情報保護學會論文誌
第11券 第6號, 2001. 12

AES(Advanced Encryption Standard) 안전성 평가에 대한 고찰*

조 용 국**, 송 정 환**, 강 성 우***

Criteria for Evaluating Cryptographic Algorithms based on Statistical Testing of Randomness

Yong Kuk Cho**, Jung Hwan Song**, Sung Woo Kang***

요 약

본 논문에서는 미국 NIST(National Institute of Standards & Technology)의 AES(Advanced Encryption Standard) 선정기준 중 안전성 평가인 난수검정에 대하여 고찰하고자 한다. 암호 알고리즘의 안전성 평가는 입출력문과 키의 크기, 평문과 암호문 및 키와 암호문의 상관성, 평문과 키의 변화에 따르는 암호문의 변화 그리고 구조적 특이성 등이 고려대상이 된다. 주어진 암호 알고리즘에 대한 안전성 필요충분조건 만족여부를 평가하는 것은 어려우며 객관적인 평가를 위해서는 정량적인 평가결과가 도출되어야 한다. 본 논문에서는 NIST에서 실시한 AES 안전성 평가항목들과 기준에 대하여 고찰하며, 국내 암호 알고리즘 표준인 SEED등 여러 암호 알고리즘과 난수발생기를 AES 평가기준에 맞추어 새롭게 분석해 보고자 한다.

ABSTRACT

In this paper, we investigate criteria for evaluating cryptographic strength based on randomness testing of the advanced encryption standard candidates, which have conducted by NIST(National Institute of Standards & Technology). It is difficult to prove that a given cryptographic algorithm meets sufficient conditions or requirements for provable security. The statistical testing of random number generators is one of methods to evaluate cryptographic strength and is based on statistical properties of random number generators. We apply randomness testing on several cryptographic algorithms that have not been tested by NIST and find criteria for evaluating cryptographic strength from the results of randomness testing. We investigate two criteria, one is the number of rejected samples and the other is the *p*-value from p-values of the samples.

Key words : *randomness testing, AES, p-value, evaluating cryptographic algorithm*

Ⅰ. 서 론

암호 알고리즘 평가기술은 안전하고 실용적인 암호 기술사용을 유도한다. 암호 알고리즘 평가에는 안전성과 효율성 평가가 있으며 안전성 평가는 암호 분석기술과 함께 연구되고 있다. 암호 알고리즘 평가에 있어서는 암호 알고리즘의 안전성 필요충분 조건에 대한 다양하고 객관적인 기준 연구가 이루어지고 있다. 암호 알고리즘의 안전성 충분조건으로는 암호문으로부터 평문 또는 키에 관한 어떠한 정보도

* 본 연구는 한국정보보호진흥원 연구과제(2001-S-073) 지원으로 수행하였습니다.
** 한양대학교 자연과학대학 수학과(dragonsoup@ihanyang.ac.kr, camp123@hanyang.ac.kr)
*** 한국정보보호진흥원(www.kisa.or.kr)

유도할 수 없어야 하며, 임의의 평문, 암호문, 키를 알고 있다고 하여도 현재의 암호문에 대응되는 평문 또는 키에 관한 정보를 유도할 수 없어야 한다. 이는 암호 키를 모르는 상태에서 평문과 암호문의 어떠한 연관관계도 도출할 수 없어야 하며 키와 암호문 사이의 연관 관계 또한 찾을 수 없어야 한다. 그러나 알고리즘에 의해 암호문이 생성되기 때문에 평문과 암호문, 키와 암호문 사이에는 특별한 연관 관계가 존재할 수 밖에 없다. 따라서 키 전수 조사보다 적은 계산량으로 미지의 암호문으로부터 키나 평문의 어떠한 정보도 이끌어 낼 수 없을 때 주어진 암호알고리즘은 안전하다라고 한다. 이러한 안전성 충분조건 만족여부를 조사하기 위하여서는 암호 알고리즘의 구조 그리고 적용된 세부 논리등을 고려하여야 하기 때문에 여러가지 암호 알고리즘을 공통된 하나의 기준으로 평가하기에는 어려움이 존재한다.

암호 알고리즘을 통해 생성된 암호문은 의사 난수성을 만족한다. 이러한 관점에서 볼 때 암호 알고리즘은 의사난수 생성기(Pseudo-Random Number Generator)이다. 일반적인 의사 난수성 검정은 다양한 통계적 방법을 사용하고 있으며, 검정대상 암호 알고리즘에 대한 표본선택방법과 검정방법 및 유의수준설정에 대해 정해진 정량적 기준에 따라 각기 다른 암호 알고리즘의 안전성을 난수적 관점의 안전성 기준에 의해 평가한다.

본 논문에서는 AES 블록 암호 알고리즘 안전성 평가방법인 의사난수검정 방법을 고찰하며, 기존의 안전성 평가방법을 구현하여 AES 블록 암호 알고리즘인 RIJNDAEL 뿐 아니라 국내 암호 알고리즘 표준인 SEED 그리고 일본의 KASUMI 등 실제 사용되는 여러 암호 알고리즘과 난수발생기를 AES 평가기준에 맞추어 새롭게 분석 한다.

