THE TRANSACTIONS OF THE IEICE, VOL. E 72, NO. 7 JULY 1989

PAPER

Several Theorems on Probabilistic Cryptosystems

Yuliang ZHENG†, *Nonmember*, Tsutomu MATSUMOTO†
and Hideki IMAI†, *Members*

SUMMARY This paper proves several theorems on probabilistic cryptosystems. From these theorems it follows directly that a probabilistic cryptosystem proposed by the authors[7], whose security is based upon the (supposed) infeasibility of γ^{th}-Residuosity Problem, is polynomially secure. Techniques developed in the paper are of independent interest.

1. Introduction

The authors proposed in Ref. (7) a probabilistic cryptosystem based upon the intractability of a famous Number-Theoretic problem called γ^{th}-Residuosity Problem (γ^{th}-RP). This cryptosystem can be viewed as a byte-by-byte generalization of a bit-by-bit probabilistic cryptosystem (the GM cryptosystem) discovered by Goldwasser and Micali in Ref. (4)*.

Let m_1 and m_2 be two messages. Intuitively, a cryptosystem is polynomially secure if no polynomial size bounded adversary, when given two encryptions one of which is for m_1 and the other for m_2, can tell which encryption corresponds to which message.

Under Quadratic Residuosity Assumption, the GM cryptosystem was shown to be polynomially secure[4]. Benaloh and Yung claimed[2] that the theorems proved in Ref.(4), which were concerned with the binary case of $\gamma=2$, can be "directly" generalized to the case of $\gamma>2$. They, however, did not present any hint on how to "directly" generalize the theorems. To the authors' knowledge, no published paper has suggested how to treat the general case of $\gamma>2$. It seems to the authors that it is not so obvious to generalize the theorems proved in Ref.(4).

The main purpose of this paper is to give proofs for several theorems on probabilistic cryptosystems, from which it follows directly that our probabilistic cryptosystem is polynomially secure. Techniques developed in the paper are of independent interest.

The remaining part of the paper is organized as follows. First, we review concisely the notion of polynomial security and our generalized probabilistic cryptosystem based upon γ^{th}-RP (Sect. 2). Then we extend the definition of unapproximable trapdoor predicates [4] to that of unapproximable trapdoor functions (Sect. 3), and construct a probabilistic cryptosystem $\mathcal{C}_{UTF}$ based upon any unapproximable trapdoor function (Sect. 4). We proceed to prove that under γ^{th}-Residuosity Assumption, the set of class-index functions $\boldsymbol{I}=\bigcup \boldsymbol{I}_k$ is an unapproximable trapdoor function (Theorems 1 and 2, Sect. 5), and the cryptosystem $\mathcal{C}_{UTF}$ is polynomially secure (Theorem 3, Sect. 6). From Theorems 2 and 3 it follows that under γ^{th}-Residuosity Assumption, our generalized cryptosystem is indeed polynomially secure (Theorem 4, Sect. 6).

2. Preliminaries

This section briefly reviews the notion of polynomial security and an efficient probabilistic cryptosystem whose security is based upon the intractability of γ^{th}-RP. See Refs.(4) and (7) for details.

2.1 Polynomial Security

Denote by $\mathcal{N}$ the set of positive integers. For any $n\in\mathcal{N}$, define $Z_n=\{0, 1, \cdots, n-1\}$, and $Z_n^*=\{x|x\in Z_n, \gcd(x, n)=1\}$. The number of elements in a finite set S is denoted by $\#S$, and in particular $\#Z_n^*$ is denoted by $\phi(n)$.

The concatenation of two sequences x and y over some alphabet is represented by $x\|y$. When we want to emphasize that an algorithm A (circuit C, respectively) receives t inputs, we write $A(\overbrace{\cdot, \cdot, \cdots, \cdot}^{t \text{ dots}})$ ($C[\overbrace{\cdot, \cdot, \cdots, \cdot}^{t \text{ dots}}]$, respectively).

Define K be the set of all security parameters [4]. For simplicity assume that K is an infinite subset of $\mathcal{N}$. We will use a $k\in K$ as an input to a cryptosystem defined below to create a pair of encryption/decryption algorithms. k determines many quantities such as the plaintext length and the security strength of a pair of

Manuscript received November 11, 1988.
Manuscript revised March 8, 1989.

† The authors are with the Faculty of Engineering, Yokohama National University, Yokohama-shi, 240 Japan.

* Before us, Benaloh and Yung [2] used a probabilistic cryptosystem presented in the appendix of Ref. (2) in constructing an election scheme. The form of this cryptosystem can be viewd as a byte-by-byte generalization of the GM cryptosysem, but as we will mention later in this paper, no proof on the security of the cryptosystem can be found in Ref. (2). Our cryptosystem looks like Benaloh-Yung's (see Ref. (7) for details) but is proved to be polynomially secure under γ^{th}-Residuosity Assumption.

encryption/decryption algorithms. For every $k\in K$ there is a set M_k called the message space associated with k.

A cryptosystem is a polynomial time algorithm $\mathcal{C}$ which on input a $k\in K$, outputs a pair $\langle E, D\rangle$ of encryption/decryption algorithms. Note that generally there are a lot of pairs $\langle E, D\rangle$ corresponding to a given k, and on input k the algorithm $\mathcal{C}$ outputs one of the pairs in a randomized and uniform way. The set of all pairs generated by $\mathcal{C}$ on input k is denoted by $\mathcal{C}(k)$, and for any $\langle E, D\rangle\in\mathcal{C}(k)$, the set of all encryptions of $m\in M_k$ is denoted by $E(m)$.

A cryptosystem $\mathcal{C}$ is called probabilistic (or randomized) if E is a probabilistic (or randomized) algorithm for all $k\in K$ and all $\langle E, D\rangle\in\mathcal{C}(k)$. Note that for a probabilistic cryptosystem $\mathcal{C}$, $\#E(m)$ is typically quite large.

The notion of polynomial security is set up by means of two kinds of polynomial size circuits—line tappers and message finders[4].

A line tapper is a family of polynomial size circuits $T=\{T_k|k\in K\}$. Each T_k has one Boolean output and two inputs: one for code(E) with $\langle E, D\rangle\in\mathcal{C}(k)$ and the other for an encryption $\alpha\in E(m)$, where m belongs to M_k and code(E) denotes a suitable encoding of the encryption algorithm E. For any $m\in M_k$, let $p(m)$ denote the probability that T_k outputs 1 on input code(E) and $\alpha\in E(m)$. $p(m)$ is taken over all the encryptions $\alpha\in E(m)$.

A message finder is also a family of polynomial size circuits $F=\{F_k|k\in K\}$. For each $k\in K$ and $\langle E, D\rangle\in\mathcal{C}(k)$, the circuit F_k outputs two messages $m_1^k, m_2^k\in M_k$, on input code(E).

[Definition 1] Let Q_1 and Q_2 be two polynomials. A cryptosystem $\mathcal{C}$ is polynomially secure if there is no message finder F such that for infinitely many $k\in K$, for a line tapper T, for any $\langle E, D\rangle\in\mathcal{C}(k)$, F_k finds with probability greater than $1/Q_2(k)$ two messages $m_1^k, m_2^k\in M_k$ such that

$$|p(m_1^k)-p(m_2^k)|\geq\frac{1}{Q_1(k)}.$$ □

2.2 A Cryptosystem Based on γ^{th}-RP

Let γ and n be positive integers. An integer x with $\gcd(x, n)=1$ is called a γ^{th}-residue mod n if there exsts an integer w such that $x\equiv w^\gamma(\mathrm{mod}\ n)$, or a γ^{th}-nonresidue mod n if there doesn't exist such a w.

Assume that γ is an odd integer with $\gamma\geq 3$. Call $n\in\mathcal{N}$ a hard integer about $k\in K$ if $n=pq=(2\gamma p'+1)(2q'+1)$ such that $|p'|=|q'|=k$, $\gcd(\gamma, q')=1$ and the four integers p, q, p', q' are all primes. Denote by H_k^γ the set of all hard integers about k.

Let $n=pq\in H_k^\gamma$, and let h_p, h_q be primitive roots mod p and mod q, respectively, such that $h_p\equiv 1(\mathrm{mod}\ q)$ and $h_q\equiv 1(\mathrm{mod}\ p)$. For any $y\in Z_n^*$, we call the triple (n, γ, y) acceptable if y can be written as $y\equiv h_p^{b_p\gamma+e}\cdot h_q^{b_q}(\mathrm{mod}\ n)$, where $0<e<\gamma$, $\gcd(e, \gamma)=1$, $1\leq b_p\leq\phi(p)$ and $1\leq b_q\leq\phi(q)$.

