Abstract — Theory of “embodied cognition” suggests that a variety of mental activities are reflected in states of the body, such as postures, arm movements and facial expressions. The present study investigates the extent to which profiles of computer users—their gender, feelings, and emotional experience—can be assessed from movements of computer cursors. In one experiment, participants (N = 372) saw three film clips for two minutes each, rated their feelings afterward, and carried out simple perception tasks three times where our program traced participants’ cursor trajectories every 20 milliseconds. We investigated the extent to which the extracted cursor trajectory features could reveal participants’ profiles. Results indicated that a small number of trajectory variables were helpful to identify which film participants saw, how they felt while viewing the film, and their gender. We suggest that cursor motions provide rich information for dynamic user profile mining.

Keywords—Cursor Trajectory Analysis; User Modeling

I. INTRODUCTION

Who are you? This question conveys both trivial and profound connotations. “I am a 32-year-old mom of two beautiful sons” can be a quick response. Which school you went to, where you are from, what you like, and who circles of friends you are in can also serve as answers in other contexts. These facts about yourself can tell a great deal about you; yet, they do not show how you behave, and how you feel at a given moment. Are you a responsible person? Are you vulnerable to emotional ups and downs? “Who are you” is also subject to what you experienced right before. After viewing an action film, your confidence may soar high, but a sad movie may put you down to the bottom. This simple observation suggests that the notion of the self, “who are you?”, is multi-faceted.

Research on user profile mining has focused on factual information, such as ad click behavior [1], log file analysis in e-learning [2], and on-line query inspections [3]. These explicit data collection is important but in some occasions these measures are not available and hard to obtain. Here, we investigate user profile mining involving non-verbal behavior. Specifically, we examine the extent to which cursor trajectory features extracted from a large body of data (more than 10 million data points) help identify a psychological profile of a computer user.

A. Theoretical Rationale: Embodied Cognition

Movements of a computer cursor can be viewed as extensions of body motions, and psychological theory of “embodied cognition” offers a conceptual framework for trajectory-based user profile mining. Theory of “embodied cognition” suggests that a variety of mental activities are reflected in states of the body, such as postures, arm movements and facial expressions; in turn these bodily conditions also result in affective states. Barsalou suggests that these bidirectional relationships between mental activities and bodily states arise because the core of social and cognitive information processing stems from the “reenactment” (simulation) of original information [4, 5].

Findings in neuroscience give a neurological rationale for why tacit psychological states can be embodied in trajectories of a computer cursor. The basal ganglia, the control center of voluntary hand movement, receive excitatory input from almost all cortical areas and send the information back to the same areas through the thalamus. This feedback loop controls voluntary hand movements as well as emotion, motivation and decision making [6-9]. A deficiency in the basal ganglia results in neurological movement disorders such as Parkinson’s disease and Tourette syndrome. These motor disorders often accompany affective and motivational difficulties such as irritability, depression and apathy. More than 40% of the people suffering from Tourette syndrome are known to experience symptoms of obsessive-compulsive disorder (OCD), which is an anxiety disorder [8].

The conceptualization of the self is also related to the ability to monitor one’s bodily movements [10]. For example, schizophrenic patients with “first rank symptoms” frequently mistake their own thoughts for someone else’s (they may mistake their internal voices as someone else’s voices); these patients also display difficulties in distinguishing their own hand motions from those made by others [11].
Given the finding that voluntary motor control is associated with emotion control and self-awareness, various psychological states (e.g., temperament, character, and attitude) can be reflected even in subtle trajectories of a computer cursor.

Psychological research measuring cursor motions in high-order cognitive judgments such as reasoning, categorization, and lexical processing suggests that cognitive decision making unfolds dynamically and its mental processes can be captured by the movement of cursors [12-16]. Anecdotal evidence abounds regarding the impact of heightened anxiety and hand movement such as highly acclaimed athletes failing to carry out seemingly routine tasks due to performance anxiety.

To summarize, psychological and neurological evidence suggests that a variety of tacit user profiles can be mined from cursor trajectory data.