Ⅱ. NIST(National Institute of Standards and Technology)의 난수성 테스트(Randomness Testing)

각각의 AES 후보알고리즘들을 대상으로 안전성 기준항목 검정을 위한 최대 9개의 데이터 집합을 구성하였고, 이들 데이터 집합들의 난수성 검정을 위하여 189개의 통계테스트들을 각각의 데이터 집합에 적용, '가설검정(Hypothesis Testing)'을 실시하게 된다.

모든 데이터 집합과 통계테스트는 NIST산하 '컴퓨터 보안부(Computer Security Division)'와 '통계 공학부(Statistical Engineering Division)' 공동으로 '난수생성과 테스트(Random Number Generation and Testing : RNG Testing)'에서 정의하고 체계화시킨 것이다. 이는 난수발생기(Random Number Generator)의 결과값들이 의사난수 성질을 만족하고 있는지를 검정할 수 있는 수학적, 통계학적 기준과 표준안 제정의 목적을 가지고 있다. 현재 그들의 프로젝트 결과를 담은 문서 'Special Publication 800-22 : A Statistical Test Suite for Random and Pseudorandom Number Generators for Cryptographic Applications'[4]와 'the NIST Statistical Test Suite'의 모든 소스파일 등이 인터넷에 공개되어있다. 또한 암호학적 난수발생기(Cryptographic RNG Standard) 미국표준안인 ANSI X9.82(the financial community를 위한 난수발생기 미국표준안)에도 포함 될 예정이다.

Ⅲ. NIST의 AES 데이터 집합

NIST는 9개의 데이터 집합을 구성하여 다양한 안전성기준 항목의 검정을 도모하고 있다. 15개 AES 후보들 중 5개 후보를 선정하기 위한 AES 1 라운드에서는, 키와 블록크기를 128비트로서 고정하였으며 9개의 데이터 집합들을 구성하여 통계테스트에 적용시켰다. 5개 AES 후보들의 검정을 위한 두번째 라운드(AES2)에서는, 전체 암호화 과정(Full Round)에 대한 테스트로 192-비트와 256-비트의 키 크기를 사용하였으며, 부분 암호화 과정(Partial Round)에 대한 테스트로는 128-비트 키 크기를 사용하였다. 데이터 집합은 'Random Plaintext/Random 128-Bit Keys' 데이터 집합이 제외된 총 8개의 데이터 집합들을 구성한 후 통계테스트에 적용시키는 과정을 거친다. 자세한 데이터 집합의 구성내용은 참고문헌을 살펴본다[1,2,3,4,8].

AES 데이터 집합

1. 키 쇄도(Key Avalanche)
2. 평문 쇄도(Plaintext Avalanche)
3. 평문/암호문 상관관계(Plaintext/Ciphertext correlation)
4. CBC 모드(Cipher Block Chaining)

5. 랜덤 평문/랜덤 키(Random Plaintext/Random Keys)
6. 저밀도 키(Low Density Keys)
7. 저밀도 평문(Low Density Plaintext)
8. 고밀도 키(High Density Keys)
9. 고밀도 평문(High Density Plaintext)

Ⅳ. 통계테스트(The Statistical Testing of Random Number Generators

NIST는 AES 후보 알고리즘에 대한 난수성을 테스트를 하기 위해, 준비되어진 각각의 9개 데이터 집합들을 189개의 통계테스트에 적용하였다. 이러한 통계 테스트는 가설검정(Hypothesis Test)의 형태를 띄고 있으며 각각의 통계테스트의 이해를 위한 자세한 내용은 참고문헌[4]를 살펴본다. 이진수열의 가설검정을 통한 평가절차는 다음의 [표 1]과 같다.

[표 1] 가설검정을 통한 평가절차

순서	절 차	주 석
1	귀무가설 (Null Hypothesis) 선언	이진수열이 난수성질을 만족한다고 가정한다
2	통계량 계산	모든 통계량 계산은 비트(Bit) 단위수준에서 수행된다
3	p-value 계산	p-value은 0에서 1
4	p-value와 유의수준(α)비교	고정된 α값($\alpha \in (0.001, 0.01)$)에 대하여 p-value이 유의수준보다 크거나 같을 때 채택(Success)하며, 그렇지 않은 경우에는 기각(Failure)을 선언한다

통계테스트의 종류는 다음과 같다.

