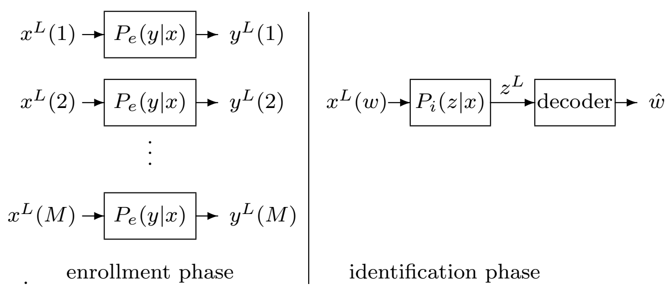
ISIT 2003, Yokohama, Japan, June 29 – July 4, 2003

On the Capacity of a Biometrical Identification System

Frans Willems, Ton Kalker, Jasper Goseling, and Jean-Paul Linnartz
Philips Research Laboratories, Eindhoven, The Netherlands
f.m.j.willems@tue.nl ton.kalker@ieee.org j.goseling@ieee.org j.p.linnartz@philips.com

Abstract — **We investigate fundamental properties of biometrical identification systems. We focus on the capacity, i.e. a measure for the number of individuals that can be reliably identified. It can be expressed using standard information-theoretical concepts.**

I. Model description, result



We assume that there are M individuals with indices $w \in \{1, 2, \cdots, M\}$. To each individual there corresponds a biometrical data sequence $x^L = (x_1, x_2, \cdots, x_L)$ with components x_l for $l = 1, L$ that assume a value from an alphabet $\mathcal{X}$. Sequence $x^L(w)$ is the sequence for individual w for $w \in \{1, 2, \cdots, M\}$. It is supposed to be *generated at random* according to

$$\Pr\{X^L(w) = x^L\} = \prod_{l=1,L} Q(x_l), \text{ for all } x^L \in \mathcal{X}^L, \quad (1)$$

hence each biometrical data sequence is produced by an i.i.d. source with distribution $\{Q(x) : x \in \mathcal{X}\}$.

In the **enrollment phase** all biometrical data sequences are observed via a memoryless enrollment channel $\{\mathcal{Y}, P_e(y|x), \mathcal{X}\}$. $\mathcal{Y}$ is the enrollment output-alphabet and

$$\Pr\{Y^L(w) = y^L | X^L(w) = x^L(w)\} = \prod_{l=1,L} P_e(y_l|x_l(w)) \quad (2)$$

for all $y^L = (y_1, y_2, \cdots, y_L) \in \mathcal{Y}^L$ and individuals w. The resulting enrollment output sequences $y^L(w)$ are all stored in a database.

In the **identification phase** the biometrical data sequence $x^L(w)$ of an unknown individual $w \in \{1, 2, \cdots, M\}$ is observed via a memoryless identification channel $\{\mathcal{Z}, P_i(z|x), \mathcal{X}\}$. $\mathcal{Z}$ is the identification output-alphabet and

$$\Pr\{Z^L = z^L | X^L(w) = x^L(w)\} = \prod_{l=1,L} P_i(z_l|x_l(w)) \quad (3)$$

for all $z^L = (z_1, z_2, \cdots, z_L) \in \mathcal{Z}^L$. The resulting identification output sequence z^L is used by a decoder that can access all enrollment output sequences $y^L(1), y^L(2), \cdots, y^L(M)$ in the database. This decoder produces an estimate of the index of the unknown individual, i.e.

$$\hat{w} = d\left(z^L, y^L(1), y^L(2), \cdots, y^L(M)\right). \quad (4)$$

We assume that the estimate $\hat{w} \in \{\varepsilon, 1, 2, \cdots, M\}$, thus an erasure ε is also a valid decoder output. The two relevant system parameters are the *maximal error probability*[1]

$$P_e^{\max} \triangleq \max_{w=1,M} \Pr\{\hat{W} \neq w | W = w\} \quad (5)$$

and the *rate*

$$R \triangleq \frac{1}{L} \log M. \quad (6)$$

The capacity of a biometrical system is C if for any $\delta > 0$ there exist, for all large enough L, decoders that achieve

$$\begin{aligned} R &\geq C - \delta, \\ P_e^{\max} &\leq \delta. \end{aligned} \quad (7)$$

Theorem 1 *The capacity C of a biometrical identification system is equal to $I(Y;Z)$ where $P(y,z) = \sum_{x \in \mathcal{X}} Q(x) P_e(y|x) P_i(z|x)$ for all $y \in \mathcal{Y}$ and $z \in \mathcal{Z}$.*

The achievability proof is based on typicality (see e.g. [1]). The converse is Fano-type.

II. An example

Let X be Gaussian, zero-mean, with variance P. Moreover let

$$Y = X + N_e, \qquad Z = Z + N_i, \quad (8)$$

with zero-mean Gaussian noise variables N_e and N_i having variances σ_e^2 and σ_i^2 respectively. Then

$$I(Y;Z) = \frac{1}{2} \log\left(1 + \frac{P}{\sigma_e^2 + \sigma_i^2 + \sigma_e^2 \sigma_i^2 / P}\right). \quad (9)$$

This demonstrates that the channel from Y to Z is **not** the cascade of the enrollment channel and the identification channel. A similar conclusion was obtained in [2] in which the emphasis is on detection.

III. Remark

We did not consider the probability that an individual, that did not undergo the enrollment procedure, is identified as one of the individuals that did enroll properly. For rates R smaller than $I(Y;Z)$ this probability can also be made smaller than any $\epsilon > 0$ by increasing L.

In order to achieve capacity we should increase the blocklength L. However in practise we are more interested in achieving a small error probability for a given number of individuals than to achieve capacity. Still the capacity is a fundamental limit that tells us what we can expect from a certain system.

References

[1] T.M. Cover and J.A. Thomas, *Elements of Information Theory.* Wiley, New York, 1991.

[2] J. Goseling, S. Baggen and A.H.M. Akkermans, "Optimal Verification of Partially Known Biometrics", *Proc. 24th Symp. Inform. Theory in the Benelux*, Veldhoven, 2003.

[1] The stochastic processes that play a role here are the generation of the biometrical data sequences and the transmission of these sequences over the enrollment and identification channel.