

HealthyOffice: Mood Recognition At Work Using Smartphones and Wearable Sensors

Alexandros Zenonos^{*†}, Aftab Khan[†], Georgios Kalogridis[†], Stefanos Vatsikas[†],
Tim Lewis[†], Mahesh Sooriyabandara[†]

[†]Toshiba Research Europe Ltd., Telecommunications Research Laboratory, Bristol, UK
{alexandros.zenonos, aftab.khan, george, stefanos.vatsikas, tim.lewis, m.sooriyabandara}@toshiba-trel.com

^{*}University of Southampton, UK
az2g13@ecs.soton.ac.uk

Abstract—Stress, anxiety and depression in the workplace are detrimental to human health and productivity with significant financial implications. Recent research in this area has focused on the use of sensor technologies, including smartphones and wearables embedded with physiological and movement sensors. In this work, we explore the possibility of using such devices for mood recognition, focusing on work environments. We propose a novel mood recognition framework that is able to identify five intensity levels for eight different types of moods every two hours. We further present a smartphone app (‘HealthyOffice’), designed to facilitate self-reporting in a structured manner and provide our model with the ground truth. We evaluate our system in a small-scale user study where wearable sensing data is collected in an office environment. Our experiments exhibit promising results allowing us to reliably recognize various classes of perceived moods.

I. INTRODUCTION

It is estimated that life expectancy of people in a stressful environment can be reduced by up to 3 years [1]. Related to this, health and wellbeing of employees in a work environment is known to have a direct impact on the overall productivity of a company. The cost of work-related stress, anxiety and depression in Europe is estimated by the European Commission to be €617 billion per year [2]. Similarly, according to conservative estimations, stress in the United States has an annual cost of US\$200 billion [2].

There is a growing need for understanding and uplifting the overall mood of employees, to help improve occupational health, safety, wellbeing, and subsequently productivity. For example, the Workplace Wellbeing Charter¹, which is an initiative for better workplace conditions in the UK, encourages employers to commit to the health of the people who work for them by signing up to any of the *Public Health Responsibility Deal* pledges concerning *Health at Work*.

Radical growth in smartphone technology has enabled large-scale sociology and psychological studies; e.g. in the ‘UBhave’ project personal mobile phone sensor data is unobtrusively collected and analyzed to help support digital behaviour change interventions [3]. Moving a step forward, advances in wrist-worn technology have also found applications in interdisciplinary domains. For instance, a smartwatch application has recently been proposed to automatically detect eating

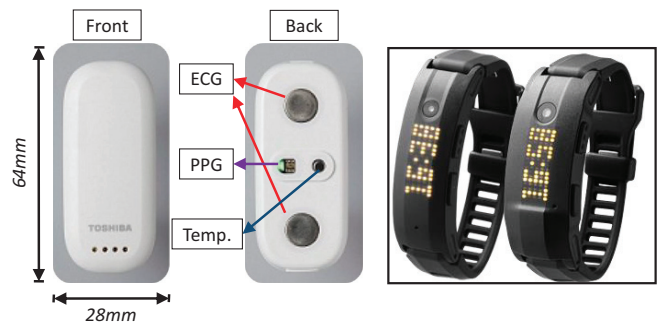


Fig. 1: Toshiba Silmee™ Bar Type (left) and Toshiba Silmee™ W20/W21 wristbands (Right).

episodes and assist in monitoring an individual’s diet [4]. Wrist-worn technology has further been used to detect self-harming activities [5], motion-based authentication [6] and recognizing social gestures [7]. Commercial smart wrist-bands are equipped with an increasing number of sensors including heart-rate monitors, skin temperature sensor and accelerometer; underpinning the motivation of this work to develop a mood recognition system on this basis.

