난수의 안전성 평가를 위한 새로운 검정

서중호, 김혜정, 이경현
부경대학교 전자계산학과

A New Test for Security Evaluation of Random Sequences

Jung-Ho Seo, Hae-Jeong Kim, Kyung-Hyune Rhee
Dept. of Computer Science
Pukyong National University

요 약

본 논문에서는 이진난수발생기들의 안전성 평가를 위한 새로운 통계적 검정을 소개한다. 검정에서 구현된 기본개념은 이진난수열이 랜덤하지 않다면 다음 발생 비트를 예측할 확률이 편향된다는 다음 비트 검정이론에 바탕을 둔다. 본 검정은 이진난수열이 아닌 임의의 d진 난수열의 안전성 검정에도 유용하게 적용될 수 있으므로 난수발생기를 이용하여 설계된 스트림 암호 시스템의 안전성 평가에 평가 측도로서 사용될 수 있다. 또한 컴퓨터 시뮬레이션을 통해 몇몇 난수발생기에 검정을 적용함으로써 검정법이 타당함을 보인다.

1. 서 론

수학적으로 완전한 난수열(truly random sequence)이란 확률변수열 $\{X_n\}$이 독립적이고 같은 분포를 가질 때 $X_n = x_n$인 실현치의 수열 $\{X_n\}$을 의미한다. 특히, $P(X_n = 0) = P(X_n = 1) = \frac{1}{2}$이고 i. i. d (independent and identically distributed) 확률변수열 $\{X_n\}$으로부터 얻어진 수열 $\{x_n\}$을 완전한 이진 난수열이라 부른다.

난수 발생기로부터 얻어지는 난수열은 모의실험(simulation), 샘플링, 수치해석, 컴퓨터 프로그래밍, 암호학 등의 여러 분야에 사용된다. 난수열의 암호학에서의 주요 응용분야는 스트림 암호계의 키(key) 수열의 소스를 제공하는 것이다. 이때 암호계의 안전성은 사용된 키수열의 randomness에 전적으로 의존하게 된다. 우리가 실제적으로 사용하는 수열은 난수발생기로부터 선형합동법(linear congruential method), linear shift register 등과 같은 결정적인 방법에 의하여 생성된다. 그러므로 이렇게 얻어진 수열을 의사랜덤수열(quasi-random sequence)이라고 부르며, 의사랜덤수열이 랜덤하다는 것을 보장할 통계적 검정이나 이론적 기준이 필요하게 된다. 이진난수열에서 발생된 비트들에 대한 다음비트검정(Next bit test)이 일반적인 검정으로서의 성질을 가진다는 것이 Yao[2]에 의해 증명되었다. 다음비트 검정은 이진난수발생기에 의해 생성되어진 스트림의 임의의 i번째 비트에 대해 1/2이상의 성공확률을 가지고 다음비트인 (i+1)번째 비트를 예측하려는 것이다. 이러한 다음비트를 효율적으로 예측하는 것이 불가능한 경우의 이진난수발생기를 의사랜덤이라 한다. 이진난수발생기가 본 논문에서 소개된 검정을 통과할 경우 기존에 알려진 통계적 검정에서 평가되는 비트들간의 일양성과 독립성을 통과한다는 관점에서 이 검정을 보다 일반적이라고 평가할 수 있다.

본 논문은 다음비트 검정의 개념을 바탕으로 실제적인 검정 알고리즘의 구현과 이 알고리즘을 여러개

의 잘 알려진 이진난수발생기에 직접적으로 적용한 결과를 도출하여 난수발생기의 안전성 평가 측도로서 사용하고자 하는데 그 목적을 두고 있다. 본 논문의 검정은 임의 이진난수발생기의 난수열에도 확대적용가능하며 임의 길이 스트림에도 적용되어질 수 있는 특징이 있다.

