Computer Usage and Multimedia Induced Emotion Relationship

AUTHOR:
Rayhan Shikder
0905060

SUPERVISOR:
Dr. A.B.M. Alim Al Islam
Assistant Professor,
Department of CSE, BUET

October 2015

DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING
BANGLADESH UNIVERSITY OF ENGINEERING AND TECHNOLOGY

Abstract

Devices, capable of detecting human emotion and interacting accordingly is an important part of building intelligent computers. Emotionally-aware systems will be able to make appropriate decisions about how to interact with the user or adapt their response. There are two main problems with current system approaches for identifying emotions that limit their applicability: they demand additional infrastructures which are often intrusive and require specialized information which are not always available. We experimented on detecting emotion by analyzing the pattern of their keyboard and mouse usage parameters. We conducted a field study where we tried to identify users different emotional states, users identity in many different ways by using two types of different classifiers. Our best result shows good result in identifying 3-class emotion, dominant user identity with data from only one emotional state.

Acknowledgements

Farzia Afroze, Kazi Zubayer Hasan, Saiyma Sarmin and all the participants of the survey.

Contents

List of Figures

List of Tables

Abbreviations

FAR	**F**alse **A**ccept **R**ate
FRR	**F**alse **R**eject **R**ate
knn	**k** **n**earest **n**eighbor

Chapter 1

Introduction

Emotion is, perhaps, the most critical attribute of living beings that is critical to detect and generate artificially. Its detection always remains a classical well-explored problem.Emotionally-aware systems would have a rich context from which to make appropriate decisions about how to interact with the user or adapt their system response. There exist many approaches for determining human emotions based on facial expression analysis [1], thermal imaging of faces [2], gesture and pose tracking [3], voice intonation [4], etc.We conducted a field study where we collected participants keystrokes, mouse usages and their emotional states via a survey system by inducing emotion through multimedia components.From this data, we extracted important features, and created classifiers for 10 emotional states. Our top results include classification of 3-level emotion(positive emotion, negative emotion and neutral emotion), 4-class dominant emotion(amusement, surprise, anger, disgust), dominant user classification etc.

Chapter 2

Related Work

Modeling affective state using typing rhythms draws from two fields: affective computing and keystroke dynamics.

2.0.1 Affective Computing

Affective computing refers to computing that relates to, arises from, or deliberately influences emotions [5]. We are interested in identifying a users emotional state, so we must first consider how emotions are described, and what other approaches have been used to classify emotion. The terms affect and emotion are often used interchangeably; we will use emotional state to refer to the internal dynamics (cognitive and physiological) that are present during an emotional episode, and emotional experience as what an individual perceives of their emotional state [5].

2.0.1.1 Describing Emotions

Two main approaches have been used to describe emotions: categorical and dimensional. The categorical approach applies specific labels to different emotional states through language (e.g. sadness, fear, joy) [6]. The dimensional approach [7] uses two orthogonal axes called arousal and valence. Arousal is related to the energy of the feeling and is typically described in terms of low (e.g. sleepiness) to high (e.g. excitement) arousal. Valence describes the pleasure (positive) or displeasure (negative) of a feeling. Labels for different emotional states can be represented in this two-dimensional space. For example, anger would be a high-arousal, low-valence state.

2.0.1.2 Sensing Emotional State

Both the categorical and dimensional models of emotion have been used in prior approaches of identifying emotional state. Some approaches use features easily discernable by other humans, such as facial expressions, gestures, vocal intonation, and language [5]. For example, face-tracking software is used to analyze facial expressions gathered from webcam images to infer users affective states [8, 10]. This approach has been extended to use thermal imaging to identify changes in blood flow patterns of the face that are synonymous with different facial expressions [9].

Other approaches use features that are less discernable to other humans, but can be measured by specialized equipment. For example, significant research has been conducted on measuring physiological changes that occur in the body during emotional episodes using sensors such as galvanic skin response, electromyography of the face, and heart activity (see [11] for an overview). In HCI, researchers have used physiological sensors to measure the affective state of a user interacting with technology. Results have been produced by studying users playing video games [12], navigating web pages [13], using video conferencing software, and using mobile technology [14].

The above approaches have two main problems that prevent their widespread use: the sensing technology is obtrusive, and requires expensive specialized equipment. For example, EKG is measured using electrodes attached directly to the users skin. In some cases, the area where the electrodes are placed needs to be shaved to prevent interference [15]. Although research is underway to integrate these sensors into interaction devices, they are currently intrusive and their mere presence may alter the users emotional state. In [9], a thermal camera is used is measure blood flow to a users face. Although unobtrusive, the equipment is specialized and not found in typical home or office settings. To eliminate the need for intrusive and costly equipment, we propose to determine affective state via typing rhythms.

2.0.2 Keystroke Dynamics

Keystroke dynamics is the study of the unique timing patterns in an individuals typing, and typically includes extracting keystroke timing features such as the duration of a key press and the time elapsed between key presses.

Much of the previous research in keystroke dynamics has been in authentication systems, with the premise that, just as with handwritten signatures, the way that an individual types can be unique enough to identify them [16]. The use of keystroke dynamics for user authentication has been an active area of research, producing many studies [16, 17, 23**?**], patents [19], and systems [20], whereby users are authenticated by providing the correct user name, password, and typing rhythm (see [21] for an overview). Anecdotal evidence suggests that strong emotional states can interfere with authentication [18]; however, little is mentioned of this and it is unclear whether the timing variance associated with these emotional states is similar between individuals.

