A NOTE ON YAO'S THEOREM ABOUT PSEUDO-RANDOM GENERATORS

STÉPHANE BALLET AND ROBERT ROLLAND

1. Introduction

The study of provable security notions can be done from different points of view:

(1) the study can be done in a static context, with given parameters. In this case, the sizes of the objects are fixed.
(2) on the contrary the study can be done in a dynamic context, namely the system depends on a variable parameter (the so-called security parameter) growing to infinity.

In the paper [Yao(1982)], the study is done in an asymptotic context an for one pseudo-random generator for each value of the security parameter. In this paper, we present modified versions of Yao's results, in particular a static version for a pseudo-random generator and for a finite family of pseudo-random generators. In each case, we give detailed proofs and we compute the exact cost of the reduction between the notion of indistinguishability and the notion of unpredictability of the next bit. We stress that this point of view can be of interest for a practical study of concrete pseudo-random generators with a fixed length. Finally we derive from the preceding results an asymptotic version which generalizes the classical Yao's theorem to a family of pseudo-random generators having the same security parameter k, as k tends to infinity.

2. Pseudo-random generators

2.1. Definition of a pseudo-random generator.

Definition 2.1. *A pseudo-random generator is a deterministic function f defined on a subset $\mathbf{U} \subseteq \{0,1\}^k$ into $\{0,1\}^l$ (where $k < l$) which maps a seed $X_0 \in \mathbf{U}$ (a secret seed) to a finite sequence of l bits:*

$$f(X_0) = (x_1, x_2, \cdots, x_l).$$

3. The security of a pseudo-random generator

3.1. Indistinguishable pseudo-random generator.

Let us consider the following pseudo-random generator:

$$f : \mathbf{U} \subset \{0,1\}^k \to \{0,1\}^l \text{ where } k < l.$$

If $\mathcal{A}$ is a probabilistic algorithm which outputs one bit, We will denote by $\mu_{\mathcal{A}}(e)$ the probability of the output 1 when the input of $\mathcal{A}$ is e.

Key words and phrases. complexity, cryptography, distinguishable, prediction, pseudo-randomness, pseudo-random number generator, Yao's theorem.

The following random experiment, involves a probabilistic algorithm $\mathcal{A}$ having for input a vector $Y \in \{0,1\}^l$ and which outputs one bit. It has for aim to recognize if a sequence Y is the output of the pseudo-random generator f or of a truly random generator.

Let us fix an integer s such that $0 \leq s \leq l$.

$\mathbf{Expt}_{\mathbf{f,s}}^{\mathbf{dist}}(\mathcal{A})$
  $X_0 \leftarrow \mathbf{U} \subseteq \{0,1\}^k$
  $X \leftarrow f(X_0)$
    (notation : $X = (x_1, \cdots, x_l)$)
  $Y_1 \leftarrow (x_1, x_2, \cdots, x_s)$
  $Y_2 \leftarrow \{0,1\}^{l-s}$
  $Y \leftarrow Y_1 || Y_2$
  $b \leftarrow \mathcal{A}(Y)$
  return b
End.

Let $q_{f,s}$ be the probability that the experiment

$$\mathbf{Expt}_{\mathbf{f,s}}^{\mathbf{dist}}(\mathcal{A})$$

returns $b = 1$.

In particular, $q_{f,0}$ is the probability of the following event: we draw at random an element of $\{0,1\}^l$ according to the uniform distribution, we run the algorithm $\mathcal{A}$ on this element, and the output is 1. The probability $q_{f,l}$ is the probability of the following event: we draw at random a seed X_0 in $\mathbf{U}$, we apply f to obtain an element of $\{0,1\}^l$ which becomes the input of the algorithm $\mathcal{A}$, and the output is 1.

Let us recall now the notion of advantage which permits to quantify the ability of $\mathcal{A}$ to distinguish f.

Definition 3.1. *The advantage of the algorithm $\mathcal{A}$ to distinguish f is*

$$Adv_f^{dist}(\mathcal{A}) = |q_{f,l} - q_{f,0}|.$$

Then we define a (T, ϵ)-distinguisher:

Definition 3.2. *Let f be a pseudo-random generator. Let T and ϵ be positive real numbers. A (T, ϵ)-distinguisher for f is a probabilistic algorithm $\mathcal{A}$ such that*

1. *the maximal running time of $\mathcal{A}$ is $\leq T$,*
2. *the input of $\mathcal{A}$ is an element of $\{0,1\}^l$,*
3. *the output of $\mathcal{A}$ is a bit b,*
4. *the algorithm $\mathcal{A}$ can distinguish the pseudo-random generator from the uniform distribution, namely*

$$Adv_f^{dist}(\mathcal{A}) > \epsilon.$$

3.2. **Unpredictable pseudo-random generator.** Let us consider the following pseudo-random generator:

$$f : \mathbf{U} \subset \{0,1\}^k \to \{0,1\}^l.$$

Let $1 \leq s < l$. The following random experiment involves a probabilistic algorithm $\mathcal{B}$ having for input a sequence of s bits and for output a bit. Roughly speaking, this

algorithm tries to predict the next bit produced by the pseudo-random generator f, namely the bit of index $s+1$.

