Mathematical Problems of Computer Science 48, 74--81, 2017.

Graphical Representations of E-Achievability Region for Identification and Secret-Key Generation System

Lilit A. Ter-Vardanyan

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

Abstract

In this paper we represent the computations of theoretical results, obtained in [1], on the example of binary symmetric channel. The computations and graphical representations are realized using the "Advanced Inftheo" module, created in 2015 [2, 3] for R environment.

Keywords: Biometric identification system, Secret key, Rate-reliability, R environment.

1. Introduction

Nowadays the research of different biometric systems is one of the most important parts of the development of safety measures.

Any biometric system may provide a certain degree of confidentiality in terms of information theory. Depending on applications, different models of biometric identification and secret-key generation were considered [4]. For each model one of the most important tasks is to determine the identification rate or the achievable rate of a secret.

The next most important task is the study of the rate-reliability or E-achievability region, which is not an easy task [5, 6]. Some outer and inner bounds of that region for various models of biometric settings were obtained in [1, 7-10].

Alongside with difficulties in theoretical research, difficulties in applying the theoretical results also arise. Since the obtained mathematical formulas are complex and difficult to calculate, a module for R was created to perform such calculations.

R is a programming language and software environment for statistical analysis, graphics representation and reporting. R was created by Ross Ihaka and Robert Gentleman at the University of Auckland, New Zealand, and is currently developed by the R Development Core Team. The core of R is an interpreted computer language which allows branching and looping as well as modular programming using functions. R allows integration with procedures written in the C, C++, .Net, Python or FORTRAN languages for efficiency.

The use of R environment in modern society is growing very fast due to a number of advantages it has, compared to other statistical tools.

The module "Advanced Inftheo" for R was developed for estimation and computation of complex formulas of Information Theory. In this paper we use this module for computation of outer and inner bounds of E-achievability region for identification and secret-key generation system.

2. The Model of Biometric Identification and Secret-Key Generation System

The Biometric identification and secret-key generation system was considered in [4, 1] (Fig.1).

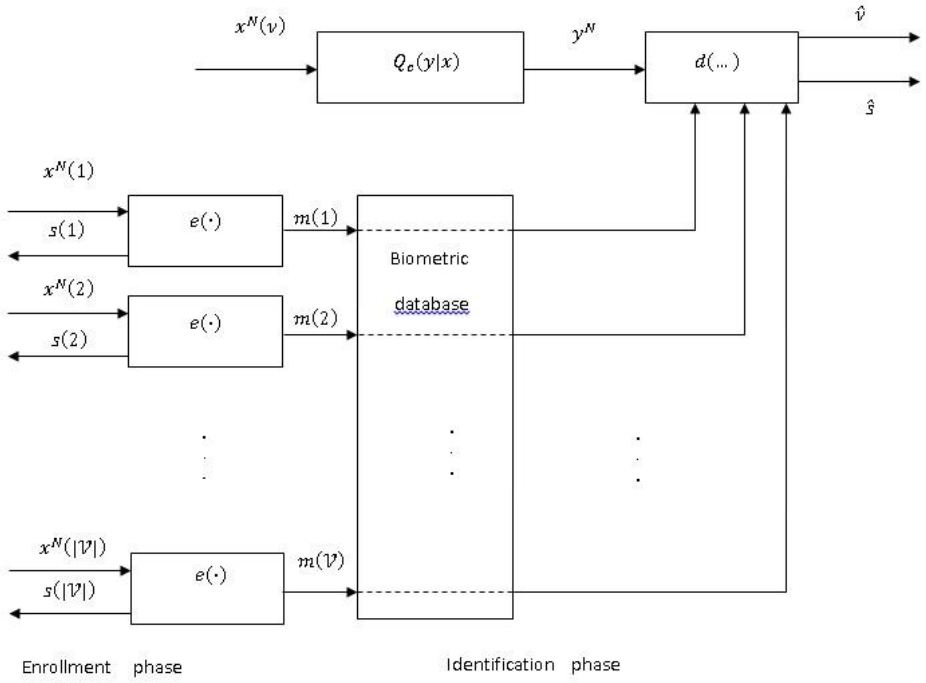

Fig. 1. The Model of Biometric Identification and Secret-key Generation System.

In [1] the inner and outer bounds of E-achievable rates region of this model were obtained. To formulate that result we remind the notations and definitions.

The model of biometric identification and secret-key generation system consists of enrollment and identification procedures.

