Eustress or Distress: An Empirical Study of Perceived Stress in Everyday College Life

Chun-Tung Li
Department of Computing
The Hong Kong Polytechnic University
Hong Kong, China
csctli@comp.polyu.edu.hk

Jiannong Cao
Department of Computing
The Hong Kong Polytechnic University
Hong Kong, China
csjcao@comp.polyu.edu.hk

Tim M. H. Li
Department of Paediatrics & Adolescent Medicine
The University of Hong Kong
Hong Kong, China
timlmh@hku.hk

Ubicomp/ISWC'16 Adjunct , September 12-16, 2016, Heidelberg, Germany
ACM 978-1-4503-4462-3/16/09.
http://dx.doi.org/10.1145/2968219.2968309

Abstract
Eustress is literally the "good stress" that associated with positive feelings and health benefits. Previous studies focused on general stress, where the concept of eustress has been overlooked. This paper presents a novel approach towards stress recognition using data collected from wearable sensors, smartphones, and computers. The main goal is to determine if behavioral factors can help differentiate eustress from another kind of stress. We conducted a natural experiment to collect user smartphone and computer usage, heart rate and survey data in situ. By correlation and principle component analysis, a set of features could then be constructed. The performance was evaluated under leave-one-subject-out cross-validation, where the combined behavioral and physiological features enabled us to achieve 84.85% accuracy for general stress, 71.33% one kind of eustress as an urge for better performance, and 57.34% for eustress as a state of better mood. This work provided an encouraging result as an initial study for measuring eustress.

Author Keywords
Eustress, Stress, mHealth, Ubiquitous Computing

ACM Classification Keywords
H.1.2. [User/Machine Systems Subjects: Software psychology]: Miscellaneous.

Introduction

Stress as one of the major attribute to mental health has received growing interest from both industry and academia. Numerous studies suggest that stress is a health crisis, which associated with several diseases such as cardiovascular diseases, anxiety, and depression. A recent survey found that about half of the Americans experienced major stressful events in the last year [11]. Many of them reported they suffer from stress-related behavioral responses including lack of sleep, losing appetite and desire to exercise. Nowadays, the term stress is generally referring to negative stress (distress) in our daily conversation. The adverse impact of stress has been studied extensively, whereas the positive aspect of stress has also attracted rising attention. For example, the business and management community aims at maximizing individual productivity by managing work stress. However, the concept of positive stress (eustress) is incomplete. Lacking of knowledge about eustress obstructed the development of positive stress.

Typically, stress was assessed through questionnaires or clinical assessment by a psychiatrist. In the last two decades, researchers tried to measure stress through physiological marker including heart rate, blood pressure, galvanic skin response, etc. The result of these methods is promising in a rigorous laboratory environment, however, not applicable to detect stress in daily life. Moreover, the concept of eustress has been overlooked in the past decades. In the light of advanced mobile and wearable technology, data can be collected ubiquitously and less obtrusively, that enabled continuous stress assessment using ubiquitous sensing technology. To address these problems, we conducted a natural experiment and evaluated the classification result on the features extracted. We showed that ubiquitous computing is a potential method for evaluating eustress.

Background

The word stress was coined by Selye back in 1965, who defined stress as "the non-specific responses of the body to any demand for change" [13]. In general, it refers to the physiological responses caused by any stressful event (stressor). These responses are triggered by the Autonomic Nervous System (ANS), which influence internal organs and regulating heart rate, respiratory rate, blood vessel, galvanic skin response, and so on. ANS is divided into two subsystems, namely Sympathetic Nervous System (SNS) and Parasympathetic Nervous System (PNS). When stressful event arises, higher activity rate in SNS, which signals the adrenal glands to release stress hormones (e.g. adrenaline and cortisol). These hormones led to physiological changes, also known as the "*fight or flight*" response. Alternatively, activity in PNS increases during the restful event.

Selye introduced the concept of positive stress, namely eustress in 1974 [14]. He extended his work in stress to distinguish eustress and distress in terms of adaptiveness toward stress response, where eustress is "healthy, positive, constructive results of stressful events and stress response" [8]. Lazarus considers eustress as a positive cognitive response to a stressor, which associated with positive feelings and a healthy physical state [9].

