A

Stress Assessment Using Smartphones

Alban Maxhuni, Venet Osmani, Oscar Mayora, CREATE-NET
Pablo Hernandez-Leal, L. Enrique Sucar, Eduardo F. Morales, Instituto Nacional de Astrofísica, Óptica y Electrónica

Nowadays, economic competition is constantly increasing and employees perceive stress of different degrees, affecting their health, decreasing their job-control in organisational performance and reducing their quality of life in general. Stress assessment is a complex issue and numerous studies have examined factors that influence stress in working environments. Most of current approaches for detecting stress use facial and voice recognition algorithms, while others have used physiological sensors. However, research studies have shown that monitoring individuals' behaviour parameters during daily life can help to assess stress levels. In this study, we examine the effects of work-related stress events and other personality traits (e.g. behaviour and routine changes) on working environments, using features derived from smartphones. In particular, we use information from physical activity levels, location, social-interactions, social-activity (e.g. phone calls and SMS) and application usage during working days. Participants in this study were 30 employees chosen from two different private companies, monitored over a period of 8 weeks in real-work environments. Our first contribution is to apply correlation analysis, hierarchical clustering and multi-linear regression analysis to find patterns, behaviours and features associated with stress. The findings suggest that information from phone usage shows important correlation with employees perceived stress level. Secondly, we used machine learning methods to classify perceived stress levels based on the analysis of the information provided by the smartphones, as indicated above. We used decision trees obtaining 67.57% accuracy and 71.73% after applying a semi-supervised method. Our results show that stress levels can be monitored in unobtrusive manner, through analysis of smartphone data.

Additional Key Words and Phrases: Stress, smartphone-computing, behaviour monitoring, behaviour patterns, semi-supervised learning, smartphone use at work

1. INTRODUCTION

Social competition is becoming increasingly stronger, which together with the rapid economic transformation have changed the dynamics of workplace environments. Due to these changes, enterprise employees are experiencing a period of intense job-insecurity, increased work-loads, and long working hours. All these factors engender work-related stress of different degrees, affecting the physiological and psychological functioning of the employees. According to recent reports from the *European Agency for Safety and Health at Work* (EU-OSHA), stress was found to be the second most common work-related health problem across 27 Member states of the European Union (EU). Overall, 22% of EU employees reported work-related stress. This was accompanied by physical, psychological or social complaints, and was shown to be associated with the inability to bridge a gap with the requirements placed on them [Schneider et al. 2010]. These patterns are similar in the U.S. where the National Institute for Occupational Safety and Health (NIOSH), have reported that over 40%

Author's addresses: Alban Maxhuni, Venet Osmani, Oscar Mayora, CREATE-NET, Via alla Cascata 56/D Povo, Trento, Italy; Pablo Hernandez-Leal (current address) Centrum Wiskunde & Informatica, Science Park 123, Amsterdam, The Netherlands; L. Enrique Sucar, Eduardo F. Morales, Instituto Nacional de Astrofísica, Óptica y Electrónica, Luis Enrique Erro # 1, Sta. María Tonantzintla, Puebla, México.

DOI: http://dx.doi.org/10.1145/0000000.0000000

of employees suffer from workplace activities being "very or extremely" stressful [NOISH 1999].

Long-term exposure to stress can lead to serious health problems, causing physical illness through its physiological effects (e.g., fatigue, decreased sleep quality), behaviour changes (e.g., addiction, attention deficit), and social isolation issues (e.g., anger) [Bongers et al. 1993; Glanz et al. 2008; Korabik et al. 1993; Maslach et al. 2001; Paoli and Parent-Thirion 2003; Sultan-Taïeb et al. 2013]. As a consequence, these negative effects have been shown to decrease well-being at workplace and employees' work effectiveness. Moreover, long-term exposure to stress typically leads to job-burnout, a state that leads to mental and physical exhaustion [Maslach et al. 2001]. Because of the previously mentioned reasons it is important to measure stress as a way of monitoring individual's well-being. However, unlike other mental and physical problems, stress is not easy to measure [NOISH 1999]. Thus, its assessment represents a current open problem.

To date, physiological measurements and self-reported questionnaires are the most common methods used to infer work-related stress. However, limited research has been conducted within real-life workplace environments, as shown through the literature review, where a substantial number of studies recruit college students rather than employees of organisations. Therefore, monitoring the affect changes of employees and other personality traits (e.g., behavioural aspects) should be of great interest for both healthcare institutions and organisations.

A number of studies have investigated detecting stress and emotions based on facial expressions [Valstar et al. 2011]. Other mood and stress detectors have used individual physiological parameters, these include heart rate and the galvanic skin response (GSR) [Bakker et al. 2011; Muaremi et al. 2013]. Lastly, other studies have analysed voice acquired from individuals to detect stress in laboratory or clinical settings [He et al. 2009]. However, their limitation is that laboratory settings are often an inadequate environment compared to the complexity of real-day environment monitoring at diverse scales (i.e., physically and socially). In this regard, another aspect that has to be considered when it comes to long-term monitoring, is that sensors need to be as least intrusive as possible trying to minimize the impact on workers' routines and their natural behaviour.

Smartphones are becoming more powerful (in terms of sensors capabilities) and every year the number of these devices is increasing. For these reasons, smartphones are excellent candidates to be used for monitoring everyday activities including activities in working environments. Thus, the challenge is to use the sensor capabilities of the smartphones to detect stress-related behaviour of a person in an unobtrusive manner. Then, this could be communicated to the person in order to take pre-emptive actions and alleviate high stress levels [Sanches et al. 2010].

In this study, the following research questions were put forth:

— Is there a correlation between the subjects' behavioural characteristics, extracted from smartphone sensor data, and their self-reported stress levels?
— Is it possible to improve prediction accuracy of work-related stress based on smartphone sensor data by combining limited labelled data and unlabelled data?

Several factors can affect employees' stress at work, however our approach focuses on behaviour changes that can be directly measured using smartphones: location changes, physical activities, social interactions and phone application usage. With the processed data from 30 subjects: i) we used hierarchical clustering of perceived stress and identify the diversity and similarity of self-reported stress level within subjects and organisations; ii) we applied multiple linear regression analysis to detect the most important variables for predicting stress; iii) we used decision tree classifiers to predict the stress level of the subjects using only data obtained from the smartphones (achieving a prediction accuracy of 67.5%). iv) Since, in this study, there were a high percentage of missing values in the self-reported ques-

tionnaires (20.03%), semi-supervised learning (SSL) methods from machine learning [Zhu 2005] were used to address these issues and exploit unlabelled samples. We have specifically made use of a method called self-training [Triguero et al. 2015] from the SSL methods for our analysis.

After applying SSL, we obtained an accuracy of 71.73 %; 67.8% for precision and 71.4% for recall in average, for all subjects.

In summary, the main contributions of the paper are:

— We provide a comprehensive qualitative and quantitative analysis of the association between objectively measured data –physical activity, location, social-interaction and social-activity– and demographic information, with subjective self-assessment of work-related stress.
— We evaluate our analysis and methods with a real-world dataset collected for a period of two months from 30 employees in working environments.

The rest of paper is organized as follows: Section 2 summarises previous research with smartphones and concerning the monitoring of stress in working environments. Section 3 provides information about the participants demographics, how the data was collected and a detailed description of the attributes that were selected for predicting the perceived stress level of the subjects. Section 4 presents correlation analysis and results of our study. Section 5 presents results regarding the prediction achieved by our models and discuss the potential benefits of the approach. Finally, conclusions and future research directions are given in Section 6.

2. RELATED WORK

The use of technological devices in health-care has shown an important role to help people manage their emotions. Furthermore, these devices have enhanced the capability to detect stress responses and assess levels of stress [Healey and Picard 2005; Jerritta et al. 2011; Liu 2004]. However, capturing what causes emotional changes that derives from work-related stressors and detect the onset of stress can be quite challenging. Several methods have tried to infer stress based on physiological signals, such as heart-rate variability, blood pressure, body temperatures and respiration [Bakker et al. 2011; Jerritta et al. 2011; Roh et al. 2012]. The work in [Wijsman et al. 2011] goes further by not only inferring stress, but also by providing short-term feedback to the users in everyday activities. However, the use of physiological sensors has reported several limitations: i) sensors require a large size to cover many signal types [Bakker et al. 2011], ii) during measurement period, movement is limited because of skin conductance sensor [Roh et al. 2012], and iii) sensors increase the discomfort [Wijsman et al. 2011] since they need to be carried at all times (in specific places on the body) in order to allow long-term continuous monitoring.

The miniaturization of wearable sensors has made it possible to be included in smartphones. Recently, there has been a growing interest in inferring stress using those sensors, since they are a personal and common accessory among people. Lu et al. [2012] proposed a method for detecting stress based on speech analysis and the variation of speech articulation using smartphones. The authors have reported a predictive accuracy of stress of 81% and 76% for indoor and outdoor environments, respectively, using the vocal production of 14 subjects. However, in real-life activities (crowded environments) this approach may lead to misinterpretation of speech and therefore of emotion.

In order to infer relationship dynamics of people and behaviour changes in daily activities, smartphones have been suggested as a promising candidate to obtain user's context. Research using smartphones for long-term stress monitoring [Maurer et al. 2006; Garcia-Ceja et al. 2016] has proposed to collect many types of contextual data (e.g., physical activities, social activities and locations) that could help in inferring stress from behaviour changes. In this line, Moodscope is a self-tracking system to help users manage their mood

[Likamwa et al. 2013]. The system detects user's mood from smartphones usage data, such as email messages, calls and SMS logs, application usage, web browsing histories and location changes. The authors reported an initial 66% accuracy of subjects' daily mood, improving to 93% after two months of training. It should be noted that the work on MoodScope focuses on two narrow dimensions of mood, namely pleasure and activeness, which may explain the high prediction accuracy. Thus, the work presented in MoodScope is closer to our previous work on predicting depression [Grünerbl et al. 2015; Maxhuni et al. 2016] (where we achieved over 97% precision and recall), rather than the work on stress detection presented in this paper.

[Bauer and Lukowicz 2012] performed stress recognition from 7 students before and after an exam period. The assumption is that students are likely to be under stress during the exam sessions. They acquired data from smartphones (location, social proximity through Bluetooth, phone calls and SMS logs) reporting an average accuracy of 53% during the exam session. Sano and Picard [2013] monitored 18 subjects for a period of 5 days. In addition to smartphone usage features they included a wrist sensor. In order to recognise stress levels, the authors applied correlation analysis and reached a 75% accuracy using machine learning techniques to classify stress moments. Similarly, Muaremi et al. [2013] measured smartphone mobility data (phone-calls, SMS, location and physical activity) and wearable Heart Rate Variability (HRV) sensor data to classify perceived stress. The authors emphasise the importance of recording human voice as a potential source for non-intrusive stress detection. Furthermore, Bogomolov et al. [2014] have used context information from the environment, such as weather condition, social proximity obtained by Bluetooth scanning, call logs, SMS logs, and self-reported surveys about personality traits to predict stress events. Finally, another work that studies stress based on wearable physiological signals such as inter-beat interval, respiratory sinus arrhythmia (RSA), and IE ratio is cStress [Hovsepian et al. 2015]. The approach builds models from measured features into segments (1 minute window size) and compare them with a physiological response to stressors. The models proposed were trained using physiological data collected from a controlled laboratory study (n=21) and tested on an independent dataset in laboratory (n=26) and in real-life settings (n=20). They report 89% recall and 72% accuracy with a dataset of 1,060 stress self-reports provided by 20 independent participants.

