情報保護學會論文誌
第11券 第3號, 2001. 6

땋임군에서의 안전성이 증명 가능한 유사난수 생성기

이 언 경*, 한 상 근**

A Provably secure Pseudorandom generator from Braid groups

Eonkyung Lee*, Sang Geun Hahn**

요 약

유사난수성(pseudorandomness)은 진정한 난수성(true randomness)을 대신하여 실제 상황에서 사용되는 개념으로서 현대 암호학의 중요한 한 분야이다. 본 논문은 땋임이론(braid theory)에서의 어려운 문제 중 하나인 공액문제(conjugacy problem)에 기반하여 단순하고 실용적인 유사난수 생성기(pseudorandom generator)를 설계한다. 그리고 그 생성기가 암호학적으로 공액문제를 변형한 또 하나의 어려운 문제 만큼 안전함을 증명한다.

ABSTRACT

The notion of pseudorandomness plays an important role in modern cryptography as well as computer science. We show a simple and practical construction of a pseudorandom generator based on the intractability of the problem in braid groups. The generator is proved as secure as a hard instance of a variant of the conjugacy problem.

keyword : *Pseudorandom Generator, Provable Security, Braid Group, Conjugacy Problem*

Ⅰ. 서 론

난수(random number)는 암호학에서 뿐만 아니라 일반 공학에서도 꼭 필요한 요소이나, 난수를 효율적으로 생성하기 위해서는 지나치게 많은 비용이 들고, 실험자가 동일한 난수를 생성할 수 없어 실험을 그대로 반복할 수 없는 등의 단점 때문에, 유사난수(pseudorandom number)가 대신 사용된다. 유사난수 생성기(pseudorandom generator)는 가장 기본적인 유사난수 원시 요소(primitive)로서, 짧은 길이의 진정한 난수(true random number)를 유사난수(pseudorando mness) 분포를 갖는 긴 길이의 비트열(bit string)로 확장시키는 효율적인 알고리즘이다. 어떤 분포가 유사난수 분포를 따른다는 것은, 간단히 말해서, 이 분포와 균등 확률분포를 구별할 수 있는 효율적인 알고리즘이 존재하지 않는다는 의미이다[4]. 유사난수 생성기가 특히 보안 분야에서 사용되기 위해서는 안전성이 중요한데, 주어진 유사난수 생성기가 완전히 안전한(perfect secure)지를 수학적으로 증명하는 방법은 구체적으로 알려진 것이 없다. 대신 몇 가지 통계적 검증을 통과한 것을 현실에서 사용하고 있다. 그러나 기존의 검증을 통과했다고 해서 새로운 검증을 통과하리란 보장은 없기 때문에 이들이 암호학적으로 안전하다고 할 수는 없다. 현존하는 모든 통계적 검증 뿐 만이 아니라, 현실적인 미지의 새로운 모든 검증도 통과

* 본 연구는 과학기술부 국가지정연구실 '99 과제 지원 하에 수행되었습니다.
* 한국과학기술원 수학과 (eklee@math.kaist.ac.kr)
** 한국과학기술원 수학과 (sghahn@mathx.kaist.ac.kr)

하는 유사난수 생성기를 안전성이 증명 가능(provably secure)하다고 하며, 최근 20 여 년 동안 이에 대한 연구가 활발히 진행되어왔다. 몇 가지 예가 제안되어 왔는데, 이들의 안전성은 어렵다고 알려진 문제들의 난이도에 의존한다. 예를 들어, RSA 유사난수 생성기는 RSA 문제가 어렵다는 가정하에서 안전성이 증명 가능한 유사난수 생성기이다.

최근 땋임군(braid group)이 새로운 연구 분야로서 암호학계에 소개되었다[1~5]. 땋임군 B_n은 n개의 가닥으로 이루어진 기하학적인 땋임들로부터 자연스럽게 발생하는 무한 비가환군이다. 땋임이론에서 가장 유명한 문제 중 하나인 공액문제(conjugacy problem)는 다음과 같다.

Conjugacy problem: Given $(\alpha, \beta) \in B_n \times B_n$, find (or determine whether there exists) $\chi \in B_n$ such that $\beta = \chi^{-1}\alpha\chi$.

이 문제는 1920년대에 제시되었는데, $n \geq 5$에 대해서는 다항시간 해법 알고리즘이 현재 알려져 있지 않다.

공액 문제의 변형으로서, $\alpha, \chi_1^{-1}\alpha\chi_1, \chi_2^{-1}\alpha\chi_2 \in B_n$이 주어져 있을 때 $\chi_1^{-1}\chi_2^{-1}\alpha\chi_2\chi_1 \in B_n$을 구하는 문제가 공개키 암호시스템을 설계하기 위하여 제안되었다[5]. 여기에서, χ_1과 χ_2는 B_n의 서로 다른 부분군의 원소로서 $\chi_1\chi_2 = \chi_2\chi_1$이다. 앞으로 편의상 이 문제를 Ko-Lee problem이라 부르겠다. Ko-Lee problem의 난이도는, 공액문제를 좀 더 작은 땋임군으로 한정하는 다음의 (n,m)-generalized conjugacy problem(GCP)의 난이도에 의존한다: $(\alpha, \beta) \in B_n \times B_n$와 m $(\leq n)$이 주어져 있을 때 $\beta = \chi^{-1}\alpha\chi$을 만족하는 $\chi \in B_m$를 찾는다. 공액문제와 마찬가지로 GCP와 Ko-Lee problem에 대해서도 다항시간 해법 알고리즘이 현재 알려져 있지 않다.

우리는 Ko-Lee problem으로부터 안전성이 증명 가능한 유사난수 생성기를 만드는 방법에 대하여 논의하기로 한다. 여기에서 Ko-Lee problem이 어렵다는 가정을 KL-Assumption이라 부르자.

본 논문은 먼저 공개키 암호시스템의 안전성을 언급할 때 이미 소개된 바 있는 결정적 KL-Assumption (DKL-Assumption)[5]을 확률적으로 정의하고, 이 DKL-Assumption 하에서 유사난수 생성기를 설계하고 안전성이 증명 가능함을 보인다.

1.1 땋임군 개요

본 논문의 이해를 돕기 위해서 여기에서는 간단히, 땋임군에 대하여 살펴본다.

