Continuous Authentication Using Behavioral Biometrics

Shambhu Upadhyaya, SUNY at Buffalo

March 24, 2017
Acknowledgments

NSF Grant No. CNS-1314803 (SaTC Medium)
Clarkson University (Stephanie Schuckers and Daqing Hou)
Yan Sun, PhD student at UB
Hayreddin Ceker, PhD student at UB
Motivation

Desktop system authentication
  • Password based
    - Creation, memorization and management daunting task
    - Typical systems do not guarantee the legitimacy of the person at the console
    - Leads to masquerade/impersonation attacks

Mobile device authentication
  • Pin or patterns
    - Easily revealed by shoulder surfing

What is the solution?
  • Just get rid of it!
  • A single ignorant person can risk the entire system!
DARPA’s Initiative in 2012

It all started at DARPA

DARPA: Dump passwords for always-on biometrics

- By Kathleen Hickey
- Mar 21, 2012

The Defense Advanced Research Projects Agency wants to eliminate passwords and use an individual’s typing style and other behavioral traits for user authentication.

Why so many bad passwords? Because the rules allow them.

- By Kevin McCane
- Mar 12, 2012
DARPA’s Active Authentication Program

Active Authentication BAA-12-06

• March 6, 2012 for Phase I
  - Develop novel ways to authenticate using unique aspects of individual (Biometrics)
  - Use observables on how we interact with the world (Behavioral Biometrics)
  - Use of software-based biometrics
  - As a first step, do not use any additional hardware

Concept of “cognitive fingerprint”

• Pattern based on how our mind processes information
  - Use multiple modalities
  - Accuracy, robustness and transparency
DARPA’s Advanced Program

Thrust 1
• Goal is to deploy the new authentication platform on a DoD desktop or laptop

Thrust 2
• Securing mobile devices

Where DARPA is Going, You Don’t Need Passwords
Active Authentication program investigates behavioral biometrics for mobile devices
DARPA’s Vision of Continuous Authentication

Continuous authentication using:
- Multiple modalities in a rotating fashion
- Multiple authentications initiated each minute
- Open architecture to bring in future modalities

You

- Data from your experiences
  - Computational linguistics
    - (How you use language)
  - Structural semantic analysis
    - (how you construct sentences)
    - Forensic authorship

- The context you exist in
  - Keystroke pattern; Mouse movement

- How you interact with technology
  - Fingerprint; Iris pattern; Vein pattern; DNA; Facial geometry

- Physical aspects of you
  - Transparent validation of the person at the computer
  - Without passwords
  - Without proxies
  - Without hassle

Courtesy: DARPA

Start here
Biometrics

Metrics related to human characteristics

- Physical
  - Fingerprint
  - Face
  - Iris, etc.

- Behavioral
  - Keystroke
  - Gestures
  - Voice
  - How user searches for information
  - How user reads material, etc.
DARPA’s Funded Programs

Iowa State
Stylometry focused on thought processing time

Drexel
Stylometry augmented by author classification and verification

NY IT
Stylometry how a user types – ignoring the words

NPS
Behavioral manifestations of human thought processes

BehavioSec
Keystroke, mouse, in context

NRL
Identification of users through Web browsing behavior

UMD
Information processing from computer screens

Coveros
User Behavior Patterns as seen from the Operating System

SWRI
Use covert games disguised as computer anomalies

Allure Security
User search behavior characteristics

AA Application
Validates Level of trust in Identity
DARPA’s Phase I Results

<table>
<thead>
<tr>
<th></th>
<th>TP</th>
<th>FAR</th>
<th>Time</th>
</tr>
</thead>
<tbody>
<tr>
<td>Allure Security</td>
<td>95.0%</td>
<td>1.0%</td>
<td>5m</td>
</tr>
<tr>
<td>Iowa State U (KRR)</td>
<td>92.7%</td>
<td>5.5%</td>
<td>29sec</td>
</tr>
<tr>
<td>NYIT</td>
<td>92.0%</td>
<td>4.0%</td>
<td>1min</td>
</tr>
<tr>
<td>Drexel U</td>
<td>92.0%</td>
<td>5.0%</td>
<td>50sec</td>
</tr>
<tr>
<td>University of MD</td>
<td>82.8%</td>
<td>20.0%</td>
<td>83sec</td>
</tr>
<tr>
<td>NRL</td>
<td>82.0%</td>
<td>6.0%</td>
<td>4hrs</td>
</tr>
<tr>
<td>Coveros</td>
<td>80.0%</td>
<td>10.0%</td>
<td>30sec</td>
</tr>
<tr>
<td>SWRI</td>
<td>75.0%</td>
<td>25.0%</td>
<td>8min</td>
</tr>
<tr>
<td>Alanka Brown</td>
<td>10.1%</td>
<td>10.0%</td>
<td>1min</td>
</tr>
</tbody>
</table>