Most of the authentication systems [16, 18, 23] use fixed-text models that is, they use the same static piece of text (entered during authentication) that the model was trained on. There have been fewer approaches [17, 18, 22] that use models based on free text (text that is not prescribed to the user), as they do not perform as well as fixed-text models [18]. The length of the required training text varies between different studies; some require a few words [23] or full pages of text [24], which can create better performing models.

Although fixed-text models generally perform better than free-text models, the potential applications of free-text models are desirable. Recent work has explored free-text models for use in continuous verification, where users are continually monitored to identify masqueraders at any time (not just during authentication), and have shown potential given enough samples of sufficient length [22]. Free-text models have even been able to identify individuals typing in different languages [25] as long as the two languages have enough similar valid digraphs. Most free-text studies require users to enter any valid text as sample text [22]; however, in [17] keystroke activity was monitored as a background process during normal computer use. This method had three benefits: the user was less disturbed by the collection method, the data was obtained unobtrusively, and it reduced the cognitive load on the user by avoiding situations where they must think of something to type.

Classification algorithms for the analysis of keystroke dynamics for user authentication include neural networks [26], distance measures [16, 18], decision trees [27], and other statistical methods [17, 18, 23]. Due to the differences in data collection approaches and classification methods, a comparison of performance across studies is difficult [23].

2.0.2.1 Keystroke Dynamics and Affective Computing

There has been very little previous work applying keystroke dynamics to affective computing. Zimmerman et al. [29] describe a method to correlate user interactions (keyboard and mouse) with affective state. Affective states were induced using films. Physiological sensors were used in conjunction with the Self-Assessment Manikin (SAM) [28], a method of subjectively expressing affective state. The authors found significant differences between the neutral state and other emotional states, but were unable to distinguish between the induced states. Recent work by Vizer et al. [30] used keystroke timing features of free text in conjunction with linguistic features to identify cognitive and physical stress. They achieved correct classifications of 62.5for cognitive stress (for 2 classes), which they state is comparable to other affective computing solutions. They also state that their solutions should be tested with varying typing abilities and keyboards, with varying physical and cognitive abilities, and in real-world stressful situations.

Chapter 3

Motivation behind our work

Existing emotion detection approaches usually demand additional infrastructures such as webcam, body mounted hardware, etc., which are often intrusive. This approaches also require specialized information such as voice, gestures, facial expression, etc., which are not always available. We propose a novel emotion detection system that detects emotion from widely-used electronic devices demanding no additional infrastructures, and exploiting conventionally available usage data.

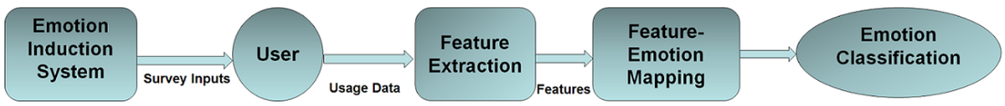

FIGURE 3.1: Our proposed approach for emotion detection

Chapter 4

Emotion Induction

We used movie clips or short films of around 2 minutes length to induce emotion into users. There were related still pictures and texts also. The videos, still pictures and texts were selected after reviewing by around 10 individuals. All the videos, still pictures and texts were presented among 10 persons and they were asked to tell their emotional states after watching each video, still picture and text. We developed a induction and survey system for our study, where we presented these selected videos, images and texts to induce certain emotions and then prompted the user with some related questions. To answer the questions user have to use both keyboard and mouse of the computer.

Chapter 5

Tracking usage parameters

While answering the question, users certain usage attributes are logged automatically in every 5 seconds. There are 17 attributes that are logged. The attributes are average mouse left click, average mouse right click, average mouse double click, average mouse scroll, average cursor x-distance, average cursor y-distance, average key down to up time, average key up to down time, average key down to down time, average regular key press, average enter key press, average arrow key press, average backspace key press, average function key press etc. We also collected some meta data from the user. These are the name, age, occupation and birth year of the survey participant.

