Network Game Design:

User Identification based on Game-Play Activity Patterns

Chun-Yang Chen, Academia Sinica
Li-Wen Hong, Academia Sinica
Motivation

- **Password-based User Identity**

  - **Vulnerability**
    - *Account hijacking (Identity Theft)*
      - Severity & prevalence
      - No general solution until the victim appears
    - *Account sharing*
      - Increase the difficulty of demographical studies of game
User Identity: Current Solutions

- Digital signature
  - Smart card

- Biometrical signature
  - fingerprint
  - voice
  - keystroke
Our Solution

A novel biometric:

**Game-Play Activity Patterns**
Outline

- Motivation
- Data Collection
- Player Activity Analysis
- Proposed Scheme
- Performance Evaluation
- Contribution & Future Work
Observation

More Regular

More Unpredictable

Motivation  Data Collection  Player Activity Analysis  Proposed Schemes …
Data Collection

- A MMORPG -- Angel’s Love
  - A commercial game in Taiwan
  - 40 thousands of players online

- The player activity logs we use
  - Trace period of 3 days
  - 287 randomly chosen accounts
  - Remove logs shorter than 200 minutes
Definitions

- **Active period**

  An active period of a game character is defined as a time interval \((t_1, t_2)\) in which the character *continuously* moves, with a tolerance of discontinuity up to 1 second.

- **Idle period**

  An idle period of a game character is defined as a time interval \((t_1, t_2)\) in which the character has *no movements*, where \(t_2 - t_1 \geq 1\) second.
## Data Summary

<table>
<thead>
<tr>
<th>Player # : 287</th>
<th>Data Length</th>
<th>Activity Rate</th>
<th>Active Period</th>
<th>Idle Period</th>
</tr>
</thead>
<tbody>
<tr>
<td>5%</td>
<td>7 hr</td>
<td>0.35 cycle/min</td>
<td>3 sec</td>
<td>7 sec</td>
</tr>
<tr>
<td>50%</td>
<td>51 hr</td>
<td>2.28 cycle/min</td>
<td>6 sec</td>
<td>18 sec</td>
</tr>
<tr>
<td>95%</td>
<td>67 hr</td>
<td>5.12 cycle/min</td>
<td>9 sec</td>
<td>181 sec</td>
</tr>
</tbody>
</table>

---

**Data Collection**  **Player Activity Analysis**  **Proposed Schemes**  **Performance Evaluation**  …
Distribution of Active / Idle Period

Data Collection  Player Activity Analysis  Proposed Schemes  Performance Evaluation  …
Idle time is much more diverse than active time

Data Collection  Player Activity Analysis  Proposed Schemes  Performance Evaluation  ...
Players’ active/idle patterns can be very different candidate features for user identification.
Why choosing idle time rather than active time

- Idle time distribution captures more variability
- Idle time process has smaller degree of autocorrelations

Data Collection  Player Activity Analysis  Proposed Schemes  Performance Evaluation  …
Idle Time Distribution of Random Players

Data Collection  Player Activity Analysis  Proposed Schemes  Performance Evaluation  …
KL Distances

- **ITD**: Idle time distribution
- **RET**: Relative Entropy Test
  - relative entropy between two ITDs
  - based on the KL distance
- **KL distance**: Kullback-Leibler distance
  - \[ D_{KL}(P \parallel Q) = \sum_i P(i) \log \frac{P(i)}{Q(i)} \]
  - \[ D_{SKL}(P \parallel Q) = D_{SKL}(Q \parallel P) = D_{KL}(P \parallel Q) + D_{KL}(Q \parallel P) \]
KL distances of Players

Player Activity Analysis  Proposed Schemes  Performance Evaluation  …
Identification Scheme: Consistency Test

- Perform consistency test
  - KLD: distribution of KL distance
  - 2 KLDs for each player
  - KLDs are tested by two-sided Wilcoxon test
Identification Scheme:
Discriminability Test

- Perform discriminability test
  - $KLD_{i,j}$: distribution of KL distances between $n_i$ ITDs of player $i$ & $n_j$ ITDs of player $j$
  - KLDs are tested by one-sided Wilcoxon test
Factor Consideration

Consideration: effect of the detection time & the history size

- $T_{rec}$: how long the player history kept in database (in minutes)

- $T_{obs}$: the detection time once the player log in (in minutes)
Evaluation Result

Proposed Schemes  Performance Evaluation  Contribution & Future Work
Performance Evaluation

- **Effect of** $T_{\text{rec}}$
  - mean of activity cycles is one minute $\Rightarrow T_{\text{rec}} = \text{idle times}$
  - assuming one million user accounts, $T_{\text{rec}} = 200$ minutes, each idle time uses 4 bytes $\Rightarrow$ storage space = 800 MB

- **Effect of** $T_{\text{obs}}$
  - assuming 10,000 players are online, $T_{\text{obs}} = 20$ minutes
    $\Rightarrow$ main memory = 0.8 MB
Contribution & Future Work

**Contribution:**

Propose the **RET scheme** for user identification from the aspect of **idle time**.

- With a 20-minute detection time period given a 200-minute history size ➔ achieve **higher than 90% accuracy**.

**Future Plan**

- Cut down the detection time
- Utilizing more aspects of game-play activities.
- Analyzing from the way users control the character.
Thank you!

Chun-Yang Chen
Li-Wen Hong