TP = True Positive Rate  
FAR = False Accept Rate  
Time = time before decision

How you think
- Look for information
- Interactions with applications
- Your personal demographics
- Time to think when writing
- Word choice/use

How you use your device
- How your phone moves on you
- How you move your phone when using it
- How you swipe/type
- How you compose language (written and spoken)
- How device is used

Physical aspects of you visible to the device
- Passive Fingerprint Detection
- Passive Facial Recognition
- Passive Heartbeat Detection

Courtesy: DARPA
Focus of the Talk

Keystroke dynamics as behavioral biometrics

- Short text
- Long text

Short text keystroke dynamics

- Generally useful for one-time authentication

Long text keystroke dynamics

- Necessary for continuous authentication

Rest of the talk will focus on this
Outline of the Talk

Introduction
- General approach to continuous authentication

Keystroke dynamics and mouse movements
- Feature selection
- Methodology - Gaussian model, SVM, transfer learning
- Datasets and anonymization

Results
- GMM, SVM, transfer learning

Research directions
- Secondary features
- Deep learning
- Adversarial learning
- Extension to network of smart devices
Current Authentication Schemes

The standard methods
- PIN/Password
- Security Questions
- Fingerprint
- Retina Scanner

They are all obtrusive!
Popular Behavioral Biometrics

Humans recognize people by who they are and how they behave

- Rather than by the secrets that they know

Cues for recognition

- Typing patterns
- Gait
- Word/phrase choices

Displayed image

Worker A

A couple holding hands on a beach during sunset. The sun is creating an orange glow which reflects into the water.

Worker B

The image shows a sunset. The image shows a beach. The image shows two people holding hands.
A General Approach to Active Authentication

Some call it “Continuous Authentication”, “Implicit Authentication”, “Transparent Authentication”

Users identify themselves at a console and simply start working
Authentication process occurring in the background
  • Invisible and free of interruptions and no loss of performance

Device recognizes the user
  • Adapts to changes
A Typical Desktop Scenario

- User logs in for the first time
- Normal activities at console
  - Unusual activities flag system
- Inactivity at console
- Activities continue

(Continuous) Cognitive and behavioral biometrics-based authentication

(Active) Physical biometrics verification

Activities at console
Background processes
The Big Picture

Transfer learning

Our objective is to design a standalone active authentication mechanism that can adapt to changing environmental conditions by using behavioral biometrics with respect to specific system requirements and certain standards. For instance, the European Stan-

Long-text data

GMM, SVM, Fusion

Keystroke dynamics
Outline of the Talk

Introduction
  • General approach to continuous authentication

Keystroke dynamics and mouse movements
  • Feature selection
  • Methodology - Gaussian model, SVM, transfer learning
  • Datasets and anonymization

Results
  • GMM, SVM, transfer learning

Research directions
  • Secondary features
  • Deep learning
  • Adversarial learning
  • Extension to network of smart devices
Why Keystrokes?

Keystroke dynamics in psychology
- Human computer interaction play a key role
- Salthouse [1] proposed a model for the steps taking place during typing
- It is a 4-stage process

Keystroke dynamics as a behavioral biometric
- Manner and rhythm of typing – Idiosyncratic
- It can be used as a means for authentication
- Low implementation and deployment cost; non-invasive, transparent
- Many methods and classifiers have been proposed

Rhythm in Keystrokes

John  Tom  Mike

Fast typist  Medium typist  Slow typist
Feature Selection

Keystroke features
- Digraphs

Mouse movements
- Clicks
- Distance
- Speed
- Angle

- $f_1$: flight time
- $f_2$: dwell time
- $f_3$: dwell time
- $f_4$: dwell time
- $f_5$: flight time
- $f_6$: dwell time
Methodology

Classification

- Keystroke dynamics recognition is a pattern recognition problem
- Three categories of algorithms [1]
  - Statistical (61%) – probabilistic, cluster analysis
  - Machine learning (37%) – Neural network, decision tree, SVM
  - Others (2%)

