Transdisciplinary Engineering for Complex Socio-technical Systems K. Hiekata et al. (Eds.) © 2019 The authors and IOS Press. This article is published online with Open Access by IOS Press and distributed under the terms of the Creative Commons Attribution Non-Commercial License 4.0 (CC BY-NC 4.0). doi:10.3233/ATDE190127

Design of Data Collection and Analysis Method for a Pleasant and Safe User Experience of Personal Mobility Device

Sebastian EIO^a, Jo-Yu KUO^{b,1}, Chun-Hsien CHEN^a and Pai ZHENG^a

^a School of Mechanical and Aerospace Engineering, Nanyang Technological University ^b School of Electrical and Electronic Engineering, Nanyang Technological University

Abstract. In recent years, the idea of personal mobility devices (PMD) has gained prominence globally for different contexts, for diverse types and extent of uses. The advantages of owning a PMD allows users to cover the short distance in between stops where they have access to long distance transportation, establishing a full end to end transport system for many. The rise in usage of PMDs also came along the rise in accidents. One of the reasons that could result in this phenomenon is the lack of calibration of PMD towards how users use it. Currently, most user experience (UX) methodologies are based on subjective questionnaires rather than by objective quantitative data. While there exists a few that studies wheelchair and electronic bicycles, UX concerning this specific device is a field not many studies have delved into. Therefore, in this project, we seek to propose a data-driven model to explore electronic scooter user's riding profile based on psychophysiological data such as galvanic skin response (GSR) and kinematics data such as the speed and acceleration. Upon retrieving the stress status of the user when he or she is riding, the dataset undergoes a data analysis pipeline that cleans, process and analyse data with Random Forest machine learning algorithms. With the ability to create customised profiles, the model can be adopted to serve the needs of PMD sharing service stakeholders or PMD design companies to ensure good user experience for their customers in the future.

Keywords. Data-driven, User Experience (UX), Personal Mobility Devices (PMD), Psychophysiological Model, Personalised Product Design.

Introduction

Personal Mobility Device (PMD) refers to 'Smart Mobility' or 'Micro Mobility' for one or two persons, and it is appreciated as an environmental-friendly means operated by electric resources [1]. They are commonly regarded as the 'first and last mile' solution, with increasing adoption of 2-wheel electric-scooter (e-scooter) sharing services in many countries. Therefore, user experience (UX) for PMDs has become a prominent feature to encourage the use of PMDs, with ongoing research that are revolving mostly around safety and usability. From the methodological view, user experience from PMD usage was usually evaluated subjectively rather than objectively, based heavily on selfreporting questionnaires rather than unconscious data. Nonetheless, it was shown that objective measurements can be implemented to measure PMD use satisfaction, such as

electroencephalograph (EEG), heart rate invariability (HRV) and galvanic skin response (GSR) [2]. The usage of the psychophysiological method of UX measurement has been used in cases of wheelchairs to predict emotional states and therefore user experience.

Unlike wheelchairs, e-scooters generally demand better motor control abilities of its users especially since the stability of the device depends on how users shift their weight. Acceleration, speed and rate of turn within the six degrees of freedoms simultaneously determines the ability of the user to sustain a comfortable ride that is well-balanced and safe. Each user will consequently have different riding profiles that they are comfortable with since they each have different centres of gravity. Therefore, a holistic solution that improves and customise riding profiles for everyone must be invented to ensure a good user experience.

1. Literature Review

Psychophysiological methods offer objective data during the experiential process, replacing subjective reports that are deemed to be disadvantageous due to their inherent obstructiveness. McCarthy and Wright's [3] theory of enhancement describes user experience as four threads: sensual, emotional, compositional and spacio-temporal. Psychophysiology uses physiological signals that are detected by electrodes attached to the skin, quantifying the psychological state of the subject. Because such measurement is not intrusive nor obtrusive, they are suitable in the case of e-scooters where user experience has to be quantified and measured. Specifically, we are interested in the users' stress level through Galvanic Skin Responses (GSR) during the riding process. GSR also known as electrodermal activity, is regulated by production of sweat in the eccrine sweat glands, where increased sweat gland activity is related to electrical skin conductance level and is associated with physical arousal and often used as an indicator of attention, cognitive and effort [4].