- Frequency(Monobits) Test
- Test for Frequency within a Block
- Runs Test
- Test for the Longest Run of Ones in a Block
- Random Binary Matrix Rank Test
- Discrete Fourier Transform(Spectral) Test
- Non-overlapping(Aperiodic) Template Matching Test
- Overlapping(Periodic) Template Matching Test
- Maurer's Universal Statistical Test
- Lempel-Ziv Complexity Test
- Linear Complexity Test
- Serial Test
- Approximate Entropy Test
- Cumulative Sum(Cusum) Test
- Random Excursions Test
- Random Excursions Variant Test

Ⅴ. NIST 안전성 평가 검정

NIST의 난수성 검정 방법은 AES Round 1과 Round 2에서 상이하게 적용되었다. AES Round 1에서 15개 후보알고리즘의 난수성 검정시는 각 난수 수열의 p-value을 구해 유의 수준인 0.01보다 값이 적은 수열의 개수를 알아내 최대 허용 기각수와 비교하는 방법을 적용하였다. AES Round 2에서는 최대 허용 기각수와의 비교와 더불어 p-value들의 Uniform 분포를 확인하는 검정방법을 적용한다. 카이스퀘어 분포를 가정하여 계산한 p-value가 0.0001보다 같거나 크면 주어진 통계테스트를 통과하는 방법을 적용한다.

5.1 통계테스트 기각 수열 개수

5.1.1 AES Round 1

NIST는 Round 1에서 데이터 집합을 알고리즘에 각각 적용시켜 난수성을 검정할 때, 유의수준(The significance level)과 신뢰구간(the confidence interval)을 이용하였다[2]. 그러나 이것은 이론상의 기대치이며 실제적용은 유의수준 0.01 수준에서의 합당한 신뢰구간(CI)을 사용하게 된다. 대상수열의 난수성을 가늠하게 해주는 기준으로 유의수준(α)오차 범위를 사용한 신뢰구간을 다음과 같이 정의한다.

$$s\left(\alpha + 3\sqrt{\frac{\alpha(1-\alpha)}{s}}\right)$$ s : 이진수열의 크기, α : 유의수준

위의 식은 정규분포를 가정한 후 표준편차의 3배(3-Sigma)인 단측(양수) 검정(One-sided test)의 형태를 띄게 됨을 알 수 있다. 위의 식을 통해 이진수열의 크기에 따른 최대기각 수를 정해보면 다음 [표 2]와 같다.

이를 적용하여 각 데이터집합과 이진수열의 크기에 따른 기각 수를 결정하게 된다.

〔표 2〕 이진수열의 크기에 따른 최대기각 수

이진수열의 수	유의수준 $\alpha=0.01$	신뢰구간 적용의 유의수준
128	1.28	4.657
300	3.00	8.170
384	3.84	9.689

5.1.2 AES Comment

AES Round 1 후보알고리즘에 대한 난수성 평가에서 일부 알고리즘이 비난수성 성향을 보이는 것은, 최대허용 기각수에 대한 통계적 편차(Statistical Anomalies)로 인해 일부 알고리즘이 비난수성 성향을 보이는 것이라 밝히고 있다[2]. 이러한 통계적 편차는 난수성 검정에서 적용된 신뢰구간(Confidence Interval)이 유의 수준(Significance Level)의 정규 근사(Normal approximation)로부터 구해졌다는 점에서 그 이유를 찾을 수 있는데, Sean Murphy[5]와 RSA Security의 RC6 통계 분석[6]에서 이러한 지적이 나타나있다. NIST에서는 각 표본들의 1종 오류확률이 0.01일 때 단측검정을 실시하여, 3-sigma의 정규근사로 계산했을 때의 유의수준 0.0013 수준을 설정하였다. 그러나 표본크기 $n(n=128, 300, 384)$에 대하여 확률 $p(n, k)$를 계산한 것은 다음 〔표 3〕과 같다.

〔표 3〕 확률 $p(n, k)$

n	k	$p(n, k)$	n	k	$p(n, k)$	n	k	$p(n, k)$
128	4	0.00961	300	8	0.00360	384	9	0.00594
128	5	0.00192	300	9	0.00102	384	10	0.00196
128	6	0.00033	300	10	0.00026	384	11	0.00060

$p(n, k)$는 표본크기 n개 중 검정에 대하여 k번까지 기각될 수 있는 확률을 뜻한다. 이항분포를 따르며 single test에 대한 유의수준은 0.01이다.

표에서 볼 수 있듯이, 정규 근사 유의수준 0.001 수준에서 볼 때, AES Round 1에서 실시한 통계테스트의 최대허용 기각수보다 1개씩의 허용 기각수가 늘어남을 알 수 있다. 따라서 표본크기 300개에 대한 최대허용 기각수는 8개에서 9개로, 표본크기 384개에 대한 최대허용 기각수는 9개에서 10개로 늘어난다. AES Round1에서 적용한 이항분포로의 정규 근사는 표본크기가 1000개 이상이 될 때 어느 정도 정확성을 보장할 수 있다[4].

5.1.3 AES Round 2

AES Round2에서는 2가지의 검정을 실시하였다.

최대허용 기각수의 확률과 유의확률(p-value)의 분포를 고려한 Round2는 〔표 4〕와 같다[3].

위의 확률은 모두 이항분포를 가정하여 계산되었으며 B Column의 기각수를 k라 하고 표본크기를 n이라 했을 때 D Column과 E Column의 확률분포는 다음과 같이 계산된다.