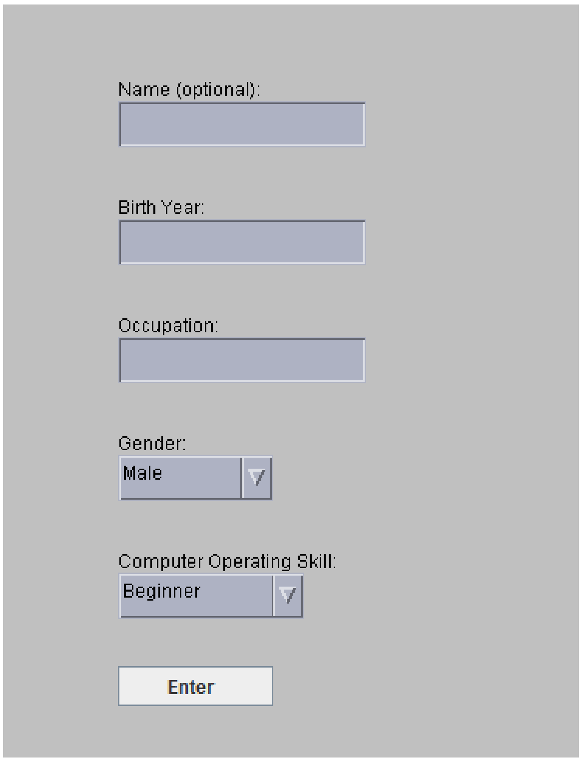

FIGURE 5.1: Start page of the survey

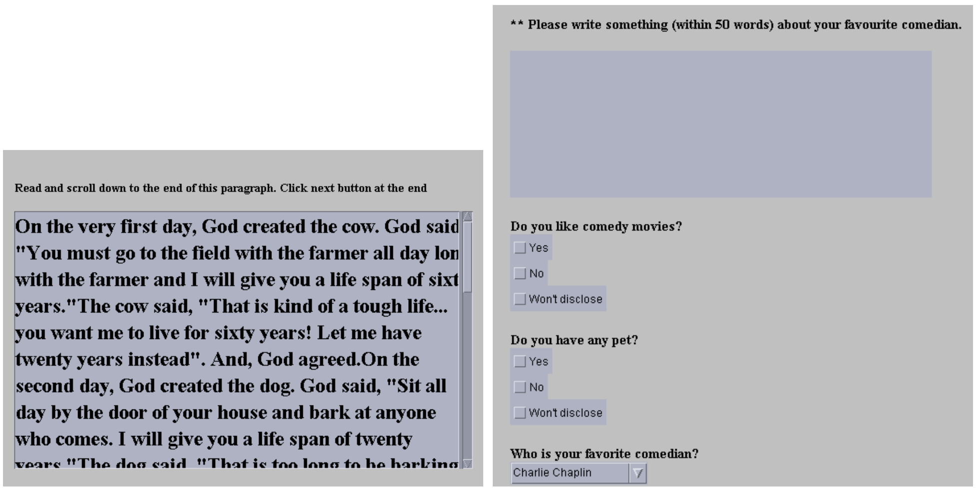

FIGURE 5.2: Text and various questions to capture scrolling and other usage attributes

Chapter 6

Emotion Usage Relationship

should write something here

6.0.3 Demography of participants

There were a total of 26 participants. 19 of them were male and 7 were female. Most of them were students and some are from other occupations.

6.0.4 Classification methods

6.0.4.1 Bounded k-means clustering

In these approach we used a variant of k-means clustering approach, where we put a constraint on the cluster size. The cluster size was varied to find out the optimal cluster size so that the total false acceptance and false rejection are minimized. The size was varied with percentages of mean and standard deviation.

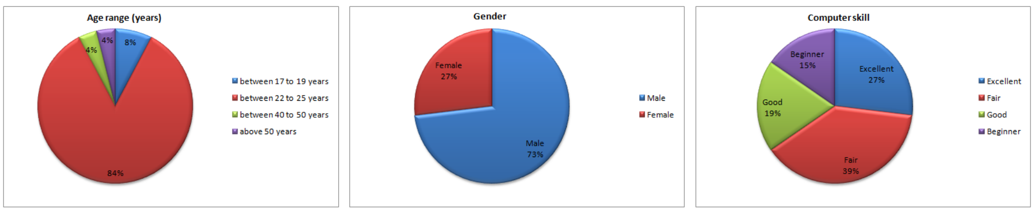

FIGURE 6.1: Category of users based on age range, gender, and computer operating skill

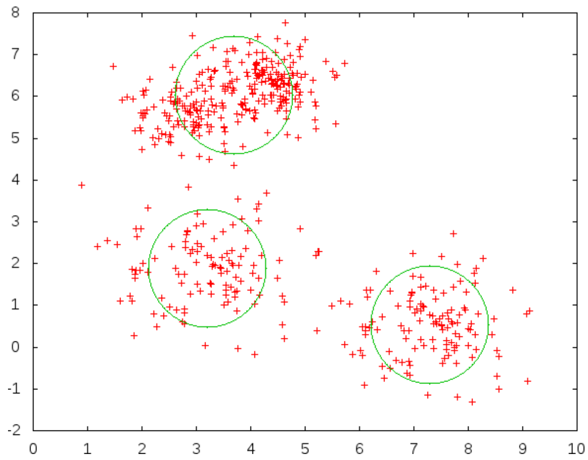

FIGURE 6.2: bounded k-means clustering

method	false accept rate	false reject rate
bounded k-means	0.43	0.34
knn	0.76	0.0

TABLE 6.1: bounded k-means and knn for 10 class emotion classification

6.0.4.2 k-nearest neighbor

We used this well known classification method where a test point is classified by examining its neighbor points. We varied both the number of neighbors and the number of attributes.

6.0.5 Identification of emotional states

At first we tried to classify human emotions based on the usage attributes we captured in each step of the survey process. Every captured data is like the following format: Att1, att2, att3, , emotion label, user label.

6.0.5.1 10 classes of emotions

We tried to classify 10 different emotional states. These are amusement, happiness, inspiration, surprise, sadness, sympathy, anger, disgust, fear and neutral emotion. We tried two different classification approaches: bounded k-means clustering and k-nearest neighbor method.