To understand the signal characteristics, statistical correlations of different signal characteristics and/or task-events are often performed to strengthen direct interpretations and/or confirm task response behaviour [5, 6]. Another popular analysis approach is supervised classification, where many relevant signal measures are characterized and used as features to discriminate between two or more internal states [7]. Correlation and supervised classification approaches are effective for classical stimulus-driven controlled tasks, where stimulus versus non-stimulus (i.e. presence of a stressor versus absence of a stressor) periods are defined. It is noted that during complex real-life situations, such as driving, there are no predefined stressor/non-stressor periods. Similarly, everyone may experience stressors differently at different periods of time. Hence, the absolute values is neglected and the deviations from the session average should be used instead.

In the field of psychophysiology behaviour, R-squared values (linear regression) are typically lower than 50%. Manually removing non-linear segments of the time series can result in the introduction of bias into the analysis. Therefore, given the stochastic nature and variability of psychophysiological data, many of the methods used in studies are under the umbrella of machine learning [8]. One of the most essential advantages Random Forest has is its versatility, capable of running both regression and classification tasks. It is an ensemble of Decision Trees, increasing overall result because of the aggregation of learning models, otherwise known as the 'bagging' method. Another benefit that it offers is that random decision forests can mitigate independent decision trees' tendency to overfit to the training set. Cross-validation with questionnaires are

eventually employed to evaluate the accuracy of models compare to actual perceived stress.

To collect psychophysiological data, a biosignal measurement system is required. Many studies employ devices readily available commercially, constructed by established companies. While regarded as reputable and user-friendly, these devices tend to have predefined configurations that cannot be modified for data manipulation. Alternatively, studies have also developed their own stress measurement systems. For example, to maintain their continuous operability and portability, systems based on Zigbee [9] and Arduino platforms have been used. Most psychophysiological measurements have been made inside an indoor laboratory or controlled space. Even for personal mobility wheelchairs [10], it is possible for various equipment to be continuously attached to the user for medical purposes. For the context of e-scooters, these setups can also be intuitively infeasible because they are designed to be compact, thin and obtrusive to the regular e-scooter rider. Nevertheless, there had not been existing studies concerning the measurement of psychophysiological signals of e-scooter users.

To measure the dynamics of the e-scooter while the user is operating, the usage of Inertial Measurement Unit (IMU) should be considered. The IMU is an electronic device that measures the acceleration and angular rate of the body along three axes using a combination of accelerometer and gyroscope. Integration of acceleration data to compute speed is possible. However, depending on the accuracy and precision of the accelerometer, a bias in acceleration becomes a linear error in speed and speed drift is compounded over long time and distances. Additional filtering such as the Kalman filter can be implemented to address these problems for speed computation. With an integral GPS module, the usage of the smartphone is considered for speed measurement. Study has shown that speed measurement using smartphone's GPS data can prove to be a good alternative, using the OBD2 speed measurement of a vehicle as a benchmark [11].

2. Methodology

2.1. Phase 1: Assembly of the stress measurement device

The stress measurement device is assembled using Seeeduino V4.2 as the core microcontroller, which has been in the context of measuring gaming experience producing reliable results [8]. It consists of 8 components, including a 6x AA battery holder to power, the sensors, the shields (Base Shield V2, SD Card Shield V4.0) and a GPS module that allows time syncing of data with the smartphone (Figure 1). The sampling frequency is programmed to be at 2 Hz.



Figure 1. Components of the stress measurement device.

