

Going Beyond Traits: Multimodal Classification Of Personality States In The Wild

Kyriaki Kalimeri
University of Trento, CIMEC
Bruno Kessler Foundation
kalimeri@fbk.eu

Bruno Lepri
Bruno Kessler Foundation
MIT Media Lab
lepri@fbk.eu

Fabio Pianesi
Bruno Kessler Foundation
pianesi@fbk.eu

ABSTRACT

Recent studies in social and personality psychology introduced the notion of personality states conceived as concrete behaviors that can be described as having the same contents as traits. Our paper is a first step towards addressing automatically this new perspective. In particular, we will focus on the classification of excerpts of social behavior into personality states corresponding to the Big Five traits, rather than focusing on the more traditional goal of using those behaviors to directly infer about the personality traits of the person producing them. The multimodal behavioral cues we exploit were obtained by means of the Sociometric Badges worn by people working at a research institution for a period of six weeks. We investigate the effectiveness of cues concerning acted social behaviors as well as of other situational characteristics for the sake of personality state classification. The encouraging results show that our classifiers always, and sometimes greatly, improve the performances of a random baseline classifier (from 1.5 to 1.8 better than chance). At a general level, we believe that these results support the proposed shift from the classification of personality traits to the classification of personality states.

Categories and Subject Descriptors

H.1.2 [Information Systems]: MODELS AND PRINCIPLES—*User/Machine Systems*; J.4 [Computer Applications]: SOCIAL AND BEHAVIORAL SCIENCES

Keywords

SVM; personality states; classification; mobile sensors; social computing

1. INTRODUCTION

In everyday life, people constantly describe others as being more or less talkative, bold or sociable; more or less angry or vulnerable to stress; more or less self-determined or influenced by external situations. Moreover, people exploit these descriptors in their everyday life to explain and/or predict others' behavior, attaching them to well-known as well as new acquaintances, friends, colleagues,

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than the author(s) must be honored. Abstracting with credit is permitted. To copy otherwise, or to publish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.

ICMI '13, December 9–13, 2013, Sydney, Australia

Copyright is held by the owner/author(s). Publication rights licensed to ACM.

ACM 978-1-4503-2129-7/13/12 ...\$15.00.

<http://dx.doi.org/10.1145/2522848.2522878>.

bosses. The attribution of stable personality traits to others and their usage in predicting and explaining people's behavior is a fundamental characteristic of human naive psychology. As agents that in increasingly many and varied ways participate in, and affect, the lives of humans, computing systems too need to understand and predict their human parties' behavior. Arguably, the achievement of this goal requires something similar to the type of knowledge and expectations that personality traits encode.

Scientific psychology has developed a view of personality as a higher level abstraction encompassing traits, sets of stable dispositions towards action, belief and attitude formation. Personality traits differ across individuals; are relatively stable over time and influence behavior. Between-individual differences in behavior, belief and attitude can therefore be captured in terms of the dispositions/personality traits that are specific to each individual, this way providing a powerful descriptive and predictive tool that has been widely exploited by, for example, clinical and social psychology, educational psychology and organizational studies. A well-known and very influential example of this approach is the Big Five [6], which owes its name to the five traits it takes as constitutive of people's personality: Extraversion vs. Introversion (sociable, assertive, playful vs. aloof, reserved, shy); Emotional stability vs. Neuroticism (calm, unemotional vs. insecure, anxious); Agreeableness vs. Disagreeable (friendly, cooperative vs. antagonistic, faultfinding); Conscientiousness vs. Unconscientiousness (self-disciplined, organized vs. inefficient, careless); Openness to experience (intellectual, insightful vs. shallow, unimaginative). Over the last 50 years the Big Five has become a standard in psychology and experiments using the Big Five have repeatedly confirmed the influence of personality traits on many aspects of individual behavior including leadership, general job performance, attitude toward machines, sales ability, teacher effectiveness, etc. More recently, Big Five traits have been shown to influence the human/technology relationship, affecting attitudes towards computers in general as well as towards specific technologies such as adaptive systems[11], conversational agents [26], tutoring systems [30] etc. In the end, finding means to automatically obtain information about people's personalities has become crucial to let machines act proactively and to endow them with the folk-psychological capability of explaining/predicting people's behaviors [1].

For all these reasons, several works have recently been exploring automated personality analysis [16, 2, 20] often targeting the Big Five model of personality [6]. All of them have been exploiting excerpts of a person's behavior (e.g., thin slices of behavior) to provide the machine-equivalent of judgments about his/her personality. The idea underlying this approach is that actual behavior is largely determined by personality so that excerpts of actual behavior can be (more or less easily) classified as belonging to an intro-

vert/extravert, neurotic/stable person, etc. This formulation of the personality problem suffers of fundamental shortcoming that such straightforward relation between personality and behavior does not hold: traits are stable and enduring properties but people's behavior is ever changing. An extravert might, on occasions, be less talkative or attempt less to attract social attention; a neurotic person need not always react anxiously, etc. One might even go as far as to admit that, in the end, all individuals routinely express in their behavior all the levels of a given trait. Such tension between the invariance of personality traits and the natural variability of behavior in concrete situations can be resolved by considering within-person variability as noise that has to be canceled out by, e.g., employing larger behavioral samples; an approach commonly employed both in psychological and computational works on personality. Although this move is surely recommended, it can be argued that it cannot by itself solve the problem because within-person behavioral variability is more than just noise; on the contrary, stemming from the interaction between enduring traits and changing *situational properties*, it can give a valuable contribution to personality prediction and to the understanding of the personality/behavior relationship [8].