In the current work we use the Toshiba Silmee™ Bar Type, a prototype wearable sensor [8], to collect physiological data in order to predict mood in the work environment. Toshiba Silmee™ Bar Type (shown in Figure 1) is a chest sensor that is able to measure heart rate from electrocardiogram (ECG) sensor, Pulse rate from photoplethysmogram (PPG) sensor, skin temperature and 3-axial acceleration [8]. The sensor is mounted to the chest using a special gel pad. Our framework will be extended with the use of the new Silmee™ W20/W21², which is a wristband sensor (Figure 1). Our main contributions in this paper are as follows:

- We introduce the HealthyOffice framework for learning mood recognition models and aggregating results in a privacy-preserving manner;
- We introduce the HealthyOffice smartphone application and a ground truth collection protocol used to furnish the physiological data collected from the Silmee™ wearables during our trials;

¹<http://www.wellbeingcharter.org.uk>

²http://www.toshiba.co.jp/healthcare/silmee_w20/

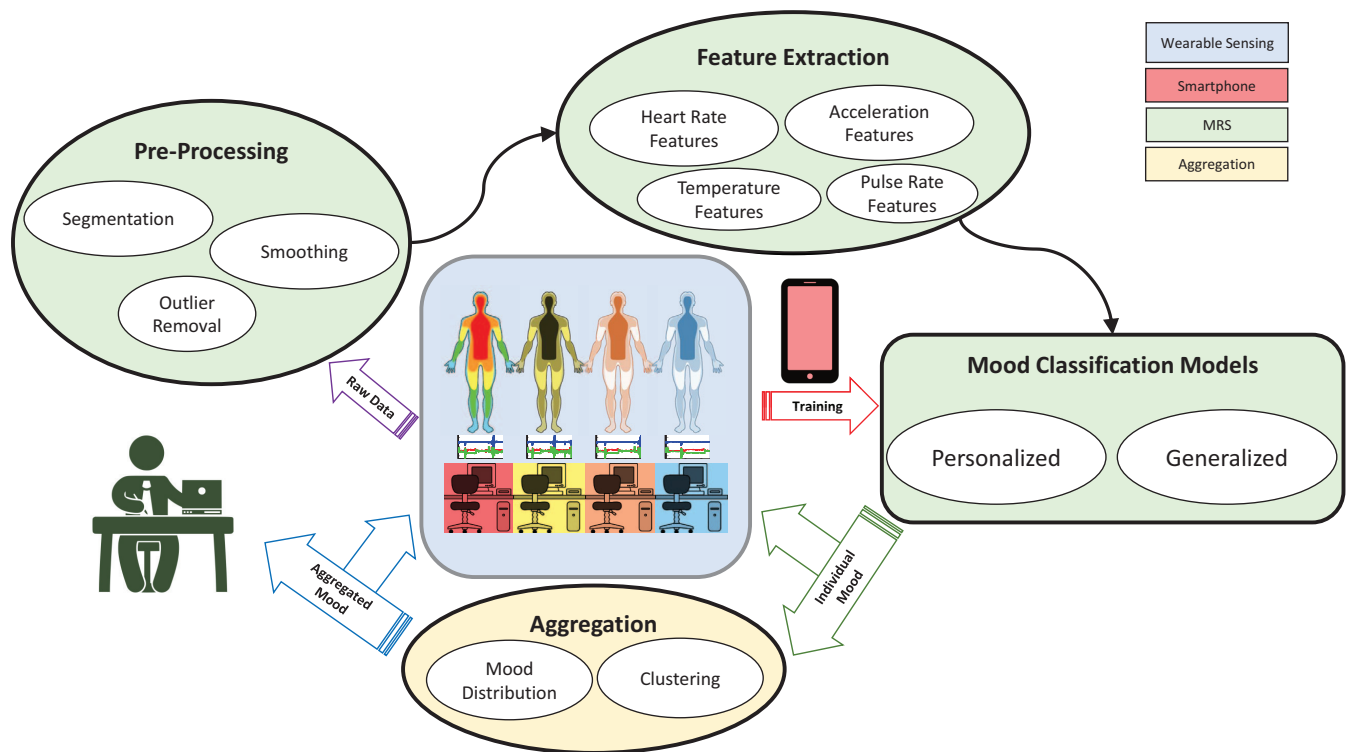


Fig. 2: Mood Recognition Framework.

- We develop separate mood models for personalized and generalized mood recognition and discuss their performance in a small-scale user study.
- We achieve an average classification accuracy of 70.6% over 8 moods and 5 mood categories, in 2 hour time intervals, and we make interpretations on the efficacy of our framework and a research direction.

II. RELATED WORK

Mood is a lasting affective state of mind which has been largely researched in psychology [9]. In our work, we focus on the *mood* rather than on *emotion* [9]: the mood of a person may last longer than emotion, persisting for hours or days rather than minutes or seconds. It can be described as a reaction to cumulative sequences of events while emotion is spontaneous. It is also more internal to a person while emotion is visible to others. Various studies have been conducted in recent times for understanding and tracking moods and emotions. For instance, EmotionSense [10] is a mobile sensing platform for sensing individual emotions using mobile phones. In particular, they analyzed participants’ voices and classified them in one of the following categories: Happy, Sad, Fear, Anger, and Neutral.