2. 일반적 통계검정

주어진 수열에 대한 randomness의 정의를 *i.i.d.*의 관점에서 볼 때 여러가지 통계적 검정의 주된 원리는 그 수열이 가지고 있는 각 항들 사이의 독립성을 보장할 수 없는 상호연관성(correlation)이나 일양분포를 따르지 않는 빈도의 편향성(bias)등과 같은 통계적 약점을 찾아내는데 있다.

기존의 시뮬레이션이나 암호적인 관점에서 난수발생기의 안전성 평가도구로서 사용되어온 대표적인 검정법[5]들은 다음과 같다.

2.1 도수검정(Frequency test)

대상수열에 대해 0과 1의 수가 일양적으로 분포하고 있는지를 결정하는데 사용하는 검정이다.

2.2 포커검정(Poker test)

대상수열에서 임의의 m 비트의 패턴을 고려하는 검정으로서 2^m의 서로 다른 패턴이 존재한다.

2.3 계열검정(Serial test)

대상수열에서 00, 01, 10, 11의 비트쌍이 고려되어지는 검정으로 검증하려는 대상 이진 수열에서 한 비트가 그 다음비트로 전이되는 전이확률을 검정하는 것이다.

2.4 자기상관검정(Autocorrelation test)

이진 수열 S_n이 주어졌을 때 (S_n)에서 d비트 만큼 전이시켜 생성한 수열 (S_{n+d})과의 상관관계를 조사하는 검정이다.

2.5 연검정(Run test)

수열에서 '0'이나 '1'이 연속하여 나타나는 것을 run이라 한다. 이 검정은 수열에서 다양한 길이의 run들의 수가 난수에 대해 예측되어지는 것과 같은지를 결정한다.

이러한 검정들 중 도수, 포커, 연은 편향성 관점의 계열과 자기상관은 독립성 관점의 검정으로 평가될 수 있다.

3. Universal 검정의 이론

난수열이 서로 독립이고 일양분포를 따르는지를 동시에 검증하는 것이다.

Schrift와 Shamir[1]의 결과에서 사용되어진 표기에 따라 본 논문에서도 아래와 같은 기호와 정의들이 사용되어진다.

s_1^n : $\{0,1\}^n$상에서 길이 n 을 갖는 이진 스트링

s_i : 스트링의 i 번째 비트

s_j^k : j번째 비트에서부터 k 번째 비트까지의 스트링

$O(v(n))$: $v(n)$의 복잡도를 가지는 Big-Oh 표기

아래에는 확률적 다항식 시간 알고리즘의 용어를 사용하여 이론적인 결과를 표현하고 있다.

정의 3.1 Next Bit 검정

만일 임의의 i와 $(1 < i \le n)$ 모든 확률적 다항식 시간 알고리즘 $A : \{0,1\}^{i-1} \to \{0,1\}$, $| prob_s\{A(s_1^{i-1}) = s_i\} - \frac{1}{2} | \le O(v(n))$이면 입력 스트링 S는 다음 비트 검정을 통과한다고 정의한다.

정의 3.2 POP(Predict or Pass) Test

만일 모든 $i(1 < i \le n)$, 모든 고정값 c, 모든 확률적 다항식시간 알고리즘 $A : \{0,1\}^{i-1} \to \{0,1,?\}$에 대해서, $\Pr ob_s\{A(s_1^{i-1}) \neq ?\} \ge \frac{1}{n^c}$ 이면

$\Pr ob_s\{A(s_1^{i-1}) = s_i \mid A(s_1^{i-1} \neq ?)\} - b \le O(v(n))$

이고 편향된 입력 스트링 S는 POP 검정을 통과한다고 정의한다.

다음 비트로 1이 예상되는 A는 1을 출력하고, 다음 비트로 0이 예상되면 A는 0을 출력한다. A가 예상치를 가질 수 없다면 ?를 출력하게 된다.

정의 3.3 Extended POP 검정

각 i, l $(1 < i, l \le n)$, 고정된 상수 c, 모든 확률적 다항식 시간 알고리즘

$A : \{0,1\}^{i-1} \to \{\{0,1\}^l, ?\}$에 대해 편향된 소스 S는 만약 $\Pr ob\{A(s_1^{i-1}) \neq ?\} \ge \frac{1}{n^c}$ 일 때

$| \Pr ob\{(A(s_1^{i-1}) = 1) = s_i^{i-1+l} \mid A(s_1^{i-1}) \neq ?\} \neq ? | \le O(v(n))$

이 성립한다면, 확장된 POP 검정을 통과한다.

