On the Security of Picture Gesture Authentication

Ziming Zhao†‡, Gail-Joon Ahn†‡, Jeong-Jin Seo†, Hongxin Hu§

†Arizona State University ‡GFS Technology §Delaware State University
Picture Gesture Authentication (PGA)

• A built-in feature in Microsoft Windows 8

• 60 million Windows 8 licenses have been sold

• 400 million computers and tablets will run Windows 8 in one year
Welcome to picture password

Picture password is a new way to help you protect your touchscreen PC. You choose the picture -- and the gestures you use with it -- to create a password that's uniquely yours.

When you've chosen a picture, you "draw" directly on the touchscreen to create a combination of taps, straight lines, or circles. The size, position, and direction of your gestures become part of your picture password.

Choose picture
How PGA Works

• Autonomous picture selection by users

Welcome to picture password

Picture password is a new way to help you protect your touchscreen PC. You choose the picture -- and the gestures you use with it -- to create a password that's uniquely yours.

When you've chosen a picture, you "draw" directly on the touchscreen to create a combination of taps, straight lines, or circles. The size, position, and direction of your gestures become part of your picture password.
How PGA Works

• Three types of gestures are allowed
  • Tap
  • Circle
  • Line
Research Questions

1. How to understand user-choice patterns in PGA?
   • Background Pictures
   • Gesture Location
   • Gesture Type
   • Gesture Order

2. How to use these patterns to guess PGA password?
Outline

Part 1: Analysis of more than 10,000 PGA passwords collected from user studies

Part 2: A fully automated attack framework on PGA

Part 3: Attack results on collected passwords
Part 1: User Studies

1. Web-based PGA system
   • Similarity to Windows PGA
     • Workflow
     • Appearance

2. Data collection

3. Analysis: survey and results
Part 1: User Studies

- Dataset-1
  - ASU undergraduate computer security class (Fall 2012)
  - 56 participants
  - 58 unique pictures
  - 86 passwords
  - 2,536 login attempts
Part 1: User Studies

- Dataset-2
  - Scenario: The password is used to protect your bank account
  - Amazon MTurk
  - 15 pictures selected in advance
  - 762 participants
  - 10,039 passwords
Part 1: User Studies

- Survey questions
  - General information of the subject
  - General feeling towards PGA
  - How she/he selects a background picture
  - How she/he selects a password
Part 1: User-choice Patterns

Background Picture

People, Civilization, Landscape, Computer-generated, Animal, Others

Dataset-1

Dataset-2 Survey
Part 1: User-choice Patterns
Why or why not picture of people

- Advocates:
  i) it is more friendly
     ‘The image was special to me so I enjoy seeing it when I log in’
  ii) it is easier for remembering passwords
     ‘Marking points on a person is easier to remember’
  iii) it makes password more secure
     ‘The picture is personal so it should be much harder for someone to guess the password’

- Others:
  i) leak his or her identify or privacy
     ‘revealing myself or my family to anyone who picks up the device’
Part 1: User-choice Patterns

Background Picture

People, Civilization, Landscape, Computer-generated, Animal, Others

Dataset-1

Dataset-2 Survey
Part 1: User-choice Patterns

Why computer-generated pictures

• Dataset-1 population characteristics:
  • 81.8% Male
  • 63.6% Age 18-24, 24.0% Age 25-34
  • 100% College students

• Survey answers:
  • ‘computer game is something I am interested [in] it’
  • ‘computer games picture is personalized to my interests and enjoyable to look at’
Part 1: User-choice Patterns

Why computer-generated pictures

- Dataset-1 population characteristics:
  - 81.8% Male
  - 63.6% Age 18-24, 24.0% Age 25-34

The background picture tells much about the user's identity, personality and interests.

- Survey answers:
  - 'computer game is something I am interested [in] it'
  - 'computer games picture is personalized to my interests and enjoyable to look at'
## Part 1: User-choice Patterns

### Gesture Locations

Which of the following best describes what you are considering when you choose locations to perform gestures?

