난수열에 대한 새로운 통계적 검정

김혜정 , 이경현
부경대학교 전자계산학과

A new statistical test for random sequences

Hea-Jeong Kim , Kyung-Hyune Rhee
Department of Computer Science, Pukyong National University

요 약

본 논문에서는 여러 난수열 발생기들의 안전성 평가를 위한 새로운 통계적 검정법을 소개한다. 검정에서 구현된 기본 개념은 다음 비트 검정 이론을 바탕으로 하였으며 전체 스트링과 스트링의 일부분에 관한 확률적 통계치가 주어진다면 이를 이용하여 추측할 수 있는 다음 비트들에 관한 정보를 얻을 수 있게 된다는 점을 이용하였다. 본 검정에서는 난수 발생기의 랜덤성 평가시 입력되는 스트링 크기의 크고 작음에 관계없이 모든 임의 길이의 스트링에 적용될 수 있도록 하였으며 이는 난수 발생기를 이용한 암호 시스템의 안전성 평가에 있어서 매우 유용하게 사용될 수 있을 것이다.

1. 서론

의사 랜덤 비트 생성기는 마치 랜덤하게 생성되는 것 같으나 실제로는 결정적으로 비트 스트링을 생성시키는 비트 생성기이다. 즉, 의사 랜덤 비트 생성기는 일양적 분포를 가지며 확률적으로 독립인 성질을 만족한다.

비트 생성기들에 대해 다음 비트 검정(next bit test)이 일반적인 검정으로서의 성질을 가진다는 것이 Yao[1]에 의해 증명되었다. 다음 비트 검정은 비트 발생기에 의해 생성되어진 스트링의 임의의 i 비트들에 대해 1/2 이상의 성공 확률을 가지고 i 비트 스트링의 다음 비트를 예측하려는 것이다. 이러한 다음 비트를 효율적으로 예측하는 것이 불가능한 경우의 비트 발생기를 의사 랜덤이라고 한다. 비트 발생기가 본 논문에서 소개된 검정을 통과할 경우 비트 발생기들에 대한 어떠한 다른 검정도 통과한다는 관점에서 이 검정을 일반적이라고 평가할 수 있다.

실제로, 많은 다른 검정들이 비트 발생기에 의해 생성된 스트링들에 관한 수행 평가를 위해 적용되고 있다. 이와 같은 실제적인 검정들은 크게 복잡성 검정과 확률적 검정의 두 그룹으로 나뉘어 질 수 있다. 복잡성 검정은 전체 스트링을 재구성하기 위해 생성된 스트링의 얼마만한 길이가 요구되어지는 지를 평가하는 것이고 확률적 검정은 비트 발생기가 특정한 확률 모델에 의해 수행되는지 안 되는지를 평가하기 위한 것이다. 이러한 두 가지 검정을 모두 만족한다면 랜덤성이 좋은 비트 발생기로서 평가되어질 수 있다.

이 논문은 다음 비트 검정의 개념을 바탕으로 한 실제적인 검정의 구현과 여러 난수 발생기에 대한 직접적인 적용 결과를 분석한 것이다. 또한 검정 적용 시에 반드시 많은 양의 임의 비트들이 필요한 것은 아니며 임의의 길이를 가진 모든 스트링에 적용되어 질 수 있도록 하였다.

2. 일반적인 통계 검정

어떤 비트 스트링이 스트링 난수 발생기에 의해 생성되었는지 아닌지 즉, 각 비트들이 서로 독립적이고 편향되었는지 아닌지를 평가하기 위해 여러 가지 종류의 검정들이 스트링에 대해 수행된다.

