한국지능시스템학회 논문지 2010, Vol. 20, No. 1, pp. 76-80

이진 알고리즘을 이용한 변형 시리얼테스트 설계에 관한 연구

Design variation serial test using binary algorithm

최진석 · 이성주+

Jin-suk Choi and Sung-Joo Lee+

조선대학교 전자계산학과, +컴퓨터공학부

요 약

급변하는 정보의 홍수 속에서 정보의 보안과 이를 가공하고 전송하는 것이 중요한 과제로 떠오르고 있다. 초기 정보보호이론과 암호화 전송단계에서는 간단한 치환과 수학적 계산 알고리즘을 적용한 암·복호화 과정을 이용하였다. 완벽한 정보보호는 One-time pad를 이용하는 것이나 이를 적용하기에는 하드웨어와 금전적 손실이 너무 크기에 실난수가 아닌 난수성을 만족하는 의사난수를 사용하고 있다. 본고에서 제안하는 변형 시리얼 테스트는 의사난수성을 입증하는 테스트 중 시리얼테스트에서 변형된 것으로 연산속도와 효율성 면에서 보다 더 강력한 난수성임을 입증하고 있다.

키워드 : 알고리즘, 의사난수, 시리얼테스트, 암호화/복호화

Abstract

It is floating to security of information and the early assignment that it is important it processes and to transmit in inundations of information that I changed suddenly. I used the encryption/decryption process that applied simple substitution and mathematical calculation algorithm at theory and encryption transmission steps protective early information. Hardware and financial loss are using spurious random number to be satisfied with the random number anger that isn't real random number to size so much perfect information protection using One-time pad for applying this. I was transformed into serial test under a test to prove spurious random number anger, and it is into random number anger stronger, and the transformation serial test that proposes is proving it in algorithm speed and efficiency planes.

Key Words : algorithm, Pseudo-random numbers, serial test, encryption/decryption

1. 서 론

현대 암호 알고리즘은 크게 공개키 방식과 비밀 키 방식으로 나누어진다. 비밀 키 방식은 암호문을 전달하기 전에 먼저 안전한 통로를 사용하여 비밀 키를 전달해야 하는 약점이 있다. 그러나 공개키 방식의 경우 키의 전달이 필요 없이 암호문을 전달한다. 즉 자기만의 개인키를 사용하여 암호문을 해독할 수 있다. 이런 공개키 계의 생각은 1976년에 처음 발표된 후 1977년에 처음으로 RSA(Rivest, Shamir, Adleman) 암호계가 실현되었다. 이후에 여러 공개키 암호계가 수학적 안전성을 기반으로 발표되었다.[1,2,3,4]

일반적으로 많이 사용되는 알고리즘은 다음과 같다.

(1) DES(The Data Encryption Standard)는 IBM에서 LUCIFER를 수정하여 개발한 비밀 키 알고리즘이다.

DES는 길이 64인 비트문자열 암호문을 만들기 위하여 길이 56비트문자열인 적당한 키를 사용하여 길이 64인 평문비트문자열을 암호화한다. DES에서 수행되는 유일한 산술은 비트문자열의 배타적논리합이기 때문에 하드웨어적으로 또는 소프트웨어적으로 매우 효율적으로 실행할 수 있다. 하지만 키 공간의 크기가 256bit로 너무 작아 전수조사(exhaustive search)등 여러 가지 공격법들이 나와 있다.[5,8]