Table I summarises related works aiming at detecting the occurrence of stress using smartphone. It can be seen in the table that there are few works (only two) that use only smartphones on realistic conditions with a reasonable number of subjects. The first one [Likamwa et al. 2013] is focussed on mood, not in stress and the second one [Bogomolov et al. 2014] is focussed mainly on prediction and makes no analysis of perceived stress based on different demographic variables, social interaction, activity level, job-related variables, phone-usage, and location nor includes a correlation and regression analysis as we do in this paper.

In contrast to previous works, we aim at enabling measurement of relevant aspects of well-being at work that can be derived from employees' behaviour in an unobtrusive way. While laboratory settings are always subject to criticism of ecological validity, our trial was based on real fieldwork for the period of two months. In addition, we performed a systematic analysis for each sensor (physical activity level, social-interaction, locations and phone usage) and its relationship with stress levels. We provide a detailed overview of subjects behaviour patterns on a daily basis, taking into account their demographics (i.e., age, gender, education and family status). Using this information, we were able to investigate the association between objectively measured data with subjective self-assessment of work-related stress into various data groups, based on demographic information. We assume that (grouping) this information is important for the organisations to identify effective interventions and strategies in order to reduce perceived stress and job-demands for their employees. Finally, we build a classification model to predict stress of each subject. Using

Table I: Main related works in Stress Detection showing the features used and details about the study.

Study	Items measured	Study details
[Kim et al. 2008]	Heart Rate Variability (HRV)	Controlled, constrained; 68 subjects. Accuracy reached from 63.2% to 66.1%
[Lu et al. 2012]	Smartphone (Speech analysis)	In-lab and uncontrolled setting; 14 subjects; accuracy from 76% and 81%
[Bauer and Lukowicz 2012]	Smartphone (Location, bluetooth, phone-call and SMS)	Uncontrolled; 7 subjects, accuracy of 53%
[Likamwa et al. 2013]	Smartphone focussed on smooth (E-mails, Call and SMS logs, application usage, web browsing histories and location changes)	32 subjects; Reached overall accuracy from 66% to 93%
[Sano and Picard 2013]	Wrist sensors and smartphone (Phone usage duration, phone calls and SMS logs)	Uncontrolled; 18 subjects; accuracy of 75%
[Muaremi et al. 2013]	HRV and smartphone (Phone- and SMS logs, Location, Audio Stress Response)	Uncontrolled, 35 subject from three different companies. Accuracy of 53% from smartphones and 61% after fusion
[Bogomolov et al. 2014]	Smartphone (Call logs, SMS, bluetooth and weather)	Unconstrained, unknown ; 117 subjects. Overall accuracy of 72.39%
[Hovsepian et al. 2015]	Inter-beat interval, respiratory sinus arrhythmia (RSA), and IE ratio	Controlled (n=26) and uncontrolled (n=20), 89% recall and 72% accuracy

semi-supervised machine learning methods we increased the accuracy from 67.57% to 71.73% for predicting the stress level.

3. DATASET

Data was collected from a group of 30 subjects in the course of 8 weeks. Considering the fact that the data collection period covered the months of November and December (where the employees have to finalize yearly objectives), we could ensure that the data contained behavioural changes from elevated stress levels. Our data collection framework was based on a server-client architecture built around the Samsung Galaxy S3 mini 32GB smartphone[1]. During the study, subjects used the smartphone in daily basis as their own phone (including working hours). There were no restrictions placed on users regarding the handling of their smartphones, so our analysis is framed under usual/realistic conditions. The application developed to collect data was running continuously as a background application. The application started automatically at 9am at working days (Monday-Friday) without any interaction with the user. In order to understand users' mood and stress levels, the app prompted users to fill in a questionnaire at three different times of the day: at 9am (at the beginning of the work hours), at 2pm (after lunch break) and at 5pm (at the end of the work hours). The questionnaires appeared automatically and the user had the option to answer the questions or snooze the questionnaire for later. The questionnaire consisted of 14 questions that were answered is around one minute. Some examples of screenshots of the questionnaire are shown in Figure 1.

Even when questionnaires appeared automatically no compulsory actions (such as blocking the phone until answering) were taken, therefore users had the possibility to ignore them. This resulted in incomplete information of two types: missing questionnaires in a day (possibly because users decided to ignore them) and missing questionnaires for a complete day (possibly because high work load). From the complete set of 30 users feature extraction was performed for two types of variables.

[1]We did not consider using other devices like smart watches are they are currently more expensive and less available among the population.

Fig. 1: Examples of screen shots of the questionnaire.

Table II: Study demographics of the subjects in our study.

Variable	Characteristics	Nr.	(%)	Variable	Characteristics	Nr.	(%)
Gender	Male	**18**	(60.00%)	**Marital status**	Married	**15**	(50.00%)
	Female	**12**	(40.00%)		Never married	**15**	(50.00%)
Education	High-school graduate	**9**	(30.00%)	**No. of children**	None	**17**	(56.67%)
	Bachelor degree	**11**	(36.67%)		1-2	**10**	(33.33%)
	Graduate degree	**10**	(33.33%)		3-4	**3**	(10.00%)
Age	26-30	**5**	(16.67%)				
	31-40	**18**	(60.00%)				
	>40	**7**	(23.33%)				
	Mean (±SD)	**37.46**	(±7.15)				

— The first group of variables includes information of user's behaviour during work hours, these are called *objective variables*.
— The second group contains subjective information obtained from the questionnaires which reflects the mood, work-demands/control and perceived stress of the user, these are called *subjective variables*.

Extracted data for everyday~~every day~~ was divided into two intervals: from 9am to 2pm, and from 2pm to 5pm, referring to the subjective variables (considered as ground truth) acquired from questionnaires.

Now we present a summary of the demographics of the 30 subjects in the study. Then, we present the variables that correspond to stress and mood (subjective variables). Finally, we present the features extracted from smartphone usage (objective variables).

3.1. Study Demographics

In total, 30 employees from two different organisation in Trento, Italy, were selected for the study. Table II provides the summary of employees' demographics characteristics. We can note that there is a fairly balanced mix of gender, age and education level, marital status and number of children among the subjects.

3.2. Subjective variables

The first type of data includes subjective information related to subjects' stress and mental state. In order to get insights in the working environments and job-demands of employees during working days, we developed a questionnaire in a smartphone application to assess several psychological working variables related to work stress. The questionnaire is clinically validated to capture subjects perceived stress and mood states of the employees at work. The questionnaire was derived from the POMS (Profile of Mood State) scale [McNair et al. 1971]

which has two dimensions related to affect of mood states, including, "Positive Affect (PA)" (e.g., Cheerful, Energetic, Friendly) and "Negative Affect (NA)" (e.g., Tensed, Anxious, Sad, Angry) and the rest measures disengagement from work. The PA, NA and disengagement from work items were presented in mixed order.

In our study, each question had five response alternatives, which assessed five stress-related factors on a scale ranging from 1 (absolutely agree) to 5 (absolutely disagree). The answers were stored on the mobile device and constituted part of the analysis. For the purpose of our analyses score distribution has been segmented into three regions, which in our case correspond to three ordinal classes: (*"low"* or *"poor"*), when $\mathbf{score} < \mathbf{3}$; (*"moderate"* or *"fair"*), when $\mathbf{score} = \mathbf{3}$; and (*"high"* or *"sufficient"*), when $\mathbf{score} > \mathbf{3}$.

The first section of the questionnaire, collected information about occupational health outcomes of the participants: i) job induced stress, ii) job-control, iii) job-demand and iv) energy perceived during working days. The second section contained several widely used scales to measure mood: the existence of tensions and pressures growing out of job requirements, feelings of anxiety, cheerfulness, friendliness, sadness, angriness, and quality of sleep.

In Table III we provide overall response rates of completed questionnaires on work-relevant stress from all participants throughout the entire study using the 3-point scale defined earlier. We obtained 1455 completed questionnaires, which represented a response rate of 79.97%. It is worth mentioning that in this study we included only self-reported questionnaire items obtained at ∼2pm and ∼5pm, since we are interested in exploring the relation of stress, moods, and job-performance with respect to the objective variables measured in the previous working hours. We did not include data of the questionnaires at 9 am because we started to get data from the smartphones exactly at that time and could not relate (almost) any information to this questionnaire. It can be noted that employees perceived increased workload and stress, since almost all of the respondents perceived a *moderate* (35.15%) to *high* (22.18%) stress level throughout the entire monitoring period. Regarding how stress impaired productivity of the employees, almost all of them (29 out of 30) reported that at some point their job tasks and job responsibilities were highly demanding (50.58%) throughout the entire monitoring period (marked with red-color in Table III). This is important since prolonged exposure to certain job-demands has been shown to lead employees to variety of health issues, such as mental and physical disorder [Maslach et al. 2001]. In response to work-related stress, 19 employees felt themselves *High - Tensed* at some point of the study, 18 respondents felt *High - Anxious*, 11 of respondents have reported *High - Angriness* (5.67%), which shows that a large group of subjects showed negative moods. Finally, a relevant physical reaction to stress is a *Poor - Sleep Quality*, which was reported by 24 of the respondents.

3.3. Objective variables

The second type of data which provides objective measures associated with users' behaviour was collected from sensors embedded on the smartphones used in this study. From the analysis presented in Section 2 we concluded that 4 categories were needed to perform a proper assessment of subjects stress: physical activity, location, social interaction and social activity. From these categories we extracted 18 features using 9 sensors, as shown in Table IV. In the following sections we discuss each category in detail.

3.3.1. Physical Activity Level - (pACL). The potential role of physical activity (and its relation with sedentary behaviour) in the development of psychological complaints has received increased attention during the last decades [Bernaards et al. 2006; Fleshner 2005; Penedo and Dahn 2005]. On the one hand, psychological stress has been reported as a factor in reducing frequency, intensity, and duration of physical activity [Lutz et al. 2010] by inducing specific physical responses such as tiredness, weakness, and fatigue [Spielberger et al. 2003].

Table III: Subjective variables: overall percentage Self-Reported Questionnaires (exhaustion and disengagement from work) by Perceived Level (High, Moderate, Low) and Number of Subjects.

Variable	Level	Nr.Resp.(%)	Nr.Subj.	Variable	Level	Nr.Resp.(%)	Nr.Subj.
Perceived Stress	*High*	325 (22.18%)	27	**Perceived Job-control**	*High*	612 **(41.77%)**	30
	Moderate	515 (35.15%)	30		*Moderate*	604 (41.23%)	30
	Low	625 **(42.66%)**	30		*Low*	249 (17.00%)	27
Perceived Job demand	*High*	741 **(50.58%)**	29	**Perceived Energy**	*High*	357 (24.37%)	28
	Moderate	357 (24.37%)	30		*Moderate*	756 **(51.60%)**	30
	Low	367 (25.05%)	24		*Low*	352 (24.03%)	28
Tensed	*High*	118 (8.06%)	19	**Anxious**	*High*	128 (8.74%)	18
	Moderate	280 (19.11%)	28		*Moderate*	279 (19.04%)	3
	Low	1067**(72.83%)**	30		*Low*	1058**(72.22%)**	30
Cheerful	*High*	274 (18.70%)	28	**Friendly**	*High*	463 (31.60%)	27
	Moderate	756 (51.60%)	30		*Moderate*	692 **(47.23%)**	30
	Low	435 **(29.70%)**	30		*Low*	310 (21.16%)	29
Angry	*High*	83 (5.67%)	11	**Sad**	*High*	28 (1.91%)	10
	Moderate	186 (12.70%)	5		*Moderate*	112 (7.65%)	30
	Low	1196 **(81.63%)**	30		*Low*	1325 **(90.44%)**	12
Sleep quality	*Sufficient*	886 **(60.48%)**	30				
	Fair	313 (21.37%)	28				
	Poor	266 (18.15%)	24				

***red-color** - represent risk-factor;

Table IV: Objective variables divided in four categories. Sensors and features extracted from smartphone usage on every subject in the study.