(1) Gaussian Model

Classification and authentication

- Every digraph latency exhibits a Gaussian distribution
- 26 X 26 digraphs – Flight time
  - E.g. TH, AB …
- Create profile for each user
- Measure similarity score

Zone of acceptance

\[
\begin{align*}
\mu - \delta \sigma & \quad \mu \quad \mu + \delta \sigma
\end{align*}
\]

- \( \mu \): mean
- \( \sigma^2 \): variance
- \( \delta \): distance
(2) Support Vector Machine (SVM)

A highly utilized classifier [1]

- Generates a region that separates majority of feature data related to a particular class
- By mapping the input vector into a high-dimensional feature space via the kernel function - linear, polynomial, sigmoid, or radial basis function
- Low energy consumption and high performance

(2) One Class Support Vector Machines (SVM)

What fits our authentication goal?

• Legitimate data assumed as positive class (+1)
• Anything else as negative class (-1)
• Gaussian Radial Base Kernel Function (RBF)
• Optimal kernel scale

\[ K(x, x') = \exp \left( -\frac{||x - x'||^2}{2\sigma^2} \right) \]

Where \( \sigma \in R \) is a kernel parameter and \( ||x - x'|| \) is the dissimilarity measure.
(3) Transfer Learning

Session I (Source Task)
- keystrokes
- features
- Source Classifier
- Model parameters

Session II (Target Task)
- keystrokes
- features
- Target Classifier
- Adapted model parameters

Knowledge Transfer
Shared Keystroke Dataset

Why?

• Generalization of results
• Benchmarking various algorithms

What is missing?

• Very few high quality datasets in the public domain
• Those available are mostly on short texts
• Some are not accessible

Long text datasets are fundamental for continuous authentication
### Related Work

**Characteristics of current publicly available datasets – long text**

<table>
<thead>
<tr>
<th>#Subject</th>
<th>#Sessions</th>
<th>Duration</th>
<th>Gap b/w Sessions</th>
<th>Clock Resolution</th>
<th>Keyboard variability</th>
<th>Gender (M:F)</th>
<th>Age</th>
</tr>
</thead>
<tbody>
<tr>
<td>Clarkson</td>
<td>39</td>
<td>2</td>
<td>1 hour</td>
<td>Mostly 1 or 2 month</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>MSU P1</td>
<td>51</td>
<td>2</td>
<td>10 – 16 min</td>
<td>Same day</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>MSU P2</td>
<td>30</td>
<td>Around 5</td>
<td>60 sec</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Ours</td>
<td>289</td>
<td>3</td>
<td>50 min</td>
<td>28 days</td>
<td>15 ms</td>
<td>Yes</td>
<td>204 : 85</td>
</tr>
</tbody>
</table>

“-” symbolizes a feature not present in the original paper
Desirable Characteristics

• Large subject number
• Characterized to reflect
  − Temporal aspects of typing patterns
  − Effect of keyboard layout variability
• Textual data included
• Mouse movements and system events data

Institutional Review Board (IRB) permission
Overview of Experiments

- A large scale data collection campaign
  - 4 months in two campaigns from Sept. to Dec. 2015 and Sept. to Dec. 2016
- 157 + 143 volunteers recruited
- 2 keystroke activities involved
  - Transcribed and free text
- 3 sessions for each participant
- Approximately 1 month between sessions
- 50 minutes for each session
- 4 different types of keyboards utilized
Dataset Design - 1

Keyboard variability
- Baseline subset
- Keyboard variation subset

<table>
<thead>
<tr>
<th></th>
<th>#subjects</th>
<th>#sessions</th>
<th>#keyboard types</th>
</tr>
</thead>
<tbody>
<tr>
<td>Baseline</td>
<td>157</td>
<td>3</td>
<td>1</td>
</tr>
<tr>
<td>Variation</td>
<td>132</td>
<td>3</td>
<td>3</td>
</tr>
</tbody>
</table>
Dataset Design - 2

Typing activities

- Transcription of fixed text
  - Steve Jobs' commencement speech at Stanford University, 2005

- Free composition
  - Survey questions and picture description
  - Routine tasks, e.g. checking and sending email, web surfing

"Describe your experience in school and college or high school. How do you think it compares to what you think is more interesting? What do you think is more challenging? What do you think is more fun? What do you think is more boring? What do you think is more rewarding?"