Fleeson [8, 9] exploited the notion of *personality states* conceived as concrete behaviors (including ways of acting, feeling and thinking) that can be described as having the same contents as traits. A personality state is, therefore, a *specific behavioral episode* wherein a person behaves more or less introvertedly/extravertedly, more or less neurotically etc. Personality traits can then be reconstructed as density distributions over personality states conditioned on situational properties. According to this perspective, between-person differences (that is, personality) amount to differences among personality state distributions: e.g., an introvert does not differ from an extravert because he/she never engages in extravert behaviors, but because he/she does so in a different manner than an introvert. Such an approach opens interesting prospects for, while broadening, the task of automatically computing personality. In the first place, by emphasising situational characteristics (together with personality) as one of the key determinants of actual behavior, it provides the necessary flexibility to ground the behavior-personality relationship. Secondly, such flexibility can be expected to ease not only the task of predicting personality from behaviors, but also the converse task of predicting/explaining behaviors from people's personality.

This paper is a first step towards addressing this new perspective, starting from the automatic classification of personality states. In particular, we will focus on the classification of excerpts of social behavior into personality *states* corresponding to the Big Five traits, rather than dealing with the more traditional goal of using behaviors to infer about personality *traits*. The behavioral cues we will exploit were obtained by means of Sociometric Badges [22] worn by people working at a research institution in Italy for a period of six weeks. By using them we will investigate the effectiveness of cues concerning acted social behaviors, for instance the number of interacting people and the number of people in close proximity, as well as of other situational characteristics such as time spent in the canteen or in meetings and so on, for the sake of personality state classification.

2. RELATED WORKS

Recently, researchers in social and ubiquitous computing have started exploring the wealth of behavioral data made available by cameras and microphones in the environment [23, 16, 20, 3, 2], smartphones [5, 7], wearable sensors [22, 21] in order to automatically classify personality traits.

Pianesi et al. [23] provided evidence about the feasibility of multimodal analysis of personality traits and the influence of the social context. In particular, they proposed an automatic prediction of two personality traits (Extraversion and Locus of Control), based on simple non-verbal features, in a group interaction context. They started from the assumption that: a) personality shows up in the course of social interaction and, b) 1-minute-long sequences of social behavior are enough to classify personality traits. On the same line, Lepri et al. [16] investigated the suitability of medium-grained meeting behaviors, namely speaking time and social attention, for the automatic classification of the Extraversion personality trait. Later on, Kalimeri et al. [15] presented the first attempt to model the influence of the social context expressed in terms of the personality trait's effect, of interaction partners on the behavior of the target subject using bayesian networks. Experimental results confirm that these behaviors are indeed effective for the automatic detection of Extraversion.

Mohammadi et al. [20] showed that prosodic features may be used to predict personality assessments of human experts on a collection of 640 speech samples. Exploiting sociometric badges, Olguin et al. [22] found that Extraversion and Neuroticism were positively correlated with degree, closeness, betweenness, and eigenvector centrality measures. Moreover, they found a negative correlation between Conscientiousness and betweenness centrality. Using mobile phone data, Chittaranjan et al. [5] exploited actor-based features (e.g. number and duration of calls, BT hits, etc.) in order to automatically classify personality traits. Their results revealed some interesting trends: extroverts were more likely to receive calls and to spend more time on them, while features pertaining to outgoing calls were found to be not predictive of the Big Five traits.

The general approach of all these previous works is to isolate promising correlates of the targeted traits for classification or regression. All these works adopted the so-called person-perspective on personality and target personality traits prediction or classification and not personality states prediction or classification.

In a study close to our objectives, Mehl et al. [19] attempted to examine the expression of personality in its natural habitat tracking 96 participants over 2 days using EAR (Electronically Activated Recorder), a portable electronic recorder which captures snippets of ambient sounds in participants' immediate environments. Participants' Big Five scores were correlated with EAR-derived information on their daily social interactions, locations, activities, moods, and language use; these daily manifestations were generally consistent with trait definition and (except for Openness) often gender-specific. To identify implicit folk theories about daily manifestations of personality, the authors correlated the EAR-derived information with impressions of participants based on their EAR sounds. Their findings point to the importance of naturalistic observation studies on how personality is expressed and perceived in the natural stream of everyday behavior. EAR fills an important gap in the psychological study of person-environment interactions (see for a review [29]); however, this method has obvious limitations. For example, it allows only the "acoustic observation" of daily life, and many interesting social phenomena can be grasped by other means. Using the Sociometric Badges, we are able to provide a multi-view observation and description of daily life capturing from co-presence in a place to face-to-face interaction and speech behaviors. More recently, Staiano et al. [28] presented a first attempt to predict the personality states on a multimodal meeting scenario corpus validating generative and discriminating models. The limited size of the corpus however, as well as its nature did not provide a wide range of social situations, preventing a full exploitation and understanding of the personality states.

3. SOCIOMETRIC BADGES CORPUS

For this study we exploited the SocioMetric Badges Corpus [17], a multimodal corpus specifically designed to capture the psychological and situational aspects of the daily lives of employees in an organizational structure. The data were collected in a research institute for six weeks and involved a sample of 54 subjects (both administrative employees and researchers) during their working hours. Males predominated (90.8%) while the average age was 36.83 years with a SD of 8.61 years. Out of the 54 subjects, 37 subjects were researchers and employees, 7 had a leading role while 10 were doctoral students. At the beginning of the six week period, the subjects provided information regarding their ties to the other participants, such as friendship level, trust etc.

3.1 Personality Data

Participants' personality states were sampled by means of the ten-item (two items per state) personality inventory TIPI [12] administered online three times per day during the morning (11:00 AM), the afternoon (2:00 PM) and the evening (5:00 PM) using 1-7 point Likert scale. The subjects were instructed to use the questionnaire to describe the way they behaved/felt during the previous 30 minutes. The scores for each state were calculated by summing the raw scores of the two corresponding items, with proper inversion when needed.

