Mathematical Problems of Computer Science 39, 88–93, 2013.

Investigation of E-Capacity for Biometric Identification Protocol with Random Parameter[*]

Mariam E. Haroutunian and Lilit A. Ter-Vardanyan

Institute for Informatics and Automation Problems of NAS RA
e-mail: armar@ipia.sci.am, lilit@sci.am

Abstract

In recent years biometrics is widely used in different tasks in the field of security. In this paper we investigated the biometric identification system from an information-theoretical point of view. We investigate the exponentially high reliability criterion in biometric identification systems. The biometric identification system with random parameter is considered, which is more realistic for application. The lower and upper bounds of identification E-capacity of the model with random parameter for maximal and average error probabilities are constructed. When $E \to 0$ we derive the corresponding bounds of the capacity of the biometric identification system with random parameter, which coincide and hence, as a corollary we obtain the identification capacity for this model.

Keywords: Biometric identification system, identification capacity, E-capacity bounds, error exponents, channel with random parameter.

1. Introduction

In recent years biometrics is widely used in different tasks in the field of security. Biometrics is being used for physical access control, computer log-in, international border crossing and ID cards, e-passports, based on biological features of any person, such as fingerprints, an eye iris [2].

In this paper we consider the protocol of biometric identification. The identification system with random parameter is given in Fig.1. A typical protocol for identification consists of two steps: enrollment and identification. During the enrollment, the biometric data of M subjects are captured and analyzed, after that, for each individual, a record is added to a database. A perfect system would always recognize an individual and reject an impostor. To build an ideal channel is impossible because biometric data are gathered from individuals under environmental conditions and the channels are exposed to noise. Such a system is not perfectly secure, it leads to some errors, and it can be information-theoretical secure up to a certain level. In this paper we investigated the biometric identification system from an information-theoretical point of view. The enrollment-data in a database is a noisy version of the biometrical data corresponding to the individual.

[*] Research was supported by Armenian national grant 11-1b255.

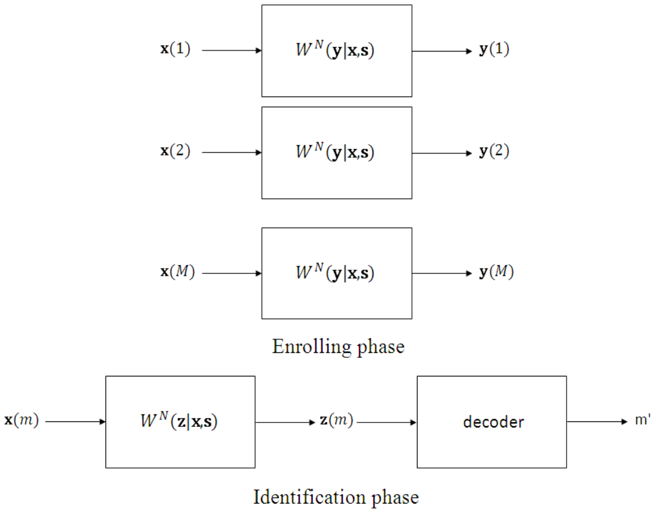

Fig. 1. Model of biometric identification system with random parameter

In the identification phase an unknown individual is observed again and another noisy version of the biometric data is compared to the enrollment data in the database. The system has to come up with an estimate of the individual.

Willems et al [1,3] investigated the fundamental properties of biomertic identification system. It has been shown that it is impossible to reliably identify more persons than capacity which is an inherent characteristic of any identification system. They derived the capacity C of this model.

In [4] the E-capacity new concept for biometrical identification system was introduced. We investigated the exponentially high reliability criterion in biometric identification systems. In other words, we introduced a new performance concept of biometric identification E-capacity, which takes into account a stronger requirement on identification fault events with extremely small probability (2^{-NE} instead of ε). In terms of practical applications an exponential decrease in error probability (namely, in unwanted identification faults) is more desirable.

In practice the individuals are observed in various places and at various times. That is why it is more interesting from the practical point of view to assume that the considered model channels depend on a random parameter of the biometric model. In this paper we investigate the biometric identification system with random parameter, which is more realistic for application.

