Hold & Sign: A Novel Behavioral Biometrics for Smartphone User Authentication

Attaullah Burio\textsuperscript{1}, Bruno Crispo\textsuperscript{2}, Filippo Del Frari\textsuperscript{1}
Konrad Wrona\textsuperscript{3}

\textsuperscript{1}University of Trento
\textsuperscript{2}DistriNet - KULeuven
\textsuperscript{3}NCIA - NATO

\{attaullah.burio, filippo.delfrari\}@unitn.it
bruno.crispo@cs.kuleuven.be

November 15, 2017
Handwritten Biometric Recognition

- Process of identifying the author of a given text.
  - Handwritten signature is a specific instance.

Barack Obama
Handwritten Biometric Recognition

- Process of identifying the author of a given text.
  - Handwritten signature is a specific instance.
- Wide social and legal acceptance in daily life.
  - Witness intentions (signing a contract)
  - Indicate physical presence (signing in for work)

Barack Obama
Handwritten Biometric Recognition

- In static systems, scanned image is processed and compared with the stored image to authenticate.

Handwritten Biometric Recognition

- In static systems, scanned image is processed and compared with the stored image to authenticate.
- In **dynamic/online** systems, signature is acquired (by stylus or pen) and processed in real time and extracted global features are compared to authenticate.

---

¹http://www.sutisoft.com/sutisign/
Handwritten Biometric Recognition

- In static systems, scanned image is processed and compared with the stored image to authenticate.
- In dynamic/online systems, signature is acquired (by stylus or pen) and processed in real time and extracted global features are compared to authenticate.
  - spatial coordinates $x(t)$
  - spatial coordinates $x(t)$
  - pressure $p(t)$
  - inclination angles $in(t)$, etc.

Handwritten Biometric Recognition

- In static systems, scanned image is processed and compared with the stored image to authenticate.
- In dynamic/online systems, signature is acquired (by stylus or pen) and processed in real time and extracted global features are compared to authenticate.
  - spatial coordinates $x(t)$
  - spatial coordinates $x(t)$
  - pressure $p(t)$
  - inclination angles $\text{in}(t)$, etc.

---

Handwritten Biometric Recognition

- In static systems, scanned image is processed and compared with the stored image to authenticate.
- In dynamic/online systems, signature is acquired (by stylus or pen) and processed in real time and extracted global features are compared to authenticate.
  - spatial coordinates $x(t)$
  - spatial coordinates $x(t)$
  - pressure $p(t)$
  - inclination angles $\text{in}(t)$, etc.

- iSignOn, SignEasy and eSignature (sutisoft\textsuperscript{1}) use signature image to authenticate.

\textsuperscript{1}http://www.sutisoft.com/sutisign/
Design Requirements

- Must be able to run on the smartphones

---

Design Requirements

- Must be able to run on the smartphones
- **Usability is very important for user acceptance**
  - Must be quick
  - easy to use
  - “I Feel Like I’m Taking Selfies All Day!”
  - “It can be a bit embarrassing to hold your phone up in a weird position while in public just to unlock it!”

---

Design Requirements

- Must be able to run on the smartphones
- **Usability is very important for user acceptance**
  - Must be quick
  - easy to use
  - “I Feel Like I’m Taking Selfies All Day!”
  - “It can be a bit embarrassing to hold your phone up in a weird position while in public just to unlock it!”\(^2\)
- **Must be secure, i.e., should provide acceptable security.**

Our Proposed Solution

Hold & Sign
Hold & Sign

- Bi-Modal System
Hold & Sign

- Bi-Modal System
  - Sign - touch (points pressed while signing)
Hold & Sign

- Bi-Modal System
  - Sign - touch (points pressed while signing)
  - Hold - phone’s micro-movements while signing

No additional hardware required.
Hold & Sign

- **Bi-Modal System**
  - Sign - touch (points pressed while signing)
  - Hold - phone’s micro-movements while signing
- **No additional hardware required.**
Smartphone Sensors

