SoK: Keylogging Side Channels

Sean Brandenburg

Outline

- Goals
- Variation
- Classifications of Side Channels
  - Spatial
  - Temporal
- Attacks
- Defenses
- Future Works
Has a Key been Pressed?

Which Key?
Has a Key Been Pressed?

- Binary Classification
- How?
  - Audio
  - Visual
  - Other Sensor Data
- Precision
Which Key?

- More complicated problem
- Pattern Recognition and Trends
- How?
  - Time attacks
  - Spatial attacks
  - More!
Outline

- Goals
- Variation
- Classifications of Side Channels
  - Spatial
  - Temporal
- Attacks
- Defenses
- Future Works
Keyboard Variation

- **Logical Layout**
  - QWERTY vs. DVORAK

- **Physical Layout**
  - Shape and arrangement of keys

- **Relative Key Dimensions**
  - Letter Key 1 U
  - Tab 1.5 U
  - Caps 1.75 U

- **Switches**
Human Variation

- Typing Style
  - Pecking
  - Touch Typing
- Typing Proficiency
  - Speed
  - Parallelizing

Image: http://www.godine.com/2016/01/12/superior-persons-tuesday-monoplegia/
Outline

- Goals
- Variation
- Classifications of Side Channels
  - Spatial
  - Temporal
- Attacks
- Defenses
- Future Works
Spatial Side Channels

- **1st Order**
- **2nd Order**
Information Gains

- 672 (9.39 Bits) 3 Letter Words
- 4 (2 Bits) Words in Error range
  - con, com, von, vpn
- 9.39 - 2.0 = 7.39 Bits of Info Gained
- Relative gain = 7.39/9.39 = .79
Temporal Side Channels

- 2 Components of Keystroke
  - $T_p$ Time Pressed
  - $T_R$ Time Released
- Use $T_p$ and $T_R$ to guess which key
- $T_p - T_R = $ Duration
  - May differ between keys
Latency

Press-Press (PP): $\tau_{i}^{PP} = t_{i}^{P} - t_{i-1}^{P}$

Press-Release (PR): $\tau_{i}^{PR} = t_{i}^{R} - t_{i-1}^{P}$

Release-Press (RP): $\tau_{i}^{RP} = t_{i}^{P} - t_{i-1}^{R}$

Release-Release (RR): $\tau_{i}^{RR} = t_{i}^{R} - t_{i-1}^{R}$

Key-press latency

User A

$\mu_\delta = -0.088\delta + 5.280$
$\rho = -0.725$

User B

$\mu_\delta = -0.066\delta + 5.271$
$\rho = -0.653$

Inter-key distance
Practice!
Common Combos are Quick

Analyze latency between two character pairs (Bigrams) to gain info on what may have been typed
Users

- Highly user dependent side channel
- Slower typers leak more information
- Most effective on touch typists
Outline