Table 1 reports the sample's average, SD, total variance, within-person variance and between-subject variance for each state. Importantly, within-subject variance tends to be always higher than the between-subject one, a trend that is stronger with Extraversion. This pattern is in line with findings in the literature on personality states [8] and emphasizes the importance of shifting the attention to within-subject variations and their dependence on situational factors when addressing the interplay between personality and actual behavioral manifestations.

	<i>E</i>	<i>A</i>	<i>C</i>	<i>E.S</i>	<i>O</i>
<i>Average</i>	4.06	5.12	5.53	5.54	4.50
<i>SD</i>	1.23	0.98	0.99	1.07	1.13
<i>Variance</i>	1.52	0.97	0.98	1.16	1.29
<i>WithinVariance</i>	1.01	0.73	0.69	0.74	0.85
<i>BetweenVariance</i>	0.64	0.61	0.66	0.70	0.72

Table 1: Descriptive Statistics on Personality States from the experience sampling

3.2 Sociometric Data

The Sociometric Badge sensors [22] are equipped with an accelerometer, microphone, bluetooth and infrared sensors that can be used to capture (i) body movements, (ii) prosodic speech features, (iii) proximity to/ colocation with other people and (iv) face-to-face interactions respectively. In this work, we exploit information from all those sensors, but the accelerometer. In detail:

Speech Data The speech signal was sampled with sampling frequency $f_s = 8$ kHz. A number of basic speech measurements such as the signal's amplitude, its standard deviation, minimum and maximum values, mean and variance, were recorded over intervals of 50ms.

Infrared Data The detection of another Infrared (IR) sensor can be used a good proxy for face-to-face interaction. For the IR sensor of one badge to be detected by the IR of another badge, in fact, the two individuals must have a direct line of sight and the receiving badge's IR must be within the transmitting badge's IR signal cone

of height $h \leq 1$ meter and a radius of $r \leq h \tan \theta$, where $\theta = \pm 15^\circ$ degrees. Infrared transmission rate (TR_{ir}) was set to 1Hz.

Bluetooth Data Bluetooth (BT), and in particular the radio signal strength indicator (RSSI), can be used as a coarse indicator of proximity between devices, hence people. In particular, by analyzing our data we found that a BT hit with an RSSI value greater than, or equal to, -80 corresponded to a physical distance between the two sensors, hence the two subjects, of less than 3 meters ("strong signal"). Those BT hits can, therefore, be taken as a good cue for small groups of people gathering at a conversational distance, as in meetings. We therefore distinguished between people being in *close* and in *intermediate* proximity, where the former corresponds to an RSSI range of $[-80, -60]$ (less than one meter, according to our data) and the latter to an RSSI range of $[-85, -80]$ (one to three meters).

Besides measuring colocation and proximity between people, we also addressed spatial localization by means of 17 badges placed at fixed locations of common interest such as the organization's canteen, the cafeteria and meeting rooms. All Sociometric Badges, including base stations, broadcast their ID every five seconds using a 2.4 GHz transceiver ($TR_{radio} = 12$ transmissions per minute).

3.3 E-mail Data

Electronic communication (e-mails) between the participating subjects was recorded. In order to preserve privacy, no information about message content was stored but only information regarding the length of the message body, the time when the email was sent; sender and/or addressees were kept only in case they were participant in the study and anonymized otherwise.

4. FEATURE EXTRACTION

The sociometric and the e-mail data described in the previous section provided the behavioral sequences that would be classified into each personality state. Behavioral sequences were aligned to the ground truth for personality states obtained from the experience samplings (see below) by restricting the former to 30 minute-long temporal windows ending at the time each experience sampling occurred. In this section, we present and discuss the features we identified to represent those behavioral sequences, clustering them according to the sensor type they are based on. Given that personality states are concrete behavioral sequences that can be described as having the same content as the corresponding trait, our initial choices were mainly driven by findings in the psychological and social computing literature concerning "typical" behaviors associated with traits.

4.1 Bluetooth Sensor

In their exploratory analyses Mehl et al. [19], reported that extrovert people spent longer time in conversations and shorter time alone than introvert ones; in the same study, conscientious individual from a student sample were observed to spend more time in class and less time in canteens, bars, or cafeterias.

For each time window and for each subject, we therefore extracted: the number of people in close proximity ($F1$ - see above for our operational definition of close proximity); the number of people at "intermediate proximity" ($F2$); the mean colocation time with others at close and intermediate proximity ($F3$, $F4$); the mean physical distance from other subjects ($F5$) according to the aforementioned operational definitions.

Bluetooth devices were also used to track people's localization with respect to three broad categories of activity places: "Meeting", "Coffee" and "Canteen". Combining this information with the sig-

nal's strength, we extracted the amount of time spent at the canteen ($F6$), at the coffee breaks ($F7$) and at meetings ($F8$).

Moreover, exploiting information subjects had provided in the initial survey about their acquaintances and friends, for each participant and for each window we extracted: the number of friends each participant interacted with ($F9$); the amount of time spent with them ($F10$); the level of global friendship of a given situation/window, computed as the fraction of friends who were present over the total number of present people ($F11$).

4.2 Infrared Sensor

The infrared sensors provide a much finer-grained picture of social interactions than BT, yielding a very good proxy for "face to face" (F2F) interactions. For each subject and for each time window, we extracted: the number of people F2F interacting with the subject ($F12$); the mean duration of the interactions ($F13$); the number of friends the participant F2F interacted with ($F14$); the amount of time spent with them ($F15$); the overall level of the F2F interactions, computed as the fraction of friends over the total number of people the subject had F2F interacted with ($F16$).