<table>
<thead>
<tr>
<th>Description</th>
<th>Dataset-1</th>
<th>Dataset-2</th>
</tr>
</thead>
<tbody>
<tr>
<td>I try to find locations where special <strong>objects</strong> are.</td>
<td>72.7%</td>
<td>59.6%</td>
</tr>
<tr>
<td>I try to find locations where some special <strong>shapes</strong> are.</td>
<td>24.2%</td>
<td>21.9%</td>
</tr>
<tr>
<td>I try to find locations where <strong>colors</strong> are different from their surroundings.</td>
<td>0%</td>
<td>8.7%</td>
</tr>
<tr>
<td>I randomly choose a location to draw without thinking about the background picture.</td>
<td>3.0%</td>
<td>10.1%</td>
</tr>
</tbody>
</table>
## Part 1: User-choice Patterns

### Gesture Locations

Which of the following best describes what you are considering when you choose locations to perform gestures?

<table>
<thead>
<tr>
<th>Gesture Locations</th>
<th>Dataset-1</th>
<th>Dataset-2</th>
</tr>
</thead>
<tbody>
<tr>
<td>I try to find locations where special <strong>objects</strong> are.</td>
<td>72.7%</td>
<td>59.6%</td>
</tr>
<tr>
<td>I try to find locations where some special <strong>shapes</strong> are.</td>
<td>24.2%</td>
<td>21.9%</td>
</tr>
<tr>
<td>I try to find locations where <strong>colors</strong> are different from their surroundings.</td>
<td>0%</td>
<td>8.7%</td>
</tr>
<tr>
<td>I randomly choose a location to draw without thinking about the background picture.</td>
<td>3.0%</td>
<td>10.1%</td>
</tr>
</tbody>
</table>
Part 1: User-choice Patterns

Gesture Locations

Which of the following best describes what you are considering when you choose locations to perform gestures?

<table>
<thead>
<tr>
<th></th>
<th>Dataset-1</th>
<th>Dataset-2</th>
</tr>
</thead>
<tbody>
<tr>
<td>I try to find locations where special <strong>objects</strong> are.</td>
<td>72.7%</td>
<td>59.6%</td>
</tr>
<tr>
<td>I try to find locations where some special <strong>shapes</strong> are.</td>
<td>24.2%</td>
<td>21.9%</td>
</tr>
<tr>
<td>I try to find locations where <strong>colors</strong> are different from their surroundings.</td>
<td>0%</td>
<td>8.7%</td>
</tr>
<tr>
<td>I randomly choose a location to draw without thinking about the background picture.</td>
<td>3.0%</td>
<td>10.1%</td>
</tr>
</tbody>
</table>
Part 1: User-choice Patterns

Gesture Locations

Which of the following best describes what you are considering when you choose locations to perform gestures?

<table>
<thead>
<tr>
<th>Description</th>
<th>Dataset-1</th>
<th>Dataset-2</th>
</tr>
</thead>
<tbody>
<tr>
<td>I try to find locations where special <strong>objects</strong> are.</td>
<td>72.7%</td>
<td>59.6%</td>
</tr>
<tr>
<td>I try to find locations where some special <strong>shapes</strong> are.</td>
<td>24.2%</td>
<td>21.9%</td>
</tr>
<tr>
<td>I try to find locations where <strong>colors</strong> are different from their surroundings.</td>
<td>0%</td>
<td>8.7%</td>
</tr>
<tr>
<td>I randomly choose a location to draw without thinking about the background picture.</td>
<td>3.0%</td>
<td>10.1%</td>
</tr>
</tbody>
</table>
Part 1: User-choice Patterns

Gesture Locations

Which of the following best describes what you are considering when you choose locations to perform gestures?

<table>
<thead>
<tr>
<th>Description</th>
<th>0%</th>
<th>8.7%</th>
</tr>
</thead>
<tbody>
<tr>
<td>I try to find locations where colors are different from their surroundings.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>I randomly choose a location to draw without thinking about the background picture.</td>
<td>3.0%</td>
<td>10.1%</td>
</tr>
</tbody>
</table>

Most users tend to draw passwords on Points-of-Interest (Pols) in the background picture.
Part 1: User-choice Patterns

Gesture Locations (Picture of People)

• Dataset-1
  • 22 subjects uploaded 27 pictures of people
  • 31 passwords (93 gestures)

