audience and should be interpreted as a snapshot of personality traits from this precise sample. It remains unclear whether what we observed is sustained on the long run, and whether our sample is influenced by the novelty effect for participating on a study such as Carat. One deeper challenge is the sustainability of actual Big Five data collection and how often we collect such data. Answering 50 questions on a regular basis is unlikely to be appreciated for a prolonged period of time. Hence using smartphones' app usage to predict personality traits is enticing. Nonetheless, given the recent interest in reducing the amount of time spent around the smartphone (e.g., Google's Digital Wellbeing, Apple's Screen Time) one must also consider other means of profiling that may be beyond a smartphone, e.g., wearables or IoT sensors at home or within a city. But that is a challenge for another day.

Future work. The open challenge here is to find a prediction model that can capture differences in personality-driven smartphone usage in more accurate way. In addition, considering the limitations of the Big Five model to define reasons behind person's everyday behaviour, also other factors, such as demographics, cultural values, mood, daily routines, health factors, and individual deeds, should be studied together with the personality traits. However, it may be challenging to have a data set wide enough to provide all the possible factors influencing smartphone usage. In any case, our results together with the previous work we have discussed in this paper, show that smartphones can accurately be used to study and describe people — in our case, especially personality traits.

9. Conclusion

In this paper, we concluded analysis with 739 Android users who answered the 50-item IPIP standardised Big Five questionnaire and provided their app usage from six months, including 7852 apps belonging into 41 Google Play categories. We compared the users' Big Five answers on how they used their apps, both in terms of apps alone and as app categories. We found out that whether using the apps themselves or app categories, the prediction fit of the results is largely the same. This indicates that studying aggregated category usage is, indeed, accurate enough for determining differences in the app usage. The prediction model presented in this papers shows that the Big Five personality traits can be predicted from smartphone app usage on average with 9%–14% error indicating prediction fit 86%–91%. For the best quarter of predictions, we get 94%–95% fit (relative error 4%–6%), however, for the worst quarter of predictions, our model is barely better than the random guess. Our results indicate that, indeed, there is an under-laying effect of personality on people's app choices and smartphone usage.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

This research is supported by Academy of Finland 6Genesis Flagship (grant 318927), SENSATE (grants 316253, 320089), Infotech Emerging Project, and Nokia Foundation (Jorma Ollila Grant for Dr Ella Peltonen).

References

[1] D.T. Wagner, A. Rice, A.R. Beresford, Device analyzer: Large-scale mobile data collection, SIGMETRICS Perform. Eval. Rev. 41 (4) (2014) 53–56.
[2] E. Peltonen, E. Lagerspetz, J. Hamberg, A. Mehrotra, M. Musolesi, P. Nurmi, S. Tarkoma, The hidden image of mobile apps: geographic, demographic, and cultural factors in mobile usage, in: 20th Int. Conference on Human-Computer Interaction with Mobile Devices and Services, ACM, 2018.
[3] I.R. Felix, L.A. Castro, L.-F. Rodriguez, O. Banos, Mobile sensing for behavioral research: A component-based approach for rapid deployment of sensing campaigns, Int. J. Distrib. Sens. Netw. 15 (9) (2019).
[4] A.J. Oliner, A.P. Iyer, I. Stoica, E. Lagerspetz, S. Tarkoma, Carat: Collaborative energy diagnosis for mobile devices, in: 11th ACM Conference on Embedded Networked Sensor Systems, ACM, 2013, p. 10.
[5] D. Ferreira, V. Kostakos, A.K. Dey, AWARE: mobile context instrumentation framework, Front. ICT 2 (2015) 6.
[6] Y. Jia, B. Xu, Y. Karanam, S. Voida, Personality-targeted gamification: A survey study on personality traits and motivational affordances, in: CHI Conference on Human Factors in Computing Systems, in: CHI '16, ACM, New York, NY, USA, 2016, pp. 2001–2013.
[7] M. Tkalcic, L. Chen, Personality and recommender systems, in: Recommender Systems Handbook, Springer, 2015, pp. 715–739.
[8] M. Naftali, L. Findlater, Accessibility in context: Understanding the truly mobile experience of smartphone users with motor impairments, in: 16th Int. ACM SIGACCESS Conference on Computers & Accessibility, in: ASSETS '14, ACM, New York, NY, USA, 2014, pp. 209–216.
[9] C. Stachl, S. Hilbert, J.-Q. Au, D. Buschek, A. De Luca, B. Bischl, H. Hussmann, M. Bühner, Personality traits predict smartphone usage, Eur. J. Personal. 31 (6) (2017) 701–722.
[10] L.R. Goldberg, J.A. Johnson, H.W. Eber, R. Hogan, M.C. Ashton, C.R. Cloninger, H.G. Gough, The international personality item pool and the future of public-domain personality measures, J. Res. Personal. 40 (1) (2006) 84–96.
[11] L.R. Goldberg, et al., A broad-bandwidth, public domain, personality inventory measuring the lower-level facets of several five-factor models, Personal. Psychol. Eur. 7 (1) (1999) 7–28.
[12] M. Böhmer, B. Hecht, J. Schöning, A. Krüger, G. Bauer, Falling asleep with angry birds, facebook and kindle: a large scale study on mobile application usage, in: 13th Int. Conference on Human Computer Interaction with Mobile Devices and Services, ACM, 2011, pp. 47–56.

