4-18-2014

Estimating Drivers Stress from GPS Traces

Sudip Vhaduri

ESTIMATING DRIVERS' STRESS FROM GPS TRACES

by

Sudip Vhaduri

A Thesis

Submitted in Partial Fulfillment of the

Requirements for the Degree of

Master of Science

Major: Computer Science

The University of Memphis

May, 2014

Acknowledgements

First of all, I would like to thank my advisor Dr. Santosh Kumar - who always wanted nothing short of excellence from all his students, and pushed me to reach that level. Without him this thesis would not be possible. He has guided me and motivated me through all the times of weakness and when I wanted to give up. Secondly, I would like to thank all my committee members Dr. King Ip (David) Lin and Dr. Scott Fleming for taking time out of their busy schedules to review and advise me on my thesis. Next, I would like to thank Dr. Moushumi Sharmin for her continuous help in analyzing stress data during driving as well as reviewing and advising me on my thesis.

Part of this thesis is from [51]. I would like to thank Amin Ahsan Ali for his contribution in modeling and Md. Mahbubur Rahman and Syed Monowar Hossain for their contribution in collecting data from a week long mobile health study [38]. For my thesis and driving stress project, we borrowed data from that project. I am grateful to Dr Karen Hovsepian and Dr Kurt Plarre for their physiological stress model that generates continuous stress values of a participant based on his/her physiological data. The binary stress model, "Continuous Inference of Psychological Stress from Sensory Measurements Collected in the Natural Environment", was published in IPSN, 2011. A special thanks for my lab members and all the related project associates for their contributions to the projects that have made this thesis possible.

I would also like to mention two special mentors - Dr. A. K. M. Ashikur Rahman and Dr. Mohammad Kaykobad at Bangladesh University of Engineering and Technology, Bangladesh who inspired the love of research and academia in me. Last, but not the least, to my friends and family who helped me get to this point.

— Sudip Vhaduri

Abstract

Vhaduri, Sudip. MS. The University of Memphis. MS in Computer Science. May, 2014. Estimating Drivers' Stress from GPS Traces. Major Professor: Dr. Santosh Kumar.

Driving is known as a daily stressor and measurement of driver's stress in real-time can improve the awareness of stress for drivers, their cars, and their phones. Integrating sensors in future cars can help assess driver's stress, but it requires either wearing sensors by the driver or instrumenting the car. In this thesis, we present "GStress", a model to estimate driver's stress using only Smartphone GPS traces. By obviating any burden on the driver or the car, our approach has a better chance of wider adoption worldwide. The GStress model is developed and evaluated from data collected in a mobile health user study where 10 participants wore physiological sensors for 7 days (for more than 10 hours) in their natural environment, including during driving. Each participant had 10 or more driving episodes over the course of the study (for a total of 37 hours of driving data). This being the first work of its kind, provides a correlation of over 0.7 between the actual and estimated driving stress by identifying some major factors such as stops, turns and brakings that contribute to the stress of a driver. Incorporation of other factors in the model as well as use of more advanced modeling approaches can further improve the accuracy of the model.

Contents

List of Tables

List of Figures

Chapter 1

Introduction

Driving is known to be a daily stressor [20], and stress during driving can result in immediate adverse events such as accidents [23]. As a result, stress during driving can result in traffic fatalities in the short-term [13] and repeated occurrences of stress can cause or worsen cancer [43], heart diseases [8], hypertension [2], aging [32], shrinking of brain [19], fatigue, problem sleeping, depression, rage, among others [26, 27]. Consequently, there has been tremendous interest in both the scientific community as well as technology and car industry in coming up with methods to measure stress, enhance stress awareness of drivers, and find ways to reduce stress. Awareness of driver's stress in real-time can be used to trigger stress interventions [24] or passively via changing music being played [31].

Driver stress has traditionally been measured at a gross level via self-reports [20] or biofluids [11]. With the emergence and use of physiological sensors such as electrocardigram (ECG), galvanic response (GSR), and video, and their instrumentation on drivers or in the car, it became feasible now to collect a continuous measure of stress. Research on developing an accurate measure of stress from physiological sensors to assess the stress during driving has been continuing for more than a decade [17, 47] with improving accuracy. More recently, the physiological measures have been supplemented with data from the car about driving events such as steering wheel movements, braking episodes, etc. [42] and information about the driving conditions from the environment [41]. Video has also been used to detect driver's stress by pointing at the face of the driver while driving [34]. Another direction of research has been to assess the cognitive load that may be caused by technology in smart phone or in the car (e.g., texting, navigation). These works also make use of physiological sensors and sensors in the car [48]. Encouraged by the positive impact of sensor-collected data in the research setting, car industry is now beginning to include some of these sensors in the car [12, 49, 5, 6].

While inclusion of physiological sensors (such as galvanic skin response on steering wheel) and video cameras in future cars can provide a continuous measure of stress, monitoring of driver's stress today requires either having the drivers wear physiological sensors or instrumenting the car with video cameras. Methods that can immediately be used by drivers to measure their stress during driving widely is still missing.

In this thesis, we develop a model called *GStress* to estimate the stress level of a driver from GPS traces. Given that GPS sensors are readily available in navigation systems and are increasingly integrated in smartphones, obtaining GPS trace in real-time is becoming increasingly feasible worldwide. The GStress model can be used in a variety of ways. The driver can become more aware of their daily stress during driving. They can overlay the stress data on the map to determine their most frequent stress occurrences during driving and use it to make changes to their route of commute, time of commute, or driving behaviors. GStress model can help inform design of technologies that are used in vehicle. For example, calls or texts could be blocked or postponed if the driver is found to be stressed. If GStress is adopted widely, real-time data from GStress models used by several drivers on various routes can be used to annotate traffic map with current stress levels being experienced by drivers on various routes, similar to real-time traffic update displayed by navigation systems today. Data collected from population can also be used by city planners to identify pain points in a city's road network (e.g., difficult intersections) that induces more stress for drivers than others.

To develop the GStress model, we use physiological data collected from 30 human volunteers who wore the sensors daily (for 10+ hours per day) for a week in their natural environment. Of the 30 volunteers, we find 11 participants, all of whom had 10+ driving episodes as drivers. The GStress model development and evaluation uses data from these participants. A stress model that has been validated in both lab and field environment [36] was applied to the physiological data collected from these participants to obtain a continuous measurement of stress for each 30 second segment. The model provides a continuous measure that is normalized to be between 0 and 1. Our method can use other existing stress models or future improvements to stress models. GPS and self-report data were used to identify driving episodes. To assess the utility of the stress model in measuring driver's stress, we compare the average stress experienced during driving from rest of the day. We find that driving is 83% more stressful compared to the rest of the day.

For model development, we first analyze the entire traffic episode to identify events that have been shown to be stressful. These include stops, braking, and turns. Next, given wide variability across individuals in their stress reactivity, we developed a Generalized Linear Mixed Model (GLMM) to separate out the effects of between person variability. The GLMM model also permits exploiting non-linear relationships while retaining the simplicity of linear regression. By only using three factors (stops, turns, and braking) from the GPS data, the GStress model obtains an r value of 0.72. We then obtain a population estimate of the person-specific biases and obtain a person independent model. Via leave-one-subject-out evaluation, the GStress model provides a median (across the 10 participants) r value of 0.687. We quantify the contribution of each factor on the overall stress and find that stops are the most stressful events.

Chapter 2

Related Work

Healey and Picard [16] described an experimental protocol for data collection. According to their protocol, each driver drove a predefined route containing 15 different events, from which they have crated four stress level categories according to the subjects' self-report questionnaires. In total, 545 one minute segments were classified. Based on recognition performance, they rank their individual feature using a linear discriminant function, and find an optimal set of features for recognizing driver stress using a sequential forward floating selection algorithm. In another study, Healey and Picard [17] presented a method for data collection and analysis under real driving conditions for the detection of the state of driver stress. They collected real driving data from 24 trip of at least 50-min in duration. The data were analyzed in two ways. In the first case, they used features from 5-min intervals of data during rest, highway, and city driving conditions to distinguish three levels of driver stress for multiple drivers over several days. In the second case, they compared continuous features calculated at 1-s intervals throughout the entire drive with a metric of observable stressors created by independent coders from videotapes. The results showed that skin conductivity and heart rate metrics are most closely correlated to driver stress level. The participants used camera and a computer on the car seat to collect their data. They need to wear physiological sensor to detect their stress level. A continuous rating of stress was also obtained, but as noted in [36], in addition to the limited dataset, the stress model developed here was not validated. Moreover, participants were required to wear physiological sensors to detect their stress level. They also used camera and a computer on the car seat to collect their data. Overall experiment setup can be a significant burden and uncomfort for the participants. Whereas our final stress detection model is not stringent to contact any sensor to the participants body, though we have collected our training data using wearable sensor and a mobile phone.

Gulian et al. [14] have chosen to focus on the general stress experiences in their research and developed a diary in order to explore the personal and situational factors, both within and outside the driving environment, that might increase the general tendency to view driving as stressful. The diary consisted of questions about the journey (e.g. congestion level, types of roads), current problems at home/work, current health, and quality of sleep. Drivers were then asked about their feelings while driving and before leaving work, and then allowed to provide open ended comments about their daily experiences with driving, work, and leisure. Their results indicated that driver stress was related to negative driving experiences, such as traffic jams and being in a hurry, but also to problems outside the driving environment, including lack of sleep and work fatigue.