(2) AES(Rijndael)은 벨기에의 암호학자인 Joan Daemen과 Vincent Rijmen이 개발한 알고리즘으로 DES와 3DES가 Feistel network이었던 것과는 달리 substitution permutation network 구조를 가지고 있으며 하드웨어나 소프트웨어적으로 구현했을 때 모두 좋은 성능을 보이는 특성을 가지고 있다. 구현이 쉽고 메모리를 적게 소모하는 장점으로 Rijndael은 암호알고리즘 분류 상 대칭형암호알고리즘으로 분류된다. 대칭형 암호알고리즘에도 두 가지 종류가 있는데 Rijndael은 이 중에서 블록 암호알고리즘 방식이며 128bit 블록 단위로 암호화를 하고, 사용되는 키의 사이즈는 128bit, 192bit, 256bit등이 있다. 라운드 수는 각각 10, 12, 14라운드를 사용하고 각 라운드 마다 SubBytes, ShiftRows, MixColumns, AddRoundKey의 4단계를 거치게 된다. 현재 AES의 취약점에 대한 연구가 활발하게 이루어지고 있으며 관련하여 XSL Attack이나 cache timing attack 등에 관련된 논문이 발표되었다.[9,10]

접수일자 : 2008년 10월 22일
완료일자 : 2009년 12월 20일
+ : 교신저자
이 논문은 2009학년도 조선대학교 연구비의 지원을 받아 연구되었음

(3) RSA(Rivest, Shamir, Adleman)는 1978년 MIT에서 개발된 공개키 알고리즘이다. 이 알고리즘은 큰 소수가 소인수분해하기 어렵다는 수학적 사실에 기반을 두고 있다. 주로 암호화와 전자서명에 사용된다.[5]

(4) DSA(Digital Signature Algorithm)는 1991년 NIST(National Institution of Standard and Technology)가 미국 전자서명 표준 (DSS, Digital Signature Standard)에서 사용하기 위하여 발표한 정부용 공개키 알고리즘으로 전자서명에만 사용된다.[7]

2. 의사난수

의사난수는 거의 모든 암호학적 알고리즘에서 뼈대와 같이 사용되는 가장 중요한 암호학적 알고리즘 중의 하나이다. 의사난수는 스트림 암호의 원천을 이루고 또한 암호 프로토콜의 초기 벡터 또는 비밀 키, 전자 서명 및 전자 결제 시스템의 비밀 파라미터, 각종 키 관리/인증메커니즘에서의 세션키나 랜덤 첼린지(random challenge)의 생성 등에 사용된다. 난수에는 크게 실 난수(True random) 와 의사난수(Pseudo random) 두 가지로 볼 수 있는데 그 차이점은 표 1 과 같다.

표 1. 난수의 비교
Table 1. Comparison of random number

실난수(True random numbers)	의사난수(Pseudo-random numbers)
비결정적	컴퓨터는 논리적이고 결정적이므로 실 난수를 산출 하지 못함
예측 불가능한 어떤 물리적인 소스로부터 획득 : 반도체, 방사선 붕괴 등으로부터 전자소음 혹은 열 소음 등	S/W에 기반 한 RNG는 최상의 경우에 의사난수를 생성 가능
	PRNG : 길이가 짧은 랜덤 비트 열(seed)을 길이가 긴 랜덤에 근접한 비트열로 출력하는 알고리즘

또한 난수성을 만족하는 좋은 난수의 특징은 다음과 같다.

- 무주기성(가능하면 긴 주기)
- 등확률성
- 무규칙성
- 재현성: 다시 만들어 낼 수 있어야 함
- 계산시간이 짧다

이러한 난수성을 만족하고 구현상의 용이함을 위하여 암호학적 알고리즘을 이용한 의사난수 생성기를 사용한다. 암호학적 알고리즘을 이용하면 먼저 암호학적 키의 안전도에 의하여 난수를 선택하는 기준을 정할 수 있고 암호학적 알고리즘(encryption algorithm)이나 보안 프로토콜(security protocols)에는 관심이 많으나 실상 키를 담당하는 의사난수 생성에는 소홀할 수 있는 단점을 가지고 있다. 하지만 현대의 공격자나 해커들은 복잡한 알고리즘보다 랜(ran) 공격에 주력하기 때문에 암호학적 알고리즘을 통한 의사난수 생성기는 기본적인 안전에 대한 믿음을 제공한다고 볼 수 있다.