4.3 Microphones

Many studies on personality have often emphasized the relevance of non-verbal acoustic features, such as intensity [10, 27]. Certain personality traits have been associated with particular speaking styles; extraverts speak more and louder, while conscientious individuals speak more fluently [10, 27]. Driven by these findings, for each time window and each subject we extracted: the percentage of his/her speaking time ($F17$); the mean energy ($F18$); average ($F19$) and the standard deviation ($F20$) of the amplitude; the average minimum ($F21$) and the average maximum value ($F22$) of the amplitude. The speaking time was estimated as the percentage of the speaking activity in the given time window. The latter was isolated from the background ambient noise using the long-term spectral divergence algorithm, used to discriminate between speaking/non speaking regions, under the assumption that the most significant information is contained in the time-varying spectral magnitude of the signal. According to [25], the long-term spectral envelope (LTSE) as well as the long term spectral divergence (LTSD) were estimated, in order to formulate the decision rule for the voice activity detection. The N-order long-term spectral divergence between speech and noise is defined as the deviation of LTSE with respect to the average noise spectrum magnitude $N(k)$ for the $k - th$ band, with $k = 0, 1, \dots, NFFT-1$, where NFFT is the length of the Fast Fourier Transform.

4.4 E-mail communication

For each subject and for each time window, the following features were extracted: the number of e-mails he/she sent ($F23$); the number of people he/she contacted ($F24$); the consistency of the communication, defined as the average number of the emails sent per recipient ($F25$); the mean length of e-mails ($F26$) measured by the number of characters used in the body text; the mean number of recipients ($F27$).

5. METHODOLOGY

5.1 Ground-truth

In this study the ground truth was obtained considering the self-assessed questionnaires filled by the subjects without additionally labelled material by external judges. The TIPI scores described above were quantized into three classes (Low, Medium, High) using as upper and lower thresholds the median of each state ± 0.5

of the standard deviation. Each class was then randomly sampled, preserving for each of them a number of samples equal to the size of the smallest class. The baseline for our classification experiment (uniform probabilities) is therefore 0.33 accuracy.

5.2 Feature Selection

Feature selection is an important issue for any pattern classification system. In this paper, the feature selection is not urged by dimensionality issues but emerges from the need to automatically select the best determinants of each personality state. In this way, we can derive a qualitative description of the state characteristics while at the same time the noisier features are excluded from the classification task. According to Jain and Zongker [13], amongst the most effective feature selection techniques are the sequential floating search methods (SFS) [24]. There are two main categories of floating search methods: forward (SFFS) and backward (SFBS). In this paper, we employed the forward search (SFFS) to select the optimum feature set for our classification task. The algorithm starts with a null feature set and, for each step, the best feature that satisfies some criterion function is included with the current feature set, e.g., one step of the sequential forward selection (SFS) is performed. The algorithm also verifies the possibility of improvement of the criterion if some feature is excluded. In this case, the worst feature (concerning the criterion) is eliminated from the set, that is, one step of sequential backward selection (SBS) is performed. Therefore, the SFFS proceeds dynamically increasing and decreasing the number of features until the desired dimension is reached.

Table 2 reports the results of the feature selection. For extraversion, the selected cues were related to F2F interaction, speech-related behavior and e-mail communication; with agreeableness, e-mail and speech retain their importance, F2F-related cues are less present and there is an increased importance of proximity- and location-related cues. The latter, though with a partially different distribution, are important also for Conscientiousness along with a few features from e-mail behavior. Such role played by the working environment (e.g., the canteen and the cafeteria) or by working situations (e.g., meetings and breaks) with respect to Conscientiousness states seems to accord well with their being associated with work-oriented responsible, careful and self-disciplined behavior. Emotional Stability was found to be related mostly with speech features and electronic communication, while Openness features extend also to F2F interaction. Following the description of the Agreeableness trait, the agreeable state could be expected to correspond to sympathetic and warm behavior, including generosity, cooperation, willingness to accommodate others wishes, etc. [6]. And indeed, the features selected for this state refer to such qualitative aspects of the social context as friendship encoded by means of (i) number of friends the participant interacted with and (ii) level of friendship in the overall social interactions of the participant. Other selected features describe other aspects of social behaviors such as the time spent interacting in meetings and the number of people who are in proximity.

5.3 Experimental Design

The classification of behavioral sequences into personality states was performed by means of one Support Vector Machine (SVM) [4] with RBF kernel per state. The cost parameters C and the RBF kernel parameters G were estimated by 10-fold cross validation of the training set, while the "one-against-one" multi-class method was used to address the three classes, with six binary models per state each having the same parameters. For each classifier the leave-one-subject-out cross-validation strategy was employed. Hence, for each personality dimension the model was interactively

Indicators of Personality States						
ID	Description	E	A	C	ES	O
F1	#PeopleCloseDist	*	*		*	*
F2	#PeopleInterDist			*		*
F3	DurationCloseDist					
F4	DurationInterDist					
F5	MeanDistance	*		*		
F6	TimeInCanteen		*	*		
F7	TimeInCafeteria	*		*		
F8	TimeInMeetings		*	*		
F9	NumberOfFriends	*	*	*		
F10	TimeWithFriends		*	*		*
F11	LevelFriendship	*	*	*	*	
F12	#People	*	*			*
F13	TimeofInteraction				*	
F14	NumberOfFriends	*		*		
F15	TimeWithFriends	*				*
F16	LevelFriendship			*		*
F17	Speakingtime%	*			*	*
F18	MeanEnergy	*				
F19	MeanAmplitude	*	*	*		*
F20	StandardDeviation		*	*	*	*
F21	MinimumAmp	*	*	*	*	*
F22	MaximumAmp	*	*	*	*	*
F23	EmailsSent	*			*	*
F24	#PeopleContacted	*	*	*	*	*
F25	Consistency	*		*	*	*
F26	MeanEmailLength	*	*		*	
F27	MeanRecipients	*				

Table 2: Indicators chosen for the predictions of each personality dimension.

trained on 53 subjects, evaluated on the remaining subject and the results were eventually averaged.