In a paper-and-pencil based study, Meschtscherjakov et al. [28] investigated User Experience (UX) factors and their relation to context factors. In terms of UX factors they were interested in the driver's general feelings during a trip and to what extend drivers enjoyed the trip. They wanted to know how distracted and stressed drivers are and what causes these experiences. Additionally, they were interested in perceived eco-friendly driving behavior and if participants had the feeling of loosing control over their vehicle. Finally, they wanted to know participants' estimation of trip costs. In terms of context factors, they wanted to know whether predominant weather and light conditions, number of passengers in the vehicle, type of road, traffic volume, trip purpose, trip length, tripduration, and average speed had an influence on the above mentioned UX factors. One of their question was "How stressed were you during the trip?". Participant had to rate this question on a 7-point Likert scale (0=not, 6=very) in order to get differentiated answers. Thereafter they should select answers why they were stressed on a multiple-choice set of predefined answers. They also had the possibility to add their own answer. In order to reduce the effort for participants we tried to keep open answers to a minimum. The researchers tried to sample the participants experience in variety of trips rather than capturing only commuting between home and work place. The researchers found that 19.3% times, the participants were stressed. Among their stress factors, they were stressed by traffic density 21% times, by time pressure 20.6% times. Among the stressful trips, 55.9% was for traffic density, 49.5% was for time pressure. Moreover, the participants are more stressed during their business trips and time pressure. This study is about stress during driving, however, this paper-and-pencil based method is rather time-consuming both for the people who collect the data and for the people who analyze them. Furthermore, data are collected (immediately) after the trip and not during the trip. Thus people might have forgotten what they actually experienced. In addition, people might have had several experiences during a trip. For example, the weather might have changed during a trip. Thus the method is inaccurate up to a certain extend. Unlike them, our study is not retrospective paper and pencil based. We have collected data using mobile phone during the driving period.

A study in [53] evaluated salivary amylase activity (sAMY) as an indicator of the acute psychological effects of driving. The psychological effects of driving were examined using sAMY analysis, oculomotor angle and subjective evaluation with a questionnaire, and the methods were compared. The change in sAMY over time was analysed before and during driving. The results indicate that the psychological effect of driving-induced stress is quickly quantified using sAMY. However, they did the study in using a driving simulator in the lab, and performance of stress detection in lab environment is always easier than detecting stress in natural environment. Unlike their study, our effort is to detect stress during their real driving in natural environment.

Next, we describe approaches reported in the literature concerning stress detection that are not related to the driving task. In the following approaches, the experiments were performed in a laboratory setting, where it is relatively easy to detect stress, since the sources and the number of stimulations are restricted, and the increase of sympathetic activity is related to a specific stimulation. However, in non-restricted environments, such as driving, the frequency and the sources of stimulations significantly vary, making the monitoring and, consequently, stress event detection more difficult.

Zhai and Barreto [55] developed a system for stress detection using blood volume pressure, skin temperature variation, electrodermal activity (EDA), and pupil diameter. Data were collected from 32 healthy subjects, demonstrating significant correlation between stress and the aforementioned physiological signals; the classification of stress was performed using a support vector machine (SVM). Rani et al. [39] presented a realtime method for stress detection based on heart rate variability (HRV) using Fourier and wavelet analysis. Ji et al. [18] presented a probabilistic model for detecting fatigue, which was extended by Li and Ji, allowing the detection of Nervous and Confused affective states [22]. The recognition of subject's affective state was based on probabilistic inference from features extracted from multiple sensors. These features include physiological measures, physical appearance, and performance measures. The main outcome of this paper is that the Bayesian framework is suitable for information fusion and provision of a reliable stress metric.

Rigas et al. [42] claimed real-time drivers stress event detection from physiological signal and the vehicle's CAN-bus the provides vehicle information e.g., speed, RPM, and throttle. They combined the physiological stress response with the driving behavior like overtake, hard braking, to improve the classification accuracy. They collected data from 13 subjects though most of the data come from first subject. Though they mainly used data from the first subject, sensitivity and specificity of their stress detection during driving is very low. Moreover, their system is stringent to wear physiological sensor attach to the subjects body. Their stress measurements only consider the driving sessions, not consider non-driving period of data. Finally none of the above method or system finds a route for given source and destination pair which is more relaxed.

Miller in his master's thesis [29], examined whether drivers' stress level various across various roadway conditions. The study was on 60 drivers from three age groups with scripted driving. He evaluated stress patterns across age and gender groups. In the study, he considered short and long interval stress to assess trends in stress from travel distance and roadway characteristics respectively. From evidence he proposed rough pavements and tunneled roadway segments are associated with an increase of cognitive load. He also found older age group faced largest incremental changes in physiological responses.

All the foregoing findings indicate that physiological signals can be exploited to provide a metric of driver stress in the car of the near future and to perform real-time driver stress monitoring. Stress monitoring could serve the management of noncritical in-vehicle information systems and provide a continuous measure of the way that road and traffic conditions affect drivers. However, a number of limitations deteriorate the applicability of the reported approaches in real-life driving conditions. The first limitation lies on the processing of physiological signals. An important issue, which is not tackled in many of the aforementioned works, is the real-time estimation of the signal baseline. The most common approach used in the literature is normalization using an initial phase, where the driver is supposed to be relaxed [17] to estimate the baseline of the signals.

The stress levels obtained from self-report studies do not have sufficient granularity to develop a model of stress estimation for a driving episode. The stress levels obtained from physiological monitoring studies have sufficient granularity, but most existing works lack a validated stress model due to insufficient dataset. Furthermore, the drivers' stress estimation system takes into account randomness due to person variability and obviates the need to wear any physiological sensor by developing a model that associates driving events such as stop, turn, braking etc. with stress.

In summary, assessment of driver's stress continues to be a very active area of research. Most of the existing works, however, focus on measuring stress from physiological measures [17, 47], video [34] and self-report [44, 14]. More recently, these measurements have been supplemented with driving and traffic related information [42, 48]. As acknowledged in very recent works [48], measurement of driver stress has usually been confined to simulators due to the difficulty, effort, and risk involved in collecting data in the natural environment. For those studies that are conducted in the natural environment, they were usually conducted along scripted routes under supervision, for very limited duration. To the best of our knowledge, this is the first study to collect continuous stress data in natural unscripted driving episodes in participants' own vehicle, where each participant contributes at least 10 driving episodes. Finally, while most existing works used driving event measurements together with physiological measurements to improve the accuracy of stress measurement, this is the first work to present a model for estimating stress from GPS data alone.

Chapter 3

Detecting Potentially Stressful Events during Driving from GPS traces

All the GPS data and participants' self-report used to detect the stressful driving events are not part of this project. We borrowed the dataset from a week long mobile health user study.

3.1 Driving Episode

3.1.1 Definition of Driving Episode

Vehicular movements are usually sandwiched between walking segments. Start of a commuting (vehicular) episode is defined when the speed, obtained from GPS samples, is over the maximum gait speed of 2.53 meter/sec [4]. A commuting episode is considered as driving episode when the person sits in the driver seat i.e. drive the car by himself/herself. A driving episode consists of various driving events such as stops, turns, congestions, braking etc.

3.1.2 Approach to Detect Driving Episode

In order to determine whether the participant was driving or riding in a commuting episode, we analyze the participant's response to the self-report question: *If you commuted since the last interview, what type?* with possible answers — *Driving, Biking, Walking, Riding as a Passenger, Riding Public Transportation, and Did not commute.* We select pre and post self-reports of a commuting episode and look at the participant's response to determine whether s/he drove or rode. Table 1 lists the participant statistics with the "Driver" pool that is composed of those who always responded "Driving" for the EMAs triggered during commute and the "Mixed" pool is composed of those who not always responded "Driving" (i.e., sometimes responded "Riding"). We consider these two pools of 25 participants as our potential subjects. Some of the commuting episodes detected do not have a self-report in their vicinity but we include them in our driving dataset, if they come from the "Driver" group. Otherwise, we include a commuting episode in the driving dataset only if there is a self-report explicitly confirming so.

Commuter Type	Female	Male
Driver (always drove)	8	6
Mixed (sometime riding)	4	7
Passenger (never drove)	3	2

Table 1: Summary statistics of the subjects.

3.2 Stop Segment

3.2.1 Definition of Stop Segment

Stop segments refer to parts of the driving episode when the vehicular speed obtained from GPS reaches zero. In most cases, stops occur when the vehicle encounter road intersections and the traffic signal is red, or a stop sign, or when they try to move in or out of a drive-way. It can consist of multiple consecutive stops or momentary stops (stopped just for a second - only one GPS sample as from the plot in Google Map we have seen many times participants stopped in intersections with a clear deceleration segment followed by just one zero speed GPS sample). From the CDF of zero speed segment duration of stops Figure 1, we also obtained more than 5% stop segments consists of only one zero speed gps sample. The entire stop segment and hence the total stop time is the time it takes from the point when the driver starts to decelerate until s/he starts accelerating after the final (the latest) zero speed sample. After a stop segment a driver usually start to increase the speed up to an almost constant speed. For our analysis, we are considering this segment of speed up, after a stop, as an acceleration segment.