Emotion classes	C-1	C-2	C-3	C-4	C-5	C-6	C-7	C-8	C-9
C-1	23.36	5.14	15.42	13.08	11.21	14.49	9.35	2.34	5.61
C-2	18.67	12.05	12.95	14.46	9.64	6.63	9.04	6.93	9.64
C-3	19.43	9.87	16.4	12.58	8.92	6.05	14.97	5.1	6.69
C-4	14.29	15.48	8.33	28.57	10.71	8.33	4.76	7.14	2.38
C-5	18.52	14.81	7.41	11.11	22.22	0	18.52	3.7	3.7
C-6	17.39	10.28	6.72	11.86	13.04	9.49	15.81	9.49	5.93
C-7	12.97	5.95	14.86	15.68	12.43	10.27	18.92	4.59	4.32
C-8	13.31	9.73	10.87	11.71	9.96	10.87	12.17	11.03	10.34
C-9	16.24	11.12	9.94	12.97	8.71	8.96	11.74	8.34	11.98

FIGURE 6.3: Matrix of domination for 9-emotion classes(without the neutral emotion)

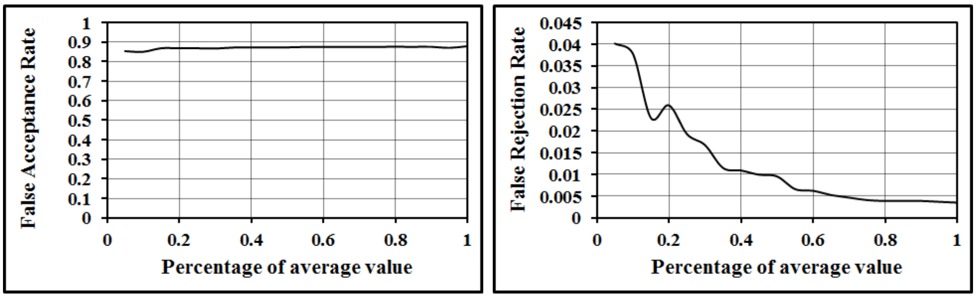

FIGURE 6.4: percentage of average vs false accept rate and false reject rate in bounded k-means for 10 class emotion classification

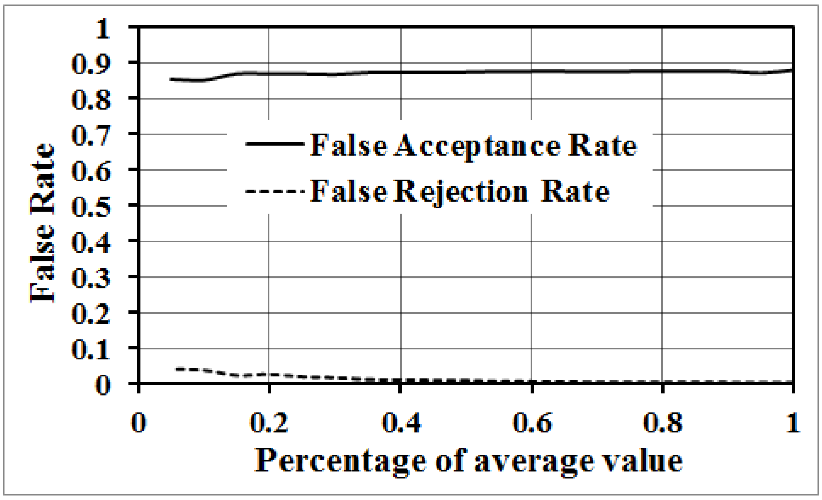

FIGURE 6.5: percentage of average vs false accept rate and false reject rate in bounded k-means for 10 class emotion classification

method	false accept rate	false reject rate
bounded k-means	0.25	0.17
knn	0.47	0.0

TABLE 6.2: bounded k-means and knn for 3 class emotion classification

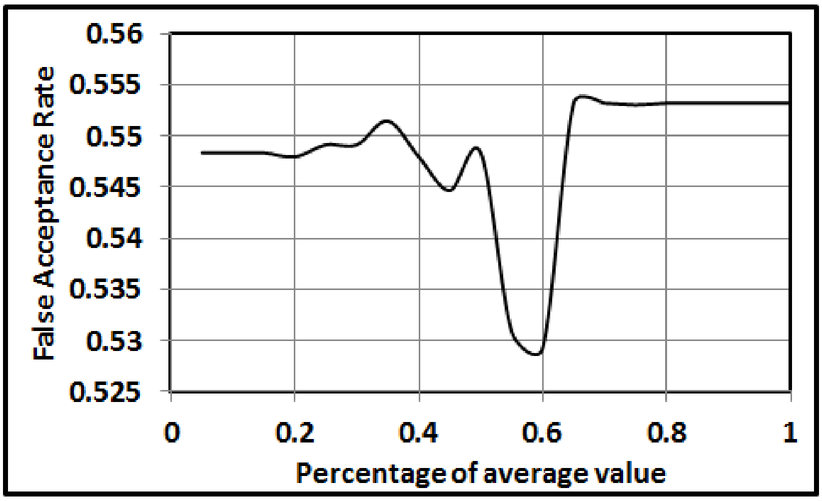

FIGURE 6.6: percentage of average vs false accept rate and false reject rate in bounded k-means for 3 class emotion classification

6.0.5.2 3 classes of emotions

We grouped the 10 classes into 3 groups. The positive emotions such as amusement, happiness, inspiration and surprise were grouped into one class. The negative emotions such as sadness, sympathy, anger, disgust and fear were grouped into one class. Neutral emotion was considered as the third class.