Another dominating approach for understanding eustress was developed on the Yerkes-Dodson Law [3]. It suggests that stress is beneficial to performance until some optimal level is reached, after which performance will decline, which follow the inverted U shape diagram.

Related Work

Owing to the unclear criteria to distinguish eustress from others, existing analysis focused on general stress. Various stress measurement methods using computational technology have been proposed in the last two decades [15]. These methods can be classified into two categories: *physiological measures* and *physical measures*. The former one evaluate mental stress by monitoring different physiological responses including skin conductivity, heart activity, brain activity, blood pressure, etc. The later one collect physical characteristics (e.g. body gesture, facial expression, voice, etc.) that are sensitive to stress, and using machine learning methods to develop a computational model for stress recognition. Among all different types of input, Sharma & Gedeom suggested that heart rate variability (HRV) rank the top among different primary measure for assessing mental stress in terms of accuracy and non-intrusiveness.

Sun et al. consider stress assessment as a detection problem, which takes accelerometer data into account to filter the effect of motion artifact [6]. In [12], the authors collected data from wearable sensors and mobile phone in situ, which accuracies range from 75-87.5% for 2-class classification problem with the different feature set. Their work was extended in [5] with a larger population and longer period, which achieving classification accuracies range from 67-92%, showing that behavioral features are possible to recognize mental stress on a daily basis.

Existing methods investigated the pattern of physical and physiological sensory data under general stress. In our work, we study the feasibility of measuring eustress by HRV, smartphone and computer usage data. To the best of our knowledge, we are the first who proposing classification model toward eustress.

Research Questions

On the basis of previous work, general stress can be recognized by physiological signal with high accuracy, and suggested that stress is related to a number of behavioral factors such as multitasking, application usages, and physical activities. Recall one of the explanations of eustress regarding performance, multitasking lead to the task-switching cost which associated with a decrement in performance [10]. It is obvious that smartphone and computer use are the major sources of interruption, and closely related to multitasking. Therefore, we designed the experiment to investigate the possibility of using physiological and behavioral signal together to build an accurate classifier of eustress recognition. Since there has no single domination definition towards eustress, we assess eustress in twofold: 1) Higher self-reported performance along with moderate stress level [3]; 2) Higher self-reported mood along with moderate stress level [9].

It is not difficult to realize that too much or too few stress might not trigger "eustress" in terms of the definitions mentioned in the previous section. Therefore, we assume eustress must be under a moderate physiological stress level. To answer these questions, we study the pattern whether these behavioral features is able to correlate to these situations.

Study Protocol

We designed the in situ study and recruited 7 physically healthy subjects (5 males and 2 females) with ages ranging from 22 to 28, in which all of them are either research students and staffs. We collected data from each participant on 5 days during their waking hours. During the study, three sources of data are collected from (1) sensor and application on smartphone, (2) application on the personal computer, and (3) wearable heart rate

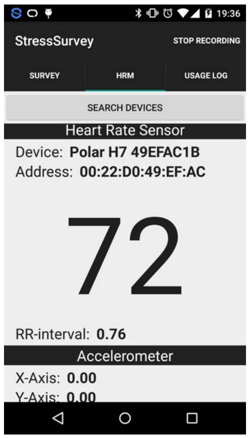

Figure 1: Control Panel for Heart Rate Measurement.

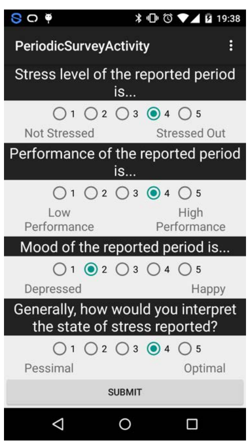

Figure 2: An Example of Periodic Survey.

sensor. These data can be categorized into heart rate, usage, and survey measure respectively.

We developed StressSurvey, which is an application for Android smartphone to collect smartphone activities and other sensory data. It connects the heart rate sensor automatically in the background, and recording heart rate data transmitted. It also captures smartphone screen and call activities. Every hour in between 8 AM to 12 PM, the application reminds the participant to report the survey by notification. The detail of data acquisition process is described in the following section.