The experimental setup was designed in order to assess and compare: (i) the predictive power of unimodal behavioral features (e.g., only F2F interaction features from IR; only proximity-based features from BT, and so on); (ii) the role of the social context in terms of the non-verbal behavior of the other subjects; (iii) the effectiveness of feature-level multi-modal fusion and (iv) the role of the social context in terms of non-verbal behavior(after multi-modal fusion in feature-level) of the other subjects. Four experimental configurations are exploited to these extends:

- *Schema I*. Unimodal features are used to predict the state of a given subject, yielding the following conditions: BT, IR, Speech, E-mail. For example, for the classification of extraversion in the BT alternative, we used (see Table 2) the MeanDistance (F5), TimeInCafeteria (F7), Number of Friends(F9) and LevelFriendship (F11).
- *Schema II*. Unimodal + Context. It adds to the unimodal features of *Schema I*'s condition the encoding of the social context. To exemplify, in the BT condition of *Schema II* the feature for subject *X* at a certain time window *T* will consist of the same features as in *Schema I* plus the average values of those same features computed over all the subjects whom *X* F2F-interacted with during *T*.

- *Schema III*. Multimodal. It consists of the following conditions: *Speech - BT*, *Speech-IR*, *Speech - Email*, *IR - BT*, *IR - Email*, *BT - Email*. The personality state of a given subject is predicted from the fusion of the features pertaining to the relevant modalities.
- *Schema IV*. Multimodal + Context. This schema has the same conditions as *Schema III* and adds the social context in the same way as *Schema II* does with *Schema I*.

6. EXPERIMENTAL RESULTS

Table 3 reports the results of our classification experiments, using accuracy as a figure of merit. Most of the results are significantly higher than the baseline classifier that exploits the observed frequencies of the three classes (0.33 per class).

Starting from *Schema I* (see also Figure 1), the Email modality has the highest average accuracy, mainly due to its good performances on Emotional Stability and Openness. The other modalities have similar, though somewhat lower, average accuracies. Of some importance is the 0.54 accuracy obtained by BT modality with Emotional Stability. Concerning personality states, the most noticeable thing is the (comparatively) high average accuracy for Emotional Stability, due to the already mentioned high performance of the Email and the BT modalities with this state.

The addition of unimodal social context by means of *Schema II* does not change much the picture, see Figure 2 and Table 3, the most noticeable results being: a) the confirmation of the good performance of email information for the classification of Emotional Stability and Openness states and of BT for Emotional Stability; b) the increased impact of BT and Email for Agreeableness states.

Feature fusion, as per *Schema III* (see Figure 3 Table 3) improves performances in a number of cases: the Speech+Email combination with Extraversion, Emotional Stability and Openness; the IR+BT combination with Agreeableness; the IR+Email combination, which further increases the accuracy on Emotional Stability up to an interesting 0.7. Even with this schema, Emotional Stability confirms itself as the best predictable state, with an average precision of 0.57.

Finally, when the multimodal social context is considered, as per *Schema IV*, the IR+Email combination confirms its very good predictive power for Emotional Stability and increases its performance over those of *Schema III* with Conscientiousness (0.59) and Extraversion (0.60). We now continue the discussion focusing on the various personality states.

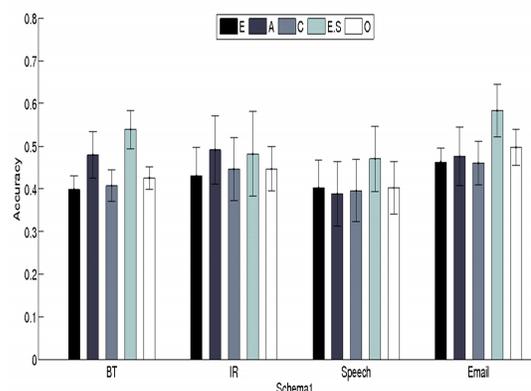


Figure 1: Schema I: Classification accuracies per personality dimension as predicted from each modality.

	<i>E</i>	<i>A</i>	<i>C</i>	<i>E.S</i>	<i>O</i>	<i>Av.</i>
<i>Schema I</i>						
<i>BT</i>	0.40	0.48	0.41	0.54	0.43	(0.45)
<i>IR</i>	0.43	0.49	0.45	0.48	0.45	(0.46)
<i>SP</i>	0.38	0.46	0.43	0.46	0.42	(0.43)
<i>EM</i>	0.46	0.48	0.46	0.58	0.50	(0.50)
<i>Av.</i>	(0.42)	(0.48)	(0.44)	(0.52)	(0.45)	
<i>Schema II</i>						
<i>BT</i>	0.45	0.52	0.40	0.57	0.51	(0.49)
<i>IR</i>	0.46	0.47	0.45	0.51	0.44	(0.47)
<i>SP</i>	0.43	0.34	0.42	0.53	0.43	(0.43)
<i>EM</i>	0.45	0.52	0.40	0.58	0.51	(0.49)
<i>Av.</i>	(0.45)	(0.46)	(0.42)	(0.55)	(0.47)	
<i>Schema III</i>						
<i>SP – BT</i>	0.41	0.50	0.41	0.49	0.42	(0.45)
<i>SP – IR</i>	0.47	0.44	0.42	0.52	0.47	(0.46)
<i>SP – EM</i>	0.50	0.48	0.47	0.59	0.56	(0.52)
<i>IR – BT</i>	0.41	0.55	0.48	0.57	0.43	(0.49)
<i>IR – EM</i>	0.44	0.13	0.29	0.70	0.54	(0.42)
<i>BT – EM</i>	0.30	0.44	0.46	0.52	0.49	(0.44)
<i>Av.</i>	(0.42)	(0.42)	(0.42)	(0.57)	(0.49)	
<i>Schema IV</i>						
<i>SP – BT</i>	0.43	0.53	0.43	0.48	0.41	(0.46)
<i>SP – IR</i>	0.49	0.54	0.45	0.60	0.51	(0.52)
<i>SP – EM</i>	0.48	0.51	0.51	0.61	0.57	(0.54)
<i>IR – BT</i>	0.42	0.52	0.47	0.59	0.44	(0.49)
<i>IR – EM</i>	0.60	0.35	0.59	0.71	0.54	(0.56)
<i>BT – EM</i>	0.30	0.43	0.46	0.40	0.48	(0.41)
<i>Av.</i>	(0.45)	(0.48)	(0.49)	(0.57)	(0.49)	