이에 본고에서는 의사난수들의 난수성 만족을 입증하는 테스트 중 시리얼 검정에 대해 알아보고 이를 좀 더 보완하여 검정 프로토콜을 간소화 시키고 보다 강한 난수성을 입증할 수 있는 검정방법을 구현하였다.

3. 시리얼 검정

이 검정 법은 한 항이 그 다음에 0 또는 1 로 바뀌어 나가는 과정이 랜덤 한지를 조사하는 방법이다. 이진 수열 (s_t) 의 N 개의 항 s_0, s_1, s_2, s_{N-1} 중에서 0 인 것의 개수와 1인 것의 개수를 각각 a_0, a_1이라고 하자. 또 이들 N 을 연이은 2 개의 항을 차례로 연이어 두 항씩 묶어 놓은

$$s_0s_1,\ s_1s_2,\ s_2s_3,\ \cdots,\ s_{n-2}s_{n-1} \tag{1}$$

중에서 00, 01, 10, 11과 같은 것의 개수를 각각 a_{00}, a_{01}, a_{10}, a_{11}이라고 할 때, 다음의 등식이 성립한다.

$$\begin{aligned} a_0 + a_1 &= N \\ a_{00} + a_{01} &= a_0 \text{ 또는 } a_0 - 1 \\ a_{10} + a_{11} &= a_1 \text{ 또는 } a_1 - 1 \\ a_{00} + a_{01} + a_{10} + a_{11} &= N-1 \end{aligned} \tag{2}$$

한편, 각 a_i의 기대 값은 $\frac{N-1}{4}$이다.

따라서 통계량

$$P = \sum_{i,j=0}^{1} \frac{\left(a_{ij} - \frac{N-1}{4}\right)^2}{\frac{N-1}{4}} - \sum_{i=0}^{1} \frac{\left(a_i - \frac{N}{2}\right)^2}{\frac{N}{2}} \tag{3}$$

는 근사적으로 자유도가 2인 χ^2 분포를 따르므로, 위 식은 다음과 같이 단순화 될 수 있다.

$$\begin{aligned} P = {} & \frac{4}{n-1}(a_{00}^2 + a_{01}^2 + a_{10}^2 + a_{11}^2) \\ & - \frac{2}{N}(a_0^2 + a_1^2) + 1 \end{aligned} \tag{4}$$

이 통계량을 이용하여 난수성을 검정하는 방법을 시리얼 검정법이라 한다.

예를 들어 길이 10인 이진 수열 0011011101에 대하여 비트길이를 3까지로 제한한다면 3비트씩 나누어야 하므로 끝에 00을 첨가한 001101110100을 이용하여 빈도를 측정한다. 같은 방법으로 2비트열에 대한 빈도 측정은 00110111010을 사용하고 단일비트에 대한 빈도수 측정은 프리컨시 검정과 동일한 방법으로 시행한다.

n : 총 도수
m : 비트열 수(여기서는 3)
v : 각 비트의 패턴일 때

각 비트열에 대한 빈도수 측정에 대한 확률 값은 다음의 식에 의해 산출된다.

3비트:

$$\phi_m^2 = \frac{2^m}{n}\sum_{i_1\cdots i_m}\left(v_{i_1\cdots i_m}-\frac{n}{2^m}\right)^2 = \frac{2^m}{n}\sum_{i_1\cdots i_m} v_{i_1\cdots i_m}^2 - n \tag{5}$$

2비트:

$$\phi_{m-1}^2 = \frac{2^{m-1}}{n}\sum_{i_1\cdots i_{m-1}}\left(v_{i_1\cdots i_{m-1}}-\frac{n}{2^{m-1}}\right)^2 = \frac{2^{m-1}}{n}\sum_{i_1\cdots i_{m-1}} v_{i_1\cdots i_{m-1}}^2 - n \tag{6}$$