3.2.2 Approach to Detect Stop Segment

To find the point when the driver starts decelerating, we look at prior 5 seconds from the final stop mark and check the speed difference ($dv = v_{t_{start}} - v_{t_{end}}$, where t_{end} and t_{start} (i.e., $t_{end} - 5$) are the end and start timestamps of the 5 seconds window respectively, $v_{t_{end}}$ and $v_{t_{start}}$ are the speeds at time t_{end} and t_{start} respectively, dv is the difference between instantaneous speed of 2 samples that are 5 seconds apart). We keep moving backward at 5 seconds interval, with 4 seconds overlap (i.e. 1 second sliding), until the speed difference dv at t_{end} is less than 10% of the speed at t_{start} i.e., $dv \leq 0.1 \times v_{t_{start}}$ and both of the two

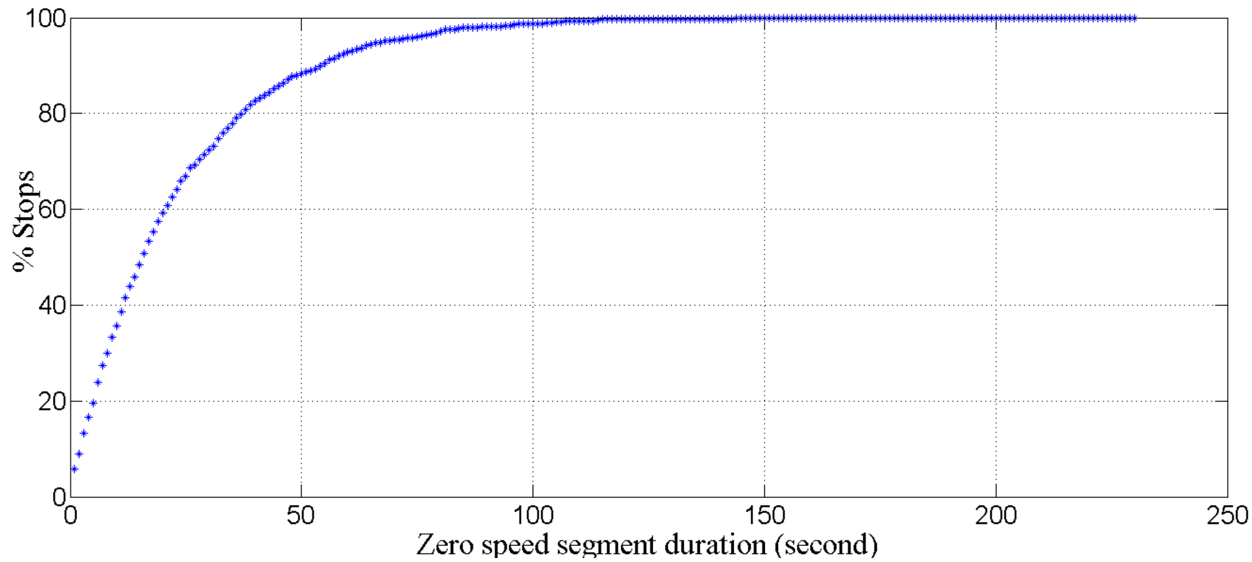

Figure 1: CDF of stop segment durations with zero (0) speed.

speeds (i.e., $v_{t_{start}}$ and $v_{t_{end}}$) are above the maximum gait speed and mark t_{start} as the start point of deceleration. See Figure 2 for the details.

3.2.3 Merging Closely spaced Stops

We merge multiple stops appear closely with slow moving driving segments among them. To merge these intermittent stops, we move backward by 5 seconds from the start (t_{start}) of each stop and compute the area under the speed curve as below

$$Area_i = \int_{t_{start}-5}^{t_{start}} dv(t)dt \tag{3.1}$$

where, $Area_i$ is the area of the i^{th} stop under the speed curve from $t_{start} - 5$ to t_{start}, dv is the change in speed during that time window. If $Area_i < \epsilon$, then we consider this slow moving driving segment as part of a stop and replace the speed values for this time period with zero. The backward propagation of the algorithm allows the detection and merging of stop-slow moving patterns until the vehicle starts moving at a faster pace. The value of the

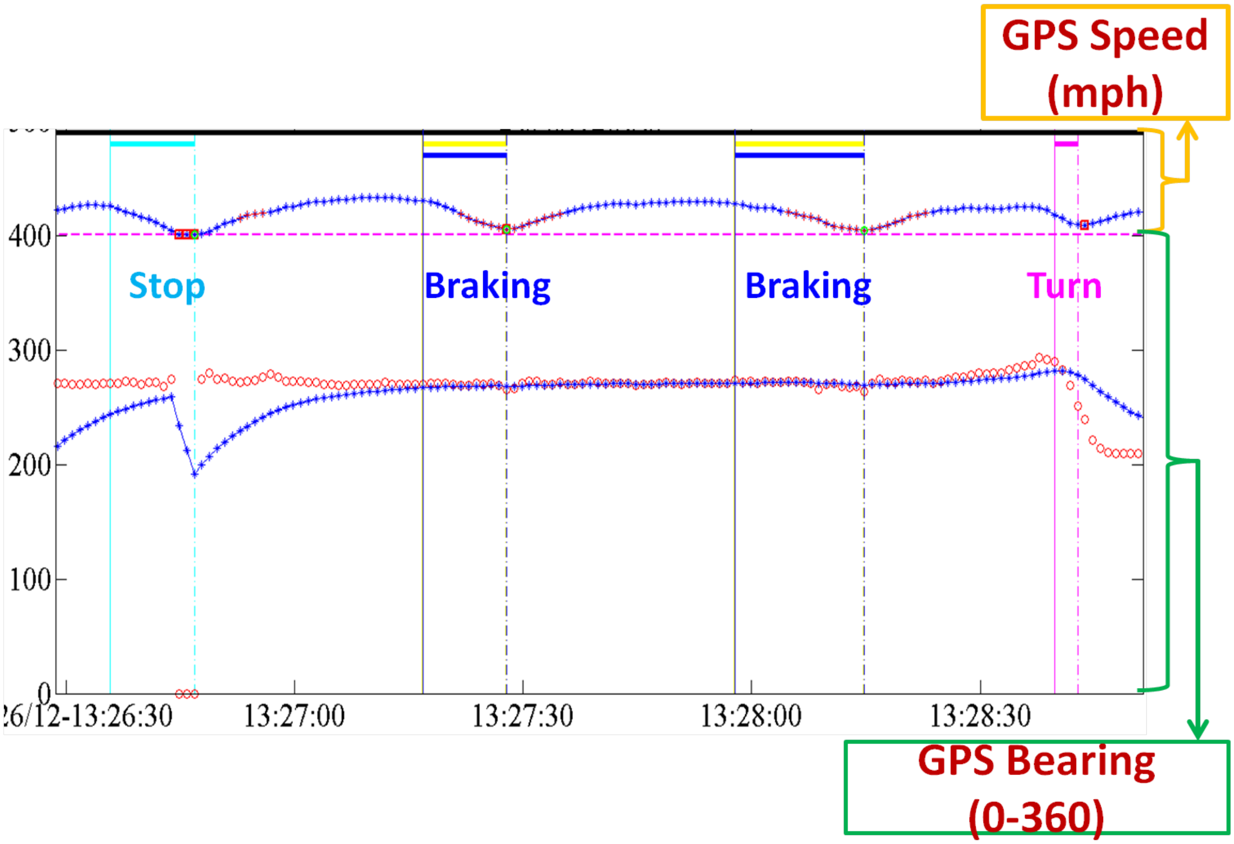

Figure 2: Part of a regular driving episode consisting of driving events such as stop, turn, and braking. In the y-axis, 0-360 is for GPS bearing and 400-500 is for GPS speed measured in mph. Cyan, blue and magenta lines corresponds to stop, braking and turn segments respectively. Solid and dashed lines are for start and end of a segment.

ϵ can be derived by taking the average of areas of all first backward windows across all stops. In our experiment, $\epsilon = 10.87$.

3.3 Driving Segments

We consider the parts of a driving episode as driving segments that remain after removing the stop segments. The driving segments consists of turn (left-right), sudden braking, congestion and various other maneuvers like overtaking, lane change, lane merge etc.

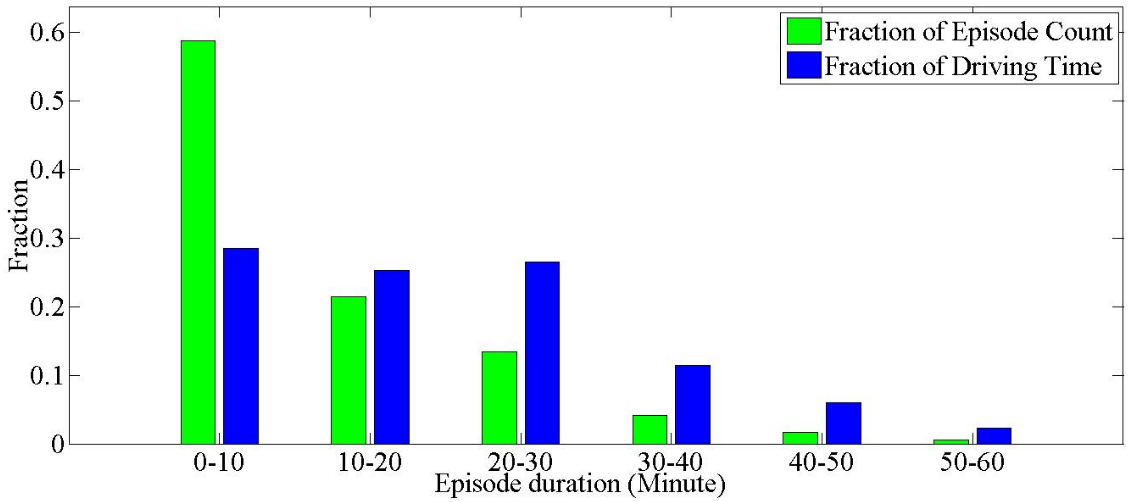

Figure 3: Distribution of driving episode counts and driving durations over 10 minute intervals. From the distribution we see majority of the driving episodes are shorter than 30 minutes but in terms of time contribution most of episodes fall between $0 - 30$ minutes.