"Describe your experience in learning the concept of..."
Data Acquisition Tool

Active system logger

• Collect system events such as keyboard activity, mouse movement coordinates and mouse clicks
• Desktop based vs. web-based
• Clock resolution: ~15 ms

<table>
<thead>
<tr>
<th>CPU Ticks</th>
<th>Event</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>634564465190625000</td>
<td>Mouse Coordinates</td>
<td>464, 348</td>
</tr>
<tr>
<td>634564465190625000</td>
<td>Left Click</td>
<td></td>
</tr>
<tr>
<td>634564462834375000</td>
<td>New Process</td>
<td>chrome.exe</td>
</tr>
<tr>
<td>634564462895937500</td>
<td>Key Down</td>
<td>G</td>
</tr>
<tr>
<td>634564462897187500</td>
<td>Key Up</td>
<td>G</td>
</tr>
</tbody>
</table>
Data Anonymization and Quality Assurance

Privacy protection
  • Rule-Based sanitization tool

Secure Transportation
  • Data transmission tool

Quality Assurance
  • Incomplete data files removed
    - Several subjects removed
    - 300 subjects -> 289 subjects
Evaluation

Statistics
  • Parameterize various experiments

Experiments
  • Show overall quality
Statistics - 1

Number of raw keystrokes
- 5,800 keystrokes each subject per session
- 17,600 keystrokes each subject
- Minimum 10,000 keystrokes per subject

<table>
<thead>
<tr>
<th></th>
<th># keys Session 1</th>
<th></th>
<th># keys Session 2</th>
<th></th>
<th># keys Session 3</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Task 1</td>
<td>Task 2</td>
<td>Sum</td>
<td>Task 1</td>
<td>Task 2</td>
</tr>
<tr>
<td>Average</td>
<td>3729</td>
<td>2082</td>
<td>5811</td>
<td>3666</td>
<td>2101</td>
</tr>
<tr>
<td>Stdev</td>
<td>467</td>
<td>650</td>
<td>891</td>
<td>393</td>
<td>750</td>
</tr>
<tr>
<td>Min</td>
<td>2334</td>
<td>393</td>
<td>3426</td>
<td>2012</td>
<td>175</td>
</tr>
<tr>
<td>Max</td>
<td>5332</td>
<td>5235</td>
<td>9506</td>
<td>6611</td>
<td>8751</td>
</tr>
</tbody>
</table>
Statistics - 2

Time intervals between sessions
• 28 days in average

Gender information
• Female 85
• Male 204

<table>
<thead>
<tr>
<th></th>
<th>S1 to S2 (days)</th>
<th>S2 to S3 (days)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Average</td>
<td>28.83</td>
<td>27.35</td>
</tr>
<tr>
<td>Stdev</td>
<td>5.99</td>
<td>5.11</td>
</tr>
<tr>
<td>Max</td>
<td>47</td>
<td>42</td>
</tr>
<tr>
<td>Min</td>
<td>18</td>
<td>14</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Subset</th>
<th># Male</th>
<th># Female</th>
</tr>
</thead>
<tbody>
<tr>
<td>Baseline subset</td>
<td>115</td>
<td>43</td>
</tr>
<tr>
<td>Keyboard variation subset</td>
<td>89</td>
<td>42</td>
</tr>
<tr>
<td>Sum</td>
<td>204</td>
<td>85</td>
</tr>
</tbody>
</table>
### Dataset

#### File ID

<table>
<thead>
<tr>
<th>Sequence No.</th>
<th>Assignment</th>
<th>Value</th>
<th>5th</th>
<th>6th</th>
<th>7th</th>
</tr>
</thead>
<tbody>
<tr>
<td>1-2-3-4th</td>
<td>Use ID</td>
<td>0000 ~ 9999</td>
<td>0 ~ 2</td>
<td>0 ~ 4</td>
<td>0 ~ 1</td>
</tr>
<tr>
<td>5th</td>
<td>Session #</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6th</td>
<td>Keyboard type code</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7th</td>
<td>Task #</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

[Type: TXT File
Size: 226 KB
Date modified: 2/9/2016 2:20 PM](http://cubs.buffalo.edu/research/datasets)
Outline of the Talk

Introduction
- General approach to continuous authentication

Keystroke dynamics and mouse movements
- Feature selection
- Methodology - Gaussian model, SVM, transfer learning
- Datasets and anonymization