[13] J.E. Klobas, T.J. McGill, S. Moghavvemi, T. Paramanathan, Compulsive youtube usage: A comparison of use motivation and personality effects, Comput. Hum. Behav. 87 (2018) 129–139.
[14] D. Ferreira, A.K. Dey, V. Kostakos, Understanding human-smartphone concerns: a study of battery life, in: Int. Conference on Pervasive Computing, Springer, 2011, pp. 19–33.
[15] K. Athukorala, E. Lagerspetz, M. Von Kügelgen, A. Jylhä, A.J. Oliner, S. Tarkoma, G. Jacucci, How carat affects user behavior: implications for mobile battery awareness applications, in: SIGCHI Conference on Human Factors in Computing Systems, ACM, 2014, pp. 1029–1038.
[16] D.A. Rohani, M. Faurholt-Jepsen, L.V. Kessing, J.E. Bardram, Correlations between objective behavioral features collected from mobile and wearable devices and depressive mood symptoms in patients with affective disorders: Systematic review, JMIR mHealth uHealth 6 (8) (2018) e165.
[17] K. Opoku Asare, A. Visuri, D.S.T. Ferriera, Towards early detection of depression through smartphone sensing, in: ACM Int. Joint Conference on Pervasive and Ubiquitous Computing, in: UbiComp/ISWC '19, ACM, New York, NY, USA, 2019, pp. 1158–1161.
[18] R. Wang, W. Wang, A. daSilva, J.F. Huckins, W.M. Kelley, T.F. Heatherton, A.T. Campbell, Tracking depression dynamics in college students using mobile phone and wearable sensing, in: Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies, Vol. 2, (1) ACM, 2018, p. 43.
[19] S. Aghanavesi, D. Nyholm, M. Senek, F. Bergquist, M. Memedi, A smartphone-based system to quantify dexterity in parkinson's disease patients, Inform. Med. Unlocked 9 (2017) 11–17.
[20] E. Kuosmanen, V.K. Kan, A. Visuri, A. Boudjelthia, L. Krizou, D. Ferreira, Measuring parkinson's disease motor symptoms with smartphone-based drawing tasks, in: Int. Workshop on Computing for Well-Being, ACM, 2019.
[21] S. Aram, A. Troiano, E. Pasero, Environment sensing using smartphone, in: IEEE Sensors Applications Symposium, IEEE, 2012, pp. 1–4.
[22] K. Opoku Asare, T. Leikanger, C. Schuss, S. Klakegg, A. Visuri, D. Ferreira, S3: environmental fingerprinting with a credit card-sized NFC powered sensor board, in: Int. Conference on Human-Computer Interaction with Mobile Devices and Services Adjunct, ACM, 2018, pp. 298–305.
[23] E. Nemati, C. Batteate, M. Jerrett, Opportunistic environmental sensing with smartphones: a critical review of current literature and applications, Curr. Environ. Health Rep. 4 (3) (2017) 306–318.
[24] J. Venkatanathan, D. Ferreira, M. Benisch, J. Lin, E. Karapanos, V. Kostakos, N. Sadeh, E. Toch, Improving users' consistency when recalling location sharing preferences, in: Human-Computer Interaction – INTERACT 2011, Berlin, Heidelberg, 2011, pp. 380–387.
[25] I. Andone, K. Błaszkiewicz, M. Eibes, B. Trendafilov, C. Montag, A. Markowetz, How age and gender affect smartphone usage, in: ACM Int. Joint Conference on Pervasive and Ubiquitous Computing, ACM, 2016, pp. 9–12.
[26] A. Berenguer, J. Goncalves, S. Hosio, D. Ferreira, T. Anagnostopoulos, V. Kostakos, Are smartphones ubiquitous?: An in-depth survey of smartphone adoption by seniors, IEEE Consum. Electron. Mag. 6 (1) (2016) 104–110.
[27] S. Zhao, J. Ramos, J. Tao, Z. Jiang, S. Li, Z. Wu, G. Pan, A.K. Dey, Discovering different kinds of smartphone users through their application usage behaviors, in: ACM Int. Joint Conference on Pervasive and Ubiquitous Computing, ACM, 2016, pp. 498–509.
[28] A. Hiniker, S.N. Patel, T. Kohno, J.A. Kientz, Why would you do that? predicting the uses and gratifications behind smartphone-usage behaviors, in: ACM Int. Joint Conference on Pervasive and Ubiquitous Computing, ACM, 2016, pp. 634–645.
[29] M. Salehan, A. Negahban, Social networking on smartphones: When mobile phones become addictive, Comput. Hum. Behav. 29 (6) (2013) 2632–2639.
[30] S. Bødker, Third-wave HCI, 10 years later-participation and sharing, Interactions 22 (5) (2015) 24–31.
[31] Y.-K. Lee, C.-T. Chang, Y. Lin, Z.-H. Cheng, The dark side of smartphone usage: Psychological traits, compulsive behavior and technostress, Comput. Human Behav. 31 (2014) 373–383.