2.1 빈도 검정(Frequency test)

가장 일반적이고 실제적인 확률적 검정으로는 난수 발생기에 의해 발생된 스트링이 편향되었는지 아닌지를 결정하는데 사용되는 빈도 검정이다. 빈도 검정은 비트 발생기에 의해 발생된 스트링이 확률적으로 독립이고 일양적으로 분포된 확률 변수를 가지게 되는 binary memoryless source(BMS) 모델을 기초로 하고 있으며 1의 발생 확률이 p= $\frac{1}{2}$로 규정되어 있다. 여기서는 0과 1의 수를 각각 n_0과 n_1로 표현한다. 그러면 빈도 검정에서 검정 통계량 χ^2의 값은 다음과 같이 계산된다.

$$\chi^2 = \frac{(n_1 - n_0)^2}{n}$$

만일 $n_0 = n_1 = \frac{n}{2}$이라면 $\chi^2 = 0$이 되며 관측도수와 기대도수의 출현 빈도간의 차이가 클수록 χ^2의 값은 더 커지게 된다. 만일 유의수준 5%에 대해 χ^2값이 3.84보다 더 커지지 않는 다면 스트링은 난수 스트링으로 받아들여지게 된다. 이 때의 χ^2값은 자유도 1을 가지는 chi-squared 테이블을 참조한 것이다. χ^2= 3.84라는 것은 $\frac{(n_1 - n_0)^2}{n}$ = 3.84임을 의미한다.

$n_1 \geq n_0$일 경우, $n_1 = n - n_0$이므로 $\frac{n_1}{n} = \frac{1+\sqrt{\frac{\chi^2}{n}}}{2}$이 된다. n=1000 비트일 경우를 고려하면 최대 $\frac{n_1}{n}$ = 0.53으로 계산되어질 수 있다. 편향값 b가 $0.47 \leq b \leq 0.53$의 값을 가지는 임의의 1000-비트 이진 스트링이 이 검정을 통과하는 것은 의미가 있다.

2.2 계열 검정(Serial test)

또 다른 일반적인 검정은 비트쌍이 고려되어지는 계열 검정이다. 비트쌍은 00, 01, 10, 11중의 어느 하나가 될 수 있다. 이 검정에서는, 스트링상에서 이러한 패턴의 발생 횟수가 측정되고 각각 n_{00}, n_{01}, n_{10}, n_{11}로 표현된다. 그러면 χ^2은 다음과 같이 계산되어진다.

$$\chi^2 = \frac{4}{n-1}\sum_{i=0}^{1}\sum_{j=0}^{1}(n_{ij})^2 - \frac{2}{n}\sum_{i=0}^{1} n_i^2 + 1$$

만일 $n_{00} = n_{01} = n_{10} = n_{11} = \frac{n}{4}$이면 χ^2= 0이되며 관측 도수와 기대도수의 차이가 크면 클수록 χ^2의 값은 더욱 커진다. 만일 유의수준 5%에 대해 χ^2값이 5.99보다 더 크지 않다면 스트링은 난수 스트링으로 받아들여진다.

$n_{11} \geq n_{01}$인 경우를 고려해 보자.

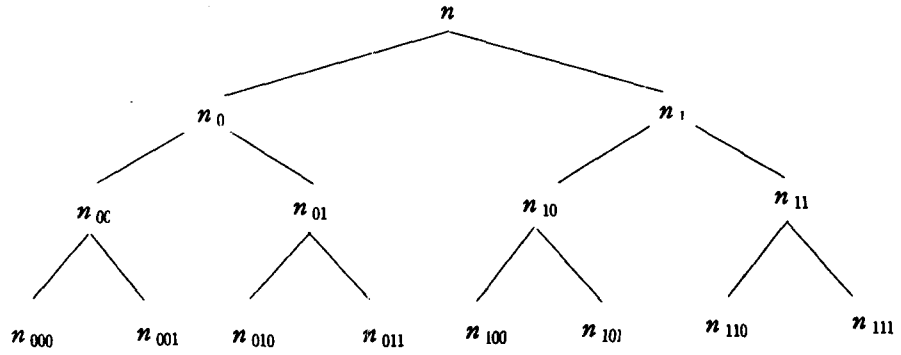
χ^2=5.99일 때, 기껏해야 $\frac{n_{11}}{n} = \frac{\sqrt{4.49n - 4.99}}{2n} + \frac{1}{4}$이 된다.