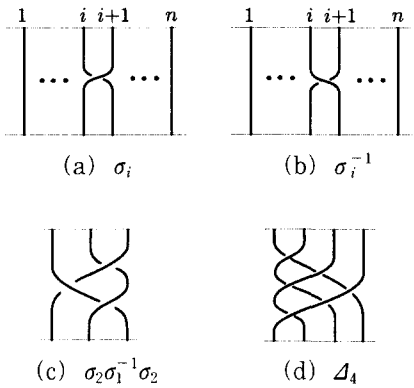
땋임군 B_n은 $\sigma_1, \ldots, \sigma_{n-1}$을 생성자로 하고 다음 관계식을 정의 관계(defining relation)로 하는 군표현(group presentation)을 가진다.

- $\sigma_i\sigma_j = \sigma_j\sigma_i$ if $|i-j| \geq 2, 1 \leq i, j \leq n-1$.
- $\sigma_i\sigma_{i+1}\sigma_i = \sigma_{i+1}\sigma_i\sigma_{i+1}$ if $1 \leq i, j \leq n-1$.

여기에서 자연수 n은 땋임지수(braid index), B_n의 원소는 n-땋임이라고 부른다. 기하학적으로 n-땋임은 동일한 방향으로 향하는 n개 가닥(strand)의 집합으로서 설명된다 (편의상 우리는 수직 방향으로 생각한다). 두 땋임 α와 β에 대하여 군연산인 곱 $\alpha\beta$는 β 위에 α를 놓음으로써 얻어진다. 땋임군에서의 항등원 e_n은 n개의 수직 가닥으로 구성된 땋임이며, 땋임 α의 역땋임 α^{-1}은 상하를 뒤집음으로써 얻어진다 (그림 1(a,b,c) 참조). 앞으로 σ_i들은 땋임군의 생성자를 표시하는 데에만 사용하기로 한다.

B_n의 생성자와 정의 관계로 정의되는 모노이드인 B_n^+은 생성자 σ_i들의 양의 지수를 갖는 땋임들로만 구성되며 그 원소들은 양의 n-땋임이라 부른다.

순열 $\pi = b_1 b_2 \cdots b_n$에 대하여 윗면의 i와 아랫면의 b_i를 직선으로 연결한 양의 n-땋임을 대응시키는 함수를 생각할 수 있는데, 이렇게 얻어진 n-땋임을 순열 π로부터 얻어진 순열땋임(permutation braid)이라 부른다. 순열 $(n)(n-1)\cdots(2)(1)$에 대응되는



(a) σ_i (b) σ_i^{-1}

(c) $\sigma_2\sigma_1^{-1}\sigma_2$ (d) Δ_4

[그림 1] 땋임들의 예

n-순열땋임을 특히 기본땋임(fundamental braid)라 부르며 Δ_n라 쓴다 (그림 1(d) 참조).

모든 땋임 $\chi \in B_n$은 $\chi = \Delta_n^u \chi_1 \cdots \chi_p$와 같은 형태로 유일하게 표현된다. 여기에서 u는 정수이고 χ_i들은 특정한 규칙을 만족하도록 배열된 e_n과 Δ_n을 제외한 순열땋임이다. 그리고, u와 p, $u+p$를 각각 χ의 infimum, canonical-length, supremum이라 부르며, $\inf(\chi)$, $\mathrm{len}(\chi)$, $\sup(\chi)$라 표기한다. 이렇게 χ를 분해한 것을 χ의 좌표준형(left canonical form)이라 부른다. 앞으로 본 논문에서 모든 땋임은 그들의 좌표준형으로 표현되었다고 간주한다. 따라서, $\alpha, \beta \in B_n$에 대하여 곱 $\alpha\beta$는 $\alpha\beta$의 좌표준형을 의미하므로 $\alpha\beta$로 부터 원래의 인자인 α 또는 β를 알아내기는 어렵다.

$m < n$에 대하여, B_m은 B_n의 생성자들 중 $\sigma_1, \ldots, \sigma_{m-1}$으로만 생성된 B_n의 부분군으로 간주될 수 있다. 따라서 $\Delta_m (\in B_n)$은 순열 $(m)(m-1)\cdots(2)(1)(m+1)(m+2)\cdots(n)$에 대응되는 n-순열땋임으로 간주될 수 있다.

땋임군의 주요 연산들(곱연산, 역연산, 좌표준형 변형연산)은 컴퓨터에 의해 효율적으로 계산된다[2~3].

1.2 기 호

1.2.1. 기본 기호

N은 자연수의 집합, Z는 정수의 집합, $F_{n,m}$은 $\{0,1\}^n \to \{0,1\}^m$인 모든 함수들의 집합을 의미한다. $\{0,1\}^*$는 모든 유한 길이의 비트열(bit string)들의 집합이다. $x \in \{0,1\}^*$에 대하여, $\|x\|$는 x의 비트열 길이를 의미한다. 유한 집합 S에 대하여, $|S|$는 S의 원소의 개수를 의미하며, $\|S\|$는 S의 원소들 중 가장 긴 길이를 가진 원소의 비트열 길이를 의미한다.

1.2.2. 확률적 기호

D가 유한집합 S상에서의 확률 분포라 하자. $[D]$는 D의 받침(support) (양의 확률을 갖는 표본들의 집합)을 의미하고, x가 S 상에서 확률 분포 D를 따르는 확률 변수일 때, $x \leftarrow_D S$라 표기한다. 문맥상 집합 S가 명백할 때, $x \leftarrow_D S$는 $x \leftarrow D$로 표기하기도 한다. $x_1, \ldots, x_n \leftarrow D$은 $x_1, \ldots, x_n$가 각각 독립적으로 D의 분포를 따르는 확률변수들임을 의미한다. $D, E, \ldots$가 확률분포일 때,

$$\Pr[p(x, y, \ldots) \mid x \leftarrow D;\ y \leftarrow E;\ \cdots]$$

는 $x \leftarrow D$; $y \leftarrow E$; $\cdots$을 차례로 실행시키고 난 후의 $p(x, y, \ldots)$이 성립할 확률을 의미한다. 확률 분포 U는 해당 집합 상에서의 균등 분포를 의미한다. 즉, 유한집합 S에 대하여, $x \leftarrow_U S$는 x가 집합 S 상에서 균등 분포를 갖는 확률변수임을 의미한다. A가 확률적 알고리즘일 때, 입력 $x, y, \ldots$에 대하여 $A(x, y, \ldots)$는 내부의 임의성(internal random coin tosses)에 의한 확률 분포를 의미한다.