Moodscope [11] also infers mood based on how the mobile phone is used. In particular, six pieces of usage information was collected, namely, text messages, emails, phone calls, applications usage, web browsing, and location. The mood recognition system in [11] was able to predict the mood intensity (at five levels of granularity) of two dimensions (pleasure and activeness) on a Circumplex model, which can

be further mapped into mood classes. For example, a high level of pleasure combined with a high level of activeness corresponds to happiness while a low level of pleasure but a high level of activeness corresponds to stress.

Other researchers have focused only on one specific mood state; stress. Particularly, in [12] stress was assessed by means of a wrist-worn sensor of electrodermal activity (EDA) and acceleration (ACC), and mobile phone usage data such as messages, calls, location and screen activity (on/off). Similarly, stress recognition was the main focus in [13]–[16]. Researchers have also shown that stress could be assessed by skin temperature [17] and Heart Rate variability [18].

In this paper we are using existing wearable technology and methods to enable the unobtrusive collection of physiological data, focusing only on the work environment. We propose an advanced multi-classification model of 8 distinct moods and 5 levels of intensity, operating on 2-hour time windows. This is based on the 2-dimensional circumplex model of [11] expanded in to 8 separate dimensions of individual moods. The overall objective is to design and validate a system capable of capturing detailed user mood information using low-cost, pervasive sensor devices. This can be part of a solution addressing various industrial use cases such as the ‘HealthyOffice’ application presented in section IV.

III. MOOD RECOGNITION FRAMEWORK

Our proposed framework for mood recognition in a work environment is illustrated in Figure 2 and includes four main components:

- A wearable device like Silmee™ (Fig. 1). This is used to collect physiological data from employees, including heart rate, skin temperature and acceleration.
- A smartphone application, like HealthyOffice. This app is used to collect the ground truth for mood.
- A mood recognition system (MRS) that makes mood predictions given the bespoke sensor data and provide feedback to employees (via the HealthyOffice app).
- An aggregation component which produces anonymized aggregated results and statistics for the employer.

In our framework, employees (users) may wear smart wristbands and carry out their normal daily routines. These wristbands unobtrusively collect physiological data from the users. Users should ideally train the system by providing their mood via the HealthOffice app for a certain period of time. Otherwise, a generalized model may be applied, further detailed in Section V-D.

In order to prepare the raw sensor data for analysis, a pre-processing step involves a) filtering noisy data, b) removing outliers, and c) smoothing the physiological signals as necessary. Further, data collected for a period of time are segmented to reflect the intervals of the available ground truth. Then the feature extraction process is called to make meaningful representations of the processed signals. Finally, the system analyzes the data representations (using machine learning techniques) and it provides feedback to the users to promote awareness. An anonymized aggregation of mood predictions and statistics can be reported to the employer in order to take actions to improve office wellbeing.

The system uses machine learning techniques in two ways. First, to develop a ‘personalized model’ where each individual gets prediction about their mood based on their past physiological data, which model is learned from the personalized ground truth data points. The alternative way is to get predictions about one’s mood based on a ‘generalized model’, which is learned from physiological data and ground truth collected from multiple users.

IV. HEALTHYOFFICE APPLICATION

The ground truth for our study was collected with the use of a smartphone mood self-reporting app, called HealthyOffice [19], which was originally created for the collaborative EU project IES Cities [20] and subsequently extended to accommodate the present work. HealthyOffice records mood information using a slider for each mood, which takes values between 0 and 100; the larger the number, the higher the intensity of a particular mood. We divide this interval in 5 categories: 0–19 / 20–39 / 40–59 / 60–79 / 80–100, corresponding to: *Not at all* / *Not much* / *Slightly* / *Fairly* / *Extremely* and we instruct the users accordingly. During initial trials, the users would pick a value for all of the 8 moods at least 5 times a day, every two hours. The in-built reminder functionality of the app was set to timely notify the users to provide their next mood annotation input.