2.3 포커 검정(Poker test)

포커 검정은 스트링 상에서 m 비트의 패턴이 고려되어 지는 검정이다. 포커 검정에서는 2^m의 서로

다른 패턴이 존재한다. 의사 랜덤 스트링에 대해서는 각각의 패턴들이 약 $\frac{1}{2^m}$의 확률로 발생하는 것을 예측할 수 있다.

위 검정들에서의 발생 횟수에 대해서는 아래와 같은 트리 구조로 살펴볼 수 있다. 여기서 각 노드들은 각 패턴의 발생 횟수를 나타내고 있다.



(그림 1) 발생 횟수에 따른 트리 구조

각 노드의 수는 하위 계층에서 각 노드에 연결된 두 개의 하위 노드에 나타나는 수의 합에 의해 계산되어 질 수 있다. 예를 들어 $n_0 = n_{00} + n_{01}$ 또는 $n_{00} + n_{01} + 1$이 된다. [8]에서 언급되었듯이 스트링 검정에 대해 필요한 결과를 얻기 위해서는 각 패턴이 적어도 5회 정도는 나오도록 스트링의 길이가 선택되어 져야 한다.

3. Universal 검정의 이론

Schrift와 Shamir의 결과 [7]에서 사용되어진 표기에 따라 본 논문에서도 아래와 같은 기호와 정의들이 사용되어 진다.

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

s_i : 스트링의 i번째 비트

s_j^k : j번째 비트에서부터 k번째 비트까지의 스트링

$O(\nu(n))$: $\nu(n)$의 복잡도를 가지는 Big-Oh 표기

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

정의 1 : 수열 S_n이 $\{0,1\}^n$상에서의 확률 분포이면 앙상블 S는 수열 $\{S_n\}$이 된다.

정의 2 : S_n이 모든 n에 대해 일양 확률 분포이면 앙상블 S는 일양적이라 한다. 즉, 모든 $\alpha \in \{0,1\}^n$에 대해 $Prob\{s_1^n = \alpha\} = \frac{1}{2^n}$이다.

단 $Prob_s(E)$: 확률 분포가 앙상블 입력 스트링 S에 의해 정의되어 질 때 사건 E가 발생할 확률.

정의 3 : Next Bit Test

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

Schrift와 Shamir[7]는 편향된 입력 스트링에 대해 "예측 또는 검정 통과(Predict or Pass Test)"라는 검정을 소개했는데 여기에서 편향된 입력 스트링은 다음과 같이 정의된다.

정의 4 : 모든 i에 대해 $Prob_S\{s_i=1\}=b$이라면 입력 스트링 S는 고정된 편향 b를 지니고 1의 방향으로 편향되어진다($\frac{1}{2}\leq b<1$). 만일 모든 비트들이 독립적이라면 입력 스트링은 독립된 편향이라 한다.

정의 5 : POP(Predict or Pass) Test

만일 모든 i($1<i\leq n$), 모든 고정 값 c, 모든 확률적 다항식 시간 알고리즘 A: $\{0,1\}^{i-1}\rightarrow\{0,1,?\}$에 대해서, $Prob[A(s_1^{i-1})\neq ?]\geq\frac{1}{n^c}$이면

$|Prob_S\{A(s_1^{i-1})=s_i \mid A(s_1^{i-1}\neq ?)\}-b|\leq O(\nu(n))$이고 편향된 입력 스트링 S는 POP 검정을 통과한다고 정의한다.

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

정리 1 : 입력 스트링이 완전히 독립적인 편향된 스트링이 될 필요충분 조건은 POP 검정을 통과하는 것이다.