- Goals
- Variation
- Classifications of Side Channels
  - Spatial
  - Temporal
- Attacks
- Defenses
- Future Works
<table>
<thead>
<tr>
<th>Modality</th>
<th>Attack Vulnerability</th>
<th>Proximity</th>
<th>Channel Type</th>
<th>Typing Speed</th>
<th>Requires Training</th>
<th>Performance</th>
<th>Ref.</th>
</tr>
</thead>
<tbody>
<tr>
<td>User</td>
<td>EEG</td>
<td>Headset</td>
<td>S1</td>
<td>-</td>
<td>-</td>
<td>63% key ACC</td>
<td>[48]</td>
</tr>
<tr>
<td></td>
<td>Motion Hand location over keyboard</td>
<td>Smartwatch</td>
<td>S1/T</td>
<td>-</td>
<td>-</td>
<td>57% TPR</td>
<td>[49]</td>
</tr>
<tr>
<td></td>
<td>Motion Hand movement over keyboard</td>
<td>Smartwatch</td>
<td>S1/S2</td>
<td>-</td>
<td>-</td>
<td>80% PIN ACC</td>
<td>[50]</td>
</tr>
<tr>
<td></td>
<td>Motion Hand movement+key acoustics</td>
<td>Smartwatch</td>
<td>S1/S2</td>
<td>-</td>
<td>-</td>
<td>55% word ACC5</td>
<td>[51]</td>
</tr>
<tr>
<td></td>
<td>Motion Hand movement+key acoustics</td>
<td>Smartwatch</td>
<td>S1/S2</td>
<td>-</td>
<td>-</td>
<td>51% word ACC10</td>
<td>[52]</td>
</tr>
<tr>
<td></td>
<td>Video Line of sight to keyboard</td>
<td>&lt; 1 m</td>
<td>S1</td>
<td>-</td>
<td>-</td>
<td>46% word ACC</td>
<td>[40]</td>
</tr>
<tr>
<td></td>
<td>WiFi WiFi CSI distortion patterns</td>
<td>4 m</td>
<td>S1</td>
<td>-</td>
<td>-</td>
<td>98% TPR</td>
<td>[5]</td>
</tr>
<tr>
<td></td>
<td>WiFi WiFi multipath localization</td>
<td>5 m</td>
<td>S1</td>
<td>-</td>
<td>-</td>
<td>92% key ACC5</td>
<td>[53]</td>
</tr>
<tr>
<td>Keyboard</td>
<td>Acoustic Keyboard acoustics</td>
<td>1 m</td>
<td>S1</td>
<td>-</td>
<td>-</td>
<td>79% key ACC</td>
<td>[54]</td>
</tr>
<tr>
<td></td>
<td>Acoustic Keyboard acoustics</td>
<td>-</td>
<td>S1</td>
<td>-</td>
<td>-</td>
<td>64% key ACC</td>
<td>[55]</td>
</tr>
<tr>
<td></td>
<td>Acoustic Keyboard acoustic differences</td>
<td>-</td>
<td>S2</td>
<td>-</td>
<td>-</td>
<td>73% word ACC5</td>
<td>[37]</td>
</tr>
<tr>
<td></td>
<td>Acoustic Bootstrapped keyboard acoustics</td>
<td>-</td>
<td>S1/S2</td>
<td>-</td>
<td>-</td>
<td>90% word ACC</td>
<td>[56]</td>
</tr>
<tr>
<td></td>
<td>Acoustic TDoA localization (3 mics)</td>
<td>&lt; 1 m</td>
<td>S1</td>
<td>-</td>
<td>-</td>
<td>72% key ACC</td>
<td>[39]</td>
</tr>
<tr>
<td></td>
<td>Acoustic TDoA localization (2 mics)</td>
<td>&lt; 1 m</td>
<td>S1/S2</td>
<td>-</td>
<td>-</td>
<td>94% key ACC</td>
<td>[57]</td>
</tr>
<tr>
<td></td>
<td>Acoustic Keyboard acoustics through VoIP</td>
<td>Remote</td>
<td>S1</td>
<td>-</td>
<td>-</td>
<td>83% key ACC</td>
<td>[58]</td>
</tr>
<tr>
<td></td>
<td>Acoustic Keyboard acoustics through VoIP</td>
<td>Remote</td>
<td>S1</td>
<td>-</td>
<td>-</td>
<td>74% key ACC</td>
<td>[9]</td>
</tr>
<tr>
<td></td>
<td>EM Cap PS/2 wire crosstalk</td>
<td>15 m</td>
<td>S1</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>[42]</td>
</tr>
<tr>
<td></td>