만일 소스가 확장된 POP 검정을 통과할 수 없다면 주어진 스트링 블록이 다음 블록을 효율적으로 추측할 수 있는 확률적 다항식 시간 검정이 존재함을 의미한다. 따라서, 검정 A는 주어진 이전 블록으로부터 s_i를 예측할 수 있고 소스는 완전 독립적 편향된 소스가 아니다.

따라서 빈도 검정을 통과하는 의사 랜덤 발생 스트링에 대해 스트링 당 다음 비트의 확률은 계산된 편향 값을 넘지 않아야 한다.

이것은 필요조건이므로 계산된 편향 값이 확률적 다항식 시간 알고리즘 A에 대한 결정 문턱치가 되

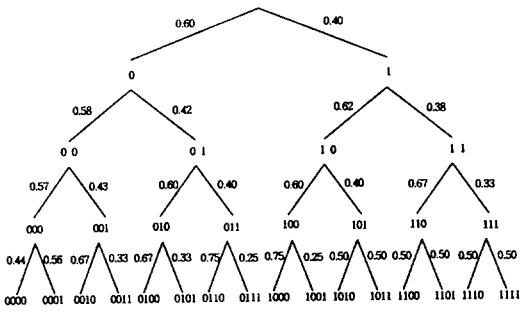
도록 정의할 수 있다. 스트링 상의 다음 비트가 편향치보다 더 높은 확률을 가지고 나타날 때, 즉 $b \le p$ or $p \le 1-b$, 알고리즘은 스트링 상의 그 비트를 예측할 수 있다. 스트링 상의 다음 비트가 편향치보다 적은 확률을 가지고 나타날 때, 즉 $1-b \le p \le b$일 때 알고리즘은 ?을 출력하며 이것은 검정대상인 의사랜덤발생기가 확장된 POP검정을 통과함을 의미한다.

4. 새로운 Universal 검정

앞에서 살펴 본 확률적 검정과 POP 검정이 서로 연관되어질 수 있다는 데에 바탕을 두고 지금부터 실제적인 새로운 일반적 검정[3]을 소개하고자 한다. 위에서와는 다른 방식으로 트리를 정렬해 보자. 여기서는 가중 트리를 고려하는데, 각 노드에는 각 계층 상에서 발생하는 패턴의 발생 횟수를 적고 각 노드에서 바로 아래에 연결되는 간선들에는 상위 계층 상에서 발생하는 패턴의 수에 대한 하위 계층에서 발생하는 패턴의 수의 확률 비를 적어내려 가면서 트리를 구성해 간다.

예제 4.1 다음과 같은 이진 스트링이 주어졌을 경우 위의 새로운 검정에서 제안한 방식으로 비트의 발생 패턴과 발생 횟수에 따른 확률 비를 적어내려 가면서 트리를 구성해 보자($n=80$)

11000010111010100110000110110110001000100001101000100101111100001010000100000100



<그림1>예제 4.1의 스트링에 대한 트리 구성

비트들이 독립적으로 생성되어질 때, 스트링의 임의의 부분이 주어진다면 다음 비트는 상수 확률로 1이 될 것이다.

위 트리에서 아래로 이동함에 따라 확률들이 $\frac{1}{2}$에서 0과 1로 편차가 생기게 됨을 알 수 있으며 이것은 임의의 스트링에 대해서 일반적인 성질이다. 즉, 트리에서 더 낮은 계층으로 내려감에 따라 패턴들의 발생횟수는 점점 더 적어지게 된다. 이러한 성질은 포커 검정처럼 큰 m에 대해서는 통계적 검정의 수행 결과를 무의미하게 만든다. 하지만 이것은 더 낮은 계층에서 통계량이 주어진다면 큰 m을 가지는 많은 패턴들의 다음 비트를 예측할 수 있음을 알려준다.