classes	Amusement	Surprise	Sadness	Anger
Amusement	36.31	25.34	16.75	21.60
Surprise	28.83	31.08	21.62	18.47
Sadness	27.60	25.04	22.78	24.59
Anger	24.04	22.70	18.43	34.83

TABLE 6.3: matrix of cluster with only dominant emotions

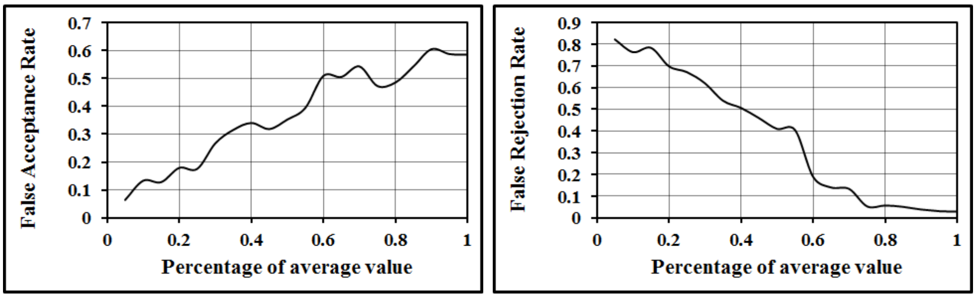

FIGURE 6.7: percentage of average vs false accept rate and false reject rate in bounded k-means for dominant emotion classfication

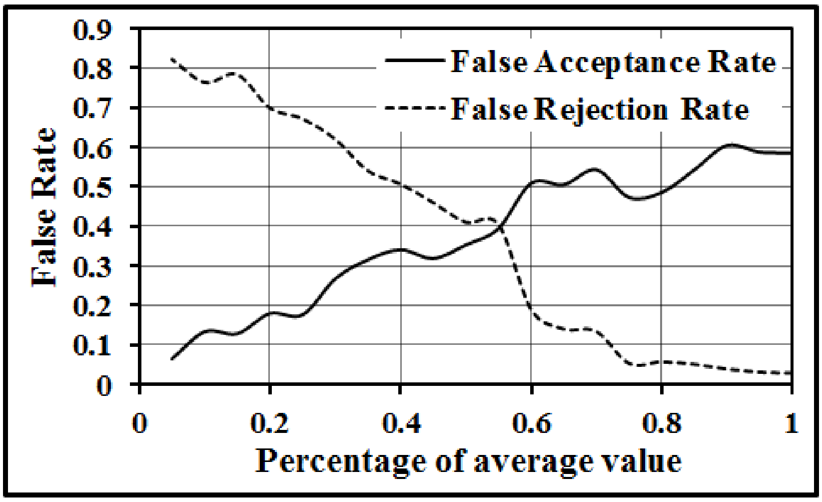

FIGURE 6.8: percentage of average vs both false accept rate and false reject rate in bounded k-means for dominant emotion classfication

6.0.5.3 Dominant classes of emotions

From the unsupervised bounded k-means clustering of 10 emotions. We found 10 different clusters, but this clusters were not dominated by only one class of emotion. Rather we found that there are some emotions which dominates not only in his own cluster but also in the clusters of others. While trying to classify these dominant clusters only we found some improvement. The experimental results are as followed from bounded k-means clustering and k-nearest neighbor classification:

classes	c-1	c-2	c-3	c-4	c-5
c-1	26.32	26.32	16.67	12.72	17.98
c-2	18.69	38.13	17.93	11.62	13.64
c-3	21.31	21.31	26.23	17.49	13.66
c-4	18.34	20.92	20.63	20.77	19.34
c-5	22.09	19.75	17.79	16.56	23.80

TABLE 6.4: matrix of cluster with less dominant emotions

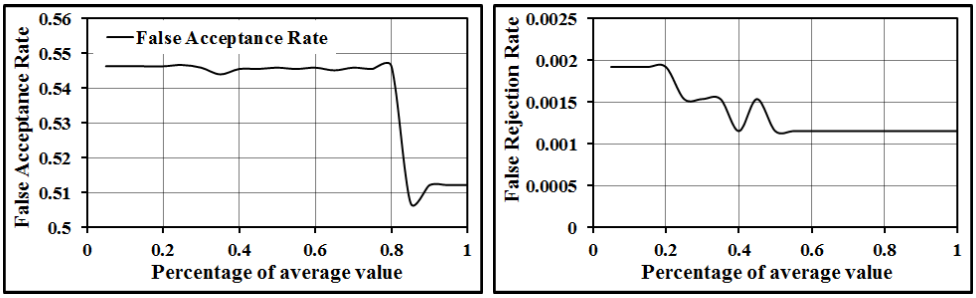

FIGURE 6.9: percentage of average vs false accept rate and false reject rate in bounded k-means for less dominant emotion classification

