S-box 형태의 다 수열 발생기에 관한 연구

이 훈 재†

요 약

출력 수열의 수가 스트림 암호의 새로운 평가요소로서 제안된 바 있으나 일반적으로 발표된 대부분의 이진 수열 발생기는 출력 수열이 1개 뿐인 것으로 알려져 있다. S-box는 비선형성이 높아서 블록 암호에서 많이 사용되고 있는 암호 요소이며, 이를 스트림 암호에 적용시킬 경우 효율적으로 비도 요소를 개선시킬 수 있다. 본 논문에서는 SAC 특성 등 비선형성이 뛰어난 S-box를 사용하여 출력 키 수열의 수가 여러 개인 다 수열 발생기를 제안하고, 이 발생기의 주기, 선형복잡도, 랜덤 특성 및 출력 수열의 수를 분석하였다.

On a Multiple-Cycle Binary Sequence Generator Based on S-box

Hoon-Jae Lee†

ABSTRACT

The number of keystream cycle sequences has been proposed as a characteristic of binary sequence generator for cryptographic application, but in general the most of binary sequence generators have a single cycle. On the other hand, S-box has been used to block cipher for a highly nonlinear element and then we apply it to the stream cipher with a high crypto-degree. In this paper, we propose a multiple-cycle binary sequence generator based on S-box which has a high nonlinearity containging SAC property and analyze its period, linear complexity, randomness and the number of keystream cycle sequences.

1. 서 론

Beker와 Piper[1, 2]는 스트림 암호에 사용되는 출력 키 수열에 대하여 주기의 최소값이 보장되어야 하고 좋은 랜덤 특성 (randomness)을 가져야 하며, 충분히 큰 선형 복잡도를 가져야 한다는 조건을 제시하였다. 그리고 Sigenthaler[3]는 출력 키 수열이 충분한 상관 면역 차수를 가져야 하며, Golic[4]은 키 수열 수(키 수열 사이클 수)가 충분히 많아야 함을 제시하였다.

† 정 회 원 : 경운대학교 컴퓨터전자정보공학부 교수
논문접수 : 1999년 10월 25일, 심사완료 : 2000년 5월 8일

키 수열 수는 대부분의 이진 수열 발생기에서는 단 하나 뿐인 것으로 알려져 있으며[4], 이 경우 키 수열 발생기는 출발점이 다를 뿐 항상 동일한 사이클 속에서만 출력을 발생한다. 그러나 키 수열 사이클이 두 개 이상일 경우 초기 값(키)을 변경시키면 다른 사이클에서 출력이 발생될 수 있기 때문에 사이클 수가 많아지면 암호 분석이 그 만큼 더 어려워진다. 여기서 사이클 수는 초기 키에 따라 변경될 수 있으며, 긴 주기, 큰 선형 복잡도 및 랜덤 특성을 골고루 갖춘 여러 겹의 서로 다른 출력 사이클 수열을 의미한다. 이런 관점으로 참고 문헌[5]에서 LFSR의 탭 선택을 조절하

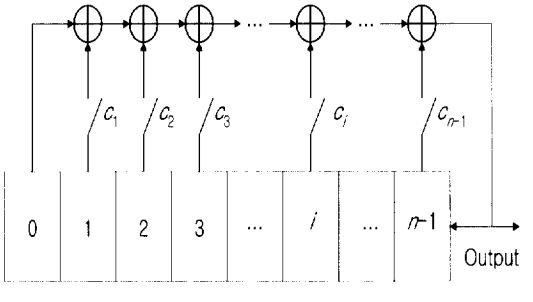
여 키 수열의 수를 높이는 방안이 제시된 바 있으나 이 방법은 선형 입력부의 변경에 의한 방법이며, 한편으로는 비선형 함수의 가변에 의하여 키 수열 수를 높이는 방법도 가능하다. S-box는 비선형성이 높아서 블록 암호의 핵심 요소로 사용되고 있으며, 이를 스트림 암호에 적용시킬 경우 키 수열 수를 효율적으로 개선시킬 수 있다.

본 논문에서는 스트림 암호 시스템에서 키 수열 사이클 수에 대한 구조 모델을 제안하고, 제안된 모델에 따라 기존 다 수열 발생기를 검토한다. 또한 스트림 암호 시스템에서 키 수열 동기에 따라 키 수열 사이클이 변경되지 않을 경우 Dawson 공격[5]에 해독될 가능성이 있음을 보여 준다. 그리고 키 수열 사이클 수를 늘릴 구체적인 방안으로서 S-box를 이용한 다 수열 발생기를 제안하여 주기, 랜덤 특성, 선형 복잡도 및 키 수열 사이클 수 등의 비도 특성을 분석한다. 마지막으로 비도 특성 비교를 위하여 유사 형태의 발생기를 선정하고, 작은 단수에 대한 시뮬레이션 결과를 비교 분석한다.