정의 6 : 임의의 확률적 다항식 시간(구별) 알고리즘 D : $\{0,1\}^n\rightarrow\{0,1\}$ 에 대해서 $|Prob(D(S_1)=1)-Prob(D(S_2)=1)|\leq O(\nu(n))$ 이라면 두 개의 입력 스트링 S_1과 S_2는 다항식적으로 구분 불가하다.

정의 7 : 편향 b를 가지는 독립적으로 편향된 입력 스트링 B로부터 소스 S가 다항식적으로 구분 불가라면 스트링 S는 고정 편향 b를 가지는 완전 독립 편향된 입력(perfect independence biased source)이 된다.

4. 새로운 검정

위에서 살펴 본 확률적 검정과 POP 검정이 서로 연관되어질 수 있다는 데에 바탕을 두고 지금부터 실제적인 새로운 일반적 검정을 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

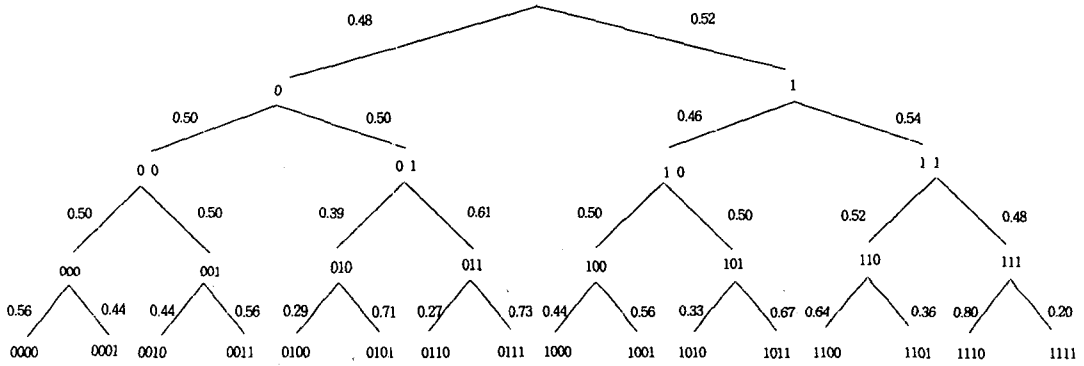
상에서 발생하는 패턴의 수에 대한 하위 계층에서 발생하는 패턴의 수의 확률 비를 적어내려 가면서 트리를 구성해 간다.

계산상의 복잡성을 피하기 위해 계층 l 에서의 패턴 비교를 위해 스트링의 처음 l-1 비트를 스트링의 끝 부분에 덧붙여 준다. 상위 계층에서 각 패턴들의 수는 그 아래 두 노드들 간에 직접적으로 관련된 패턴들의 수의 합으로부터 계산되어질 수 있다.

충분히 큰 랜덤 스트링에 대해서, 모든 확률 비들이 $\frac{1}{2}$에 가까워 질 것이라고 예상한다. 그러나, 의사 랜덤 스트링에 대해서는 항상 약간의 편차가 허락되어 진다. 예제1의 스트링에 대해 이 트리의 구성이 (그림2)에 주어졌다.

예제 1. 다음과 같은 이진 스트링이 주어졌을 경우 위의 new 검정에서 제안한 방식으로 비트의 발생 패턴과 발생 횟수에 따른 확률 비를 적어내려 가면서 트리를 구성해 보자.(n=75)

010101110010111001111011100110000100100111000110101110000010111101100000111



(그림2) 예제1의 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 발생 횟수는 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 알려 준다.

예를 들어, 패턴 1010이 예제1의 스트링에서 나타난다면 1의 확률로 다음 비트는 1이 됨을 알 수 있다.

75-비트 스트링에 대해서 만일 편차 값이 $0.3829 \le b \le 0.6171$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 8 : s_1^n은 길이 n을 가지는 스트링이다. 결정 문턱치(threshold decision) α는 다음과 같이 정의된다.