1비트:

$$\phi_{m-2}^2 = \frac{2^{m-2}}{n}\sum_{i_1\cdots i_{m-2}}\left(v_{i_1\cdots i_{m-2}}-\frac{n}{2^{m-2}}\right)^2 = \frac{2^{m-2}}{n}\sum_{i_1\cdots i_{m-2}} v_{i_1\cdots i_{m-2}}^2 - n \tag{7}$$

위 식을 이용하면 다음과 같은 결과를 얻을 수 있다.

$$\phi_3^2 = \frac{2^3}{10}(0+1+1+4+1+4+4+1)-10 = 12.8-10=2.8 \tag{8}$$

$$\phi_2^2 = \frac{2^2}{10}(1+9+9+9)-10=1.2 \tag{9}$$

$$\phi_1^2 = \frac{2}{10}(16+36)-10=10.4=0.4 \tag{10}$$

여기서, χ^2 분포에서 유의수준 5 %, 자유도 2의 한계 값은 5.99이다.

따라서 P의 통계량의 값이 5.99보다 큰 경우에, 시리얼 검정에 대하여 유의수준 5 %인 이진 수열은 난수성이 없다고 판단되어 기각한다.

4. 난수성 검정 방법 제안 및 결과 고찰

제안하는 의사난수 생성기에 의한 출력물에 대한 난수성 테스트는 먼저 난수의 기본성질인 예측 불가능성에 대한 안전도를 확보하는데 주안점을 두었다. 난수성을 만족하는 요소로써 구별불능성이 있고 선행비트들에 대한 후위 비트의 예측을 할 수 없어야 하며 기존에 나와 있는 트래픽으로부터 이후 출력 값들의 분석을 할 수 없어야 한다. 또한 주어진 난수열의 주기가 판별할 수 없어야 한다. 이를 실험적으로 검증하기에는 무리가 따르기에 확률적 방법을 이용하여 난수성에 근접하는 데이터에 대한 신뢰도를 인증한다. 기존에 나와 있는 의사난수 테스트들 중에서 프리컨시 검정, 런 검정, 시리얼검정 들은 출력 비트들에 빈도수를 측정하여 이를 통계적 분포도로 산출한다는 비슷한 특징을 지니고 있다.

이에 본고에서 제안하는 의사난수 생성기에 의한 출력 데이터의 통계적 테스트는 프리컨시, 런, 시리얼 검정의 압축된 형태의 구현방식으로 난수성에 대한 검증을 진행하였다. 제안하는 통계적 난수테스트는 선행 비트들의 출력에 대한 후위비트의 형태양상을 빈도수를 통한 측정으로 구한 다음 통계적 분포도로써 난수성 검증을 완성한다. 사용 예는 다음과 같다.

122 이진 비트열 A에서 각각의 비트열의 형태를 파악하여 이후에 나오는 비트의 상태를 평가한다.

표 2. 이진 비트열
Table 2. bit string

122 이진 비트열 A 의 8 비트 형태	
11010101	00000001
01001110	01001011
01100000	11101111
00101011	10101000
10110110	00100001
00011011	01000000
01100010	10111010
00110010	00

표 2 에서 선행 비트 1이 나온 후 1이 나올 비트들의 수를 $a1$ 이라하고 0이 나온 후 다시 1이 발생할 경우를 $a0$ 라면

$$a1 = 20$$

$$a0 = 32$$

가 된다. 이는 시리얼 테스트의 $m-2$에 해당하는 2비트 데이터의 빈도수에 대한 검정에서 “11”, “10”의 빈도수만을 측정하는 것과 마찬가지라고 볼 수 있다. 또한 선행비트 “10”, “11”, “01”, “00”이 나온 후 다시 이들 후에 1이 나올 경우의 수를 계산하면 다음과 같다.

$$a11 = 6$$

$$a10 = 18$$

$$a01 = 13$$

$$a00 = 14$$

이는 시리얼테스트의 $m-3$에 해당하는 테스트들 중에서 비트열 “111”, “101”, “011”, “001” 만의 빈도수를 측정하는 것과 마찬가지다. 또한 여기서 비트열 “111”에 대한 빈도수의 측정은 런 테스트의 3단 진행을 의미하기도 하므로 본 테스트는 런 테스트와 시리얼 테스트와 밀접한 관계가 있음을 알 수 있다. 이를 시리얼 테스트의 함수에 적용하면

다음과 같이 나타낼 수 있다.