B. Related Work

1) Nonverbal cues for mind-mining: The emerging fields of social signal processing and affective computing posit that nonverbal behavioral cues, such as vocal tones, facial expressions, gestures and body motions, license automatic analysis of human behavior [17-19]. Bodily motions are particularly indicative of people’s emotional states. For example, Thrasher and colleagues [20] demonstrated that movements of heads, shoulders, and trunks differ systematically between happy and sad moods induced by music. Glowinski and colleagues [21] introduced sample entropy (SampEn) of body movements as a real-time method of emotion detection, where a high value of sample entropy corresponds to disorderly behavior and smaller values indicate greater regularity and reservation. Body movements are shown to be useful for the prediction of personality traits (e.g., extraversion and locus of control) in a group meeting [22]. Head movements reveal a great deal of normal and perturbed parent-infant interaction [23]. Overall movements of body parts also serve as an indicator of depression [24].

2) Cursor motions for emotion analysis: Although a bulk of research studies have shown a convincing link between bodily movements and underlying psychological states, few studies have investigated cursor motions as a tool for user-data mining. Cursor movement analysis came to light in the human computer interaction research as early as in the 1970s [25-28]; yet only a handful of research studies have investigated cursor motions as a mind-mining tool.

In the Mueller and Lockerd study, subjects (N = 17) carried out an online shopping task while their cursor activities were tracked [29]. The recorded cursor activities were later reproduced for observational analysis, and the researchers reported “similarities” of cursor activities relative to users’ interest. Guo and Agichtein assessed users’ intention in queries from their cursor movement patterns [30]. The researchers judged the “intent of a user” manually and suggest that the average trajectory length of navigational queries was shorter than that of informational queries.

Zimmermann employed a film-based emotion elicitation technique and investigated the impact of arousal and valence on cursor motions in an online shopping task (N = 76) [31, 32]. The study showed that the total duration of cursor movement and the number of velocity changes were related to arousal. However, no evidence linking valence (e.g., positive and negative affects) and cursor activities was corroborated.

Kaklauskas et al. [33, 34] developed a comprehensive mouse sensor system that evaluates users’ psychological, physiological and behavioral input for the detection of stress. The system records the location, speed and distance, hand shaking, and force pressure from motions of a mouse. However no empirical results linking cursor activities and stress levels have been reported to date.

Kapoor et al. [35] integrated a pressure-sensitive mouse into their multichannel automatic affect detection system. The researchers measured mean, variance, and skewness of mouse pressure while subjects (middle school students, N = 24) learned to solve a Tower of Hanoi puzzle. The mouse pressure was as discriminable as the skin conductance measure for the detection of frustration. Azcarraga and Suarez [36] evaluated EEG signals and cursor activities (the number of mouse clicks, distance traveled, click duration) during algebra learning in an intelligent tutoring system (ITS) to predict subjects’ emotions (N = 25). Prediction rates based solely on EEG were 54 to 88%. When cursor activity data were augmented to the EEG data, accuracy rates increased up to 92%, indicating that cursor activity data can supply useful information for ITS learning on top of EEG data. Finally, Yamauchi [37] studied the relationship between cursor trajectories and generalized anxiety in human subjects. The researcher found that an assortment of about 20 trajectory features related to controlling and targeting cursors can predict users’ anxiety levels.

In cognitive science, a number of researchers have employed cursor motion to investigate cognitive mechanisms underlying reasoning, categorization, lexical processing and unconscious priming [12-16]. These studies show that cognitive decision making occurs dynamically and its mental processes can be captured by the movement of cursors.

In summary, these studies all suggest an important link between psychological states and cursor activities; yet, there are still several critical problems to be clarified. First, nearly all the studies involving cursor motion analysis have focused on emotion assessment; no systematic studies have been conducted for the general link between cursor trajectories and person profiles, such as the identification of
gender, personality traits and event experience (e.g., identifying the type of a film clip that a user viewed). Second, most studies employed a small number of subjects (mostly N = 15–40 with data points of about 1000); thus statistical power of these studies was miniscule (see [37] for an exception). Third, the features extracted in these experiments were relatively uniform (speed, duration, number of clicks); thus, the applicability of these analyses is limited.