Results
- GMM, SVM, transfer learning

Research directions
- Secondary features
- Deep learning
- Adversarial learning
- Extension to network of smart devices
Metrics

Evaluation criteria

• False Reject Rate (FRR) – falsely denied genuine users – (Type 1 error)
• False Accept Rate (FAR) – falsely accepted unauthorized users – (Type 2 error)
• Equal Error Rate (EER) – overall accuracy – (Cross-over error rate)
• Receiver Operating Characteristic (ROC) – true positive rate vs. false positive rate
• Area under the curve (AUC) – scalar representation of ROC
(1) GMM as the Classification Algorithm

- GMM can represent complex and hard-to-map data to an understandable and distinguishable format
- Perturbations can be acquired
- Easy to implement
Traditional Approaches

- Considering all digraphs: ab, ac, ..., zz
- Various distance (δ) values
- GMMs with different number of components
- Leave-one-out cross-validation

Gaussian Mixture Model

\[ \mu: \text{mean} \]
\[ \sigma^2: \text{variance} \]
\[ \delta: \text{distance} \]
Keystroke Dynamics with GMM

Hard to separate with 1G
Somewhat better
More separable
Results

- Clarkson dataset (39 users) is used
- Word-initiation effect
- Curse of dimensionality
- Presence of singularities
False Accept/Reject Rate

<table>
<thead>
<tr>
<th>1G</th>
<th>1.3%</th>
</tr>
</thead>
<tbody>
<tr>
<td>2G</td>
<td>0.08%</td>
</tr>
<tr>
<td>3G</td>
<td>5.88%</td>
</tr>
<tr>
<td>4G</td>
<td>3.36%</td>
</tr>
<tr>
<td>5G</td>
<td>5.28%</td>
</tr>
</tbody>
</table>
Is GMM Enough?

- No winner model
- Consolidating the strengths
- Anomalous characteristics are suppressed

We take one step further!
Information Fusion

- Multiple sources, modalities or decisions
- Parameters from different classifiers are consolidated
- More refined set of criteria
Fusion of GMMs

Naïve Bayes
\[ A_{ij}^{-1} = B_{i}^{-1} + C_{j}^{-1} \]
\[ a_{ij} = A_{ij}(B_{i}^{-1}b_i + C_{j}^{-1}c_j) \]
\[ r_{ij} = \frac{p_i q_j}{\sum_{k=1}^{M_b} \sum_{l=1}^{M_c} p_k q_l} \]

Covariance Intersection
\[ A_{ij}^{-1} = \omega_{ij} B_{i}^{-1} + (1 - \omega_{ij}) C_{j}^{-1} \]
\[ a_{ij} = A_{ij}\left(\omega_{ij} B_{i}^{-1}b_i + (1 - \omega_{ij}) C_{j}^{-1}c_j\right) \]
\[ r_{ij} = \frac{\omega_{ij} p_i + (1 - \omega_{ij}) q_j}{\sum_{k=1}^{M_b} \sum_{l=1}^{M_c} \omega_{kl} p_k + (1 - \omega_{kl}) q_l} \]

Chernoff Information
\[ A_{ij}^{-1} = \omega B_{i}^{-1} + (1 - \omega) C_{j}^{-1} \]
\[ a_{ij} = A_{ij}\left(\omega B_{i}^{-1}b_i + (1 - \omega) C_{j}^{-1}c_j\right) \]
\[ r_{ij} = \frac{p_i^{\omega} q_j^{(1-\omega)}}{\sum_{k=1}^{M_b} \sum_{l=1}^{M_c} p_k^{\omega} q_l^{(1-\omega)}} \]
Fusion Results

- Lower error rates
- Regular trend-lines
- Robust classifier
Time Performance

![Graph showing time performance with different numbers of components and configurations. The x-axis represents the number of components (1comp, 2comp, 3comp, 4comp, no limit), and the y-axis represents elapsed time (s). Different configurations are represented by various markers and colors: 1G & 4G, 1G & 5G, 2G & 3G, 2G & 4G, 2G & 5G, 3G & 4G, 3G & 5G, 4G & 5G.]
### Feature alignment method

- Each observation holds a single row
- Observations from different features in different columns
- Rest cells filled with 0