It is proved in Ref.(7) that for any acceptable triple (n, γ, y), every element $x\in Z_n^*$ can be represented as $x\equiv y^i\cdot w^\gamma(\mathrm{mod}\ n)$ with a unique $0\leq i<\gamma$ and some $w\in Z_n^*$. This unique i is called the class-index of x with respect to (n, γ, y). Thus we obtain a function $I_{n\|\gamma\|y}: Z_n^*\to Z_\gamma$, which we call a class-index function and is defined as $I_{n\|\gamma\|y}(x)=i$ iff $x=y^i\cdot w^\gamma \bmod n$ for some $w\in Z_n^*$. Clearly, from the definition of $I_{n\|\gamma\|y}$, an element $x\in Z_n^*$ is a γ^{th}-residue mod n iff $I_{n\|\gamma\|y}(x)=0$.

It is also proved in Ref.(7) that for an acceptable triple (n, γ, y) where $\gamma=O(k^c)$†, the following three closely related problems are polynomially equivalent:

(1) γ^{th}-RP: Given a randomly selected element $z\in Z_n^*$, decide whether or not z is a γ^{th}-residue mod n.

(2) Class-Index-Comparing Problem: Given two randomly selected elements $z_1, z_2\in Z_n^*$, judge whether or not z_1 and z_2 have the same class-index with respect to (n, γ, y).

(3) Class-Index-Finding Problem: Given a randomly selected element $z\in Z_n^*$, find the class-index of z with respect to (n, γ, y).

There are strong evidences[7] which support the conjecture that γ^{th}-RP is intractable unless the factorization of n is known. This conjecture is formally stated as:

[Definition 2] (γ^{th}-Residuosity Assumption)†† Let Q_1, Q_2 be polynomials, and $C_{k,RP}[\cdot,\cdot]$ a circuit with one Boolean output. Call an integer $n\in H_k^\gamma$ easy with respect to $C_{k,RP}[\cdot,\cdot]$ if for a fraction $1-1/Q_1(k)$ of $z\in Z_n^*$, $C_{k,RP}[n, z]=1$ iff z is a γ^{th}-residue mod n. Denote by $H_k^\gamma(C_{k,RP})$ the set of all easy integers in H_k^γ with respect to $C_{k,RP}[\cdot,\cdot]$. Then for any polynomial size circuit $C_{k,RP}[\cdot,\cdot]$, for any polynomials Q_1, Q_2, and for all sufficiently large k,

$$\frac{\#H_k^\gamma(C_{k,RP})}{\#H_k^\gamma}<\frac{1}{Q_2(k)}.$$ □

Let L be a polynomial in k. For each $k\in K$ define $M_k=\{m_l\|m_{l-1}\|\cdots\|m_1|m_i\in Z_\gamma, l=L(k)\}$. Based on γ^{th}-Residuosity Assumption, a probabilistic cryptosystem is constructed in the following way: Let (n, γ, y) be an acceptable triple selected by Alice, where $\gamma=O(k^c)$. Alice makes n, γ and y public, but keeps the factorization (p, q) of n secret. Now suppose Bob wants to securely send a message $m\in M_k$ to Alice. The encryption algorithm for Bob and the decryption algorithm for Alice are as follows:

ENCRYPTION ALGORITHM $E(n, \gamma, y, m)$

From $i=1$ to l, randomly choose a $w_i\in Z_n^*$, and

† By $\gamma=O(k^c)$, we mean that γ is bounded by a polynomial in k.

†† This assumption is more formal, but a little stronger, than that given in Ref.(7).

compute $c_i := y^{m_i} \cdot w_i^\gamma \bmod n$. Let $c = c_l \| c_{l-1} \| \cdots \| c_1$ be an encryption of the message $m = m_l \| m_{l-1} \| \cdots \| m_1$.

DECRYPTION ALGORITHM $D(p, q, \gamma, y, c)$

For each $c_i, 1 \le i \le l$, do as follows: (*) Randomly select an $f \in Z_\gamma$ as well as an $x \in Z_n^*$, and compute $z := y^f \cdot x^\gamma \cdot c_i \bmod n$. If $z^{(p-1)/\gamma} \equiv 1 (\bmod p)$ then let $m_i := (\gamma - f) \bmod \gamma$, or otherwise return to (*). $m = m_l \| m_{l-1} \| \cdots \| m_1$ is the message concealed in the encryption $c = c_l \| c_{l-1} \| \cdots c_1$.

3. Unapproximable Trapdoor Functions

We now generalize unapproximable trapdoor predicates defined in Ref.(4) to unapproximable trapdoor functions.

Let P_1, P_2 be two polynomials. For any $k \in K$, let S_k denote a subset of integers each of which is of $P_1(k)$-bit long. For any $n \in S_k$, let Ω_n be a subset of $P_2(k)$-bit sequences, i. e., $\Omega_n \subseteq \{0, 1\}^{P_2(k)}$.

Assume that $\Sigma = \{0, 1, \cdots, \beta - 1\}$, where $\beta \in \mathcal{N}$. For any $n \in S_k$, define a function $Y_n : \Omega_n \to \Sigma$. Denote the set of functions indexed by integers of length $P_1(k)$ by $\boldsymbol{Y}_k = \{Y_n | n \in S_k\}$. Also let $\boldsymbol{Y} = \bigcup \boldsymbol{Y}_k$, where $\bigcup$ is the union operation about all security parameters in K.

[Definition 3] An approximater is a family $A = \{A_k | k \in K\}$ of circuits, where every A_k is a polynomial size circuit with two inputs and one output. On input an $n \in S_k$ and $x \in \Omega_n$, A_k outputs the binary representation of an integer $i \in \Sigma$.

[Definition 4] Let $A = \{A_k | k \in K\}$ be an approximater, let $\boldsymbol{Y} = \bigcup \boldsymbol{Y}_k$, and let $\varepsilon : K \to [0, 1/2]$ be a function. For $n \in S_k$, we say that A_k $\varepsilon(k)$-approximates the function $Y_n : \Omega_n \to \Sigma$ if $A_k[n, x] = Y_n(x)$ for a fraction at least $1/\beta + \varepsilon(k)$ of the x in Ω_n. □

[Definition 5] Let $A = \{A_k | k \in K\}$ be an approximater, and let $\boldsymbol{Y} = \bigcup \boldsymbol{Y}_k$. For any $k \in K$, and for any function $\varepsilon : K \to [0, 1/2]$, denote by $S_k(A_k, \varepsilon(k))$ the set of integers $n \in S_k$ for which A_k $\varepsilon(k)$-approximates the function $Y_n : \Omega_n \to \Sigma$. $\boldsymbol{Y}$ is called unapproximable if for any approximater A, for any polynomials Q_1 and Q_2, for all sufficiently large $k \in K$,

$$\frac{\#S_k\left(A_k, \frac{1}{Q_1(k)}\right)}{\#S_k} < \frac{1}{Q_2(k)}. \qquad \square$$

For any $k \in K$, $n \in S_k$, and $i \in \Sigma$, let $\Omega_n^i = \{x | x \in \Omega_n, Y_n(x) = i\}$. Thus Ω_n^i is the set of elements in Ω_n which have the same image i under the function Y_n.

[Definition 6] Let $\boldsymbol{Y} = \bigcup \boldsymbol{Y}_k$. $\boldsymbol{Y}$ is called an unapproximable trapdoor function if:

1. $\boldsymbol{Y}$ is unapproximable.
2. $\boldsymbol{Y}$ is trapdoor in the following sense:
   2.1. (Encryptability Condition) Given $n \in S_k$ and $i \in \Sigma$, it is easy (i. e., can be done in probabilistic polynomial time) to select an x randomly and uniformly from Ω_n^i.
   2.2. (Decryptability Condition) There is a function $\delta : \bigcup S_k \to \mathcal{N}$ such that $\delta(n)$ is of polynomial size in k for all n, and such that for all $n \in S_k$, and for any $x \in \Omega_n$, it is easy to compute $Y_n(x)$ from x, n and $\delta(n)$. $\delta(n)$ is called the secret trapdoor of n.
   2.3. (Constructability Condition) For any $k \in K$, it is easy to select randomly and uniformly an $n \in S_k$ and hence its secret trapdoor $\delta(n)$. □

The definition of unapproximable trapdoor predicates is obtained by restricting β to 2, i. e., by setting $\Sigma = \{0, 1\}$ in Definitions 3, 4, 5 and 6.