With these issues in mind, the current study had several aims. We applied cursor trajectory analysis for general user profile mining. We extracted 159 trajectory variables from more than 100,000 trials in a behavioral experiment involving 372 subjects (nearly 10 million data points). Using this rich data source, we examined the extent to which these cursor trajectory variables help assess profiles of computer users.

II. EXPERIMENT

We conducted a large scale experiment involving simple film viewing. The experiment consisted of three parts. In each part, participants saw a film clip for two minutes, rated their feelings afterward, and then carried out a similarity perception task (96 trials). This cycle was repeated three times with participants viewing different film clips (fearful, funny or neutral clips) (Fig. 1). During the similarity perception task, we traced trajectories of cursor motions and examined the extent to which these trajectories help identify participants’ gender, the type of film they saw, and their feelings while watching the film clip. In the first segment, either fearful or funny film was shown. In the second segment, all participants saw a neutral film. In the third segment, participants who saw the fearful film in the first segment saw a funny film, and vice versa.

The assignment of fearful or funny film in the first segment was determined randomly. To rate their feelings, participants filled the Positive and Negative Affect Schedule-Expanded (PANAS-X) [38] electronically.

A. Method

1) Participants: A total of 372 undergraduate students (female = 238, male = 134) participated in the experiment for course credit. The data from one participant were not analyzed because he did not complete the experiment.

2) Materials and procedure for film viewing and emotion rating: For the film clips, we selected excerpts of three films, When Harry Met Sally (1998), The Grudge (2004), a mundane winter scene taken from YouTube (https://www.youtube.com/watch?v=-kpEOQWN39s). The details are listed in Table I. Each clip lasted approximately two minutes. These movie clips were selected from movieloads.com, which provides scenes of films, trailers, and previews. This website also classifies movie clips in terms of related moods, such as fearful, angry, brainy, funny, and sentimental.

Shortly after viewing a film, participants rated their feelings using the questionnaire—Positive and Negative Affect Schedule-Expanded (PANAS-X) [38]. In PANAS-X, two broad affects, positive and negative affects, are further subdivided into smaller units of emotions—fear, sadness, guilt, hostility, shyness, fatigue, surprise, joviality, self-assurance, attentiveness, and serenity. For our data analysis, we focused on four emotion categories—attentiveness, joviality, fear and fatigue.

3) Materials and procedure for the similarity perception task: Shortly after emotion rating, participants carried out a visual perception task pertaining to judgments of similarities of geometric figures [37, 39, 40]. Participants were presented with a triad of geometric figures on a computer monitor. The task was to select which choice figure, shown at the top-left or top-right corner of the monitor, was more similar to the base figure shown at the bottom. Participants indicated their choices by clicking a “left” or “right” button placed at the top of each choice figure (Fig. 2). In each trial, our program recorded the x-y coordinates of the cursor location every 20 milliseconds from the onset of a trial (participants pressing the “Next” button) until the end of the trial (participants pressing an either left- or right-choice button).

<table>
<thead>
<tr>
<th>Category</th>
<th>Name</th>
<th>Length</th>
<th>Portions in the film</th>
</tr>
</thead>
<tbody>
<tr>
<td>Funny</td>
<td>When Harry Met Sally</td>
<td>2:06</td>
<td>0:45-0:47:25</td>
</tr>
<tr>
<td>Fearful</td>
<td>The Grudge</td>
<td>2:00</td>
<td>0:05:35-0:07:35</td>
</tr>
<tr>
<td>Neutral</td>
<td>Winter Scene</td>
<td>2:02</td>
<td>0:1:01-0:03:03</td>
</tr>
</tbody>
</table>

Figure 1. A schematic illustration of the experiment.
To start each trial, participants pressed the “Next” button, and a triad stimulus was presented. Participants indicated which choice figure, shown at the top-left or top-right corner of the monitor, was more similar to the base figure shown at the bottom by pressing a “left” or “right” button (Fig. 2). After the response, the “Next” button appeared again. This cycle was repeated 96 times.

There were no correct/incorrect answers in this task, and participants were instructed to make a selection based on their personal preference and no incentive to make “correct” responses was provided.

We selected this simple similarity judgment task because the perception of similarity is one of the most fundamental psychological functions that mediate decision making, memory, generalization, impression formation and problem solving [41]. Thus, the basic characteristics of our similarity perception task are likely to speak to more complex and realistic situations, such as comparing and selecting consumer products at an online shopping site.

