Predicting Online Shopping Search Satisfaction and User Behaviors with Electrodermal Activity

Yingying Wu[‡†], Yiqun Liu[†], Ning Su[†], Shaoping Ma[†], Wenwu Ou[♮]
[‡]Department of Mathematics, The University of Texas at Austin, USA
[†]Tsinghua National Laboratory for Information Science and Technology,
Department of Computer Science and Technology, Tsinghua University, Beijing 100084, China
[♮]Alibaba Group
yiqunliu@tsinghua.edu.cn

ABSTRACT

Electrodermal activity (EDA) delineates positive and negative emotions, especially in the lower arousal range, which reflects small variations. This study derived formulas to extract features from the EDA graph and found they are effective to predict search satisfaction and explain mobile shopping behaviors including *add to cart, purchase,* and *abandonment* (added to cart without purchase). The best features found in the generalized linear mixed model with the binomial family agree with known physiological results. To the best of our knowledge, this study is the first to use physiological methods to study satisfaction and user behavior.

Keywords

Mobile product search; Satisfaction; Physiological signals

1. INTRODUCTION

Studies show that humans have lower electrodermal activity (EDA) levels with positive emotions and higher EDA with negative ones [2] and that EDA can reveal unnoticeable emotions in browsing experiences [1]. To use the advantages of EDA to compensate for the limitations of current methods of studying search satisfaction and to study mobile shopping behavior, we defined the discretized slope and curvature for plane curves to characterize EDA graphs, based on the curvature of discrete surfaces by Gu and Yau [3]. We found that these characterizations are effective, which can be explained by known physiological results. We found that the features fit the *abandonment* scenario (merchandise added without purchase) best, and the negative slope and the minimum EDA are the best predictors. Linear regression performs well using these features. This study makes three contributions. First, we constructed an experimental shopping search system on mobile phones that collects users' satisfaction feedback and shopping behaviors with physiological signals recorded. Second, we derived mathematical formulas to describe the EDA graphs. Third, to the best of our knowledge, this attempt is the first to use physiological methods to study search satisfaction and shopping behavior.

©2017 International World Wide Web Conference Committee (IW3C2), published under Creative Commons CC BY 4.0 License.
WWW'17 Companion, April 3–7, 2017, Perth, Australia.
ACM 978-1-4503-4914-7/17/04.
http://dx.doi.org/10.1145/3041021.3054226

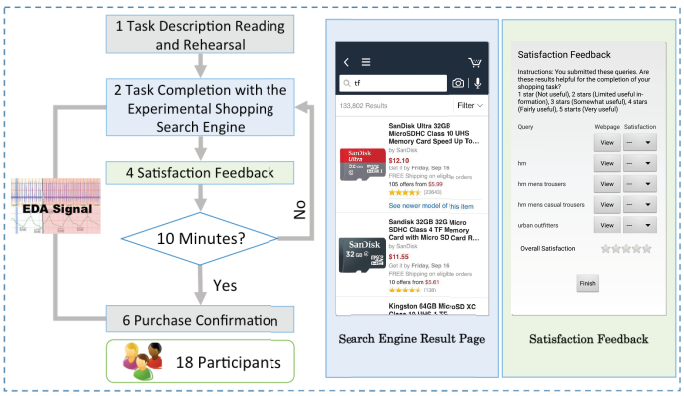

Figure 1: Experimental procedure

2. DATA COLLECTION AND ANALYSIS

We recruited 18 undergraduate students to perform online shopping tasks on a mobile phone with EDA signals recorded, as Fig. 1 shows. Participants submitted 178 shopping queries and added 85 items to their carts. They confirmed purchasing 47 of the 85 added goods at the end.

2.1 Definitions and Features

Consider the graph $(x, f(x))$ of a differentiable function f, and take the natural discretization given by the system, where $(x_i, f(x_i))$ is the EDA level $f(x_i)$ at time x_i.

Definition 1 (Average Slope)**.** *The average slope is the integration of the absolute value of the slope divided by the length of the interval:*

$$\bar{k}_{x_0,x_n} = \frac{1}{x_n - x_0} \int_{x_0}^{x_n} |k(x)| \,.$$

Definition 2 (Discrete Slope)**.** *Consider a continuous function $F(x_i) = Y_i$, which is the discretization of f. Define the discrete slope of the graph by forward differencing:*

$$K(x_i) = \frac{Y_{i+1} - Y_i}{x_{i+1} - x_i}.$$

Feature 1 (Discrete Average Slope)**.** *The discrete average slope is defined as the average of the absolute value of discrete slopes, where n is the number of points:*

$$\overline{K}_{x_0,x_n} = \frac{1}{n-1} \sum_{i=0}^{n-2} |K(x_i)| \,.$$

Table 1: Odds ratio between variables and dichotomized *satisfaction, add to cart, purchase* and *abandonment* (**: significant at 0.01, *: significant at 0.05, .: significant at 0.1)

	aveSlope	posSlope	negSlope	aveCurv	posCurv	negCurv	maxCurv	aveEDA	maxEDA	minEDA
Satisfaction	0.350.	1.932.	2.139**	–	–	–	–	–	1.268*	–
Add to Cart	–	–	–	0.347.	–	1.659.	–	2.067*	–	0.424**
Purchase	–	–	–	–	–	–	–	–	–	0.612.
Abandonment	–	–	–	0.186.	2.767*	–	0.736*	2.112.	1.359.	0.292**

Feature 2 (Discrete Positive Slope)**.** *The discrete positive slope is defined as the average of the positive discrete slopes:*

$$\overline{K}^{+}_{x_0,x_n} = \frac{1}{n^+} \sum_{K(x_i)>0} K(x_i),$$

where n^+ is the number of points with positive slopes.