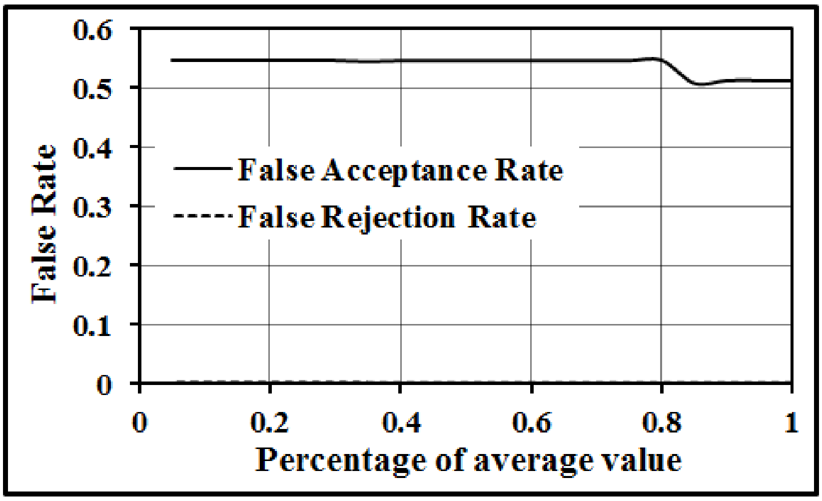

FIGURE 6.10: percentage of average vs false accept rate and false reject rate in bounded k-means for less dominant emotion classification

6.0.5.4 Less Dominant classes of emotions

From the cluster matrix we can identify the clusters(emotions) which has less influence over their own clusters and others clusters. When we tried to classify these less dominant emotions, we found the following results:

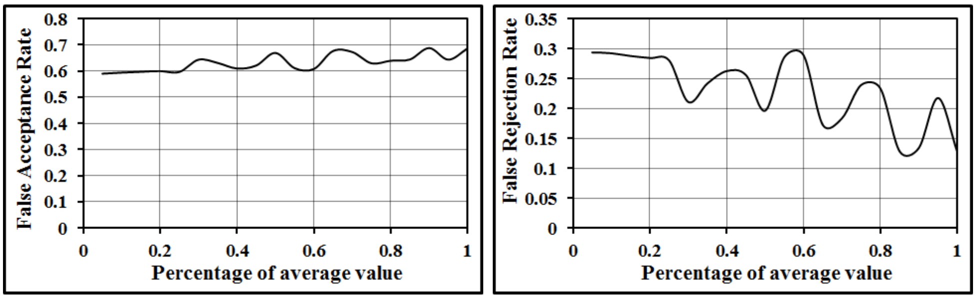

FIGURE 6.11: percentage of average vs false accept rate and false reject rate in bounded k-means for dominant user classification

6.0.6 User classification

6.0.6.1 All user classification

What about classifying a user based on his usage attributes. In this case we considered the data without any normalization. At first we tried to cluster the data with our unsupervised clustering approach (bounded k-means clustering). But we found it difficult to successfully identify users based on their usage attributes. The experimental results are as follows

6.0.6.2 Dominant user classification

From matrix of users clusters, we found that some of the users heavily dominates others clusters. So we made an attempt to classify these dominant users. The results we found was much better than the general case. The results are illustrated below:

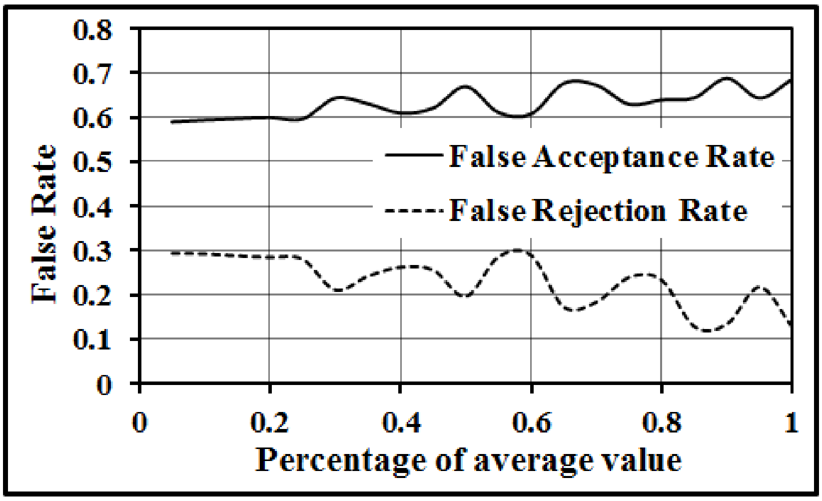

FIGURE 6.12: percentage of average vs both false acceptance rate and false rejection rate in bounded k-means for dominant user classification

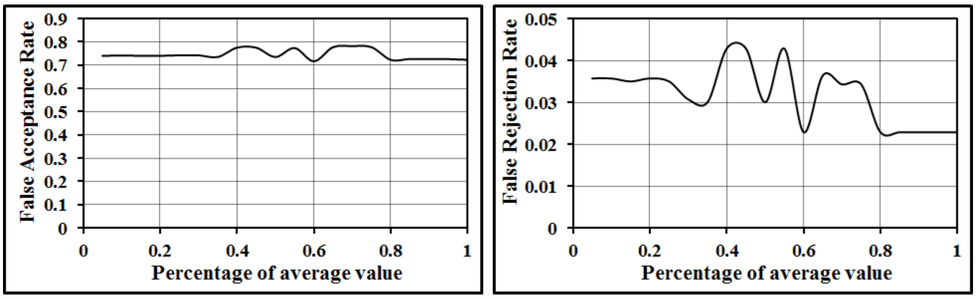

FIGURE 6.13: percentage of average vs false accept rate and false reject rate in bounded k-means for dominant user classification with neutral data only