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

단 χ^2는 검정에서 요구되는 유의수준에 대응되는 값.

de bruijn 수열과 같이 높은 복잡도를 가지는 랜덤 수열들에 대해서 $n = 2^l$일 때, $l = \log_2(n)$인 계층에서 각 패턴은 한번씩 발생한다. 하지만 좀더 적은 복잡도를 지니는 수열들에 대해서는 몇몇 패턴들은 계층 l 에서 발생하지 않고 나머지 패턴들은 한 번이나 한 번 이상 발생한다. 이것은 계층 l-1의 노드에서 계층 l 까지의 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.

재구성에 관한 문턱치 길이(threshold length)를 정의하기 위해 POP 검정의 확장인 아래와 같은 정의와 정리를 고려해 보자.

정의 9 : Extended POP Test

각 i, $l(1<i, l\leq n)$, 고정된 상수 c, 모든 확률적 다항식 시간 알고리즘 A : $\{0,1\}^{i-1} \rightarrow \{\{0,1\}^l, ?\}$에 대해 편향된 소스S는 만약 $Prob\{A(S_1^{i-1}) \neq ?\} \geq \frac{1}{n^c}$일 때,

$|Prob\{A(s_1^{i-1}) = s_i^{i-1+l} \mid A(s_1^{i-1}) \neq ?\}| \leq O(\nu(n))$이 성립한다면,확장된-POP 검정을 통과한다.

정리 2 : 다음의 조건들은 동치이다.

i) 편향된 소스는 완전 독립적 편향된 소스이다.

ii) 편향된 소스는 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 되도록 정의할 수 있다. 스트링 상의 다음 비트가 편향치 보다 더 높은 확률을 가지고 나타날 때, 즉 $b \leq p$ or $p \leq 1-b$, 알고리즘은 스트링 상의 그 비트를 예측할 수 있다. 스트링 상의 다음 비트가 편향치보다 적은 확률을 가지고 나타날 때, 즉 $1-b \leq p \leq b$,일 때 알고리즘은 ?을 출력한다.

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

Step 1. 결정 문턱치 α 값을 다음과 같이 계산한다.

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

Step 2. $l = round(\log_2(n))$을 계산한다.

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

Step 4. 각각의 블록을 비교해 나가면서 길이 l 을 갖는 각 패턴의 발생 횟수를 계산하라.

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

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

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

위의 알고리즘을 사용하여 스트링이 국소적으로 랜덤하지 않은 경향과 전체적으로 랜덤하지 않은 경향을 다음의 방법으로 평가할 수 있다.

(1) Local non-random 성향 : 만약 계층 l - 1에서 l + 1보다 많은 비트들이 예측되어 질 수 있는 임의의 노드가 존재한다면 다음 블록이 예측될 수 있는 길이 l 의 블록이 존재함을 의미한다. 따라서, 스트링 상에 local non-randomness가 존재하고, 스트링은 요구되는 성질을 만족하는 생성기로는 기각되어질 수도 있다.

위의 검정은 상위 계층의 가지들에 나타나는 확률이 $\frac{1}{2}$로부터의 편차가 더 낮은 가지들에 더 많은 편차를 유발시킬 것이라는 관점에서 만능 검정의 의미가 있다. 즉, 새로운 검정에서 편차는 스트링에 가중되어지는 local non-random behavior로서 나타내어질 것이며 반면에 기존의 통계적 검정의 각 종류는 트리의 상위 계층 가지들에서 $\frac{1}{2}$로부터의 편차의 측정치에 대한 것이다. 따라서 스트링이 새로운 검정을 통과한다면, 기존의 빈도, 계열 및 포커 검정과 같은 표준적인 통계적 검정을 통과하게 된다.