In total, 16 basic triads were produced by varying the number of local shapes—figures made of 3 or 4 local shapes (3-4), 9 or 10 local shapes (9-10), 15 or 16 shapes (15-16), and 36 shapes (36) (Fig. 3). In the experiment, 32 triads were produced from the 16 basic triads by swapping the locations of the choice figures, and participants received 32 triads three times.

Feature selection: To cope with the high dimensionality of the data (159 feature dimensions) relative to individual cases (N = 371), we applied random forest for feature selection [42] because it is suitable for the selection of relevant features when the data contain a large number of input variables with fewer observations (cases). Random forest employs “ensemble” learning; 500 or more decision trees are formed by randomly selecting observations and variables. By aggregating “votes” cast by these random decision trees, the algorithm generates estimated likelihoods of a dependent variable together with the importance scores of features. For our film classification study, we first randomly separated training (n = 258) and test data (n = 113) (training : test data = approximately 70:30). For each subject we applied random forest to

4) Data analysis: To pre-process the cursor trajectory data, we first applied a linear interpolation method and standardized cursor trajectories of all trials to 100 equally-spaced time steps starting from the onset of the first cursor move to the time slice of the final move (at which a choice button, either left or right, was pressed) [12, 14]. We extracted trajectory properties of time (e.g., inception, end time, move duration), space (e.g., displacement, attraction_in, attraction#, attraction_out), tuning (direction change, motion change, mid-line cross, x-overshoot, y-overshoot) and distributions of speed (e.g., mean, standard deviation of speed) (Fig. 4). Each movement trajectory was also segmented into quadrants with its temporal layouts (1–25th, 26–50th, 51–75th, 76–100th time steps) and the values of these features were calculated both for the entire trial (1–100 time steps) and each quadrant (e.g., 1–25th time steps). Trajectory features were also segmented into different spatial regions. Specifically, we analyzed trajectory features when the cursor deviated away from the final selection (away region) or toward the final selection (toward region) with respect the ideal path (Fig. 4). All variables were standardized by subtracting the means and dividing by the standard deviations. Trials that took more than 6 seconds were excluded for the data analysis and the data from participants who had fewer than 90 trials were not analyzed.

Figure 3. Illustrations of stimuli used in the similarity perception task. 16 basic triads were produced by varying the number of local shapes—3 or 4, 9 or 10, 15 or 16, and 36 shapes.
classify individual trials. The algorithm generates estimated likelihoods of a dependent variable together with the importance scores of features, which we adopted for feature selection. We obtained feature importance scores for the 159 trajectory features, and selected 10 high-ranked features. The validity of the selected features was tested with a separate set of 111 subjects.

6) Classifier training and test: We employed both random forest and Support Vector Machine (SVM) with a radial basis function (RBF) kernel as classifiers. An SVM-RBF classifier has two parameters, gamma and cost. To identify appropriate parameter values, we employed the grid search algorithm combined with a 10-fold cross-validation procedure applied to the training data. This parameter search yielded parameter values of gamma = 0.097 and cost = 5.032. These fixed parameter values were applied in our all test conditions. No parameter tuning was made with the test data. For the random forest classifier, we employed default parameter values set in the R package “randomForest” [42]. The validity of the selected features was tested with the test data set. For random forest, The prediction performance of the selected features was measured by Out of Bag (OOB) cases—cases that were not used for training. Thus, our OOB classification measure was equivalent to a bootstrap cross validation method [42]. For SVM-RBF, we calculated prediction accuracy by a 10 fold cross-validation method. Specifically, the test data set was segmented into 10 groups, an SVM-RBF model was formed with the data from 9 groups, and the prediction accuracy of the trained model was tested with the data from the remaining untrained group. This procedure was repeated 10 times for every group segment. The overall prediction performance was calculated by the total prediction accuracy taken from all “untrained” groups.
B. Results

1) Film classification: Our first attempt is to analyze the cursor trajectory data to identify the types of films (fearful, funny, or neutral films) that a participant viewed. For each participant, we analyzed 288 trials (96 trials taken from the similarity perception task made three times in the three segments). Our program extracted 159 trajectory features from each trial and classified each trial whether it came from the segment of the funny, neutral, or fearful film (Fig. 5). For this analysis, we first selected 260 subjects randomly from the pool of 371, and selected high-ranked features from the importance scores generated by random forest. The validity of the selected features was tested with a separate set of 113 subjects.