예를 들어, 패턴 0011 이 예제 4.1의 스트링에서 나타난다면 1의 확률로 다음 비트는 1이 됨을 알 수 있다.

80비트 스트링에 대해서 만일 편차 값이 $0.3905 \le b \le 0.6095$라면 스트링은 빈도 검정을 통과하게 되고 그렇지 않다면 의사 난수 스트링이 되는 것이 기각될 것이다. 이러한 검정은 다른 길이를 가지는 임의의 다른 스트링에 대해서도 확장할 수도 있다.

정의 4.1 s_1^n은 길이 n을 가지는 스트링이다.

결정 문턱치(threshold decision) α는 다음과 같이 정의된다.

$$\alpha = \frac{1+\sqrt{\frac{x^2}{n}}}{2}$$

단 χ^2는 검정에서 요구되는 유의수준에 대응되는 값.

De bruijn수열과 같이 높은 복잡도를 가지는 랜덤 수열들에 대해서 $n=2^l$일 때, $l=\log_2(n)$인 계층에서 각 패턴은 한 번씩 발생한다. 하지만 좀더 적은 복잡도를 지니는 수열들에 대해서는 몇몇 패턴들은 계층 l에서 발생하지 않고 나머지 패턴들은 한 번이나 한번이상 발생한다. 이것은 계층 $l-1$의 노드에서 계층 l까지의 노드에 놓여있는 간선들 몇몇은 확률이 1(또는 0)이 되게 된다. 또한 이러한 확률들이 특정 패턴의 발생과 관련되어 있음을 고려한다면 스트링의 일부분을 재구성할 수도 있다.

<검정 알고리즘> 길이 n을 갖는 스트링에 대한 새로운 검정 알고리즘

Step 1. 결정 문턱치 α값을 다음과 같이 계산한다.

$$\alpha = \frac{1+\sqrt{\frac{x^2}{n}}}{2}$$

Step 2. $l = round(\log_2(n))$을 계산한다.

Step 3. 스트링의 꼬리에 스트링의 처음 부분에 나타나는 $l-1$개의 비트들을 덧붙이고 스트링을 서로 겹쳐 가면서 l비트의 단위로 나눈다.

Step 4. 각각의 블록을 비교해 나가면서 길이 l을 갖는 각 패턴의 발생횟수를 계산한다.

Step 5. 계층 l과 $l-1$에서 트리를 형성해 나가면서 각 간선에 대응되는 확률을 구한다.

Step 6. 계층 $l-1$에 있는 각 노드에 대해 만일 다음비트가 α보다 더 높은 확률을 가지고 나타난다면 다음 비트는 예측되어질 수 있으며 그렇지 않은 경우 다음 비트는 결정될 수 없다.

Step 7. 계층 $l-1$에 있는 각 노드에 대해 이후에 예측되어질 수 있는 스트링의 길이를 계산하다.

위의 알고리즘을 사용하여 스트링이 국소적으로 랜덤하지 않은 성향과 전역적으로 랜덤하지 않은 성향을 다음의 방법으로 평가할 수 있다.

(1) 국소적 non-random 성향

만약 계층 $l-1$에서 $l+1$보다 많은 비트들이 예측되어 질 수 있는 임의의 노드가 존재한다면 다음 블록이 예측될 수 있는 길이 l의 블록이 존재함을 의미한다. 따라서 스트링상에 국소적 non-randomness가 존재하고, 스트링은 요구되는 성질을 만족하는 발생기로는 기각되어질 수도 있다.

(2) 전역적 non-random 성향

스트링의 전역적 성향에 대한 평가를 위해 새로운 검정의 결과를 이용할 수 있다. 이것은 다음에 예측되어지는 비트들의 수에 대한 노드들의 수가 주어지는 히스토그램을 형성함으로써 가능해 진다

구체적인 히스토그램은 다음과 같은 절차로서 형성된다.