In an **enrollment phase** $\mathcal{V}$ individuals are observed. For each individual $v = \{1, 2, \dots, |\mathcal{V}|\}$ in the system the biometric source produces a biometric enrollment sequence $\mathbf{x}(v) = (x_1, x_2, \dots, x_N) \in \mathcal{X}^N$. All these sequences are supposed to be generated at random with a given probability distribution (PD)

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

During the enrollment procedure the biometric sequence $\mathbf{x}(v)$ of individual $v = \{1, 2, \dots, \mathcal{V}\}$ is encoded into helper data $m(v) \in \{1, 2, \dots, |\mathcal{M}|\}$ and a secret key $s(v) \in \{1, 2, \dots, |\mathcal{S}|\}$, hence the encoder mapping is $\mathrm{e}(\mathbf{x}(v)) = (m(v); s(v))$ for $v = \{1, 2, \dots, |\mathcal{V}|\}$.

The helper data $m(v)$ is then stored in a (public) database at position v and the generated secret key $s(v)$ is handed over to the individual. The helper data makes the reliable identification possible.

During the **identification procedure** a biometric identification sequence $\boldsymbol{y} = (\mathrm{y}_1, \mathrm{y}_2, \ldots, \mathrm{y}_\mathrm{N}) \in \mathcal{Y}^N$ is observed. If individual v was observed, the sequence $\boldsymbol{y}(v)$ occurs with probability

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

since the biometric channel $W^N(\mathbf{y}|\mathbf{x})$ is memoryless.

During identification, upon observing the biometric identification sequence **y**, the decoder forms an estimate $\hat{v}$ of the identity of the observed individual as well as an estimate of his secret key $\mathrm{s}(\hat{v})$,

$$\big(\hat{v}, \mathrm{s}(\hat{v})\big) = d(\mathbf{y}, m(1), m(2), \ldots, m(\mathcal{V})),$$

where d is the decoder mapping.

Definition (E-achievability): *For $E > 0$ an identification and secret-key rate pair R_I, R_S with $R_I \geq 0$ and $R_I \geq 0$ is E-achievable in a biometric identification setting, if for all $\delta > 0$ for all N large enough, there exists an encoder and a decoder such that*

$$\Pr\{(\hat{v}, \hat{s}) \neq (v, s)\} \leq \exp\{-N(E - \delta)\},$$

$$\frac{1}{N}\log|\mathcal{V}| \geq R_I - \delta,$$

$$\frac{1}{N}\log|\mathcal{S}| \geq R_S - \delta,$$

$$\frac{1}{N} I(S \wedge M) \leq \delta.$$

We use the following PD in the formulation of result:

$$P = \{P(x), x \in \mathcal{X}\};\ V = \{V(y \mid x), y \in \mathcal{Y}, x \in \mathcal{X}\}.$$

For information-theoretic quantities, such as entropy $H_P(X)$, mutual information $I_{P,V}(X \wedge Y)$, divergence $D(V||W|P)$ we refer to [5, 6, 11].

The main result of [1] is the following theorem.

Theorem 1: *For the biometric identification and secret-key generation system with the given Q, W and for all $E > 0$*

$$\mathcal{R}_r(E, Q, W) \subseteq \mathcal{R}(E, Q, W) \subseteq \mathcal{R}_{sp}(E, Q, W),$$

where

$$\mathcal{R}_r(E, Q, W) = \{(R_I, R_S): R_I + R_S \leq \min_{P,V: D(P\circ V||Q\circ W)\leq E} |I_{P,V}(X \wedge Y) + D(P \circ V||Q \circ W) - E|^+\}$$

$$\mathcal{R}_{sp}(E, Q, W) = \{(R_I, R_S): R_I + R_S \leq \min_{P,V: D(P\circ V||Q\circ W)\leq E} I_{P,V}(X \wedge Y)\}.$$

The notation $|a|^+$ is used for $max(a, 0)$.

3. Graphical Representation of Computations

Let's consider a binary symmetric channel with the parameter d, it means that $W(y|x)$ conditional distribution matrix is defined as

$$W(1|1) = W(0|0) = 1 - d,$$

$$W(1|0) = W(0|1) = d.$$

We have performed computations of above formulas in the theorem using “Advanced Inftheo” module.
Figure 2 shows the dependence of the R_I+R_S from E for the parameter $d = 0.2$. The calculations are realized with the step 0.003.