<table>
<thead>
<tr>
<th>Attributes</th>
<th># Gesture</th>
<th># Password</th>
<th># Subject</th>
</tr>
</thead>
<tbody>
<tr>
<td>Eye</td>
<td>36 (38.7%)</td>
<td>20 (64.5%)</td>
<td>19 (86.3%)</td>
</tr>
<tr>
<td>Nose</td>
<td>21 (22.5%)</td>
<td>13 (48.1%)</td>
<td>10 (45.4%)</td>
</tr>
<tr>
<td>Hand/Finger</td>
<td>6 (6.4%)</td>
<td>5 (18.5%)</td>
<td>4 (18.2%)</td>
</tr>
<tr>
<td>Jaw</td>
<td>5 (5.3%)</td>
<td>3 (11.1%)</td>
<td>3 (13.7%)</td>
</tr>
<tr>
<td>Face</td>
<td>4 (4.3%)</td>
<td>2 (7.4%)</td>
<td>2 (9.1%)</td>
</tr>
</tbody>
</table>
Part 1: User-choice Patterns

Gesture Locations (Civilization)

- Dataset-1
  - Two versions of *Starry Night* uploaded by two participants
Part 1: User-choice Patterns

Gesture Locations (Civilization)

- Dataset-1
  - Two versions of *Starry Night* uploaded by two participants

Gesture 1: Tap a star
Gesture 2: Tap a star
Gesture 3: Tap a star

(a) (b)
Part 1: User-choice Patterns

Gesture Locations (Civilization)

- Dataset-1
  - Two versions of *Starry Night* uploaded by two participants

Users have the tendencies to choose Pols with the same attributes to draw on.
Part 1: User-choice Patterns

Windows PGA Advertisements

Asia
Part 1: User-choice Patterns

Windows PGA Advertisements

Asia

Circle an eye

Circle an eye

Circle an ear
Part 1: User-choice Patterns
Windows PGA Advertisements

Asia

Circle an eye
Circle an eye
Circle an ear

South America
Part 1: User-choice Patterns

Windows PGA Advertisements

Asia

Circle an eye

Circle an eye

Circle an ear

South America

Line an arm

Circle a head

Line an arm
Part 1: User-choice Patterns

Windows PGA Advertisements

Asia
- Circle an eye
- Circle an eye
- Circle an ear

South America
- Line an arm
- Circle a head
- Line an arm

Europe
Part 1: User-choice Patterns

Windows PGA Advertisements

Asia
- Circle an eye
- Circle an eye
- Circle an ear

South America
- Line an arm
- Circle a head
- Line an arm

Europe
- Line an arm
- Circle a head
- Circle a head
Part 2: Attack Framework

• To generate dictionaries that have potential passwords
  • Picture-specific dictionary
  • Rank passwords with likelihood
  • Work on previously unseen pictures

• Our approach
  • Automatically learns user-choices patterns in the training pictures and corresponding passwords
  • Then applies these patterns to the target picture for dictionary generation
Part 2: Attack Framework

Selection Function

- Selection function
  - Models the **password creating process** that users go through
  - Takes two types of parameters
    - Gesture type, such as tap, circle, line
    - Pol attribute, such as face, eye, …
  - Generates a group of gestures
Part 2: Attack Framework

Selection Function (Examples)

\[ s : \{\text{tap, circle, line}\} \times \text{Pol Attributes}^* \]

\[ s(\text{circle}, \text{face}) \]
Circle a face in the picture

\[ s(\text{line}, \text{nose}, \text{nose}) \]
Line a nose to another nose in the picture

\[ s(\text{tap}, \text{nose}) \]
Tap a nose in the picture
Part 2: Attack Framework

Extract Selection Functions

Password

Points-of-Interest
Part 2: Attack Framework

Extract Selection Functions

circle Password

face Points-of-Interest
Part 2: Attack Framework

Extract Selection Functions

Password

Points-of-Interest

Function 1: $s(circle, face)$
Part 2: Attack Framework

Extract Selection Functions

Password

Points-of-Interest
Part 2: Attack Framework

Extract Selection Functions

Password

Points-of-Interest

Function 2: \( s(\text{line}, \text{nose}, \text{nose}) \)
Part 2: Attack Framework