5. 시뮬레이션 및 결과 분석

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 J-K 플립플롭

J-K 플립플롭은 m-LFSR 2개의 출력을 J-K 플립플롭에 의해 조합하여 출력수열을 발생하는 시스템이다.

J-K플립플롭을 수식화하면

$q_n = a_n \oplus q_{n-1}(1 \oplus a_n \oplus b_n)$ (단, $q_{-1}=0$)이다.

여기서 (a_n):m-LFSR1의 출력수열, (b_n):m-LFSR2의 출력수열이다.

J-K플립플롭:

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

key : initial states $a_0^{(1)}$, $a_0^{(2)}$ of the 2LFSRs.

For i = 0,1,2,...do

a. Shift each LFSR

b. Compute kth J-K flip-flop function for correspond-ing pair of LFSRs

$y_i = a_i^{(1)} \oplus y_{i-1}(1 \oplus a_i^{(1)} \oplus a_i^{(2)})$

c. Collect four keystream bits as $z_{i+k}=y_i^{(k)}$

Output : the sequence of z_i, l=1,2,...

5.2 SUMMATION Generator

Summation Generator[2]는 입력 수열은 2진수이며, SUM 함수 $f: Z^N \rightarrow Z$, $Z=\sum_{i=1}^{N} x_i$로 정의하며 최하위 비트에서부터 비트 연산을 한다. 정수 덧셈은 F_2상에서 높은 비선형성을 가지며, 상관관계 공격에 대한 강인한 면역성을 갖는다는 사실에 근간을 두고 연구하였다.

SUMMATION Generator:

Input : parameters : 2 LFSR $< L_j, C_j(D)>$

Key : initial states of the N LFSRs and carry C_0

For i=1,2,.....do

1. Step each shift register once to produce
   $x_{1i}, x_{2i}, \ldots, x_{Ni}$.
2. Compute the integer sum
   $$S_i=\sum_{k=1}^{N} x_{ki}+ C_{i-1}$$
3. Set
   $$z_i= S_i \mod 2$$
   $$C_i= \lfloor \frac{S_i}{2} \rfloor$$

Output: the sequence z_i, $i=1,2,\ldots$

5.3 MUX 시스템

MUX 시스템[2]은 멀티플렉서와 2개의 m-LFSR로 구성된다. m-LFSR1과 m-LFSR2의 차수가 각각 m, n이고 m-LFSR1의 단이 $A_0, A_1, \ldots, A_{m-1}$, m-LFSR2의 단이 $B_0, B_1, \ldots, B_{n-1}$, 그리고 m-LFSR1의 출력 수열을 (a_n), m-LFSR2의 출력 수열을 (b_n)이라고 하자. 먼저 1 < h<m인 정수 h 를 선택한 후, m-LFSR2의 n 단 중에서 2^h단을 선택한다. 시간이 t 일때 MUX 시스템의 출력 u_t 는 m-LFSR1의 h 단의 내용에 의해 m-LFSR2의 2^h단 중 한 단의 내용으로 결정된다.

Mutiplexer generator:

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

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

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

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
   $$a_t=\sum_{k=0}^{h-1} 2^k s_i^{(1)}(j_k)$$
3. Extract
   $z_i= s_i^{(2)}(\theta(a_i))$ (θ 는 {0,1,..., 2^h-1)에서 {0,1,....,
   L_2-1 }로의 대응 함수)

Output : the sequence of z_i, i=1,2,...

5.4 BRM 시스템

BRM 시스템[2]은 2개의 m-LFSR과 BRM(BinaryRated Multiplexer)으로 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

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

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

   Shift LFSR2

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

5.5 난수 발생기들의 성능 평가 결과

아래에는 각각의 난수 발생기의 종류에 따라 다음 비트를 예측 가능한 비트들의 발생 횟수를 표로 나타내었다. 입력 스트링의 길이는 2500으로 잡았으며 유의 수준 5%(결정 문턱치 값 : 0.52)와 유의 수준 1%(결정 문턱치 값 : 0.53) 각각에 대한 경우를 <표1>과 <표2>에 나타내었다.