Heart rate measure. Heart rate variability is collected using Polar H7 heart rate sensor [1], wearing with a chest band to record beat-to-beat interval and average heart rate. The heart rate data is measured by ECG sensor and preprocessed within the H7 device. Then it transmits the record in 1000ms via Bluetooth to the Android smartphone. Since the connection is using Bluetooth 4.0 (BLE), the smartphone is required at least Android version 4.3 with BLE enabled (e.g. Nexus 5, Galaxy S3). The data transmitted complies with the BLE specification, where the characteristic specified the format of the record. Each record is either 8 or 16 bit int format, indicated by the first bit of data (0 for 8-bit int, 1 for 16-bit int). Bit 1 and 2 indicate whether sensor contact feature supported and the sensor contact status. Bit 3 is the indicator for energy status that indicates if energy expended data is presented. Bit 4 indicate if RR-interval data is presented, and the interval is represented in 1/1024 sec. We shift the reading byte by checking the flag data. Each record is stored with UNIX timestamp on the smartphone in common separated values (CSV) format.

In order to eliminate the effect of heart rate due to the human artifact, motion data was collected along with heart rate measure, obtained from the accelerometer on the android smartphone. Each motion data contains a three-dimensional vector, which was calculated after removing the influence of the force of gravity.

Smartphone and computer usage measure. The usage log is collected via commercial application RescueTime [2]. Participants are asked to install the RescueTime client application on both computer and smartphone, each of them is assigned to seven prepared user accounts: hkpu.stresssurvey.#@gmail.com where # is an integer id from one to seven. Data can be downloaded through the public API, each row contains the timestamp, application name, category, duration, and estimated productive index ranging from -2 to 2. We collected the most fine-grained record in five minutes' interval for each participant. Screen on and off events and the state of smartphone call are collected directly by StressSurvey. Each record comes with an event indicator and timestamp and stored locally in CSV format.

Survey measure. This study using experience sampling method (ESM) to capture self-reported survey from time to time. During the day time, the application sends out the notification to remind participant to complete a survey every hour. The survey consists of several questions and provided an integer scale ranging from one to five, asking the perceived stress, performance, and mood. Participant completed the end-of-day survey rated the same scale according to the daily basis.

Data Overview

Over 7 participants, one was excluded from the analysis because the heart rate sensor was disconnected most of the time. We collected 5,058,233 accelerometer data, 1,410,109 heart rate data, 10,851 screen activity data, 878 call activity, 14,746 smartphone and computer usage and 252 self-reported survey data in raw format.

By removing incomplete data, there are 143 survey data combined with sensory, usage and survey data aggregated in hourly basis. Statistic of the reported survey is shown in table 1, where we found that each participant has their own preference of reporting their values.

Table 1: Statistic of each participant

Subj.	Survey	# of reports (1-5)					Total
	Stress	11	1	3	1	0	
1	Mood	1	1	3	3	8	16
	Performance	2	5	5	1	3	
	Stress	10	3	2	2	0	
3	Mood	0	1	6	8	2	17
	Performance	4	7	4	2	0	
	Stress	1	6	6	1	0	
4	Mood	0	1	7	6	0	14
	Performance	0	4	9	1	0	
	Stress	3	9	22	4	2	
5	Mood	0	5	19	14	2	40
	Performance	0	10	20	9	1	
	Stress	7	17	5	1	0	
6	Mood	0	2	17	11	0	30
	Performance	4	9	9	8	0	
	Stress	2	3	1	9	11	
7	Mood	9	13	1	3	0	26
	Performance	5	8	12	0	1	

In general, perceived stress is positively associated with performance and inversely for mood as shown in figure 3. The average use of smartphone and computer increase starting from 6AM to 11AM and reach the first peak in the morning. After which it slightly drop during 12PM to 1PM. The use of computer and smartphone at night decreased significantly.