다음은 두 확률 분포의 차이를 측정하는 개념에 관한 정의이다.

정의 1.1: D과 E은 S 상에서의 확률 분포라 하자. 이 때, D과 E의 통계적 거리(statistical distance)는 다음과 같이 정의된다.

$$\mathrm{dist}(D, E) = \frac{1}{2} \cdot \sum_{x \in S} |\Pr[x = x : x \leftarrow_D S] - \Pr[y = x : y \leftarrow_E S]|$$

D과 E이 최대 ϵ 통계적 거리를 갖는다는 것은 $\mathrm{dist}(D, E) \le \epsilon$임을 의미한다.

II. The DKL-Assumption

처음 제안되었을 당시의 KL-Assumption은 다음과 같다[5].

n-땋임 $\alpha, \chi^{-1}\alpha\chi, \phi^{-1}\alpha\phi$이 있을 때,
$\phi^{-1}\chi^{-1}\alpha\chi\phi$를 계산하는 것은 어렵다.

여기에서, $\chi \in \langle \sigma_1, \ldots, \sigma_{\lfloor n/2 \rfloor - 1} \rangle$이고 $\phi \in \langle \sigma_{\lfloor n/2 \rfloor + 1}, \ldots, \sigma_{n-1} \rangle$이다.

위 가정은 무한 집합 상에서 정의되어 있는데, 안전성이 증명 가능한 시스템을 설계하기 위해서는 우선 이 가정이 성립하는 유한 집합을 정의하여야 한다. KL-Assumption은 $(n, n/2)$-GCP의 난이도에 의존한다. $n = 2m$인 (n, m)-GCP에서 $(\alpha, \chi^{-1}\alpha\chi) \in B_n \times B_n$가 주어져 있을 때, $\inf(\chi) = 0$이고 $\mathrm{len}(\chi) = p$인 $\chi \in B_m$를 전수조사에 의해 찾기 위해서는 적어도

$$\left(\lfloor \frac{m-1}{2} \rfloor !\right)^p = \left(\lfloor \frac{n-2}{4} \rfloor !\right)^p$$

개의 m-땋임들을 조사해야 한다[5]. 따라서, 땋임지수 (n)와 좌표준형 길이 (p)를 안전도 매개변수(security parameter)로 간주한다.

다음은 유사난수 생성기를 설계하기 위한 가장 기본 단계인 색인(index)에 관한 정의이다.

정의 2.1:

- 모든 $n \in N$과 $i \le j \in Z$에 대하여, B_n의 부분집합 $[i,j]_n$를 $\{\chi \in B_n \mid \inf(\chi) \ge i,\ \sup(\chi)(\chi) \le j\}$라 하자.
- 색인 집합 I를 $\{(n,p) \mid n,p \in N\}$라 하자.
- 각 $k=(n,p) \in I$에 대하여, $I_{1,k} = \{\alpha \in B_n^+ - \Delta_n B_n^+ \mid \mathrm{len}(\alpha) = p\}$라 하자.
- IG는 입력 $k \in I$에 대하여 $I_{1,k}$의 원소를 출력하는 확률적 알고리즘이라 하자.

본 정의에서 $\Delta_n B_n^+$는 B_n^+의 정의에 의해서, $\{\alpha \in B_n^+ \mid \inf(\alpha) \ge 1\}$을 의미한다. 따라서, $I_{1,k} = \{\alpha \in B_n \mid \inf(\alpha) = 0,\ \mathrm{len}(\alpha) = p\}$이다.

$m(n) = \lfloor \frac{n}{2} \rfloor$라 하자. 편의상 $k=(n,p) \in I$와 $m = m(n)$를 고정시키고, 다음과 같이 단사 군 준동형사상 τ를 정의하자.

$$\tau : \begin{array}{ccc} B_{n-m} & \to & B_n \\ \sigma_i & \mapsto & \sigma_{m+i} \end{array} \quad \forall i.$$

따라서 $\tau(B_{n-m}) = \langle \sigma_{m+1}, \ldots, \sigma_{n-1} \rangle$은 B_n의 부분군으로서 B_{n-m}과 동형이 된다.

다음은 KL-Assumption에서 χ와 ψ의 유한 영역(finite domain)을 정의한다.

정의 2.2: 각 $k=(n,p) \in I$에 대하여 $LD_k = [-p,p]_m$, $RD_k = \tau([-p,p]_{n-m})$라 하고, 각 $\alpha \in I_{1,k}$에 대하여 $R_{k,\alpha} = \{\zeta^{-1}\alpha\zeta \mid \zeta \in [-p,p]_n\}$라 하자.

위에서 $m(n)$을 편의상 $\lfloor n/2 \rfloor$라 정하였으나, 대략 $\lfloor n/2 \rfloor$ 근처의 값이면 된다. LD_k와 RD_k의 정의로부터, 모든 $k=(n,p) \in I$와 $(\chi,\psi) \in LD_k \times RD_k$에 대하여 다음 두 가지 사실이 성립한다: (i) $\chi\psi = \psi\chi$, (ii) $\chi\psi \in [-p,p]_n$. (i)은 당연하고, (ii)는 $\Delta_n = \Delta_m \tau(\Delta_{n-m})\zeta$를 만족하는 $\zeta \in B_n^+$이 존재하다는 사실을 이용하여 증명한다. 위 정의들로 부터 DKL-Assumption은 다음과 같이 정의 된다.