3비트:

$$\phi_m^2 = \frac{\frac{2^m}{2}}{n}\sum_{i_1\cdots i_m}\left(v_{i_1\cdots i_m}-\frac{n}{\frac{2^m}{2}}\right)^2$$
$$= \frac{2^{m-1}}{n}\sum_{i_1\cdots i_m}\left(v_{i_1\cdots i_m}^2-\frac{n}{2^{m-1}}\right)^2 \quad (11)$$
$$= \frac{2^{m-1}}{n}\sum_{i_1\cdots i_m} v_{i_1\cdots i_m}^2 - n$$

2비트:

$$\phi_{m-1}^2 = \frac{\frac{2^{m-1}}{2}}{n}\sum_{i_1\cdots i_{m-1}}\left(v_{i_1\cdots i_{m-1}}-\frac{n}{\frac{2^{m-1}}{2}}\right)^2$$
$$= \frac{2^{m-2}}{n}\sum_{i_1\cdots i_{m-1}}\left(v_{i_1\cdots i_{m-1}}-\frac{n}{2^{m-2}}\right)^2 \quad (12)$$
$$= \frac{2^{m-2}}{n}\sum_{i_1\cdots i_{m-1}} v_{i_1\cdots i_{m-1}}^2 - n$$

1비트:

$$\phi_{m-2}^2 = \frac{\frac{2^{m-2}}{2}}{n}\sum_{i_1\cdots i_{m-2}}\left(v_{i_1\cdots i_{m-2}}-\frac{n}{\frac{2^{m-2}}{2}}\right)^2$$
$$= \frac{2^{m-3}}{n}\sum_{i_1\cdots m-2}\left(v_{i_1\cdots i_{m-2}}-\frac{n}{2^{m-3}}\right)^2 \quad (13)$$
$$= \frac{2^{m-2}}{n}\sum_{i_1\cdots i_{m-2}} v_{i_1\cdots i_{m-2}}^2 - n$$

이는 시리얼 테스트에서 한 단계 아래 값들의 연산 즉 ϕ_{m+1}^2, ϕ_m^2, ϕ_{m-1}^2 의 값과 같음을 알 수 있다. 다시 말하여 전위 비트에 의한 후위 비트의 빈도수를 측정하는 것은 전위 데이터의 영향을 받으므로 통계적 결과 치는 전위 데이터에 종속되며 검정 량은 $m+1$비트의 검정을 시행한 것과 동일한 효과를 볼 수 있다.

이를 적용 시에는 3비트 연산을 시행하면 이전 2비트까지의 시리얼검정을 정확히 통과했다는 가정 하 에 비트형의 데이터 분포에서 $\frac{1}{2}$가지의 시행만 하면 된다. 탐색 알고리즘에 있어서 n비트의 데이터에 대하여 작은 블록인 m비트열로 나눈다면 이에 대한 검색 타임은$\frac{m}{3}\times\frac{8}{2}\times\frac{n}{m}=\frac{4}{3}\times n$ 회의 타임 절약이 발생한다. 보편적으로 n = 100,000 비트 이상의 데이터에 대한 난수성을 테스트 하므로 약 13,000,000 타임의 탐색시간을 절약할 수 있다. 이를 일반화 시켜

n = 100,000 으로 놓으면

$$\frac{m}{t}\times\frac{2^t}{2}\times\frac{n}{m} = \frac{1}{t}\times 2^{t-1}\times n \quad (14)$$

(단 t : 최종 비트 수,

n : 랜덤 스트링, m : 블록길이)

이므로 최종 8비트까지의 절약되는 연산수를 계산하면 표 3 과 같다.