Category	Sensors	Features
1. **Physical Activity Level**	*Accelerometer*	1) 3-axis Magnitude
		2) Variance Sum [FUNF 2015]
2. **Location**	*Cellular*	3) CellID and LACID (Number of clusters - DBSCAN) [Birant and Kut 2007]
	WiFi	4) Access Points (Number of clusters - DBSCAN) [Birant and Kut 2007])
	Google-Maps	5) Latitude and Longitude (Number of clusters - DBSCAN)[Birant and Kut 2007], Haversine [Robusto 1957])
3. **Social Interaction**	*Microphone*	6) Proximity based on verbal interaction (Pitch [Hedelin and Huber 1990], Mel-MBSES [Harris 1978])
	Phone Calls	7) Number of Incoming Calls
		8) Number of Outgoing Calls
		9) Number of missed Calls
		10) Duration of Incoming Calls
		11) Duration of Outgoing Calls
		12) Most common Contact-Calls
	SMS	13) Number of Incoming SMS's
		14) Number of Outgoing SMS's
		15) Length of Incoming SMS's
		16) Length of Outgoing SMS's
		17) Most common Contact-SMS
4. **Social Activity**	*App usage*	18) Number of used applications (Social and system)
		19) Duration of used applications (Social and system)

On the other hand, research studies have acknowledged physical activity as a psychological de-stressor since an active lifestyle is associated with health benefits [Proper et al. 2003; Fleshner 2005]. Most related research has used mainly self-reported questionnaires to address the association between physical activity and psychological well-being. In contrast, we wanted to investigate the association between objectively measured physical activity and perceived psychological stress.

We assume that most forms of physical activity (such as mini-breaks and lunch breaks) could reduce the level of stress and increase the positive mood of the subjects. To analyse physical activity, we measure it using accelerometer signals from the smartphones. For this study, we captured 3-axial linear acceleration continuously at a rate of 5Hz, which was sufficient to infer physical activity levels of subjects. Similar to the work in [FUNF 2015], we measured the variance sum (varSum) of 26 seconds (non-overlapping fixed length windows

of *n=128* samples) accelerometer readings, providing the activity levels of *high*, *low*, and *none* using the magnitude of the signal.

We define three ranges of *percentage of physical activity level* (*pACL*) as follows: *high*-(h) when varSum≥7, *low*-(l) when 3≤varSum≤7, and *none*-(n) when varSum<3; using Equation 1:

$$\mathbf{pACL}_{(h,l,n)} = \frac{\textit{Number of High Activities (h)}}{\textit{Total Classified Activities (h,l,n)}} \times 100\% \quad (1)$$

3.3.2. Location. Additional sources of stress can produce behaviours such as frequent smoking, caffeine consumption and skipping lunch [Conway et al. 1981], which are known to contribute to health issues. For this reason, we analyse locations of subjects with the focus in understanding frequent locations changes during working hours. For example, we assume that during the days with high job-demands and high-stress, subjects tend to reduce changing locations or skip lunches due to their responsibilities or deadlines for delivering their work.

In order to measure location changes, we retrieved 3 important sources: (i) the list of WiFi networks available with their respective BSSID address, (ii) cell tower locations (CID, LAC-ID) and (iii) Google Maps locations information (latitude, longitude).[2] Using the location information we clustered locations from each source using the DBSCAN algorithm [Birant and Kut 2007], which is an algorithm mainly used for clustering spatio-temporal locations. For Google location information, we clustered locations with maximal diameter of 300 meters (using latitudinal and longitudinal coordinates and the Haversine distance equation [Robusto 1957]) where the subjects stay for more than 15 minutes and measured the amount of locations in each day. For cell tower information and WiFi networks we clustered location information on an hourly basis. Our objective was to test whether subjects show changes of location in each interval (9am-2pm and 2pm-5pm). For this we compared locations every hour counting +1 when different clusters appear with respect to the previous hour.

The first two categories described are related to activity and location, ignoring the social components, the next two feature categories consider that aspect.

3.3.3. Social Interaction (SI). Perceiving stress in everyday activities evokes a number of emotional responses that may affect interpersonal relations and social ties. Several works have reported that continuous stress may reduce social well-being in the long-term [Cohen and Wills 1985]. As a result, lowered social functioning may predict decreased mental and physical health [Singh-Manoux et al. 2005]. For example, social withdrawal has been used as one of the diagnostic criteria for post-traumatic stress disorder. On the contrary, being socially active has been found to reduce stress by providing a sense of security, enhancing self-confidence, and buffering the impacts of a stressful situations on individuals [Cohen and Wills 1985].

In the last decades, monitoring social interaction has attracted significant attention [Vinciarelli et al. 2009]. Previous studies monitored speech articulation aiming at inferring stress using smartphones. However, these works have been performed on controlled experimental (laboratory) studies [Lu et al. 2012]. In contrast, in this study we investigate the effects of stress on social behaviour derived from continuously recorded and classified human voice (from smartphone's microphone) in real working environments. Moreover, since social interaction includes not only face to face conversations but also phone conversations and messages, another important social aspect that we have taken into account are the employees phone conversations and SMS logs. For this, we investigated the number of conversations (incoming, outgoing and missing), SMS messages (incoming and outgoing), and the com-

[2]We have intentionally not used the GPS sensor, in order to preserve the battery life of the smartphones.

mon *called* and *calling* contacts, compared with the perceived stress on a daily basis.[3] In detail, we measured two aspects of social interaction:

— **Speaker Recognition:** Recent work in stress detection suggest to use Bluetooth embedded sensor on smartphones for measuring social-proximity [Bogomolov et al. 2014]. However, this method poses several disadvantages since the users may not carry the phones all the time. Second, Bluetooth scans have time limits, which restricts the estimation of social-interaction. In contrast, in this study we use the microphone embedded on the smartphones for better and accurate recognition of verbal interactions (namely social-interactions).[4] We have extracted two main audio features: Pitch [Hedelin and Huber 1990] and Mel-MultiBand Spectral Entropy Signature (Mel-MBSES) [Harris 1978] to obtain a higher accuracy in speech activity recognition. A previous study [Ferdous et al. 2015] has described two conditions required for processing audio on smartphones: i) measuring pitch within the range of human voice (40 Hz to 600 Hz), and ii) recognising human voice from the captured frames using the MEL-MBSES coefficients and Support Vector Machine (SVM) classifiers [Vapnik et al. 1997]. We built a SVM using MEL-MBSES coefficients trained on frames coming from 3 minutes of voiced data and 3 minutes of background data. The training set for the SVM consisted of positive vectors (speech) and the negative vectors (non-speech or background). We sampled audio frequency of 8000Hz and set a frame every 256 samples where we calculated Pitch and Mel-MBSES features for each frame, then each frame is labelled either as human voice or not human voice. Approximately every 0.7 second (7 out of 30 frames) must be detected as voice in order to indicate voice activity in that audio segment. We measured percentage of social-interaction based on the total duration (hourly, daily, weekdays) of conversations as shown in Equation 2:

$$\boldsymbol{Social\text{-}Interaction} = \sum_{i=1}^{n} \frac{true - classified}{total - classified} \times 100\% \tag{2}$$

It should be noted that since there were no restrictions on the use of the smartphones, in some cases these were placed inside pockets. The smartphone can still be used to recognise voice in these cases, although the information is less reliable and only works at reduced distances. This may result on underestimating our results for social interaction.

— **Phone-Call and SMS behaviour:** Since calling and texting messages (SMSs) behaviour could be an important source to infer stress-relevant factors we considered phone calls in terms of: *number*, *duration* and *most frequent number* (on a daily basis) of incoming, outgoing and missed. Furthermore, for SMSs, we measured the *number* and *length* (incoming and outgoing). These features may serve as a source of stress, for example understanding phone-call behaviour from subjects that contact different persons more frequent during stress-less periods in comparison with stress-full times. In order to find the most common called/calling ID in each interval (9am-2pm and 2pm-5pm) we used $\text{argmax}_{(Call)} = \sum_{i=1}^{n} countmax(CallID)$ and $\text{argmax}_{(SMS)} = \sum_{i=1}^{n} countmax(SMSID)$ for most frequent Call and SMSs respectively. In order to remove ties among IDs that have the same number of calls, we proposed a scoring model *Score* for both calls and SMSs:

$$\mathbf{Score}_{(Call)} = \frac{duration(CallID)}{countmax(CallID)} \text{ and } \mathbf{Score}_{(SMS)} = \frac{length(SMSID)}{countmax(SMSID)}$$

[3]In order to protect users privacy, all phone call events where anonymised where we register only the five last numbers of each calling or called contact.

[4]The application did not store conversations. In situations like incoming and outgoing phone calls, accepting or dialling a phone call; we stop the recognition service on the phone and restart service after the phone-calls are ended.

3.3.4. Social-Activity. Finally, another aspect that may have impact on the stress levels is application usage. To capture this information, each time an employee used an app, our software stored the event together with the duration and time-stamp. With this information we were able to extract the following data: number of application used per interval and duration of their usage. Applications were divided in two categories:

— System apps: pre-installed apps like Camera or Calendar, Web-browsing, E-Mail client.
— Social apps: such as Viber, WhatsApp, Facebook, Skype and other user downloaded apps (e.g., games and other entertainment apps).

We have described the extracted features from smartphones used to predict mood and in particular perceived stress level. Next, we analyse the retrieved data from the 30 subjects in the study.

4. ANALYSIS OF INFORMATION

Using the features presented in Section 3 we retrieved the data from all the participants in the study. First, data was filtered discarding information from weekends and hours not in the range 9:00am-5:00pm (representing the working hours). Recall that this range is closely related with the ground-truth information acquired from self-assessments (Section 3.2). After the data was filtered, different techniques were used to perform a thorough analysis: (i) we started using hierarchical clustering (Section 4.1), (ii) then correlation analysis (Section 4.2), and (iii) finally, we performed variable importance analysis (Section 4.3).