$\mathbf{Expt}_{\mathbf{f,s}}^{\mathbf{pred}}(\mathcal{B})$
  $X_0 \leftarrow \mathbf{U} \subseteq \{0,1\}^k$
  $X \leftarrow f(X_0)$
    (notation : $X = (x_1, \cdots, x_l)$)
  $Y \leftarrow (x_1, x_2, \cdots, x_s)$
  $b \leftarrow \mathcal{B}(Y)$
  if $b = x_{s+1}$
    then return 1
    else return 0
  fi
End.

Let $r_{f,s}$ be the probability that the experiment $\mathbf{Expt}_{\mathbf{f,s}}^{\mathbf{pred}}(\mathcal{B})$ returns 1.

Definition 3.3. *The advantage of the algorithm $\mathcal{B}$ to predict the bit of index $(s+1)$ computed by f is:*

$$Adv_{f,s}^{pred}(\mathcal{B}) = \left| r_{f,s} - \frac{1}{2} \right|.$$

We can now define the notion of (T, s, ϵ)-prediction algorithm.

Definition 3.4. *Let f be a pseudo-random generator. Let T and ϵ be positive real numbers and s be an integer such that $1 \leq s < l$. A (T, s, ϵ)-prediction algorithm $\mathcal{B}$ is a probabilistic algorithm such that:*

(1) *the maximal running time of $\mathcal{B}$ is $\leq T$,*
(2) *the input of $\mathcal{B}$ is an element of $\{0,1\}^s$,*
(3) *the output of $\mathcal{B}$ is a bit,*
(4) *the algorithm $\mathcal{B}$ can predict the next bit, namely*

$$Adv_{f,s}^{pred}(\mathcal{B}) > \epsilon.$$

3.3. **Yao's theorem, static version.** The Yao's theorem relates the notion of security to the notion of unpredictability of the next bit. We express it in its non-asymptotic form. In this case, we give two results which can be considered respectively as a necessary condition and a sufficient condition to have the security of a generator f.

Theorem 3.5. *We consider the following pseudo-random generator:*

$$f : \mathbf{U} \subset \{0,1\}^k \rightarrow \{0,1\}^l.$$

If we have a (T, s, ϵ)-prediction algorithm for f, we can build a $(T+c, \epsilon)$-distinguisher where c is the constant time needed to compare two bits.

Theorem 3.6. *Let f be a pseudo-random generator:*

$$f : \mathbf{U} \subset \{0,1\}^k \rightarrow \{0,1\}^l.$$

Let us suppose that for all s such that $1 \leq s < l$, it does not exist any (T, s, ϵ)-prediction algorithm. Then f is $(T - (c_1 l + c_2), l\,\epsilon)$-secure where c_1 is the constant time needed to draw one bit at random, and c_2 is the constant time needed to test the value of a bit and then depending upon the value of this bit to return a bit or its complementary.

4. The security of a family of pseudo-random generators with same given size

We have considered the case of one pseudo-random generator f. But even in the non-asymptotic case where k and l are fixed, we have to study not only one, but a family (a finite family because k and l are fixed) Γ of function f defined on a subset $\mathbf{U}_f$ (which can depend on f) of $\{0,1\}^k$ with images in $\{0,1\}^l$.

4.1. Revisiting the previous notions in the case of a family of pseudo-random generators with same size.

In a realistic situation we must, in the random experiment which defines the attacker's advantage, draw at random the function f in the family Γ according to a probability δ. So, we replace now the algorithms $\mathcal{A}$ and $\mathcal{B}$ of the previous section by algorithms whose inputs are a function $f \in \Gamma$ and an a vector. The random experiments $\mathbf{Expt}^{\mathbf{dist}}_{\mathbf{f,s}}(\mathcal{A})$ and $\mathbf{Expt}^{\mathbf{pred}}_{\mathbf{f,s}}(\mathcal{B})$ are replaced by the random experiments $\mathbf{Expt}^{\mathbf{dist}}_{\mathbf{\Gamma,s}}(\mathcal{A})$ and $\mathbf{Expt}^{\mathbf{pred}}_{\mathbf{\Gamma,s}}(\mathcal{B})$ where we draw at random not only the seed X_0, but also the function f itself. Then we extend the previous definitions for the case of a family of pseudo-random generator and we get a an adapted version of the Yao's theorem.