〔The DKL-Assumption〕

모든 확률적 다항시간 알고리즘 A와, 양의 다항함수 $P(\cdot)$, 충분히 큰 모든 $k \in I$에 대하여 다음 식이 성립한다.

$$\begin{aligned} &|\Pr[A(\alpha, \chi^{-1}\alpha\chi, \psi^{-1}\alpha\psi, \psi^{-1}\chi^{-1}\alpha\chi\psi) = 1: \\ &\qquad \alpha \leftarrow IG(k); \chi \leftarrow_U LD_k; \psi \leftarrow_U RD_k] \\ &- \Pr[A(\alpha, \chi^{-1}\alpha\chi, \psi^{-1}\alpha\psi, \beta) = 1: \alpha \leftarrow IG(k); \\ &\qquad \chi \leftarrow_U LD_k; \psi \leftarrow_U RD_k; \beta \leftarrow_U R_{k,\alpha}]| \\ &< \frac{1}{P(k)}. \end{aligned}$$

앞으로 '큰 k'라고 할 때는, '큰 n과 큰 p'를 의미한다. 위에서 LD_k로부터 χ를 균등하고 임의로 선택하는 확률적 다항시간 알고리즘은 실제로 알려진 것이 없다. LDG를 $k \in I$가 입력되었을 때 LD_k의 원소를 출력하는 확률적 알고리즘이라 하자. DKL-Assumption이 성립하기 위해서는, 참고로 $|[LDG(k)]|$는 충분히 커야한다 (자세히 설명해서, 다항 개수보다 커야한다)[4]. Ko[5]의 Theorem 3의 증명으로부터 다음의 따름정리를 얻을 수 있다.

따름정리 2.3: 모든 $n \ge 2, p \in N$에 대하여 다음 세 가지 조건을 만족하는 (n,p)에 대한 확률적 다항시간 알고리즘 A가 존재한다.

(i) $[A(n,p)] \subseteq \{\chi \in B_n^+ - \Delta_n B_n^+ \mid \mathrm{len}(\chi) = p\}$.

(ii) $A(n,p)$는 $[A(n,p)]$ 상에서 균등 확률 분포를 갖는다.

(iii) $|[A(n,p)]| \ge \left(\lfloor \frac{n-1}{2} \rfloor !\right)^p$.

위의 따름정리 (i)로부터 정의 2.1에서 정의된 IG는 확률적 다항시간 알고리즘으로서 간주될 수 있다. 그리고 이와 별개로, 위의 따름정리 (i), (ii), (iii)으로부터, LDG는 위의 세 조건을 만족하는 확

률적 다항시간 알고리즘으로서 간주될 수 있다. 따라서 이후로는 $\chi \leftarrow_U LD_k$는 다음의 이중적인 의미를 갖는다: (i) 위 세 가지 조건을 만족하는 LDG가 존재한다. (ii) $\chi \leftarrow LDG(k)$이다. 즉, LD_k는 확률적 관점에서 $[LDG(k)]$를 의미한다. 마찬가지로 $\chi \leftarrow_U RD_k$ 와 $\chi \leftarrow_U R_{k,\alpha}$를 동일한 맥락에서 이해하자.

Ⅲ. 설 계

여기서는 유사난수 도구 중 가장 기본이 되는 유사난수 생성기를 DKL-Assumption 하에서 설계한다. 먼저, 유사난수 생성기의 기본 정의를 살펴본다.

정의 3.1: 아래 두 가지 조건을 만족하는 (결정적) 다항시간 알고리즘 $G:\{0,1\}^* \to \{0,1\}^*$를 유사난수 생성기(pseudorandom generator)라고 부른다.

(i) 모든 $n \in N$에 대하여 $l(n) > n$이고, 모든 $x \in \{0,1\}^*$에 대하여 $\|G(x)\| = l(\|x\|)$를 만족하는 확장 함수 $l : N \to N$이 존재한다.

(ii) 모든 확률적 다항시간 알고리즘 A와 양의 다항함수 P, 충분히 큰 모든 $n \in N$에 대하여 다음 식이 성립한다.

$$|\Pr[A(G(x))=1 : x \leftarrow_U \{0,1\}^n] - \Pr[A(r)=1 : r \leftarrow_U \{0,1\}^{l(n)}]| < \frac{1}{P(n)}.$$

본 논문이 제안하고자 하는 유사난수 생성기는, $\alpha \in B_n$, $\chi \in B_m$에 대하여 $(\alpha, \chi^{-1}\alpha\chi)$이 주어져 있을 때, 비록 $\tau(B_{n-m})$에서 임의로 선택된 많은 (다항개수) ψ_i들에 대해 $(\psi_i^{-1}\alpha\psi_i, \chi^{-1}\psi_i^{-1}\alpha\psi_i\chi)$들을 알더라도 χ를 알아내기는 어려울 것이라는 점에 착안하여 설계된다.

기호: 모든 $k \in I$와 $\alpha \in I_{1,k}$에 대하여,

$$LR_{k,\alpha} = \{\chi^{-1}\alpha\chi \mid x \in LD_k\}$$

라 정의하자.

다음은 개개의 유사난수 생성기를 결정하는 생성자에 관한 정의이다.

정의 3.2($PGIG_{KL}$): $PGIG_{KL}$는 확률적 다항시간 알고리즘으로서 입력 (k, l)에 대하여 다음과 같이 설계된다. 여기에서 $k \in I$이고 $l \in N$이다.

i) $\alpha \leftarrow IG(k)$,
ii) $\alpha_1, \ldots, \alpha_l \leftarrow_U LR_{k,\alpha}$,
iii) output $(\alpha, \alpha_1, \ldots, \alpha_l)$.

IG의 정의로 부터, $PGIG_{KL}$는 당연히 k에 관해 다항시간 안에 수행된다.

정의 3.3(G_{KL}): $l : I \to N$를 다항함수라 하자. 모든 $k \in I$, $\alpha \in I_{1,k}$, $\vec{\alpha} = (\alpha_1, \ldots, \alpha_l) \in (LR_{k,\alpha})^l$에 대하여 $g_{\alpha,\vec{\alpha}}$를 다음과 같이 정의하자.

$$g_{\alpha,\vec{\alpha}} : RD_k \to (R_{k,\alpha})^l, \quad \psi \mapsto (\psi^{-1}\alpha_1\psi, \ldots, \psi^{-1}\alpha_l\psi).$$

G_k를 $PGIG_{KL}(k, l)$의 확률 분포를 따르고 $G_k(\alpha, \vec{\alpha}) = g_{\alpha,\vec{\alpha}}$로 정의되는 확률 변수라 하고, $G_{KL} = \{G_k\}_{k \in I}$라 하자.