2. 키 수열 사이클 수

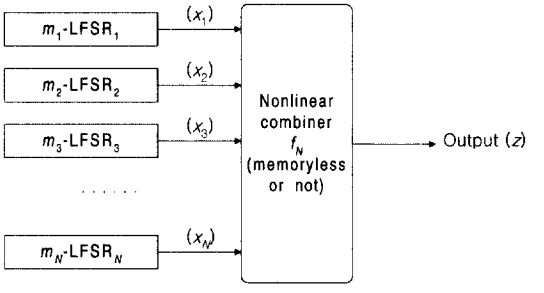
2.1 이진 수열 발생기

(그림 1)의 n단 LFSR (linear feedback shift register)은 선형 출력을 발생하기 때문에 $2n$ 비트만 알면 귀환 탭 (feedback tap)을 유추할 수 있으며[1, 2], 일반적으로 이를 방지하기 위해 (그림 2)와 같이 여러 개의 LFSR을 비선형적으로 조합함으로서 출력 수열의 선형 복잡도 (linear complexity)를 증가시킨다.



$$g(x)=x^n+c_{n-1}x^{n-1}+c_{n-2}x^{n-2}+\ldots+c_1x+1$$

(그림 1) n단 LFSR



(그림 2) 일반형 키 수열 발생기

[정리 1]

(1) n단 LFSR은 모든 초기 상태 벡터가 영 벡터가 아닐 (non-null) 때 (2^n-1)인 최대 주기를 갖는다[1, 2]. 이 때 n단 LFSR은 원시 다항식 (primitive polynomial)으로부터 얻어진다.

(2) n차 원시 다항식의 총 갯수는 $\lambda(n)=\frac{\Phi(2^n-1)}{n}$ 이며, 이는 최대 주기를 만족하는 n단 LFSR의 feedback 함수의 총 갯수를 의미한다[2]. 여기서 $\Phi(\cdot)$는 오일러 함수이다.

Rueppel 등[9]과 Golic[10]은 비메모리 (memoryless)형 비선형 결합 함수 f_N (그림 2) 함수의 선형 복잡도를 계산하기 위한 star-등식 f_N^*를 다음과 같이 정의하였다. 여기서, $a_0, a_i, a_{ij}, \cdots, a_{12\cdots N} \in \{0,1\}$이고, $a_0^*=0$ ($a_0=0$일때) 또는 1 ($a_0\neq 0$일때), $a_i^*=0$ ($a_i=0$일때) 또는 1 ($a_i\neq 0$일때), $\cdots$, $a_{12\cdots N}^*=0$ ($a_{12\cdots N}=0$일때) 또는 1 ($a_{12\cdots N}\neq 0$일 때)이며, f_N의 계산 영역은 $GF(2)$ 영역, f_N^*의 계산 영역은 실수 영역이다.

$$f_N(x_1, x_2, \cdots, x_N)= a_0+\sum a_i x_i+\sum a_{ij}x_ix_j+\cdots+a_{12\cdots N}x_1x_2\cdots x_N \quad (1)$$

$$f_N^*(x_1, x_2, \cdots, x_N)= a_0^*+\sum a_i^* x_i+\sum a_{ij}^* x_ix_j+\cdots+a_{12\cdots N}^* x_1x_2\cdots x_N \quad (2)$$

[정리 2]

(Rueppel-Staffelbach) 비메모리형 키 수열 발생기 (그림 2)의 출력 z에 대한 선형 복잡도는 다음과 같이 star-등식으로 주어진다[9, 10].

$$LC(z)=f_N^*(L_1, L_2, \cdots, L_N) \quad (3)$$

여기서, $L_i = \deg[g_i(x)]$, $i = 0, 1, \cdots, N$ 이고, $g_i(x)$는 각 LFSR에 대한 원시 다항식이다.

2.2 키 수열 사이클 구조 모델

스트림 암호에서는 송·수신 키 수열 발생기에서 발생되는 키 수열 출력이 일치되지 않으면 평문을 원래대로 복호할 수 없기 때문에 송·수신 키 수열 발생기의 동기를 일치시키는 문제가 필수적이다. 키 수열 동기 방식이란 송·수신 키 수열 발생기에서 각각 동일한 키 수열을 발생시킨 다음 이들의 시작점을 서로 일치시키는 방식을 말한다. 이 때 키 수열 동기를 일치시키기 위해서는 세션 키를 포함한 동기 신호를 안전하게 상호 교환할 필요가 있다.