The HealthyOffice app input is uploaded to a remote database using the REST API of IES Cities citizen-centric

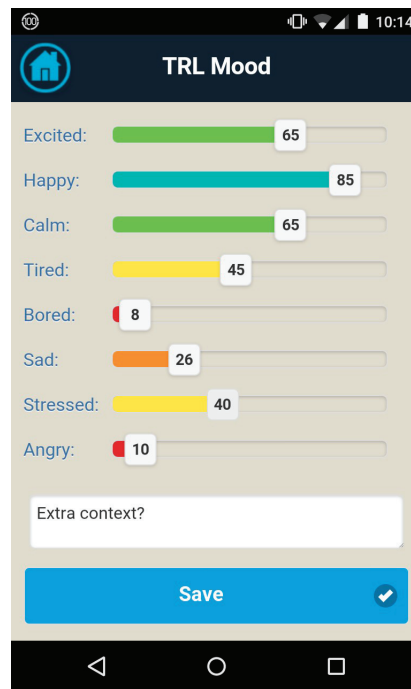


Fig. 3: Screenshot of the HealthyOffice application.

open data platform [21]. HealthyOffice further links to a secure database where wearable sensor data are stored and processed with the proposed MRS. The results of MRS, are also stored in the secure database, and may be retrieved by the HealthyOffice app and presented to the user in various forms, such as charts, graphics or even short textual messages. The most appropriate and engaging way to present mood information to the users is a research topic in its own right in the area of User eXperience (UX) and is still ongoing work for HealthyOffice.

V. METHODOLOGY

We provide implementation details of our MRS approach and we evaluate it by means of a small-scale user study in an office environment.

A. Data Collection Protocol

In this study, we collected biophysical user data using Toshiba Silmee™ wearable device in conjunction with the HealthyOffice app, introduced in Section IV. We recruited 4 users (researchers) to take part in this study which was conducted in an office environment. Researchers were provided with the sensors to wear during office hours and instructions were provided for recording ground truth mood data every two hours. The work of each researcher was not affected by this experiment. Participants were instructed to choose how they felt for each one of the 8 moods. Specifically, we were interested in Excitement, Happiness, Calmness, Tiredness, Boredom, Sadness, Stress and Anger. The question that was posed for each one of the moods was the following: *How have you been feeling for the last 2 hours?* The study ran from Monday to Friday during working hours (9:00–17:00) with a

total duration of 11 working days, resulting in 44 segments per user. In total, we collected 352 hours of physiological data.

B. Data Pre-processing and Feature Extraction

Annotations every two hours were not problem-free. For example, some participants missed some annotations and others had delayed recordings. For this reason, we allowed a tolerance window of an hour so that any annotation recorded with at most one hour delay from a pre-defined annotation point, would be matched to that point. We consider that this adjustment does not compromise the ground truth. Users are assumed to only reflect their mood in the predetermined time intervals regardless of the tolerated delay; any potential mood change during the delay period should be accounted by users for the subsequent time interval. More advanced techniques such as [22] for automatic annotation correction can also be deployed at this stage.

Feature extraction for the heart-rate variability was performed using the HRVAS toolbox [23] and acceleration features were calculated based on the standard Human Activity Recognition (HAR) features detailed in [24] and [25]. Pulse Wave Transit Time (PWTT) is defined for each beat as the difference in time between a cardiac wave peak and the next pulse wave valley. PWTT is essentially a byproduct of the ECG and PPG signals, which has been correlated with the pulse amplitude and, subsequently, with stress [8].

Since no other study has verified which features are best to predict mood, we start with an over-complete set of features. Starting with the heart rate, this was represented by inter-beat intervals (IBI) which are the intervals between two successive heart beats. IBI was obtained immediately from the SilmeeTM sensor. We analyzed the IBI signals in the time domain, the frequency domain and in the time-frequency domain.

In particular, for IBI's time domain features we calculated both statistical and geometrical features as recommended from the Guidelines of standards of measurement, physiological interpretation, and clinical use [26]. Statistical features include mean inter-beat interval, standard deviation of inter-beat interval (SDNN), root mean square of successive differences of the inter-beat intervals (RMSSD), and the percentage of total intervals that successively differ by more than 50 milliseconds (pNN50). Geometric features include the heart rate variability index (HRVi) which is the number of inter-beat intervals divided by the height of the histogram of all IBIs. Also, we perform triangular interpolation on the IBI's histogram and calculate the difference in milliseconds between the two angles of the base of the triangle (TINN).

