Satellite Workshop II – Template Protection Testing

Friday, March 5, 2010, 9am – 11am
NIST, Gaithersburg, MD
Admin Building 101, Employee Lounge
Point of Contact: Elaine.Newton@nist.gov

Agenda

1. Welcome and Introductions – Elaine Newton, NIST
2. Template Protection Background – "Security and Privacy in Biometric Systems - The purpose of Biometric Encryption," Tom Kevenaar, priv-ID
3. NIST's Cryptographic Module Validation Program, Randall Easter (CMVP Director)
4. ISO 24745 Biometric Template Protection – Christoph Busch, Fraunhofer
5. Twiki Set-up – Ross Micheals, NIST
6. Panel on Testing of Template Protection Schemes

Moderator: Elaine Newton, NIST
Panelists:
Terry Boult, UCCS
Christoph Busch, Fraunhofer
Jean-Christophe Fondeur, Sagem
Tom Kevenaar, priv-ID
Nalini Ratha, IBM
Xuebing Zhou, Fraunhofer
Others TBD

- Presentations on "NIST Biometric Authentication Testing, Data, and Principles" and "Measures and Trade-offs of Biometric Template Protection Schemes" – Elaine Newton, NIST
- Reaction from Panelists
- Discussion on How to Work with the TURBINE Project
- Attendee Questions and Comments

7. Wrap-up – Newton and Micheals from NIST

Background for Panel Discussion

NIST Special Publication 800-63 contains guidelines for remote authentication over open networks. Per an OMB memo (M-04-04), NIST SP 800-63 provides the technical requirements for meeting four levels of authentication assurance for remote access to government systems. These four qualitative levels range from little or no confidence in the claimant's identity (assurance level 1) to very high confidence in the asserted identity's validity (assurance level 4). Since its publication, this document has been referenced and adopted widely by non-USG organizations, and a version is currently being standardized by ISO and ITU-T.

NIST SP 800-63 was first published in 2006 and currently being revised. Both the original and the draft revision make nearly identical statements regarding the use of biometrics for remote authentication (bold type added here for emphasis):

"**Biometrics do not constitute secrets suitable for use in the conventional remote authentication protocols addressed in this document.** In the local authentication case, where the Claimant is observed and uses a capture device controlled by the Verifier, authentication does not require that biometrics be kept secret."

"Biometrics are unique personal attributes that can be used to verify the identity of a person who is physically present at the point of verification. They include facial features, fingerprints, DNA, iris and retina scans, voiceprints and many other characteristics. This publication recommends that biometrics be used in the registration process to later prevent a Subscriber who is in fact registered from repudiating the registration, to help identify those who commit registration fraud, and to unlock tokens. Biometric characteristics are not recommended for use directly as tokens in this document."

This document contains guidelines for identity management problems for each of the four levels of assurance, including "token and credential revocation and destruction."

Biometrics are not secret; nor are they easily destroyed and re-issued, unlike authentication factors such as passwords or smartcards with digital certificates. The ability to detect a live, unspoofed biometric sample at a sensor is a critical component of the integrity of biometrics for reliable authentication. Short of an attendee being present at the point of collection, liveness detection methods that can be independently evaluated could potentially be applied to aid the security of remote authentication over open networks. Depending upon the training and efficacy of personnel overseeing the authentication process, attended applications may also opt to employ liveness detection methods.

As the use of biometrics becomes more wide-spread, the inability to destroy the factor and re-issue a new digital identity could present heightened security weaknesses for organizations and compromise the ability of individuals to maintain confidentiality. The use of a common pool of developed biometrics, such as fingerprints, faces, and irises, could give rise to the ability to link data across domains. Imagine using the same password to log onto your personal email; your work email with sensitive, proprietary, and/or classified data; your online banking applications; your online travel agent; and all the websites you log into to make purchases or read news articles. Imagine not changing that password every 90 days, and imagine that every system administrator, professional to amateur hacker, curious intern, disgruntled former employee, etc. knew that this was the case. This would be analogous to wide-spread use of biometrics as an authentication factor without implementing measures to combat their weaknesses.

Methods of protecting biometric data to solve this problem – to renew, revoke, or cancel the credential – are an active area of research and development. Here they will be generally referred to as *template protection algorithms*. The goal of these algorithms is to produce templates for an individual that accurately match, while severing the link to the individual's.

<u>Current related NIST testing</u>

NIST has been conducting evaluations of fingerprint, face, iris, and speaker recognition algorithms to measure their accuracy. These evaluations largely report one-to-one matching error rates – false match and false non-match rates for algorithms submitted by companies and universities.   NIST has the largest sequestered databases for independent testing of these modalities.

For the case of remote authentication, authentication factors which are passed through a network would be encrypted to protect the secret information being communicated. NIST standards and guidelines published in FIPS 140-2 and SP 800-63 specify how this can be done. To support validation of products, NIST partners with the Communication Security Establishment Canada (CSEC) to run two programs: Cryptographic Algorithm Validation Program (CAVP) and the Cryptographic Module Validation Program (CMVP).