Let $\boldsymbol{Y} = \bigcup \boldsymbol{Y}_k$ be an unapproximable trapdoor function, and Q_1, Q_2 be two polynomials. Then there is an important fact which will be used later in the proof of Theorem 3: For all sufficiently large k, for a fraction greater than $1 - 1/Q_2(k)$ of the n in S_k, and for all $i \in \Sigma$, we have $|\#\Omega_n^i / \#\Omega_n - 1/\beta| < 1/Q_1(k)$.

If the fact was not true, then for infinitely many $k \in K$ there would be trivial circuits A_k each of which could $1/((\beta - 1)Q_1(k))$-approximate $Y_n : \Omega_n \to \Sigma$ for a fraction $1/Q_2(k)$ of the n in S_k. A_k works as follows: On input an $n \in S_k$ and an $x \in \Omega_n$, it outputs i whenever $\#\Omega_n^i / \#\Omega_n - 1/\beta \ge 1/Q_1(k)$ for some $i \in \Sigma$, outputs a randomly and uniformly selected $j \in \Sigma - \{i\}$ whenever $1/\beta - \#\Omega_n^i / \#\Omega_n \ge 1/Q_1(k)$ for some $i \in \Sigma$, and outputs 0 otherwise. So the fact is true.

Several functions have been proved to be unapproximable trapdoor ones (under some reasonable assumptions). One of them is based on the difficulty of breaking the RSA cryptosystem and has now become well-known in the literature. For details see Ref.(1).

We now consider another candidate for unapproximable trapdoor functions: Let $\beta = \gamma$, where $\gamma > 2$ is an odd. Thus $\Sigma = \{0, 1, \cdots, \gamma - 1\} = Z_\gamma$. Let H_k^γ be the hard integer set. Also let $S_k = \{n \| \gamma \| y | n \in H_k^\gamma, y \in Z_n^*$ such that (n, γ, y) is an acceptable triple$\}$, and let $\Omega_{n\|\gamma\|y} = Z_n^*$ for any $n \| \gamma \| y \in S_k$. Finally, let $\boldsymbol{Y} = \boldsymbol{I} = \bigcup \boldsymbol{I}_k$ where $\boldsymbol{I}_k = \{I_{n\|\gamma\|y} | n \| \gamma \| y \in S_k\}$, and $I_{n\|\gamma\|y}$ is the class-index function, i.e., $I_{n\|\gamma\|y}(x) = i$ iff $x = y^i \cdot w^\gamma \bmod n$ for some $w \in Z_n^*$. (Note: We have simplified the definitions for S_k and $\Omega_{n\|\gamma\|y}$. Rigorous definitions are troublesome, and lacking in readability: First an invertible function J mapping $n \| \gamma \| y$ into a positive integer should be introduced. S_k should be defined as $S_k = \{J(n \| \gamma \| y) | \cdots\}$, and instead of $\Omega_{n\|\gamma\|y} = Z_n^*$, we should have $\Omega_{J(n\|\gamma\|y)} = \{$binary representation of $x | x \in Z_n^*\}$.)

It is easy to see that $\boldsymbol{I}$ is trapdoor, i. e., it satisfies all the encryptability, decryptability and constructability conditions. In Sect. 5 we will prove that $\boldsymbol{I}$ is also unapproximable under γ^{th}-Residuosity Assumption.

4. Probabilistic Cryptosystem $\mathcal{C}_{UTF}$

Let L be a polynomial in k. And let $\boldsymbol{Y} = \bigcup \boldsymbol{Y}_k$ be an unapproximable trapdoor function, $M_k = \{m_l \| m_{l-1} \| \cdots \| m_1 | m_i \in \Sigma, l = L(k)\}$ be the message space

associated with k. Along the same line as Ref.(4), we obtain from $\boldsymbol{Y}$ a probabilistic public-key cryptosystem which we call $\mathcal{C}_{UTF}$.

On input $k\in K$, $\mathcal{C}_{UTF}$ outputs a randomly and uniformly selected element $n\in S_k$, its secret trapdoor $\delta(n)$, and descriptions of encryption/decryption algorithms $E_{UTF,n}$ and $D_{UTF,n}$ specified below.

ENCRYPTION ALGORITHM $E_{UTF,n}(m)$

For $i=1$ to l, select randomly and uniformly a c_i from Ω_n such that $Y_n(c_i)=m_i$. Output $c=c_l\|c_{l-1}\|\cdots\|c_1$ as an encryption of $m=m_l\|m_{l-1}\|\cdots\|m_1$.

DECRYPTION ALGORITHM $D_{UTF,n}(\delta(n), c)$

For $i=1$ to l, determine the function value $Y_n(c_i)=m_i\in\Omega_n$ by the use of the secret trapdoor $\delta(n)$. Output $m=m_l\|m_{l-1}\|\cdots\|m_1\in M_k$.

When $\Sigma=\{0, 1\}$, i. e., when $\boldsymbol{Y}$ is an unapproximable trapdoor predicate, polynomial security of $\mathcal{C}_{UTF}$ was proved by Goldwasser and Micali (see Ref.(4), Theorem 5.1). They considered furthermore an implementation of $\mathcal{C}_{UTF}$ by a predicate $\boldsymbol{B}=\cup\boldsymbol{B}_k$ (see Ref.(4) for details). Especially, they showed that under Quadratic Residuosity Assumption, $\boldsymbol{B}$ is an unapproximable trapdoor predicate, from which it follows that the concrete implementation of $\mathcal{C}_{UTF}$ by the predicate $\boldsymbol{B}$ is indeed polynomially secure.

Our generalized probabilistic cryptosystem is a concrete implementation of $\mathcal{C}_{UTF}$ by the set of class-index functions $\boldsymbol{I}=\cup\boldsymbol{I}_k$ defined in Sect. 3. We now prove the polynomial security of this cryptosystem by establishing a series of theorems (Theorems 1, 2, 3 and 4).

5. Unapproximable Trapdoorness of $\boldsymbol{I}=\cup\boldsymbol{I}_k$

Theorems 1 and 2 to be proved below are related to the unapproximable trapdoorness of $\boldsymbol{I}=\cup\boldsymbol{I}_k$.

First we recall a classic but invaluable proposition in Probability Theory—the Weak Law of Large Numbers (WLLN)[5]:

[Proposition 1] Let α be an event which occurs with probability p in a given experiment. Let $X_1, X_2, \cdots, X_n$ be a sequence of random variables defined as

$$X_i=\begin{cases}1, & \text{if } \alpha \text{ occurs at the } i\text{-th trial};\\ 0, & \text{otherwise.}\end{cases}$$

Also let $\hat{X}_n=(X_1+X_2+\cdots+X_n)/n$. Then for any $\varepsilon>0$,

$$\Pr\{|\hat{X}_n-p|<\varepsilon\}\geq 1-\frac{1}{4n\varepsilon^2}\longrightarrow 1, \quad \text{as } n\to\infty. \quad \square$$

The WLLN is the basis of the extensively applied Monte Carlo experiment[5].

We are ready to prove Theorem 1. This theorem says that even a small amount of stochastic advantage in correctly guessing the residuosity of elements in Z_n^* can be amplified, so that the residuosity of any element in Z_n^* can be efficiently decided with probability arbitrarily close to 1.

Though one can view Theorem 1 as the generalization of Theorem 5.1 of Ref.(4), it seems difficult to apply the technique for proving Theorem 5.1 of Ref.(4) to the proof of Theorem 1, since here one has to handle as many as $\gamma\geq 3$ different cases. The technique we use to prove Theorem 1 differs quite from that for proving Theorem 5.1 of Ref.(4), and in designing it we obtain some hints from the template-matching method, which is an intuitive pattern recognition technique and is well explained in Ref.(3).

[Theorem 1] Let Q_1 and Q_2 be any polynomials, and let H_k^γ be the hard integer set where $\gamma=O(k^c)$. Suppose there exists a circuit $C_k[\cdot,\cdot,\cdot]$ such that $C_k[x, y, n]=I_{n\|\gamma\|y}(x)$ for any $n\in H_k^\gamma(C_k)$, for any $y\in Z_n^*$ with (n, γ, y) an acceptable triple, and for a fraction greater than $1/\gamma+1/Q_1(k)$ of the $x\in Z_n^*$, where $H_k^\gamma(C_k)$ is a subset of H_k^γ. Then we can obtain a probabilistic polynomial time algorithm A_{RP} which, by taking C_k as an oracle, finds with probability greater than $1-1/Q_2(k)$, the class-index of z with respect to (n, γ, y), for any $n\in H_k^\gamma(C_k)$, any $y\in Z_n^*$ with (n, γ, y) an acceptable triple, and any $z\in Z_n^*$.