Base Shield V2 allows for neat organisation of the hardware assembly especially since our setup will involve three sensors, namely GSR sensor, HR sensor and IMU. The shield removes the need of breadboard and jump wires, which reduces the complexity of wiring and therefore eliminates potential connection issues. It comprises of the necessary

analog connectors to be used for the GSR sensor which is an analog reader and I2C buses for HR sensor and IMU. One limitation of using a microcontroller is that since it possesses a single processing chip, multithreading will not be possible and only one line of code can be run at any point of time. This will introduce the processing latency during the datalogging, reducing sampling frequency to be less than the stipulated 2 Hz. Nonetheless, the variables used in the study are not time dependent, eliminating the concerns of computations requiring time as part of the function.

2.2. Phase 2: Data collection for e-scooter ride

The experiment involved eight healthy male students aged from 21 to 26 years, with monetary voucher incentives for participation. Participants had little or no experience in riding e-scooter. This is so that greater variances of emotional stress were recorded because of the inexperience in riding. The experiment is divided into two parts, where the first aims to collect a psychophysiological stress template of each participant, while the second aims to collect real-time data during the ride.

First, participants had a resting (non-stress) period of 1 minute. During this minute, participants are also briefed about the task, which is part of the Maastricht Acute Stress Test [12] for inducing psychological stress. It involved backwards counting starting at 1000 in steps of 13 as fast and accurately as possible within the time limit of 3 minutes. The timer was shown to be running on a screen, with the experimenter reminding the participant of time pressure every 30 seconds after the 1-minute mark. Negative feedback by halting the process was also given to participants when they make an error, and participants were forced to start over at 1000.



Figure 2. Predefined routes and the e-scooter used in the experiment.

Next, participants rode the e-scooter in a predefined route scouted by the author. It consisted of various route forms that come in 5 upslopes and downslopes, 2 right-angled turns, 1 roundabout and 2 different types of path (Tartan and concrete). They were meant to encompass a large range of kinematic quantities produced by e-scooter manoeuvres. The distance of the route is approximately 5.6 km, which takes an estimated 20 to 25 minutes per trip completion. The experiment sought to retrieve ride data based on the user's natural riding habit. Participants were asked to keep up with the route navigation personnel. A minimum distance between the navigator and the participant was maintained to minimise ride obstruction.

• The e-scooter maximum motor speed goes up to 25 km/h, reaching a terminal speed of 28 km/h when going down slopes belonging to the route. The stress sensor assembly was attached at the middle of the e-scooter. The IMU was fixed

in the middle of the handles and on the axis parallel to the vertical support of the scooter. The directional vectors of the scooter are illustrated in Figure 2.

- The psychophysiological sensors were attached to the participant and GPS data is also recorded using the smartphone (left with the participant) throughout the ride. They were removed at the end of the ride, and the participant returned to the lab to complete a stress perception questionnaire.
- The stress perception questionnaire sought to retrieve user perceived stress when they experience certain dynamics. These include unexpected acceleration/braking, making turns and moving at high speeds. The results of the questionnaire will be used as a validation for the data model derived in the analysis.

2.3. Phase 3: Data analysis and imputation

This phase is responsible for converting raw data into a clean and ready-to-process dataset, used to establish the relationship between user stress and riding profile with the aid of machine learning algorithm. For this study where inputs are essentially the force and rotation within 6 degrees of freedom coupled with speed data. In order to improve the accuracy of actual test data from e-scooter riding, stress test under a controlled environment will be performed to collect a stress template prior to the test itself. Subsequently, controlled data will serve as the training data required as part of supervised learning methods that will be explored in this subsection. In addition, a partial dependence plot (short PDP or PD plot) will be generated when using Random Forest Regressor to be able to observe how each feature relates to continuous values of GSR. Also, applying K-fold Cross Validation (CV) by dividing the data into folds and ensuring that each fold is used as a testing set at some point.

The identification of erroneous data is done by setting a constraint for both GSR and HR values in the test dataset. The GSR analog value has a negative correlation against stress, while HR value has a positive correlation. Since the calibration test was conducted in a laboratory setting, the first minute of relaxation will record the GSR and HR value when the participant in his most relaxed state. Riding an e-scooter is an activity that requires physical exertion and mental concentration. It is therefore surmised that GSR and HR values will not be able to go beyond a value such that they suggest that the participant is in a greater state of relaxation. These erroneous data are classified as artefacts and given a null value.