난수 스트링에 대해, 계층 $l-1$과 계층 l 사이의 가지들에서 나타나는 확률은 (0,1)사이의 값을 가지는 확률변수가 된다. 이러한 확률이 $(1-\alpha, \alpha)$의 범위를 초과한다면 다음 비트를 예측할 수 있게 된다. 그러나, 만일 확률이 위의 범위 내에 존재한다면 다음 비트를 결정하거나 예측할 수 없다. 이러한 $(2\alpha-1)$의 범위를 β라 부르기로 하자. 각 노드에 대해 다음 비트가 예측되어진다면 만족하는 노드들의 경로를 그려보자. 예를 들어, 추측되는 다음 비트가 1인 노드 00101로부터 노드 01011를 고려하게 된다. 이러한 단계의 노드들을 단계-1 노드라 부른다. 다시 노드 01011에 대한 다음 노드는 10111이 되고 이러한 단계에서의 노드들을 단계-2 라 부른다. 이러한 방법으로 서로 다른 단계에 있는 노드들의 수를 셀 수 있게 된다.

De bruijn수열과 같은 복잡한 수열에 대해, 단계-0 노드의 수는 $N_0 = 2^{round(\log_2(n))-1}$이 된다. 많은 노드들에 대해서, 모두 다음 비트를 예측할 수 없는데, 그러한 노드들의 수는 $\aleph_0 = \beta N_0$와 동치이다. 다음 비트를 예측할 수 있는 노드들에 대해서는 다음 비트를 추측하고 한 단계씩 더 확장할 수 있다.

일반적으로, 단계-i 노드들의 수는

$N_i = N_{i-1}(1-\beta)(1-\frac{\gamma}{2})$가 된다.

더 이상 확장할 수 없는 단계-i 노드들의 수는 $\aleph_i = \beta N_i$이고 단계-0 노드들에 대응하는 노드들의 수는 $NA_i = \frac{\aleph_i}{\left(1-\frac{\gamma}{2}\right)^i}$가 된다. 다시 한 번, 이후의 비트들이 예측되어 질 수 있는 계층 $l-1$ 노드들의 수는 NA_i가 됨을 유의하자.

여기서 γ는 두 노드가 다음단계의 같은 노드에 도달할 확률이다. 예를들면, 예제4.1의 스트링에서 노드 10011과 노드 00011의 뒤에 비트 0이 추측되어 질 수 있다. 따라서 양쪽 노드들은 모두 노드 00110에 도달한다.

5. 시뮬레이션 및 결과분석

5.1 시뮬레이션용 PRBG

새로운 검정을 4가지 종류의 난수 발생기를 이용하여 시뮬레이션 하였다. 시뮬레이션에 사용된 각각의 난수 발생기의 구성과 특성에 대해 간략히 소개하고 있으며 이들을 이용한 실제적인 시뮬레이션 결과 분석을 나타내었다.

5.1.1 LCG(Linear Congruential Generator)

선형합동발생기[4]는 $x_{i+1} = a x_i + b \pmod m$와 같은 형태를 갖는 의사 난수열 발생기이다.(여기서, x_n는 수열의 n번째 수이고, x_{n-1}은 그 이전 수이다.) 선형 합동 발생기의 성질은 parameter 값에 크게 의존하는데, 본 논문에서는 SUN workstation에서 gcc 컴파일러에 의한 random 함수를 사용하였다.

5.1.2 m-LFSR(Maximum Length Linear Feedback Shift Register) Generator

m-LFSR[4]은 원시 다항식 $f(x) = 1 + c_1 x + c_2 x^2 + \ldots + c_{n-1}x^{n-1} + x^n, c_0 = 1$을 특성 다항식으로 갖고 초기 상태가 $s_0 s_1 s_2 \ldots s_{n-1}$일 경우, 생성되는 이진 수열 (s_t)의 주기가 최대 주기인 $2^n - 1$이 되는 선형 쉬프트 레지스터라 하고 이를 최대 주기를 갖는 선형 쉬프트 레지스터라 한다.