We did not impose any constraint on route, time, direction or even on car on our participants during the study. It consists of driving events like stop, turn, braking etc; having an average duration of 12.51 minutes (SD = 9.15 minutes) and we observe that driving episodes of below 30 minutes are more frequent, see Figure 3.

Data missing and data quality were big issues while collecting data from our participants in natural environment without imposing constraint on them. We obtain 372 driving episodes from 25 subjects and for our stress modeling, we consider participants having more than 10 driving episodes. We found 11 out of 25 participants and together they contributed 295 driving episodes out of 372 i.e. nearly 80% driving episodes. Then we discard those episode that doesn't have stress data at all as well as the driving episodes from subject ID#31 as for him driving was a pleasant activity. We remain with 215 driving episodes from 10 subjects. We then discard driving episodes with poor data quality and obtain 181 (37.05 hours) out of 215 i.e. nearly 84% driving episodes to build our stress model. We found 637 stops, 1120 turns, 840 braking, 1477 acceleration segments from this 181 driving episodes.

We observed five congestion segments. This is because we did not impose any constraint on route selection and therefore, they picked their familiar and less congested routes. Table 2 presents a summary statistics of the stressful event durations.

Event Type	Mean (minutes)	SD (minutes)
Stop segments	2.04	1.72
Turn segments	0.28	0.18
Braking segments	0.20	0.32
Acceleration segments	0.44	0.51
Congestion segments	5.95	1.04

Table 2: Summary statistics of the stressful event durations.

3.4 Turn

3.4.1 Definition of Turn

A turn is associated with a change in driving direction (more than 30° [37]) and speed of the vehicle. The geometric properties [15] of the curve determines the amount of change in driving direction and speed while making a smooth transition between roadways or pass this section of a road. To reduce centrifugal forces and hence smoothly pass the curve, drivers usually reduce the speed significantly or make a complete stop.

3.4.2 Approach to Detect Turn

We detect turn from change in driving direction obtained from GPS bearing using a modified approach proposed in [37, 54] as we do not have external Dead Reckoning devices that consists of Gyroscope and Odometer. For slow movement, (i.e. speed lower than 3m/s) GPS bearing data is very inaccurate. In that case, we consider the heading at current and fifth last fix. For speed greater than 3m/s we consider the heading at current and third last fix. We detect a turn when the absolute value of direction change, i.e. difference between the two fixes is more than 35°. Sign of the direction change defines the turn type (right or left). We do not perform turn detection when the vehicle is not moving i.e. GPS speed is zero. See Figure 2 for details.

3.5 Braking and Acceleration

3.5.1 Definition of Braking and Acceleration

Braking is a driving event that causes an immediate deceleration segment in order to avoid unwanted scenario (e.g. stop suddenly before red light or stop to avoid hitting the front car that stopped suddenly) while driving. Braking can result in a full stop, however, here we are considering only the braking segment followed by an acceleration segment. Braking leading to a stop are considered as part of the corresponding stop. This way, some of our braking in a driving episode are considered inside stop segments. Intensity of deceleration defines the category of braking such as moderate, severe, negligible [21, 9]. After braking a driver usually starts to increase the speed up to an almost constant speed. For our analysis, we are considering this segment of speed up after braking as an acceleration segment.

3.5.2 Approach to Detect Braking and Acceleration

Braking detection consists of two steps. First, to find the local minima in each driving segment and then to identify the point from where deceleration starts. We follow two approach to find local minima - first, “PeakFinder” [35] to find local minima and second, “imregionalmin” [40] to find Regional minima. Then, to find the point when the driver starts decelerating, we move backward and check the speed difference similar as we did to find the deceleration start point for stops. See Figure 2 for details. To detect acceleration from the end of braking we follow same approach as braking detection.

3.6 Congestion

3.6.1 Definition of Congestion

When the demand of a road network exceeds its finite capacity, then network imposes additional travel cost to all users of the network and this situation is known as vehicular congestion. It can happen on regular, cyclic basis which reflects social and economic activities of a area. It can also happen irregularly in certain points in the network due to irregular occurrence of road work, breakdown, and/or accidents [50].

3.6.2 Approach to Detect Congestion

We use Density-Based Spatial Clustering of Applications with Noise (DBSCAN) [10] to detect traffic congestion which can be improved by incorporating surrounding traffic and geographical information [7, 3, 33], but our dataset doesn't have this information. If a driver moves slowly nearly 5 m/s, then in 3secod he will pass 15 meter of haversine distance and by that time the GPS receiver will gather 3 samples at 1Hz sampling rate. To apply DBSCAN, we define our core points as those points/smaples that have a neighborhood radius of 15 meter and have at least 3 points/samples within that radius. After we find the core points, we consider them together to make segments. We are considering only those segments that have at least a 5-minute duration. However, we found few congestion instances as our subjects took routes that were less congested.

Chapter 4

G-Stress Model

All the physiological data used to get continuous stress values are not part of this project. We borrowed the dataset from a week long mobile health user study. We first describe the stress computation from physiological data. We then describe the development of the GStress model and its evaluation.

4.1 Obtaining Stress from Physiological Data

We divide the entire day into 30 second segments (or windows) and compute stress for each segment as a continuous measure in terms of posterior probability of being stressed using the model presented in [36]. We present the overall pattern obtained using this model. We find that the average driving stress is 0.3992 while average stress during the rest of the day is 0.2178 (with standard error of 0.0041 and 0.0013 respectively). Thus, driving is 83% more stressful than rest of the day which is consistent with existing literature [17, 42, 47, 46]. Figure 4 presents average stress level of 11 participants during driving and during the rest of the day.

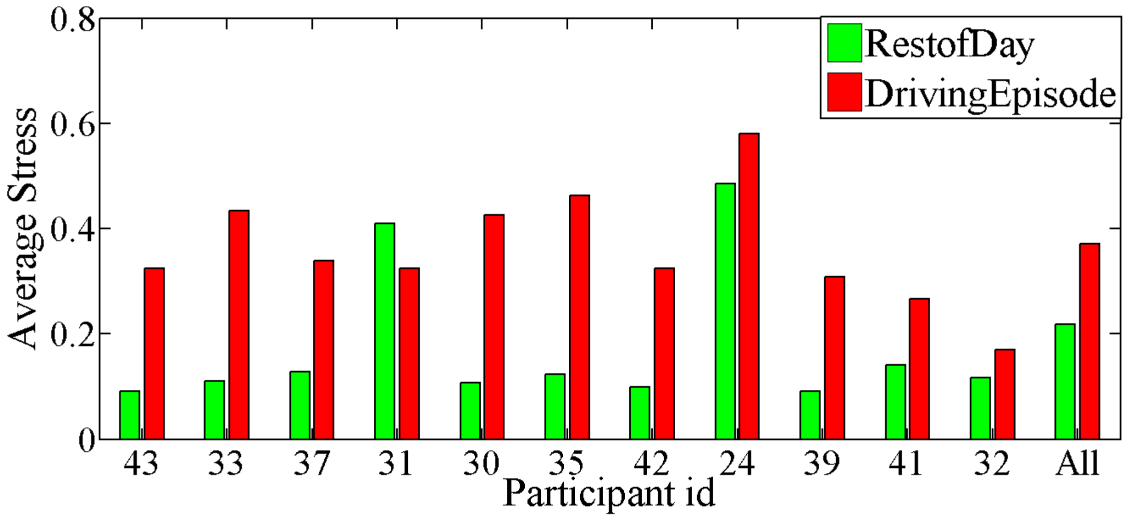

Figure 4: Average stress baseline for driving and rest of the day across different persons.

To test whether the stress during driving is indeed different from rest of the day, we use a two sample t-Test to test the null hypothesis, $H_0 : \mu_D = \mu_{RD}$, where μ_D and μ_{RD} are average stress during driving and during rest of the day. We perform the test for both individual level and population level, i.e. all subject together. At $\alpha = 0.05$ level of significance, we reject the null hypothesis for both individual level and population level with a $p - value < 0.001$. Therefore, participants' average driving and average non-driving stress are not equal. We observed an anomaly for Participant ID#31. This participant had three exams and a couple of deadlines during the study week. Further, he mentioned in his interview that he enjoyed driving and used it to relax. Hence, his average stress level for rest of the day is higher than that during driving. He enjoyed driving and considered driving as his method of stress reduction. This participant is not used in developing GStress model since driving is not a stressful activity for him/her. We are not covering such uncommon incidents in our stress estimation model.

4.2 Development of G-Stress

Figure 5 shows a snapshot of several driving episodes plotted in Google Earth. From the plot, we can see different parts of the driving episodes are stressful (red colored marker). We examined this representation to identify factors that are contributing to driving stress. Then we randomly select a driving episode from a random participant (Figure 6) and have observed that parts of the episode are stressful. We start investigating in reason for the parts of the entire driving episode to be stressful. We have found there was a left turn and a stop that makes that part of driving stressful (Figure 7). Therefore, we need to come up with a method to account for these events while assessing/estimating driving stress.

The goal of GStress model is to estimate the stress level in a driving episode from GPS trace. For this purpose, it can use any data that can be inferred from GPS. We used the duration of stops, turns, brakings, and driving segments within a driving episode to estimate the stress level.