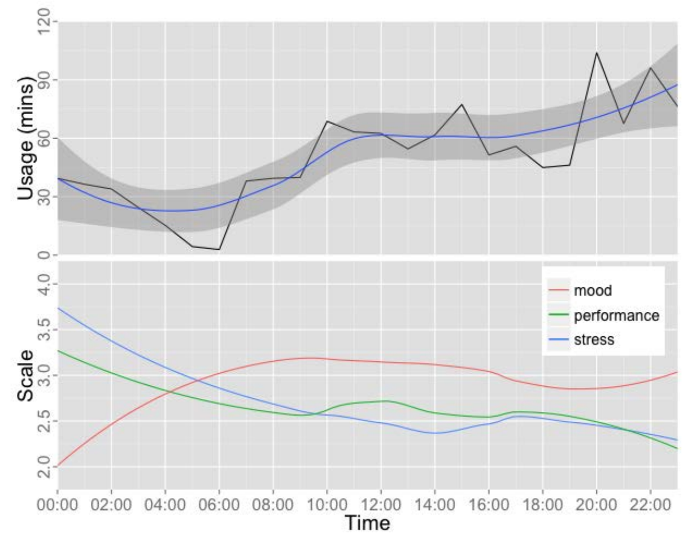

Figure 3: Average of inter-subject computer and smartphone usage (duration) and survey value.

Feature Extraction

Data especially heart rate measure requires cleaning and transformation prior to classification. First of all, we remove obvious error (e.g. heart rate < 40), and RR-interval that is more than 20% different from the previous one. Then, the value is interpolated by the moving average. The summary of features extracted is shown in table 2.

Heart Rate Measure

Heart rate measure (HRM) including average heart rate data and actual R-R interval obtained from the heart rate sensor. The average heart rate data were aggregating in 60-minute windows, in which standard deviation of heart rate (SDHR) and the average of heart rate (AVHR) were derived. Heart rate variability features can also be extracted from the windows including standard deviation of NN-interval (SDNN), average of NN-interval (AVNN), percentage of adjacent NN-intervals differing by more than 50ms (pNN50), and root-mean-square differences of successive R-R intervals (RMSSD). For frequency domain features, since the sampling rate deviate because of the system operation, and the number of samples is not necessary to be the product of two. Therefore, we employ the Lomb-Scargle Periodogram [7] that is capable of analyzing unevenly sampled time-series and data sets with missing values.

Table 2: Summary of extracted features

Modality	Features
Heart rate measure	AVHR, SDHR, AVNN, SDNN, RMSSD, PNN50, VLF, LF, HF, LF/HF
Motion	AVMI, SDMI
Screen	Duration of screen on time (secs), frequency of screen on event
Call	Number of call, answered call; Duration of off-hook
Application	Duration of each category: social, entertainment, internet, communication, study, email

Then the power spectrum obtained is sum up to three separate bin, grouped by very low frequency (VLF) < 0.04 Hz, low frequency (LF) 0.04 - 0.15 Hz and high frequency (HF) 0.15 - 0.4 Hz respectively. In addition, the accelerometer data was collected during heart rate measure is available. Then the motion intensity (MI) was defined by $\frac{1}{3}(|ACC_x| + |ACC_y| + |ACC_z|)$, where average and standard deviation of MI were calculated in the 60-minute windows aligned to the HRM features.

Smartphone and Computer Usage

Usage log including smartphone screen, call state, and application used are captured from smartphone and computer. For screen and call activities, duration and frequency are extracted from raw data. For the application usage, records are aggregated in hourly basis. Each record consists of the name of application, time of the usage recording, and duration of each application. Some other information such as category and estimated productivity provided by RescueTime were not used. The usage records are then labeled manually into the following categories: internet, email, social, communication, study, and entertainment. Then the sum of duration of application used from the same category were calculated. In order to eliminate the individual difference among different participants, the categorized data was used to derive three ratios namely: social, productive and non-productive ratio. Then we perform the dimension reduction by using Principle Component Analysis (PCA), to further eliminate linearly dependent features.

Classification Result

In this section, we present the process of training classifier and the result of different approaches. We use the R (programming language) to build various classifiers using well-known learning methods: Multinomial Logistic Regression (MLR), Support Vector Machine (SVM) and Random Forest (RF), to evaluate the predictive power of linear classifier, non-linear classifier, and ensemble classifier respectively.