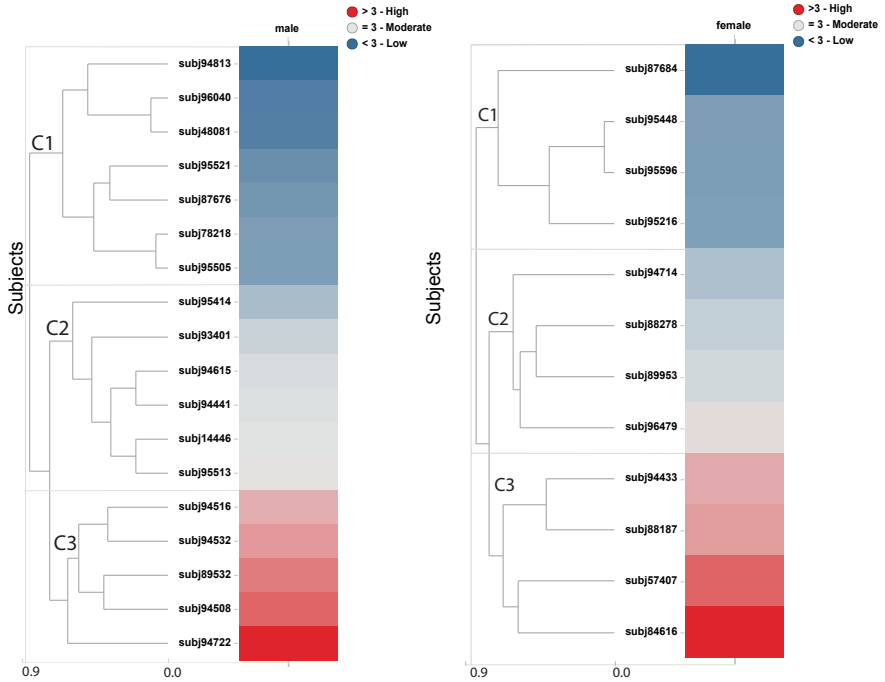
4.1. Diversity and Similarity of Stress Level within Subjects

Hierarchical clustering was used to analyse the participants self-reported stress on a daily basis. We used Ward's method [Ward Jr 1963] to perform the hierarchical clustering of self-reported stress using the half-square euclidean distance[5] between subjects. Figure 2 presents dendrograms about the perceived stress level divided by gender and organisation. Each dendrograms is ordered by clusters, and inside each cluster they are ordered by mean values of perceived stress level.

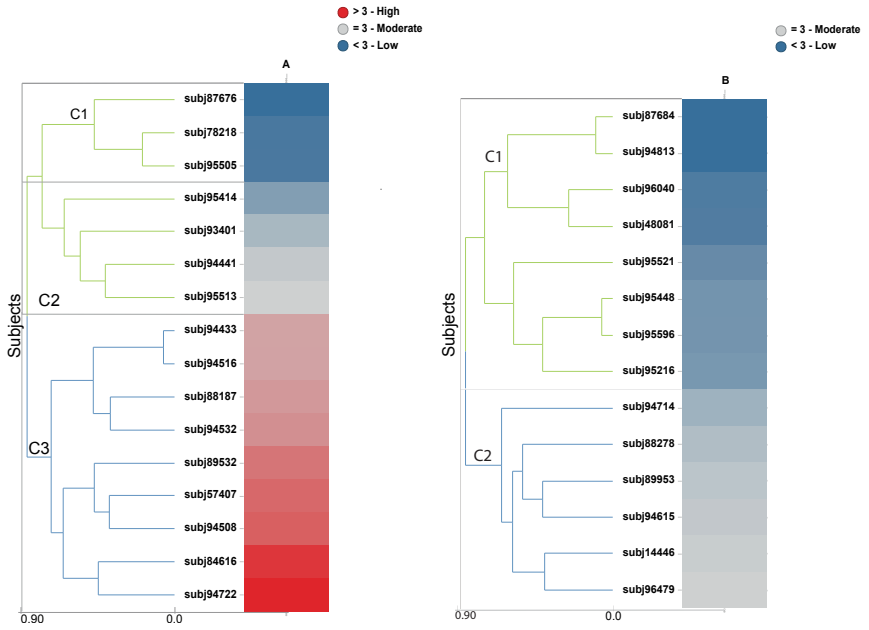
From these figures we note that gender do not easily determine the stress level since both of them show a great variation of perceived stress, however, as we will see, at least in these experiments, there is a higher percentage of women in the high stress group. In contrast, when clustering by organisation we can see that subjects in organisation A showed in average a higher stress than those in organisation B. It is interesting to note that organisation A is an IT organisation, while B is a social support organisation. In Table V, we provide an overview of clustering results based on gender. Cluster analysis yielded 3 distinct clusters (C1, C2 and C3) which represent *low*, *moderate*, and *high* stress levels. Note that women show a uniform distribution across stress levels and men showed slightly more subjects with low stress. We also performed clustering within the organisations, which is shown in the Table VI. The results show that stress was different between organisations. For example, in organisation A, all women (4) showed high stress levels. In contrast, in organisation B, half of the women showed low stress and half of the women showed moderate stress levels. Again, in this company, there are slightly more men with low level of stress.

Finally, we clustered self-reported stress changes within intervals (9am-2pm and 2pm-5pm) as shown in Table VII. For example, low $\longleftrightarrow$ moderate, means that subjects in the clusters showed low stress levels in the first interval and then changed to moderate in the second interval or that moderate changed to low. In this case, 23.33% of the subjects showed at least a high level of stress in their daily activities (*high*$\longleftrightarrow$*moderate* or *high*$\longleftrightarrow$*high*)

[5]Euclidean distance is always greater than or equal to zero. Measurements would be ≈ 0 for identical subjects and ≈ 1 for subjects that show less similarity.



(a)



(b)

Fig. 2: Dendrograms obtained by computing similarities and diversity between perceived stress level of each subject (a) by Gender and (b) by organisation. Three major clusters can be noted, color boxes correspond to average stress for different subjects.

Table V: Perceived Stress Level from dendrogram analysis by gender. Three major clusters can be noted based on perceived level of stress.

Cluster (Stress-Level)	Men (Nr./%)	Women (Nr./%)
C1 (*low* < 3)	7/18 (38.89%)	4/12 (33.33%)
C2 (*moderate* = 3)	6/18 (33.33%)	4/12 (33.33%)
C3 (*high* > 3)	5/18 (27.78%)	4/12 (33.33%)

Table VI: Perceived Stress Level from dendrogram analysis by gender within organisations. Three major clusters can be noted based one perceived stress.

Cluster (Stress-Level)		Org. A	Org. B
C1 (*low*<3)	**Men:**	3/12 (25.00%)	4/6 (66.67%)
	Women:	0/4 (0.00%)	4/8 (50.00%)
C2 (*moderate*=3)	**Men:**	4/12 (33.33%)	2/6 (33.33%)
	Women:	0/4 (0.00%)	4/8 (50.00%)
C3 (*high*>3)	**Men:**	5/12 (41.67%)	0/6 (0.00%)
	Women:	4/4 (100.00%)	0/8 (0.00%)

Table VII: Perceived Stress Level from dendrogram analysis by response intervals ([9am-2pm], [2pm-5pm]). Three major clusters can be noted based on perceived level of stress and transition of perceives stress into intervals ([9am-2pm] and [2pm-5pm]).

Cluster (Stress-Level)	Intervals
C1 *low*⟷*low*; *low*⟷*moderate*	11/30 (36.67%)
C2 *moderate*⟷*low*; *moderate*⟷*moderate*	12/30 (40.00%)
C3 *high*⟷*moderate*; *high*⟷*high*	7/30 (23.33%)

Table VIII: Overall average percentage of Physical Activity Level (pACL) by intervals (9am-2pm and 2pm-5pm) and perceived Stress Level (SL) [High, Moderate, Low].

Distribution of pACL by (Gender, Age, Education, Marital Status and organisation)	**pACL** [9-2]	**pACL** [2-5]	**High (SL)**	**Moderate (SL)**	**Low (SL)**
– **Male**	**18.03**	**21.34**	16.29 (*)	16.68	**23.60**
– **Women**	15.66	18.74	10.57 (**)	15.37	**18.89**
– **26-30** (28.6±1.95)	12.89	15.48	12.45	13.65	**17.83**
– **31-40** (35.33±2.4)	17.50	21.00	12.87	16.22	**21.97**
– **>40** (49±2.52)	**18.69**	**21.66**	17.61	18.20	**21.90**
– **High school graduate**	17.01	21.40	16.84	16.84	**18.77**
– **Bachelor degree**	**19.22**	**23.52**	11.70	17.48	**29.19**
– **Graduate degree**	14.78	15.54	12.64	14.86	**16.51**
– **Married**	**20.51**	**25.48**	17.71	19.53	**26.73**
– **Never married**	13.36	14.78	10.23	13.39	**16.31**
– **A.**	12.17	15.50	12.21	10.77	**17.33**
– **B.**	**22.45**	**25.49**	18.39	23.93	**24.21**
– **Overall (Mean±SD) of pACL (%)**	17.06 (±12.01)	**20.14** (±13.12)	16.43 (±16.42)	16.46 (±12.30)	**19.65** (±12.85)

(*) **16/18** - male subjects perceived high stress.
(**) **11/12** - female subjects perceived high stress.

and 2/3 of the subjects (63.33%) showed levels between moderate and high. It is important to note that employees did not perceive drastic changes of stress, from *low*⟷*high*.

As a summary of this first set of experiments, we can note that with our current data: (i) there is a slight bias in men towards lower levels of stress in their working environments, (ii) there is a clear difference between stress levels in companies, where an IT company showed higher stress levels than a social support company, (iii) about 2/3 of the employees perceived moderate to high stress and 23.33% perceived high stress, and (iv) there were no drastic changes between levels of stress.

Now we present a more detailed analysis for each category of objective variables and its relation with mood, and specifically with perceived stress levels.

4.1.1. Physical Activity Levels. Table VIII presents overall percentage of physical activity level with respect to perceived stress level (*High*, *Moderate*, and *Low*) on a daily basis for all 30 participants compared with demographic characteristics (age, gender, education,

Table IX: Overall average percentage of activity level (mean ± std.dev.) during working days and perceived level (SL) of Stress (h-*high*, m-*moderate*, l-*low*) by Gender.

	Men			***Women***		
	H (SL)	**M** (SL)	**L** (SL)	**H** (SL)	**M** (SL)	**L** (SL)
Monday:	**24.3±22.2**	16.2±16.2	21.6±18.0	12.3±12.1	13.0±7.0	**21.6±22.4**
Tuesday:	10.0±6.5	17.5±16.6	**22.2±14.4**	6.2± 3.1	12.3± 6.3	**16.5±7.7**
Wednesday:	18.0±19.8	19.8±18.3	**22.5±18.5**	12.6±8.4	13.2±7.7	**14.6±7.6**
Thursday:	19.0±20.7	20.7±18.6	**24.3±18.7**	9.6±8.0	**17.9±12.4**	14.3±13.9
Friday:	14.9±17.3	15.9±19.6	**20.4±19.5**	11.4±12.4	**17.7±13.6**	13.8±8.0

Table X: Overall average percentage of activity level (Mean ± std. dev.) by Job-Demands, Job-Control, Energy and Sleep-Quality perceived level (PL) with respect to gender.

	Men			***Women***		
	H (PL)	**M** (PL)	**L** (PL)	**H** (PL)	**M** (PL)	**L** (PL)
Job-Demand (%)	20.0 ± 16.7	17.9±14.8	**22.3 ±22.1**	**16.1±6.0**	13.1±8.4	11.7±6.2
Job-Control (%)	**19.0±14.6**	18.0±14.8	18.9±16.0	14.2± 6.5	**16.7±6.2**	13.3 ± 8.7
Energy (%)	**23.7±15.6**	19.9±14.8	17.8± 14.6	14.1± 9.7	15.3± 6.1	**16.8±9.7**
Sleep-Quality (%)	20.9± 14.3	22.1±15.8	**22.7± 17.4**	**16.6±6.9**	15.9±6.8	15.8± 6.9

marital status, number of children and organisations). Activity levels were normalized for each interval (9am-2pm and 2pm-5pm) or for a complete day. Based on this analysis we can see that:

- pACL during lower perceived stress times was associated with higher activity (19.65% of activity). In contrast, a high perceived stress showed less activity (16.43%).
- Subjects were more active during the second interval (2pm-5pm), with 20.14% pACL compared to 17.06% in the first interval.
- Male subjects had higher physical activity levels than female subjects.
- Following age, education level, and marital status, participants that reported *high* and *moderate* stress levels were associated mostly with lower pACL than when they have *low* stress.
- The age group of (≥40) showed more activity level than the rest when they perceived high stress level.
- The group of married subjects showed more activity than the never married group no matter their perceived stress level.
- Subjects from company B has less stress and were physically more active than subjects from company A.