[32] M. Takao, S. Takahashi, M. Kitamura, Addictive personality and problematic mobile phone use, CyberPsychol. Behav. 12 (5) (2009) 501–507.
[33] K. Katevas, I. Arapakis, M. Pielot, Typical phone use habits: Intense use does not predict negative well-being, in: Int. Conference on Human-Computer Interaction with Mobile Devices and Services, ACM, 2018, p. 11.
[34] A. Ehrenberg, S. Juckes, K.M. White, S.P. Walsh, Personality and self-esteem as predictors of young people's technology use, Cyberpsychol. Behav. 11 (6) (2008) 739–741.
[35] C. Montag, K. Błaszkiewicz, R. Sariyska, B. Lachmann, I. Andone, B. Trendafilov, M. Eibes, A. Markowetz, Smartphone usage in the 21st century: who is active on whatsapp? BMC Res. Notes 8 (1) (2015) 331.
[36] O.P. John, S. Srivastava, et al., The big five trait taxonomy: History, measurement, and theoretical perspectives, in: Handbook of Personality: Theory and Research, Vol. 2, (1999) Guilford, 1999, pp. 102–138.
[37] L.R. Goldberg, The structure of phenotypic personality traits, Amer. Psychol. 48 (1) (1993) 26.
[38] T.A. Judge, R. Klinger, L.S. Simon, I.W.F. Yang, The contributions of personality to organizational behavior and psychology: Findings, criticisms, and future research directions, Soc. Personal. Psychol. Compass 2 (5) (2008) 1982–2000.
[39] D.Y. Wohn, R. LaRose, Effects of loneliness and differential usage of facebook on college adjustment of first-year students, Comput. Educ. 76 (2014) 158–167.
[40] D. Quercia, M. Kosinski, D. Stillwell, J. Crowcroft, Our Twitter profiles, our selves: Predicting personality with Twitter, in: IEEE 3rd Int. Conference on Privacy, Security, Risk and Trust, IEEE, 2011, pp. 180–185.
[41] A. Ortigosa, R.M. Carro, J.I. Quiroga, Predicting user personality by mining social interactions in facebook, J. Comput. Syst. Sci. 80 (1) (2014) 57–71.
[42] A. Mehrotra, V. Pejovic, J. Vermeulen, R. Hendley, M. Musolesi, My phone and me: understanding people's receptivity to mobile notifications, in: CHI Conference on Human Factors in Computing Systems, ACM, 2016, pp. 1021–1032.
[43] L.R. Goldberg, The development of markers for the big-five factor structure., Psychol. Assess. 4 (1) (1992) 26.
[44] L.H. Gilpin, D.M. Olson, T. Alrashed, Perception of speaker personality traits using speech signals, in: CHI Conference on Human Factors in Computing Systems, ACM, 2018, p. LBW514.
[45] L. Gou, M.X. Zhou, H. Yang, Knowme and shareme: understanding automatically discovered personality traits from social media and user sharing preferences, in: SIGCHI Conference on Human Factors in Computing Systems, ACM, 2014, pp. 955–964.
[46] G. Chittaranjan, J. Blom, D. Gatica-Perez, Who's who with big-five: Analyzing and classifying personality traits with smartphones, in: 2011 15th Annual Int. Symposium on Wearable Computers, IEEE, 2011, pp. 29–36.
[47] G. Chittaranjan, J. Blom, D. Gatica-Perez, Mining large-scale smartphone data for personality studies, Pers. Ubiquitous Comput. 17 (3) (2013) 433–450.
[48] R. de Oliveira, A. Karatzoglou, P. Concejero Cerezo, A. Armenta Lopez de Vicuña, N. Oliver, Towards a psychographic user model from mobile phone usage, in: CHI'11 Extended Abstracts on Human Factors in Computing Systems, ACM, 2011, pp. 2191–2196.
[49] Y.-A. de Montjoye, J. Quoidbach, F. Robic, A.S. Pentland, Predicting personality using novel mobile phone-based metrics, in: Int. Conference on Social Computing, Behavioral-Cultural Modeling, and Prediction, Springer, 2013, pp. 48–55.
[50] C. Stachl, Q. Au, R. Schoedel, S.D. Gosling, G.M. Harari, D. Buschek, S.T. Völkel, T. Schuwerk, M. Oldemeier, T. Ullmann, et al., Predicting personality from patterns of behavior collected with smartphones, Proc. Natl. Acad. Sci. 117 (30) (2020) 17680–17687.
[51] J. Staiano, B. Lepri, N. Aharony, F. Pianesi, N. Sebe, A. Pentland, Friends don't lie: inferring personality traits from social network structure, in: ACM Conference on Ubiquitous Computing, ACM, 2012, pp. 321–330.