The null values that replaced erroneous data will then be imputed using the MICE package in R. Among the several methods of imputation available, the author decided to use predictive mean matching (PMM) to impute these missing values. PMM can impute data that resembles more closely to real values which is what psychophysiological data demands. Empirical evaluations have determined that PMM can produce results that have minimal bias and in greater precision than weaker imputation models [13]. To model against the stress stimulus during the ride, it is important to consider the response time of both HR and GSR signals. GSR and HR response time for a stress stimulus takes up to 5 seconds for onset [14]. A 10-second (to include the possibilities of different response time among different individuals) moving average will therefore be used in this case to smooth out short-term fluctuations and examine more accurately the relationship between the input stimulus and output psychophysiological signals.

For analysis, the use of both random forest classifier and regression algorithms will be considered. For classification, stress template retrieved in the first part of data collection will be labelled into binary values, where 0 is non-stressed and 1 is stressed. This set of data will be used as a training set, in which the features are equivalent to HR and GSR values. This data is further split up randomly into train: test in the ratio of 3:1 to verify the trained model against the test model for robustness. Once the model is trained, it will be applied to predict the stress value according to the same features in the data from the ride. The prediction values (0 or 1) will be appended to the dataset. The second set of data will then undergo the same process of splitting up to 3:1 ratio and training again. In this iteration, the classifier score, feature importance and finally a partial dependence plot of each feature will be determined.

For regression, first we pass the features and the GSR values of the dataset to the method created for the random forest regression model. We then use the grid search cross validation method from the sklearn library to determine the optimal values to be used for the hyperparameters of our model from a specified range of values. Here, we have chosen the two hyperparameters; max_depth and n_estimators, to be optimized. According to sklearn documentation, max_depth refers to the maximum depth of the tree and n_estimators, the number of trees in the forest. Ideally, more trees translate to better modelling. After creating a random forest regressor object, we pass it to the cross_val_score() function which performs K-Fold cross validation on the given data and provides as an output, an error metric value, which can be used to determine the model performance. Here we have used 10-fold cross validation where the negative Mean Absolute Error (MAE) is taken as the error metric (specified using the scoring parameter) to measure model performance. The lower the MAE is, the better. Feature importance and partial dependence plot will also be captured.

3. Experimental Results and Discussion

The goal of this study was to design a data analysis method to model user experience in terms of psychological stress during a ride. The stress values labelled as part of the calibration are binary. The independent variables in our modelling consist of 7 features extracted from our kinematic sensors and including speed data based on our smartphone's GPS. To represent it as a regression model, we seek to predict GSR values that will be as close to the real values as possible given a certain set of input values. To represent it as a classification model, we seek to predict stress (0 or 1) based on instantaneous HR and GSR and explore each independent variable's importance in predicting stress. The predicted class will also be matched up against the actual class to determine the accuracy of the classification model. To illustrate what is discussed, the result of a random participant will be used below as an example.

3.1. Random Forest Regressor

For the regression model, the K-fold cross validation method provides us with the optimal hyperparameters, max_depth and n_estimators, such that MSE is minimised (Figure 3). These hyperparameters are then replicated in the RandomForestRegressor and trained to determine the robustness of the model. It is easy to derive that max_depth has a positive relationship with regression score (Figure 4). This is mostly attributed to the fact that the greater the max_depth, the closer the model fits to existing data. One can

increase the max_depth indefinitely to improve on modelling score. However, this introduces the problem of overfitting, models itself so closely on existing data, it fails to consider anomalous data and therefore models according to them as illustrated on the right model. This results in the inability of the model to run on new datasets to give a similarly 'good' score. K-fold cross validation prevents this from happening by tuning the hyperparameters enough to optimise the model by validating k subsets of data against each other.





