Poster: Inference of Big-Five Personality Using Large-scale Networked Mobile and Appliance Data

Catherine Tong[†], Gabriella M. Harari[§], Angela Chieh[‡],
Otmane Bellahsen[‡], Matthieu Vegreville[‡], Eva Roitmann[‡], Nicholas D. Lane[†*]
[†]University of Oxford, [§]Stanford University, [‡]Nokia Digital Health - Withings, [*]Nokia Bell Labs

ABSTRACT

We present the first large-scale (9270-user) study of data from both mobile and networked appliances for Big-Five personality inference. We correlate aggregated behavioral and physical health features with personalities, and perform binary classification using SVM and Decision Tree. We find that it is possible to infer each Big-Five personality at accuracies of 75% from this dataset despite its size and complexity (mix of mobile and appliance) as prior methods offer similar accuracy levels. We would like to achieve better accuracies and this study is a first step towards seeing how to model such data.

CCS CONCEPTS

• **Human-centered computing** → *Ubiquitous and mobile devices*; *Ubiquitous and mobile devices*; • **Applied computing** → *Psychology*; *Psychology*;

1 INTRODUCTION

Personality traits describe a person's characteristic patterns of thinking, feeling, and behaving. The most widely used model for measuring personality focuses on the Big-Five personality traits - extraversion, agreeableness, conscientiousness, emotionalstability and openness. From an applied perspective, knowing a user's personality could be useful for device customization (e.g. personalization based on psychological characteristics) and for understanding contributing factors to their wellbeing. Traditionally, personality traits are measured using self-report surveys, which can be difficult to scale up in commercial settings. However, mobile sensing technologies permit unobtrusive collection of real-world behavioral patterns [3]. Such technologies may be used to classify personality traits passively, without requiring surveys. Prior research examing this topic [1] indicates the capabilities of using mobile data for their classification. However, these studies used small datasets (under 100 users) and focused on social data (e.g. SMS logs) from only mobile devices. In this study, we aim to understand if methods used in prior studies [1] can be applied to an increasingly popular form of data - large-scale datasets with sparsely collected passive behavioural data.

MobiSys '18, June 10–15, 2018, Munich, Germany

ACM ISBN 978-1-4503-5720-3/18/06.
https://doi.org/10.1145/3210240.3210823

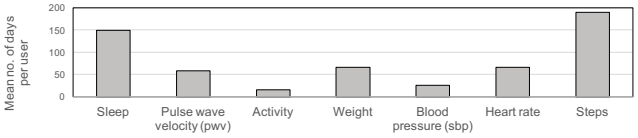

Figure 1: Data per modality. Left to right: modalities most strongly correlated with Big-Five traits to the least.

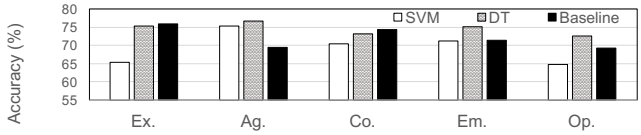

Figure 2: Classification accuracies.

2 METHODOLOGY

Data. 9270 Withings-device users completed a TIPI survey [2] to measure their Big-Five personality traits. We collect sensor data from all Withings devices each user owns (e.g. weight scale, sleep tracker) in daily resolutions over a window of up to 1 year. The data is *passive* with some inferred behavioral statistics (e.g. sleep-stage durations). Each modality is in varying amounts depending on individual usage (Fig. 1).

Data Analysis. We split sensor data across months, aggregating events to extract features, which we use to perform Pearson's correlation with personality traits.

Classification. We consider a binary classification task for each Big-Five personality, using the median to group each into 2 classes. We use features most strongly correlated with personality and at least 500 user-months as input features to SVM and Decision Tree.

3 RESULTS

Correlations. We find a number of strong and significant correlations between features and traits ($|r|>0.1$ and $p<0.01$), e.g. mean and variance in bed-in time with co. and em.; mean weight with ag.

Classification. We use 17 input features from sleep, pwv, weight and sbp. The current study achieves similar accuracies to those observed in [1] (Fig. 2).

Summary. Our study replicates the personality classification analysis done in [1]. We show the application of SVM and Decision Tree onto a new type of large-scale complex dataset to binarily classify Big-Five personality traits is able to achieve accuracies of 74.6%.

REFERENCES

[1] Chittaranjan, G., et al. Who's who with big-five: Analyzing and classifying personality traits with smartphones. *ISWC '11*, 29–36.
[2] Gosling, S., et al. A very brief measure of the big-five personality domains. 504–528.
[3] Harari, G., et al. Using smartphones to collect behavioral data in psychological science: Opportunities, practical considerations, and challenges. 838–854.