<td>EM Rad PS/2 signal radiation</td>
<td>15 m</td>
<td>S1</td>
<td>-</td>
<td>-</td>
<td>95% key ACC</td>
<td>[6]</td>
</tr>
<tr>
<td></td>
<td>EM Rad Matrix scan delay position</td>
<td>5 m</td>
<td>S1</td>
<td>-</td>
<td>-</td>
<td>95% key ACC5</td>
<td>[6]</td>
</tr>
<tr>
<td></td>
<td>Seismic Vibration sensed by acoustic laser</td>
<td>30 m</td>
<td>S2</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>[42]</td>
</tr>
<tr>
<td></td>
<td>Seismic Vibration sensed by smartphone</td>
<td>50 mm</td>
<td>S2</td>
<td>-</td>
<td>-</td>
<td>43% word ACC10</td>
<td>[59]</td>
</tr>
<tr>
<td>procs</td>
<td>procs procs stats (ESP and EIP)</td>
<td>Kernel</td>
<td>T</td>
<td>-</td>
<td>-</td>
<td>100% TPR</td>
<td>[60]</td>
</tr>
<tr>
<td>procs</td>
<td>procs procs schedstat</td>
<td>Kernel</td>
<td>T</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>[61]</td>
</tr>
<tr>
<td>Host</td>
<td>CPU Shared event loop time differences</td>
<td>Browser</td>
<td>T</td>
<td>-</td>
<td>-</td>
<td>98% TPR</td>
<td>[62]</td>
</tr>
<tr>
<td></td>
<td>CPU Instruction throughput differences</td>
<td>Browser</td>
<td>T</td>
<td>-</td>
<td>-</td>
<td>79% word ACC</td>
<td>[29]</td>
</tr>
<tr>
<td></td>
<td>CPU rdtsc differences</td>
<td>Core</td>
<td>T</td>
<td>-</td>
<td>-</td>
<td>100% TPR</td>
<td>[10]</td>
</tr>
<tr>
<td></td>
<td>CPU XTI event duration</td>
<td>Core</td>
<td>S2</td>
<td>-</td>
<td>-</td>
<td>2.5 bits/key</td>
<td>[63]</td>
</tr>
<tr>
<td></td>
<td>CPU Keyboard interrupt duration</td>
<td>Core</td>
<td>T</td>
<td>-</td>
<td>-</td>
<td>1 bit/key</td>
<td>[63]</td>
</tr>
<tr>
<td></td>
<td>Memory Prime+Trigger+Probe</td>
<td>Cache</td>
<td>T</td>
<td>-</td>
<td>-</td>
<td>95% TPR</td>
<td>[64]</td>
</tr>
<tr>
<td></td>
<td>Memory Cache-hit ratio pattern</td>
<td>Cache</td>
<td>S1</td>
<td>-</td>
<td>-</td>
<td>3.3 bits/key</td>
<td>[7]</td>
</tr>
<tr>
<td></td>
<td>Memory clflush latency</td>
<td>Cache</td>
<td>T</td>
<td>-</td>
<td>-</td>
<td>92% TPR</td>
<td>[65]</td>
</tr>
<tr>
<td></td>
<td>Memory DRAM row buffer latency</td>
<td>DRAM</td>
<td>T</td>
<td>-</td>
<td>-</td>
<td>100% TPR</td>
<td>[66]</td>
</tr>
<tr>
<td></td>
<td>Memory Multi-Prime+Probe</td>
<td>Cache</td>
<td>T</td>
<td>-</td>
<td>-</td>
<td>92% TPR</td>
<td>[10]</td>
</tr>
<tr>
<td></td>
<td>EM Cap USB hub crosstalk</td>
<td>USB hub</td>
<td>S1</td>
<td>-</td>
<td>-</td>
<td>97% key ACC</td>
<td>[67]</td>
</tr>
<tr>
<td></td>
<td>HTTP HTTP response size</td>
<td>Remote</td>
<td>S1</td>
<td>-</td>
<td>-</td>
<td>3.6 bits/key</td>
<td>[68]</td>
</tr>
<tr>
<td></td>
<td>SSH Packet timing (interactive shell)</td>
<td>Remote</td>
<td>T</td>
<td>-</td>
<td>-</td>
<td>1 bit/key</td>
<td>[8]</td>
</tr>
<tr>
<td></td>
<td>VoIP RTP event packet timing</td>
<td>Remote</td>
<td>T</td>
<td>-</td>
<td>-</td>
<td>1.7 bits/key</td>
<td>[69]</td>
</tr>
</tbody>
</table>
Attacks on the User