Figure 3. Modelling GSR and predicted GSR with cross-validated hyperparameters.

Figure 4. Modelling GSR and predicted GSR using score-maximising hyperparameters.

Table 1 shows the feature importance of the model determines and how much each prediction is attributed to each independent variable. It can be observed that Z-axis acceleration which is basically gravity acting on IMU, is of least importance. X-axis rate of turn, which is the rolling of the e-scooter does not significantly affect the stress level of this participant as well. The larger effectors of the participant's emotional state are speed, Z-axis rate of turn (yaw/turning of the e-scooter direction) and X-axis acceleration (braking/acceleration). The relationship can be validated as we are comparing these features according to the stress perception questionnaire done by the participant at the end of the experiment.

Table 1. Feature importance.	
Variable	Feature Importance
X Acceleration	0.158
Y Acceleration	0.120
Z Acceleration	0.045
X rate of turn	0.068
Y rate of turn	0.112
Z rate of turn	0.165
Speed	0.332

The partial dependence plot also reveals how each variable varies with GSR while keeping all other variables constant (Figure 5). The red triangles illustrated attempts to highlight the inverse relationship between GSR and the variables. In these graphs, GSR is in analog form and in units of Ohm. Higher GSR translates to greater resistance, lower skin conductance and therefore lower psychophysiological stress. The partial dependence plots show that GSR tends towards the minimum as it moves towards the left and right of the equilibrium point. This could serve as a potential diagnostic tool to design personalised riding profiles based on user stress minimisation. Through the

Regressor model, we can establish the relationship between the features and GSR value in terms of the model scores, feature importance and partial dependence plots. Removal of features could potentially improve the modelling as well.



Figure 5. Partial dependence plots of each variable.

3.2. Random Forest Classifier

For classification, the accuracy metric is defined as the percentage of the correctly predicted samples among the two classes. Like Regressor model, K-fold cross validation is used for hyperparameter tuning. The scoring method for cross validation is changed to AOC and cross validation scores ranged from 0.463 to 0.972. With the optimal max_depth and n_estimators at 4 and 100 respectively, the RandomForestClassifier score is at 0.9123.



Figure 6. Confusion matrix with 2 labels.

The confusion matrix depicts how many times the predicted label matches the true label (Figure 6). Accuracy of 541 out of 591, the number of true positive is 535, true negative is 6 and false positive is 52. This model is therefore skewed towards making predictions that the user is stressed. Naturally, given a greater max_depth as with in Regressor will allow it to get a higher accuracy score. However, the author actively seeks to prevent overfitting since the data is understood to have existing artefacts. The score of 0.9123 is satisfactory and cross-validation when k = 10 has accuracy ranging from 0.463 to 0.972 with a standard deviation and mean of 0.148 and 0.723 respectively. Binary classification model of stress can therefore be evaluated as satisfactorily predictable. While both regressor and classification models have similarly good performances in predicting user stress when riding the e-scooter, they can be further optimised with other configurations unemployed in this study.

4. Conclusions

The studies and findings from this research have proven the added potential of using psychophysiological data to model user experience for different users. Machine learning methods Random Forest Regressor and Classifier have proven themselves to be suitable models to establishing the relationship between the kinematic capabilities of the escooter and psychophysiological signals that translate to user stress. The use of k-fold cross-validation method allowed hyperparameter tuning to assure the robustness and optimisation of the learnt model while avoiding the overfitting of the model for futureproofing against new data. This pipeline of data analysis, starting from raw data collection, cleaning, processing and finally analysis can be easily replicated under different contexts to review in future psychophysiological studies for the usage of escooter or even other forms of PMDs.