Table 3: Schema I,II,III and IV: Classification accuracies per personality dimension from Bluetooth(BT), Infrared(IR) Speech (SP) and Emails(EM)

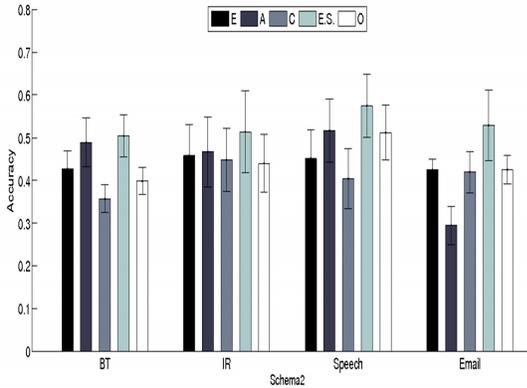


Figure 2: Schema II: Classification accuracies per personality dimension as predicted from separate modalities including the social context.

Extraversion. The highest performance (0.6 - an improvement of 0.27 over the baseline) is obtained by combining IR and Email information from both the target subject and the people he/she F2F interacted with (*Schema IV*). All the other feature assemblages never exceed 0.5 accuracy. This is of some interest for at least two reasons: in the first place, one could have expected a greater role of, e.g., Speech features, given the relevance that both the psycho-

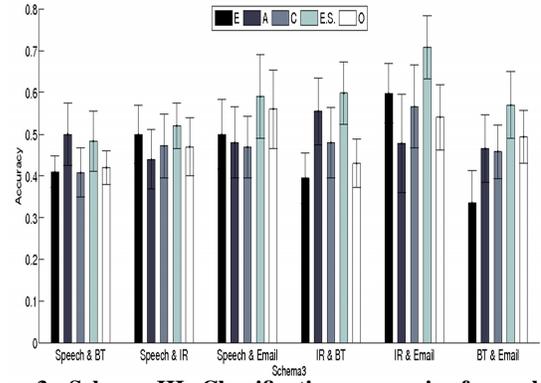


Figure 3: Schema III: Classification accuracies for each personality dimension as predicted from fusion of different modalities.

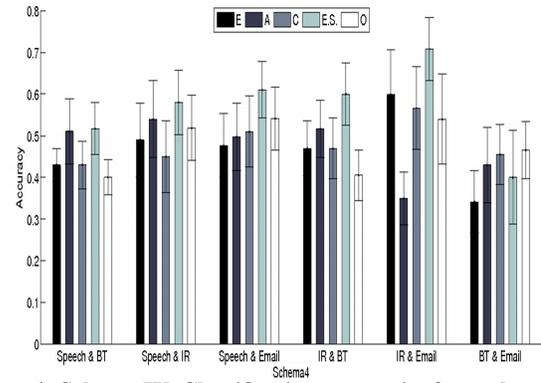


Figure 4: Schema IV: Classification accuracies for each personality dimension as predicted from fusion of different modalities considering the social context.

social and the computational literature have assigned them for the Extraversion *trait*; in our case, neither in isolation nor under multimodal fusion or with contextual features, Speech ever attains accuracy values higher than 0.5 for the Extraversion *state*. If confirmed by further studies, this datum could show that the behavioral fabric of states can be, at least partially, different from that of traits. Secondly, this result emphasizes the role that the communicative behavior (amount and quality of F2F interaction, amount and quality of electronically mediated communication, typology of interaction targets, etc.) of the other people has for the prediction of Extaversion states. (*Parameters used C=1, gamma=0.01 and C=1, gamma=0.03 respectively*)

Agreeableness. Our results are less conclusive concerning Agreeableness. Apart from a marked drop in accuracy for condition *IR - Email* of *Schema III*, all the other conditions yield similar results, comprised in a range between 0.43 (condition *BT - Email* of *Schema IV*) and 0.55 (condition *IR - BT* of *Schema III*). (*Parameters used C=0.25, gamma=0.06 and C=0.25, gamma=0.03 respectively*)

Conscientiousness. For Conscientiousness the highest accuracy score (0.59, an improvement of 0.17 compared to the baseline) was obtained by combining information concerning F2F interaction (in particular, number of the friends with whom the subject interacted and average level of friendship) and Email communication both for the target subject and his/her interaction parties. Noticeably, the same condition without the social context yields the lowest accu-

racy, this way strongly pointing towards an important role of the social context for this state. (*Parameters used $C=0.25$, $\gamma=1$ and $C=0.25$, $\gamma=0.01$ respectively*)