Table III summarizes the classification prediction performance of the 10 high-ranked features obtained from the random forest classifier. The numbers shown in the tables denote the number of individual trials taken from the test data. Overall prediction accuracy of the classification performance is 0.60, which is significantly higher than chance-level performance of 0.33; \( t(112) = 39.9, p < 5.1 \times 10^{-68} \), suggesting that a small number of cursor trajectory features are indeed useful to identify the type of a film clip the participant viewed.

The performance of the SVM-RBF classifier was lower than that of random forest; mean accuracy = 0.52, standard deviation = 0.07. Yet, its prediction performance was still far above chance-level performance of 0.33; \( t(112) = 26.7, p < 1.1 \times 10^{-49} \). Table III summarizes the classification prediction performance of the 10 high-ranked features obtained from the random forest classifier. The numbers shown in the tables denote the number of individual trials taken from the test data. Overall prediction accuracy of the classification performance is 0.60, which is significantly higher than chance-level performance of 0.33; \( t(112) = 39.9, p < 5.1 \times 10^{-68} \), suggesting that a small number of cursor trajectory features are indeed useful to identify the type of a film clip the participant viewed.

The performance of the SVM-RBF classifier was lower than that of random forest; mean accuracy = 0.52, standard deviation = 0.07. Yet, its prediction performance was still far above chance-level performance of 0.33; \( t(112) = 26.7, p < 1.1 \times 10^{-49} \). Table III summarizes the classification prediction performance of the 10 high-ranked features obtained from the random forest classifier. The numbers shown in the tables denote the number of individual trials taken from the test data. Overall prediction accuracy of the classification performance is 0.60, which is significantly higher than chance-level performance of 0.33; \( t(112) = 39.9, p < 5.1 \times 10^{-68} \), suggesting that a small number of cursor trajectory features are indeed useful to identify the type of a film clip the participant viewed.

2) Selected features for film classification: What trajectory features were particular useful for the film classification? Factor analysis applied to the 10 selected features show that approximately 60% of the variance of these features was explained by six factors (Table 2). Factor 1, which explains 17% of the variance, is dominated by overall time, inception time (i.e., the time when a cursor...
motion starts), and the time spent in the “away” region (the region where a cursor veers away from the final destination, and see Fig. 4). Factor 2 is related to speed in the “away” region, and Factor 3 is associated with the speed of cursor motion in the early temporal segment (1Q and see Fig. 4). Trajectory features pertaining to speed dominated the selected features; seven out of the 10 features were related to speed. Earlier time segments (1Q and 2Q) were also critical.

Our additional analysis suggests that the three film clips indeed generated different cursor trajectories. Overall cursor motion time and inception time were significantly longer in trials given after the fearful film and the funny film than those given after the neutral film; t’s > 6.0, p’s <10−8 (Fig. 6). In contrast, the fearful film produced lower velocity trials in the away region than the neutral film and the funny film did; t’s > 2.40, p’s <0.02.

These results suggest that the types of films that participants saw were indeed reflected in their cursor motions.

![Film classification and relevant features](image)

Figure 6. Trajectory feature values extracted right after each of the three film viewings. The values of the individual features (the y-axis) were standardized with respect to their means and standard deviations.

3) Emotion classification: The analysis described above indicates that our cursory trajectory method is valid to detect emotionally loaded events that a user experienced shortly before. Can the same method be applied for the identification of actual feelings of the user? That is, is our analysis sensitive enough to predict feelings that our participants actually felt after viewing these film clips?

To address this question, we first collected emotion ratings made in the questionnaire given after each film viewing, calculated the median of ratings for each emotion category (joviality, attentiveness, fear, and fatigue) and then classified individual rating scores as high or low; emotion rating scores above the median score were classified as “high” and rating scores equal to or below the median value were classified as “low” for each participant. For example, given a participant A, if his rating of joviality is above the median of the entire subject ratings, then his joviality is classified as “high” in that segment. If his rating was below the median in the neutral film segment, then his joviality is classified as “low” in the neutral film segment.