#### Feature matrix

<table>
<thead>
<tr>
<th>Feature 1</th>
<th>Feature 2</th>
<th>\cdots</th>
<th>Feature N</th>
</tr>
</thead>
<tbody>
<tr>
<td>$f^1_1$</td>
<td>0</td>
<td>\cdots</td>
<td>0</td>
</tr>
<tr>
<td>\vdots</td>
<td>\vdots</td>
<td>\ddots</td>
<td>\vdots</td>
</tr>
<tr>
<td>$f^k_1$</td>
<td>0</td>
<td>\cdots</td>
<td>0</td>
</tr>
<tr>
<td>0</td>
<td>$f^1_2$</td>
<td>\cdots</td>
<td>0</td>
</tr>
<tr>
<td>0</td>
<td>\vdots</td>
<td>\ddots</td>
<td>\vdots</td>
</tr>
<tr>
<td>0</td>
<td>$f^k_2$</td>
<td>\cdots</td>
<td>0</td>
</tr>
<tr>
<td>\vdots</td>
<td>\vdots</td>
<td>\ddots</td>
<td>\vdots</td>
</tr>
<tr>
<td>0</td>
<td>0</td>
<td>\cdots</td>
<td>$f^1_n$</td>
</tr>
<tr>
<td>\vdots</td>
<td>\vdots</td>
<td>\ddots</td>
<td>\vdots</td>
</tr>
<tr>
<td>0</td>
<td>0</td>
<td>\cdots</td>
<td>$f^k_n$</td>
</tr>
</tbody>
</table>

(2) SVM as the Classification Algorithm
Experiment Setting

- Data partition
  - Training sets (80%) and Testing set (20%)
- SVM packages on MATLAB
- Model trained with genuine data (positive class)
- One vs. All testing strategy
- Optimal kernel scale
SVM Results

<table>
<thead>
<tr>
<th>Digraph set</th>
<th>Kernel Scale</th>
<th>Training Time</th>
<th>Testing Time</th>
<th>AUC</th>
<th>EER %</th>
</tr>
</thead>
<tbody>
<tr>
<td>hrtate/eehn</td>
<td>0.36</td>
<td>0.20</td>
<td>0.0128</td>
<td>0.9973</td>
<td>2.58</td>
</tr>
<tr>
<td>hnhrtate/adeehnor</td>
<td>0.46</td>
<td>0.28</td>
<td>0.0169</td>
<td>0.9979</td>
<td>2.94</td>
</tr>
<tr>
<td>12 digraphs</td>
<td>0.54</td>
<td>0.55</td>
<td>0.0275</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>14 digraphs</td>
<td>0.56</td>
<td>0.67</td>
<td>0.0379</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>16 digraphs</td>
<td>0.65</td>
<td>0.84</td>
<td>0.0448</td>
<td>1</td>
<td>0</td>
</tr>
</tbody>
</table>
(3) Transfer Learning as Classification Algorithm

Authenticate an enrolled user in a different environment with least amount of re-training

- Knowledge acquired in previous session is transferred via parameters that contain classifier info
- There is a source system and a target system
- Use two different adaptive SVMs with linear and Gaussian kernels
- Source profile works as a regularizer of target profile in the SVM cost function
- Uses a small no. of samples from the target system

Transfer learning in other domains

- Concept drift in data mining
- Incremental learning
- Cross-domain learning
Intra-User Variability

John

Normal conditions

Keyboard

Time

Angry
Separating Hyperplanes
SVM Projection

Classic SVM
\[
\min J(w, b, \xi) = \frac{1}{2} \|w\|^2 + C \sum_{i=1}^{l} \xi_i \\
\text{s.t. } y_i (w^T x_i + b) \geq 1 - \xi_i \\
\xi_i \geq 0, \ i = 1, \cdots, l
\]

A-SVM [1]
\[
\min \left( \frac{1}{2} \|w_t - \Gamma w_s\|^2 + C \sum_{i=1}^{l} \xi_i \right)
\]

Deformable Adaptive SVM [2]
\[
\min \left( \frac{1}{2} \|w_t - \Gamma f(w_s)\|^2 + \lambda \cdot \Delta + C \sum_{i=1}^{l} \xi_i \right)
\]

Projective Model Transfer SVM [2]
\[
\min \left( \frac{1}{2} \|w_t\|^2 + \Gamma \|P w_t\|^2 + C \sum_{i=1}^{l} \xi_i \right)
\]

