

StressAware: An App for Real-Time Stress Monitoring on the Amulet Wearable Platform

George Boateng, David Kotz
Dartmouth College

Abstract—Stress is the root cause of many diseases and unhealthy behaviors. Being able to monitor when and why a person is stressed could inform personal stress management as well as interventions when necessary. In this work, we present StressAware, an application on the Amulet wearable platform that classifies the stress level (low, medium, high) of individuals continuously and in real time using heart rate (HR) and heart-rate variability (HRV) data from a commercial heart-rate monitor. We developed our stress-detection model using a Support Vector Machine (SVM). We trained and tested our model using data from three sources and had the following preliminary results: PhysioNet, a public physiological database (94.5% accurate with 10-fold cross validation), a field study (100% accurate with 10-fold cross validation) and a lab study (64.3% accurate with leave-one-out cross-validation). Testing the StressAware app revealed a projected battery life of up to 12 days. Also, the usability feedback from subjects showed that the Amulet has a potential to be used by people for monitoring their stress levels. The results are promising, indicating that the app may be used for stress detection, and eventually for the development of stress-related intervention that could improve the health of individuals.

Keywords—stress; machine learning; mobile health (mHealth); wearables; heart rate; heart-rate variability; support vector machine

I. INTRODUCTION

The American Medical Association has noted that stress is the underlying cause of more than 60 percent of all human illness and disease [1]. Stress can trigger onset or recurrence of addictive behaviors like unhealthy eating, smoking, or drug use. There is a need to measure stress, unobtrusively and continuously and in the field, because stress is so often tied to these challenging behaviors. Knowing when, where, and why a person is under stress can help health professionals develop mechanisms to intervene “in the moment,” in a way appropriate to the person and the condition, to help that person deal with the stress and avoid the unhealthy behavior or seek out healthy stress-reducing activities. Additionally, having more information about stress can help individuals manage their own stress levels.

In this work, we built StressAware, an application that continuously monitors the stress level of individuals in real time using data from a commercial heart-rate chest strap (such as the Zephyr [2]) and runs on the Amulet, a low-power wrist-worn device [3]. We developed an SVM-based machine-learning model to detect stress using data from PhysioNet, a public database with various physiological data, and two sets of studies approved by Dartmouth’s Institutional Review Board in which we collected heart rate (HR) and heart-rate variability (HRV) data – an in-lab study in which subjects performed various stress-inducing activities and a field study where volunteers wore the Amulet and Zephyr during waking hours. We then built an app for the Amulet that implements our stress-detection model, continuously streams data from the Zephyr, classifies the stress level of an individual, logs the stress level, and then displays it as a graph on the screen. In the remainder of this work, we focus on the StressAware machine-learning model, the components of the StressAware app, the energy efficiency results and usability feedback from the user study.

II. BACKGROUND

In this section, we describe the necessary background to understand the work described. First, we describe the Amulet platform on which the StressAware app runs and why it is suitable for running the app. We then describe the science of stress and its relation to stress measurement.

A. Amulet Wearable Device Platform

The Amulet is a hardware and software platform for writing energy and memory-efficient sensing applications that achieve long battery life [3]. The Amulet hardware is a wrist-worn device that has two microcontrollers: an MSP430 running applications, and an nRF51822 for communicating with peripheral Bluetooth Low Energy (BLE) devices such as a heart-rate monitor. It has several built in sensors: microphone, accelerometer, etc. The Amulet platform is hence useful for creating and running mHealth apps that need to continuously run for hours and days to monitor the physiological and behavioral health of its wearer.

B. Science of Stress

Stress is a physiological response to mental, emotional, or physical challenges we encounter [4]. When a person is stressed, the hypothalamus signals two systems in the body, namely, the Hypothalamic Pituitary-Adrenal system, which deals with the release of the stress hormone cortisol, whose levels could be used as an indicator of a person’s stress level [5], and the Sympathomedullary Pathway system, which deals with the two parts of autonomic nervous system: the sympathetic nervous system (SNS), which increases HR when under stress and the parasympathetic nervous system (PNS), which reduces HR when the stressful situation ends [6]. Since there is a cardiac response to stress, we may be able to infer whether a person is stressed by looking at their cardiac
activity, which can be captured electrically with an electrocardiogram (ECG) (Figure 1). The HR is the number of R peaks in a minute and RR interval (RRI) is the time interval between two R peaks. HRV, which is the variability in RRI, provides information about the relative activation between the SNS and PNS and may be used to distinguish the stress level of individuals [5].