- Monitor Brain Activity with EEG
  - 63% key identification accuracy
- Compromised Smart Accessories
  - Mostly Watches: Accelerometer
  - Pin Entry: 80% Accuracy
  - Typing: 50% Word Accuracy
Attacks on the User

- Video
  - Hard for humans to process quickly
  - Computer Vision

- Wifi
  - Place an SDR and multiple antenna nearby
  - Detect changes in signal delay
Attacks on the **Keyboard**

- **Acoustic Side Channels**
  - A majority of side channel attacks
  - Create a key profile
    - Duration
    - Drum Effect
  - Triangulation with multiple mics
    - High Sampling Rate
    - Timing differences
Attacks on the Host

- Process Footprint
  - Monitor procfs on linux for stack pointers values to known locations in memory

- CPU Load
  - Monitor for spikes in CPU Load due to interrupts
  - Temporal Attack
Attacks on the **Network**

- **Payload**
  - Server Response
  - Real time auto-complete
- **Timing**
  - Victim types and creates bursts of network traffic
  - Use Temporal Attacks to gain info
Outline

- Goals
- Variation
- Classifications of Side Channels
  - Spatial
  - Temporal
- Attacks
- Defenses
- Future Works
Defenses

- Impediment: Restrict sensor data
- Obfuscation: Decrease signal to noise ratio
- Concealment: Add random noise

---

**TABLE III**

**Keylogging Side Channel Defenses.** DET = Keystroke Detection, ID = Key Identification.

<table>
<thead>
<tr>
<th>Modality</th>
<th>Defense Description</th>
<th>Method</th>
<th>Target</th>
<th>Channels Protected</th>
<th>Noticable to User?</th>
<th>Ref.</th>
</tr>
</thead>
<tbody>
<tr>
<td>User</td>
<td>Induce covert responses to irrelevant stimuli</td>
<td>Obfuscate</td>
<td>ID</td>
<td>✓</td>
<td>✓</td>
<td>[71], [81]</td>
</tr>
<tr>
<td>EEG</td>
<td>Filter keystroke-identifying features</td>
<td>Impede</td>
<td>DET</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Motion</td>
<td>Limit sensor permissions during typing</td>
<td>Impede</td>
<td>DET</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Acoustic</td>
<td>Reduce keyboard acoustic emissions</td>
<td>Impede</td>
<td>DET</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Acoustic*</td>
<td>Keys produce homomorphic sounds</td>
<td>Obfuscate</td>
<td>ID</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Acoustic*</td>
<td>Emit synthetic keyboard sounds</td>
<td>Conceal</td>
<td>DET</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>EM Rad/Cap.</td>
<td>Filter/shield EM emissions</td>
<td>Impede</td>
<td>DET</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>EM Rad.</td>
<td>Randomly delay matrix scan routine</td>
<td>Obfuscate</td>
<td>DET</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>EM Rad.</td>
<td>Randomize matrix scan pattern</td>
<td>Obfuscate</td>
<td>ID</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Host</td>
<td>Generate spurious key press/release events</td>
<td>Conceal</td>
<td>DET</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>CPU/Memory</td>
<td>Decrease timer resolution</td>
<td>Obfuscate</td>
<td>DET</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>HTTP</td>
<td>Obfuscate packet size through padding</td>
<td>Obfuscate</td>
<td>ID</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>SSH/TEL</td>
<td>Randomly delay key press/release events</td>
<td>Obfuscate</td>
<td>ID</td>
<td>✓</td>
<td>✓</td>
<td>Maybe</td>
</tr>
</tbody>
</table>

*Not including acoustic TDoA localization attacks.
Outline

- Goals
- Variation
- Classifications of Side Channels
  - Spatial
  - Temporal
- Attacks
- Defenses
- Future Works
Future Work

- Other Mediums
- Continue to tune temporal attacks to be effective on a larger range of users
- Search for defenses that stop attacks without affecting user experience