The simplest model that can be used to model stress data is a linear regression model with the assumption that errors are Normally distributed. This assumption does not hold for cases where the response variable (Y) is count, proportion or positive continuous data. Our response variable(Y) is average stress of a driving episode which is a positive continuous data. Hence, we considered Generalized Linear Models (GLM) which assume that data comes from some distribution other than Normal distribution and a linear function ($\eta = g(.)$) of the mean ($\mu = E[Y]$) of response variable is related to the predictors i.e. $\eta = g(\mu) = X\beta$, where, X stands for predictor variables, β for fixed-effects regression coefficients. We used *Gamma* distribution as it is suitable for cases where the response variable takes positive continuous data such as ours and *identity* as our transformation/link function (i.e. $\mu = (E[Y]) = X\beta$, where, $Y = X\beta + \epsilon$). An advantage of using such a descriptive model is that it helps us to determine the relative importance of each factor in measuring the response variable.

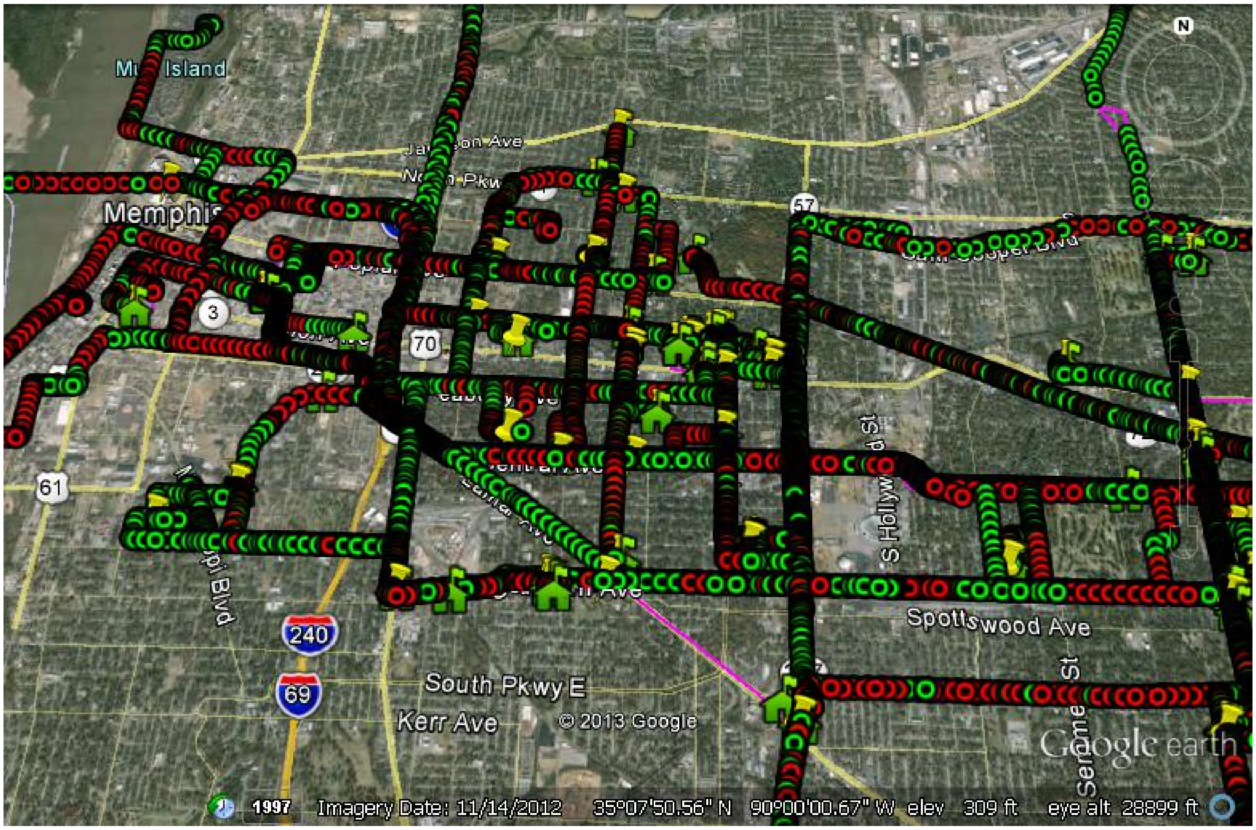

Figure 5: Snapshot of several driving episodes plotted in Google Earth. Yellow pin and green flagged house are start and end marker of a driving episode. Circular markers are for GPS traces plotted at 0.2Hz i.e. gap between successive traces is 5 second. Red and green markers are for stressed and not stressed samples.

Driving in a natural environment involves several factors such as phone call, bad weather etc. in addition to the major factors we detected from GPS traces, and also there exists randomness due to wide between-person variation in stress-reactivity that restricts the use of regression models like GLM that relies on only fixed effects. Hence, we need some modeling scheme that takes into account the random effects in addition to the fixed effects and it should be generalizable. Generalized Linear Mixed Model (GLMM) is, therefore, widely used in health research since it takes into account both the fixed effects and the random

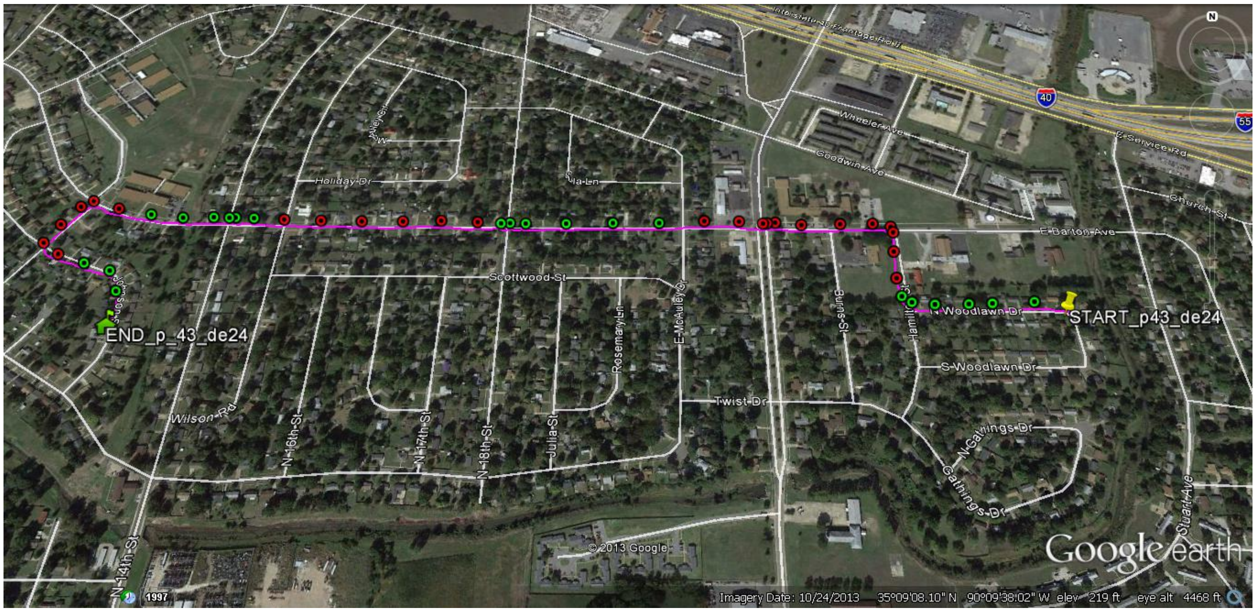

Figure 6: A regular driving episode from participant#43. Parts of this driving episode is stressful i.e. red colored.

effects that we cannot estimate with fixed effects or errors. The general form of the model is

$$Y = X\beta + Z\gamma + \epsilon \tag{4.1}$$

where, Y stands for the response variable, X stands for predictor variables, β for fixed-effects regression coefficients, Z stands for design matrix for random effects, γ for random effects, ϵ stands for random errors or residuals.

Our predictor variables for the fixed effects are the amount of time, in terms of 30 second segment counts, affected by different events (stop(x_3), braking(x_4), turn(x_5), acceleration after braking(x_6) and after stop(x_7), congestion(x_8)). We also consider the stress level prior to driving(x_1) and the amount of driving time(x_2), in terms of 30 second segment counts, that is not affected by stop, braking, turn, acceleration and congestion as our predictor variables. We scale all the predictor variables except x_1 in a 0 to 1 scale by dividing the