5.1.3 BRM(Binary Rated Multiplexer) Generator

BRM 시스템[8]은 2개의 m-LFSR과 BRM 로직으로 구성된다.

Input : parameter:2 LFSRs $< L_j,\ C_j(D) >$,

k and control vector $j = (j_0, j_1, \ldots, j_{k-1})$ such that

$o \le j_0 \le j_1 < \ldots < j_{k-1} \le L_1$.

key: initial states $s_0^{(1)}, s_0^{(2)}$ of the 2 LFSRs.

For i = 1,2,...do

1. Shift $LFSR_1$, $LFSR_2$
2. Compute the integer

$$a_t = \sum_{l=0}^{k-1} 2^l s_l^{(1)}(j_l)$$

3. For j=0,1,... a_t do

Shift LFSR2
Output : the sequence of $s_i^{(2)}$, I=0,1,2,...

5.1.4 BBS(Blum-Blum-Shub) Generator

BBS 발생기[6]의 기본적 구성은 다음과 같다. p, q가 $p \equiv q \equiv 3 \bmod 4$를 만족하는 소수일 경우에, $N_{Blum} = pq$가 된다. N_{Blum}의 제곱 잉여수를 $QR(N_{Blum})$로 표현하고 초기값 x_0는 $QR(N_{Blum})$의 임의의 요소라고 한다.
이때 $i = 0,1,2,\ldots$이 되어,

$$x_{i+1} = x_i^2 \bmod N_{Blum}$$
$$y_i = lsb(x_i)$$

에 의해 얻어지는 수열 $\{y_i\}_{i\geq 1}$이 BBS 발생기의 출력이 된다. 여기서 lsb(x_i)는 x_i의 최하위 bit를 의미한다. 본 논문에서 사용된 $QR(N_{Blum})$로는 192649가 사용되었으며 $p=383$, $q=503$, 그리고 초기값은 $x_0 = 101355^2 \bmod n = 20749$가 사용되었다.

N_{Blum}의 인수분해와 $f(x) = x^2 \pmod{N_{Blum}}$의 역함수의 계산은 등가이며, N_{Blum}의 인수분해는 계산량적으로 어렵다는 것이 알려져 있으므로, f 는 일방향 치환으로 된다.

5.2 결과분석

아래에는 각각의 난수 발생기의 종류에 따라 다음 비트를 예측 가능한 비트들의 발생 횟수와 각 스트링의 전역적인 성향에 대한 평가(l =6일 경우)를 표로 나타내었다. 입력 스트링의 길이는 2500으로 잡았으며 유의 수준 5%(결정 문턱치 값 : 0.53)에 대한 경우를 표에 나타내었다.

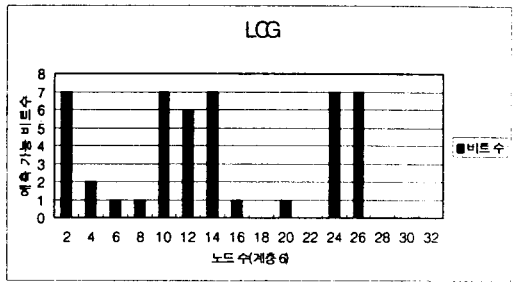
구 분	LCG	m-LFSR	BRM	BBS
계층1	0	0	1	0
계층2	1	0	2	0
계층3	4	0	4	2
계층4	5	4	7	7
계층5	10	11	10	12
계층6	24	20	26	28
계층7	54	52	57	57
계층8	110	105	115	119
계층9	230	221	221	225
계층10	412	407	370	392
계층11	751	756	545	570

<표1> 예측 가능한 다음 비트 발생횟수

계층 6에 대한 각 노드의 예측 가능한 다음 비트 블록의 수(각각의 난수 발생기들에 대한 적용 결과 비교)

노 드 수	2	4	6	8	10	12	14	16	18	20	22	24	26	28	30	32
비 트 수	7	2	1	1	7	6	7	1	0	1	0	7	7	0	0	0