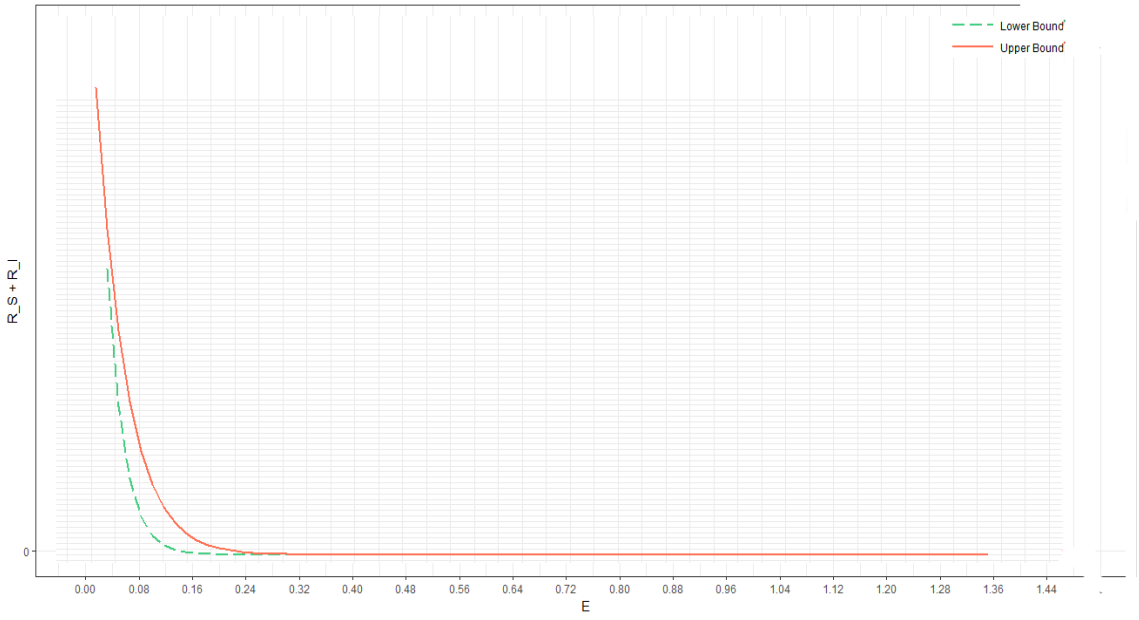

Fig.2. The bounds of E-achievable rate region, when d=0.2 with step 0.003

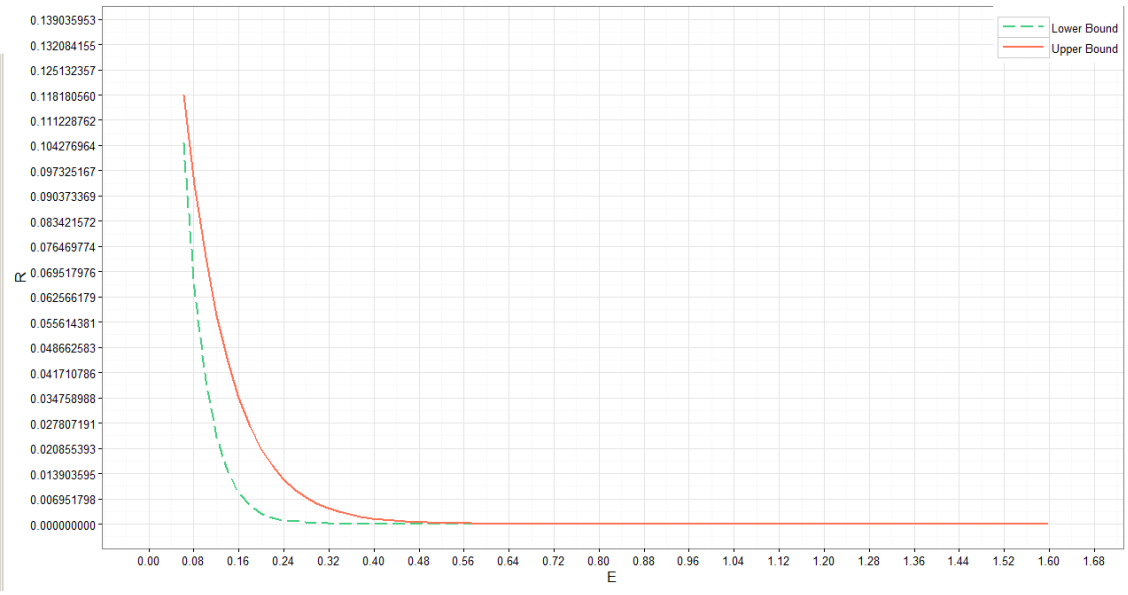

Fig. 3. The bounds of E-achievable identification rate

Figure 3 shows the dependence of identification rate R_I on E and Figure 4 shows the dependence of secret-key rate R_S on E for the same parameter d=0.2.