Furthermore, separating overall activities into working days allowed us to compare pACL in different days of the week (as shown in Table IX). Results show that men have a higher pACL on Mondays when they perceive high stress, while the rest of the week they showed higher percentage of activity with low perceived stress. In contrast to women that showed more activity with low levels of perceived stress at the beginning of the week and more activity with moderate levels of stress at the end of the week.

As described in Section 3.2, it is important to collect information about occupational health, such as job-demands and job-control. In this regard, Table X shows mean scores on perceived job-demand, perceived job-control, perceived stress, and perceived energy for the respondents. From the table we can observe that low perceived job-demand was associated with higher physical activity level (22.3%) for male participants. In contrast, women showed increased activity levels when they perceive high job-demands. Similarly, men participants with higher pACL perceived higher energy. In contrast, women with higher pACL showed lower energy. Also men with lower pACL showed higher percentage of sleep-quality. In

summary, this table shows that in general men and women show opposite results in terms of perceived job-demand and energy with respect to their activity levels throughout the day.

4.1.2. Social Interaction. In contrast to our previous work [Ferdous et al. 2015], where we explored the correlation of total amount of verbal interaction per day with self-reported stress, in this study we expand that analysis, since now we explore the distribution of the verbal-interaction in an hourly basis and working intervals.

In Table XI we present a summary of social-interaction levels, with respect to different characteristics of the employees. Some findings are the following:

— Both male and female subjects showed higher social interaction in moments of high stress.
— However, analysing this data by age group we observe that older (and married) employees showed the opposite behaviour, they increased their social interaction during low levels of perceived stress.
— There is in general more social interaction in the afternoons than in the mornings.
— Another interesting behaviour appears across organisations. In this case subjects in organisation A showed higher social interaction than those in organisation B.

We explored further these measurements. We depict in Figure 3 social-interaction as a) percentage in hourly basis in a day, b) per day day of week, c) per hour within organisations and d) per day of the week by gender.

— A notable result is an homogeneous behaviour (similar shapes of the curves) of social interaction across stress levels (Figure 3 (a)), with higher interaction in the morning for moderate perceived stress and a higher interaction in the afternoon for high perceived stress.
— Another homogeneous behaviour is shown across organisations, where people decrease their social interaction near lunch time (12-13 hrs), see Figure 3 (a).
— The peak of the verbal interaction in High-Level of stress is achieved in the afternoon (one hour before the end of the working day).
— When subjects perceive high stress, social interaction drops on Thursdays and then increases again on Fridays, see Figure 3 (b).
— The social interaction varies with the perceived stress during the week, except on Mondays where it has similar values with the different perceived stress levels.
— With respect to gender, men showed a more stable social interaction across the weekdays. In contrast, women then to increase their interaction near the weekend, see Figure 3 (d).

Social interaction also include phone calls and SMS behaviours. Using the self-reported stress level, we were able to compare the phone activeness from 5767 phone calls and 5911 SMSs.[6] In Tables XII and XIII we explored the relation of phone calls and SMSs with respect to perceived level of stress. From these tables it can be seen that the number of phone-placed *Outgoing*, phone received *Incoming* and missing calls, was higher when subjects perceive less stress.

We also analysed the duration and length of calls and SMSs and some interesting observations are the following:

— In stress-full days, in most of the cases *Outgoing* calls have in average shorter duration.
— Longer duration of *Incoming calls* were associated with high perceived stress level.
— Almost in all cases a high number (and length) of *Incoming-SMS* and *Outgoing-SMS* were also related to *high* stress.
— Analysing the conversations by weekdays, high perceived stress was associated with longer duration of *Incoming-Calls* and the length of *Incoming-SMS's*, which in contrary to du-

[6] All marketing SMS's or responses from the GSM operators were excluded in this study.

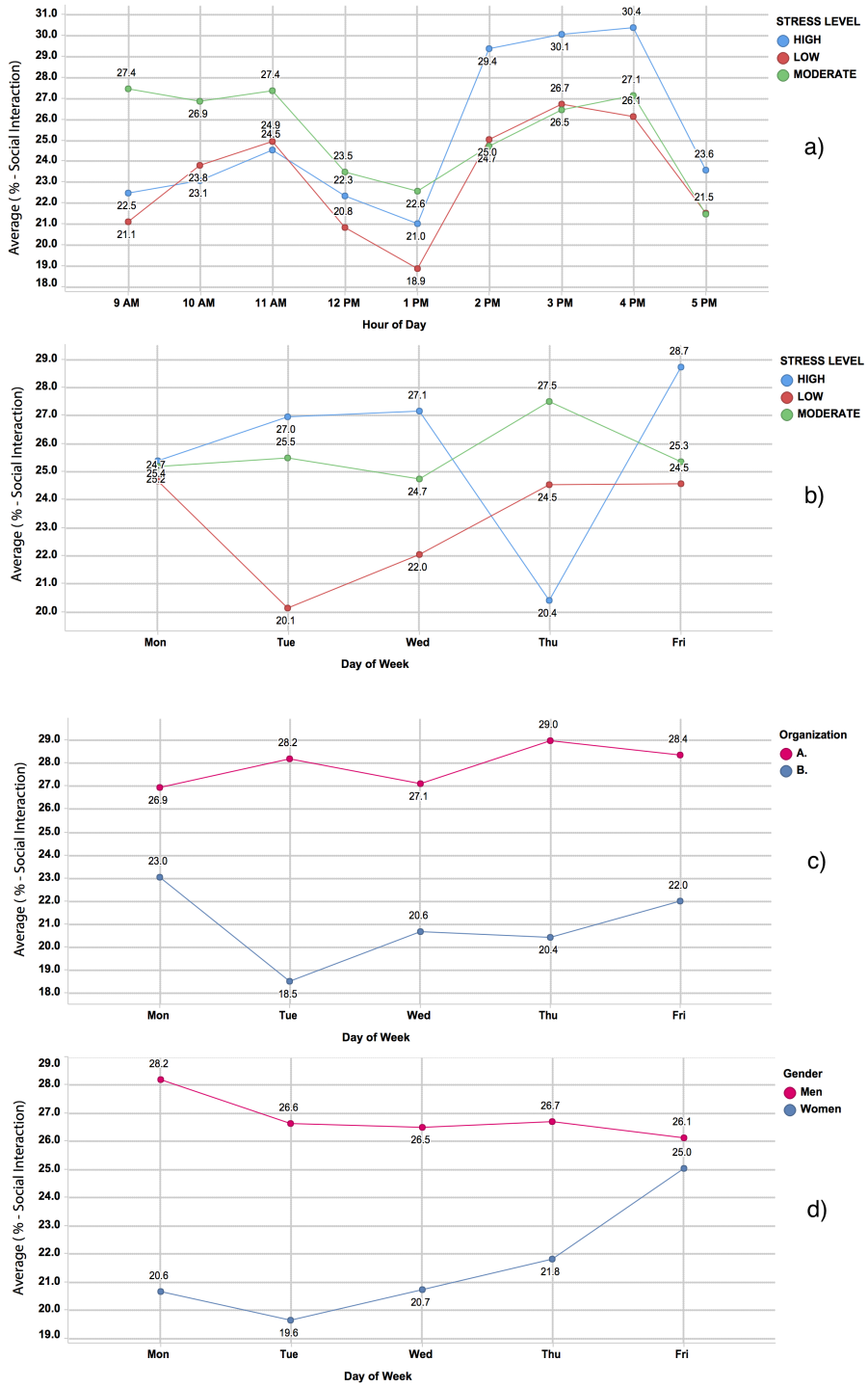


Fig. 3: Overall mean percentage of Social Interaction and Stress Level a) by working hours, b) by week days, c) by organisations and d) by gender

Table XI: Distribution of Social-Interaction (SI) by response intervals ([9am.-2pm.],[2pm. - 5pm.]) and Stress-Level (SL)

Distribution of SI by Gender, Age, Education, Marital Status, organisation	**SI [9-2]**	**SI [2-5]**	**High (SL)**	**Moderate (SL)**	**Low (SL)**	**Nr.** Employee
– **Male**	**25.67**	**28.75**	**27.88**(*)	27.54	25.74	18
– **Women**	20.17	23.83	**22.88**(**)	22.72	19.79	12
– **26-30** (28.6±1.95)	**25.46**	**29.53**	**28.57**	26.02	26.44	5
– **31-40** (35.33±2.4)	22.96	26.61	24.90	**26.67**	22.34	18
– **>40** (49±2.52)	22.73	24.84	22.97	22.54	**24.32**	7
– **High school graduate**	20.63	25.09	22.97	**26.95**	26.22	11
– **Bachelor degree**	24.23	**28.16**	**29.49**	26.28	23.12	10
– **Graduate degree**	**25.30**	26.94	22.81	**23.37**	21.24	9
– **Married**	21.75	25.02	22.91	21.92	**23.56**	15
– **Never married**	**24.68**	**28.07**	27.61	**28.45**	22.89	15
– **A.**	**26.40**	**30.41**	27.96	**29.64**	25.67	16
– **B.**	20.07	22.49	18.20	20.21	**21.80**	14
– **Overall (Mean±SD) of SI (%)**	23.61 (±10.53)	**26.93** (±11.04)	23.47 (±11.02)	24.58 (±10.47)	**25.28** (±11.67)	**30**

(*) **16/18** - male subjects perceived high stress.
(**) **11/12** - female subjects perceived high stress.

Table XII: Number of Phone-Calls by perceived Stress Level (SL).

	Nr. Phone Calls	**High SL**	**Moderate SL**	**Low SL**
Incoming:	1696 (100%)	355 (20.9%)	511(30.1%)	**830 (48.9%)**
Outgoing:	2912 (100%)	547 (18.7%)	839 (28.8%)	**1526 (52.4%)**
Missing:	1159 (100%)	220 (18.9%)	405 (34.9%)	**534 (46.1%)**

Table XIII: Number of SMS's by perceived Stress Level (SL).

	Nr. SMS	**High SL**	**Moderate SL**	**Low SL**
Incoming:	3767 (100%)	1067 (28.3%)	801 (21.2%)	**1899 (50.4%)**
Outgoing:	2144 (100%)	697 (32.5%)	710 (33.1%)	**737 (34.3%)**

ration of *Outgoing-Calls* and length of *Outgoing-SMS's* is lower when the employees perceive high stress. Similarly, having high job-demands was associated with lower duration of phone-call and length of SMS's in all categories.

Moreover, in Figures in 4(a) and 4(b) we depict the frequency of the most common contact for phone calls and SMSs (blue line) for every subject. From these figures we note a higher frequency of phone-calls and SMS's with the most common contacted number when they perceive high stress levels (average frequency of most frequent contacts is shown with red line). In contrary, in low and moderate stress the frequency of phone-call is in average lower. These results shows that a higher frequency of the phone-calls and SMS's can be an indicator of stress during the working times.

4.1.3. Location Changes. To analyse location changes we measured the number of clusters obtained from different locations throughout the entire monitoring interval (see Table XIV). From all three sources it is evident that overall subjects tend to reduce visiting different places or going further away from work environments when they perceive High-Stress level during working days. We obtained more Cell-Tower and WiFi clusters from both parameters due to frequent scanning. Changes of clusters in WiFi represent changes of indoor locations, such as changing environments, areas, departments or either having mini-breaks in specific hours. In contrast, using Google Maps locations show distance and most visited places outdoors.