(Proof) The proof can be roughly divided into two stages. In the first stage, we construct two matrices $\hat{W}$ and V, where $\hat{W}$ depends on a triple (n, γ, y) and V depends not only on (n, γ, y) but also on an element $z\in Z_n^*$ whose class-index is to be determined. In the second stage, we compare V with $\hat{W}^s$, where $\hat{W}^s$ is obtained by shifting $\hat{W}$ downward and cyclically for s times, and examine whether or not they are close to each other. Here the template-matching method is used. If we find that for some $s\in Z_\gamma$, V and $\hat{W}^s$ are quite close to each other, we decide that this s is the class-index of z. Error probability introduced by such a decision procedure will be proved to be negligible.

Now we describe the proof in detail. For any $n\in H_k^\gamma$, and any acceptable triple (n, γ, y), define $R^i=\{x|x\equiv y^i\cdot w^\gamma(\mathrm{mod}\ n),\ x\in Z_n^*, w\in Z_n^*\}$. Thus $x\in R^i$ is equivalent to $I_{n\|\gamma\|y}(x)=i$. Also let $\varepsilon=1/Q_1(x)$ and $\sigma=1/Q_2(k)$.

Assume for simplicity that Z_n^* is uniformly distributed. (Discussions made below can be immediately generalized to the case when Z_n^* is arbitrarily distributed). Then for any $n\in H_k^\gamma(C_k)$, and any randomly and uniformly selected $x\in Z_n^*$, we have $\Pr\{C_k[x, y, n]=I_{n\|\gamma\|y}(x)\}>1/\gamma+\varepsilon$.

Now for each $(i, j)\in Z_\gamma\times Z_\gamma$, let $w_{i,j}=\Pr\{C_k[x, y, n]=j|x\in R^i\}$, i. e., $w_{i,j}$ is the probability with which C_k outputs j on input a randomly selected element $x\in R^i$. Thus we have the following stochastic matrix:

$$W=\begin{bmatrix} w_{0,0} & w_{0,1} & \cdots & w_{0,\gamma-1}\\ w_{1,0} & w_{1,1} & \cdots & w_{1,\gamma-1}\\ \vdots & \vdots & \vdots & \vdots\\ w_{\gamma-1,0} & w_{\gamma-1,1} & \cdots & w_{\gamma-1,\gamma-1}\end{bmatrix}$$

Three obvious properties of the matrix are:

P1: For each $i\in Z_\gamma$, $\sum_{j\in Z_\gamma} w_{i,j}=1$. And hence, $\sum_{i\in Z_\gamma}\sum_{j\in Z_\gamma} w_{i,j}=\sum_{j\in Z_\gamma}\sum_{i\in Z_\gamma} w_{i,j}=\gamma$.

P2: $\sum_{i\in Z_\gamma} w_{i,i}>1+\gamma\varepsilon$.

P3: For any $i\in Z_\gamma$, there must exist at least one $j\in Z_\gamma$ with $j\neq i$ such that $|w_{i,j}-w_{j,j}|>\gamma\varepsilon/(\gamma-1)$. This implies that for each row i, there is at least one row j which is very different from the row i.

Proof for the property P3: Suppose that the property is not true. Then for any $i\in Z_\gamma$,

$$\begin{aligned}\sum_{j\in Z_\gamma} w_{j,j}-\sum_{j\in Z_\gamma} w_{i,j}&=\sum_{j\in Z_\gamma}(w_{j,j}-w_{i,j})\\&=\sum_{j\in Z_\gamma, j\neq i}(w_{j,j}-w_{i,j})\\&\le\sum_{j\in Z_\gamma, j\neq i}|w_{i,j}-w_{j,j}|\\&\le(\gamma-1)\cdot\max\{|w_{i,j}-w_{j,j}||j\in Z_\gamma,\ j\neq i\}\\&\le(\gamma-1)\cdot\frac{\gamma\varepsilon}{\gamma-1}=\gamma\varepsilon.\end{aligned}$$

On the other hand, from the properties P1 and P2, $\sum_{j\in Z_\gamma} w_{j,j}-\sum_{j\in Z_\gamma} w_{i,j}>\gamma\varepsilon$.

Let $m=3[\log_2\gamma]+1$, and $t=\gamma^{3m+3}/4\varepsilon^2\sigma=\gamma^{3m+3}Q_1^2(k)Q_2(k)/4$. Now we estimate the probability $w_{i,j}$, for each $(i,j)\in Z_\gamma\times Z_\gamma$, by the following Monte Carlo experiment.

ENTIMATING-1:

For each row i, where $i\in Z_\gamma$, do the following steps.

0. Set $Counter_{i,0}:=0,\cdots,Counter_{i,(\gamma-1)}:=0$.
1. Repeat for t times: Select randomly and uniformly a $u\in Z_n^*$, form $x_i:=y^i\cdot u^\gamma \bmod n$, and query the oracle circuit C_k with parameters x_i, y and n. Increment $Counter_{i,j}$ by 1 iff $C_k[x_i, y, n]=j$.
2. For each $j\in Z_\gamma$, let $\hat{w}_{i,j}:=Counter_{i,j}/t$.

Denote by $\hat{W}$ the matrix $(\hat{w}_{i,j}|(i,j)\in Z_\gamma\times Z_\gamma)$. From the WLLN, we know that

$$\Pr\left\{|\hat{w}_{i,j}-w_{i,j}|<\frac{\varepsilon}{\gamma^{m+1}}\right\}\ge 1-\frac{1}{4t(\varepsilon/\gamma^{m+1})^2}=1-\frac{\sigma}{\gamma^{m+1}}.\qquad(**)$$

Consequently, $\hat{w}_{i,j}$, and hence $\hat{W}$ are very favorable estimates of $w_{i,j}$ and W, respectively.

Now suppose we are given an element $z\in Z_n^*$ whose class-index I_z should be determined, where $n\in H_k'(C_k)$.

From this element z, we construct a matrix $V=(v_{i,j}|(i,j)\in Z_\gamma\times Z_\gamma)$ as follows:

Replace $x_i:=y^i\cdot u^\gamma \bmod n$ in the step 1 of ESTIMATING-1 with $x_i:=z\cdot y^i\cdot u^\gamma \bmod n$ and, $\hat{w}_{i,j}$ in the step 2 of ESTIMATING-1 with $v_{i,j}$. Then do the newly designed Monte Carlo experiment.

Clearly, the class-index of $x_i:=z\cdot y^i\cdot u^\gamma \bmod n$ is $(I_z+i) \bmod \gamma$. For each $(i,j)\in Z_\gamma\times Z_\gamma$, $v_{i,j}$ is a satisfactory estimate of $w_{(I_z+i)\bmod\gamma,j}$, since by the WLLN, we have

$$\Pr\left\{|v_{i,j}-w_{(I_z+i)\bmod\gamma,j}|<\frac{\varepsilon}{\gamma^{m+1}}\right\}\ge 1-\frac{\sigma}{\gamma^{m+1}}.$$

Denote by $W^s(\hat{W}^s$, respectively) the matrix obtained by shifting $W(\hat{W}$, respectively), downward and cyclically, for s times, and by $w_{i,j}^s(\hat{w}_{i,j}^s$, respectively) the element of $W^s(\hat{W}^s$, respectively) at the position (i,j).