The channel with random parameter with additional information on the encoder was first considered by Shannon [5] and studied by Gelfand S. I. and Pinsker M. S. [6]. They found the capacity of this channel for the average error probability $\overline{C}$ in a situation when the state sequence is known at the encoder. Ahlswede R. F. [7] showed that the capacity for the maximum error probability C is the same. The E-capacity for the channel with random parameter, E-capacity and capacity for the multiple-access channel with random parameter was investigated by Haroutunian E. A., Haroutunian M. E. [8,10].

The channel can be considered in four cases, when the state sequence is known or unknown at the encoder and decoder. Proceeding from the applications in the identification protocol, the situation, when the state sequence is unknown at the encoder and decoder, is possible.

In this paper the lower and upper bounds of the identification E-capacity of the model with random parameter for maximal and average error probabilities are constructed. When

$E \to 0$ we derive the corresponding bounds of the capacity of the biometric identification system with random parameter, which coincide and hence, as a corollary we obtain the identification capacity for this model.

2. Notations and Definitions

Let $\mathcal{X}, \mathcal{Y}, \mathcal{Z}, \mathcal{S}$ be finite sets and W be a family of discrete memoryless channels $W_s : \mathcal{X} \to \mathcal{Y}$, with an input alphabet $\mathcal{X}$ and an output alphabet $\mathcal{Y}$. The s is the channel state, varying independently at each moment of the channel action with the same known PD $Q(s)$ on S. There are M individuals and each individual has an index $m = \{1, 2, \cdots, M\}$. A biometric data sequence $\mathbf{x}(m) = \{x_1, x_2, \cdots, x_N\}$, where $x_n \in \mathcal{X}, n = \overline{1, N}$ corresponds to each individual m. All these sequences are suposed to be generated at random with a given probability distribution

$$P^N(x) = \prod_{n=1}^{N} P(x_n), \quad x \in \mathcal{X}^N.$$

Enrollment phase. Let us have the stationary and discrete memoryless channel $W(y|x, s)$ with random parameter. In this phase all biometric data sequences $\mathbf{x}(m)$ are observed via this channel. The state of the channel is changed by the following probability distribution $Q(s)$, it means

$$W^N(\mathbf{y}|\mathbf{x}, \mathbf{s}) = \prod_{n=1}^{N} W(y_n|x_n, s_n), \;\; Q^N(s) = \prod_{n=1}^{N} Q(s_n) \;\; \mathbf{x} \in \mathcal{X}^N, \;\; \mathbf{y} \in \mathcal{Y}^N, \;\; \mathbf{s} \in \mathcal{S}^N.$$

The resulting $\mathbf{y}(m)$ enrollment output sequences for all $m = \{1, 2, \cdots, M\}$ are stored in the database (define it as Y_{DB}).

Identification phase. In the identification phase the biometric data sequence of an unknown individual is observed via the same memoryless channel $W(z|x, s)$ with random parameter.

$$W^N(\mathbf{z}|\mathbf{x}, \mathbf{s}) = \prod_{n=1}^{N} W(y_n|x_n, s_n), \;\; \mathbf{z} \in \mathcal{Z}^N, \;\; \mathbf{x} \in \mathcal{X}^N, \;\; \mathbf{s} \in \mathcal{S}^N.$$

The resulting identification output sequence $\mathbf{z}$ is compared to the sequences $\mathbf{y}(m)$, $m = 1, 2, \cdots, M$, from the database and the identification function

$$g_N : \mathcal{Z}^N \to \{0, 1, 2, \cdots, M\}$$

produces the index of the unknown individual $m' = g_N(\mathbf{z})$, here 0 stands for the case, when the unknown individual has not been observed by the enrollment phase. If the state sequence is unknown at the encoder and decoder, let us denote

$$W^*(y|x) = \sum_{s \in S} Q(s) W(y|x, s), \quad W^*(z|x) = \sum_{s \in S} Q(s) W(z|x, s).$$

