ISSN:0975-9646

Objective of Keystroke Dynamics for Identifying Emotional State

Shivshankar Rajput and Priyanka Vijayavargiya

Shri Vaishnav Institute of Technology and Science,
Indore, India

***Abstract*— This paper describes the concept based on using standard input devices, such as keyboard and mouse, as sources of data recognition of user's emotional states. Emotions are generally considered as products of evolution. Conventional emotional identification methods (facial expressional analysis, voice intonation analysis, thermal imaging of the face) use intrusive, lab-based, and expensive tools which are unsuitable for real-world situations. The idea behind keystroke dynamics is that people have different typing styles and by analysing the timings of keystrokes, a person's mental state can be identified. A possible solution for emotion identification is to determine user emotion by analyzing the rhythm of their typing pattern on a keyboard.**

***Keywords*— emotion, keystroke dynamics, biometric, emotion detection**

I. Introduction

Traditionally, Machines like computer do not adapt according to user's emotional states. If the computer system were capable of extracting the emotional state that the user is going through in a particular period of time they would have many benefits for intelligent computers.

A form of emotional intelligent would provide a richer context from which computer could change its behavior accordingly.There are two categories of Biometric: physiological (fingerprints) and behavioural (handwriting) [1]. Keystroke dynamics falls within the category of behavioural biometrics. The idea behind keystroke dynamics is that people have different typing styles and by analysing the timings of keystrokes, a person can be identified. A benefit of this metric is that measuring keystrokes can be done through a keyboard, thus negating the cost of typical physiological biometric systems which require expensive hardware to measure physical attributes.

"Keystroke dynamics is the study of unique timing patterns in the individual's typing and it includes extracting keystroke timing features such as the duration of key press and the time elapsed between key presses"

There have been various methods for evaluating emotional states like facial expression analysis, voice intonation analysis, change in breathing and physiological sensors attached to the skin etc. that have varying rates of success, but they still exhibit one or both of two main problems preventing wide scale use: they can be intrusive to the user, and can require specialized equipment that is expensive and not found in typical home or office environments. This technique is relatively straightforward to apply and is one of the least expensive biometrics. No additional devices need to be purchased, installed, or integrated.

.

II. Related Work

Creating an emotion recognition system is a challenging task. It requires solving dozens of problems, which may be grouped. The first research work in this domain was realized in 1980 with the report of the *Rand Corporation* [2]. The RAND report used a digraph representation for the keystrokes and conducted experiments on a small population of users. The majority of studies in keystroke dynamics are for authentication and verification purposes.

A. Emotion

According to [6], "*two generally agreed-upon aspects of emotion are:*

(a) Emotion is a reaction to events deemed relevant to the needs, goals, or concerns of an individual and

(b) Emotion encompasses physiological, affective, behavioral, and cognitive components.

Fear, for example, is a reaction to a situation that threatens (or seems to threaten) an individual's physical well-being, resulting in a strong negative affective state, as well as physiological and cognitive preparation for action. Joy, on the other hand, is a reaction to goals being fulfilled and gives rise to a more positive, approach-oriented state. "

Emotional state is an attribute of certain states. e.g. – "His voice was tinged with emotion".

B. Sensing and recognizing emotion

Criticism 1: The range of means and modalities of emotion expression is so broad, with many of these modalities being inaccessible (brain activity, neurotransmitters), and many others being too non-differentiated. This makes it unlikely that collecting the necessary data will be possible or feasible in the near future.

Criticism 2: People's expression of emotion is so idiosyncratic and variable, that there is little hope of accurately recognizing an individual's emotional state from the available data.

With emotion, as with weather, one can build sensors for measuring the physical equivalents of temperature, pressure, humidity, etc. One can also build successful algorithms for combining patterns of such measures, and thus recognize the emotional equivalents of a tornado or a blizzard. At this level, I expect emotion recognition to be solvable by machine, at least as well as people can label such patterns. However, I do not expect researchers will have success matching human labels when such labels may not exist: just like we do not have special names for most of the states of weather but rather only for its extreme states. In the in-between cases, we tend to use adjectives like, ‘‘ok day’’ or ‘‘partly cloudy’’ referring only to some quality of the weather. Similarly, not all aspects of affective state will be relevant or useful to observers. Sometimes just reporting the quality of the stateit’s ‘‘ok’’ or it’s ‘‘not so great’’ will suffice. Not that computer couldn’t label any state that they could represent, only that such labels might not matter much unless they are in service of some greater purpose.

C. Emotion Detection

There have been two main approaches for describing emotions: categorical and dimensional. The categorical approach applies labels to emotions with some languages or words (e.g. sadness, fear, joy) [3]. Dimensional approach uses two orthogonal axes called arousal and valence to describe emotions [4].Arousal is related to the energy of the feeling and is typically described in terms of low (i.e., sleepiness) to high (e.g. excitement) arousal. Valence describes the pleasure and displeasure of a feeling.