한편, Dawson[6]은 동일 키를 사용한 스트림 암호 시스템에서 과거 암호문과 현재 암호문을 알고 있을 경우 서로 XOR시키면 암호해독이 가능함을 보였다. 즉, 과거 평문 $P' = p_0', p_1', p_2', \ldots$, 과거 키 수열 $K' = k_0', k_1', k_2', \ldots$, 과거 암호문 $C' = c_0', c_1', c_2', \ldots$, 현재 평문 $P = p_0, p_1, p_2, \ldots$, 현재 키 수열 $K = k_0, k_1, k_2, \ldots$, 현재 암호문 $C = c_0, c_1, c_2, \ldots$,이라 두자. 이 때 동일 키를 사용한다는 가정($K = K'$)으로 부터

$$C' = p_0' \oplus k_0, p_1' \oplus k_1, p_2' \oplus k_2, p_3' \oplus k_3, p_4' \oplus k_4, p_5' \oplus k_5, \ldots \quad (4)$$

$$C = p_0 \oplus k_0, p_1 \oplus k_1, p_2 \oplus k_2, p_3 \oplus k_3, p_4 \oplus k_4, p_5 \oplus k_5, \ldots \quad (5)$$

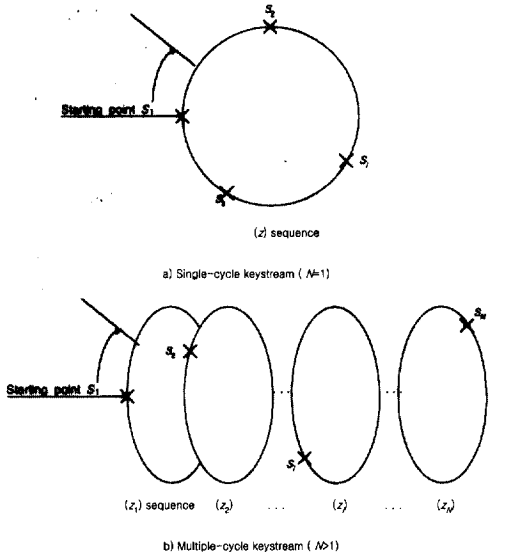
$$C \oplus C' = p_0' \oplus p_0, p_1' \oplus p_1, p_2' \oplus p_2, p_3' \oplus p_3, p_4' \oplus p_4, p_5' \oplus p_5, \ldots \quad (6)$$

이 된다. 결국 암호문 2개를 XOR하면 평문 2개의 XOR 형태로 남기 때문에 Dawson의 방법대로 평문의 잉여도를 이용하면 암호문이 해독될 수 있다. 이러한 공격 방법에 안전하려면 동기를 확립할 때마다 항상 새로운 세션 키(session key)로 초기화시켜야 한다.

[정의 3]

어떤 키 수열 발생기에서 초기 값(키) 변경에 따라 생성되어지는 동일 길이의 상이한 출력 수열 사이클의 총 갯수를 **키 수열 사이클 수(number of keystream cycle sequences)** 또는 **키 수열 수**라 한다.

(그림 3)은 스트림 암호를 이용한 암호 통신에서 키 수열 사이클 수와 안전성과의 관계를 나타내는 모델을 제안한 것이다. (그림 a)는 키 수열 사이클이 하나 뿐인 경우(대부분의 발생기), 초기 키가 지정되면 출발점(starting point) S_1으로부터 고정된 키 수열 사이클을 따라 키 수열이 발생되고, 후속 통신 시에도 동일한 키 수열 사이클 속에서 출발점만 S_2로 바뀌어 발생된다. 이렇게 중첩되어 사용된 키 수열은 Dawson 공격에 취약해진다[6]. 그러나 키 수열 사이클이 둘 이상 ($N \geq 2$)이고 세션 키에 따라 서로 다른 키 수열 사이클이 선택될 수 있다면 (최상의 경우, 세션 키 설정 수만큼의 출력 키 수열이 존재함) 출력 키 수열 사이클이 서로 겹치지 않으므로 Dawson 공격에 안전하게 된다. 게다가 일부분의 키 수열로부터 나머지 키 수열의 유추가 어렵고 세션 키 설정 수(키 공간)가 충분히 크다고 가정하면 가장 안전한 암호로 알려진 원 타임 패드 (one-time pad)[1-2]와 유사한 수준의 완전 안전성 (perfect secrecy)[1-2]을 갖는 키 수열을 발생할 수도 있다. 여기에서 원 타임 패드란 Gilvert Vernam에 의하여 제안된 암호 방식으로서 평문의 길이와 동일한 크기의 키를 생성하여 비밀 채널로 은밀히 전달하고 암호문을 해독하는 원리이며, 완전 안전성을 갖는 것으로 알려져 있다. 완전 안전성은 암호 해독가가 암호문을 획득하건 아니건 평문을 해독할 확률에 영향을 미치지 않는 가장 이상적인 안전성을 말하며, 원 타임 패드는 완전 안전성 조건을 갖고 있지만 키 길이의 상한선이 없기 때문에 키 관리가 어려운 단점이 있다.