좌표준형 n-땋임들 χ_1, χ_2 ($\mathrm{len}(\chi_i) = O(p)$, $i = 1,2$) 로부터 $\chi_1\chi_2$의 좌표준형을 구하는 과정의 복잡도는 $O(pn^2 \log n)$이고, χ_1^{-1}의 좌표준형을 구하는 과정의 복잡도는 $O(pn)$이다 ([5] 참고). 따라서, 키 $(\alpha, \vec{\alpha})$가 고정되어 있을 때, $\psi \in RD_k$로부터 $g_{\alpha,\vec{\alpha}}(\psi)$를 구하는 속도는 $k = (n, p)$에 대한 다항시간 $O(l(k)pn^2 \log n)$이다.

Ⅳ. 안전성 분석

여기서는 G_{KL}의 안전성을 분석한다.

정리 4.1: 모든 확률적 다항시간 알고리즘 A와, 양의 다항함수 $P(\cdot)$, 충분히 큰 모든 $k \in I$에 대하여 DKL-Assumption 하에서 다음 식이 성립한다.

$$|\Pr[A(g_{\alpha,\vec{\alpha}}(\psi)) = 1 : (\alpha, \vec{\alpha}) \leftarrow PGIG_{KL}(k, l);\ \psi \leftarrow_U RD_k] - \Pr[A(\beta_1, \ldots, \beta_l) = 1 :$$

$\alpha \leftarrow IG(k); \beta_1, \ldots, \beta_l \leftarrow_U R_{k,a}]|$
$< \frac{1}{P(k)}$.

여기서, $l = l(k)$이다.

본 정리는 표준 hybrid 기법(standard hybrid technique)[4~6]에 의해 다음과 같이 증명된다.

증명: 귀류법에 의해, 확률적 다항시간 알고리즘 A와, 양의 다항함수 $P(\cdot)$, N의 무한 부분 집합 F가 존재해서, 모든 $k \in F$에 대하여 다음 식이 성립한다고 가정하자.

$|\Pr[A(g_{a,\vec{a}}(\psi)) = 1:$
$(a, \vec{a}) \leftarrow PGIG_{KL}(k, l);\ \psi \leftarrow_U RD_k]$
$- \Pr[A(\beta_1, \ldots, \beta_l) = 1:$
$\alpha \leftarrow IG(k); \beta_1, \ldots, \beta_l \leftarrow_R R_{k,a}]|$
$\geq \frac{1}{P(k)}$.

만약, 모든 $k \in F$에 대해서 다음 식을 만족하는 확률적 다항시간 알고리즘 M이 존재하면, 이는 DKL-Assumption에 모순이 되어 본 정리가 증명된다.

$|\Pr[M(\alpha, \chi^{-1}\alpha\chi, \psi^{-1}\alpha\psi, \psi^{-1}\chi^{-1}\alpha\chi\psi) = 1:$
$\alpha \leftarrow IG(k); \chi \leftarrow_U LD_k; \psi \leftarrow_U RD_k]$
$- \Pr[M(\alpha, \chi^{-1}\alpha\chi, \psi^{-1}\alpha\psi, \beta) = 1: \alpha \leftarrow IG(k);$
$\chi \leftarrow_U LD_k; \psi \leftarrow_U RD_k; \beta \leftarrow_U R_{k,a}]|$
$\geq \frac{1}{l(k)P(k)}$.

편의상 $k \in F$를 임의로 고정하고, $l = l(k)$라 하자. 입력 $\langle \alpha, \chi^{-1}\alpha\chi, \psi^{-1}\alpha\psi, \beta' \rangle$에 대하여, M를 다음과 같이 설계하자. 여기에서, $\alpha \in I_{1,k}$, $\chi \in LD_k$, $\psi \in RD_k$, $\beta' \in R_{k,a}$이다.

1. $J \leftarrow_U \{1, \ldots, l\}$.
2. $\chi_1, \ldots, \chi_{J-1}, \chi_{J+1}, \ldots, \chi_l \leftarrow_U LD_k$;
   $\beta_{J+1}, \ldots, \beta_l \leftarrow_U R_{k,a}$.
3. $H = \langle \chi_1^{-1}\psi^{-1}\alpha\psi\chi_1, \ldots, \chi_{J-1}^{-1}\psi^{-1}\alpha\psi$
   $\chi_{J-1}, \beta', \beta_{J+1}, \ldots, \beta_l \rangle$라 하자.
4. $A(H)$를 출력한다.

위의 알고리즘으로부터, M은 균등 확률 변수 J에 대한 함수이다. 따라서, 다음 식 (1)이 성립한다.

(1) $\Pr[M(\alpha, \chi^{-1}\alpha\chi, \psi^{-1}\alpha\psi, \psi^{-1}\chi^{-1}\alpha\chi\psi)$
$= 1; \alpha \leftarrow IG(k); \chi \leftarrow_U LD_k; \psi \leftarrow_U RD_k]$
$= \sum_{j=1}^{l} \Pr[(M(\alpha, \chi^{-1}\alpha\chi, \psi^{-1}\alpha\psi, \psi^{-1}\chi^{-1}\alpha\chi\psi)$
$= 1: \alpha \leftarrow IG(k); \chi \leftarrow_U LD_k; \psi \leftarrow_U RD_k)$
and $(J = j: J \leftarrow_U \{1, \ldots, l\})]$
$= \sum_{j=1}^{l} \Pr[J = j: J \leftarrow_U \{1, \ldots, l\}] \cdot \Pr[$
$(M(\alpha, \chi^{-1}\alpha\chi, \psi^{-1}\alpha\psi, \psi^{-1}\chi^{-1}\alpha\chi\psi) = 1:$
$\alpha \leftarrow IG(k); \chi \leftarrow_U LD_k;\ \psi \leftarrow_U RD_k)|$
$(J = j: J \leftarrow_U \{1, \ldots, l\})]$
$= \frac{1}{l} \sum_{j=1}^{l} \Pr[(M(\alpha, \chi^{-1}\alpha\chi, \psi^{-1}\alpha\psi,$
$\psi^{-1}\chi^{-1}\alpha\chi\psi) = 1: \alpha \leftarrow IG(k); \chi \leftarrow_U LD_k;$
$\psi \leftarrow_R RD_k)|(J = j: J \leftarrow_U \{1, \ldots, l\})]$.

마찬가지로, 다음 식 (2)도 성립한다.