First of all, we perform inter-subject z-score normalization

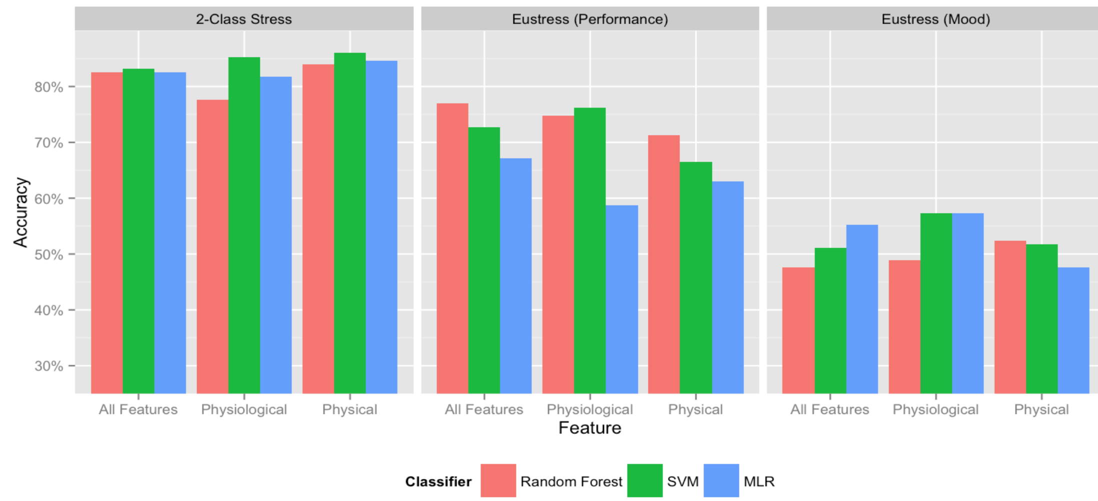

Figure 4: Classification Result

on the features in order to increase generality of the model. Then we calculate the correlation matrix to eliminate redundant features, which has coefficient greater than 0.75. Then the features were selected by exhaustive search with 10-fold cross-validation using Random Forest. Then we apply Synthetic Minority Over-Sampling Technique (SMOTE) [4] to the training data set to avoid over-fitting and deal with unbalanced data distribution.

For each classification problem, we partition all features into to two subsets of features: physiological features, physical features. We tested every problem with any set of features before and after dimension reduction using PCA. The performance was evaluated under leave-one-subject-out cross-validation. For each learning method, the model was built using repeated cross-validation. We also fine-tuning the parameters of the model using greedy approach in terms of accuracy.

General Stress Recognition
Prior to the eustress recognition, we tested our features on two-class general stress recognition with the above setting. Whereas the self-reported survey collected during the study was ranging perceived stress from one to five, then the value was normalized within subject and the class "stressed" is defined by z-score > 0 where the alternative is "not stressed".

On average, we achieved 82.75% accuracy and 96.93% recall for two-class stress recognition problem using all features by applying PCA; More specifically, the best result was obtained by Support Vector Machine with 83.22% accuracy and 97.9% recall. For physiological features alone the accuracy reached 81.59% and 96.27%

recall, where behavioral features obtained 84.85% accuracy and 99.03% recall. Our results show that we achieved competitive classification accuracy comparing to the state of the art.

Eustress Recognition

In this study, we have several assumptions: 1) eustress is the "right" amount of stress that improves performance [3]; 2) eustress associated with positive feeling. Therefore, we define eustress in twofold: Eustress is the combination of moderate stress with high performance, and eustress is the combination of moderate stress with high mood. We consider moderate stress as 1 standard deviation away from 0 (both positive and negative direction) for z-score normalized stress. Mood as a subjective measure as stress was applied the same normalization technique as stress, where the distribution of performance is more consistent over different subjects, we considered high performance strictly greater than 3.

For eustress in terms of perceived performance, the accuracy achieved 67.13% with recall only 42.75% using all PCA features. For eustress defined by perceived mood, the accuracy has only 55.25% and recall 56.22% using physiological PCA features. It shows that the highly unbalanced data result in a poor recall rate on eustress classification.