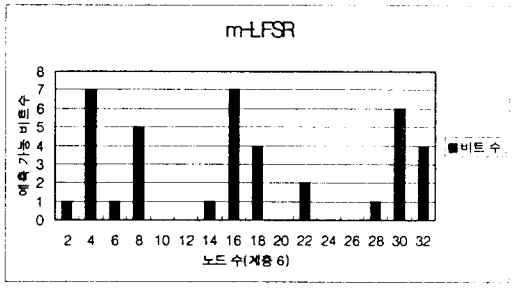
<표2> LCG의 예측가능한 다음비트수



<그림2> LCG의 전역적 성향평가

노 드 수	2	4	6	8	10	12	14	16	18	20	22	24	26	28	30	32
비 트 수	1	7	1	5	0	0	1	7	4	0	2	0	0	1	6	4

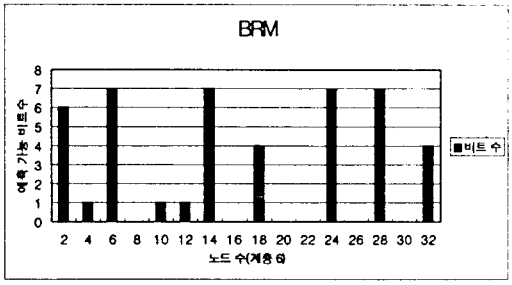
<표3> m -LFSR의 예측가능한 다음비트수



<그림3> m -LFSR의 전역적 성향평가

노 드 수	2	4	6	8	10	12	14	16	18	20	22	24	26	28	30	32
비 트 수	6	1	7	0	1	1	7	0	4	0	0	7	0	7	0	4

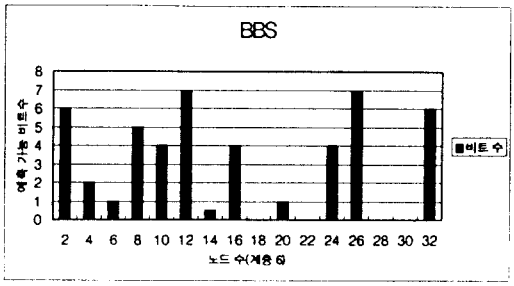
<표4> BRM의 예측가능한 다음비트수



<그림4> BRM의 전역적 성향평가

노 드 수	2	4	6	8	10	12	14	16	18	20	22	24	26	28	30	32
비 트 수	6	2	1	5	4	7	1	4	0	1	0	4	7	0	0	6

<표5> BBS의 예측가능한 다음비트수



<그림5> BBS의 전역적 성향평가

위의 검정에는 LCG, m-LFSR, BRM, BBS시스템이 사용되었으며 BRM이 계층1에서 예측 가능한 비트 패턴의 발생이 처음 나타났으며 LCG는 계층2에서, BBS는 계층3에서 나타났다. 그리고 m-LFSR 계층4에서 4개의 예측가능한 비트가 나타났다. LCG와 BRM은 각 계층마다의 예측 가능한 비트 패턴의 발생 횟수와 증가 비율이 서로 비슷하게 나타남을 알 수 있다. 여기서 LCG는 예측 가능한 다음 비트의 발생 횟수는 적게 나타나지만 발생 시 다음 비트를 예측할 수 있는 확률이 매우 높으므로 성능이 좋은 난수열 발생기라고 볼 수 없다. 다른 난수열 발생기들과는 달리 비트 패턴의 예측 가능한 다음 비트의 발생 횟수가 매우 적게 나타나고 있음을 볼 수 있는데, BRM은 다른 난수열 발생기들에 비해서 각 발생 횟수에 대해 다음 비트를 예측할 수 있는 확률이 상위 계층에서부터 매우 높게 나타나고 있다. 즉, 하위 계층으로 내려갈수록 예측 가능한 비트 패턴에 대해 다음 비트를 예측할 수 있는 확률이 높아지게 되나 하위 계층으로 내려갈수록 실제 그러한 비트 패턴의 발생이 어렵다는 점에 유의하자. 그리고 각 스트링의 전역적인 성향에 대한 평가표에서는 각 스트링이 모두 국소적인 비난수 성향을 가짐을 알 수 있었으며 LCG가 가장 강한 국소적인 비난수 성향을 보였고 m-LFSR, BBS가 비교적 강한 난수 성향을 보였다. 본 논문에서 사용된 m-LFSR는 128 차수가 적용되었는데 이는 다른 난수 생성기(BRM)에 사용된 것보다 매우 긴 주기를 갖도록 설계되었으므로 일반적으로 계산량적으로 안전하다고 하는 BBS 발생기와 비슷한 성향을 보이고 있다.