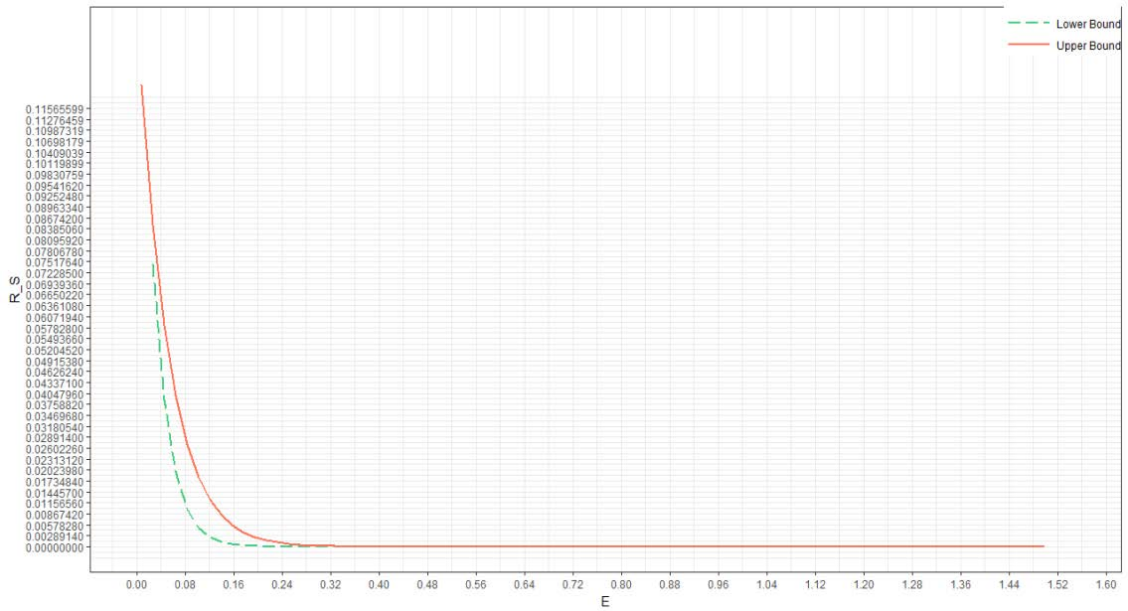

Fig. 4. The bounds of E-achievable secret key rate

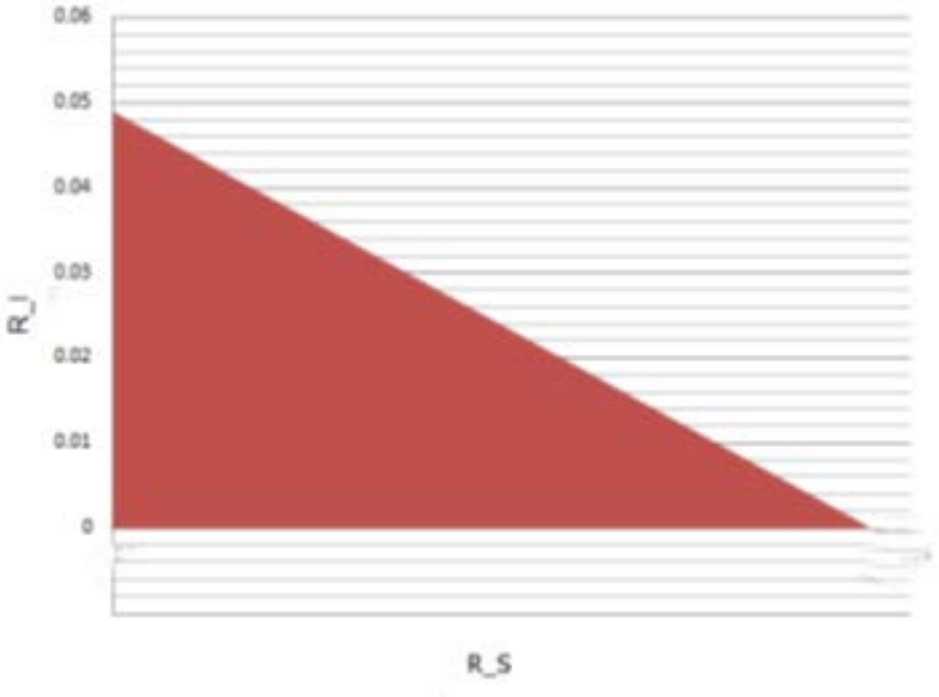

Fig. 5. Identification versus secret-key rate projection, *d=0.2*

In Figure 5 the dependence of R_I onR_S is represented for various E. In other words Figure 5 shows the trade-off between identification and secret-key rates. It means that the more individuals would like to be able to reliably identify, the smaller secret keys can be assigned to individuals for identification purposes and the identification system becomes less secure. It means that it becomes easier to get access to the systems that deploy biometric identification, since smaller secrets are easier to guess and also require less biometric information for their reconstruction.

4. Conclusion

The outer and inner bounds for E-achievability region for the model of biometric identification system and secret-key generation is investigated by constructing graphics for binary symmetric channel.

References

[1] M. Haroutunian, L. Ter-Vardanyan**,** "Investigation of E-achievability region for identification and secret-key generation system*"*, *CSIT-2013 The 9th International Conference on Computer Science and Information Technologies (Dedicated to the 70th anniversary of the National Academy of Sciences of Armenia)*, Yerevan, Armenia, pp. 107-110 , 2013.

[2] N. Pahlevanyan and M. Haroutunian, "Technical solutions of developing Advanced Inftheo new module for R," *CSIT-2015 The 10th International Conference on Computer Science and Information Technologies*, Yerevan, Armenia, pp. 306-309, 2015.

[3] N. Pahlevanyan and M. Haroutunian, "Results of performance analysis of Advanced Inftheo new package for R," *Transactions of IIAP of NAS of RA, Mathematical Problems of Computer Science*, vol. 45, pp. 5-13, 2016.

[4] T. Ignatenko and F. M. 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.

[5] 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. 23, pp. 97–263, 2008.