Discussion and Conclusion

Existing work studied general stress in both laboratory and natural environment. However, there are only a few works contributed to eustress since the concept has been proposed in the 70's. Our work studies the possibility of using ubiquitous sensing technologies for eustress recognition. We conducted a natural experiment and recruited 7 participants over 5 days. With an Android based application developed, heart rate and smartphone usage data were collected to constructed a set of features using correlation and principle component analysis. We estimated the robustness of the features by three standard learning algorithms.

The result showed that heart rate variability, computer and smartphone usage can be used for general stress classification as literature suggested. The recognition accuracy also remains consistent over different learning algorithms. On the other hand, the accuracy of eustress in terms of performance is higher than mood, since perceived performance is highly related to application used on smartphone and computer. However, the recall rates are low for both cases showing that the generality of the model still requires further study. The gap between general stress and eustress mainly due to the solid background of general stress that facilitated the feature engineering process and results in better classification performance.

Notice that the accuracy comparing to the existing work may seem quite low, however, is reasonable since the previous studies assess mental stress in rigorous laboratory or aggregated the data by days. In contrast, our natural experiment approach and finer granularity of time-series result in noisier data which leads to decrement of performance. We agreed that there is room for improvement, further study is required to achieved better recognition accuracy and recall rate.

To conclude, eustress as a widely accepted psychological phenomenon should receive more interest from the academia. As an initial study, our work provided encouraging result of eustress recognition, which can facilitate research on this problem in the near future.

Limitation

This work as a preliminary study of eustress has several limitations. Firstly, the sample size is limited to 6, where a larger scale study is required for further study. Secondly, self-report surveys are considered as ground truth in this work, where it may suffer from inconsistent between different subjects. Lastly, the concept of eustress is unclear, where a more accurate model can be achieved by introducing a more concrete definition of eustress.

References

[1] Polar h7 heart rate sensor. http://www.polar.com/hk-en/products/accessories/H7_heart_rate_sensor, 2016. Accessed: 2016-03-29.

[2] Rescuetime : Time management software for staying productive and happy in the modern workplace. http://www.rescuetime.com, 2016. Accessed: 2016-01-30.

[3] Benson, H., and Allen, R. How much stress is too much? *Harvard Business Review 58*, 5 (1979), 86–92.

[4] Chawala, N., Bowyer, K., Hall, L., and Kegelmeyer, W. Smote: Synthetic minority over-sampleing technique. *J. Artif. Intell. Res 16* (2002), 321–357.

[5] et al., A. S. Recognizing academic performance, sleep quality, stress level, and mental health using personality traits, wearable sensors and mobile phones. In *Wearable and Implantable Body Sensor Networks (BSN), 2015 IEEE 12th International Conference on*, IEEE (2015), 1–6.

[6] et al., F.-T. S. Activity-aware mental stress detection using physiological sensors. In *Mobile computing, applications, and services*. Springer, 2012, 211–230.

[7] İşler, Y., and Kuntalp, M. Combining classical hrv indices with wavelet entropy measures improves to performance in diagnosing congestive heart failure. *Computers in Biology and Medicine 37*, 10 (2007), 1502–1510.

[8] Kupriyanov, R., and Zhdanov, R. The eustress concept: Problems and outlooks. *World Journal of Medical Sciences 11*, 2 (2014), 179–185.

[9] Lazarus, R. From psychological stress to the emotions: A history of changing outlooks. *Personality: Critical Concepts in Psychology 4* (1998), 179.

[10] Meiran, N., Chorev, Z., and Sapir, A. Component processes in task switching. *Cognitive psychology 41*, 3 (2000), 211–253.

[11] NPR, Foundation, R. W. J., and of Public Health, H. S. The burden of stress in america, 2014.

[12] Sano, A., and Picard, R. Stress recognition using wearable sensors and mobile phones. In *Affective Computing and Intelligent Interaction, 2013 Humaine Association Conference on*, IEEE (2013), 671–676.

[13] Selye, H. The stress syndrome. *AJN The American Journal of Nursing 65*, 3 (1965), 97–99.

[14] Selye, H. Md stress without distress. *New York: The New Ameri-can Library* (1974).

[15] Sharma, N., and Gedeon, T. Objective measures, sensors and computational techniques for stress recognition and classification: A survey. *Computer methods and programs in biomedicine 108*, 3 (2012), 1287–1301.