III. RELATED WORK

Several research projects aimed to collect continuous measures of stress both in and outside the laboratory setting. A study conducted by Healey and Picard involved collecting and analyzing physiological data like ECG and skin conductance during real-world driving tasks to determine a driver’s relative stress level [7]. They developed a classifier that could distinguish between three stress levels – low stress, medium stress and high stress – with an accuracy of over 97% with leave-one-out cross-validation. Another study, conducted by Ertin et al., used a custom suite of wireless sensors called Autosense to measure ECG and galvanic skin response to infer the stress of subjects [5]. Using data from 20+ participants via a lab study and a field study, they obtained an accuracy of 90% on in-lab data. These studies show that physiological signals can be used to determine stress levels.

IV. SOLUTION: STRESSAWARE

Given the need to continuously measure and monitor the stress levels of individuals, we developed the StressAware app for the Amulet wearable device. StressAware monitors a person’s stress level on a scale of low, medium, and high, using data from a heart-rate monitor such as the Zephyr, logs that information, and then displays the stress level over the past hours as a bar graph on the Amulet screen. The app uses an implementation of a machine-learning model that is trained offline.

V. STRESS DETECTION MODEL - MACHINE LEARNING OFFLINE

SVM is a classifier that constructs a high-dimensional hyper-plane and uses it to perform classification [4]. SVM is recently popular for mining physiological data because of its ability to handle high-dimensional data using minimal training features [8]. We focus on using SVM because it uses a subset of the training set – support vectors – for its prediction function as compared to other models like k-nearest neighbor (KNN), which will need to store all the data points in memory for prediction. It is hence memory efficient and ideal for low-memory platform like the Amulet. We trained two SVM models, one a Linear SVM and the other a Radial Basis Function (RBF) SVM, using the scikit-learn library [9] to distinguish low, medium, and high stress levels.

A. Data Extraction I: PhysioNet

We obtained the data used for training the machine-learning model from PhysioNet, a public database with various physiological data [10]. We extracted HR and RRI from the data and grouped them into three stress levels resulting in training dataset containing 73 data points; see our technical report [11] for more information.

B. Data Extraction II: Field Study

We ran a field study from which we collected data from a total of 10 subjects. The participants wore the Amulet and Zephyr for one day, lasting between 8 and 12 hours. The Zephyr transmitted HR and RRI data to the Amulet throughout the day. The Amulet also recorded acceleration. The app on the Amulet logged 1 minute of data (and 5 minutes for some subjects) every 10 minutes. The app then prompted the subjects to answer two questions via the Ecological Momentary Assessment (EMA) component of the Amulet app. The app asked subjects to rate their stress levels and their activity levels between low, medium and high at the moment. There were 4 EMAs per hour and at least 32 EMAs per day. After we collected the data, we used data from four subjects to obtain a training dataset containing 120 data points.

C. Data Extraction III: Lab Study

We collected data from two subjects in the lab study. These participants were subjected to mild stressors for about 80 minutes that previous experiments have shown to induce stress [4]. During the lab study, each subject wore a Zephyr chest strap. The Zephyr transmitted HR and RRI data to the Amulet throughout the duration of the study. The subjects were periodically asked to rate their stress level between low, medium and high for that session, which we later used for the stress annotation. We built a training dataset containing 14 data points from the lab study.

D. Feature Extraction

From each subject’s data, we extracted 14 HR and HRV features that previous studies have shown to be relevant for stress detection. The features are detailed in our report [11]: HR features – mean_hr, std_hr, median_hr, percentile_20_hr, percentile_80_hr and HRV features – mean_rri, std_rri, rMSSD, NN50, pNN50, median_rri, max_rri, min_rri, percentile_80_rri [4][5].