(그림 3) 키 수열 사이클의 구조 모델

그림에서 주기 P인 두 출력 수열 사이클 (z_i), (z_j) $(i \neq j)$를 다음과 같이 나타내기로 한다.

$$(z_i) = z_{i0}, z_{i1}, z_{i2}, z_{i3}, \ldots, z_{ik}, \ldots, z_{i,P-1}, z_{i0}, z_{i1}, \ldots, i = 1, 2, \ldots, N \quad (7)$$

$$(z_j) = z_{j0}, z_{j1}, z_{j2}, z_{j3}, \ldots, z_{jk}, \ldots, z_{j,P-1}, z_{j0}, z_{j1}, \ldots, j = 1, 2, \ldots, N \quad (8)$$

이 때 임의의 정수 $k(0 \leq k \leq P)$와 j에 대하여 (z_j) 수열을 k만큼 순회 (cyclic rotate)시킨 수열 $(\widetilde{z_j})$와 어떤 수열 (z_i)가 같지 않다면 키 수열 사이클 수는 N이라고 할 수 있다.

$$(\widetilde{z_j}) = Rot((z_j), k) \neq (z_i) \quad (9)$$

여기서 $Rot((z), k)$은 (z) 수열에 대하여 k비트 만큼 이동 순회(cycle rotate)시킨 것이다. 그리고 본 논문은 일반형 수열 발생기에서 키 수열 사이클 수를 늘리기 위해서 SAC (strict avalanche criterion) 특성을 만족하는 S-box를 이용한 발생기를 제안코자 한다.

3. 다 수열 발생기 제안

3.1 S-box의 특성

Z를 정수의 집합, Z_2를 유한체 GF(2), Z_2^n을 Z_2상에서의 n차원 벡터, $\bigoplus$를 Z_2^n 또는 비트 단위의 이진 가산 연산 XOR (exclusive-OR)이라고 둔다.

[정의 6]

양의 정수 n에 대하여, $c_1^{(n)}$, $c_2^{(n)}$, $\cdots$, $c_n^{(n)} \in Z_2^n$을 다음과 같이 정의한다.

$$c_1^{(n)} = [0, 0, \cdots, 0, 0, 1]$$
$$c_2^{(n)} = [0, 0, \cdots, 0, 1, 0]$$
$$\vdots$$
$$c_n^{(n)} = [1, 0, \cdots, 0, 0, 0]$$

직관적으로 $c_i^{(n)}$는 i번째 위치에서 해밍중(Hamming weight)이 1이 되는 n차원 벡터이다. 여기에서 해밍 중이란 주어진 벡터의 이진표현에서 "1"의 개수를 의미한다.

[정의 7]

함수 $f: Z_2^n \rightarrow Z_2^m$ 는 $Z_2^n \rightarrow Z_2$인 함수 $f_j (1 \leq j \leq m)$를 이용하여 다음과 같이 표시할 때,

$$f(x) = (f_m(x), f_{m-1}(x), \cdots, f_2(x), f_1(x))$$

Z_2^k의 한 원소 $z = (z_k, z_{k-1}, \cdots, z_2, z_1)$를 정수 $\sum_{i=1}^{k} z_i 2^{i-1}$로 나타낼 수 있다. 또한 함수 $f: Z_2^n \rightarrow Z_2^m$를 이진 n-벡터(tuple)로 나타내면 다음과 같다.

$$\langle f \rangle = [f(0), f(1), \ldots, f(2^n - 1)]$$

이를 함수 f의 정수 표현 (integer representation)이라 부른다. 이 표현은 $\langle f_m \rangle$, $\langle f_{m-1} \rangle$, $\cdots$, $\langle f_2 \rangle$, $\langle f_1 \rangle$를 결합한 식 $\langle f \rangle = \sum_{j=1}^{m} \langle f_j \rangle \cdot 2^{j-1}$ 로 표현한다.