<u>Potential Measures of Template Protection for Remote Authentication</u>

The suitability of template protection algorithms for remote authentication could be comprised of three measurements:

1. Authentication Accuracy
    a. This is the traditional measure of error rates from one-to-one comparisons, producing false match and non-match rates (and true match and non-match rates) that can be plotted on Receiver Operating Characteristic (ROC) curve.

    b. Measure: True Match Rate (how often transformed enrollment of Alice matches transformed capture of Alice at another time)
        i. Must be measured at a specific point on the ROC curve, at a false match rate that can be supported by the size of the data set

2. De-Identification
    a. Face De-Identification was defined by Newton, Sweeney, and Malin[1] as

> "Let $\mathbf{H}$ and $\mathbf{H_d}$ be face sets, $\Gamma\in\mathbf{H}$, $\Gamma_d\in\mathbf{H_d}$, f:$\mathbf{H}\rightarrow\mathbf{H_d}$ be a function that attempts to conceal the identity of the subject of the original face image; and, $f(\Gamma) = \Gamma_d$ but $\Gamma\neq\Gamma_d$ (element-wise). f is termed **<u>face de-identification</u>** ("de-identification", "de-identification function"). Γ_d is a de-identified image."

[1] E.M. Newton, L. Sweeney, B. Malin. "Preserving Privacy by De-identifying Facial Images." in *IEEE Transactions on Knowledge and Data Engineering*, February 2005, 17(2).

They also define Effective De-Identification[2] as

"Let $\mathbf{H}$ be a person-specific face set; $\mathbf{H_d}$ be a face set; f:$\mathbf{H}\to\mathbf{H_d}$ be the transformation function used in face de-identification, such that $f(\Gamma)=\Gamma_d$ where $\Gamma\in\mathbf{H}$ and $\Gamma_d\in\mathbf{H_d}$; g be a face identification relation g:$\mathbf{H_d}\to\mathbf{H}$; and, C be a provable claim about f's ability to restrict face identification (or face recognition) by g. The function f provides **effective de-identification** with respect to C and f is said to be **effective**. If f_1 and f_2 are effective with respect to the same C, then f_1 and f_2 are considered **equally effective** with respect to C."

While Newton et al. developed de-identification methods to apply to original biometric data to enable privacy-preserving data sharing (as opposed to *a priori* transformation at the sensor), the definitions for de-identification and effective de-identification can be generalized and used here to describe the goal of template protection algorithms, as follows:

Let $\mathbf{H}$ be a person-specific biometric set; $\mathbf{H_d}$ be a biometric set; f:$\mathbf{H}\to\mathbf{H_d}$ be the transformation function used in biometric de-identification, such that $f(\Gamma)=\Gamma_d$ where $\Gamma\in\mathbf{H}$ and $\Gamma_d\in\mathbf{H_d}$; g be a biometric identification relation g:$\mathbf{H_d}\to\mathbf{H}$; and, C be a provable claim about f's ability to restrict biometric identification (or biometric recognition) by g. The function f provides **effective de-identification** with respect to C and f is said to be **effective**.

b. Measure: False Non-Match Rate
   i. How often original of Alice's biometric does not match transformed template of Alice
   ii. The FNMR Must be measured at a specific point on the ROC curve, at a false match rate that can be supported by the size of the data set

3. Security Strength
   a. Preimage resistance – It should be very difficult to recover the original template or image from a protected template.
   b. Measure: Entropy, bits.

[2] Ibid.

Figure 1 shows the three measures, each increasing as you move away from the origin. Algorithms that are able to perform further from the origin would be considered better than those closer to the origin.

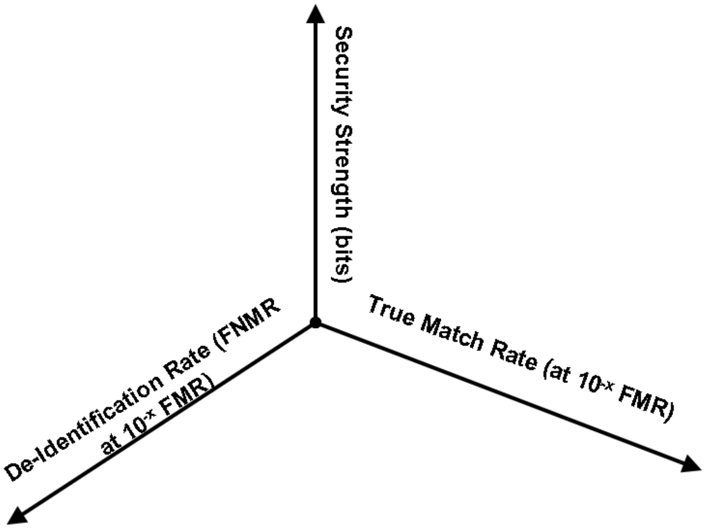
Security Strength (bits)
True Match Rate (at 10^{-x} FMR)
De-Identification Rate (FNMR at 10^{-x} FMR)

Figure 1: 3-dimensions of template protection algorithm performance