E. Training/Classification

We trained two SVM models – Linear and RBF SVM – each of which classified the data into three stress levels: low, medium and high. We ran experiments to test the two classifiers.

F. Testing and Results I: PhysioNet Dataset

We experimented with different sets of HR and HRV features. To evaluate the importance of subsets of the feature set, we ran 10-fold cross validation with four sets of features: all features, only HR features, only RRI features, and features that represent an aggregate of several HR and RRI values.
(mean_hr, std_hr, mean_rri, std_rri, rMSSD, NN50, pNN50). The RRI feature set and HR feature set had the least accuracy (71.2% and 76.7% respectively) for RBF SVM. The “all features” set did better with 90.4% for RBF. The aggregate features did best with 94.5% for RBF. Overall, RBF did much better than Linear SVM. This result shows that it is best to use features that aggregate several HR and RRI values rather than features that are directly chosen from the HR and RRI values such as maximum, minimum, median and percentiles.

G. Testing and Results II: Field Study

We ran 10-fold cross-validation on the data from the field study using Linear and RBF SVM. We also tested the accuracy of the two models with and without acceleration data. To test the effect of including accelerometer data, we computed an additional feature for each window: mean of acceleration. We included this feature in the aggregate features set, which had produced the highest accuracy in previous experiments. We used data from 4 subjects (120 data points) and then ran 10-fold cross-validation. RBF had an accuracy of 89.2% without acceleration, which increased to 100% when acceleration is included. Since acceleration captures a person’s activity level, it is useful in distinguishing between an increase in HR stemming from increased activity and that stemming from stress. We found that including acceleration helps in accurately classifying the stress level of a person.

H. Testing and Results III: Lab Study

We ran leave-one-out cross-validation on the data from the lab study using Linear SVM and RBF SVM. We were unable to run 10-fold cross validation because we had only 14 data points. Linear SVM had an accuracy of 50% and RBF had an accuracy of 64.29%. RBF once again performed better than Linear SVM. Because of the limited amount of data, we could not infer much from the results.

VI. STRESSAWARE - APP ON AMULET

The StressAware app consists of five components: EMA, Data Collector, Stress Detector, Stress Level Graph, and Data Logger.

A. EMA

The EMA component of the app is responsible for intermittently asking the user about her/his stress and activity level (Figure 2). The EMA results are used as ground truth of a person’s stress level. This component of the app is only used for data collection during the user study.

B. Data Collector

The data collector is responsible for getting 5 minutes and 1 minute worth of HR and RRI data from the Zephyr (Figure 2). It also collects acceleration data from the Amulet. The data is used by the stress detection model and the data logger.

C. Stress Detector

The stress detector determines the stress level of the user. We compute the feature vector from the 60-second HR and RRI data obtained by the data collector. Then, we scale the feature vector using the scaling factors from the trained model. The stress classifier then uses the feature vector for the prediction. The stress classifier is an implementation of the prediction equation of a Linear SVM. The equation is:

\[ y = wx + b \]  

Here, \( y \) is the vector that holds the result of the evaluation for the three stress classes, \( x \) is the computed feature vector, \( w \) is the coefficient matrix and \( b \) is the intercept vector. The values for \( w \) and \( b \) are obtained from the linear model that was trained offline. Since this is a multi-label classification, we implemented the “one-vs-the-rest” approach for multi-label classification since the scikit learn Linear SVM function used this approach. In this approach, three classifiers are trained for each of the classes and the result of solving the equation is a vector containing a value for each of the three classes. The class with the maximum value is the predicted class.

D. Stress Level Graph

The stress-level graph displays the stress level of the user over the past two hours as a bar graph (Figure 2). This information could provide better insight to users about their stress pattern on a particular day.

E. Data Logger

The data logger logs the HR, RRI and acceleration to a file on the SD card. It also logs the stress level as indicated by the user. We use the logged data to develop the stress-detection model. This component of the app is also only used during the user study.