With regards to frequency analysis, we quantify fluctuations of the IBI times series by calculating the power spectral density (PSD), i.e., the distribution of variance as a function of frequency. To do so we used three approaches, namely, Welch, Lomb-Scargle periodogram and Autoregression, which are commonly used for frequency analysis for heart rate variability [23]. Then PSD is divided in to 3 bands: very low frequency (VLF), 0.0033–0.04 Hz; low frequency (LF), 0.04–0.15 Hz; and high frequency (HF), 0.15–0.4 Hz. In this work

TABLE I: Number of Features Extracted from 5 Modalities and Related Sampling Rates.

Modality	No. of Features	Sampling Rate (Hz)
Heart Rate	120	125Hz (ECG)
Pulse Rate	9	62.5Hz (PPG)
PWTT	5	N/A
Accelerometer	137	15.625Hz
Skin Temperature	10	0.1Hz

we used frequency domain features based on the normalized values of LF, normalized values of HF and LF/HF ratio.

Frequency domain features are dealing with signal power distributed only in the frequency domain. The signal however can be also be examined in the time-frequency domain. Similarly to frequency-domain analysis, time-frequency analysis quantifies VLF, LF, and HF related measures. However, to achieve this analysis we use windowed versions of Lomb-Scargle periodogram, Autoregression and wavelet transforms which are explained in [27].

We note that heart rate signal may contain non-linear components. So, we further perform non-linear analysis to extract such features. Concretely, we capture sample entropy (sampen) which quantifies the signal's complexity. Also, we perform Detrended Fluctuation Analysis (DFA) based on the concept that a system or a shape can be decomposed into smaller pieces where each piece resembles one another but on different scales. So, DFA analysis finds self-similar properties in non-stationary time series. Another non-linear component comprises the coefficients of Poincare plot (SD1, SD2, SD1/SD2). Poincare plot is a plot of IBIs versus the previous IBIs which captures the self-similarities in the signal. It assumes that the current IBI is affected by the previous one. An ellipse is fitted with its centre at the mean of the data. SD1 is the standard deviation over the minor axis of the ellipse and SD2 the standard deviation over the major axis.

C. Classification Mechanisms

In this paper, we use 3 standard classifiers to automatically recognize 8 recorded moods namely, *i*) k -Nearest Neighbour (k -NN), *ii*) Decision Tree (DT) and *iii*) Ensemble approach (in particular Bagged Ensembles of Decision Trees (BE-DT)) [28]. As detailed in Section IV, all of these 8 moods are discretized into 5 categories in equal proportions. Categorical mood data is used as ground truth for evaluating the mood recognition framework. We benchmark these classifiers against two baselines, namely, personalized baseline (P-BL) and generalized baseline (G-BL). The P-BL is calculated by taking the most frequent class or intensity of each mood for each individual and dividing the number of occurrences of that class with the total number of each individual's data-points. The G-BL is calculated by taking the most frequent class among all the users, count the occurrences of that class and divide with the total number of all data-points.

D. Experimental Results

We evaluate the MRS using two approaches.

TABLE II: Classification Results in Accuracy for 8 Individual Moods Using Personalized and Generalized approaches. BLD is the Difference Between Baselines (P-BL or G-BL) and the Best Algorithm (in Bold).

Mood	Personalised					Generalized				
	kNN	DT	BE-DT	P-BL	BLD (rank)	kNN	DT	BE-DT	G-BL	BLD (rank)
Excited	0.6420	0.5625	0.6648	0.4986	0.1662 (3)	0.4886	0.5284	0.5114	0.4414	0.0870 (7)
Happy	0.6818	0.6477	0.7500	0.5795	0.1705 (2)	0.5795	0.5511	0.6477	0.5856	0.0621 (8)
Calm	0.5455	0.6080	0.6023	0.5493	0.0587 (7)	0.5341	0.5057	0.5398	0.4414	0.0984 (6)
Tired	0.6080	0.5511	0.6875	0.5095	0.1780 (1)	0.5000	0.5057	0.4886	0.3694	0.1363 (3)
Bored	0.6023	0.6420	0.6534	0.5014	0.1520 (4)	0.5909	0.5682	0.5909	0.4054	0.1855 (2)
Sad	0.7670	0.6818	0.7727	0.6551	0.1176 (6)	0.6705	0.7159	0.7670	0.5676	0.1994 (1)
Stressed	0.6705	0.6193	0.6989	0.5583	0.1406 (5)	0.5852	0.5625	0.6307	0.4955	0.1352 (4)
Angry	0.7727	0.7727	0.8182	0.7925	0.0257 (8)	0.7273	0.6648	0.7955	0.6757	0.1198 (5)
Average	0.6612	0.6356	0.7060	0.5805	0.1255	0.5845	0.5753	0.6214	0.4977	0.1237

