Attacks at the Sensor:
Standards Development on Presentation Attacks and Liveness Detection

Elaine Newton, PhD
Deputy Standards Liaison
Information Technology Laboratory
NIST
All of these scenarios rely on secure, trustworthy identity credentials. How can we reliably reject spoof attempts, while giving access to legitimate users?
Biometric Security Issues

Figure by Nalini Ratha, IBM, 2001
We’re not in Kansas anymore…

• Increasing use of online and mobile apps and need for more complex & secure ID management
  – Exemplified by the National Strategy for Trusted Identities in cyberspace, released April 2011
• Recognized need by groups of potential users:
  – Financial Services Technology Consortium
  – The Drug Enforcement Administration (DEA)
Authentication Use Case Comparison

For law enforcement, immigration, etc.
- Enrollment and subsequent recognition attempts
  - Highly controlled
  - Supervised / Attended
- Successful recognition
  - Answers the question, “Has this person been previously encountered?”
  - Is a unique pattern

For online transactions, e.g. banking, health, etc.
- Enrollment
  - Less controlled
  - Probably not in person
- Subsequent recognition attempts
  - Unattended
- Successful recognition
  - Answers the question, “How confident am I that this is the actual claimant?”
  - Is a tamper-proof rendering of a distinctive pattern
Existing International Standards or Best Practices Documents on Liveness Detection or Countering “Fake Biometrics”
Examples of Points of Attack in a Biometric System

From the 1st Committee Draft of IS Project 30107-1, inspired by figure by Nalini Ratha from 2001 and Standing Document 11 of ISO/IEC JTC1 SC37.
High Confidence in factors available to consumers for authentication (and access) over open networks

Presentation Attack Detection Standards Project

Threats/Attacks

Countermeasures

Data/Data Formats (for Metrics)

Accreditation, QPL, Validation, and/or Certification Program (e.g. NVLAP in the US)

Measurement/Evaluation
• Part 1 establishes terms and definitions that are useful in the specification, characterization and evaluation of presentation attack detection methods.

• Part 2 establishes a common data format for conveying the type of approach used and the assessment of presentation attack in data formats.

• Part 3 establishes principles and methods for performance assessment of presentation attack detection algorithms or mechanisms, and it includes an informative annex with a classification of known attacks types.
Outside the scope (of all parts) are
- standardization of specific PAD detection methods;
- detailed information about countermeasures (i.e. anti-spoofing techniques), algorithms, or sensors; and
- overall system-level security or vulnerability assessment.

The attacks to be considered in this standard will take place at the sensor during the presentation and collection of the biometric characteristics. Any other attacks are considered outside the scope of this standard.
General biometric framework with Presentation Attack Detection

From the 1st Committee Draft of IS Project 30107-1
Where can PAD take place?

The sub-system which detects attack presentations may be located:
- following the data capture subsystem
- within the data capture subsystem,
- following the signal processing sub-system, and/or
- after the comparison or decision subsystems (not shown) or at several points in the system.

The different components could be in different locations (client versus server or front end versus back end).

From the 1st Committee Draft of IS Project 30107-1
## Examples of Presentation Attacks

<table>
<thead>
<tr>
<th>Category</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gummy Finger</td>
<td>Video of face</td>
</tr>
<tr>
<td>Glue on Finger</td>
<td>Sunglasses, artificial/patterned contact lens</td>
</tr>
<tr>
<td>Cadaver Part</td>
<td>Severed finger/hand</td>
</tr>
<tr>
<td>Mutilation</td>
<td>Surgical switching of fingerprints between hands and/or toes</td>
</tr>
<tr>
<td>Facial Expression</td>
<td>Extreme, tip or side of finger</td>
</tr>
<tr>
<td>Unconscious</td>
<td>Under duress</td>
</tr>
<tr>
<td>Zero Effort Impostor Attempt</td>
<td></td>
</tr>
</tbody>
</table>

*From the 1st Committee Draft of IS Project 30107-1*
Terms (1 of 3)

- **Presentation attack**
  presentation of an artefact or human characteristic to the biometric capture subsystem in a fashion that could interfere with the intended policy of the biometric system.

- **Presentation attack instrument (PAI)**
  biometric trait or object used in a presentation attack.
  
  NOTE  The set of PAI includes artefacts but would also include lifeless biometric characteristics (i.e. stemming from dead bodies) or altered biometric characteristics (e.g. altered fingerprints) that are used in an attack.

- **Presentation attack detection (PAD)**
  automated determination of a presentation attack.