VII. Usability Feedback

After participating in the field study, each participant completed a survey to evaluate the usability of the Amulet in monitoring the stress level of individuals in the wild. All 10 participants (3 males, 7 females) filled the survey. Overall, participants thought the study was a good experience. Eighty percent (80%) of participants agreed that it was easy to answer the questions on the Amulet, which demonstrates the potential of using wrist-worn devices such as the Amulet for EMA studies. People enjoyed being able to see their HR and stress level, and 70% of participants mentioned that they would wear the Amulet and Zephyr if it automatically measured and monitored their stress level for several hours and days. The only concerns were about the bulkiness of the Amulet, the

Figure 2: Data collection (left), EMA questions (middle) and stress level graph (right)
frequent EMAs and disconnection of the Amulet from the Zephyr. The responses from the survey show that the Amulet and Zephyr have a potential to be used by people for stress monitoring.

VIII. ENERGY EFFICIENCY OF STRESSAWARE

We tested the energy efficiency of the StressAware app by running it for 8 hours as it computed stress levels every five minutes. We logged the battery voltage level over the 8-hour period. The graph of the battery level shows battery percentage as the y-axis and time (seconds) as the x-axis. The battery level dropped linearly from 100% to 98% over the 8-hour period indicating a 2% loss in battery life (Figure 3). However, after 8 hours, the battery level on the screen indicated a 92% battery level. This means that StressAware consumes 8% of battery life per 8 hours: 1% per hour. StressAware is expected to run for about 8 hours a day during which period a person is awake. This result means that the app is projected to run for about 2 weeks (12 days) before the Amulet needs to be recharged.

IX. LIMITATIONS AND FUTURE WORK

Our experiments have several limitations that imply the need for more work in the future. We used only time-based features and no frequency-based features for training the classifiers because they are computationally intensive to derive. The Zephyr is a commercial heart-rate monitor and hence the values used for computing the features would not be as accurate as those obtained from using a medical-grade device. Also, we implemented a Linear SVM model in the StressAware app rather than an RBF SVM model, which had the highest accuracy because RBF is more computationally intensive and requires more memory to store all the support vectors. And after developing the stress detection model, we did not test model with new data from subjects in the wild. In the future, computing frequency-based features, using medical-grade devices, implementing an RBF model in StressAware and testing on new data from the wild could lead to more accurate real-time predictions. We did not focus on collecting data that could be used to infer the causes of a person’s stress. In the future, tracking contextual information like location, noise level, sleep duration, etc., could be used to diagnose the cause of stress. Finally, our results are based on a small number of subjects and must be repeated with larger populations before final conclusions can be drawn.

X. CONCLUSION

In this work, we present StressAware, an application on the Amulet wearable platform to measure the stress levels of individuals continuously and in real time. The app implements a stress detection model, continuously streams data from a commercial heart-rate monitor such as a Zephyr, classifies the stress level of an individual, logs the stress level, and then displays it as a graph on the screen. The machine-learning results show an accuracy of 94.5% for the PhysioNet dataset, 100% for the field study dataset, and 64.3% for the lab study with RBF SVM. Testing the StressAware app revealed a projected battery life of up to 12 days. Also, the usability feedback from subjects revealed an interest in stress monitoring and showed that the Amulet and Zephyr have a potential to be used by people for monitoring their stress levels. The machine learning, energy efficiency, and usability results are promising, and show that StressAware has the potential to be used for stress measurement and monitoring and could eventually inform the development of stress-related intervention and personal stress management that could improve the health of individuals.

ACKNOWLEDGMENT

We thank Professor Ryan Halter for his advice throughout this work and Dr. Sarah Lord who helped us to write the IRB protocol to run the study. We also thank the members of the Amulet team for helping to resolve issues while using the Amulet: Ron Peterson, Tianlong Yun, Josiah Hester, and Travis Peters. We thank Gunnar Pope who ran the lab study and the volunteers who participated in the study. This work was funded by a grant from the National Science Foundation.

REFERENCES


Figure 3: Graph of Battery Percentage over 8 hours running StressAware.