a) *Personalised Models*: 8 mood models are separately trained for all 4 users individually resulting in a total of 32 models. These models are then evaluated using a leave-one-out cross-validation protocol. Table II shows the average classification accuracy using various algorithms.

b) *Generalized Models*: 8 mood models are trained for all users combined. All of these models are separately evaluated using a leave-one-participant-out cross-validation approach. Since we are testing for generalization, a random k -fold cross-validation approach is therefore not appropriate in this setting as some folds might contain samples from the same user in both train and test sets. Table II shows the average accuracy for all moods using various classification algorithms.

Our results in Table II show that all models perform better than the baseline which indicates the potential of our models to predict the mood of different users. However, personalized models work better than generalized ones on average. Also, the Bagged Ensembles of Decision Trees classifier has a higher accuracy compared with the rest of the algorithms in both the personalized and generalized models. The most predictable mood, in terms of classification accuracy of the personalized model, is Anger followed by Sadness, Happiness, Stress, Tiredness, Boredom and the least is Calmness. However, the most predictable in terms of accuracy difference from the baseline is Tiredness followed by Happiness, Excitement, Boredom, Stress, Sadness, Calmness and the least is Anger. Concerning the generalized model, Anger is the most predictable as well followed by Sadness, Happiness, Stress, Boredom, Calmness, Excitement and Tiredness. However, the most predictable in terms of accuracy difference from the baseline is Sadness, Boredom, Tiredness, Stress, Anger, Calmness, Excitement and Happiness. These results suggest that different people perceive their moods in a relatively different manner. For example, Happiness is very predictable in the individual models but not so much in the generalized case. The opposite happens with sadness where the generalized models perform better compared with the personalized models.

Even though our accuracies are always higher than the baseline a larger study is required to confirm the significance of our results. Since we are dealing with the perceived intensity of each mood, each individual had their own internal model of annotating the ground truth which is evident in the overall

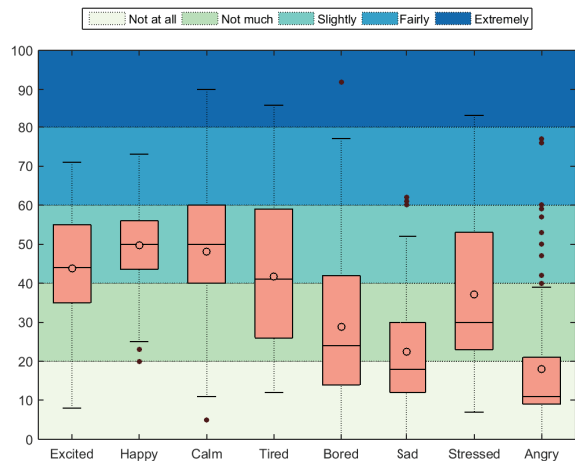


Fig. 4: Summarized mood results in a box plot for the employer.

predictability of the Generalized models.

Figure 4 shows the average self-reported moods for all participants over the course of the study. However, in practice, instead of using the ground truth to create this plot we would use the predictions over a given interval. The employer can use this information to understand the general feeling of the work-environment at any given time without explicitly asking any employees. Based on this information, the employer can take decisions to increase positive (e.g. happiness) and reduce the negative moods of the employees (e.g. stress and tiredness).

VI. CONCLUSION AND FUTURE WORK

In this paper we presented our vision for a fully automated mood recognition system that is applicable in the work environment to benefit employees' health and productivity. We made use of the SilmeeTM device to capture physiological and accelerometer data to recognize 8 different types of perceived moods every 2 hours, at 5 intensity levels.