Feature 3 (Discrete Negative Slope)**.** *The discrete negative slope is defined as the average of the negative discrete slopes:*

$$\overline{K}^{-}_{x_0,x_n} = \frac{1}{n^-} \sum_{K(x_i)<0} K(x_i),$$

where n^- is the number of points with negative slopes.

Definition 3 (Curvature)**.** *Consider a twice differentiable plane curve $y = f(x)$, and denote the first- and second-order derivatives as $f'(x)$ and $f''(x)$. The curvature is defined as*

$$\kappa(x) = \frac{f''(x)}{(1 + f'(x)^2)^{3/2}}.$$

Definition 4 (Average Curvature)**.** *Denote the curvature of a graph by $\kappa(x)$. The average curvature is the integration of the absolute value of the curvature divided by the interval:*

$$\bar{\kappa}_{x_0,x_n} = \frac{1}{x_n - x_0} \int_{x_0}^{x_n} |\kappa(x)|.$$

Definition 5 (Discrete Curvature)**.** *Let $F(x_i) = y_i$ be the discretization of a twice-differentiable function f. Define the discrete curvature by second-order forward differencing:*

$$\mathcal{K}(x_i) = \frac{K'(x_i)}{(1 + K(x_i)^2)^{3/2}}, \text{ where } K'(x_i) = \frac{K(x_{i+1}) - K(x_i)}{x_{i+1} - x_i}$$

and $K(x_i)$ is the discrete slope of $f(x)$ at x_i.

Feature 4 (Discrete Average Curvature)**.** *The discrete average curvature is the average of the absolute value of the discrete curvature, where n is the number of points:*

$$\overline{\mathcal{K}}_{x_0,x_n} = \frac{1}{n-1} \sum_{i=0}^{n-2} |\mathcal{K}(x_i)|.$$

Feature 5 (Discrete Positive Curvature)**.** *Discrete positive curvature is the average of the positive discrete curvature:*

$$\overline{\mathcal{K}}^{+}_{x_0,x_n} = \frac{1}{n^+} \sum_{\mathcal{K}(x_i)>0} \mathcal{K}(x_i),$$

and n^+ is the number of points with positive curvatures.

Feature 6 (Discrete Negative Curvature)**.** *The discrete negative curvature is defined as the average of the negative discrete curvature:*

$$\overline{\mathcal{K}}^{-}_{x_0,x_n} = \frac{1}{n^-} \sum_{\mathcal{K}(x_i)<0} \mathcal{K}(x_i),$$

and n^- is the number of points with negative curvatures.

We calculated these six features defined above along with the maximum, minimum, and average EDA and the maximum discrete curvature.

2.2 Analysis

We used these features to predict dependent variables including dichotomized satisfaction and shopping behaviors using the generalized linear mixed model with the binomial family. Table 1 reports the exponentiated coefficients of significant variables; we found *abandonment* has the most significant predictors. In our experiment, users purchased 66.67% of added items of merchandise but 21.97% of the total value. The negative slope and the minimum EDA are the predictors with the smallest p-values. We found that a highly negative slope positively correlates with a satisfying search experience, and a high minimum EDA negatively correlates with satisfaction, agreeing with physiological results that EDA decreases during positive emotions and increases with negative ones [2]. EDA is an indicator of the arousal index, which explains some of the other correlations.

Linear regression performs well with the aforementioned features, especially in predicting *add to cart,* as Table 2 shows. This result suggests that *add to cart* reflects an impulse to purchase, which is captured by the change of EDA. We performed 10-fold cross-validation and compared the results for each task with chance using a linear mixed model with fixed effects for *tasks* and the AUC of linear regression *vs.* chance. The chance fixed effect is significant with $F(1, 66) = 104.61, p < 3.11 \times 10^{-15}$.

Table 2: Satisfaction and Shopping Search Behavior Prediction Results

	Satisfaction	Add to Cart	Purchase	Abandon
AUC	0.62	0.80	0.76	0.73

3. CONCLUSION

This study validates the effectiveness of features extracted from EDA graphs to fit and predict satisfaction and user behavior in a mobile shopping environment. Experiments revealed three points. First, the features fit the *abandonment* scenario best. Next, the negative slope and the minimum EDA are the most effective indicators of satisfaction and shopping behaviors. Third, predictions made by linear regression are significantly better than chance, and predicting *add to cart* achieves the best result.

4. ACKNOWLEDGMENTS

Wu was supported in part by NIH grant 5U54CA193313; this work was supported by Natural Science Foundation of China (61622208, 61532011, 61672311) and National Key Basic Research Program (2015CB358700).

5. REFERENCES

[1] M. Barreda-Ángeles, I. Arapakis, X. Bai, B. Cambazoglu, and A. Pereda-Baños. Unconscious physiological effects of search latency on users and their click behaviour. In *SIGIR*, pages 203–212, 2015.

[2] Wolfram Boucsein. *Electrodermal activity.* Springer Science & Business Media, 2012.

[3] Xianfeng David Gu and Shing-Tung Yau. In *Computational conformal geometry*, pages 238–239. International Press Somerville, MA, USA, 2008.