D. Keystroke Dynamics

Keystroke dynamics is the process of analyzing the way a user types on a keyboard and identify him based on his habitual typing rhythm. Keystroke dynamics is not what you type, but how you type.

Much of the previous research in keystroke dynamics has been for authentication and verification purposes. Since Gaines et al. [4] first proposed an approach using keystroke dynamics to verify users' identity; typing patterns have been studied extensively for security applications, to enhance the authentication process by comparing the current typing pattern with a previously constructed typing pattern.

Monorose and Rubin [5] made an automated classifier that used keystroke features including keystroke timing and key latency to detect unsuitability in users' typing patterns to enhance the authentication process. The typing features were extracted from both predefined text (fixed text) and spontaneously generated text (free text). Their proposed method yielded a 48.9% accuracy-recognition rate for a population of 31 users.

Monorose et al. [8], in another study, suggested that individuals' typing patterns are not stable, and change according to their environment, stress level, and cognitive function.

Classification algorithms for the analysis of keystroke dynamics include neural network [8], distance measures [7, 10], decision tree [11], and other statistical methods [7].

Affective Computing and Keystroke Dynamics

Recent work by Vizer, et al. [12] diagnosed individuals' cognitive and physical functions using keystroke dynamics. Their approach utilized the users' everyday interactions to detect changes in their cognitive and physical functions. The experiment consisted of control (no stress) condition, cognitive stress condition, and physical stress condition where the participants were asked to provide a text sample under each condition. The extracted features included timing, keystroke and linguistic features. The collected data was analyzed using several machine learning techniques: decision trees, support vector machine, k-nearest neighbor, artificial neural network They achieved correct classification of 62.5% physical stress and 75% for cognitive stress. They also tested their experiment with varying physical and cognitive abilities and with varying typing habits and keyboard.

Zimmermann et al. [13] described a method which uses keyboard and mouse interactions to detect users' emotions. This study used the categorical approach that uses emotional valence and arousal dimensions. Participants' emotions were assessed using physiological sensors that measured their respiration rate, pulse rate, and skin conductance. They were also asked to self-report their emotional states by using the Self-Assessment-Manikin (SAM) [3], which consisted of graphical manikin that each represents score in the valance and arousal dimension. Zimmermann et al's classiffer was able to distinguish between neutral and other emotional states, but was not able to distinguish between the other four induced states.

III. Different Method For Identifiyng Emotions

To build emotional classifier, typically researchers have to collect users’ behavioral or physiological patterns which are then mapped to emotional categories. To accomplish these goals, there would be two method: They either induce participants' emotions in a laboratory setting using one of the Mood Induction Procedures (MIPs) e.g., video or story, or in a real-world setting in which participants use their personal computers in their daily lives [15]. Both approaches have advantages and disadvantages: a laboratory-based study using mood induction procedures will yield more cleanly labelled data. However, the induced emotions do not necessarily represent the emotions that users experience in the real world.

On the other hand, using a real-world approach generates a greater amount of data compared to a time-limited laboratory-based approach, but with more noise and more in-complete data points. In this study, we chose to gather spontaneously generated interaction data without using any Mood Induction Procedures (MIPs), using same laboratory setting, computer application, and computer settings.

We are working on six basic emotional classes – confidence, sadness, nervousness, happiness, tiredness, hesitation. Our aim is to develop a web application to recognize human emotional states.

A. **The Data Collection Process:** It consists of gathering and labeling users’ keystroke and mouse data. This process runs in the background, gathering keystrokes regardless of the application that is currently in focus. The only visible sign that the application is running has an icon in the desktop system tray. The data collection server is allowed us to preserve participant anonymity and supported remote users.