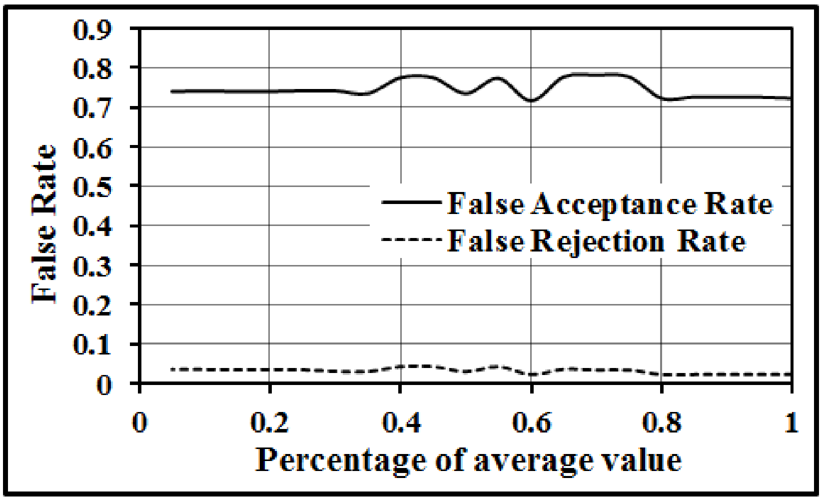

FIGURE 6.14: percentage of average vs false accept rate and false reject rate in bounded k-means for dominant user classification with neutral data only

6.0.6.3 Dominant user classification with neutral emotion only

From matrix of users clusters, we found that some of the users heavily dominates others clusters. So we made an attempt to classify these dominant users. The results we found was much better than the general case. The results are illustrated below:

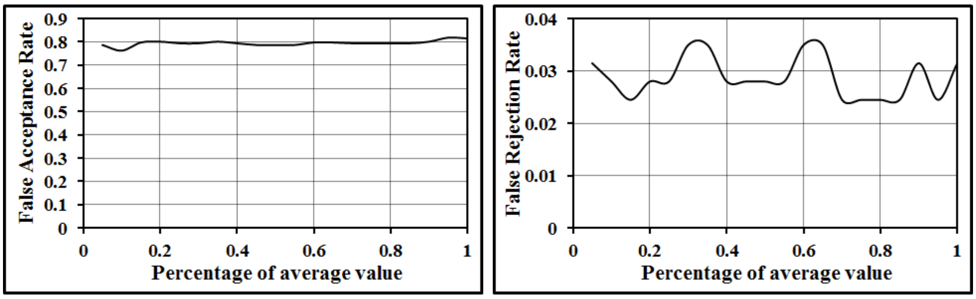

FIGURE 6.15: percentage of average vs false accept rate and false reject rate in bounded k-means for less dominant user classification with only neutral emotion's value

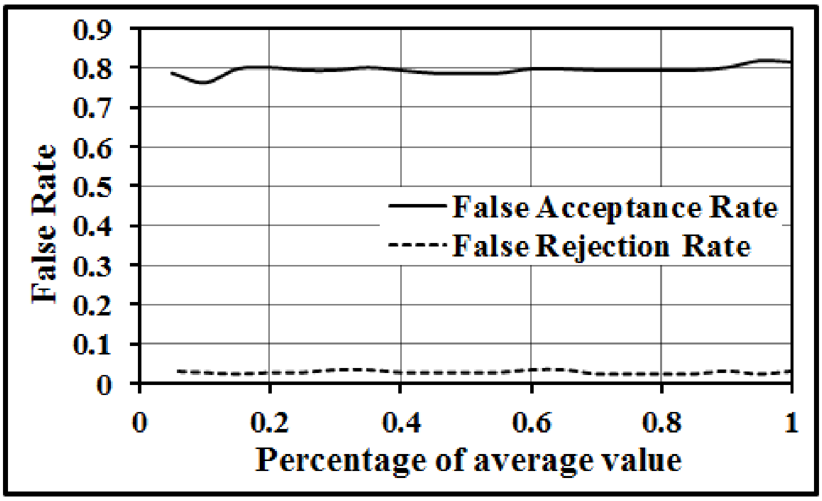

FIGURE 6.16: percentage of average vs false accept rate and false reject rate in bounded k-means for less dominant user classification with only neutral emotion's value

6.0.6.4 Less Dominant user classification with neutral emotion only

From matrix of users clusters, we found that some of the users heavily dominates others clusters. So we made an attempt to classify these dominant users. The results we found was much better than the general case. The results are illustrated in figure 6.16.

Chapter 7

Conclusion

Devices, capable of detecting human emotion and interacting accordingly is an important part of building intelligent computers. An emotionally aware system will be much user friendly, and less frustrating for the users as it will be able to make appropriate decisions about how to interact with the user or adapt their response. There are two main problems with current system approaches for identifying emotions that limit their applicability: they demand additional infrastructures which are often intrusive and require specialized information which are not always available. We experimented on detecting emotion by analyzing the pattern of their keyboard and mouse usage parameters. In our study, we tried to identify users different emotional states, users identity in many different ways by using bounded k-means clustering and k-nearest neighbor approach. Our best result shows good result in identifying 3-class emotion, dominant user identity with data from only one emotional state.

Bibliography

[1] S. Ioannou, A. Raouzaiou, V. Tzouvaras, T. Mailis, K. Karpouzis and S. Kollias: "Emotion Recognition through Facial Expression Analysis Based on a Neurofuzzy Method", *Neural Networks,* vol. 18, pp. 423-435, 2005.