(2) $\Pr[M(\alpha, \chi^{-1}\alpha\chi, \psi^{-1}\alpha\psi, \beta) = 1: \alpha \leftarrow IG(k);$
$\chi \leftarrow_U LD_k; \psi \leftarrow_U RD_k; \beta \leftarrow_U R_{k,a}]$
$= \frac{1}{l} \sum_{j=1}^{l} \Pr[(M(\alpha, \chi^{-1}\alpha\chi, \psi^{-1}\alpha\psi, \beta) = 1:$
$\alpha \leftarrow IG(k); \chi \leftarrow_U LD_k; \psi \leftarrow_R RD_k;$
$\beta \leftarrow_R R_{k,a})|(J = j: J \leftarrow_U \{1, \ldots, l\})]$.

각각의 $i \in \{1, \ldots, l\}$에 대하여, i번째 확률 분포 $H_i^{k,l}$를 다음과 같이 정의한다.

$H_i^{k,l} = \langle \psi^{-1}\alpha_1\psi, \ldots, \psi^{-1}\alpha_i\psi, \beta_{i+1}, \ldots, \beta_l \rangle$.

여기에서, $(\alpha, \alpha_1, \ldots, \alpha_i) \leftarrow PGIG_{KL}(k, i)$;
$\psi \leftarrow_U RD_k;\ \beta_{i+1}, \ldots, \beta_l \leftarrow_U R_{k,a}$이다.

그러면, $PGIG_{KL}$과 $H_i^{k,l}$, M의 정의로부터 다음 식 (3), (4)가 성립한다.

(3) $\Pr[(M(\alpha,\chi^{-1}\alpha\chi,\phi^{-1}\alpha\phi,\phi^{-1}\chi^{-1}\alpha\chi\phi)=1:$
$\alpha\leftarrow IG(k);\ \chi\leftarrow_U LD_k;\ \phi\leftarrow_R RD_k)|$
$(J=j:J\leftarrow_U\{1,\ldots,l\})]$
$=\Pr[A(H_j^{k,l})=1:(\alpha,\alpha_1,\ldots,\alpha_j)\leftarrow$
$PGIG_{KL}(k,j);\ \phi\leftarrow_U RD_k;$
$\beta_{j+1},\ldots,\beta_l\leftarrow_U R_{k,a}]$.

(4) $\Pr[(M(\alpha,\chi^{-1}\alpha\chi,\phi^{-1}\alpha\phi,\beta)=1:$
$\alpha\leftarrow IG(k);\ \chi\leftarrow_U LD_k;\ \phi\leftarrow_U RD_k;$
$\beta\leftarrow_U R_{k,a})|\ (J=j:J\leftarrow_U\{1,\ldots,l\})]$
$=\Pr[A(H_{j-1}^{k,l})=1:(\alpha,\alpha_1,\ldots,\alpha_{j-1})$
$\leftarrow PGIG_{KL}(k,j-1);\ \phi\leftarrow_U RD_k;$
$\beta_j,\ldots,\beta_l\leftarrow_U R_{k,a}]$.

$H_i^{k,l}$와 $g_{a,\vec{a}}$의 정의로부터 다음 식 (5), (6)이 성립한다.

(5) $\Pr[A(H_l^{k,l})=1:(\alpha,\alpha_1,\ldots,\alpha_l)\leftarrow$
$PGIG_{KL}(k,l);\phi\leftarrow_U RD_k]$
$=\Pr[A(g_{a,\vec{a}}(\phi))=1:(\alpha,\vec{a})$
$\leftarrow PGIG_{KL}(k,l);\ \phi\leftarrow_U RD_k]$.

(6) $\Pr[A(H_0^{k,l})=1:\alpha\leftarrow IG(k);$
$\beta_1,\ldots,\beta_l\leftarrow_U R_{k,a}]$
$=\Pr[A(\beta_1,\ldots,\beta_l)=1:\alpha\leftarrow IG(k);$
$\beta_1,\ldots,\beta_l\leftarrow_U R_{k,a}]$.

(1)-(6)에 의해서, 모든 $k\in F$에 대하여 다음 식이 성립한다.

$|\Pr[M(\alpha,\chi^{-1}\alpha\chi,\phi^{-1}\alpha\phi,\phi^{-1}\chi^{-1}\alpha\chi\phi)=1:$
$\alpha\leftarrow IG(k);\ \chi\leftarrow_U LD_k;\ \phi\leftarrow_U RD_k]$
$-\Pr[M(\alpha,\chi^{-1}\alpha\chi,\phi^{-1}\alpha\phi,\beta)=1:\alpha\leftarrow IG(k);$
$\chi\leftarrow_U LD_k;\phi\leftarrow_U RD_k;\beta\leftarrow_U R_{k,a}]|$
$=\frac{1}{l}\,|\sum_{j=1}^{l}\Pr[A(H_j^{k,l})=1:(\alpha,\alpha_1,\ldots,\alpha_j)$
$\leftarrow PGIG_{KL}(k,j);\phi\leftarrow_U RD_k;$
$\beta_{j+1},\ldots,\beta_l\leftarrow_U R_{k,a}]$
$-\sum_{j=1}^{l}\Pr[A(H_{j-1}^{k,l})=1:$
$(\alpha,\alpha_1,\ldots,\alpha_{j-1})\leftarrow PGIG_{KL}(k,j-1);$
$\phi\leftarrow_U RD_k;\beta_j,\ldots,\beta_l\leftarrow_U R_{k,a}]|$
$=\frac{1}{l}\,|\Pr[A(H_l^{k,l})=1:(\alpha,\alpha_1,\ldots,\alpha_l)$
$\leftarrow PGIG_{KL}(k,l);\phi\leftarrow_U RD_k]$
$-\Pr[A(H_0^{k,l})=1:\alpha\leftarrow IG(k);$
$\beta_1,\ldots,\beta_l\leftarrow_U R_{k,a}]|$
$\geq\frac{1}{lP(k)}$. Q.E.D.

V. 유사난수 생성기

정리 4.1로부터, G_{KL}은 유사난수 땋임들의 수열을 생성함을 알 수 있다. 여기서는 G_{KL}로부터, 유사난수 비트들의 수열을 생성하는 유사난수 생성기를 설계한다.