[정의 8]

(1) 임의의 입력 한 비트가 변할 때마다 각각의 출력 비트가 1/2의 확률을 가지고 변하는 함수를 애벌런시 효과 (avalanche effect)를 만족한다고 한다.

(2) 임의의 정수 $i(1 \leq i \leq n)$에 대하여 함수 $f: Z_2^n \rightarrow Z_2^m$이 애벌런시 효과를 만족할 필요 충분 조건은 다음과 같다.

$$\sum_{x \in Z_2^n} wt(f(x) \oplus f(x \oplus c_i^{(n)})) = m 2^{n-1}$$

여기서 $wt()$는 해밍 중 함수를 나타낸다.

즉, 애벌런시 효과란 입력 비트 중 1비트가 변할 때 출력 비트의 절반이 변한다는 것을 의미한다.

[정의 9]

(1) 함수 $f: Z_2^n \rightarrow Z_2^n$ 가 주어졌을 때, 임의의 (i, j) $1 \leq i, j \leq n$에 대하여 다음 조건을 만족하는 $x, x' \in Z_2^n$이 존재할 때, "f 는 complete하다" 라고 말한다.

① x와 x'는 i번째 비트에서만 서로 다르고 다른 비트는 동일하다.

② $f(x)$와 $f(x')$는 j번째 비트가 서로 다르다.

(2) 임의의 정수 $i(1 \leq i \leq n)$에 대하여 함수 $f: Z_2^n \rightarrow Z_2^m$ 가 complete할 필요 충분 조건은 다음과 같다.

$$\sum_{x\in Z_2^n}(f(x)\oplus f(x\oplus c_i^{(n)}) > (0,\ 0,\ \cdots,\ 0)$$

여기서 덧셈은 Z^m에서 비트별 덧셈을 나타낸다.

[정의 10]

임의의 정수 $i(1\le i\le n)$에 대하여 함수 $f: Z_2^n \to Z_2^m$가 다음 식을 만족할 때 "f는 SAC (strict avalanche criterion)를 만족한다" 또는 "f는 강력한 S-box이다"라고 말한다.

$$\sum_{x\in Z_2^n} f(x)\oplus f(x\oplus c_i^{(n)}) = (2^{n-1},\ 2^{n-1},\ \cdots,\ 2^{n-1})$$

SAC를 만족하는 함수 f에 대하여, 단 하나의 입력 비트가 변화할 때마다 각각의 출력비트가 변화될 확률은 1/2이다. 참고문헌[8]에서는 상기 SAC 설계 조건 뿐 만 아니라 S-box의 출력간 상관 계수 (correlation coefficient) 설계 조건을 다음과 같이 요약하고 있다.

(1) 입력 n비트에 대하여 2^n 입력이 모두 가능하여야 한다.

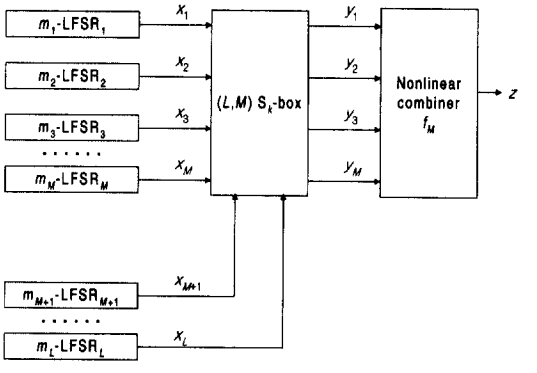
(2) 임의의 입력 비트 i를 변화시켜 출력 비트 j가 변화될 확률은 $p_{ij}\approx 0.5$이며, 본 판정의 수용범위는 $0.375\le p_{ij}\le 0.625$ 이다.

(3) 모든 가능한 입력 벡터 x에 대하여 출력 i, j의 상관 계수는 0에 근사하여야 한다.