And

$$P^* = \{P^*(y) = \sum_{x} W^*(y|x) P(x), \;\; x \in X, \;\; y \in Y\},$$

$$W^*(z|y) = \frac{\sum_x W^*(y|x) W^*(z|x) P(x)}{P^*(y)}.$$

The channel $W^*:\mathcal{Y}\to\mathcal{Z}$ is memoryless:

$$W^N(\mathbf{z}|\mathbf{y})=\prod_{n=1}^{N}W(z_n|y_n),\quad \mathbf{z}\in\mathcal{Z}^N,\quad \mathbf{y}\in\mathcal{Y}^N.$$

Denote by R the following rate

$$R=\frac{1}{N}\log_2 M.$$

The error probability of the identification of the person m is

$$e(N,m)=W^N(\mathcal{Z}^N\backslash g_N^{-1}(m)|\mathbf{y}(m)),$$

where

$$g_N^{-1}(m)=\{\mathbf{z}:g_N(\mathbf{z})=m\}.$$

We consider the **maximal** and the **average error probabilities**

$$e(N)=\max_{m\in M}e(m),$$

$$\overline{e}(N)=\frac{1}{M}\sum_{m\in M}e(m).$$

The E-capacity function for the given $E>0$ is defined as

$$C(E,P^*,W^*)=\overline{\lim_{N\to\infty}}\frac{1}{N}\log M(E,P^*,W^*,N),$$

where

$$M(E,P^*,W^*,N)=\sup_{g_N}\{M:e_Q(N)\le\exp(-NE)\}.$$

We denote by $\overline{C}(E,P^*,W^*)$ the E-capacity for the average error probability.
We shall use the following PD in the formulation of results:

$$P=\{P(y),y\in\mathcal{Y}\},$$

$$V=\{V(z|y),z\in\mathcal{Z},y\in\mathcal{Y}\}.$$

For information-theoretic quantities, such as entropy $H_P(Y)$, mutual information $I_{P,V}(Z\wedge Y)$, the divergence $D(V||W|P)$ and for the notion of the type we refer to [9]– [15].

3. Formulation of the Result

To define the lower bound (*random coding bound*) of the identification of E-capacity for the channel with random parameter let us denote:

$$R_r(E,P^*,W^*)\triangleq$$

$$\triangleq\min_{P,V:D(P\circ V||P^*\circ W^*)\le E}\left|I_{P,V}(Z\wedge Y)+D(P\circ V||P^*\circ W^*)-E\right|^+.\tag{1}$$

For the formulation of the upper bound (*sphere packing bound*) of the indentification of E-capacity let us introduce the following function:

$$R_{sp}(E,P^*,W^*)\triangleq\min_{P,V:D(P\circ V||P^*\circ W^*)\le E}I_{P,V}(Z\wedge Y).\tag{2}$$

Theorem. *For the biometric identification system with random parameter for the given P^*, W^* and for all $E > 0$*

$$R_r(E, P^*, W^*) \leq C(E, P^*, W^*) \leq \overline{C}(E, P^*, W^*) \leq R_{sp}(E, P^*, W^*).$$

The proof of the theorem is similar to the proof exposed in [4].

Corollary. *When $E \to 0$ we derive the lower and upper bounds capacity of the channel with random parameter, which coincide and hence, we obtain the capacity*

$$C = I_{P^*,W^*}(Z \wedge Y).$$

References

[1] F. Willems, T. Kalker, J. Goseling, and J.-P. Linnartz, “On the capacity of a biometrical identification system”, *International Symposium on Information Theory*, Yokohama, Japan, p. 82, 2003.

[2] S. Pankanti, R. M. Bolle and A. Jain, “Biometrics-The Future of Identification”, *IEEE Computer*, vol. 33, no. 2, pp. 46 49, February, 2002.

[3] T. Ignatenko and F. Willems, ”Biometric security from an Information-Theoretical perspective”, *Foundations and Trends in Communications and Information Theory*, vol. 7, no 2-3, pp. 135-316, 2012.

[4] M.E.Haroutunian, A. Muradyan and L. Ter-Vardanyan, ”Upper and lower bounds of biometric identification E- capacity”, *Transactions of IIAP of NAS of RA, Mathematical Problems of Computer Science*, V36, pp.1-10, 2012.