- IMU combo
- Magnetometer
- MEMS microphones
- Pressure sensor
- Humidity + Temperature sensor
- BAW filters and duplexers
- Antenna tuner
Smartphone Sensors

- **Accelerometer** - 3 variants, i.e., RAW, LPF and HPF
  - accelerations in 3 dimensions
- **Gravity**
  - gravity forces in 3 dimensions
- **Magnetometer**
  - Strength and direction of magnetic field in 3 dimensions
- **Touchscreen**
- **Hold & Sign** utilizes the existing hardware
Smartphone Sensors

- **Accelerometer** - 3 variants, i.e., RAW, LPF and HPF
  - accelerations in 3 dimensions
- **Gravity**
  - gravity forces in 3 dimensions
- **Magnetometer**
  - **Strength and direction of magnetic field** in 3 dimensions
- **Touchscreen**
- **Hold & Sign** utilizes the existing hardware
Smartphone Sensors

- **Accelerometer** - 3 variants, i.e., RAW, LPF and HPF
  - accelerations in 3 dimensions
- **Gravity**
  - gravity forces in 3 dimensions
- **Magnetometer**
  - Strength and direction of magnetic field in 3 dimensions
- **Touchscreen**
- **Hold & Sign** utilizes the existing hardware
Validation

- **Data Collection**
  - 30 samples in each activity (total 90)
  - 3 activities (sitting, standing, walking)

- 30 participants (22 male), all of them Masters/PhD students, Google Nexus 5
Verifiers

We used a set of four conceptually different verifiers.

- 3-NN (K-Nearest Neighbor)
- 3-layer MLP (Multilayer Perceptron)
- BayesNet
- Random Forest
One Class Classification (Anomaly Detection)

- **Verifiers**
  - We used a set of four conceptually different verifiers.
    - 3-NN (K-Nearest Neighbor)
    - 3-layer MLP (Multilayer Perceptron)
    - BayesNet
    - Random Forest

- After initial experiments we selected MLP since it performed consistently across situations (TAR=79%, FAR=0.1%)
Feature Selection

- 93 features (80 hold and 13 sign)
Feature Selection

- 93 features (80 hold and 13 sign)
- Recursive Feature Elimination (RFE) Method

![Graph showing cross-validation score vs. number of features selected]

- # of optimal features for fused data (Sitting state): 10
- # of optimal features for fused data (Standing state): 11
- # of optimal features for fused data (Walking state): 11
Feature Selection

- 93 features (80 hold and 13 sign)
- Recursive Feature Elimination (RFE) Method
  - Scikit-learn, 10-fold cross validation
Results

- The results are averaged over 30 users
Results

- The results are averaged over 30 users
  - Inter-Activity

<table>
<thead>
<tr>
<th></th>
<th>TAR(%)</th>
<th>FRR (%)</th>
<th>FAR (%)</th>
<th>TRR(%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>MLP</td>
<td>94.8</td>
<td>5.2</td>
<td>3.1</td>
<td>96.9</td>
</tr>
</tbody>
</table>

![Box plots showing True Acceptance Rate for different activities.](image)
Results

- The results are averaged over 30 users
  - Inter-Activity

![Box plots showing True Acceptance Rate (%) for Sitting, Standing, and Walking activities.]

- Combined-Activities (RFE)

<table>
<thead>
<tr>
<th>Classifier</th>
<th>TAR(%)</th>
<th>FRR (%)</th>
<th>FAR (%)</th>
<th>TRR(%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>MLP</td>
<td>94.8</td>
<td>5.2</td>
<td>3.1</td>
<td>96.9</td>
</tr>
</tbody>
</table>
## Performance

### Sample Acquisition Time

<table>
<thead>
<tr>
<th>Method</th>
<th>Sample Acquisition Time (s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hold &amp; Sign</td>
<td>3.5</td>
</tr>
<tr>
<td>PIN</td>
<td>3.7</td>
</tr>
<tr>
<td>Voice</td>
<td>5.15</td>
</tr>
<tr>
<td>Face</td>
<td>5.55</td>
</tr>
<tr>
<td>Password</td>
<td>7.46</td>
</tr>
<tr>
<td>Face + Voice</td>
<td>7.63</td>
</tr>
<tr>
<td>Gesture</td>
<td>8.10</td>
</tr>
<tr>
<td>Gesture + Voice</td>
<td>9.91</td>
</tr>
</tbody>
</table>