As in the film classification study, we analyzed 288 trials (96 trials taken from the similarity judgment task made three times in the three segments) for each participant in the test data set. Our program extracted 10 trajectory features identified earlier (see Table III) from each trial and classified each trial whether it came from the segment in which the participant actually felt high / low in a given emotion (e.g., joviality, attentiveness, fear and fatigue). See Fig. 7 for an illustration.

![Figure 7. An illustration of the emotion classification study. The classifier classifies each trial whether it came from the segment when the participant was high / low in a given emotion.](image)

Table IV summarizes the classification performance for the four emotion categories obtained from a random forest classifier. In all cases, the classification accuracy was above 0.70, far above the chance level performance of 0.5; in all cases t’s > 30.0, p < 10−10.

As in the film classification analysis, the performance of the SVM-RBF classifier (gamma = 0.097 and cost = 5.032) was lower than that of random forest; mean accuracy = 0.69, 0.67, 0.68, 0.67, respectively, for joviality, attentiveness, fear, and joviality. However, its prediction performance was still far above chance-level performance of 0.50; t’s = 23.0, p < 6.0 x 10−3.

4) Gender classification: Our last analysis involves the identification of gender. That is, by simply analyzing cursor trajectories made by a participant, can we identify the gender of that person? For this analysis, we calculated the means of trajectory values for each participant over 96 trials taken from Part 2 (a neutral film was shown in Part 2 and see Fig. 1) and examined the extent to which our classifier could detect the gender of the person.

As in the previous analysis, we separated the data randomly into training (n_train = 253; n_female = 167; n_male = 86) and test sets (n_test = 113; n_female = 69; n_male = 44), selected relevant features using the training data, and evaluated the
prediction performance of our classifier with Out of Bag (OOB) cases obtained exclusively from the test data. Because far more female subjects participated in our experiment than male subjects, we balanced the male-female ratio by randomly sampling \( n_{\text{male}} \) participants from the pool of the female participants. We repeated this sampling procedure 50 times and report the confusion matrices and statistics corresponding to the medians of classification accuracy. To investigate impacts of emotion on cursor motion further, we also analyzed the performance of our classifier with four emotion variables—positive affect, negative affect, self-assurance, and attentiveness—taken from PANAS-X ratings.

### Table IV. Emotion classification performance obtained from a random forest classifier

Subjects whose emotion ratings had only one class (either high or low only) were excluded for this analysis.

<table>
<thead>
<tr>
<th>Joviality</th>
<th>Predicted</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>High</td>
<td>Low</td>
<td></td>
</tr>
<tr>
<td>Reported</td>
<td>High</td>
<td>6,614</td>
<td>2,640</td>
</tr>
<tr>
<td></td>
<td>Low</td>
<td>2,450</td>
<td>7,792</td>
</tr>
<tr>
<td></td>
<td>Total</td>
<td>19,496</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Accuracy</td>
<td>0.74</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Attentiveness</th>
<th>Predicted</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>High</td>
<td>Low</td>
<td></td>
</tr>
<tr>
<td>Reported</td>
<td>High</td>
<td>7,550</td>
<td>2,217</td>
</tr>
<tr>
<td></td>
<td>Low</td>
<td>2,900</td>
<td>6,292</td>
</tr>
<tr>
<td></td>
<td>Total</td>
<td>18,959</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Accuracy</td>
<td>0.73</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Fear</th>
<th>Predicted</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>High</td>
<td>Low</td>
<td></td>
</tr>
<tr>
<td>Reported</td>
<td>High</td>
<td>5,324</td>
<td>3,000</td>
</tr>
<tr>
<td></td>
<td>Low</td>
<td>2,019</td>
<td>9,189</td>
</tr>
<tr>
<td></td>
<td>Total</td>
<td>19,532</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Accuracy</td>
<td>0.74</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Fatigue</th>
<th>Predicted</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>High</td>
<td>Low</td>
<td></td>
</tr>
<tr>
<td>Reported</td>
<td>High</td>
<td>4,916</td>
<td>3,225</td>
</tr>
<tr>
<td></td>
<td>Low</td>
<td>1,946</td>
<td>8,575</td>
</tr>
<tr>
<td></td>
<td>Total</td>
<td>18,662</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Accuracy</td>
<td>0.72</td>
<td></td>
</tr>
</tbody>
</table>