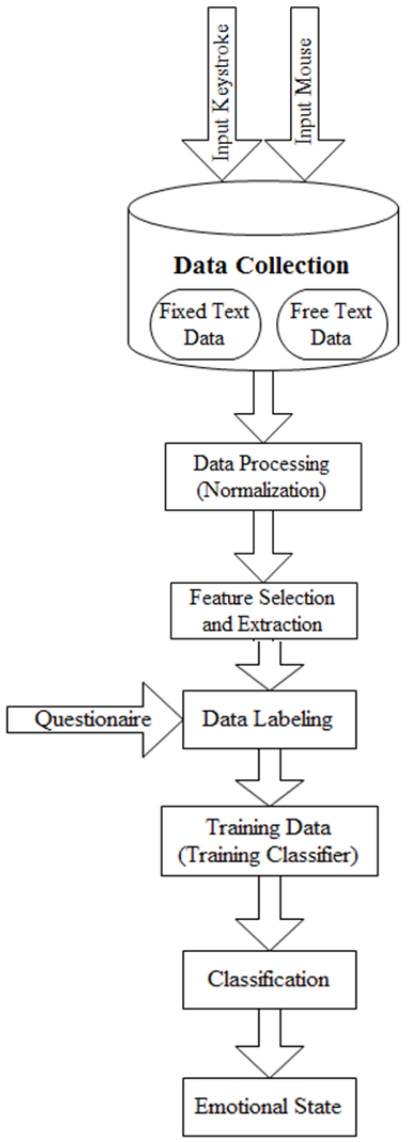

Fig. 1. Flow Chart for Emotional Classification

Fixed Text Data collection

This data collection method asked the user to type some fixed piece of text from the famous novel *"Alice's Adventures in Wonderland"* [16]. The user was unable to copying and pasting the fixed text. The user could close the data collection window if they were too busy.

We will develop a Java application to collect users' keystroke data for using in keystroke dynamics. This application will record all pressed key by the user and their press and release time. We asked the user to enter data at least once in a day. This ensured data collection under different emotional states of the same user. This took about 4 weeks to collect all the data. There are several reasons behind choosing fixed text over free text. It is very likely that while normal computer use, the user may use mouse more than keyboard. Using fixed text ensures a minimum number of keystrokes per sample and produces good results in building models. After typing these paragraphs, the user had to choose one of seven emotional states which best matched with the user's current emotional state.

Free Text Data collection

A java application which can run in background to collect keystroke timings (key press time and key release time). The user is not aware of the hidden software and thus is less bothered about the data collection process. The software prompts the user every 15 minutes to enter his/her mental state. A small window pops up with the six emotional states stated in the previous section and prompts the user to tick one of them according to their present emotional state. All collected data are stored in different tables of the database by different name.

B. Feature Extraction: Due to the large number of raw data we perform attribute selection which reduces the number of attribute to facilitate the classification process. We could select the necessary features like keystroke latency, duration of key hold, typing rate etc. These biometric features are converted into fuzzy vector which are further used for classification.There are several keystroke features, but we are going to extract following essential features

1) Session time is the total time spent by the user on the system. Session time is calculated by computing the difference between the starting time and user response times.

$$Session\ time = Starting\ time - User\ response\ time.$$

2) Keystroke latency is the time interval between the key release of the first keystroke and the key press of the second keystroke. However, the latency of longer n-graphs (n > 2) is defined as the time interval between the down key events of the first and last keystrokes that make up the n-graph.

Keystroke Latency

$$= \frac{\sum Releasing\ times\ of\ Key1 - \sum Depressing\ time\ of\ Key2}{Character\ per\ sentence}$$

3) Held Time (or Dwell Time) is the time (in milliseconds) between a key press and a key release of the same key.

4) Sequence is a list of consecutive keystrokes. For example 'REVIEW' is a Sequence. A Sequence can be of any length (minimum two). In this example, the Sequence is a valid English word, but this need not be the case. Thus, 'REV', 'IEW' are also valid Sequences from the same keystroke stream.

- A length-2 Sequence is called a digraph
- Length-3 Sequence is a trigraph; etc
- A general Sequence is therefore an n-graph.

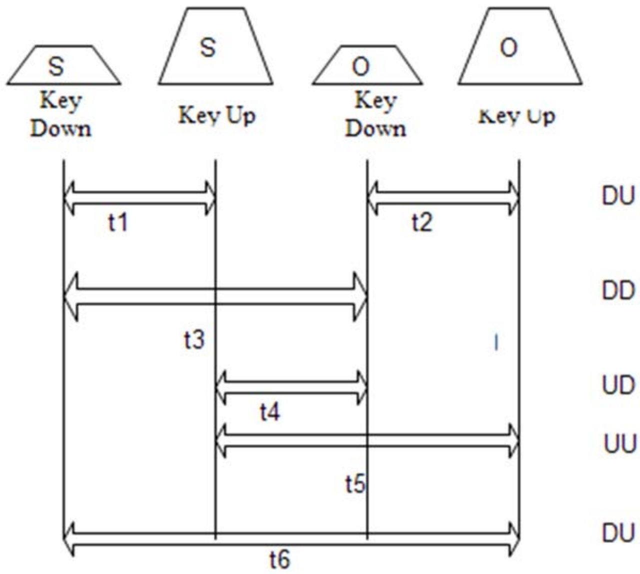

Fig. 2. Key hold time and Key flight time

5) *Typing speed* is the total number of keystrokes or words per minute. The typing speed could be an indicator of different emotions.

About the characteristic capturing, we mainly measured the Keystroke rhythm as biometric. We defined the following definition for the system captured time. The diagram show as figure 2.