Emotional Stability is by-and-large the state that seems easiest to predict in our setting: it has the highest average accuracies across the various schemas and for each schema it is the state that is best recognized, with a peak of 0.71 for the *IR - Email* condition of *Schema IV*. Email features provide the best results, both when used alone or in combination with information concerning F2F interaction (IR). Noticeably, the addition of the corresponding information about the interacting parties, as per *Schema IV*, does not further improve performances, suggesting that this state might be less sensible to the social context than other states. The findings in the literature showing that the Emotional Stability *trait* is second only to Extraversion in being reliably estimable from speech [18], and the existence of a significant correlation between phoneticians' ratings of vocal intensity and self and peer ratings of emotional stability [27], do only partially carry over to the case of the Emotional Stability *state*. According to our results, in fact, though often yielding accuracy values higher than 0.5, Speech features (and combinations thereof) are largely outperformed by information concerning electronic communication and F2F interaction. (*Parameters used $C=0.25$, $\gamma=0.01$*)

Openness. Email features are quite promising for Openness, too. Though far from yielding the high performances observed with the Emotional Stability state, Email provides 0.5 accuracy when used alone and 0.56 when used in conjunction with Speech. As with Emotional Stability, the addition of information about the social context does not improve the performances, possibly pointing, again, to a lower role of the context with this state. (*Parameters used $C=1$, $\gamma=0.03$*)

7. CONCLUSIONS, IMPLICATIONS AND LIMITATIONS

The major goals of this paper were to propose a novel perspective on personality computation and move the first steps towards approaching it. With respect to more "traditional" approaches, we have proposed an important shift of focus whereby actual behaviors are not simply ancillary to personality classification but become the actual target of the investigation themselves. The key idea is that personality *traits* can be seen as density distributions over personality *states*, the latter being actual behavioral sequences describable by the same language as traits. The other qualifying element are situational characteristics that, together with traits, determine actual behaviors/states. The importance and novelty of this conception of personality is that: a) it turns the traditional approach to computing personality upside down; rather than starting from actual behaviors to (directly) answer the question "how Extravert/Emotionally Stable/etc. *X is?*" try to understand "how much extravertedly/in an emotionally stable way/ etc. *X behaved?*" to later use it to reconstruct the traits. b) It allows for greater flexibility, thanks to the acknowledgement of the crucial role of situational factors. c) It provides for natural, and bidirectional, links between traits (stable dispositions towards behaviors) and states (transient properties of actual behaviors), this way opening interesting prospect for the task that, ultimately, motivates much of the computational research in personality: using it to predict/explain behavior. This could make it easier for automatic systems to predict outcomes related to personality such as job satisfaction at particular days, leadership, team-effectiveness(see for a review, [14]), actual spending behavior, etc.

The first, exploratory steps we have taken have in this paper consisted in: a) extracting and filtering relevant cues from signals obtained through wearable sensors and from electronic communication activity; b) use them to model actual behavioral sequences; c) exploit the results from experience sampling as a ground truth; d) classify the behavioral sequences into the appropriate levels of the various states by comparing unimodal, multimodal and two derived conditions accounting for the role of the social context.

In terms of obtained accuracies, the results are quite promising: compared to a baseline of 0.33, we have obtained the following highest accuracies figures: 0.6 for Extraversion; 0.59 for Conscientiousness; 0.55 for Agreeableness; an accuracy of 0.7 for Emotional Stability; 0.57 for Openness. In a number of cases (Extraversion and Conscientiousness) evidence was found for a role of the social context, while in others (Agreeableness, Emotional Stability and Openness) such evidence is still lacking. Other valuable results concern indications about the effectiveness of the feature extracted, all of them built from signals or information provided by nowadays widely available means (BT, wearable microphones, e-mails, infrared sensors).

From a more general point of view, the results of our experiments show the feasibility of the proposed perspective and hopefully encourage further research. Specific topics that we think need to be urgently addressed are: a) the development of much more refined and empirically grounded notion of situation that focuses on psychological efficacy, i.e., on the capability of causing a change in personality state; b) its usage towards the development of more flexible, robust and principled computational models of personality states than the one adopted here.

We conclude by acknowledging some of the many limitations of the present study, starting from the not huge sample size (54 subjects) and the (theoretically) unsophisticated computational models adopted (SVM). As already said, our work in this paper was meant as the initial step of a long term program addressing personality by trying to reconcile three aspects that have been often kept separate, at least in the computational literature: the stability of *traits*, the apparent ever-changing nature of actual behavior/*states* and the elusive role of *situations*.

8. ACKNOWLEDGEMENTS

Bruno Lepri's research was funded by the Marie Curie-COFUND-7th Framework PERSI project. The authors thanks the Sociometric Solutions and Alex "Sandy" Pentland for providing the sociometric badges.

9. REFERENCES

- [1] K. Andrews. It's in Your Nature: A Pluralistic Folk Psychology. *Synthese*, 165(1), 2008.
- [2] L. Batrinca, B. Lepri, and F. Pianesi. Multimodal recognition of personality during short self-presentations. In *Proceedings of the 2011 joint ACM workshop on Human gesture and behavior understanding*, pages 27–28. ACM, 2011.
- [3] L. M. Batrinca, N. Mana, B. Lepri, F. Pianesi, and N. Sebe. Please, tell me about yourself: automatic personality assessment using short self-presentations. In *ICMI*, pages 255–262, 2011.
- [4] C.-C. Chang and C.-J. Lin. LIBSVM: A library for support vector machines. *ACM Transactions on Intelligent Systems and Technology*, 2:27:1–27:27, 2011. Software available at <http://www.csie.ntu.edu.tw/~cjlin/libsvm>.