Our initial results are promising as our classifiers perform better than the baseline of always choosing the majority class. This motivates us to continue our research work with a large-scale user trial to verify and improve our results. A larger dataset will effectively provide more representative mood data at all intensities thereby mitigating problems associated with classification overfitting. Also, we will experiment with more

advanced machine learning algorithms, including regression, and focus on feature engineering, feature extraction and feature selection to improve the system's overall performance. Finally, we will focus on the feedback to the employees and the employer ensuring that individuals' privacy is maintained. In this study we did not consider sensor data from users' smartphone which can also be used for mood prediction. Furthermore, we plan to investigate whether environmental factors play any role in employees mood.

For a model building perspective, collecting ground truth at 2 hour intervals may be challenging. To address this, we are also interested in experimenting with hybrid models i.e., use a generalized model to make predictions until enough data points are collected for each individual followed by personalizing individual models. Additionally, we want to embed high level activities such as taking breaks, making coffee and working in our system in order to enhance its predictive capabilities; as such employ contextual information to improve these models [29]. These can also be analyzed in a sequential manner by building rule hierarchies and establishing personalized hierarchical mood models [30]. A potential future avenue would be to investigate the commuting patterns of individuals and how they affect their mood in the working environment. It would also be interesting to discover which features predict (or possibly correlate with) which mood and provide personalized feedback based on them. For example, if features from accelerometer are selected for happiness, whilst also utilizing some high level activity recognition output, we can infer whether specific activities or the lack of activities result in a specific mental state.

The applicability of the introduced mood prediction framework in different work environments, using either smartphones or wearable devices (or both), and the relations between different moods are further interesting research questions.

ACKNOWLEDGMENT

The authors would like to thank the users that participated in the trial, and the Directors of the Toshiba Telecommunications Research Laboratory for their support.

REFERENCES

- [1] J. Goh, J. Pfeffer, and S. Zenios, "Exposure to harmful workplace practices could account for inequality in life spans across different demographic groups," *Health Affairs*, vol. 34, no. 10, pp. 1761–1768, 2015.
- [2] J. Hassard, K. Teoh, T. Cox, P. Dewe, M. Cosmar, R. Grundler, D. Flemming, B. Cosemans, and K. Van den Broek, "Calculating the cost of work-related stress and psychosocial risks." Luxembourg: Publications Office of the European Union, 2014.
- [3] N. Lathia, V. Pejovic, K. Rachuri, C. Mascolo, M. Musolesi, and P. Rentfrow, "Smartphones for large-scale behavior change interventions," *Pervasive Computing, IEEE*, vol. 12, no. 3, pp. 66–73, July 2013.
- [4] S. Sen, V. Subbaraju, A. Misra, R. Balan, and Y. Lee, "The case for smartwatch-based diet monitoring," in *IEEE PerCom Workshops*, 2015, pp. 585–590.
- [5] L. Malott, P. Bharti, N. Hilbert, G. Gopalakrishna, and S. Chellappan, "Detecting self-harming activities with wearable devices," in *IEEE PerCom Workshops*, 2015, pp. 597–602.
- [6] J. Yang, Y. Li, and M. Xie, "Motionauth: Motion-based authentication for wrist worn smart devices," in *IEEE PerCom Workshops*, 2015, pp. 550–555.