- DU: means the Key Down_Up. It represents the time interval t1 or t2.
- DD: means the Key Down_Down. It represents the time interval t3.
- UD: means the Key Up_Down. It represents the time interval t4.
- UU: means the Key Up_Up. It represents the time interval t5. T5 is equal to(t2 + t5).
- TT: means the Key Total_Time. It represent the time interval t6.

C. Data Labeling:

There are two main methods to emotion labeling. One of the method is to label emotions by a human, the second approach is automatic labeling. There are different types of labels (like fear,excitement, boredom, joy, surprise etc.) where people are able to choose from a predefined list of word labels during questionnaire . The questionnaires may use scales such as the Likert scale presenting a range of responses to each question [17] or Self-Assessment-Manikin (SAM) technique, which is a graphical way of expressing valence, arousal and dominance [18].

We are going to labels the data provided by the human while questioning for the best accuracy of the trained system as shown in figure 1.

D. Training Classifier and recognizing emotions:

Numerous algorithms (Discriminate Analysis, Bayesian Analysis, k-Nearest Neighbor, Artificial neural network and Decision Trees) may be applied to train classifier of emotional states. It was difficult to evaluate which classification methods produced the best model due to the variability in experimental conditions.

Rule based systems for classification have the disadvantage that they involve sharp cutoffs for continuous attribute so used fuzzy logic as a solution to recognize users' emotional state. Fuzzy logic systems provide graphical tools to assist users in converting attribute values to fuzzy truth values (low, medium, high).

IV. CONCLUSIONS

In this work we have focused on the problem of human emotion recognition in the case of naturalistic, rather than acted and extreme, expressions. Keystroke dynamics has become an interesting research topic in the area of behavioral biometrics due to its non-intrusiveness and convenience. The current methods for detecting user emotions are voice intonation analysis, facial expression analysis, and physiological sensors attached to the skin etc. But these techniques use expensive, specialized, and invasive technologies not found in typical home or office.

Our solution is to identify user's emotional state through their typing rhythms or keystroke dynamics. The main benefit of using keystroke dynamics is that the required equipment, any standard keyboard which is in expensive and already widely used in most computer systems.

REFERENCES

[1] B. Miller, Vital sings of identity, IEEE Spectrum, 1994, pp. 22-30.
[2] R. Gaines, W. Lisowski, S. Press, and N. Shapiro, "Authentication by
keystroke timing: some preliminary results," Rand Corporation, Tech.
Rep., 1980.
[3] Picard, R.W. *Affective Computing*. MIT Press, Cambridge, 2007.
[4] P. Ekman. An argument for basic emotions. Cognition and Emotion, 6:169–200, 1992.
[5] Clayton Epp, Michael Lippold, and Regan L. Mandryk, Identifying Emotional States using Keystroke Dynamics , CHI 2011
[6] Emotion in human-computer interaction, *Scott Brave and Clifford Nass*, *Stanford University*
[7] F. Monrose and A. Rubin. Authentication via keystroke dynamics. In Proceedings of the 4th ACM Conference on Computer and Communications Security. ACM, 1997.
[8] F. Monrose and A. D. Rubin. Keystroke dynamics as a biometric for authentication. Future Generation Computer Systems, 16(4), 2000.
[9] 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.
[10] Joyce, R. and Gupta, G. Identity authentication basedon keystroke latencies. *Commun. ACM 33*, 2 (1990), 168-176.
[11] Sheng, Y., Phoha, V., and Rovnyak, S. A paralleldecision tree-based method for user authenticationbased on keystroke patterns. *IEEE Transactions on Systems, Man, and Cybernetics 35*, 4 (2005), 826-833.
[12] L. M. Vizer, L. Zhou, and A. Sears. Automated stress detection using keystroke and linguistic features: An exploratory study. International Journal of Human-Computer Studies, 67(10), 2009.
[13] P. Zimmermann, S. Guttormsen, B. Danuser, and P. Gomez. A_ective computing{arationale for measuring mood with mouse and keyboard. International Journal of Occupational Safety and Ergonomics, 9(4), 2003.

[14] M. M. Bradley and P. J. Lang. Measuring emotion: the self-assessment manikin and the semantic differential. Journal of Behavior Therapy and Experimental Psychiatry, 25(1), 1994.

[15] R. Westermann, K. Spies, G. Stahl, and F. W. Hesse. Relative effectiveness and validity of mood induction procedures: a metaanalysis. European Journal of Social Psychology, 26(4), 1996.

[16] Carroll, L. *Alice's Adventures in Wonderland*. TheGutenberg Project, 2008.

[17] Likert R (1932), A technique for the measurement of attitudes, Archives of Psychology, 22(140): 1-55

[18]] Bradley M, Lang P (1994) Measuring emotion: the self-assessment manikin and the semantic differential. Journal of behavior therapy and experimental psychiatry, 25(1):49-59