정의 5.1: $m<n$에 대하여, $H(\subseteq F_{n,m})$가 짝으로 독립인 해쉬 함수들의 집합 (a family of pairwise independent universal hash functions)이라 함은, $\forall x\neq x'\in\{0,1\}^n$, $\forall a,a'\in\{0,1\}^m$에 대하여 다음 식이 성립함을 의미한다.

$$\Pr[(h(x)=a)\wedge(h(x')=a'):h\leftarrow_U H]=\frac{1}{2^{2m}}.$$

따라서, H가 짝으로 독립인 해쉬 함수의 집합이라 함은, 서로 다른 x,x'에 대하여 $h\leftarrow_U H$일 때, $h(x),h(x')$가 독립적으로 균등 분포를 보임을 의미한다.

보조정리 5.2[6]: $n,b,e\in N$이 $3e<b<n$이라 하자. $S\subseteq\{0,1\}^n$이 $|S|\geq 2^b$이고, x는 S 상에서 균등 확률 분포를 따르는 확률 변수라 하자. H를 짝으로 독립인 해쉬 함수들, $\{0,1\}^n\rightarrow\{0,1\}^{b-3e}$의 집합이라 하자. 그러면, 2^{-e}의 비율을 제외한 모든 $h\in H$에 대해 $\{0,1\}^{b-3e}$ 상에서의 균등 분포와 $h(x)$의 분포는 최대 2^{-e}의 통계적 거리를 갖는다.

모든 $k=(n,p)\in I$와 $\alpha\in I_{1,k}$, $\vec{a}=(a_1,\ldots,a_{l(k)})\in(LR_{k,a})^{l(k)}$에 대하여, 함수 $a(k),b(k)$를 $a(k)=5pn^2l(k)$, $b(k)=pnl(k)$ 라 정의하자. 그러면 $(R_{k,a})^{l(k)}\subseteq\{0,1\}^{a(k)}$이고 $|R_{k,a}|^{l(k)}>2^{b(k)}$이 된다.

정의 5.3 ($\tilde{G}_{KL}$): 모든 $k\in I$에 대하여, H_k를 짝으로 독립인 해쉬 함수들, $\{0,1\}^{a(k)}\to\{0,1\}^{b(k)/2}$의 집합이라 하자. 각각의 $h\in H_k$에 대해 $\tilde{g}_{\alpha,\vec{\alpha},h}$를 다음과 같이 정의하자.

$$\begin{aligned}\tilde{g}_{\alpha,\vec{\alpha},h}: RD_k &\to \{0,1\}^{b(k)/2}\\ \psi &\mapsto h(g_{\alpha,\vec{\alpha}}(\psi)).\end{aligned}$$

$\tilde{G}_k$를 H_k 상에서의 균등 확률 분포와 $PGIG_{KL}(k,l)$의 확률 분포를 따르고 $\tilde{G}_k(\alpha,\vec{\alpha},h)=\tilde{g}_{\alpha,\vec{\alpha},h}$로 정의되는 확률 변수라 하고, $\tilde{G}_{KL}=\{\tilde{G}_k\}_{k\in I}$라 하자.

정리 5.4: DKL-Assumprion 하에서, $\tilde{G}KL$는 유사난수 생성기이다.

기초정리 5.5: f가 집합 S상에서의 함수이고, D,E가 S상에서의 확률 분포들이면, 항상 다음 식이 성립한다.

$$\mathrm{dist}(f(D),f(E))\le \mathrm{dist}(D,E).$$

증명:

$$\begin{aligned}&\mathrm{dist}(f(D),f(E))\\ &=\frac{1}{2}\sum_{a\in f(S)}|\Pr[f(x)=a: x\leftarrow_D S]\\ &\quad-\Pr[f(y)=a: y\leftarrow_E S]|\\ &=\frac{1}{2}\sum_{a\in f(S)}|\sum_{z\in f^{-1}(a)}(\Pr[x=z: x\leftarrow_D S]\\ &\quad-\Pr[y=z: y\leftarrow_E S])|\\ &\le\frac{1}{2}\sum_{z\in S}|\Pr[x=z: x\leftarrow_D S]\\ &\quad-\Pr[y=z: y\leftarrow_E S]|\\ &=\mathrm{dist}(D,E).\end{aligned}$$ Q.E.D.

따름정리 5.6: $A:S\to\{0,1\}$가 확률적 다항 시간 알고리즘이고, D,E가 S상에서의 확률 분포들이면, 항상 다음 식이 성립한다.

$$\begin{aligned}&|\Pr[A(x)=1: x\leftarrow_D S]\\ &\quad-\Pr[A(y)=1: x\leftarrow_E S]|\\ &\le \mathrm{dist}(D,E).\end{aligned}$$

증명의 아이디어:

(i) A를 이변수 함수 $S\times T\to\{0,1\}$로 본다. 여기서, T는 A의 내부의 임의성을 결정하는 모든 비트열들의 집합이다.

(ii)
$$\begin{aligned}&\mathrm{dist}(A(D),A(E))\\ &=\frac{1}{2}(|\Pr[A(x)=0: x\leftarrow_D S]\\ &\quad-\Pr[A(y)=0: x\leftarrow_E S]|\\ &\quad+|\Pr[A(x)=1: x\leftarrow_D S]\\ &\quad-\Pr[A(y)=1: x\leftarrow_E S]|)\\ &=|\Pr[A(x)=1: x\leftarrow_D S]\\ &\quad-\Pr[A(y)=1: x\leftarrow_E S]|.\end{aligned}$$ Q.E.D.