Using this study, future applications develop from the possibility of incorporating such portable technology for real-time usage. One area of implementation would be for e-scooter sharing services. Embedded sensors can be incorporated in rental e-scooters to provide suitable and comfortable intervention when the system senses heightened stress levels from the user. This is made possible by IoT solutions by uploading said data to cloud servers like Amazon Web Services which can then undergo a backend stress data analysis before sending suitable intervention instructions to ensure a better and safer user experience. Alternatively, the study can be utilised by businesses where e-scooter designs can be morphologically charted with specifications that translates to the kinematic quantities described (motor power, stiffness of rotation etc.). With sufficient data, businesses can determine the optimal range of kinematic settings that their escooters should offer to appeal to the majority of the consumers. The ease and reliability of extracting psychophysiological signals with embedded and cloud technology will therefore transform the future of human factors engineering in the field of personal mobility device.

Acknowledgement

The authors would like to thank the support of Schaeffler Hub for Advanced Research at Nanyang Technological University (SHARE at NTU).

References

- [1] G. W. Shin, K.-J. Lee, D. Park, J. H. Lee and M.H. Yun, Personal Mobility Device and User Experience: A State-of-the-art Literature Review, *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, 2018, Vol. 62, no. 1, pp. 1336-1337: SAGE Publications Sage CA: Los Angeles.
- [2] J.T. Cacioppo, L.G. Tassinary and G. Berntson, *Handbook of psychophysiology*. Cambridge University Press, Boston, 2007.
- [3] J. McCarthy and P. Wright, *Technology as experience*. MIT press, Boston, 2007.
- [4] D.T. Lykken and P.H. Venables, Direct measurement of skin conductance: A proposal for standardization, *Psychophysiology*, Vol. 8, No. 5, pp. 656-672, 1971.
- [5] R. L. Mandryk, K. M. Inkpen, and T. W. Calvert, Using psychophysiological techniques to measure user experience with entertainment technologies, *Behaviour & information technology*, Vol. 25, No. 2, pp. 141-158, 2006.
- [6] C.P. Latha and M. Priya, A review on deep learning algorithms for speech and facial emotion recognition, *APTIKOM Journal on Computer Science and Information Technologies*, Vol. 1, No. 3, pp. 88-104, 2016.
- [7] R.L. Diaz, U. Wong, D.C. Hodgins, C.G. Chiu, and V.M. Goghari, Violent video game players and nonplayers differ on facial emotion recognition, *Aggressive behavior*, Vol. 42, No. 1, pp. 16-28, 2016.
- [8] M. Čertický, M. Čertický, P. Sinčák, G. Magyar, J. Vaščák, and F. Cavallo, Psychophysiological Indicators for Modeling User Experience in Interactive Digital Entertainment, *Sensors*, Vol. 19, No. 5, p. 989, 2019.
- [9] M. V. Villarejo, B. G. Zapirain, and A. M. Zorrilla, A stress sensor based on Galvanic Skin Response (GSR) controlled by ZigBee, *Sensors*, vol. 12, no. 5, pp. 6075-6101, 2012.
- [10] J. Abdur-Rahim *et al.*, Multi-sensor based state prediction for personal mobility vehicles, *PloS One*, vol. 11, no. 10, e0162593, 2016.
- [11] A. Chowdhury, T. Chakravarty and P. Balamuralidhar, Estimating true speed of moving vehicle using smartphone-based GPS measurement, 2014 IEEE International Conference on Systems, Man, and Cybernetics (SMC), 2014, pp. 3348-3353: IEEE.
- [12] T. Smeets, S. Cornelisse, C. W. Quaedflieg, T. Meyer, M. Jelicic, and H. Merckelbach, Introducing the Maastricht Acute Stress Test (MAST): A quick and non-invasive approach to elicit robust autonomic and glucocorticoid stress responses, *Psychoneuroendocrinology*, Vol. 37, No. 12, pp. 1998-2008, 2012.
- [13] L. R. Landerman, K. C. Land, and C. F. Pieper, An empirical evaluation of the predictive mean matching method for imputing missing values, *Sociological Methods & Research*, Vol. 26, No. 1, pp. 3-33, 1997.
- [14] W. Boucsein, *Electrodermal activity*, Springer Science & Business Media, 2012.