Table V (A) shows a confusion matrix when 10 cursor trajectory variables were supplied to our random forest classifier. Using only cursor variables, the median of classification performance was 0.61 (mean = 0.61, SD = 0.04), which is above chance level performance of 0.50; Chi-squared = 3.81, \( p = 0.05 \). By including the four emotion variables, our classifier was capable of classifying Out of Bag cases with median accuracy of 0.65 (mean = 0.64, SD = 0.04; Chi-squared = 6.63, \( p < 0.01 \) (Table V (B)).

As in the previous analyses, the performance of the SVM-RBF classifier (gamma = 0.125 and cost = 4) was lower than that of random forest; median accuracy = 0.57 (both for with only 10 cursor features and 10 cursor features combined with four emotion variables).

### Table V. Gender classification performance with 10 cursor trajectory variables—(A), 10 cursor trajectory variables combined with 4 emotion scores—(B), and 4 emotion scores only—(C).

#### (A)

<table>
<thead>
<tr>
<th>Predicted</th>
<th>Female</th>
<th>Male</th>
</tr>
</thead>
<tbody>
<tr>
<td>Actual</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>31</td>
<td>13</td>
</tr>
<tr>
<td>Male</td>
<td>21</td>
<td>23</td>
</tr>
<tr>
<td>Total</td>
<td>88</td>
<td></td>
</tr>
<tr>
<td>Accuracy</td>
<td>0.61</td>
<td></td>
</tr>
</tbody>
</table>

#### (B)

<table>
<thead>
<tr>
<th>Predicted</th>
<th>Female</th>
<th>Male</th>
</tr>
</thead>
<tbody>
<tr>
<td>Actual</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>31</td>
<td>13</td>
</tr>
<tr>
<td>Male</td>
<td>18</td>
<td>26</td>
</tr>
<tr>
<td>Total</td>
<td>88</td>
<td></td>
</tr>
<tr>
<td>Accuracy</td>
<td>0.65</td>
<td></td>
</tr>
</tbody>
</table>

### III. DISCUSSION

1) **Summary and Conclusion:** User profile mining generally involves either a large scale observational study collecting explicit user data (e.g., online query samples) or very small experimentation with 20–40 subjects. In this study, we conducted a human behavioral experiment consisting of more than 100,000 trials and investigated the extent to which non-verbal behavioral cues—cursor motions collected in a mundane perception task—help assess users’ tacit psychological profiles.

Results indicate that emotionally loaded films indeed influence cursor motions and fine-tuned classification analysis of cursor movements helps detect the type of a film clip that the user viewed, and the feeling and gender of the user.

The user-data mining community has focused on the explicit fact-based mining to find user profiles. Our research shows that people’s “profile” can be assessed not

---

2 The data from five participants were excluded because they did not complete the PANAS-X questionnaire.
only by what they select or which ad they click but also how they click and how fast they move computer cursors.

One clear advantage of our cursor motion analysis is its versatility. Because cursor motions are subtle, unobtrusive, inexpensive and ubiquitous, this technique can be integrated into existing user interfaces relatively easily to assess people’s psychological states.

2) Limitations and Future Studies: One significant limitation is the generalizability of the current experimental setting. Here, we employed a simple perception task that required a large number of trials. This task, although basic and can be applied to different situations, is not directly related to what people do regularly on the Internet. Although some indirect link between the current experimental task and general Web-based activities is plausible, this should be tested empirically with a task that bears real-life implications. In this regard, the validity of the cursor-based data-mining should be tested further. In the same vein, it is possible that the cursor trajectory features identified in the current experiment can be specific to the task at hand. The generalizability of these features should be investigated further. Finally, the present study addresses the trajectory of a computer cursor. A touch-based user interface, such as tablets and mobile phones, do not measure the trajectories in the same way tested in this experiment. Clearly, different features need to be extracted for different interfaces.

REFERENCES