$$\rho_{ij}(x)=\frac{cov(i,j)}{\sigma(i)\sigma(j)}$$

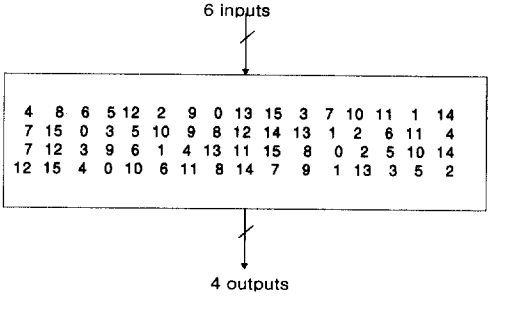
3.2 S-box 형태의 다 수열 발생기

입력 비트 수 L, 출력 비트 수 $M(M\le L)$이고, 키 k에 따라 변화되는 $(L,M)S_k$-box는 SAC 조건 및 상관 계수 특성을 만족하여야 하며, 한 가지 예로 (그림 5)와 같은 (6,4) S-box를 발생시켜 설계기준에 적합한지 여부를 판단하였다. 그리고 설계된 S-Box는 k에 따라 내용이 바뀌어 지며 (새로운 테이블 생성/체크/구성), 예제의 S-Box에 대한 SAC 및 상관 계수 분석 결과는 각각 <표 1> 및 <표 2>와 같다. 표에서는 설계된 S-box는 입출력 비트 변화 확률이 모두 $0.375\le p_{ij}\le 0.625$ 범위를 만족하였으며, 그 평균값은 0.497로 기존 DES의 평균치 0.641[8]보다 0.5에 근사함을 알 수 있다. S-Box 출력 비트간 평균 상관 계수 역시 0.024로 DES의 평균값 -0.163보다 0에 더 근사하였다. (그림 4)에서 제안된 S-box 형태의 이진 수열 발생기는 이러한 S- box를 N개 저장하여 두었다가 키 값이 바뀌면 새로운 S-box를 선택케함으로서 키 수열 사이클 수를 증가시킬 수 있게 된다. 이 때 제안 발생기는 키에 따라 S-box가 변경되므로 결과적으로 N개의 사이클을 갖게 된다.



Note: $(L,M)S_k$-box selected by key from N pre-generated boxes

(그림 4) S-box 형태의 다수열 발생기



(그림 5) (6, 4) S-Box 설계 예

<표 1> S-Box의 입출력 비트 변화 확률 p_{ij}

		Outputs p_{ij}			
	i \ j	1	2	3	4
Inputs	1	0.562500	0.500000	0.500000	0.500000
	2	0.500000	0.500000	0.500000	0.500000
	3	0.500000	0.500000	0.500000	0.500000
	4	0.500000	0.500000	0.500000	0.500000
	5	0.500000	0.500000	0.500000	0.500000
	6	0.437500	0.500000	0.500000	0.437500
Mean		0.497396			

〈표 2〉 S-Box의 입출력 상관관계 계수 $\rho_{ij}(k)$

Inputs	Outputs $\rho_{ij}(k)$					
	k	ρ_{12}	ρ_{13}	ρ_{14}	ρ_{23}	ρ_{24}	ρ_{34}
	1	0.000000	0.375000	0.000000	-0.125000	-0.125000	0.251976
	2	0.000000	-0.250000	0.125000	0.000000	-0.125000	0.000000
	3	-0.251976	-0.125000	-0.125000	0.000000	-0.125000	0.238095
	4	0.375000	0.000000	0.000000	0.000000	-0.125000	-0.125000
	5	-0.125000	-0.125000	-0.125000	-0.125000	-0.250000	0.125988
	6	-0.125000	0.000000	0.125000	-0.250000	0.000000	0.125988
Mean	-0.024581						

제안된 키 수열 발생기의 출력 수열에 대한 비도 특성을 분석한 결과는 정리 11와 같다.

[정리 11]

사용된 모든 LFSR의 크기가 $\gcd(m_i, m_j) = 1, 1 \le i, j(i \ne j) \le L$(relative prime)이고, 모든 LFSR의 초기값이 non-null이 될 때 S-box 형태의 다수열 발생기의 주기, 선형복잡도, 랜덤 특성 및 키 수열 사이클 수는 다음과 같다. 단, 사용된 f 함수는 "0"와 "1"의 발생 확률이 각각 1/2인 balance 함수이어야 한다.

(1) 주기 $P = \prod_{i=1}^{L}(2^{m_i}-1)$ 이다.

(2) 선형복잡도 $LC(z) = g_M^*(L_1, L_2, \ldots, L_M)$ 이다.
여기서, $L_i = LC(y_i)$, i = 0, 1, ..., M 이다.

(3) 랜덤 특성 : 양호함.

(4) 키 수열 사이클 수는 S-box의 개수와 같은 N 이다.

[증 명]

(1) Boolean 함수의 일반화된 ANF(algebraic normal form) 형태로 표현될 수 있는 일반화된 함수(그림 2)에서의 주기는 참고문헌[10]에 따라 $P = \prod_{i=1}^{L}(2^{m_i}-1)$로 구할 수 있다. (그림 4)에서 제안된 함수 역시 S-Box와 f_M 함수를 합하여 Boolean 함수의 일반화된 ANF 형태로 재 표현될 수 있기 때문에 주기는 변함이 없다.