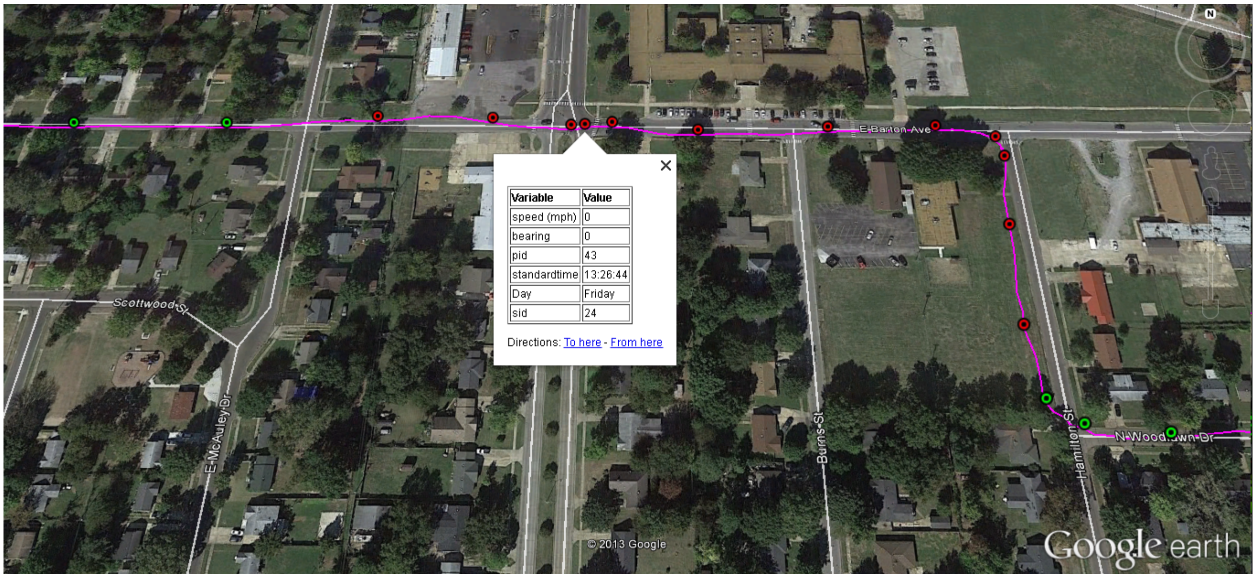

Figure 7: Stressful part of a driving episode from participant#43. The stressful part consists of a left turn and a stop.

segment count by the total number of 30 second segments in a driving episode. The output or response variable(y) is the average stress per 30 second segment in a driving episode. We consider person level variation as our random effect.

Our stress estimation model based on GPS traces, GStress is

$$y_{ij} = (\beta_0 + \gamma_j) + \beta_1 x_{1ij} + \beta_2 x_{2ij} + \beta_3 x_{3ij} + \beta_4 x_{4ij} + \beta_5 x_{5ij} + \beta_6 x_{6ij} + \beta_7 x_{7ij} + \beta_8 x_{8ij} + \epsilon \quad (4.2)$$

where, y_{ij} is the average stress of i^{th} driving episode of person j; for $k = 1 \ldots 8$, β_k is the k^{th} fixed effects, x_{kij} is the k^{th} predictor variables for fixed effects in i^{th} episode of j^{th} person; β_0 is the fixed effect on intercept and γ_j is the random effect on intercept for $j^{th} person$.

While building the "GStress" model from all 10 subjects' data, we did not find sufficient number of "Congestion" instances which is because we did not impose any route constraint on our subjects, and hence they usually took routes that are not congested. We tested the significance of both "Congestion" and "Acceleration" (both after braking and stop) i.e.

$H_0 : \beta_6 = \beta_7 = \beta_8 = 0$. We built two models – one with congestion and acceleration effect (refer, gm44) and another one without congestion and acceleration effect (refer, gm33). We performed Chi-Square test on the two models and cannot reject the H_0 for $\chi^2 = 1.5678$ and $p - value = 0.2105$ at $\alpha = 0.05$ level of significance. Also, the model (gm44), without "congestion" and "acceleration" fixed effects, has lower Akaike Information Criterion (AIC) [1] and Bayesian Information Criterion (BIC) [45] (Table 3), so the effect of "Congestion" and "acceleration" is not significant for ours.

We also tried to see whether random effect of person/subject is significant i.e. $H_0 : \gamma = 0$ and to do so we build two models – one with person random effect (refer, gm22) and a second one without person random effect (refer, gm11). We performed Chi-Square test on the two models and reject the null hypothesis ($\chi^2 = 3237.8$ and $p - value < 0.001$ at $\alpha = 0.001$ level of significance). Also, the model with "person random effect" (gm22) has lower AIC, and BIC(Table 3), so person random effect is significant.

model	AIC	BIC	Deviance	χ^2	$p - value$
gm11	-3367.1	-3341.4	-3383.1		
gm22	-3447.3	-3418.4	-3465.3	82.245	< 0.001
gm33	-95.5	-69.9	-107.88		
gm44	-95.074	-66.287	-109.35	1.5678	0.2105

Table 3: χ^2 test to check the significance of "person random effect" and fixed effect "congestion" and "acceleration".

Therefore, the final GStress model is,

$$y_{ij} = (\beta_0 + \gamma_j) + \beta_1 x_{1ij} + \beta_2 x_{2ij} + \beta_3 x_{3ij} + \beta_4 x_{4ij} + \beta_5 x_{5ij} + \epsilon \tag{4.3}$$

Table 4 presents our GStress model. Stress levels prior to driving(x_1), driving time without events(x_2), amount of driving time affected by stop(x_3) and affected by turn(x_5) are

fixed effects	β	SE	$t - value$	$p - value$
Intercept	-0.445	0.181	-2.454	0.014147
x_1	0.681	0.077	8.795	<0.001000
x_2	0.682	0.188	3.632	0.000281
x_3	0.755	0.207	3.655	0.000257
x_4	0.668	0.210	3.178	0.001480
x_5	0.703	0.200	3.523	0.000427

Table 4: G-Stress: Driving Stress Estimation Model. Here, SE stands for standrd error.

significant at $\alpha = 0.001$, amount of time affected by braking(x_4) is significant at $\alpha = 0.01$, and intercept is significant at $\alpha = 0.05$. We observe that all significant fixed effect factors have positive coefficients, i.e. they increase stress while driving.

4.3 Evaluation of G-Stress

We obtain a correlation of 0.722 (Pearson Correlation, r) between the actual and estimated driving stress while building the "GStress" model (Table 4) considering all 10 subjects' together. The variance for person variability is 0.002, variance for residual is 0.188, variance for fixed effects is 0.061. Therefore, $R^2_{GLMM}(c) = 0.252$ [30] i.e. 25% variability of data can be explained with both fixed effect and random effect.

In the Bland-Altman plot [25] (Figure 8), green, cyan, magenta and red dashed lines are for 25%, 50%, 75% and 100% difference respectively. The plot shows moderate agreement of actual stress and predicted stress, but lack of consistent agreement for values lower than 0.4 and variability above 0.4 mainly within 50% difference levels. Using paired t-test we found that the estimated stress closely matches with the actual stress for $p - value = 0.8893$ at $\alpha = 0.05$.

To evaluate the suitability of GStress model usage on participants on whom no training data has been collected, we train the model on nine participants and apply it on the remaining

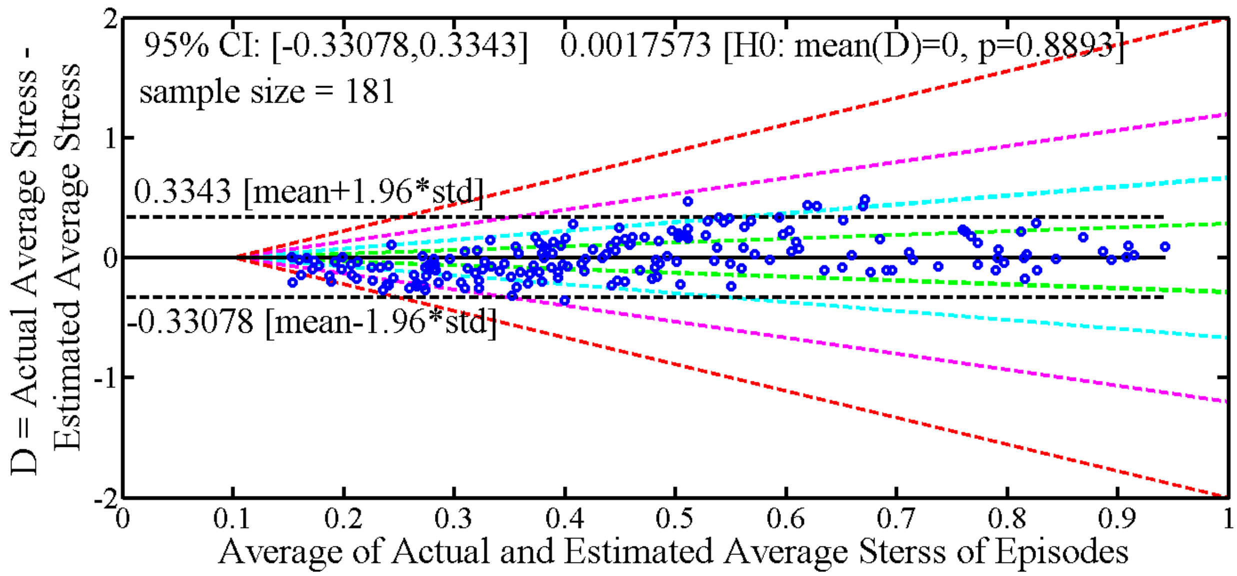

Figure 8: Bland-Altman plot for the actual and estimated average stress for all driving episodes.

one participant with a population estimate of the random effect and fixed effects. Figure 9 shows the correlation for each subject. For this leave-one-subject-out validation, we obtain a median correlation of 0.687 (see Table 5).