표 3. 비트별 효과
Table 3. The effect of bit

최종 비트수	검 색 타 임	최종 비트수	검 색 타 임
$t=3$	$\frac{4}{3}\times n = 133{,}333$	$t=4$	$\frac{8}{4}\times n = 200{,}000$
$t=5$	$\frac{16}{5}\times n = 320{,}000$	$t=6$	$\frac{32}{6}\times n = 533{,}333$
$t=7$	$\frac{64}{7}\times n = 914{,}285$	$t=8$	$\frac{128}{8}\times n = 1{,}600{,}000$

5. 결 론

본 논문에서는 if 연산자의 단조감소효과를 입증하므로 데이터양이 증가하는 현 추세의 랜덤키와 랜덤수열 생성에 있어서 보다 작은 양의 테스트 프로그램을 완성할 수 있다. 또한 스트림암호(stream cipher)를 이용한 인지감별 시스템 구축과 RFID를 이용한 유비쿼터스 시스템구축의 암호알고리즘(crypto algorithm) 설계 시 우수한 난수열을 가벼운 알고리즘 구현만으로 보다 강력한 난수열을 얻는데 탁월한 효과가 있다고 본다. 향후 실시간 데이터 암호화 기술과 통신 데이터의 암호화 구현 시 난수열의 난수성 검정을 가볍게 구현하고 프로그램 양을 감소하여 전체적인 시스템 블록의 단순화를 이룰 수 있을 것으로 보인다.

참 고 문 헌

[1] H. Feistel, “Block Cipher Cryptographic System,” U.S. Patent #3,798,359,19 Mar 1974.

[2] N.Goots, B. Izotov, A. Moldovyan, and N. Moldovyan, *Modern Cryptography: Protect Your Data with Fast Block Ciphers*, A-LIST, 2003.

[3] L. R. Knudsen, “Block Ciphers-Analysis, Design and Applications,” Ph.D Thesis, Computer Science department, Aarhus University, 1994.

[4] A. Menezes, P. van Oorschot, S. Vanstone, *Handbook of Applied Cryptography*, CRC Press, 1996.
[5] B. Schneier, *Applied Cryptography, Second Edition: Protocols, Algorthms, and Source Code in C(cloth)*, 1996.
[6] B. Schneier, "The uses and Abuses of Biometrics," *Communications of the ACM*, vol. 42, no. 8, p136, 1999.
[7] D. Stinson, *Cryptography: Theory and Practice*, CRC Press, 1995.
[8] NIST, "Data Encryption Standard(DES)," http://www.itli.nist.gov/fipspubs/fip46-2.htm
[9] NIST, "Instruction-level Parallelism in AES Candidates," http://cs-www.ncsl.nist.gov/CryptoToolkit/aes/round1/conf2/papers/clapp.pdf
[10] NIST, "Federal Information Processing Standards Publication 197 - Specification for the Advanced Encryption Standard(AES)," http://csrc.nist.gov/publications/fips/fips197/fips-197, pdf, 2001.

저 자 소 개

최진석(Jin-Suk Choi)
1998년 : 경희대학교 대학원 수학과 (이학석사)
2001년~현재 : 조선대학교 전자계산학과 박사과정

관심분야 : 퍼지 이론, 소프트컴퓨팅, 양자암호
E-male : cjspro1525@naver.com

이성주(Sung-Joo Lee)
1970년 : 한남대학교 물리학과(이학사)
1992년 : 광운대학교 전자계산학과 (이학석사)
1998년 : 대구가톨릭대학교 전자계산학과 (이학박사)
1988년~1990년 : 조선대학교 전자계산소 부소장
1995년~1996년 : 조선대학교 산업대학장
1995년~1997년 : 조선대학교 정보과학대학장
1981년~현재 : 조선대학교 컴퓨터공학부 교수

관심분야 : 소프트웨어공학, 프로그래밍언어, 객체지향 시스템 러프집합, 신경망
E-male : sjlee@chosun.ac.kr