[2] S.G. Kong, J. Heo, B. Abidi, J. Paik and M. Abidi: "Recent advances in visual and infrared face recognitiona review" *Computer Vision and Image Understanding,* vol. 97, no. 1, pp. 103-135, Jan. 2005.

[3] H. Gunes and M. Pantic, & ldquo, Dimensional Emotion Prediction from Spontaneous Head Gestures for Interaction with Sensitive Artificial Listeners, *Proc. Int',l Conf. Intelligent Virtual Agents,* pp. 371-377, 2010.

[4] O. W. Kwon, K. Chan, J. Hao, and T. W. Lee: Emotion recognition by speech signals, In Proc. EUROSPEECH, 2003.

[5] Picard, R.W. *Affective Computing.* MIT Press, Cambridge, 2007.

[6] Ekman, P. *Basic Emotions.* John Wiley and Sons, Ltd, 2005.

[7] Russell, J. Core affect and the psychological construction of emotion. *Psychological Review 110,* 1 (2003), 145-172.

[8] De Silva, L. and Suen Chun, H. Real-time facial feature extraction and emotion recognition. *Infor., Comm., and Sig. Proc. 2003 and the 4th Pac. Rim Conf. on Multimedia,* (2003).

[9] Khan, M.M., Ingleby, M., and Ward, R.D. Automated facial expression classification and affect interpretation using infrared measurement of facial skin temperature variations. *ACM Trans. Auton. Adapt. Syst. 1,*1 (2006), 91-113.

[10] Partala, T., Surakka, V., and Vanhala, T. Real-time estimation of emotional experiences from facial expressions. *Interact. Comput. 18,*2 (2006), 208-226.

[11] Fairclough, S. Fundamentals of physiological computing. *Inter. with Comp. 21,*1-2 (2009), 133-145.

[12] Mandryk, R.L. and Atkins, M.S. A fuzzy physiological approach for continuously modeling emotion during interaction with play technologies. *Int. J. Hum.- Comput. Stud. 65,* 4 (2007), 329-347.

[13] Ward, R.D. and Marsden, P.H. Physiological responses to different web page designs. *Int. J. Hum.-Comput. Stud. 59,* 1-2 (2003), 199-212.

[14] Chen, D. and Vertegaal, R. Using mental load for managing interruptions in physiologically attentive user interfaces. *CHI '04 ext. abst. on Human fac. in comp. systems,* ACM (2004), 1513-1516.

[15] Stern, R.M., Ray, W.J., and Quigley, K.S. *Psychophysiological recording.* Oxford University Press, New York, 2001.

[16] Joyce, R. and Gupta, G. Identity authentication based on keystroke latencies. *Commun. ACM 33,* 2 (1990), 168-176.

[17] Dowland, P. and Furnell, S. A Long-term trial of keystroke profiling using digraph, trigraph, and keyword latencies. In *IFIP Intern. Fed. for Infor. Processing.* Springer Boston, 2004, 275-289.

[18] Monrose, F. and Rubin, A.D. Keystroke dynamics as a biometric for authentication. *Future Gener. Comput. Syst. 16,* 4 (2000), 351-359.

[19] Bender, S. and Postley, H. Key sequence rhythm recognition system and method. .

[20] Admit One Security. *AdmitOneSecurity.* http://www.admitonesecurity.com.

[21] Epp, C. Identifying emotional states through keystroke dynamics. 2010.http://library2.usask.ca/theses/available/etd-08312010- 131027/.

[22] Gunetti, D. and Picardi, C. Keystroke analysis of free text. *ACM Trans. Inf. Syst. Secur. 8,* 3 (2005), 312-347.

[23] Bergadano, F., Gunetti, D., and Picardi, C. User authentication through keystroke dynamics. *ACM Trans. Inf. Syst. Secur. 5,* 4 (2002), 367-397.

[24] Gaines, R., Lisowski, W., Press, S., and Shapiro, N. Authentication by keystroke timing: some preliminary results. 1980.

[25] Gunetti, D., Picardi, C., and Ruffo, G. Keystroke analysis of different languages: a case study. In *Lecture Notes in Computer Science.* Springer Berline / Heidelberg, 2005, 133-144.

[26] Brown, M. and Rogers, S.J. User identification via keystroke characteristics of typed names using neural networks. *Int. J. Man-Mach. Stud. 39,* 6 (1993), 999-1014.

[27] Sheng, Y., Phoha, V., and Rovnyak, S. A parallel decision tree-based method for user authentication based on keystroke patterns. *IEEE Transactions on Systems, Man, and Cybernetics 35,* 4 (2005), 826-833.

[28] Lang, P. Behavioral treatment and bio-behavioral assessment: computer applications. *Technology in mental health care delivery systems,* (1980), 119-137.

[29] Zimmermann, P., Guttormsen, S., Danuser, B., and Gomez, P. Affective computing - a rationale for measuring mood with mouse and keyboard. *Inter. J. of Occ. Saf. and Ergo. 9,* 4 (2003), 539-551.

[30] Vizer, L.M., Zhou, L., and Sears, A. Automated stress detection using keystroke and linguistic features: An exploratory study. *Int. J. Hum.-Comput. Stud. 67,* 10 (2009), 870-886.