4.1.4. Application Usage. Another source that provides information relevant to subjects daily activities at work is the usage of the smartphone applications. Recall that we divided the type of applications subjects ran on their devices during the working days and we categorized them into system and social applications (as described in Section 3.3.4). Next, we examine the frequency (number of accesses) and the duration of the applications used

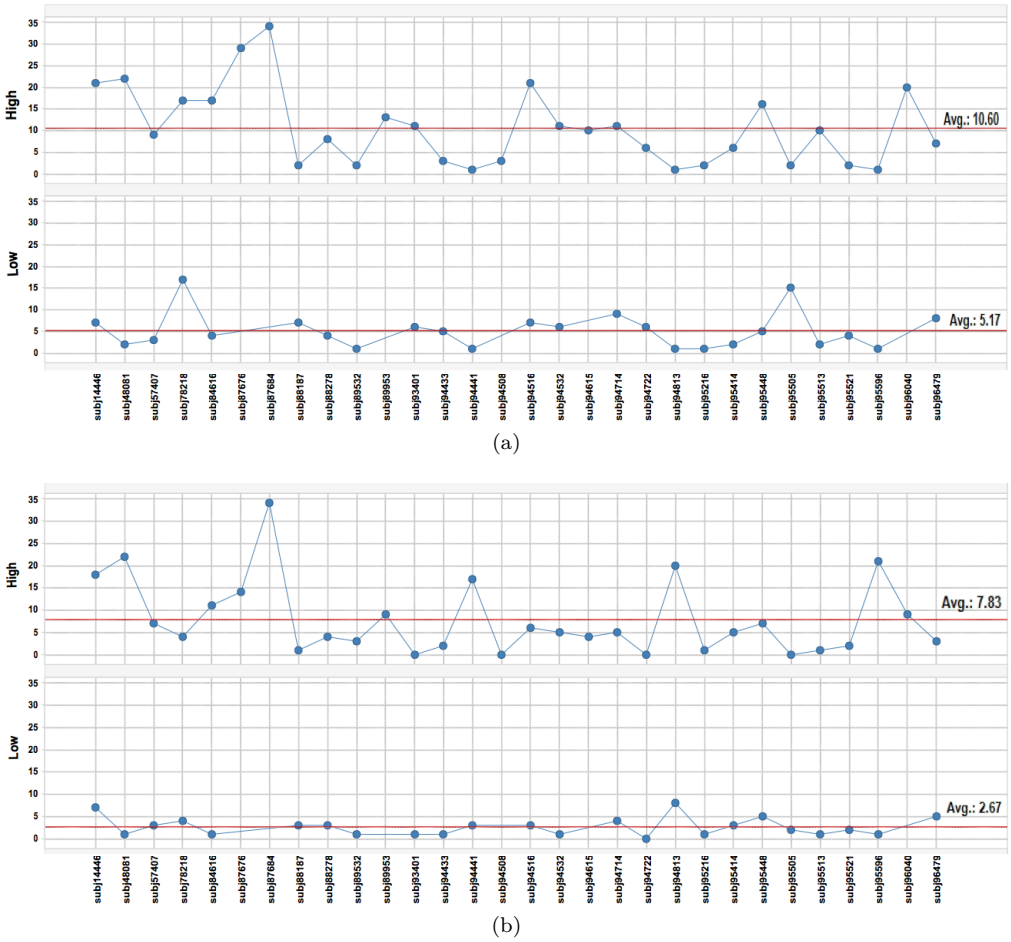


Fig. 4: Frequency of the most common contact (a) calls and (b) SMSs for each subject by perceived stress level (High and Low).

Table XIV: Overall number of clusters obtained from location using the DBSCAN algorithm by perceived Stress Level (SL). Descriptive statistics (Mean±SD) provide information of overall number of clusters retrieved from the 30-subjects throughout the entire monitoring period.

Locations	**Clusters** 9am-5pm Nr.(Mean±SD)	**High**-(SL) Nr.(Mean±SD)	**Moderate**-(SL) Nr.(Mean±SD)	**Low**-(SL) Nr.(Mean±SD)
Cell:	1383 (1.05±0.38)	230 (1.01±0.39)	349 (1.07±0.40)	**527** (1.05±0.33)
WiFi AP's:	2663 (1.40±1.38)	486 (1.42±1.41)	742 (1.49±1.35)	**961** (1.55±1.39)
Google Maps:	628 (0.48±0.78)	143 (0.63±1.01)	158 (0.48±0.90)	**234** (0.46±0.85)

Table XV: Overall Number/Duration (seconds) of phone application usage by perceived Stress Level (SL). Descriptive statistics (Mean±SD) provides overall usage of applications from 30 subjects during the entire monitoring period.

Perceived Stress Level	**Frequency System-Apps** Nr.(Mean±SD)	**Frequency Social-Apps** Nr.(Mean±SD)	**Duration System-Apps** Nr.(Mean±SD)	**Duration Social-Apps** Nr.(Mean±SD)
High	5531 (24.0±26.1)	357 (3.5±4.0)	48445 (211.0±157.2)	4621 (45.3±52.5)
Moderate	7823 (25.2±28.5)	508 (4.0±4.2)	57607 (185.2±153.3)	7420 (57.0±73.2)
Low	**13787** (31.0±28.2)	**966** (4.3±4.2)	**88782** (197.2±150.3)	**9582** (42.3±65.2)

and contrast them with the perceived self-reported stress on a daily basis (see Table XV). Results show that in stress-less times subjects tend to use longer times the smartphone (both with social and system applications). This also seems a good indicator for identifying perceived stress levels.

In summary, from these results we can draw the following conclusions:

— Activity levels changed with perceived stress and with weekdays.
— There is an opposite behaviour of activity levels in male and female in terms of job-demand and energy.
— There is more social interaction with higher stress levels except for older people that show an opposite behaviour.
— There is more social interaction during the afternoons.
— There is an increase level in social interaction by women towards the end of the week.
— There is a very different social interaction among employees of different companies. Curiously the company with higher stress levels also have higher percentages of social interaction.
— There are shorter outgoing calls and longer incoming calls during high stress levels.
— People use much more their smartphones during lower perceived stress levels.

In this section we have presented an initial analysis for each category of subjective data contrasted with the perceived stress levels. The next section explores the correlation between objective and subjective variables, as well as the importance of each variable for mood prediction.

4.2. Correlation between objective and subjective data

We conducted two correlation analyses to investigate the association between four factors: perceived stress, negative-mood, positive-mood, and overall mood score. Emotions were divided in two categories: negative-mood (tense, stress, angry, anxious and sad) and positive-mood (friendly, energetic, cheerful and being good at current activity). An overall score derived from both types of emotions was obtained by subtracting negative mood scores from positive scores.

A two-tailed Pearson correlation and multiple linear regression analysis were performed to examine the relationships among perceived stress and well-being (moods) scores with objective measurements. First, we performed the correlation tests between objective and subjective variables. The Pearson correlation coefficient ρ was used, with a statistically significant value when $\rho<0.05$ (*) and $\rho<0.01$ (**). In Table XVI we present the correlations between objective measurements (rows) and subjective measurements derived from self-reported stress, negative-mood score, positive-mood score, and overall-mood score (columns) and we can make some observations:

— For stress level, physical activity (r= -0.153,**, N=1465), number of WiFi (r= -0.087, **, N=1456) and cellular location (r= -0.070, *, N=1456), number and duration of outgoing phone calls (r= -0.098, *, N=1120), number and length of SMS responses (r= 0.090,*, N=505), and number of system apps (r= -0.129, **, N=1292) obtained statistically significant correlations.

Table XVI: Pearson correlations between objective variables and perceived Stress Level, Negative Mood Score, Positive Mood Score, and Overall Mood Score.

Objective Variables	**Stress Level**	**Negative Mood**	**Positive Mood**	**Total Mood Score**
Physical Activity Levels	-0.153**	-0.112**	0.071**	0.116**
Cellular Location	-0.070 *	-0.070*	0.033	0.065*
Google-Maps Location	0.051	0.017	0.079*	0.033
Wifi Locations	0.087**	0.039	-0.120**	-0.093**
Social Interaction (SI)	0.032	0.059*	-0.142**	-0.119**
Number-Outgoing-Calls	-0.980**	-0.112**	0.083**	0.121**
Number-Incoming-Calls	-0.005	-0.090**	-0.019	0.05
Missed-Incoming-Calls	-0.006	-0.023	-0.012	0.009
Duration-Outgoing-Calls	-0.098**	-0.097**	0.101**	0.123**
Duration-Incoming-Calls	0.037	-0.034	0.091*	0.074*
Number-SMS-Outgoing	0.090**	-0.071*	0.004	0.05
Number-SMS-Incoming	0.006	-0.012	-0.044	-0.016
Length-SMS-Outgoing	-0.154**	-0.153**	0.106*	0.156**
Length-SMS-Incoming	0.013	-0.028	0.088*	0.069
Duration-Application-System	0.008	-0.021	-0.024	0.001
Duration-Application-Social	0.067	0.067	-0.218**	-0.161**
Number-Application-System	-0.129**	-0.181**	0.194**	0.228**
Number-Application-Social	-0.060	-0.040	-0.004	0.024

Significant at levels: $\rho < 0.05$ **(*)**; $\rho < 0.01$ **(**)**.

Table XVII: Significant results from the multiple regression using Objective measurements with respect to Stress and Total Mood Score.

	Stress			***Total Mood Score***		
Objective Variables	β	**t**	ρ	β	**t**	ρ
Physical-Activity Levels	-.0111	-5.88	***0.001***	-.0111	-5.88	***0.001***
Cellular Location	-.2333	-2.29	***0.022***	.0376	2.10	***0.036***
Google-Maps Location	.0685	1.65	0.100	.0077	1.06	0.289
Wifi Location	.0057	3.34	***0.001***	-.0041	-3.58	***0.001***
Social Interaction (SI)	.0001	1.13	0.258	-.0008	-4.28	***0.001***
Number-Outgoing-Calls	-.0374	-3.31	***0.001***	.0081	4.07	***0.001***
Number-Incoming-Calls	-.0033	-0.17	0.866	.0058	1.68	0.093
Missed-Incoming-Call	-.0015	-0.19	0.847	.0004	0.29	0.769
Duration-Outgoing-Call	-.0125	-2.73	***0.006***	.0026	3.43	***0.001***
Duration-Incoming-Call	.0048	1.01	0.313	.0016	2.02	***0.044***
Number-SMS-Outgoing	.0188	3.05	***0.002***	.0018	1.68	0.092
Number-SMS-Incoming	.0003	0.19	0.850	-.0001	-0.54	0.590
Length-SMS-Outgoing	-.0015	-3.49	***0.001***	.0003	3.55	***0.001***
Length-SMS-Incoming	.0001	0.34	0.737	.0001	1.72	0.086
Duration-Application-System	.0001	0.31	0.759	.0001	0.03	0.976
Duration-Application-Social	.0001	1.43	0.153	-.0001	-3.47	***0.001***
Number-Application-System	-.0061	-4.69	***0.001***	.0020	8.42	***0.001***
Number-Application-Social	-.0189	-1.27	0.203	.0014	0.51	0.610

— In particular, for missing calls we expected to have correlation with different factors. We assume that during a stress-full day, participants are more prone to reject the phone-conversations due to responsibilities and task that they have to achieve. However, it was shown to have a weak correlations with the stress factor.