[5] C. E. Shannon ”Channel with side information at the transmitter.”, *IBM Res. Developm.*, 1958. V. 2. No. 4. P. 289-293.

[6] S. I. Gelfand , M. S. Pinsker ”Coding for channel with random parameters.”,*Problems of Control and Information Theory*, 1980. V.8. No. 1. P. 19-31.

[7] R. F. Ahlswede, ”Arbitrarily varying channels with states sequence known to the sender.”, *IEEE Transactions on Information Theory*, 1986. V. 32. No. 5. P. 621-629.

[8] M.E. Haroutunian, “The bounds for E-capacity of channel with random parameter.”, *Problems of Information Transmission*, vol. 27, no. 1, pp. 14-23, 1991.

[9] E. A. Haroutunian, M. E. Haroutunian and A. N. Harutyunyan, ”Reliability criteria in information theory and in statistical hypothesis testing”, *Foundations and Trends in Communications and Information Theory*, vol. 4, no 2-3, pp. 97-263, 2008.

[10] M. E. Haroutunian, “Estimates of E-capacity and capacity regions for multiple-access channel with random parameter”, *Lecture Notes in Computer Science*, vol. 4123, Springer Verlag, pp. 196-217, 2006.

[11] E. A. Haroutunian, ”On bounds for E-capacity of DMC”, *IEEE Transactions on Information Theory*, V53, N11, pp. 4210-4220, 2007.

[12] M. E. Haroutunian, S. A. Tonoyan, “Random coding bound of information hiding E-capacity”, *Proc. of IEEE International Symposium on Information Theory*, p. 536, USA, Chicago, 2004.

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

[14] I. Csiszár and J. Körner, *Information Theory: Coding Theorems for Discrete Memoryless Systems*, Academic Press, New York, 1981.

[15] I. Csiszár, “The method of types”, *IEEE Transactions on Information Theory*, vol. 44, no. 6, pp. 2505-2523, 1998.

Submitted 22.11.2012, accepted 14.02.2013.

E-ունակության ուսումնասիրությունը պատահական պարամետրով կենսաչափական նույնականացման արձանագրության համար

Մ. Հարությունյան և Լ. Տեր-Վարդանյան

Ամփոփում

Վերջին տարիներին կենսաչափությունը լայնորեն կիրառվում է անվտանգության ոլորտի տարբեր խնդիրներում։ Աշխատանքում հետազոտվել է կենսաչափական նույնականացման համակարգը՝ ինֆորմացիոն-տեսական տեսանկյունից։ Մենք հետազոտում ենք ցուցչային բարձր հուսալիության չափանիշը՝ կենսաչափական նույնականացման համակարգերում։ Դիտարկված է պատահական պարամետրով կենսաչափական նույնականացման համակարգը, որը կիրառությունների տեսակետից ավելի իրատեսական է։ Կառուցված են E- ունակության վերին և ստորին գնահատականները առավելագույն և միջին սխալի հավանականությունների դեպքում պատահական պարամետրով մոդելի համար։ Երբ $E \to 0$, ստանում ենք պատահական պարամետրով կենսաչափական նույնականացման համակարգի ունակության համապատասխան գնահատականները, որոնք համընկնում են և, որպես հետևանք, ստանում ենք այդ մոդելի նույնականացման ունակությունը։

Исследование E-пропускной способности для протокола биометрической идентификации со случайным параметром

М. Арутюнян и Л. Тер-Варданян

Аннотация

За последние годы биометрика широко используется в различных задачах в сфере безопасности. В данной работе мы исследовали систему биометрической идентификации с информационно-теоретической точки зрения. Мы исследуем критерий экспоненциально высокой надежности в биометрических системах идентификации. Рассмотрена система биометрической идентификации со случайным параметром, которая более реалистична с точки зрения приложений. Построены нижняя и верхняя границы для E-пропускной способности модели со случайным параметром для максимальной и средней вероятности ошибки. Когда $E \to 0$ мы получаем соответствующие оценки пропускной способности для системы биометрической идентификации со случайным параметром, которые совпадают и, как следствие, мы получаем пропускную способность идентификации для этой модели.