정리 5.4의 증명: 귀류법에 의해 $\tilde{G}_{KL}$가 유사난수 생성기가 아니라고 가정하자. 그러면, 확률적 다항 시간 알고리즘 A와 양의 다항 함수 P, I의 무한 부분 집합 F가 존재해서, 모든 $k\in F$에 대하여 다음 식이 성립한다.

$$\begin{aligned}&|\Pr[A(h(g_{\alpha,\vec{\alpha}}(\psi)))=1:(\alpha,\vec{\alpha})\leftarrow\\ &\quad PGIG_{KL}(k,l);h\leftarrow_U H_k;\psi\leftarrow_U RD_k]\\ &\quad-\Pr[A(r)=1: r\leftarrow_U \{0,1\}^{b(k)/2}]|\\ &\ge\frac{1}{P(k)}.\end{aligned}\quad(1)$$

이 식의 좌변에 대하여 다음 식이 성립한다.

$$\begin{aligned}&|\Pr[A(h(g_{\alpha,\vec{\alpha}}(\psi)))=1: h\leftarrow_U H_k;\\ &\quad(\alpha,\vec{\alpha})\leftarrow PGIG_{KL}(k,l);\psi\leftarrow_U RD_k]\\ &\quad-\Pr[A(r)=1: r\leftarrow_U \{0,1\}^{b(k)/2}]|\\ &\le|\Pr[A(h(g_{\alpha,\vec{\alpha}}(\psi)))=1: h\leftarrow_U H_k;\\ &\quad(\alpha,\vec{\alpha})\leftarrow PGIG_{KL}(k,l);\psi\leftarrow_U RD_k]\\ &\quad-\Pr[A(h(\beta_1,\ldots,\beta_{l(k)}))=1: h\leftarrow_U H_k;\\ &\quad\alpha\leftarrow IG(k);\beta_1,\ldots,\beta_{l(k)}\leftarrow_U R_{k,\alpha}]|\\ &+|\Pr[A(h(\beta_1,\ldots,\beta_{l(k)}))=1: h\leftarrow_U H_k;\\ &\quad\alpha\leftarrow IG(k);\beta_1,\ldots,\beta_{l(k)}\leftarrow_U R_{k,\alpha}]\\ &\quad-\Pr[A(r)=1: r\leftarrow_U \{0,1\}^{b(k)/2}]|.\end{aligned}\quad(2)$$

보조정리 5.2에 의해서 다음 식이 성립한다.

$$|\Pr[A(h(\beta_1,\ldots,\beta_{l(k)}))=1 : h \leftarrow_U H_k, \alpha \leftarrow IG(k); \beta_1,\ldots,\beta_{l(k)} \leftarrow_U R_{k,a}] - \Pr[A(r)=1 : r \leftarrow_U \{0,1\}^{b(k)/2}]| \leq 2^{1-b(k)/6}. \quad (3)$$

(1)-(3)에 의해서, 다음 식을 만족하는 양의 다항함수 Q가 존재한다.

$$|\Pr[A(h(g_{\alpha,\vec{a}}(\psi)))=1 : (\alpha,\vec{a}) \leftarrow PGIG_{KL}(k,l); h \leftarrow_U H_k, \psi \leftarrow_U RD_k] - \Pr[A(h(\beta_1,\ldots,\beta_{l(k)}))=1 : h \leftarrow_U H_k, \alpha \leftarrow IG(k); \beta_1,\ldots,\beta_{l(k)} \leftarrow_U R_{k,a}]| \geq \frac{1}{Q(k)}.$$

이는 따름정리 5.6에 의해서, 정리 4.1에 모순이다. Q.E.D.

유사난수 생성기, $\hat{G}_{KL}$,의 확장 성질은 함수 $l(\cdot)$에 달려있다. 즉, $l(\cdot)$은 $l(k)\log_2(|R_{k,a}|) > 2|RD_k|$을 만족해야 한다. 예 를 들어 $l(n,p)=2np$ 정도면 확장 함수로서 충분하다.

Ⅵ. 결 론

본 논문에서는 결정적 Ko-Lee 가정(DKL-Assumption) 하에서 단순하고 실용적인 유사난수 생성기를 설계하고 안전함을 증명하였다. 따라서, DKL-Assumption이 유효하다면, $\hat{G}_{KL}$은 비주기적이고, 난수도를 측정하는 모든 통계적 검증을 통과하여야 한다. 그러므로, $\hat{G}_{KL}$을 실제로 구현해서, 주기 및 통계적 검증 결과를 관찰함으로써 DKL-Assumption 자체를 분석하는 작업도 의미가 있을 것이다.

땋임군은 지금까지 암호학에서 주로 다루어 오던 다른 군들과는 매우 달라, 유사난수 생성기로부터 유사함수 생성기를 설계하는 기존 방법들이 적용되지 못한다. 따라서 본 논문의 결과로부터 다음 단계의 유사난수 도구인 유사함수 생성기를 만드는 방법을 연구하는 작업도 가치 있을 것이다.

참 고 문 헌

[1] I. Anshel, M. Anshel, and D. Goldfeld, "An algebraic method for public-key crypto graphy", *Math. Res. Lett.* 6, pp. 287-291, 1999.

[2] J.S. Birman, K.H. Ko, and S.J. Lee, "New approaches to the world and conjugacy problem in the braid groups", *Advances in Math.* 139 pp. 322-353, 1998.

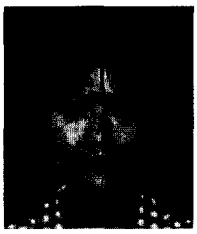
[3] D.B.A. Epstein, J.W. Cannon, D.F. Holt, S.V.F. Levy, M.S. Patterson, and W. Thurston, *Word processing in groups*, Jones and Barlett, Boston and London, 1992.

[4] O. Goldreich, Foundation of Cryptography—*Fragments of a Book*, Available at http://www.theory.lcs.mit.edu/~oded/frag.html, 1995.

[5] K.H. Ko, S.J. Lee, J.H. Cheon, J.W. Han, J.S. Kang, and C. Park, "New Public-key Cryptosystem Using Braid Groups", Proc. Crypto 2000, LNCS 1880, pp. 166-183, 2000.

[6] M. Naor and O. Reingold, "Number-Theoretic constructions of efficient pseudorandom functions", *Proc. 38th IEEE Symp. on Foundations of Computer Science*, pp. 458-467, 1997.

〈著者紹介〉



이언경 (Eonkyung Lee) 비회원
1992년 2월 : 한국과학기술원 수학과 졸업
1994년 2월 : 한국과학기술원 수학과 석사 졸업
1994년 2월~1997년 8월 : 한국전자통신연구원 연구원
1997년 9월~현재 : 한국과학기술원 수학과 박사과정
〈관심분야〉 암호학, Pseudorandomness, Provable Seurity

한상근 (Sang Geun Hahn) 종신회원
1979년 : 서울대학교 수학과 졸업
1982년 : 뉴멕시코 주립대 석사 졸업
1987년 : 오하이오 주립대 박사 졸업
1987년~현재 : 한국과학기술원 수학과 교수
〈관심분야〉 암호학, 타원곡선, 정수론