- [5] G. Chittaranjan, J. Blom, and D. Gatica-Perez. Mining large-scale smartphone data for personality studies. *Personal and Ubiquitous Computing*, 17(3):433–450, 2013.
- [6] P. T. Costa and R. R. McCrae. Four ways five factors are basic. *Personality and Individual Differences*, 13(6):653–665, 1992.
- [7] Y.-A. de Montjoye, J. Quoidbach, F. Robic, and A. Pentland. Predicting personality using novel mobile phone-based metrics. In *SBP*, pages 48–55, 2013.
- [8] W. Fleeson. Toward a Structure- and Process-Integrated View of Personality: Traits as Density Distributions of States. *Journal of Personality and Social Psychology*, 80(6):1011–1027, 2001.
- [9] W. Fleeson. Situation-based contingencies underlying trait-content manifestation in behavior. Technical Report 4, Department of Psychology, Wake Forest University, Winston-Salem, North Carolina 27109, USA. FleesonW@wfu.edu, 2007.
- [10] A. Furnham, B. Gunter, and A. Green. Remembering science: The recall of factual information as a function of the presentation mode. *Applied Cognitive Psychology*, 4(3):203–212, 1990.
- [11] D. Goren-Bar, I. Graziola, F. Pianesi, and M. Zancanaro. The influence of personality factors on visitor attitudes towards adaptivity dimensions for mobile museum guides. *User Modeling and User-Adapted Interaction*, 16(1):31–62, 2006.
- [12] S. Gosling. A very brief measure of the Big-Five personality domains. *Journal of Research in Personality*, 37(6):504–528, 2003.
- [13] A. Jain and D. Zongker. Feature Selection: Evaluation, Application, and Small Sample Performance. *IEEE Trans. Pattern Anal. Mach. Intell.*, 19(2):153–158, Feb. 1997.
- [14] T. A. Judge, J. E. Bono, R. Ilies, and M. W. Gerhardt. Personality and leadership: a qualitative and quantitative review. *Journal of Applied Psychology*, 87(4):765–780, 2002.
- [15] K. Kalimeri, B. Lepri, and F. Pianesi. Causal-modelling of personality traits: extraversion and locus of control. In *Proceedings of the 2nd international workshop on Social signal processing*, pages 41–46. ACM, 2010.
- [16] B. Lepri, S. Ramanathan, K. Kalimeri, J. Staiano, F. Pianesi, and N. Sebe. Connecting meeting behavior with extraversion - a systematic study. *T. Affective Computing*, 3(4):443–455, 2012.
- [17] B. Lepri, J. Staiano, G. Rigato, K. Kalimeri, A. Finnerty, F. Pianesi, N. Sebe, and A. Pentland. The sociometric badges corpus: A multilevel behavioral dataset for social behavior in complex organizations. In *SocialCom/PASSAT*, pages 623–628, 2012.
- [18] F. Mairesse, M. A. Walker, M. R. Mehl, and R. K. Moore. Using Linguistic Cues for the Automatic Recognition of Personality in Conversation and Text. *Journal of Artificial Intelligence Research*, Vol. 30:457–501, 2007.
- [19] M. R. Mehl, S. D. Gosling, and J. W. Pennebaker. Personality in its natural habitat: manifestations and implicit folk theories of personality in daily life. Technical Report 5, Department of Psychology, University of Arizona, Tucson, AZ 85721, USA. mehl@email.arizona.edu, 2006.
- [20] G. Mohammadi and A. Vinciarelli. Automatic Personality Perception: Prediction of Trait Attribution Based on Prosodic Features. *IEEE TRANSACTIONS ON AFFECTIVE COMPUTING*, Volume 3(no. 3):pp. 273–284, 2012.
- [21] D. O. Olguín, P. A. Gloor, and A. Pentland. Capturing individual and group behavior with wearable sensors. In *AAAI Spring Symposium: Human Behavior Modeling*, pages 68–74, 2009.
- [22] D. Olguin-Olguin, B. N. Waber, T. Kim, A. Mohan, K. Ara, and A. Pentland. Sensible organizations: technology and methodology for automatically measuring organizational behavior. *Trans. Sys. Man Cyber. Part B*, 39(1):43–55, Feb. 2009.
- [23] F. Pianesi, N. Mana, A. Cappelletti, B. Lepri, and M. Zancanaro. Multimodal recognition of personality traits in social interactions. In *ICMI*, pages 53–60, 2008.
- [24] P. Pudil, J. Novovičová, and J. Kittler. Floating search methods in feature selection. *Pattern Recognition Letters*, 15(11):1119–1125, 1994.
- [25] J. Ramirez, J. C. Segura, C. Benitez, A. D. L. Torre, and A. Rubio. Voice activity detection with noise reduction and long-term spectral divergence estimation, 2004.
- [26] B. Reeves and C. Nass. *The media equation: how people treat computers, television, and new media like real people and places*. Cambridge University Press, New York, NY, USA, 1996.
- [27] K. R. Scherer and H. Giles. Linguistics: Social markers in speech., eds. *American Anthropologist*, 83(2):437–438, 1981.
- [28] J. Staiano, B. Lepri, K. Kalimeri, N. Sebe, and F. Pianesi. Contextual modeling of personality states’ dynamics in face-to-face interactions. In *SocialCom/PASSAT*, pages 896–899, 2011.
- [29] W. Walsh, K. Craik, and R. Price. *Person-Environment Psychology: New Directions and Perspectives*. Lawrence Erlbaum Associates, Incorporated, 2000.
- [30] X. Zhou and C. Conati. Inferring user goals from personality and behavior in a causal model of user affect. In *Proceedings of the 8th international conference on Intelligent user interfaces*, IUI ’03, pages 211–218, New York, NY, USA, 2003. ACM.