D : Probability of each event in column C

$$P(x=k)=\binom{n}{k}p^k q^{n-k},$$

$$p=\alpha,\quad q=1-\alpha,\quad \alpha=0.01$$

E : p-value associated with each event

$$P(x\geq k)=1-\sum_{j=0}^{k-1}\binom{n}{j}p^j q^{n-j}$$

〔표 4〕 300개 데이터 집합에서의 확률표

A	B	C	D	E
No. of successes (p)	No. of failures (q)	Proportion of successes out of 300	Probability of each event in column C	P-value associated with each event
300	0	1.0000	0.049040894070000	1.000000000000000
299	1	0.9967	0.148608769000000	0.950959105900000
298	2	0.9933	0.224414253600000	0.802350336000000
297	3	0.9900	0.225169857100000	0.577936824000000
296	4	0.9867	0.168877392800000	0.352766225000000
295	5	0.9833	0.100985269200000	0.183888832500000
294	6	0.9800	0.050152616880000	0.082903563300000
293	7	0.9767	0.021276867770000	0.032750946400000
292	8	0.9733	0.007871366486000	0.011474078600000
291	9	0.9700	0.002579617290000	0.003602712100000
290	10	0.9667	0.000758251145300	0.001023094800000
289	11	0.9633	0.000201921792600	0.000264843700000
288	12	0.9600	0.000049120705450	0.000062921900000
287	13	0.9567	0.000010992045980	0.000013801200000
286	14	0.9533	0.000002276130732	0.000002809200000
285	15	0.9500	0.000000438365918	0.000000533100000
[284,0]	[16,300]	[0.9467, 0.0000]	[0.000000078872655 9, 0.000000000000000]	[0.000000094700000, 0.000000000000000]

AES Round2에서는 이러한 확률표를 바탕으로 검정을 통과한 표본의 비율이 0.9633이상이면 난수성 검정을 통과한다고 구분하였다. 통과표본의 비율 0.9633은 곧 누적 p-value 0.0001과 같은 수준이다. 따라서 확률표에서 보듯이 표본크기 300개에 대한 최대허용 기각수는 11개까지 늘어난다.

지금까지 AES제정을 위해 실시된 난수성 검정에서, 통계테스트 적용 시 허용되는 최대기각수를 결정하는 방법에 대하여 고찰하였고 위의 3가지 기준을 하나의 표로서 정리해 보면 〔표 5〕과 같다. 각 데이터 집합에 따른 검정을 실시했을 때, 각각의 통계테스트의 기준에 따른 기각의 수가 표본크기의 최대기각수보다 클 경우, 그 데이터 집합에서 해당 통계테스트에 대한 비 난수적 성질을 보인다고 한다.

〔표 5〕 각 데이터 집합, 표본 수, 표본 크기, 최대 허용기각 수 , 검정 수에 대한 표

데이터 집합	표본 수		표본 길이		최대허용기각수			테스트 적용횟수	
	Round 1	Round 2	Round 1	Round 2	Round 1		Round 2	Round 1	Round 2
					NIST	Comment			
128-비트 키 Avalanche	384	300	1,048,576	1,048,576	9	10	11	187	189
평문 Avalanche	384	300	1,048,576	1,048,576	9	10	11	187	189
평문/암호문 상관관계	128	300	1,048,576	1,048,576	4	5	11	59	189
CBC 모드	300	300	1,048,576	1,048,576	8	9	11	59	189
난수평문/ 난수 128-비트 키	128		1,048,576	1,048,576	4	5		59	
저밀도 평문	128	300	1,056,896	1,048,576	4	5	11	59	189
저밀도 128-비트 키	128	300	1,056,896	1,048,576	4	5	11	59	189
고밀도 평문	128	300	1,056,896	1,048,576	4	5	11	59	189
고밀도 128-비트	128	300	1,056,896	1,048,576	4	5	11	59	189

5.2 *p*-value의 Uniform 분포

AES Round2에서 새롭게 적용되었으며, 진정한 난수수열들의 *p*-value 분포는 Uniform 하다는 가정에 따른 테스트이다.

자유도 9를 가지는 χ^2분포를 바탕으로 *p*-value들의 *p*-value을 계산하게 되며, 테스트 결과값에 Goodness-of-fit-distributional test를 적용한다. 사용된 χ^2분포는 다음과 같다.

$$\chi^2 = \sum_{i=1}^{10} \frac{(F_i - \frac{s}{10})^2}{\frac{s}{10}}$$

F_i is the number of *p*-value in sub-interval i
s is sample size

따라서, 계산된 *p*-value가 0.0001 보다 적을 때 통계테스트를 기각한다[4].

Ⅵ. 암호알고리즘 안전성 분석

본 논문에서는 AES 평가기준을 바탕으로 여러 알고리즘을 새롭게 분석해 보고자 한다. 이번에 다룬 알고리즘으로는 DES, KASUMI, Rijndael, SEED, RC 2 그리고 srand와 BBS 이다.