Now we know that both V and $\hat{W}^{I_z}$ are good estimates of W^{I_z}. So the two matrices V and $\hat{W}^{I_z}$ must be quite similar to each other. In fact, for each $(i,j)\in Z_\gamma\times Z_\gamma$ we have $\Pr\{|v_{i,j}-\hat{w}_{i,j}^{I_z}|<\varepsilon/\gamma^m\}>1-\sigma/\gamma^m$: Imagine that there is a circle with radius $\varepsilon/2\gamma^m$ and center $w_{i,j}^{I_z}$. Then the distance between points $v_{i,j}$ and $\hat{w}_{i,j}^{I_z}$ is less than ε/γ^m if both of the two points are in the circle. Thus

$$\begin{aligned}&\Pr\left\{|v_{i,j}-\hat{w}_{i,j}^{I_z}|<\frac{\varepsilon}{\gamma^m}\right\}\\&\ge\Pr\left\{|v_{i,j}-w_{i,j}^{I_z}|<\frac{\varepsilon}{2\gamma^m}\right\}\cdot\Pr\left\{|\hat{w}_{i,j}^{I_z}-w_{i,j}^{I_z}|<\frac{\varepsilon}{2\gamma^m}\right\}\\&>\Pr\left\{|v_{i,j}-w_{i,j}^{I_z}|<\frac{\varepsilon}{\gamma^{m+1}}\right\}\cdot\Pr\left\{|\hat{w}_{i,j}^{I_z}-w_{i,j}^{I_z}|<\frac{\varepsilon}{\gamma^{m+1}}\right\}\\&\ge\left(1-\frac{\sigma}{\gamma^{m+1}}\right)^2>1-\frac{\sigma}{\gamma^m}.\end{aligned}$$

This suggests to us that we can decide I_z by matching V, in turns, with $\hat{W}^s$ for $s=0$ to $\gamma-1$, and examining whether or not the distance between $\hat{W}^s$ and V, measured with a predetermined distance-measure $\mathcal{M}$ for $\{\hat{W}^s|s\in Z_\gamma\}\cup\{V\}$, is negligibly small. (Note: A distance-measure for a set X is a feasible function $\mathcal{M}: X\times X\to[0,+\infty)$ such that $\mathcal{M}(x,y)=0$ whenever $x=y$. $\mathcal{M}(x,y)$ is called the distance between x and y measured by $\mathcal{M}$.)

To estimate how correct the above matching procedure is, it is necessary that the distance-measure $\mathcal{M}$ satisfies the following:

CONDITION: The distance between $\hat{W}^s$ and V measured by $\mathcal{M}$ is smaller than some constant when $s=I_z$, and not smaller than the constant when $s\neq I_z$, both with probability greater than $(1-\sigma/\gamma)$.

Notice that $\gamma=O(k^c)$. So if there exists actually such a distance-measure, it can be quickly certified. The following three lemmas assure us there exists definitely at least one distance-measure such that CONDITION is satisfied.

[Lemma 1] There exists at least one distance-measure $\mathcal{M}$ such that for some positive real number $\chi>0$, and for some $(f,h)\in Z_\gamma\times Z_\gamma$, with $f\neq h$, the distance between W^f and W^h measured by $\mathcal{M}$, that is denoted by $\mathcal{M}(f,h)$ for convenience, is greater than χ.

(Proof) Assume for contradiction that there is no distance measure $\mathcal{M}$ such that for any $(f,h)\in Z_\gamma\times Z_\gamma$, for any $\chi>0$, the distance $\mathcal{M}(f,h)$ between W^f and W^h is greater than χ.

Then every W^s, $s\in Z_\gamma$, is identical to W. Hence by the property P2 we have

$$\sum_{i=0}^{\gamma-1} w^s_{i,i}=\sum_{i=0}^{\gamma-1} w_{i,i}>1+\gamma\varepsilon, \quad \text{for every } s\in Z_\gamma, \text{ and}$$

$$\sum_{s=0}^{\gamma-1}\sum_{i=0}^{\gamma-1} w^s_{i,i}=\sum_{s=0}^{\gamma-1}\sum_{i=0}^{\gamma-1} w_{(i-s)\bmod\gamma,i}=\sum_{j=0}^{\gamma-1}\sum_{i=0}^{\gamma-1} w_{j,i}$$

$$=\gamma\cdot\sum_{i=0}^{\gamma-1} w_{i,i}>\gamma+\gamma^2\varepsilon.$$

This contradicts the property P1. □

[Lemma 2] There exists at least one distance-measure $\mathscr{A}^*$ such that for some positive real number $\chi^*>0$, and for any $(s_1, s_2)\in Z_\gamma\times Z_\gamma$ with $s_1\neq s_2$, we have $\mathscr{A}^*(s_1, s_2)>\chi^*$.

(Proof) By Lemma 1, there exists a distance-measure $\mathscr{A}$ such that $\mathscr{A}(f, h)>\chi$ for some $(f, h)\in Z_\gamma\times Z_\gamma$ and for some positive real number χ.

Assume without loss of generality that $\mathscr{A}(i, j)\leq\chi$ for any $(i, j)\in Z_\gamma\times Z_\gamma$ with $(i, j)\neq(f, h)$. Let $g=\chi^*/\chi+1$, and call g an amplification factor.

From the distance-measure $\mathscr{A}$, we can obtain a new distance-measure $\hat{\mathscr{A}}$. With this new distance-measure, we measure the distance between two matrices W^i and W^j (both i and j are not necessarily known) as follows:

We first measure the distance between W^i and W^j with the distance-measure $\mathscr{A}$. If $\mathscr{A}(i, j)>\chi$, then we know that $(i, j)=(f, h)$, and we let $\hat{\mathscr{A}}(i, j)=g\cdot\mathscr{A}(i, j)$. Otherwise, from $s=0$ to $\gamma-1$ and $t=0$ to $\gamma-1$, we shift, downward and cyclically, the matrix W^i for s times, and the matrix W^j for t times. Every time the matrices are shifted, we measure the distance between the two newly gotten matrices $W^{(i+s)\bmod\gamma}$ and $W^{(j+t)\bmod\gamma}$. There exists certainly a pair $(s, t)\in Z^\gamma\times Z^\gamma$ such that the distance between $W^{(i+s)\bmod\gamma}$ and $W^{(j+t)\bmod\gamma}$ is greater than χ. Now let $\hat{\mathscr{A}}(i, j)=|s-t|\cdot g\cdot\chi$.

From the forgoing description of measuring procedure with $\hat{\mathscr{A}}$, we know that $\hat{\mathscr{A}}(i, i)=0$ for all $i\in Z_\gamma$ and $\hat{\mathscr{A}}(i, j)>\chi^*$ for all $(i, j)\in Z_\gamma\times Z_\gamma$ with $i\neq j$.

Let $\mathscr{A}^*$ be the above constructed distance-measure $\hat{\mathscr{A}}$, which concludes the proof. □

[Lemma 3] There exists at least one distance-measure $\mathscr{A}^*$ such that for $\chi^*\geq\varepsilon=1/Q_1(k)>0$, and for any $s_1, s_2\in Z_\gamma$ with $s_1\neq s_2$, we have $\mathscr{A}^*(s_1, s_2)>\chi^*$.

(Proof) From the property P3, we know that for at least one $s\in Z_\gamma$ and for at least one position (i, i), $|w_{i,i}-w^s_{i,i}|>(\gamma/(\gamma-1))\varepsilon$. Thus by Lemma 1, we can find a distance-measure $\mathscr{A}$ such that $\mathscr{A}(s, 0)>(\gamma/(\gamma-1))\varepsilon=\chi$. And by Lemma 2, we can furthermore construct a distance-measure $\mathscr{A}^*$ such that $\mathscr{A}^*(s, 0)>\chi^*\geq\chi=(\gamma/(\gamma-1))\varepsilon>\varepsilon=1/Q_1(k)$. □

By (**), we know that with probability greater than $1-\sigma/\gamma^{m+1}$, $\hat{w}^s_{i,j}$ is a satisfactory estimate of $w^s_{i,j}$ for any $s\in Z_\gamma$ and any $(i, j)\in Z_\gamma\times Z_\gamma$. Also we know that γ is of polynomial size. So from the above Lemma 3, we can find in polynomial time a distance-measure for $\{\hat{W}^s|s\in Z_\gamma\}\cup\{V\}$ such that it both satisfies CONDITION and corresponds, with probability greater than $1-\sigma/\gamma$, to one of the distance-measures $\mathscr{A}^*$ for $\{W^s|s\in Z_\gamma\}$ observed in the Lemma 3. We say that such a distance-measure for $\{\hat{W}^s|s\in Z_\gamma\}\cup\{V\}$ is truthful with probability greater than $1-\sigma/\gamma$.

Now we outline our algorithm A_{RP}.

OUTLINE OF ALGORITHM $A_{RP}(n, y, z)$:

1. Determine $\{\hat{W}^s|s\in Z_\gamma\}$ as follows:
   1.1 Query the circuit C_k with random elements in Z_n^* whose class-indices are known to us. Get an estimate $\hat{W}$ of the stochastic matrix W from the outputs of C_k.
   1.2 For each $s\in Z_\gamma$, obtain $\hat{W}^s$ by shifting $\hat{W}$ downward and cyclically for s times.
2. Construct from z a lot of random elements in Z_n^*, and query the circuit C_k with these elements. Get a matrix V from the outpts of C_k.
3. Find some distance-measure $\mathscr{A}$ for $\{\hat{W}^s|s\in Z_\gamma\}\cup\{V\}$ such that $\mathscr{A}$ satisfies CONDITION.
4. Match V with $\hat{W}^s$ for $s=0$ to $s=\gamma-1$. If the distance between $\hat{W}^s$ and V measured by $\mathscr{A}$ is less than some constant (say, $\leq\varepsilon/\gamma$), then judge that the class-index of z is equal to s.