4.4 Contribution of Traffic Factors to Stress

We now quantify the contribution of each traffic factor to the total stress. The contribution of the factors depend on two items — how frequent they are in a typical driving episode and their weight in the traffic model. For the first, we compute the amount of time (in terms of number of 30 second segments) that are classified to be affected by a particular factor. Figure 10(a) shows the fraction of time in a driving episode that falls under each class — stops, turns, brakings, and driving. We observe that 24% of the 30 second segments are classified as stops and the rest as driving. From the 76% attributed to the driving time, 23% is affected by turns and 21% by brakings. This leaves 32% of the total driving episode to

test subject	r	r^2	ρ	τ	SSE
43	0.667	0.444	0.614	0.450	0.152
33	0.790	0.624	0.740	0.552	0.118
37	0.708	0.502	0.638	0.463	0.099
30	0.493	0.243	0.403	0.290	0.166
35	0.761	0.578	0.659	0.477	0.122
42	0.515	0.265	0.749	0.516	0.166
24	0.934	0.873	0.855	0.673	0.083
39	0.605	0.366	0.679	0.524	0.072
41	0.503	0.253	0.371	0.333	0.188
32	0.966	0.934	1.000	1.000	0.078
Median	0.687	0.473	0.669	0.497	0.120

Table 5: Different measures of relationship between the actual average stress and estimated average stress for "leave one subject out testing". Where, r is Pearson linear correlation, ρ is Spearman non-linear rank correlation and τ is Kendalls concordance measure based on rank, SSE is Sum of Squared Error. Subject-32 and subject-24 have very few driving episodes with stress data. Therefore, we observed very high correlation from them.

be classified as *driving*. There are several other factors in a driving episode that may cause stress (e.g., sharp curves in a road, potholes), but in the GStress model, they are all still accounted under the broad umbrella of *driving*. Future work can tease out these additional factors and improve the accuracy of stress estimation from GPS data. We also note that high occurrence of stops, turns, and brakings in our dataset may be due to driving in university neighborhoods where our participants stay.

Figure 10(b) shows the contribution of various factors to the overall stress. If we compare these contributions to their frequency of occurrences, we observe that stops are more stressful than driving since stops contribute 27% to the stress even though they only occur 24% of the time. On the other hand, driving accounts of 32% of the time, but only contributes 28% to the total stress. This may imply that reducing stops (e.g., highways, expressways) in a driving episode may be one approach to reducing stress during driving.

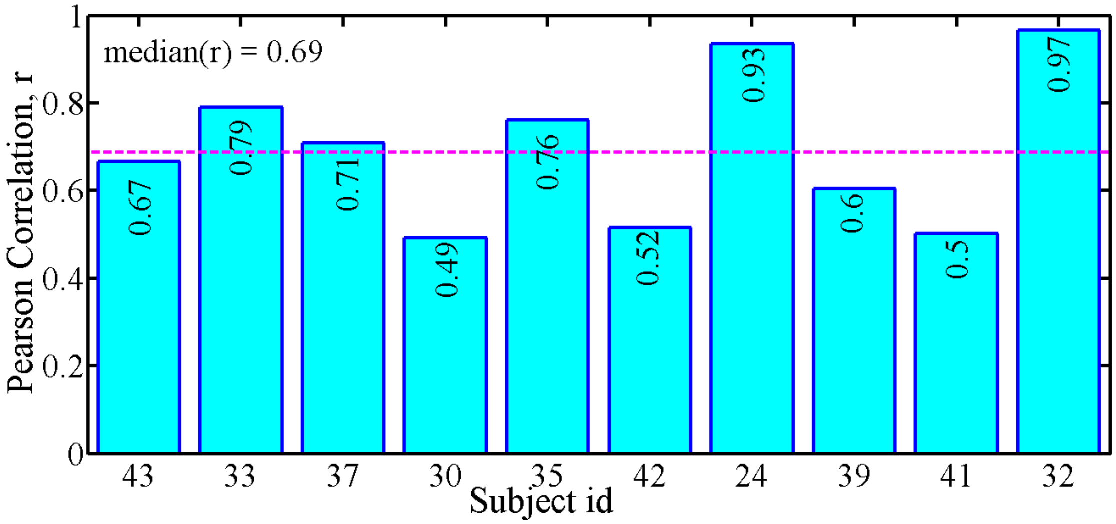

Figure 9: Pearson Correlation, r between actual and estimated average stress for "leave one subject out" validation. Horizontal magenta line corresponds to median correlation (0.687) from all 10 test subjects.

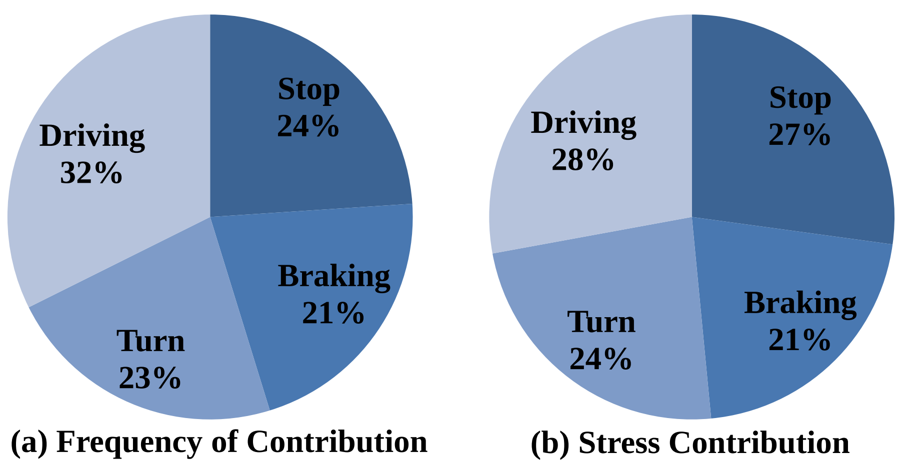

Figure 10: (a) shows the fraction of an average driving episode that is affected by various stressful factors i.e. frequency of events in an average driving episode. (b) shows the contribution of each factor to the total stress.

Chapter 5

Conclusions and Future Work

Its promising research direction to estimate driver's stress from GPS traces as it has several real life implications such as it will help us to become aware of one of the very common daily stressor, i.e. driving, can reduce the chance of an accident, can help city planners in designing better roads and intersections, in the long term can reduce chance of cancer, heart diseases, hypertension, depression, can help us delivering stress interventions in real-time, can make the car as well as the phone stress aware of the driver. The stress aware phone can block or forward calls when the driver is stressed and the stress aware car can play pleasant music to reduce driver's stress. The big advantage of such a phone based stress estimation approach is that it imposes no additional burden on the driver such as wearing sensors as well as no burden of instrumenting the car. Therefore, it can be adopted worldwide in a large scale.

This work pointed out the feasibility of estimating drivers' stress using only GPS traces collected from the driver's smart phone. This being the first work of its kind, provides a correlation of over 0.7 by identifying some major factors such as stops, turns and brakings. This correlation can be improved by fetching/finding more factors from driving in addition to stop, turn, braking i.e. by increasing information gain about the factors during driving.

We can obtain road information such as complex intersections, number of lanes, type of road, speed limit, traffic light, curvature information etc. from a geographic database such as digital map. We can also obtain environment information of the vehicle such as traffic condition from radar, steering wheel information, Anti-lock Braking Systems(ABS) etc. Availability of road and other traffic information as proposed by Woltermann et al. [52] will contribute to enhance the correlation i.e. can have better estimation of drivers' stress. Another direction can be the use of more advanced modeling approaches to further improve the accuracy of the model.

Bibliography

[1] Hirotugu Akaike. A new look at the statistical model identification. *IEEE Transactions on Automatic Control*, 19(6):716–723, 1974.

[2] M. al'Absi and D.K. Arnett. Adrenocortical responses to psychological stress and risk for hypertension. *Biomedecine & Pharmacotherapy*, 54(5):234–244, 2000.

[3] Ramon Bauza, Javier Gozalvez, and Joaquin Sanchez-Soriano. Road traffic congestion detection through cooperative vehicle-to-vehicle communications. In *IEEE LCN*, pages 606–612, 2010.

[4] Richard W Bohannon. Comfortable and maximum walking speed of adults aged 20–79 years: reference values and determinants. *Age and Ageing*, 26(1):15–19, 1997.

[5] Your Car as Stress Monitor? `http://spectrum.ieee.org/riskfactor/green-tech/advanced-cars/your-car-as-stress-monitor`.

[6] A Car That Takes Your Pulse. `http://online.wsj.com/news/articles/SB10001424127887324352004578131083891595840?mg=reno64-wsj&url=http%3A%2F%2Fonline.wsj.com%2Farticle%2FSB10001424127887324352004578131083891595840.html`.

[7] Ahmet Can Diker and Elvin Nasibov. Estimation of traffic congestion level via fn-dbscan algorithm by using gps data. In *IEEE PCI*, pages 1–4, 2012.

[8] J.E. Dimsdale. Psychological stress and cardiovascular disease. *Journal of the American College of Cardiology*, 51(13):1237–1246, 2008.

[9] Partich AM Ehlert and Leon JM Rothkrantz. Microscopic traffic simulation with reactive driving agents. In *IEEE Intelligent Transportation Systems*, pages 860–865, 2001.

[10] Martin Ester, Hans-Peter Kriegel, Jörg Sander, and Xiaowei Xu. A density-based algorithm for discovering clusters in large spatial databases with noise. In *KDD*, volume 96, pages 226–231, 1996.

[11] Gary W Evans and Sybil Carrère. Traffic congestion, perceived control, and psychophysiological stress among urban bus drivers. *Journal of Applied Psychology*, 76(5):658–663, 1991.

[12] Ford smart car locks your phone when youre stressed or distracted. `http://www.extremetech.com/extreme/132247-this-smart-car-locks-out-calls-and-texts-but-only-when-it-senses-heavy-traffic`.

[13] Global Status Report on Road Safety 2013, World Health Organization, 2013. `http://www.who.int/iris/bitstream/10665/78256/1/9789241564564_eng.pdf`.