6.1 데이터 집합 구성

AES 선정 작업 중 마지막 최종 5개의 후보들의 난수성 테스트에서 제외된 '난수평문/난수 128-비트 키' 데이터 집합과 'CBC모드'를 제외한 6개의 데이터 집합을 각각의 알고리즘 별로 구성하였다. 〔표 6〕과 같이 표본수는 모두 300개로 통일하였으며 표본의 길이는 각 알고리즘의 키길이와 불록 크기에 따라 약간의 차이가 있다. srand는 각 표본길이를 1,048,576으로 생성하였으며, BBS는 NIST가 공개한 통계테스트[4]에 포함된 수열을 사용하였다.

〔표 6〕 해당 알고리즘에 대한 표본길이, seed value

데이터 집합 (표본수)	seed value	표 본 길 이 (표본수 1개당)				
		DES	KASUMI	Rijndael	SEED	RC 2
키 쇄도(300)	Blum-Blum-Shub	1,053,696	1,048,576	1,048,576	1,048,576	1,048,576
평문 쇄도(300)	Blum-Blum-Shub	1,048,576	1,048,576	1,048,576	1,048,576	1,048,576
고밀도 키(300)	$\sqrt{3}$-extension	1,048,576	1,048,576	1,056,896	1,056,896	1,048,576
고밀도 평문(300)	e-extension	1,048,576	1,048,576	1,056,896	1,056,896	1,048,576
저밀도 키(300)	$\sqrt{2}$-extension	1,048,576	1,048,576	1,056,896	1,056,896	1,048,576
저밀도 평문(300)	π-extension	1,048,576	1,048,576	1,056,896	1,056,896	1,048,576

6.2 통계 테스트 적용과 안전성 분석

구성된 데이터 집합들을 NIST의 16개 통계테스트와 똑같이 구성된 통계테스트에 각각 적용시켰다. 189개의 통계테스트 중 기각수가 9를 넘어가는 통계테스트를 각 데이터 집합별로 정리한 것이다.

정리한 결과값들을 살펴볼 때 AES Final Round에서 실시한 난수성 검정의 최대 허용 기각수 11개를 넘는 알고리즘은 없었다. 하지만 난수발생기인 srand의 Excursion Variant 테스트에서 기각수가 12개를 나타냄으로 난수성질로부터 일탈성을 보임을 알 수 있다.

〔표 7〕 DES에 대한 검정결과

데이터 집합 (표본수)	DES	
	통계테스트	기각수
키 쇄도(300)		
평문 쇄도(300)		
고밀도 키(300)	Non-overlapping 000101111 Non-overlapping 010000011 Non-overlapping 101001000 Excursion Var. state +6	9 9 9 9
고밀도평문(300)	Overlapping template	9
저밀도 키(300)		
저밀도평문(300)	Non-overlapping 111010010 Serial state +1	10 10

〔표 8〕 KASUMI에 대한 검정결과

데이터 집합 (표본수)	KASUMI	
	통계테스트	기각수
키 쇄도(300)	Excursion state -4	11
평문 쇄도(300)		
고밀도 키(300)	Non-overlapping 111101010 Serial state +1 Excursion state -3	9 9 9
고밀도 평문(300)		
저밀도 키(300)	Non-overlapping 111110110 Lempel Ziv. Serial state +2	9 10 9
저밀도 평문(300)	Non-overlapping 110010010 Excursion state -4	10 11

〔표 9〕 Rijndael에 대한 검정결과

데이터 집합 (표본수)	Rijndael	
	통계테스트	기각수
키 쇄도(300)	Non-overlapping 110110100 Excursion Var. state -3	9 9
평문 쇄도(300)		
고밀도 키(300)	Non-overlapping 001010111	9
고밀도평문(300)	Non-overlapping 010100111 Non-overlapping 101000100	9 11
저밀도 키(300)		
저밀도평문(300)	Approximate	10

〔표 10〕 SEED에 대한 검정결과

데이터 집합 (표본수)	SEED	
	통계테스트	기각수
키 쇄도(300)	Non-overlapping 000011101 Non-overlapping 001011111 Non-overlapping 111101010 Excursion Var. state -4	9 9 9 10
평문 쇄도(300)		
고밀도 키(300)		
고밀도평문(300)	Non-overlapping 011111111 Non-overlapping 111111110	9 9
저밀도 키(300)		
저밀도평문(300)	Non-overlapping 101001000	9

〔표 11〕 RC2에 대한 검정결과

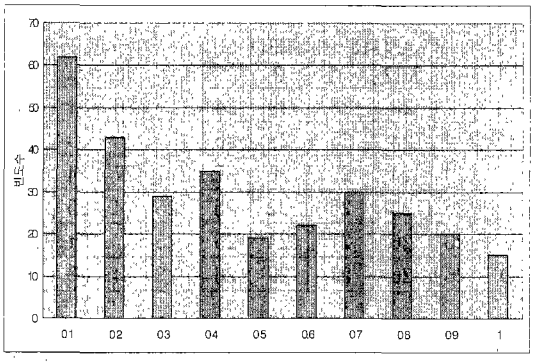
데이터 집합 (표본수)	RC 2	
	통계테스트	기각수
키 쇄도(300)		
평문 쇄도(300)	Frequency M	9
고밀도 키(300)	Non-overlapping 11111011	10
고밀도평문(300)	Non-overlapping 101010000 Non-overlapping 111010000 Excursion state +4	9 10 10
저밀도 키(300)		
저밀도평문(300)	Run test	9