4.2. Yao's theorem.

Theorem 4.1. *Let Γ be a family of pseudo-random generators having the same size where each $f \in \Gamma$ is a function*

$$f : \mathbf{U}_f \subset \{0,1\}^k \to \{0,1\}^l.$$

If we have a (T, s, ϵ)-prediction algorithm for Γ, we can build a $(T+c, \epsilon)$-distinguisher for Γ where c is the constant time needed to compare two bits.

Theorem 4.2. *Let Γ be a family of pseudo-random generators having the same size where each $f \in \Gamma$ is a function*

$$f : \mathbf{U}_f \subset \{0,1\}^k \to \{0,1\}^l.$$

Let us suppose that for all s such that $1 \leq s < l$, it does not exist any (T, s, ϵ)-prediction algorithm. Then f is $(T - (c_1 l + c_2), l\,\epsilon)$-secure where c_1 is the constant time needed to draw one bit at random, and c_2 is the constant time needed to test the value of a bit and then depending upon the value of this bit to return a bit or its complementary.

5. Asymptotic behavior

As a consequence of the previous results for fixed k and l, we can deduce results on the asymptotic theory of the pseudo-random generators. Let k be a positive integer (the security parameter) and $l(k)$ a polynomial function of k such that $l(k) > k$. For any k we have a set Γ_k of deterministic functions such that

1. if $f \in \Gamma_k$ then f is a function from a subset $\mathbf{U}_f$ of $\{0,1\}^k$ into $\{0,1\}^{l(k)}$;
2. there exist a polynomial function $t(k)$ such that for any k, any $f \in \Gamma_k$ and any $X \in U_f$ the computation time of $f(X)$ is upper-bounded by $t(k)$;
3. for any k we provide a probability δ_k on the set Γ_k.

Definition 5.1. *The family $\Gamma = (\Gamma_k)_{k>0}$ of sets of pseudo-random generators is said asymptotically secure if for any polynomial $S(k)$, any integer u and any*

distinguisher $\mathcal{A}$ with running time $\leq S(k)$, the advantage of the algorithm $\mathcal{A}_k$ is a negligible function of $\frac{1}{k^u}$, namely

$$\lim_{k\to+\infty} k^u \, Adv^{dist}_{\Gamma_k}(\mathcal{A}_k) = 0.$$

Let $s = (s_k)_{k\geq 1}$ a sequence of integers such that $1 \leq s_k < l(k)$. We define now a s-prediction algorithm to be a probabilistic polynomial algorithm $\mathcal{B}$ having for inputs the security parameter k, a function $f \in \Gamma_k$ and a vector $Z \in \{0,1\}^{s_k}$, and which outputs a bit. Let k be an integer, we will denote by $\mathcal{B}_k$ the probabilistic algorithm obtained from $\mathcal{B}$ by fixing the first entry to the value k.

Definition 5.2. *The family $\Gamma = (\Gamma_k)_{k>0}$ of sets of pseudo-random generators is said asymptotically unpredictable if for any polynomial $S(k)$, any sequence s and any s-prediction algorithm $\mathcal{B}$ with running time $\leq S(k)$, the advantage of the s_k-prediction algorithm $\mathcal{B}_k$ is a negligible function of $\frac{1}{k^u}$, namely*

$$\lim_{k\to+\infty} k^u \, Adv^{pred}_{\Gamma_k,s_k}(\mathcal{A}_k) = 0.$$

The two notions are related by the following theorem:

Theorem 5.3. *Let $l(k)$ be a polynomial function of one integer variable k such that $l(l) > k$. Let $\Gamma = (\Gamma_k)_{k>0}$ a family of sets, where any set Γ_k is a probability set of random generators mapping a subset of $\{0,1\}^k$ into $\{0,1\}^{l(k)}$ (more precisely, each $f \in \Gamma_k$ has its own definition subset $\mathbf{U}_f \subseteq \{0,1\}^k$). The family Γ is asymptotically secure if and only if it is asymptotically unpredictable.*

References

[Goldreich(1999)] Goldreich O (1999) Modern Cryptography, Probabilistic Proofs and Pseudo-randomness. No. 17 in Algorithms and Combinatorics, Springer

[Goldreich(2001)] Goldreich O (2001) The Foundations of Cryptography, Volume I. Cambridge University Press

[Luby(1996)] Luby M (1996) Pseudorandomness and Cryptographic Applications. Princeton University Press

[Stinson(2005)] Stinson D (2005) Cryptography: Theory and Practice, Third Edition. CRC Press

[Yao(1982)] Yao AC (1982) Theory and Applications of Trapdoor Functions. In: Proceedings of the 23rd IEEE Symposium on Foundations of Computer Science, IEEE Computer Society, pp 80–91

Institut de Mathématiques de Luminy and eRISCS, case 930, F13288 Marseille cedex 9, France

E-mail address: ballet@iml.univ-mrs.fr

Institut de Mathématiques de Luminy and eRISCS, Campus de Luminy, Case 907, 13288 MARSEILLE Cedex 9

E-mail address: robert.rolland@acrypta.fr