— Negative emotions show high correlation with physical activity, number of system apps, social interaction information and social activeness (number of incoming and outgoing phones calls, and outgoing SMS's).

— Social interaction information, the number of outgoing calls and the use of social applications showed high correlation with positive mood scores.

— Information from duration of social applications and location information obtained low correlation with negative emotions. This is interesting because these same two variables obtained high correlation with positive emotions. Similarly the number of incoming calls show low correlation with positive emotions but is highly correlated with negative emotions.

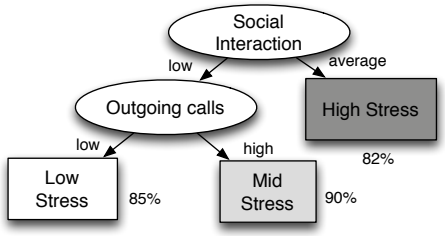


Fig. 5: An example of a decision tree, each oval represent a decision node which contain arrows to other decision nodes. Squares are leaves (terminal nodes) that give the classification value, in this case they represent Low, Mid or High level of stress.

4.3. Multiple regression

In order to obtain the best possible model for prediction of stress and total mood score we decided to use multiple linear regression. We found that regression result was significant for stress (r^2=0.3912, F(18,64)=2.28, ρ<0.008) and with total mood-scores (r^2=0.4419, F(18,64)=2.81, ρ<0.001) using all features (as shown in Table XVII, which depict the name of each feature, the regression coefficient, β, the distribution value, t, the and ρ-value for each used feature). This results show that selected features are having an effect on predicting stress (ρ<0.008) and total mood score (ρ<0.001). Similarly, several objective variables (with italic typeface in Table XVI) show significant correlation with perceived stress and total mood score of the subjects. It is interesting to note that these objective variables (physical activity, cellular and Wifi location, number and duration of outgoing calls, number and length of outgoing SMSs, and number of applications) also show significant linear correlation using Pearson.

In summary:

- Stress level is highly correlated with physical activity, WiFi location, number and duration of outgoing calls and SMS, and with social apps. These values are consistent with what was obtained with multiple linear regression.
- In contrast, negative mood is highly correlated with the number of incoming calls and is not correlated with WiFi location.
- Similarly, positive mood is highly correlated with social interaction and duration of social apps but it is not correlated with the number of outgoing SMS.

The next section deals with the problem of mood prediction based only on objective data obtained from phone usage.

5. PREDICTION

In the previous section we analysed the relation between the measured objective variables with perceived stress also we presented results showing many features correlated with stress levels. Thus, our next step is to make a model capable of predicting the stress level given the objective variables.

Predicting perceived stress of the user can be seen as a classification problem. In this case, the attributes correspond to each feature related to the objective variables and the class to predict is the self-reported stress level (low, moderate, high). Since we are interested in analysing behaviour changes or patterns that may appear in daily activities, we used decision tress [Quinlan 1993] which can be easily understood. Our approach was to build a decision tree for each subject of the study, with the idea of analysing individual behaviours and models.

In Figure 5 we present a decision tree that classifies the Stress level of a subject in the study. The subject shows low levels of stress when having an average level of social

Table XVIII: Stress Prediction using decision tress before and after applying a Semi-supervised learning approach. Overall classes represent overall number of labelled classes in supervised learning and after performing unsupervised learning methods.

Subjects	**Supervised**	**Semi-Supervised**	**Overall Increase in Prediction** (%)
Acc. Mean±SD:	67.57% (±15.60%)	**71.73%** (±15.25%)	**4.20%** (±9.52%)
Labelled Data:	79.97% (1465/**1832**)	**94.00%** (1722/**1832**)	**14.03%**
Precision (%):	65.4%	**68.9%**	
Recall (%):	68.9%	**73.0%**	
F-Score (%):	66.0%	**70.0%**	

interaction, or when the social interaction and number of outgoing calls is low. On the contrary, if this subject had low level of social interaction but a high number of outgoing calls then it is more probable to have a mid level of stress.

We performed classification of the stress variable using the C4.5 algorithm [Quinlan 1993] and 10-fold cross validation for each user. Table XVIII presents the classification accuracy and the average precision, recall and f-measure for stress level for the 30 subjects. In average the accuracy obtained was 67.57%. However, as previously mentioned, the dataset contained 20% of missing data. This is an important portion which can be exploited with a semi-supervised learning (SSL) technique.

5.1. Semi-supervised learning (SSL)

In most real-world datasets it is common to have missing data. The most basic approach is to ignore those instances. However, that information even when it is not complete can be helpful and should not be discarded. Semi-supervised learning [Zhu 2005; Longstaff et al. 2010] has been suggested as a method aiming to address this issues in machine learning. The main objective of semi-supervised learning is to learn from both labelled and unlabelled data, i.e., by exploiting unlabelled samples to improve the learning performance.

For this study we consider one of the most common methods of SSL that uses a single classifier called Self-Training [Triguero et al. 2015]. This method works by selecting the most confident unlabelled points, together with their predicted labels and then adding those to the training set. In each iteration the newly high-confidence (>80%) labelled instances are added to the original labelled data. Note that the classifier uses its own predictions to teach itself. The classifier is re-trained and the procedure repeated. In particular we have used decision tree classifiers with the C4.5 algorithm in Weka [Hall et al. 2009], with 10 fold cross-validation with default values.

In Table XVIII we present the results in terms of accuracy after applying the SSL approach on all subjects in the study. Using the Self-Training method, we were able to improve the accuracy on predicting stress to 71.73% (+4.20%). In Table XVIII we show that using Self-Training we were able to reduce the number of unlabelled data from 20% to 6%. We have also analysed accuracy results by gender. Results show that the *Male* achieved better accuracy 72%(*Precision: 73.5%; Recall: 78.5%*) for supervised approach and 76.4% (*Precision: 73.5%; Recall: 78.5%*) for SSL, in contrast to *Female* with 59.8%(*Precision: 59.0%;Recall: 60.0%*) for supervised and 64.8% (*Precision: 62.0%; Recall: 65.0%*) for SSL approach.

In this section we have shown that simple models can be generated to predict stress levels with around 70% of accuracy. Unsurprisingly, most of the models used the relevant features identified in the previous section. It is also shown that a slight improvement in the predictive performance can be achieved with a simple semi-supervised learning algorithm. It is left as future work to use other more powerful classifiers and semi-supervised techniques.

6. CONCLUSIONS

Stress at work is an important problem affecting employees' health, decreasing job-control, organisational performance, and reducing the quality of life in general. Stress assessment is a complex issue, in particular detecting stress using non obtrusive approaches.

In this work, we presented an extensive analysis based on real data from 30 users in two organisations related to stress using information derived from smartphones. We contrasted objective variables, obtained from smartphones, such as physical activity, location, social interaction and social activity with respect to perceived stress levels, considering several demographics (gender, age, education and marital status). Correlation analysis was used to analyse the possibility of using smartphones derived data to predict stress levels. Finally, we presented results using decision trees to classify stress level, including the incorporation of semi-supervised learning techniques.

A summary of the most important findings in this study.

- There is correlation between objective data such as: location information (WiFi and Google Location data), social interaction, and information from phone calls and SMS with subjective data that represents mood of the user (i.e., level of stress).
- Overall physical activity during lower perceived stress times throughout the entire monitoring period was associated with higher activity. In contrast, a high perceived stress showed lower physical activity.
- With respect to gender, men showed a more stable social interaction across the weekdays. In contrast, women then to increase their interaction near the weekend.
- Our results suggests that the more social the subject is the more stressed he gets, this can be explained because the subject is probably talking with colleagues about work which increases its stress. On the other side there is negative correlation between duration of calls and stress, the reason could be that the subject is stressed so she has no time to spend on calls.
- Based on smartphone data it is possible to predict stress using decision trees. However, missing data is an aspect to take into account. In this work using semi-supervised learning techniques we increased the accuracy from 67.57% to 71.73% for predicting stress.

Some conclusions of this work are:

- There is clearly a high to moderate perceived stress in most of the employees. This confirms some of the findings on other reported studies about stress. The possible consequences of stress motivated our work for finding unobtrusive ways to detect it, via smartphones, and analyse in ~~more~~ deep the most relevant aspects related with changes in the behaviour of employees under different stress conditions. We believe that this is an important step towards a better understanding of behaviour of employees under stress and to design remedy actions.
- It appears that women tend to present higher percentage levels of perceived stress. This does not necessarily mean that they are more stressed, but at least that they perceive it more. Whether this has to do this with higher sensitivity levels in women than men, a biased finding due to our small sample size or to a more profound reason related to gender, this requires further and deeper studies.
- Perceived stress varies among companies and this could be related to their working conditions. Identifying working conditions on companies with low levels of stress could help to establish better working policies to reduce stress among employees.
- There appears to be different behaviours in some job-related aspects in relation to stress between men and women. Although again this needs a more deeper and thorough study, if it is the case it could help to improve some working conditions based on gender.

— The use of smartphones has become part of the daily activities of people and our experiments showed that there are clear changes in their use (phone calls, SMSs, apps) under different stress conditions.
— There is a clear correlation between how people behave at work (physical activity, WiFi location, number and duration of outgoing calls and SMS, and with social apps) and stress levels. This could be easily monitored with current smartphones, as shown in this research, to detect possible stress levels and help to implement corrective measures.

6.1. Future work

As future work we would like to continue with larger scale studies for the detection and prevention of stress at work. We also plan to analyze more in depth the decision trees obtained for each subject in order to obtain clusters of people who behave similarly; this could help us to build prediction models for new users with few data. We also want to test different levels of granularity for the time dimension to see whether stress patterns appear during different time intervals. At the end we would like to have an app in smartphones that could reliably detect stress levels in working environments.

Acknowledgements

The authors would like to thank the anonymous reviewers who provided detailed feedback. Their suggestions and comments certainly improved the paper. This research was supported partially by grant 335245/234507 by Consejo Nacional de Ciencia y Tecnologia (CONACyT) Mexico.

REFERENCES

J Bakker, M Pechenizkiy, and N Sidorova. 2011. What's Your Current Stress Level? Detection of Stress Patterns from GSR Sensor Data. In *Data Mining Workshops (ICDMW), 2011 IEEE 11th International Conference on*. Vancouver, Canada, 573–580.

Gerald Bauer and Paul Lukowicz. 2012. Can smartphones detect stress-related changes in the behaviour of individuals?. In *Pervasive Computing and Communications Workshops (PERCOM Workshops), 2012 IEEE International Conference on*. Lugano, Switzerland, 423–426.

C.M. Bernaards, M.P. Jans, S.G. Van den Heuvel, I.J. Hendriksen, I.L. Houtman, and P.M. Bongers. 2006. Can strenuous leisure time physical activity prevent psychological complaints in a working population? *Occupational and environmental medicine* 63, 1 (2006), 10–16.

Derya Birant and Alp Kut. 2007. ST-DBSCAN: An algorithm for clustering spatial–temporal data. *Data & Knowledge Engineering* 60, 1 (2007), 208–221.

Andrey Bogomolov, Bruno Lepri, Michela Ferron, Fabio Pianesi, and Alex Sandy Pentland. 2014. Daily Stress Recognition from Mobile Phone Data, Weather Conditions and Individual Traits. In *Proceedings of the ACM International Conference on Multimedia*. Orlando, Florida, USA, 477–486.