〔표 12〕 SRAND, BBS에 대한 검정결과

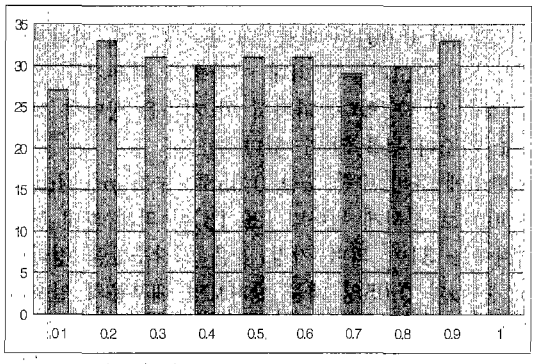
데이터 집합 (표본수)	srand		BBS	
	통계테스트	기각수	통계테스트	기각수
srand (300) BBS (300)	Frequence-M Non-overlapping 000110101 100100000 110001010 Excursion -1 Excursion +2 Excursion Var.	11 10 9 10 9 9 12	Run test	9

6.3 *p*-value의 Uniform 분포 분석

진정한 난수수열들의 *p*-value 분포는 Uniform하다는 가정에 따른 테스트이며 *p*-value들의 *p*-value을 계산하여 나온 결과값이 0.0001보다 적을 때 통계테스트를 기각하게 된다. 〔그림 1〕은 KASUMI의 저밀도 평문 데이터 집합에 대한 Entropy 통계테스트 결과의 *p*-value의 분포이며, 〔그림 2〕는 DES의 저밀도 키 데이터 집합에 대한 Excursion (+3) 통계테스트 결과의 *p*-value의 분포를 나타낸 히스토그램이다.



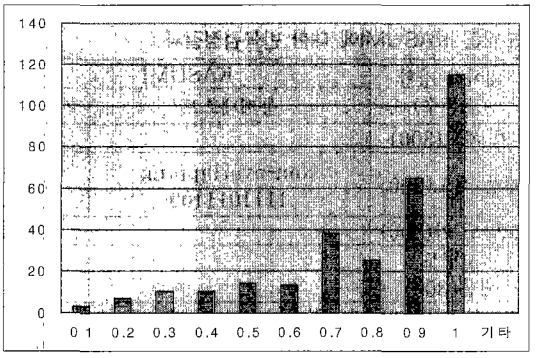
〔그림 1〕 p-value=0.0000000



〔그림 2〕 p-value=0.993404767

〔그림 2〕의 p-value가 높을 때의 히스토그램이 〔그림 1〕의 낮은 p-value을 보이는 히스토그램보다 월등히 Uniform하다는 것을 알 수 있다.

하지만, p-value의 분포가 Uniform한 것과 해당 데이터 집합의 난수성과는 아무런 상관계수가 없다. 즉, p-value의 분포가 Uniform하다고 해서 그 데이터 집합의 난수성질의 일탈성이 없다고 확신할 수 없다는 뜻이다.



〔그림 3〕 p-value=0.0000000

다음 〔그림 3〕은 srand데이터 집합에 대한 Serial 통계테스트 결과를 나타내는 히스토그램이다.

p-value가 0.0000000이므로 전혀 Uniform하지 않다는 것을 알 수 있다. 그러나 히스토그램을 자세히 살펴보면 전체 p-value의 60%가 0.9에서 1사이에 몰려있다는 점을 알 수 있다. 이는 곧 위의 데이터 집합의 난수성질이 상당히 높다는 것을 뜻하며, 데이터 집합에서의 최대 허용 기각 표본수와 p-value 분포의 Uniform 성질은 아무런 상관관계가 없다는 것을 말해준다. 따라서 p-value 분포의 Uniform 성질은 하나의 독립적인 평가기준으로 다뤄져야 한다.

6.4 유의 수준의 변화에 따른 최대 기각수 분석

AES Final Round에서는 유의 수준 α를 0.01로 고정하고 누적 p-value을 0.0001로 하여 검정을 실시하였으며 최대허용 기각수를 11개로 설정하였다. 그러나 NIST는 이러한 기준에 대한 어떠한 근거도 제시하지 않았다. 따라서 기준 제시를 위한 다양한 시도를 통해 가장 적합한 기준, 즉 최대 기각수를 찾아내기 위한 분석을 실시한다.

유의 수준을 0.001, 0.01, 0.02, 0.03, 0.04, 0.05 등 6개로 나누고 〔표 4〕과 같은 각각의 유의 수준에 대한 확률표를 구하였다.

이중 누적 p-value의 기준에 따른 각각의 최대 허용 기각수를 정리하면 다음의 〔표 13〕과 같다.