Obviously A_{RP} runs in probabilistic polynomial time. The probability with which A_{RP} outputs correctly the class-index I_z of z is:

$$\Pr\{A_{RP}\text{ correct}\}\geq\Pr\{A_{RP}\text{ correct}|\mathscr{A}\text{ truthful}\}\cdot\Pr\{\mathscr{A}\text{ truthful}\}$$

$$\geq\left(1-\frac{\sigma}{\gamma}\right)\cdot\left(1-\frac{\sigma}{\gamma}\right)>1-\sigma.$$

This completes the proof of Theorem 1. □

[Theorem 2] (Under γ^{th}-Residuosity Assumption) When $\gamma=O(k^c)$, the set of class-index functions $\boldsymbol{I}=\cup\boldsymbol{I}_k$ defined in Sect. 3. is an unapproximable trapdoor function.

(Proof) The trapdoorness of $\boldsymbol{I}$ has been discussed in Sect. 3. Now we prove that $\boldsymbol{I}$ is unapproximable, by the use of Theorem 1.

Assume that there is an approximater $A=\{A_k|k\in K\}$ such that for two polynomials Q_1 and Q_2, for infinitely many $k\in K$, and for a fraction $1/Q_2(k)$ of $n\|\gamma\|y\in S_k$, such that A_k can $1/Q_1(k)$-approximates the function $I_{n\|\gamma\|y}: Z_n^*\to Z_\gamma$.

Denote by K' the infinite set of above k's, and by $S_k(A_k, 1/Q_1(k))$ the set of $n\|\gamma\|y\in S_k$ for which A_k can $1/Q_1(k)$-approximates the function $I_{n\|\gamma\|y}: Z_n^*\to Z_\gamma$.

For any $k\in K'$, for any polynomial Q_3, and for any $n\|\gamma\|y\in S_k(A_k, 1/Q_1(k))$, we can, by Theorem 1, construct a probabilistic polynomial time algorithm A'_k which finds the class-index of any $z\in Z_n^*$ with probability greater than $1-1/Q_3(k)$. From the algorithm A'_k, we can get a circuit C'_k of polynomial size[6] which, for a fraction $1/Q_2(k)$ of integers $n\in H'_k$, solves γ^{th}-RP with probability greater than $1-1/Q_3(k)$. This contradicts γ^{th}-Residuosity Assumption. □

6. Polynomial Security of C_{UTF}

In this section we prove that the cryptosystem C_{UTF} is polynomially secure when β is of polynomial size.

[Theorem 3] Let $\boldsymbol{Y}=\cup \boldsymbol{Y}_k$ be an unapproximable trapdoor function, where $\boldsymbol{Y}_k=\{Y_n|n\in S_k\}$, $Y_n: \Omega_n \to \Sigma=\{0, 1, \cdots, \beta-1\}$, and $\beta=O(k^c)$. The cryptosystem C_{UTF} constructed from $\boldsymbol{Y}$ is polynomially secure.

(Proof) We first sketch the proof as follows: Suppose that some F_k can find two messages $m_1^k, m_2^k\in M_k$ such that some T_k behaves quite differently about the two messages. From m_1^k and m_2^k we can find other two messages $b_f^k=b_l b_{l-1}\cdots b_h\cdots b_1$ and $b_{f+1}^k=b_l b_{l-1}\cdots b_h'\cdots b_1$ such that they differ only at the position h, and T_k behaves also quite differently about them. Next for each $x\in\Sigma$, we put x in the position h of b_f^k (or b_{f+1}^k), observe and record the behavior of T_k about the message $b_l b_{l-1}\cdots x\cdots b_1$. Thus, given any element $z\in\Omega_n$, we can judge which Ω_n^i the element z belongs to by the use of the records of the behaviors of T_k.

Let Q_1 and Q_2 be any polynomials, and let $\varepsilon_k=1/Q_1(k)$ and $\eta_k=1/Q_2(k)$. Suppose there is a message finder $F=\{F_k|k\in K\}$ such that for infinitely many $k\in K$, for some line tapper $T=\{T_k|k\in K\}$, for a fraction η_k of $n\in S_k$, F_k finds two messages $m_1^k, m_2^k\in M_k$ such that $|p(m_1^k)-p(m_2^k)|>\varepsilon_k$, where $p(m)$ is the probability with which T_k outputs 1 on input code $(E_{UTF,n})$ and $\alpha\in E(m)$.

Denote by K' the set of the above infinitely many k's, and by $S_k(F_k, 1)$ the set of elements $n\in S_k$ for which F_k finds two messages $m_1^k, m_2^k\in M_k$ such that $|p(m_1^k)-p(m_2^k)|>\varepsilon_k$.

We now introduce two other polynomials Q_3 and Q_4. It is assumed that $Q_3(k)>Q_2(k)/(1-\theta)$ for some constant $0<\theta<1$ and for all $k\in K$. Let $\delta_k=1/Q_4(k)$.

From the discussion made in Sect. 3, we know that for all sufficiently large $k\in K$, for a fraction greater than $1-1/Q_3(k)$ of the n in S_k, and for all $i\in\Sigma$, $|\#\Omega_n^i/\#\Omega_n-1/\beta|<\zeta_k$ where $\zeta_k=\delta_k/\beta^2 l$. Consequently, for all sufficiently large $k\in K'$, for some line tapper T, for a fraction greater than $f(k)$ of $n\in S_k$, we have the following two things:

(1) $|\#\Omega_n^i/\#\Omega_n-1/\beta|<\zeta_k$ for all $i\in\Sigma$, and

(2) F_k finds two messages $m_1^k, m_2^k\in M_k$ such that $|p(m_1^k)-p(m_2^k)|>\varepsilon_k$,

where $f(k)=1/Q_2(k)-1/Q_3(k)>1/Q_2(k)-(1-\theta)/Q_2(k)=\theta\cdot\eta_k$, and it indicates the lower-bound on the fraction of $n\in S_k$ such that the two things hold. Denote by K'' the set of the above infinitely many $k\in K'$, and by $S_k(F_k, 2)$ the set of the above elements $n\in S_k(F_k, 1)$. (Notice that $\#S_k(F_k, 2)/\#S_k>\theta\cdot\#S_k(F_k, 1)/\#S_k>\theta\cdot\eta_k$.)

Now we show that, for any $k\in K''$, a probabilistic polynomial time algorithm A_{UTF}^k can be constructed from F_k and T_k such that A_{UTF}^k can $1/2\beta$-approximate Y_n for a fraction greater than $\theta\cdot\eta_k$ of the $n\in S_k$.

Consider a $k\in K''$ and an $n\in S_k(F_k, 2)$. Let $x=x_l\|x_{l-1}\|\cdots\|x_1$ and $y=y_l\|y_{l-1}\|\cdots\|y_1$ be two sequences over $\Sigma=\{0, 1, \cdots, \beta-1\}$. The (Hamming) distance between x and y is the number of positions where x and y are different. We say x and y are adjacent if the distance between them is one.

Recall that $m_1^k, m_2^k\in M_k$ are two messages generated by F_k such that for T_k, $|p(m_1^k)-p(m_2^k)|>\varepsilon_k$. Suppose the distance between them is $0<d<l$.

Let $b_{d-1}^k, b_{d-2}^k, \cdots, b_0^k$ be any d sequences of length l, where $b_0^k=m_1^k$, $b_{d-1}^k=m_2^k$, and b_{j+1}^k is adjacent to b_j^k for all $0<j<d-1$. Corresponding to these d sequences, there are d probabilities: $p(b_{d-1}^k), p(b_{d-2}^k), \cdots, p(b_0^k)$.