6. 결론

난수발생기로부터 얻어지는 난수열은 모의실험, 샘플링, 수치해석, 컴퓨터프로그래밍, 암호학 등의 여러 분야에 사용된다. 현실에서 실제적으로 사용되는 수열은 난수발생기로부터 얻어진 수열로서 이를 의사랜덤수열(quasi-random sequence)이라고 부르고 응용에 따라 이 난수열의 randomness는 중요한 의미를 지닌다. 따라서 의사랜덤수열이 랜덤하다는 것을 보장할 통계적 검정이나 이론적 기준이 필요하게 된다.

특히 이러한 난수열이 스트림 암호시스템과 같은 키 스트림 수열 및 키 수열로 사용될 경우 난수열의 randomness는 암호시스템의 안전성평가와 직결된다.

본 논문에서는 난수열이 가지고 있는 여러 확률 및 통계적 특성들을 살펴보고 주어진 수열의 randomness여부를 판정하는 새로운 통계적 검정법을 소개하였다.

만약 랜덤 비트 발생기에 의해 생성된 스트링이 본 논문에서 소개한 새로운 통계적 검정을 통과한다면 표준적인 통계 검정인 도수, 계열, 포커 검정을 모두 통과한다는 점에서 새로운 통계적 검정을 일반적이라 할 수 있다. 그리고 새로운 통계적 검정의 결과를 이용하여 스트링의 국소적, 전역적 성향에 대한 평가를 히스토그램을 통하여 살펴볼 수 있었다. 새로운 통계적 검정에 대하여 본 논문에서는 이진 스트링에 대해 평가가 이루어 졌으나 이를 임의의 난수열로도 확장이 가능하다. 또한 입력 스트링의 랜덤성 평가시 임의의 길이를 가지는 모든 스트링에 적용될 수 있으며 응용에 따라 다른 유의 수준을 적용할 수도 있다. 위의 시뮬레이션에서 적용된 계산들은 요구되는 결과를 얻는 데에 근사값을 사용하였으므로 정확하고 향상된 결과를 얻기 위해서는 좀더 많은 계산이 이루어져야 할 것으로 판단된다.

참 고 문 헌

[1] A. Schrift and A. Shamir, " Universal tests for nonuniform distributions " Journal of Cryptology, Vol. 6, No. 3, 1993, pp. 119-113.

[2] A. Yao, "Theory and application of trapdoor function", proc. 23rd FOCS, pp. 88-91, 1982.

[3] B. Sadeghiyan and J. Mohajeri, " A new universal test for bit strings", ACISP'96, pp.311-320, 1996.

[4] B. Schneier, "Applied Cryptography", 2nd, John Wiley & Sons, 1996.

[5] D. E. Knuth, The Art of Computer Programming, Vol. 2 Semi Numerical Algorithms, Addison-Wesley Publishing Company, 1981.

[6] Douglas R. Stinson, "Crytography : theory and practice" , CRC Press, pp. 370-377, 1995.

[7] Gustavus J.Simmons, Contemporary Cryptology, the Science of Information Integrity, IEEE Press, New York, 1992.

[8] L. Brown et al., "A generalized test-bed for analyzing block and stream ciphers", Proc. of IFIP 1991/Information Security, 1991.

[9] U. Maurer, "A universal statistical test for random bit generators", Journal of Crytology, vol. 5, No. 2, pp.89-105, 1992.