다양한 기준에 적용하기 위해 새로운 데이터 집합을 구성하였는데, 각각의 암호 알고리즘을 새로운 유의수준에 따라 기각수를 살펴보고 기각수 분포를 통해 해당 데이터 집합중 가장 높은 기각수만을 정리하여 새롭게 만든표는 다음의 〔표 14〕~〔표 19〕와 같다. KASUMI와 Rijndael의 다양한 데이터 집합은 각기 다른 seed 값을 통해 새롭게 구성하였다.

〔표 13〕 유의수준과 누적 p-value에 대한 최대 허용 기각수

		유의수준 α					
		0.001	0.01	0.02	0.03	0.04	0.05
누적 p-value	0.0001	4	11	17	22	26	31
	0.001	3	10	15	19	24	28
	0.01	2	8	12	16	21	24

〔표 14〕 α=0.001 경우의 기각수

0.001	DES2	KASUMI2	Rijndael2	SEED2	RC2	Rijn3	Rijn4	Kasu3	Kasu4
AK	3	4	3	4	2	4	3	3	4
AP	5	2	2	2	3	2	3	2	2
HDK	4	2	2	2	2	4	4	2	3
HDP	4	3	4	3	3	3	2	2	6
LDK	3	3	3	2	3	2	3	3	3
LDP	4	4	3	3	3	3	3	3	2

〔표 15〕 α=0.01 경우의 기각수

0.01	DES2	KASUMI2	Rijndael2	SEED2	RC2	Rijn3	Rijn4	Kasu3	Kasu4
AK	8	11	9	10	8	9	8	9	9
AP	8	7	8	8	9	7	9	9	10
HDK	9	9	9	8	10	11	9	10	8
HDP	9	8	11	9	10	10	9	7	9
LDK	8	10	7	8	8	8	10	11	12
LDP	10	11	10	9	9	10	9	8	11

〔표 16〕 α=0.02 경우의 기각수

0.02	DES2	KASUMI2	Rijndael2	SEED2	RC2	Rijn3	Rijn4	Kasu3	Kasu4
AK	13	14	15	14	15	13	14	14	13
AP	13	12	15	13	16	12	12	14	14
HDK	15	15	12	12	14	16	13	13	12
HDP	15	13	14	13	13	14	14	13	15
LDK	13	14	19	13	12	15	13	13	13
LDP	14	15	13	14	13	14	13	14	15

〔표 17〕 α=0.03 경우의 기각수

0.03	DES2	KASUMI2	Rijndael2	SEED2	RC2	Rijn3	Rijn4	Kasu3	Kasu4
AK	17	17	19	18	20	23	20	20	18
AP	18	15	17	18	17	17	17	20	18
HDK	20	18	18	18	17	20	21	17	18
HDP	19	17	19	16	18	18	18	18	20
LDK	19	20	22	16	16	19	18	19	19
LDP	16	18	18	19	17	19	18	17	21

〔표 18〕 α=0.04 경우의 기각수

0.04	DES2	KASUMI2	Rijndael2	SEED2	RC2	Rijn3	Rijn4	Kasu3	Kasu4
AK	22	20	24	23	22	25	22	24	21
AP	21	20	20	22	23	21	21	24	22
HDK	24	22	22	23	23	24	21	21	23
HDP	22	23	21	19	22	22	21	23	22
LDK	22	21	25	20	22	23	22	26	22
LDP	22	20	26	23	22	29	21	22	25

〔표 19〕 α=0.05 경우의 기각수

0.05	DES2	KASUMI2	Rijndael2	SEED2	RC2	Rijn3	Rijn4	Kasu3	Kasu4
AK	27	23	28	29	24	29	26	31	26
AP	24	23	25	28	26	26	25	30	26
HDK	29	24	29	26	27	27	27	24	27
HDP	26	23	25	22	25	25	24	25	25
LDK	26	26	27	24	25	25	27	31	27
LDP	26	26	27	29	26	33	24	26	29

만약 기준을 유의 수준 α는 0.001로, 누적 p-value은 0.01로 정하게 된다면 〔표 13〕에 의해 최대 허용 기각수는 2가 되며 가장 높은 기각수만을 정리한 표 중 α가 0.001일 때를 〔표 14〕와 같이 살펴보면 대부분의 알고리즘의 최대 기각수가 허용치를 웃도는 것을 알 수 있다. 따라서 이 기준은 상당히 까다로운 기준이 됨을 알 수 있다.

반면에 유의 수준 α를 0.05로, 누적 p-value은 0.0001로 정하게 된다면 〔표 13〕에 의해 최대 허용 기각수는 31이 되며 가장 높은 기각수만을 정리한 표 중 α가 0.05일 때를 살펴보면 대부분의 알고리즘의 최대 기각수가 허용치를 밑도는 것을 알 수 있다. 따라서 이 기준은 상당히 완화된 기준임을 알 수 있다.