[52] W. Wang, G.M. Harari, R. Wang, S.R. Müller, S. Mirjafari, K. Masaba, A.T. Campbell, Sensing behavioral change over time: Using within-person variability features from mobile sensing to predict personality traits, in: Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies, Vol. 2, (3) ACM New York, NY, USA, 2018, pp. 1–21.
[53] N. Gao, W. Shao, F.D. Salim, Predicting personality traits from physical activity intensity, Computer 52 (7) (2019) 47–56.
[54] C. Stachl, Q. Au, R. Schoedel, D. Buschek, S. Völkel, T. Schuwerk, M. Oldemeier, T. Ullmann, H. Hussmann, B. Bischl, et al., Behavioral patterns in smartphone usage predict big five personality traits, 2019.
[55] D. Schmitt, J. Allik, R. McCrae, V. Benet, J. Veríssimo, U.-D. Reips, The geographic distribution of big five personality traits, J. Cross-Cultural Psychol. 38 (2007) 173–212.
[56] B. Olabenjo, Applying naive Bayes classification to google play apps categorization, 2016, arXiv preprint arXiv:1608.08574.
[57] D. Azucar, D. Marengo, M. Settanni, Predicting the big 5 personality traits from digital footprints on social media: A meta-analysis, Personal. Individ. Differ. 124 (2018) 150–159.
[58] J.G. Phillips, S. Butt, A. Blaszczynski, Personality and self-reported use of mobile phones for games, CyberPsychol. Behav. 9 (6) (2006) 753–758.
[59] P. Welke, I. Andone, K. Blaszkiewicz, A. Markowetz, Differentiating smartphone users by app usage, in: ACM Int. Joint Conference on Pervasive and Ubiquitous Computing, in: UbiComp '16, ACM, New York, NY, USA, 2016, pp. 519–523.
[60] D. Yu, Y. Li, F. Xu, P. Zhang, V. Kostakos, Smartphone app usage prediction using points of interest, in: Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies, Vol. 1, (4) ACM New York, NY, USA, 2018, pp. 1–21.