Results*

Figure 4: ROC with various step sizes

<table>
<thead>
<tr>
<th>Sample Size:</th>
<th>1</th>
<th>5</th>
<th>11</th>
<th>15</th>
</tr>
</thead>
<tbody>
<tr>
<td>Classic SVM</td>
<td>71.28 ± 11.39</td>
<td>80.01 ± 9.01</td>
<td>93.71 ± 4.55</td>
<td>96.20 ± 3.42</td>
</tr>
<tr>
<td>A-SVM</td>
<td>80.11 ± 2.61</td>
<td>85.54 ± 3.49</td>
<td>93.33 ± 2.33</td>
<td>94.86 ± 1.84</td>
</tr>
<tr>
<td>PMT-SVM</td>
<td>82.64 ± 9.00</td>
<td>87.12 ± 6.95</td>
<td>95.86 ± 2.74</td>
<td>95.76 ± 2.75</td>
</tr>
<tr>
<td>DA-SVM</td>
<td>84.39 ± 4.03</td>
<td>92.53 ± 3.42</td>
<td>97.18 ± 1.10</td>
<td>97.37 ± 1.01</td>
</tr>
</tbody>
</table>

AUC values are multiplied with 100 for higher precision

Comparison of SVM Algorithms

![Graph showing comparison of SVM algorithms](image-url)
Performance

![Graph showing performance comparison between different SVM methods.](image-url)
Some Observations

- GMM provides improvement over single Gaussian
- Fusion can improve the accuracy
- SVM can be utilized for long-text data efficiently
- Intra-user variability is important
  - Can be addressed by using transfer learning
Outline of the Talk

Introduction
  • General approach to continuous authentication

Keystroke dynamics and mouse movements
  • Feature selection
  • Methodology - Gaussian model, SVM, transfer learning
  • Datasets and anonymization

Results
  • GMM, SVM, transfer learning

Research directions
  • Secondary features
  • Deep learning
  • Adversarial learning
  • Extension to network of smart devices
Secondary Features – Achieving More with Less

Punctuations
• Period
• Comma
• …

Functional keys
• Shift
• Backspace

Number and others
• 1, 2, 3 …
• Dash

Compare with primary features
• Primarily from 26 letters (A to Z)
Feature Extraction

Dwell time
- Period, Comma, Tab, Space, Enter, Backspace, Arrow keys, Number keys, Dash

Flight time
- Period – Space and Comma – Space
- Shift – [a-z 0-9]
- Ctrl – [a-z]
Available Secondary Features in the Clarkson’s dataset

<table>
<thead>
<tr>
<th>Feature</th>
<th>Feature # Type</th>
<th>Average # records</th>
<th># Occurrence (out of 34)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Backspace</td>
<td>Dwell</td>
<td>1010.47</td>
<td>34</td>
</tr>
<tr>
<td>Space</td>
<td>Dwell</td>
<td>2535.85</td>
<td>34</td>
</tr>
<tr>
<td>Number 1</td>
<td>Dwell</td>
<td>64.03</td>
<td>34</td>
</tr>
<tr>
<td>Number 2</td>
<td>Dwell</td>
<td>36.23</td>
<td>31</td>
</tr>
<tr>
<td>Number 3</td>
<td>Dwell</td>
<td>32.35</td>
<td>31</td>
</tr>
<tr>
<td>Shift_I</td>
<td>Flight</td>
<td>89.03</td>
<td>31</td>
</tr>
<tr>
<td>Shift_N</td>
<td>Flight</td>
<td>33.5</td>
<td>28</td>
</tr>
<tr>
<td>Shift_T</td>
<td>Flight</td>
<td>22.6</td>
<td>30</td>
</tr>
<tr>
<td>Shift_1</td>
<td>Flight</td>
<td>29.75</td>
<td>32</td>
</tr>
<tr>
<td>Comma</td>
<td>Dwell</td>
<td>118.38</td>
<td>34</td>
</tr>
<tr>
<td>Comma_Space</td>
<td>Flight</td>
<td>116.38</td>
<td>34</td>
</tr>
<tr>
<td>Period</td>
<td>Dwell</td>
<td>162.68</td>
<td>34</td>
</tr>
<tr>
<td>Period_Space</td>
<td>Flight</td>
<td>155.71</td>
<td>34</td>
</tr>
<tr>
<td>Dash</td>
<td>Dwell</td>
<td>19.53</td>
<td>30</td>
</tr>
<tr>
<td>LeftArrow</td>
<td>Dwell</td>
<td>30.94</td>
<td>16</td>
</tr>
<tr>
<td>RightArrow</td>
<td>Dwell</td>
<td>33.15</td>
<td>13</td>
</tr>
</tbody>
</table>
Single Feature Evaluation