	J-K 플립플롭	SUMMATION	MUX	BRM
계층1	1	0	0	1
계층2	2	0	1	2
계층3	4	3	4	4
계층4	5	7	8	7
계층5	9	11	14	12
계층6	14	24	32	26
계층7	20	54	60	58
계층8	26	114	119	119
계층9	30	232	194	222
계층10	41	390	285	370
계층11	45	591	370	545

<표1> 난수 발생기의 예측 가능한 다음 비트 발생 횟수(α : 5%)

	J-K 플립플롭	SUMMATION	MUX	BRM
계층1	1	0	1	1
계층2	2	0	1	2
계층3	4	2	4	4
계층4	6	7	8	7
계층5	9	11	14	10
계층6	14	21	31	26
계층7	20	51	58	57
계층8	26	110	119	115
계층9	31	231	193	221
계층10	41	390	285	370
계층11	47	591	370	545

<표2> 난수 발생기의 예측 가능한 다음 비트 발생 횟수(α : 1%)

위의 검정에는 J-K 플립플롭, SUMMATION Generator, MUX 시스템, BRM 시스템이 이용되었으며 J-K 플립플롭과 BRM이 계층1에서 예측 가능한 비트 패턴의 발생이 처음 나타났으며 MUX은 계층2에서 나타났다. 그리고 SUMMATION Generator에서는 계층3에서 3개의 예측 가능한 비트 패턴이 처음으로 발생하였다. SUMMATION Generator, BRM 시스템에서는 각 계층마다의 예측 가능한 비트 패턴의 발생 횟수와 증가 비율이 서로 비슷하게 나타남을 알 수 있다. 여기서 J-K 플립플롭은 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, J-K 플립플롭은 다른 난수열 발생기들에 비해서 각 발생 횟수에 대해 다음 비트를 예측할 수 있는 확률이 상위 계층에서부터 매우 높게 나타나고 있다. 즉, J-K 플립플롭을 제외한 난수열 발생기들은 하위 계층으로 내려갈수록 예측 가능한 비트 패턴에 대해 다음 비트를 예측할 수 있는 확률이 높아 지게되나 하위 계층으로 내려갈수록 실제 그러한 비트 패턴의 발생이 어렵다는 점에 유의하자.

본 검정은 입력 스트링의 랜덤성 평가 시 임의의 길이를 가지는 모든 스트링에 적용될 수 있다. 본 논문에서는 5%와 1%의 유의 수준에 대해서 각각 테스트가 이루어 졌으며 스트링 길이를 2500으로 잡은 본 검정에서는 결과적으로 큰 차이를 보이지 않았다. 애플리케이션에 따라 다른 유의 수준을 적용할 수도 있다. 위의 시뮬레이션에서 적용된 계산들은 요구되는 결과를 얻는 데에 약간의 오차를 허용하였으므로 정확하고 향상된 결과를 얻는 데에 적용되어질 수 있다.

참고 문헌

[1] A. Yao,"Theory and application of trapdoor functions", Proc. 23rd FOCS, pp. 80-91, 1982.

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

[3] H. Gustafson et al., "Comparison do block ciphers", in Lecture Notes on Computer Science, Vol.453, Proceedings of Auscrypt '90, pp.153-165, 1990.

[4] J. Ziv, "Compression tests for for randomness and estimating the statistical model of an individual sequences", Sequences, pp.366-373, 1990.

[5] H.Beker and F. Piper, "Ciper Systems", Northood Books, 1982.

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

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

[8] D. Knuth, "The Art of Computer Programming", Vol.2, Addison-Wesley, 1973.

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

[10] 성둘옥, 신상욱, 이경현, "비선형 로직의 통계적 검정", 정보처리학회 논문지, Vol. 3, No. 2, pp.225-230, 1996.

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