From the 1st Committee Draft of IS Project 30107-1
• **Liveness**
the quality or state of being alive, made evident by anatomical characteristics (e.g. skin or blood absorption of illumination), involuntary reactions or physiological functions (e.g. iris reaction to light, heart activity – pulse), or voluntary reactions or subject behaviors (e.g. squeezing together fingers in hand geometry or a biometric presentation in response to a directive cue).

• **Liveness detection**
detection of anatomical characteristics or involuntary or voluntary reactions, in order to determine if a biometric sample is being captured from a living subject present at the point of capture.

Liveness detection methods are defined to be a sub-set of presentation attack detection (PAD) methods.
Terms (3 of 3)

- **Spoof**
  to subvert a system by presentation of an artefact

- **Artefact**
  artificial object or representation presenting a copy of biometric characteristics or synthetic biometric patterns
Types of Biometric Presentation Attacks

(General Noun)

Presentation attack instrument

(Adjectives describing categories)

Artificial
Human
Other Natural

(Qualifying adjective.
Under each sub-category, these types are listed based on a general sense of increasing difficulty to detect (from left to right).)

Complete
Partial

Lifeless
Altered
Non-Conformant
Coerced
Conformant
## Types of Detection

<table>
<thead>
<tr>
<th>Through a biometric System</th>
<th>Artefact Detection</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Liveness Detection</td>
</tr>
<tr>
<td></td>
<td>Alteration Detection</td>
</tr>
<tr>
<td></td>
<td>Non-conformance Detection</td>
</tr>
<tr>
<td></td>
<td>Coercion Detection</td>
</tr>
<tr>
<td></td>
<td>Obscuration Detection</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Through system security policies</th>
<th>Failed attempt detection counter</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Geographic</td>
</tr>
<tr>
<td></td>
<td>Temporal</td>
</tr>
<tr>
<td></td>
<td>Video Surveillance</td>
</tr>
</tbody>
</table>
Examples of Data Types for Detecting Presentation Attacks

These are all options in the draft standard:

- The local PAD decision (pass/fail)
- A score between 0 and 100 provided by the PAD mechanism, with lower scores being indicative of spoofed samples
- Identifier for technique specific data (to identify a vendor and algorithm);
- Level of supervision / surveillance during capture (qualitative categories)
Data Format Topics

• WD of Part 2 contains many options for PAD data that could placed in a 19794 header in TLV format.

• A Special Group considered encoding/including extended data in 2013.
  – Questions that still need to be addressed:
    • a definition of extended data,
    • use cases, and
    • how to encode the extended data (does it belong in Part 2 or should it be handled another way such as a new data format type).
Where We Are At Today wrt Testing

- **Terms & Concepts**: Focus on the device; ignore intent of the attacker.

- **Testing Metrics and Reporting**:
  - Different approaches around the world.
    - Can we strike a balance between encouraging progress versus a hard line or exhaustive/expensive method?
  - Flexibility in the standard for different approaches and agility to deal with new threats, and use testing reporting to deal with differences.
How to Participate in the Development of 30107

• In the US, interested parties participate through INCITS M1
  – http://standards.incits.org/a/public/group/m1

• In other countries, interested parties participate in their country’s mirror committee, called a Technical Advisory Group (TAG), to ISO/IEC JTC1 SC37
## Types of Detection

<table>
<thead>
<tr>
<th>Through a biometric System</th>
<th>Through system security policies</th>
</tr>
</thead>
<tbody>
<tr>
<td>Artefact Detection</td>
<td>Failed attempt detection counter</td>
</tr>
<tr>
<td>Liveness Detection</td>
<td>Geographic</td>
</tr>
<tr>
<td>Alteration Detection</td>
<td>Temporal</td>
</tr>
<tr>
<td>Non-conformance Detection</td>
<td>Video Surveillance</td>
</tr>
<tr>
<td>Coercion Detection</td>
<td></td>
</tr>
<tr>
<td>Obscuration Detection</td>
<td></td>
</tr>
</tbody>
</table>
Future Guidelines Need to Address

• How can the vulnerability at the sensor be mitigated at different levels of risk (i.e. different security levels)?
  – Is there a way to quantitatively or qualitatively rank options?

• What are equivalent means for mitigating this risk through the biometric system versus system security policies or some mix of the two?
  – Example: can liveness detection methods based on some material testing be considered equivalent to use of second factor such as a password of strength $1:10^x$?
Thank you

Elaine Newton, PhD
Lead Editor for ISO/IEC JTC1 SC 37 Project 30107-1,
Co-Editor for Parts 2 and 3
elaine.newton@nist.gov
1-301-975-2532