AES 난수성 검정 기준인 α=0.01, 누적 p-value은 0.0001을 적용시켰을 때 최대 허용 기각수는 11개이다. 〔표 7〕~〔표 12〕에서 알 수 있듯이 허용 기각수를 넘어가는 알고리즘은 최대 기각수 12개가 존재하는 srand이다. 다른 표본군을 생성하여 반복검정을 실시하였을 때 KASUMI에서는 〔표 20〕과 같이 최대기각수가 12개가 발생되었다. 따라서 AES 난수성 검정 기준을 그대로 적용하였을 때, KASUMI는 난수성질의 일탈성이 보인다고 판단된다.

〔표 20〕 KASUMI에 대한 반복검정결과

데이터 집합 (표본수)	KASUMI	
	통계테스트	기각수
키 쇄도(300)		
평문 쇄도(300)	Non-overlapping 1111011100	10
고밀도 키(300)		
고밀도 평문(300)		
저밀도 키(300)	Excursion state -2	12
저밀도 평문(300)	Lempel-Ziv	10
	Cusum(forward)	11

또한 AES 알고리즘으로 선정된 Rijndael은 유의수준 α가 높아질수록 각 누적 p-value에 대한 최대 기각수가 허용치를 넘어가는 난수성질의 일탈성 현상이 두드러지는 점을 발견할 수 있었다.

Ⅶ. 결 론

난수성 검정(Randomness testing)은 알고리즘과 난수발생기의 안전성을 위한 필요조건인 난수성을 검정할 수 있는 정량적 기준에 의한 유일한 평가방법이라는 것에 그 의미가 있다.

각 데이터 집합의 통계테스트에 대한 기각수를 실제 히스토그램으로 보았을 때 AES Round1에서 적용한 가정의 근거인 정규분포를 따라간다는 점을 확인할 수 있었으며, AES 난수성 검정 기준을 적용하였을 때는 KASUMI의 난수성질의 일탈성을 의심할 수 있다는 점과 유의수준을 높였을 때는 AES로 선정된 Rijndael에게도 난수성질의 높은 일탈현상이 보인다는 점도 알 수 있었다. 또한 AES Round 2에 적용된 p-value의 Uniform분포 검정 방법은 최대 허용 기각수를 통한 검정법과는 별개의 독립적인 검정법으로 다루어야 한다.

이러한 결과를 통하여 더욱 다양한 암호알고리즘의 평가작업과 더불어 각각 독립인 통계테스트들의 총체적 평가기준 제시를 위한 이론적 근거 연구, 유의 수준의 세부적 기준 분할을 적용한 적절한 검정 기준 제시 등, 효율적인 평가기준을 도출할 수 있도록 해야 한다.

참 고 문 헌

[1] J. Soto. "Statistical Testing of Random Number Generators", Proceedings of the 22nd National Information Systems Security Conference, Crystal City, Virginia, October 1999.

[2] J. Soto. "Randomness Testing of the AES Candidate Algorithms", NIST 1999 (http://csrc.nist.gov/encryption/aes/round1/r1-rand.pdf).

[3] J. Soto. "Randomness Testing of the Advanced Encryption Standard Finalist Candidates", NIST 2000(http://csrc.nist.gov/encryption/aes/round2/conf3/papers/30-jsoto.pdf).

[4] Andrew Rukin, Juan Soto,, James Nechvatal, Miles Smid, Elaine Barker, James Dray, San Vo, Stefan Leigh, Mark Levenson, Mark Vangel, David Banks, Allen Heckert, "A Statistical Test Suite for Random and Pseudorandom Number Generators for Cryptographic Applications"(http://csrc.nist.gov/rng/rng2.html).

[5] Sean Murphy, "The Power of NIST's Statistical Testing of AES Candidate" (http://csrc.nist.gov/encryption/aes/round2/conf3/papers/09-smurphy.pdf).

[6] RSA Security, "On the Statistical Testing of RC6"(http://csrc.nist. gov/ encryption/aes/round2/comments/20000407-jjonsson.pdf).

[7] 이상진, 성수학, 송정환, "블록 암호 알고리즘 구조 복잡도 및 암호문 특성 분석", 한국정보보호센터 암호기술연구 00-8.

[8] 송정환, 조용국, 현진수. "AES 안전성 평가에 대한 고찰", WISC2000.

〈著者紹介〉

조 용 국 (Yong Kuk Cho) 학생회원
2000년 2월 : 한양대학교 수학과 졸업
2000년 3월~현재 : 한양대학교 수학과 석사과정
〈관심분야〉 암호학, 정보보호

송 정 환 (Jung Hwan Song) 정회원
1984년 2월 : 한양대학교 수학과 졸업
1989년 5월 : Syracuse University Mathematics 석사
1993년 5월 : Rensselaer Polytechnic Institute Mathematics 박사
1999년 3월~현재 : 한양대학교 수학과 조교수
〈관심분야〉 암호학, 최적론, 수리계획법, 정보보호

강 성 우 (Sung Woo Kang)
1996년 2월 : 중앙대학교 수학과 졸업
2001년 8월 : 서울대학교 수학과 대학원 졸업(이학석사)
2000년 12월~현재 : 한국정보보호진흥원 기술단/암호기술팀(연구원)
〈관심분야〉 암호학