Extract Selection Functions

Password

Points-of-Interest
Part 2: Attack Framework

Extract Selection Functions

Password

Points-of-Interest

Function 3: \( s(tap, nose) \)
Part 2: Attack Framework

Apply Selection Functions

Function 1: \( s(circle, face) \)
Output: 4 gestures
Part 2: Attack Framework

Apply Selection Functions

Function 1: $s(circle, face)$
Output: 4 gestures

Function 2: $s(line, nose, nose)$
Output: 12 gestures
Part 2: Attack Framework

Apply Selection Functions

Function 1:
\( s(\text{circle, face}) \)
Output: 4 gestures

Function 2:
\( s(\text{line, nose, nose}) \)
Output: 12 gestures

Function 3:
\( s(\text{tap, nose}) \)
Output: 4 gestures
Part 2: Attack Framework

Apply Selection Functions

Function 1: \( s(\text{circle, face}) \)
Output: 4 gestures

Function 2: \( s(\text{line, nose, nose}) \)
Output: 12 gestures

Function 3: \( s(\text{tap, nose}) \)
Output: 4 gestures

Number of potential passwords: \( 4 \times 12 \times 4 = 192 \)
Part 2: Attack Framework

Rank Selection Functions

1. **BestCover algorithm**
   - Derived from emts (Zhang et al., CCS’10)
   - Optimizes guessing order for passwords in the training dataset

2. **Unbiased algorithm**
   - Reduces the biased **Points-of-Interest distributions** in the training set
Part 3: Attack Results

Automatically Identify PoIs

- OpenCV as the computer vision framework
  - Object detection
    - Face, eye, nose, mouth, ear, body
  - Low-level feature detection
    - Circle
    - Color
  - Objectness measure: Alexe et al. (TPAMI’12)
    - Other standout regions
Part 3: Attack Results

Points-of-Interest Sets

\[ P^1_{A-40} \]
- Pols of Dataset-1
- Identified by OpenCV
- 40 Pols at most

\[ P^2_{A-40} \]
- Pols of Dataset-2
- Identified by OpenCV
- 40 Pols at most

\[ P^2_{L-15} \]
- Pols of Dataset-2
- Manually labeled
- 15 Pols at most
Part 3: Attack Results

Methodology

- Guessability on passwords of previously unseen pictures
- Dictionary size: $2^{19} = 524,288$
Part 3: Attack Results

Dataset-1 vs. Dataset-2

More cracked

Dataset-1: 48.8%
Dataset-2: 24.03%
Part 3: Attack Results

BestCover vs. Unbiased

~9400 training passwords

BestCover 24.03%  Unbiased 24.09%
Part 3: Attack Results

BestCover vs. Unbiased

60 training passwords

Unbiased 23.44%

BestCover 13.27%
Part 3: Attack Results

Labeled PoI set vs. OpenCV-Identified PoI set

More cracked

Labeled 29.42%

Identified 24.03%
Part 3: Attack Results

Simple Pictures (Unbiased algorithm)

- 39.0% cracked after 6,000 guesses
- 31.2% cracked after 30,000 guesses

13 passwords cracked
Part 3: Attack Results

Portraits (Unbiased algorithm)

29.0% in Total

9 Cracked
Part 3: Attack Results

Complex Picture (Unbiased algorithm)

![Diagram showing the number of passwords cracked against the number of password guesses (log-2 scale). The graph indicates a significant increase in the number of passwords cracked after a certain point, reaching 10.1%.]
Part 3: Attack Results
Online Attacks on Dataset-2

![Graph showing the number of passwords cracked vs. the number of password guesses. The graph compares different methods: BestCover $P^2_A_{-40}$, BestCover $P^2_L_{-15}$, Unbiased $P^2_A_{-40}$, and Unbiased $P^2_L_{-15}$.

Key data points:
- BestCover $P^2_A_{-40}$: Approximately 266/10K
- BestCover $P^2_L_{-15}$: Approximately 94/10K

The graph illustrates the efficiency of each method in cracking passwords, with the BestCover methods generally outperforming the Unbiased methods.
PGA Password Strength Meter

- https://honeyproject1.fulton.asu.edu/stmidx
- BestCover algorithm
- Generate dictionary and calculate strength in 20 seconds
Summary and Future Work

• We have presented an analysis of user-choice patterns in PGA passwords
• We have proposed an attack framework on PGA
• We have evaluated our approach on collected datasets

• We plan to improve online attack results by integrating shoulder-surfing and smudge attacks into our framework
Thank you!

Q & A