[6] E. Haroutunian, "On bounds for E - capacity of DMC," *IEEE Transactions on Information Theory*, vol. 53, no. 11, pp. 4210–4220, 2007.

[7] M. Haroutunian, A. Muradyan and L. Ter- Vardanyan, "On the E-capacity of a biometric identification system*"*, C*SIT-2011 The 8th International Conference on Computer Science and Information Technologies*, Yerevan, Armenia, pp. 132-134, 2011.

[8] M. E. Haroutunian, A. Muradyan and L. Ter-Vardanyan, "Upper and lower bounds of biometric identification E-capacity", *Transactions of IIAP of NAS RA, Mathematical Problems of Computer Science*, vol. 36, pp .1-10, 2012.

[9] M. Haroutunian and N. Pahlevanyan "Information theoretical analysis of biometric secret key sharing model," *Transactions of IIAP of NAS RA, Mathematical Problems of Computer Science*, vol. 42, pp. 17-27, 2014.

[10] M. Haroutunian and L. Ter-Vardanyan, "Rate-reliability for protected biometric identification system with secret generation*"*, *From Information Age to Big Data Era*, Yerevan, Armenia, pp. 54-68, 2016.

[11] T. M. Cover and J. A. Thomas, *Elements of Information Theory*, 2nd Edition, New York, NY, USA: Wiley-Interscience, 2006.

Submitted 05.08.2017, accepted 04. 12.2017.

Նույնականացման և գաղտնի բանալի գեներացման համակարգի համար E-հասանելի տիրույթի գրաֆիկական ներկայացումները

Լ. Տեր-Վարդանյան

Ամփոփում

Հոդվածում ներկայացված են [1]-ում ստացված տեսական արդյունքների հաշվարկները երկուական սիմետրիկ կապուղու օրինակով: Հաշվարկները և գրաֆիկները կատարվել են 2015թ.-ին R միջավայրի համար ստեղծված «Advanced Inftheo» մոդուլի միջոցով [2, 3]:

Графические представления области E-достижимости для системы идентификации и генерации секретного ключа

Л. Тер-Варданян

Аннотация

В статье представлены расчеты теоретических результатов, полученных в [1], на примере двоичного симметричного канала. Вычисления и графические представления реализованы с использованием модуля «Advanced Inftheo», созданного в 2015 году для среды R [2, 3].