- [7] J. Knighten, S. McMillan, T. Chambers, and J. Payton, "Recognizing social gestures with a wrist-worn smartband," in *IEEE PerCom Workshops*, 2015, pp. 544–549.
- [8] S. Fuke, T. Suzuki, K. Nakayama, H. Tanaka, and S. Minami, "Blood pressure estimation from pulse wave velocity measured on the chest," in *Engineering in Medicine and Biology Society (EMBC), International Conference of the IEEE*, July 2013, pp. 6107–6110.
- [9] C. Beedie, P. Terry, and A. Lane, "Distinctions between emotion and mood," *Cognition and Emotion*, vol. 19, no. 6, pp. 847–878, 2005.
- [10] K. K. Rachuri, M. Musolesi, C. Mascolo, P. J. Rentfrow, C. Longworth, and A. Aucinas, "Emotionsense: A mobile phones based adaptive platform for experimental social psychology research," in *Proc. ACM Conference on Ubiquitous Computing*, 2010, pp. 281–290.
- [11] R. LiKamWa, Y. Liu, N. D. Lane, and L. Zhong, "Moodscope: Building a mood sensor from smartphone usage patterns," pp. 389–402, 2013.
- [12] A. Sano and R. W. Picard, "Stress recognition using wearable sensors and mobile phones," in *Proceedings of the Humaine Association Conference on Affective Computing and Intelligent Interaction*. IEEE Computer Society, 2013, pp. 671–676.
- [13] A. Muaremi, B. Arnrich, and G. Tröster, "Towards measuring stress with smartphones and wearable devices during workday and sleep," *Springer BioNanoScience Journal*, vol. 3, no. 2, pp. 172–183, 2013.
- [14] E. Ceja, V. Osmani, and O. Mayora, "Automatic stress detection in working environments from smartphones' accelerometer data: A first step," *Biomedical and Health Informatics, IEEE Journal of*, 2015.
- [15] H. Lu, D. Frauendorfer, M. Rabbi, M. S. Mast, G. T. Chittaranjan, A. T. Campbell, D. Gatica-Perez, and T. Choudhury, "Stressense: Detecting stress in unconstrained acoustic environments using smartphones," in *Proc. ACM Conference on Ubiquitous Computing*, 2012, pp. 351–360.
- [16] M. Gjoreski, H. Gjoreski, M. Lutrek, and M. Gams, "Automatic detection of perceived stress in campus students using smartphones," in *Intelligent Environments (IE), Int. Conf. on*, July 2015, pp. 132–135.
- [17] H. Kataoka, H. Kano, H. Yoshida, A. Saijo, M. Yasuda, and M. Osumi, "Development of a skin temperature measuring system for non-contact stress evaluation," in *Engineering in Medicine and Biology Society, IEEE*, vol. 2, Oct 1998, pp. 940–943.
- [18] G. G. Berntson and J. T. Cacioppo, *Heart Rate Variability: Stress and Psychiatric Conditions*. Blackwell Publishing, 2007, pp. 57–64.
- [19] "Bristol HealthyOffice for IES Cities (2015)," <https://play.google.com/store/apps/details?id=eu.iescities.HealthyOffice>, mobile application software for Android.
- [20] "IES Cities – Internet Enabled Services for the Cities across Europe," <http://iescities.eu>, last visit: November 2015.
- [21] "IES Cities server, software documentation." <https://iescities.com/IESCities/manual/index.html>, last visit: November 2015.
- [22] R. Kirkham, A. Khan, S. Bhattacharya, N. Hammerla, S. Mellor, D. Roggen, and T. Ploetz, "Automatic correction of annotation boundaries in activity datasets by class separation maximization," in *Proc. Pervasive and ubiquitous computing adjunct publication*. ACM, 2013.
- [23] J. T. Ramshur, "Design, evaluation, and application of heart rate variability analysis software (HRVAS)," MSc thesis, Univ. of Memphis, 2010.
- [24] N. Y. Hammerla, R. Kirkham, P. Andras, and T. Ploetz, "On preserving statistical characteristics of accelerometry data using their empirical cumulative distribution," in *Proc. International Symposium on Wearable Computers (ISWC)*. ACM, 2013, pp. 65–68.
- [25] A. Bulling, U. Blanke, and B. Schiele, "A tutorial on human activity recognition using body-worn inertial sensors," *ACM Comput. Surv.*, vol. 46, no. 3, pp. 33:1–33:33, Jan. 2014.
- [26] T. F. of the European Society of Cardiology *et al.*, "Heart rate variability standards of measurement, physiological interpretation, and clinical use," *Eur Heart J*, vol. 17, pp. 354–381, 1996.
- [27] U. R. Acharya, K. P. Joseph, N. Kannathal, C. M. Lim, and J. S. Suri, "Heart rate variability: a review," *Medical and biological engineering and computing*, vol. 44, no. 12, pp. 1031–1051, 2006.
- [28] C. M. Bishop, *Pattern Recognition and Machine Learning (Information Science and Statistics)*. Springer-Verlag New York, Inc., 2006.
- [29] A. Khan, J. Nicholson, S. Mellor, D. Jackson, K. Ladha, C. Ladha, J. Hand, J. Clarke, P. Olivier, and T. Plötz, "Occupancy monitoring using environmental & context sensors and a hierarchical analysis framework," in *Proc. BuildSys*. ACM, 2014.
- [30] A. Khan, D. Windridge, and J. Kittler, "Multilevel chinese takeaway process and label-based processes for rule induction in the context of automated sports video annotation," *IEEE Transactions on Cybernetics*, vol. 44, no. 10, pp. 1910–1923, Oct 2014.