[14] E. Gulian, AI Glendon, G. Matthews, DR Davies, and LM Debney. The stress of driving: a diary study. *Work & Stress*, 4(1):7–16, 1990.

[15] Fundamentals of Transportation/Horizontal Curves/Fundamental Horizontal Curve Properties. `http://en.wikibooks.org/wiki/Fundamentals_of_Transportation/Horizontal_Curves`.

[16] J. Healey and R. Picard. Smartcar: Detecting driver stress. In *IEEE International Conference on Pattern Recognition*, volume 4, pages 218–221, 2000.

[17] J.A. Healey and R.W. Picard. Detecting stress during real-world driving tasks using physiological sensors. *IEEE Transactions on Intelligent Transportation Systems*, 6(2):156–166, 2005.

[18] Q. Ji, Z. Zhu, and P. Lan. Real-time nonintrusive monitoring and prediction of driver fatigue. *IEEE Transactions on Vehicular Technology*, 53(4):1052–1068, 2004.

[19] H.J. Kang, B. Voleti, T. Hajszan, G. Rajkowska, C.A. Stockmeier, P. Licznerski, A. Lepack, M.S. Majik, L.S. Jeong, M. Banasr, et al. Decreased expression of synapse-related genes and loss of synapses in major depressive disorder. *Nature Medicine*, 18(9):1413–1417, 2012.

[20] M. Koslowsky, A.N. Kluger, and M. Reich. *Commuting Stress: Causes, Effects and Methods of Coping*. Springer, 1995.

[21] John D Lee, Joshua D Hoffman, and Elizabeth Hayes. Collision warning design to mitigate driver distraction. In *ACM SIGCHI conference on Human factors in computing systems*, pages 65–72, 2004.

[22] X. Li and Q. Ji. Active affective state detection and user assistance with dynamic bayesian networks. *IEEE Transactions on Systems, Man and Cybernetics, Part A: Systems and Humans*, 35(1):93–105, 2005.

[23] Cejun Liu and Tony Jianqiang Ye. Run-off-road crashes: An on-scene perspective, National Highway Traffic Safety Administration. Technical Report DOT HS 811 500, 2011.

[24] Diana MacLean, Asta Roseway, and Mary Czerwinski. Moodwings: a wearable biofeedback device for real-time stress intervention. In *Proceedings of the 6th International Conference on Pervasive Technologies Related to Assistive Environments*, page 66, 2013.

[25] J Martin Bland and DouglasG Altman. Statistical methods for assessing agreement between two methods of clinical measurement. *The lancet*, 327(8476):307–310, 1986.

[26] B.S. McEwen. Stress, adaptation, and disease: Allostasis and allostatic load. *Annals of the New York Academy of Sciences*, 840(1):33–44, 2006.

[27] B.S. McEwen. Physiology and neurobiology of stress and adaptation: Central role of the brain. *Physiological Reviews*, 87(3):873–904, 2007.

[28] A. Meschtscherjakov, D. Wilfinger, S. Osswald, N. Perterer, and M. Tscheligi. Trip experience sampling: Assessing driver experience in the field. In *ACM International Conference on Automotive User Interfaces and Interactive Vehicular Applications*, pages 225–232, 2012.

[29] Erika E Miller. Effects of roadway on driver stress: An on-road study using physiological measures. Master's thesis, University of Washington, 2013.

[30] Shinichi Nakagawa and Holger Schielzeth. A general and simple method for obtaining r2 from generalized linear mixed-effects models. *Methods in Ecology and Evolution*, 4(2):133–142, 2013.

[31] Shahriar Nirjon, Robert F Dickerson, Qiang Li, Philip Asare, John A Stankovic, Dezhi Hong, Ben Zhang, Xiaofan Jiang, Guobin Shen, and Feng Zhao. Musicalheart: A hearty way of listening to music. In *ACM SenSys*, pages 43–56. ACM, 2012.

[32] A. O'Donovan, A.J. Tomiyama, J. Lin, E. Puterman, N.E. Adler, M. Kemeny, O.M. Wolkowitz, E.H. Blackburn, and E.S. Epel. Stress appraisals and cellular aging: A key role for anticipatory threat in the relationship between psychological stress and telomere length. *Brain, Behavior, and Immunity*, 2012.

[33] Rebecca Ong, Fabio Pinelli, Roberto Trasarti, Mirco Nanni, Chiara Renso, Salvatore Rinzivillo, and Fosca Giannotti. Traffic jams detection using flock mining. In *Machine Learning and Knowledge Discovery in Databases*, pages 650–653. Springer, 2011.

[34] M Paschero, G Del Vescovo, L Benucci, A Rizzi, M Santello, G Fabbri, and FM Frattale Mascioli. A real time classifier for emotion and stress recognition in a vehicle driver. In *IEEE International Symposium on Industrial Electronics (ISIE)*, pages 1690–1695, 2012.

[35] How to find local peaks or valleys in a noisy vector? `http://www.mathworks.com/matlabcentral/fileexchange/25500-peakfinder`.

[36] K. Plarre, A. Raij, S.M. Hossain, A.A. Ali, M. Nakajima, M. Al'absi, E. Ertin, T. Kamarck, S. Kumar, M. Scott, D. Siewiorek, A. Smailagic, and L. Wittmers. Continuous inference of psychological stress from sensory measurements collected in the natural environment. In *IEEE IPSN*, pages 97–108, 2011.

[37] Mohammed A Quddus, Washington Y Ochieng, and Robert B Noland. Current map-matching algorithms for transport applications: State-of-the art and future research directions. *Transportation Research Part C: Emerging Technologies*, 15(5):312–328, 2007.

[38] Md Mahbubur Rahman, Rummana Bari, Moushumi Sharmin, Amin Ahsan Ali, Syed Monowar Hossain, Mustafa al'Absi, Andrew Raij, Emre Ertin, David Epstein, Kenzie Preston, Satish Kedia, Ashley Kennedy, Kenneth Ward, Michelle Jobes, and Santosh Kumar. Are we there yet? feasibility of long-term stress assessment via physiological sensors in field. In *ACM UbiComp (under review)*, 2014.

[39] P. Rani, J. Sims, R. Brackin, and N. Sarkar. Online stress detection using psychophysiological signals for implicit human-robot cooperation. *Robotica*, 20(06):673–685, 2002.

[40] How to find local peaks or valleys in a noisy vector? `http://www.mathworks.com/help/images/ref/imregionalmin.html`.

[41] G. Rigas, Y. Goletsis, P. Bougia, and D.I. Fotiadis. Towards driver's state recognition on real driving conditions. *International Journal of Vehicular Technology*, 2011, 2011.

[42] G. Rigas, Y. Goletsis, and D.I. Fotiadis. Real-time driver's stress event detection. *IEEE Transactions on Intelligent Transportation Systems*, 13(1):221–234, 2012.

[43] R.M. Sapolsky. *Why Zebras Don't Get Ulcers*. Owl Books, 2004.

[44] Caroline Schießl, Mark Vollrath, Tobias Altmüller, Michael Dambier, and Christian Kornblum. What is strenuous? driving itself or the driving situation. In *Intelligent Transportation Systems and Services*, 2006.

[45] Gideon Schwarz et al. Estimating the dimension of a model. *The Annals of statistics*, 6(2):461–464, 1978.

[46] Mandeep Singh and Abdullah Bin Queyam. A novel method of stress detection using physiological measurements of automobile drivers. *International Journal of Electronics Engineering*, 5(2), 2013.

[47] Rajiv Ranjan Singh, Sailesh Conjeti, and Rahul Banerjee. Assessment of driver stress from physiological signals collected under real-time semi-urban driving scenarios. *International Journal of Computational Intelligence Systems*, (ahead-of-print):1–15, 2013.

[48] Erin T Solovey, Marin Zec, Enrique Abdon Garcia Perez, Bryan Reimer, and Bruce Mehler. Classifying driver workload using physiological and driving performance data: Two field studies. In *ACM CHI*, 2014.

[49] BMW Steering Wheel monitors your heart. `http://4lunchdate.wordpress.com/tag/bmw-steering-wheel-sensors-monitor-vital-signs-and-stress-levels/`.

[50] Michael AP Taylor, Jeremy E Woolley, and Rocco Zito. Integration of the global positioning system and geographical information systems for traffic congestion studies. *Transportation Research Part C: Emerging Technologies*, 8(1):257–285, 2000.

[51] Sudip Vhaduri, Amin Ahsan Ali, Moushumi Sharmin, Karen Hovsepian, and Santosh Kumar. Estimating drivers' stress from gps traces. In *ACM UbiComp (under review)*, 2014.

[52] Bernd Woltermann and Stefan Schroedl. Personalized driver stress prediction using geographical databases, July 29 2003. US Patent 6,599,243.

[53] M. Yamaguchi and J. Sakakima. Evaluation of driver stress in a motor-vehicle driving simulator using a biochemical marker. *The Journal of International Medical Research*, 35(1):91–100, 2007.

[54] Chuang-Wen You, Nicholas D Lane, Fanglin Chen, Rui Wang, Zhenyu Chen, Thomas J Bao, Martha Montes-de Oca, Yuting Cheng, Mu Lin, Lorenzo Torresani, et al. Carsafe app: alerting drowsy and distracted drivers using dual cameras on smartphones. In *ACM MobiSys*, pages 13–26, 2013.

[55] J. Zhai and A. Barreto. Stress detection in computer users based on digital signal processing of noninvasive physiological variables. In *IEEE International Conference of the Engineering in Medicine and Biology Society (EMBS)*, pages 1355–1358, 2006.