### Performance

#### Testing Time

<table>
<thead>
<tr>
<th>Scheme</th>
<th>Device</th>
<th>Classifiers</th>
<th># Users</th>
<th>Testing Time (s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hold &amp; Sign</td>
<td>Nexus 5</td>
<td>MLP</td>
<td>30</td>
<td>0.215 – 0.250</td>
</tr>
<tr>
<td>Lee &amp; Lee(^3)</td>
<td>Nexus 5</td>
<td>SVM</td>
<td>8</td>
<td>20</td>
</tr>
<tr>
<td>Li et al(^4).</td>
<td>Motorolla Droid</td>
<td>Sliding patterns</td>
<td>75</td>
<td>0.648</td>
</tr>
<tr>
<td>Nickel et al(^5).</td>
<td>Motorolla Milestone</td>
<td>KNN</td>
<td>36</td>
<td>30</td>
</tr>
</tbody>
</table>


\(^4\)Li, X. Zhao, and G. Xue, Unobservable re-authentication for smartphones. in NDSS, 2013.

In order to check the overhead resulting from use of the application (in different steps), we terminated all the running applications and all Google services, switched off WiFi, Bluetooth, and cellular radios. The screen was kept running for the entire duration of the experiment with brightness at the lowest level and automatic brightness adjustment disabled.

We used Trepn to profile power usage.

Our app consumes ≈1000mW (for 35 attempts) all in the stages of decision making (less than the power consumed in one minute of phone call (≈1054mW)) during all the stages.

---

Performance

▶ Power Consumption

▶ In order to check the overhead resulting from use of the application (in different steps), we terminated all the running applications and all Google services, switched off WiFi, Bluetooth, and cellular radios. The screen was kept running for the entire duration of the experiment with brightness at the lowest level and automatic brightness adjustment disabled\(^6\). We used Trepn\(^7\) to profile power usage.

---


\(^7\)https://play.google.com/store/apps/details?id=com.quicinc.trepn

\(^8\)https://developer.qualcomm.com/blog/mobile-apps-and-power-consumption-basics-part-1
Performance

- **Power Consumption**
  - In order to check the overhead resulting from use of the application (in different steps), we terminated all the running applications and all Google services, switched off WiFi, Bluetooth, and cellular radios. The screen was kept running for the entire duration of the experiment with brightness at the lowest level and automatic brightness adjustment disabled. We used Trepn to profile power usage.
  - **our app consumes \( \approx 1000\text{mW} \) (for 35 attempts) all in the stages of decision making (less than the power consumed in one-minute of phone call (\( \approx 1054\text{mW} \)) during all the stages.)

---

Usability

- Trade-off between training & accuracy.

---

Trade-off between training & accuracy.
Usability

- Trade-off between training & accuracy.

- System Usability Scale (SUS⁹) Evaluation - 68.3% (72% or higher is considered good.)

Hold & Sign in Action

- This one-minute video shows the working of Hold & Sign.
Open Issues

- Extend **number of testers** and run tests on the **field**.
Open Issues

- Extend number of testers and run tests on the field.
- Reduce training time without compromising (too much) security
Open Issues

- Extend number of testers and run tests on the field.
- Reduce training time without compromising (too much) security
- **Unplanned Situations**
Open Issues

- Extend number of testers and run tests on the field.
- Reduce training time without compromising (too much) security
- Unplanned Situations
- Assess its robustness to multiple attacks.
Open Issues

- Extend number of testers and run tests on the field.
- Reduce training time without compromising (too much) security.
- Unplanned Situations
- Assess its robustness to multiple attacks.
- Variability depending on OS (iOS) and on HW (different phones, sensors)
Thank You!

- attaullah.buriro@unitn.it
- bruno.crispo@cs.kuleuven.be