There must exist at least one $0\le f<d-1$ such that $|p(b_f^k)-p(b_{f+1}^k)|>\varepsilon_k/l$, and that such an f can be easily found. This can be explained by elementary mathematics: We know that

$$|p(m_1^k)-p(m_2^k)|=|p(b_0^k)-p(b_{d-1}^k)|\le\sum_{i=0}^{d-2}|p(b_i^k)-p(b_{i+1}^k)|$$
$$\le(d-1)\cdot\max\{|p(b_i^k)-p(b_{i+1}^k)||i=0, 1, \cdots, d-2\}.$$

Thus if there did not exist any $0\le f<d-1$ such that $|p(b_f^k)-p(b_{f+1}^k)|>\varepsilon_k/l$, we would have $|p(m_1^k)-p(m_2^k)|\le(d-1)\cdot\varepsilon_k/l<\varepsilon_k$. This is not true.

Suppose we have found such an f, and suppose

$$b_f^k=b_l b_{l-1}\cdots b_h\cdots b_1,$$

$$b_{f+1}^k=b_l b_{l-1}\cdots b_h'\cdots b_1,$$

where $b_h, b_h'\in\Sigma$, and $b_h\neq b_h'$. For notational simplicity, let $\xi=b_h$, $\pi=b_h'$, and let $b(x)$ be the sequence obtained by replacing the letter b_h in the sequence b_f^k with the letter $x\in\Sigma$. Also let $p_x=p(b(x))$, and assume without loss of generality that $p_\xi>p_\pi$.

Let $P=\{p_0, p_1, \cdots, p_{\beta-1}\}$. Because that $p_\xi-p_\pi>\varepsilon_k/l$, and that $\#P$ is finite $(=\beta)$, there must exist at least one real number τ, called a threshold, with $p_\pi<\tau<p_\xi$ such that the set Σ can be divided into two non-empty sets

$$\Sigma_a^\tau=\left\{i|p_i>\tau+\frac{\varepsilon_k}{2\beta l}\right\} \quad\text{and}\quad \Sigma_b^\tau=\left\{i|p_i<\tau-\frac{\varepsilon_k}{2\beta l}\right\}.$$

The reason is as follows:

Let $p_{m_{h-1}}\ge p_{m_{h-2}}\ge\cdots\ge p_{m_1}$ be probabilities in P whose sizes are between p_ξ and p_π, i.e., $p_\xi>p_{m_i}>p_\pi$ for all $1\le i<h$, where $h\le\beta-2$. Let $p_{m_h}=p_\xi$ and $p_{m_0}=p_\pi$. Then we have

$$\max\{p_{m_{i+1}}-p_{m_i}|0\le i\le h\}\ge\frac{(p_\xi-p_\pi)}{h}>\frac{\varepsilon_k}{hl}>\frac{\varepsilon_k}{\beta l}.$$

Suppose for some $0\le i\le h$ we have indeed $p_{m_{i+1}}-p_{m_i}>\varepsilon_k/\beta l$, then a threshold τ is obtained by setting $\tau=p_{m_i}+(p_{m_{i+1}}-p_{m_i})/2=(p_{m_{i+1}}+p_{m_i})/2$.

Assume, for simplicity once again, that Ω_n and Ω_n^i are uniformly distributed for all $k\in K$ and $n\in S_k$. Let $t=\beta^6 l^6/\delta_k\varepsilon_k^2=\beta^6 l^6 Q_1^2(k)Q_4(k)$. Now we estimate p_i for all $i\in\Sigma$ by the following Monte Carlo experiment.

ENTIMATING-2:
For any $i\in\Sigma$, do the following three steps.
(0) Set $Counter_i := 0$.
(1) Repeat for t times: By means of $E_{UTF,n}$, form a random and uniform encryption $a\in E_{UTF,n}(b(i))$ of the sequence $b(i)$, and input the pair (a, code ($E_{UTF,n}$)) to the circuit T_k. Increment $Counter_i$ by 1 iff the output of T_k is 1.
(2) Let $p'_i=Counter_i/t$ as an estimate of p_i.

By the WLLN, we have

$$\Pr\left\{|p'_i-p_i|<\frac{\varepsilon_k}{2\beta^2l^2}\right\}\geq 1-\frac{1}{4t(\varepsilon_k/2\beta^2l^2)^2}\geq 1-\frac{\delta_k}{\beta^2l^2}.$$

Next we show that $p'_\xi-p'_\pi>\varepsilon_k/l-\varepsilon_k/\beta^2l^2$ with high probability. Let c_ξ and c_π be two circles with radii $\varepsilon_k/2\beta^2l^2$ and centers p_ξ and p_π respectively. Recall that the distance between p_ξ and p_π is greater than ε_k/l. So the distance between p'_ξ and p'_π is greater than $\varepsilon_k/l-\varepsilon_k/\beta^2l^2$ if p'_ξ is in the circle c_ξ and p'_π is in the circle c_π. Consequently,

$$\begin{aligned}&\Pr\left\{p'_\xi-p'_\pi>\frac{\varepsilon_k}{l}-\frac{\varepsilon_k}{\beta^2l^2}\right\}\\&\geq\Pr\left\{|p'_\xi-p_\xi|<\frac{\varepsilon_k}{2\beta^2l^2}\right\}\cdot\Pr\left\{|p'_\pi-p_\pi|<\frac{\varepsilon_k}{2\beta^2l^2}\right\}\\&\geq\left(1-\frac{\delta_k}{\beta^2l^2}\right)^2\geq 1-\frac{\delta_k}{\beta l}.\end{aligned}$$

Let $P'=\{p'_0, p'_1\cdots, p'_{\beta-1}\}$. Notice that neither the summation of $p_0, p_1, \cdots, p_{\beta-1}$, nor that of $p'_0, p'_1, \cdots, p'_{\beta-1}$ is necessarily equal to 1.

Since $p'_\xi-p'_\pi>\varepsilon_k/l-\varepsilon_k/\beta^2l^2$ with probability $1-\delta_k/\beta l$, and $\#P'$ is finite ($=\beta$), we can find, with the same probability $1-\delta_k/\beta l$, a threshold τ' with

$$p'_\pi+\frac{\varepsilon_k}{2\beta^2l^2}<\tau'<p'_\xi-\frac{\varepsilon_k}{2\beta^2l^2}$$

such that the set Σ can be divided into two non-empty sets

$$\Sigma_a^{\tau'}=\left\{i|p'_i>\tau'+\frac{\varepsilon_k}{2\beta^2l^2}\right\},\ \Sigma_b^{\tau'}=\left\{i|p'_i<\tau'-\frac{\varepsilon_k}{2\beta^2l^2}\right\}.$$

Let's estimate the probability with which the pair $(\Sigma_a^{\tau'}, \Sigma_b^{\tau'})$ is equal to one of the above mentioned pairs $(\Sigma_a^{\tau}, \Sigma_b^{\tau})$. First we notice that if $(\Sigma_a^{\tau'}, \Sigma_b^{\tau'})$ does not coincide with any of $(\Sigma_a^{\tau}, \Sigma_b^{\tau})$, then $|p'_i-p_i|\geq\varepsilon_k/2\beta^2l^2$ for at least one $i\in\Sigma$. Therefore

$$\begin{aligned}&\Pr\{(\Sigma_a^{\tau'}, \Sigma_b^{\tau'}\neq\text{any of }(\Sigma_a^{\tau}, \Sigma_b^{\tau})\}\\&\leq\sum_{i=0}^{\beta-1}\Pr\left\{|p'_i-p_i|\geq\frac{\varepsilon_k}{2\beta^2l^2}\right\}\\&<\sum_{i=0}^{\beta-1}\frac{\delta_k}{\beta^2l^2}=\frac{\delta_k}{\beta l^2}<\frac{\delta_k}{\beta l},\end{aligned}$$

and hence

$$\begin{aligned}&\Pr\{(\Sigma_a^{\tau'}, \Sigma_b^{\tau'}=\text{one of }(\Sigma_a^{\tau}, \Sigma_b^{\tau})\}\\&=1-\Pr\{(\Sigma_a^{\tau'}, \Sigma_b^{\tau'})\neq\text{any of }(\Sigma_a^{\tau}, \Sigma_b^{\tau})\}>1-\frac{\delta_k}{\beta l}.\end{aligned}$$

Now we approximate $Y_n(z)$ for any $z\in\Omega_n$, where $n\in S_k(F_k)$ using the following algorithm *APPRX*.