Paulien M Bongers, Cornells R de Winter, Michiel AJ Kompier, and Vincent H Hildebrandt. 1993. Psychosocial factors at work and musculoskeletal disease. *Scandinavian journal of work, environment & health* (1993), 297–312.

Sheldon Cohen and Thomas A Wills. 1985. Stress, social support, and the buffering hypothesis. *Psychological bulletin* 98, 2 (1985), 310.

Terry L Conway, Ross R Vickers Jr, Harold W Ward, and Richard H Rahe. 1981. Occupational stress and variation in cigarette, coffee, and alcohol consumption. *Journal of Health and Social Behavior* (1981), 155–165.

R. Ferdous, V. Osmani, J. B. Márquez, and O. Mayora. 2015. Investigating correlation between verbal interactions and perceived stress. In *2015 37th Annual International Con-*

ference of the IEEE Engineering in Medicine and Biology Society (EMBC). 1612–1615. DOI:http://dx.doi.org/10.1109/EMBC.2015.7318683

F Fleshner. 2005. Physical activity and stress resistance: sympathetic nervous system adaptations prevent stress-induced immunosuppression. *Exercise and sport sciences reviews* 33, 3 (2005), 120–126.

FUNF. 2015. Open Sensing Framework-FUNF. http://funf.org/about.html.. (2015). Accessed: 2015-AUG-29.

E. Garcia-Ceja, V. Osmani, and O. Mayora. 2016. Automatic Stress Detection in Working Environments From Smartphones' Accelerometer Data: A First Step. *IEEE Journal of Biomedical and Health Informatics* 20, 4 (July 2016), 1053–1060. DOI:http://dx.doi.org/10.1109/JBHI.2015.2446195

Karen Glanz, Barbara K Rimer, and Kasisomayajula Viswanath. 2008. *Health behavior and health education: theory, research, and practice*. John Wiley & Sons.

Agnes Grünerbl, Amir Muaremi, Venet Osmani, Gernot Bahle, Stefan Oehler, Gerhard Tröster, Oscar Mayora, Christian Haring, and Paul Lukowicz. 2015. Smartphone-based recognition of states and state changes in bipolar disorder patients. *IEEE Journal of Biomedical and Health Informatics* 19, 1 (2015), 140–148.

Mark Hall, Eibe Frank, Geoffrey Holmes, Bernhard Pfahringer, Peter Reutemann, and Ian H Witten. 2009. The WEKA data mining software: an update. *ACM SIGKDD explorations newsletter* 11, 1 (2009), 10–18.

Fredric J Harris. 1978. On the use of windows for harmonic analysis with the discrete Fourier transform. *Proc. IEEE* 66, 1 (1978), 51–83.

Ling He, Margaret Lech, MC Maddage, and Nicholas Allen. 2009. Stress detection using speech spectrograms and sigma-pi neuron units. In *Natural Computation, 2009. ICNC'09. Fifth International Conference on*, Vol. 2. Tianjin, China, 260–264.

Jennifer A Healey and Rosalind W Picard. 2005. Detecting stress during real-world driving tasks using physiological sensors. In *Intelligent Transportation Systems, IEEE Transactions on*, Vol. 6. 156–166.

Per Hedelin and Dieter Huber. 1990. Pitch period determination of aperiodic speech signals. In *Acoustics, Speech, and Signal Processing, 1990. ICASSP-90., 1990 International Conference on*. IEEE, Albuquerque, NM, USA, 361–364.

Karen Hovsepian, Mustafa al'Absi, Emre Ertin, Thomas Kamarck, Motohiro Nakajima, and Santosh Kumar. 2015. cStress: towards a gold standard for continuous stress assessment in the mobile environment. In *Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing*. ACM, 493–504.

S Jerritta, M Murugappan, R Nagarajan, and Khairunizam Wan. 2011. Physiological signals based human emotion recognition: a review. In *Signal Processing and its Applications (CSPA), 2011 IEEE 7th International Colloquium on*. IEEE, Penang,Malaysia, 410–415.

Desok Kim, Yunhwan Seo, Jaegeol Cho, and Chul-Ho Cho. 2008. Detection of subjects with higher self-reporting stress scores using heart rate variability patterns during the day. In *Engineering in Medicine and Biology Society, 2008. EMBS 2008. 30th Annual International Conference of the IEEE*. IEEE, Vancouver, Canada, 682–685.

Karen Korabik, Lisa M McDonald, and Hazel M Rosin. 1993. Stress, coping, and social support among women managers. *Women, work and coping* (1993), 133–153.

Robert Likamwa, Yunxin Liu, Nicholas D Lane, and Lin Zhong. 2013. Moodscope: building a mood sensor from smartphone usage patterns. In *Proceeding of the 11th annual international conference on Mobile systems, applications, and services*. Taipei, Taiwan, 389–402.

Karen Kay-Lynn Liu. 2004. *A personal, mobile system for understanding stress and interruptions*. Master's thesis. MIT Media Arts and Science.

Brent Longstaff, Sasank Reddy, and Deborah Estrin. 2010. Improving activity classification for health applications on mobile devices using active and semi-supervised learning. In

4th International Conference on Pervasive Computing Technologies for Healthcare. IEEE, Munich, Germany, 1–7.

Hong Lu, Denise Frauendorfer, Mashfiqui Rabbi, Marianne Schmid Mast, Gokul T. Chittaranjan, Andrew T. Campbell, Daniel Gatica-Perez, and Tanzeem Choudhury. 2012. StressSense: Detecting Stress in Unconstrained Acoustic Environments Using Smartphones. In *Proceedings of the 2012 ACM Conference on Ubiquitous Computing (UbiComp '12)*. ACM, New York, NY, USA, 351–360. http://doi.acm.org/10.1145/2370216.2370270

Rafer S Lutz, Matthew A Stults-Kolehmainen, and John B Bartholomew. 2010. Exercise caution when stressed: stages of change and the stress–exercise participation relationship. *Psychology of Sport and Exercise* 11, 6 (2010), 560–567.

Christina Maslach, Wilmar B Schaufeli, and Michael P Leiter. 2001. Job burnout. *Annual review of psychology* 52, 1 (2001), 397–422.

Uwe Maurer, Asim Smailagic, Daniel P Siewiorek, and Michael Deisher. 2006. Activity recognition and monitoring using multiple sensors on different body positions. In *Wearable and Implantable Body Sensor Networks, 2006. BSN 2006. International Workshop on*. Cambridge, MA, USA.

Alban Maxhuni, Angélica Muñoz-Meléndez, Venet Osmani, Humberto Perez, Oscar Mayora, and Eduardo F. Morales. 2016. Classification of bipolar disorder episodes based on analysis of voice and motor activity of patients. *Pervasive and Mobile Computing* (2016), –. DOI:http://dx.doi.org/10.1016/j.pmcj.2016.01.008

Douglas M McNair, Maurice Lorr, and Leo F Droppleman. 1971. *Profile of mood states*. Univ.

Amir Muaremi, Bert Arnrich, and Gerhard Tröster. 2013. Towards measuring stress with smartphones and wearable devices during workday and sleep. *BioNanoScience* 3, 2 (2013), 172–183.

NOISH. 1999. Stress at Work. http://www.cdc.gov/niosh/docs/99-101/. (1999). Accessed: 2015-AUG-29.

P. Paoli and A. Parent-Thirion. 2003. Working Conditions in the acceding and candidate countries. *European Foundation for the Improvement of Living and Working Conditions, Office for Official Publications of the European Communities.* 6 (2003). www.eurofound.europa.eu/publications/htmlfiles/ef0306.htm

Frank J Penedo and Jason R Dahn. 2005. Exercise and well-being: a review of mental and physical health benefits associated with physical activity. *Current opinion in psychiatry* 18, 2 (2005), 189–193.

Karin I Proper, Vincent H Hildebrandt, Allard J Van der Beek, Jos WR Twisk, and Willem Van Mechelen. 2003. Effect of individual counseling on physical activity fitness and health: a randomized controlled trial in a workplace setting. *American journal of preventive medicine* 24, 3 (2003), 218–226.

John Ross Quinlan. 1993. *C4. 5: programs for machine learning*. Morgan Kaufmann.

CC Robusto. 1957. The cosine-haversine formula. *Amer. Math. Monthly* (1957), 38–40.

Taehwan Roh, Kyeongryeol Bong, Sunjoo Hong, Hyunwoo Cho, and Hoi-Jun Yoo. 2012. Wearable mental-health monitoring platform with independent component analysis and nonlinear chaotic analysis. In *Engineering in Medicine and Biology Society (EMBC), 2012 Annual International Conference of the IEEE*. San Diego, CA, USA, 4541–4544.

Pedro Sanches, Kristina Höök, Elsa Vaara, Claus Weymann, Markus Bylund, Pedro Ferreira, Nathalie Peira, and Marie Sjölinder. 2010. Mind the body!: designing a mobile stress management application encouraging personal reflection. In *DIS '10 Proceedings of the 8th ACM Conference on Designing Interactive Systems*. Aarhus, Denmark, 47–56.

Akane Sano and Rosalind W. Picard. 2013. Stress recognition using wearable sensors and mobile phones. In *Affective Computing and Intelligent Interaction (ACII), 2013 Humaine Association Conference on*. Geneva, Switzerland, 671–676.

Elke Schneider, Sarah Copsey, and Xabier Irastorza. 2010. *OSH in figures: work-related*

musculoskeletal disorders in the EU-facts and figures. Office for Official Publications of the European Communities.

Archana Singh-Manoux, Michael G Marmot, and Nancy E Adler. 2005. Does subjective social status predict health and change in health status better than objective status? *Psychosomatic Medicine* 67, 6 (2005), 855–861.

Charles D. Spielberger, Peter R. Vagg, and Carol F. Wasala. 2003. Occupational stress: Job pressures and lack of support. (2003), 185–200.

Hélène Sultan-Taïeb, Jean-François Chastang, Malika Mansouri, and Isabelle Niedhammer. 2013. The annual costs of cardiovascular diseases and mental disorders attributable to job strain in France. *BMC public health* 13, 1 (2013), 748.

Isaac Triguero, Salvador García, and Francisco Herrera. 2015. Self-labeled techniques for semi-supervised learning: taxonomy, software and empirical study. *Knowledge and Information Systems* 42, 2 (2015), 245–284.

Michel François Valstar, Bihan Jiang, Marc Mehu, Maja Pantic, and Klaus Scherer. 2011. The first facial expression recognition and analysis challenge. In *Automatic Face & Gesture Recognition and Workshops (FG 2011), 2011 IEEE International Conference on.* Santa Barbara, CA, USA, 921–926.

Vladimir Vapnik, Steven E Golowich, and Alex Smola. 1997. Support vector method for function approximation, regression estimation, and signal processing. *Advances in neural information processing systems* (1997), 281–287.

Alessandro Vinciarelli, Maja Pantic, and Hervé Bourlard. 2009. Social signal processing: Survey of an emerging domain. *Image and Vision Computing* 27, 12 (2009), 1743–1759.

Joe H Ward Jr. 1963. Hierarchical grouping to optimize an objective function. *Journal of the American statistical association* 58, 301 (1963), 236–244.

Jacqueline Wijsman, Bernard Grundlehner, Hao Liu, Hermie Hermens, and Julien Penders. 2011. Towards mental stress detection using wearable physiological sensors. In *Engineering in Medicine and Biology Society, EMBC, 2011 Annual International Conference of the IEEE.* IEEE, 1798–1801.

Xiaojin Zhu. 2005. Semi-supervised learning literature survey. (2005).