(2) Boolean 함수의 일반화된 ANF 형태로 재 표현된 함수를 $g_M(x_1, x_2, \cdots, x_M)$ 라 할 때 정리 2와 같이 $LC(z) = g_M^*(L_1, L_2, \ldots, L_M)$를 얻을 수 있다.

(3) 발생기 출력 함수는 "0"과 "1"에 대한 balance 함수가 선택되어야 하며, S-box 역시 balance 함수이다. 그리고 함수에 입력되는 수열 (LFSR) 출력 역시 랜덤하고 "0"-"1" balance 하기 때문에 발생기의 출력 역 "0"-"1" balance 하다[1].입력의 랜덤성이 아주 좋기 때문에 출력 수열 역시 랜덤성이 우수하다.

(4) 새로운 함수의 키 수열 사이클 수는 S-box 내용이 키의 선택에 따라 바뀌어지기 때문에 재 표현될 함수 역시 변경되며, 이러한 변화의 개수는 사전에 정의된 N이 된다.

정리 11의 검증을 위하여 짧은 단수에 대한 시뮬레이션 결과는 <표 3> 및 <표 4>에 요약되어 있다. 본 발생기 출력은 (그림 4)와 같이 여러 개(본 시뮬레이션에서는 6개)의 LFSR과 S-box를 통과한 후 $z = f_4(y_1, y_2, y_3, y_4) = y_1 \oplus y_2 \oplus y_3 \oplus y_4$를 통하여 발생된다. 비교 대상인 유사 함수는 $z = f_6(x_1, x_2, x_3, x_4, x_5, x_6) = x_1 \oplus x_2 \oplus x_3 \oplus x_4 \oplus x_5 \oplus x_6$ 형태로 단순화시켜 구성하였다. 또한 LFSR 구성을 위한 원시 다항식은 참고 문헌[11]에 따라 다음과 같이 선택하였다.

$$g_1(x) = x^2 + x + 1$$
$$g_2(x) = x^3 + x + 1$$
$$g_3(x) = x^5 + x^4 + x^3 + x + 1$$
$$g_4(x) = x^7 + x^6 + x^4 + x + 1$$
$$g_{5(x)} = x^{11} + x^{10} + x^3 + x + 1$$
$$g_6(x) = x^{13} + x^7 + x^3 + x + 1$$

<표 3>에서 랜덤 특성 분석을 위한 시험 항목은 Beker와 Piper[1]의 항목을 따랐으며, 3가지 샘플에 대한 분석 결과는 모두 3.1절에서 정의된 S-box의 출력간 상관계수 문턱 값(threshold) 이하인 조건을 만족하므로 전체적인 랜덤 특성이 양호함을 알 수 있다.

〈표 3〉 랜덤 특성 검증 결과

Test items	Threshold	Test results		
		Sample 1	Sample 2	Sample 3
1) Frequency test	3.84	0.027	0.005	2.042
2) Serial test	5.99	1.390	0.023	2.233
3) Generalized t-serial test				
t = 3	9.48	6.919	1.836	3.696
t = 4	15.50	12.459	3.412	5.462
t = 5	26.29	23.057	12.727	7.762
4) Poker test				
m = 3	14.067	6.185	7.035	6.850
m = 4	24.996	17.287	7.617	17.510
m = 5	44.654	37.304	24.487	20.592
5) Autocorrelation test	max. ≤0.05	max = 0.0060	max = 0.0063	max = 0.0072

〈표 4〉 유사 수열 발생기 비교
(작은 단수에 대한 시뮬레이션)

Items	Original function f_6	Proposed type
Period (1)	$P_1 \approx 10^{12}$	$P_2 \approx 10^{12}$
Randomness	Good	Good (Table 3)
Linear complexity	$LC_1 = 41$ (2)	$LC_2 \geq LC_1$ (4)
Number of keystream cycles	$N_1 = 1$ (3)	Able to variable choice ($1 \leq N_2 \leq$ number of key, selective)

주 : (1) $P=(2^2-1)(2^3-1)(2^5-1)(2^7-1)(2^{11}-1)(2^{13}-1)\approx 10^{12}$
(2) $LC_1 = 2+3+5+7+11+13 = 41$, from reference[3]
(3) $N_1 = 1$, no variable parameters in function.
(4) $LC_2 \geq LC_1$, because S-box contains 2~6-variable product terms.