- 8 features (groups)
- Best \(\rightarrow\) Shift group
- Worst \(\rightarrow\) Backspace
- Comma \(\rightarrow\) Comma-Space
- Period \(\rightarrow\) Period-Space

<table>
<thead>
<tr>
<th>Kernel Scale</th>
<th>Backspace</th>
<th>Space</th>
<th>Num(123)</th>
<th>Shift(INT1)</th>
<th>Comma</th>
<th>Comma-Space</th>
<th>Period</th>
<th>Period-Space</th>
</tr>
</thead>
<tbody>
<tr>
<td>AUC</td>
<td>0.5401</td>
<td>0.7307</td>
<td>0.8323</td>
<td>0.8991</td>
<td>0.7877</td>
<td>0.7599</td>
<td>0.7493</td>
<td>0.7636</td>
</tr>
<tr>
<td>EER (%)</td>
<td>47.5</td>
<td>27.8</td>
<td>23.53</td>
<td>14.71</td>
<td>28.88</td>
<td>29.41</td>
<td>32.35</td>
<td>29.41</td>
</tr>
</tbody>
</table>
Overall Evaluation

- 2 : Comma and Period-Space
- 4 : + Left & Right arrow
- 5 : + Dash
- 9 : + Shift [ I N T 1 ]
- 12 : + Number [ 1 2 3 ]
- 13 : + Space

(Data Sampling )

<table>
<thead>
<tr>
<th># Features</th>
<th>2</th>
<th>4</th>
<th>5</th>
<th>9</th>
<th>12</th>
<th>13</th>
</tr>
</thead>
<tbody>
<tr>
<td>AUC</td>
<td>0.8521</td>
<td>0.8932</td>
<td>0.9088</td>
<td>0.9530</td>
<td>0.9897</td>
<td>0.9937</td>
</tr>
<tr>
<td>EER (%)</td>
<td>21.57</td>
<td>19.61</td>
<td>16.22</td>
<td>6.95</td>
<td>3.83</td>
<td>2.94</td>
</tr>
</tbody>
</table>
## Comparison with Primary Features

<table>
<thead>
<tr>
<th>Study</th>
<th># users</th>
<th>EER (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Atam et al. [8]</td>
<td>43</td>
<td>8.77</td>
</tr>
<tr>
<td>Killourhy et al. [12]</td>
<td>51</td>
<td>10.2</td>
</tr>
<tr>
<td>Giot et al. [7]</td>
<td>100</td>
<td>6.96</td>
</tr>
<tr>
<td>Gabriel et al. [1]</td>
<td>24</td>
<td>1.57</td>
</tr>
<tr>
<td>Rahman et al. [13]</td>
<td>50</td>
<td>10</td>
</tr>
<tr>
<td>Kaneko et al. [10]</td>
<td>51</td>
<td>0.84</td>
</tr>
<tr>
<td>Ceker et al. [4]</td>
<td>30</td>
<td>0.08</td>
</tr>
<tr>
<td>Our work</td>
<td>34</td>
<td>2.94</td>
</tr>
</tbody>
</table>
Deep Learning

- Current solutions in keystroke dynamics
  - Use timing information between the keys separately (digraph, trigraph, n-graphs), fusion
  - Trial and error works, but unwieldy
  - Computational complexity increases exponentially
- Scaling up (no. of users) would mean lower accuracy
- CNNs can provide a deeper architecture and unify ML techniques by consolidating the power of various features
- CNN has been successfully applied in vision, speech, NLP
Adversarial Learning

Attack scenarios
- Adjust attack based on the feedback on where the typing was different from the legitimate users
- Synthetic forgery attacks designed to mimic the legitimate users based on their typing profiles

Possible solution
- Combine multiple biometrics features
  - E.g., Keystroke dynamic and mouse movement
Extension to Smartphone Environment

- Portable mobile devices have become ubiquitous
  - Sensitive data, business usage
  - Owner may leave this device unlocked
  - There is security risk
- What features can be extracted?
  - User activities on the mobile devices touchscreen - clicks, speed, angles of movement, number of clicks during a session, pressure on the touchscreen
  - Accelerometer, rotation vector and orientation sensor to generate the feature vectors
- We can apply a variety of ML algorithms in this context
Publications