ALGORITHM $APPRX(n, z)$
0. Let $Counter := 0$.
1. Repeat for t times: For every $j\in\{1, 2, \cdots, h-1, h+1, \cdots, l\}$, select randomly and uniformly an $x_j\in\Omega_n^{b_j}$ as an encryption of b_j. Then input the pair (x, code ($E_{UTF,n}$)) to the circuit T_k, where $x=x_l\|x_{l-1}\|\cdots\|x_{h+1}\|z\|x_{h-1}\|\cdots\|x_1$. Increment $Counter$ by 1 iff the output of T_k is 1.
2. Let $p' := Counter/t$.
3. Select, randomly and uniformly, an $i\in\Sigma_a^{\tau'}$ and a $j\in\Sigma_b^{\tau'}$, then approximate $Y_n(z)$ as follows:

$$Y_n(z)=\begin{cases}i, & \text{if } p'>\tau'+\dfrac{\varepsilon_k}{2\beta^2l^2};\\ j, & \text{otherwise.}\end{cases}$$

We have completed the description of the algorithm A_{UTF}^k. Finally, we estimate the probability with which A_{UTF}^k outputs $Y_n(z)$ correctly.

Notice that

$$\Pr\{APPRX(n, z)=Y_n(z)|\,Y_n(z)\in\Sigma_a^{\tau'}\}=\frac{1}{\#\Sigma_a^{\tau'}}$$

and

$$\Pr\{APPRX(n, z)=Y_n(z)|\,Y_n(z)\in\Sigma_b^{\tau'}\}=\frac{1}{\#\Sigma_b^{\tau'}}$$

respectively. Thus we get

$$\begin{aligned}&\Pr\{A_{UTF}^k(n, z)=Y_n(z)\}\\&\geq\Pr\{APPRX(n, z)=Y_n(z)|(\Sigma_a^{\tau'}, \Sigma_b^{\tau'})\\&\quad=\text{one of }(\Sigma_a^{\tau}, \Sigma_b^{\tau})\}\cdot\Pr\{(\Sigma_a^{\tau'}, \Sigma_b^{\tau'})=\text{one of }(\Sigma_a^{\tau}, \Sigma_b^{\tau})\}\\&>\left(1-\frac{\delta_k}{\beta l}\right)\cdot[\Pr\{APPRX(n, z)\\&\quad=Y_n(z)|\,Y_n(z)\in\Sigma_a^{\tau'}\}\cdot\Pr\{Y_n(z)\in\Sigma_a^{\tau'}\}\\&\quad+\Pr\{APPRX(n, z)=Y_n(z)|\,Y_n(z)\in\Sigma_b^{\tau'}\}\\&\quad\cdot\Pr\{Y_n(z)\in\Sigma_b^{\tau'}\}]\\&=\left(1-\frac{\delta_k}{\beta l}\right)\cdot[\Pr^1+\Pr^2],\end{aligned}$$

Where

$$\begin{aligned}\Pr^1&=\Pr\{APPRX(n, z)=Y_n(z)|\,Y_n(z)\\&\quad\in\Sigma_a^{\tau'}\}\cdot\Pr\{Y_n(z)\in\Sigma_a^{\tau'}\}\\&\geq\frac{1}{\#\Sigma_a^{\tau'}}\cdot\#\Sigma_a^{\tau'}\cdot\left(\frac{1}{\beta}-\zeta_k\right)=\frac{1}{\beta}-\zeta_k,\end{aligned}$$

and similarly

$$\begin{aligned}\Pr^2&=\Pr\{APPRX(n, z)=Y_n(z)|\,Y_n(z)\\&\quad\in\Sigma_b^{\tau'}\}\cdot\Pr\{Y_n(z)\in\Sigma_b^{\tau'}\}\geq\frac{1}{\beta}-\zeta_k.\end{aligned}$$

So, we have

$$\Pr\{A^k_{UTF}(n, z)= Y_n(z)\} > \left(1-\frac{\delta_k}{\beta l}\right)\cdot[\Pr^1+\Pr^2]$$

$$\geq\left(1-\frac{\delta_k}{\beta l}\right)\cdot\left(\frac{2}{\beta}-2\zeta_k\right)=\frac{1}{\beta}+\left(\frac{1}{\beta}-\frac{2\delta_k}{\beta^2 l}-\frac{2\delta_k}{\beta^2 l}\right.$$

$$\left.+\frac{2\zeta_k\delta_k}{\beta l}\right)=\frac{1}{\beta}+\left(\frac{1}{\beta}-\frac{4\delta_k}{\beta^2 l}+\frac{2\delta_k^2}{\beta^3 l^2}\right)\geq\frac{1}{\beta}+\frac{1}{2\beta}.$$

For all $k\in K''$, we can construct from A^k_{UTF} a polynomial size circuit C^k_{UTF} (see for example Ref.(6)) such that this circuit can $1/2\beta$-approximates Y_n for all $n\in S_k(F_k, 2)$, i. e., for a fraction greater than $\theta\cdot\eta_k$ of n in S_k, where θ is an arbitrary but fixed real number within the range $(0, 1)$. This contradicts the unapproximability of $\boldsymbol{Y}$, and the proof for Theorem 3 is completed. □

[Theorem 4] (Under γ^{th}-Residuosity Assumption) The probabilistic cryptosystem based on γ^{th}-RP is polynomially secure.

(Proof) It follows from Theorems 2 and 3. □

7. Concluding Remarks

In proving Theorem 1, we benefited a lot from fully-developed pattern-matching techniques. Partitioning Σ into two non-empty sets $\Sigma^{r'}_a$ and $\Sigma^{r'}_b$ is the key point of the proof for Theorem 3. This proof technique, however, seems not directly generalizable to the case when β is not of polynomial size. It would be nice to show that the cryptosystem C_{UTF} is also polynomially secure even when β is exponentially large.

References

(1) W. Alexi, B. Chor, O. Goldreich and C. P. Schnorr: "RSA and Rabin functions: certain parts are as hard as the whole", SIAM Journal on Computing, **17**, 2, pp. 194-209 (1988).

(2) J. Benaloh and M. Yung: "Distributing the power of a government to enhance the privacy of voters", Proc. of the 5th ACM Symposium on Principles of Distributed Computing, pp. 52-62 (1986).

(3) K. S. Fu: "Recent developments in pattern recognition", IEEE Trans. Comput., **C-29**, 10, pp. 845-854 (1980).

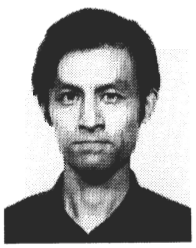
(4) S. Goldwasser and S. Micali: "Probabilistic encryption", J. Comput. & Syst. Sci., **28**, pp. 270-299 (1984).

(5) R. Y. Rubinstein: "Simulation and the Monte Carlo method", John Wiley & Sons, Inc. (1981).

(6) I. Wegener: "The complexity of Boolean functions", Wiley-Teubner Series in Computer Science, John Wiley & Sons, Inc. (1987).

(7) Y. Zheng, T. Matsumoto and H. Imai: "Residuosity problem and its applications to cryptography", Trans. IEICE, **E71**, 8, pp. 759-767 (Aug. 1988).

Yuliang Zheng was born in Jiangsu, China on February 5, 1962. He received the B. S. degree in computer science from Southeast University (formerly Nanjing Institute of Technology), Nanjing, China, in 1982, and the M. E. degree in computer engineering from Yokohama National University, Yokohama, Japan, in 1988. From 1982 to 1984, he was with Guangzhou Research Institute for Communications, Guangzhou, China. He is currently pursuing the Ph. D. degree under the supervision of Professor Hideki Imai. His research interests include cryptography, computational complexity theory and information theory. He is a student member of IEEE.



Tsutomu Matsumoto was born in Maebashi, Japan, on October 20, 1958. He received the B. Eng. and M. Eng. degrees in computer eng. both from Yokohama National University, Yokohama, Japan, in 1981 and 1983, respectively, and the D. Eng. degree in electronic eng. from the University of Tokyo, Tokyo, Japan, in 1986. Since 1986, he has been Lecturer for Electrical and Computer Engineering at Yokohama National University. He is currently working in cryptography, complexity theory, computational mathematics, and their applications to information security. Dr. Matsumoto is a member of ACM, IACR, IEEE, IPSJ, ITA, and Akarui Angou Kenkyu-kai.

Hideki Imai was born in Shimane, Japan on May 31, 1943. He received the B. E., M. E. and Ph. D. degrees in electrical engineering from University of Tokyo, Tokyo, in 1966, 1968, and 1971, respectively. He is currently a Professor in the Division of Electrical and Computer Engineering, Yokohama National University, Yokohama. His current research interests include information theory, coding theory, cryptography, and their applications. He is the author of two books and coauthor of several books. Dr. Imai is a member of IEEE, IEE of Japan, IPS of Japan, and ITE of Japan.