결과적으로 비선형성이 높은 S-Box를 스트림 암호에 적용하는 방법으로 여러 개의 S-Box를 메모리에 사전 발생 및 저장해두고 키 값이 바뀔 때마다 다른 S-Box가 선택되도록 한다면 스트림 암호에서도 비도 요소를 높이는 방향으로 적용 가능하며, 이러한 유형의 발생기는 기존의 발생기와 비교할 때 주기와 선형 복잡도는 비슷한 수준으로 유지시키면서 키 수열 사이클 수를 크게 높일 수 있게 된다.

4. 결 론

지금까지 제안된 대부분의 이진 수열 발생기의 경우 키 수열 사이클 수가 단 한 개뿐인 단일 사이클 발생기로서 Dawson 공격에 취약성을 갖고 있으며, 이를 보완하기 위해서 본 논문에서는 다 수열 발생기의 출력 사이클 수에 대한 모델을 제안하였다. 한편, S-box는 그 특성상 비선형성이 높아 블록 암호의 핵심 요소이며, 스트림 암호에 적용시 비도 요소를 개선시킬 수 있는 장점이 있음을 살펴보았다. 즉, 여러 개의 (M, L) S-Box를 미리 발생시켜서 메모리에 저장해두고 키 값이 바뀔 때마다 다른 S-Box를 선택케 하면 스트림 암호에서도 비도 요소를 높이는 용도로 적용시킬 수 있음을 보였다. 이러한 S-box의 우수한 비선형성 특성을 이용하여 본 논문에서는 키 수열 사이클 수가 여러 겹인 S-box 형태의 다수열 발생기를 제안하였으며, 그 비도 특성을 분석하였다. 마지막으로, 이론적 결과를 확인하기 위하여 유사 형태의 발생기를 선정하여 작은 단수에 대한 시뮬레이션을 실시하였으며, 그 결과 제안된 발생기는 기존의 키 수열 발생기와 비슷한 수준의 비도 요건 (주기, 랜덤 특성 및 선형 복잡도)를 유지하면서도 키 수열 사이클 수를 크게 개선시킬 수 있음을 알 수 있었다.

참 고 문 헌

[1] Henry J. Beker and Fred C. Piper, 'Cipher systems : The Protection of Communications,' Northwood Books, London, 1982.

[2] Henk C. A. van Tilborg, 'An Introduction to Cryptology,' KLUWER ACADEMIC PUBLISHERS, Boston, etc., 1988.

[3] T. Siegenthaler, "Correlation-Immunity of Nonlinear Combining Functions for Cryptographic Applications," IEEE Trans. on Infor. Theo., Vol.IT-30, No.5, pp.776-780, Sep. 1984.

[4] J. Dj. Golic, "The Number of Output Sequences of a Binary Sequence Generator," LNCS 547, Advances in Cryptology-EUROCRYPT'91, pp.160-167, 1991.

[5] 이훈재, 문상재, "다수열 출력 이진 수열 발생기", 한국통신정보보호학회 논문지, 제7권 제3호, pp.11-22, 1997.

[6] E. Dawson, L. Nielsen, "Automated Cryptanalysis of XOR Plaintext Strings," Cryptologia, Vol.XX, No.2, pp.165-181, Apr. 1996.

[7] Philip R. Geffe, "How to Protect Data with Ciphers that are really hard to Break," Electronics, pp.99-101, Jan. 1973.

[8] 최희동, 노종선, "PN 시퀀스를 이용한 S-box 설계에 관한 연구", 한국통신학회 논문지, 제20권 제2

호, pp.319-326, 1995.

[9] R. A. Rueppel and O. J. Stafflebach, "Products of Linear Recurring Sequences with Maximum Complexity," IEEE Trans. on Infor. Theo., Vol.IT-33, No.1, pp.124-131, Jan. 1987.

[10] J. Dj. Golic, "On the Linear Complexity of Functions of Periodic GF(q) Sequences," IEEE Trans. on Infor. Theo., Vol.IT-35, No.1, pp.69-75, Jan. 1989.

[11] B. Park, H. Choi, T. Chang and K. Kang, "Period of Sequences of Primitive Polynomials," Electronics Letters, Vol.29, No.4, pp.390-391, Feb. 1993.

이 훈 재

e-mail : hjlee@kyungwoon.ac.kr

1985년 경북대학교 전자공학과 졸업(공학사)

1987년 경북대학교 대학원 전자공학과 정보통신전공 (공학석사)

1998년 경북대학교 대학원 전자공학과 정보통신전공 (공학박사)

1987년~1998년 국방과학연구소 제5개발본부 선임연구원

1998년~현재 경운대학교 컴퓨터전자정보공학부 조교수

관심분야 : 암호 및 정보보호, 전자상거래, 정보통신망, 디지털 통신