수축생성기에 기반한 비선형 수열의 분석

조성진[*] · 최언숙[**] · 김한두[***] · 안현주[****]

Analysis of nonlinear sequences based on shrinking generator

Un-sook Choi[*] · Sung-jin Cho[**] · Han-doo Kim[***] · Hyun-joo An[****]

요 약

본 논문에서는 수축생성기(Shrinking Generator)의해 생성되는 비선형수열의 성질을 분석한다. 또한 수축생성기에 의해 생성되는 비선형 수열을 삽입수열로 해석하여 제어레지스터에서 생성되는 수열의 성질을 이용하여 출력된 수축수열의 위상이동차를 분석하여 가로챈 일부 수열로부터 원래 수열을 복원해내는 방법을 제안한다.

ABSTRACT

In this paper, we analyze the properties of nonlinear sequence generated by the shrinking generator. Also we propose a method for recovering the original sequence from intercepted bits by analyzing phase shifts of the output sequence using the properties of sequences generated from control register.

키워드

phase shift, 수축수열, 삽입수열, 수축생성기, 의사난수열

Ⅰ. 서 론

스트림암호는 대규모의 데이터를 매우 빠르게 암호화하기 위해 사용되는 비밀키 암호시스템이다. 스트림암호의 키스트림은 통상 LFSR들을 이용하여 생성된다. 키스트림의 부분 정보로부터 키스트림을 완전히 알아낼 수 없을 때 키스트림이 안전하다고 한다. 일반적으로 스트림암호의 키스트림은 긴 주기, 높은 선형복잡도를 가지고, 한 주기 내에서 1의 개수와 0의 개수의 차가 1이하여야 안전하다고 한다. 스트림암호는 긴 주기의 난수열을 발생시켜 전송하고자 하는 평문과 비트별 XOR 연산을 하여 암호문을 생성하는 방식으로서 유사 난수열을 발생시켜야 한다. 키스트림에 비선형성을 내재시키기 위한 시도로서 LFSR에 불규칙 클럭 신호를 보내는 방법이 있다. 한 LFSR의 클럭 작동이 다른 LFSR에 의해 제어되도록 설계한다면 특별한 용도로 사용되는 중요한 수열을 얻을 수 있다.

1994년 Coppersmith 등에 의해 제안된 수축생성기는 3개의 LFSR에 의해 구현된 교대단계생성기(alternating step generator)에 비해 작동 방법이 간단하고, 구현이 용이하며 고속 암호화가 요구되는 응용분야에 적합한 암호로 인식되고 있다[1]. 1995년 Gong 등은 삽입수열(interleaved sequence)을 정의하고 수축수열(shrunken sequence)을 삽입수열로 해석하여 분석하였다[2]. Cho 등은 의사난수열을 출력할 수 있는 최대주기 유한상태 기계의 위상이동차(phase

* 부경대학교 응용수학과(sjcho@pknu.ac.kr) ** 교신저자 :동명대학교 미디어공학과(choies@tu.ac.kr)
*** 인제대학교 컴퓨터응용과학부, 기초과학연구소(mathkhd@inje.ac.kr) **** 부경대학교 응용수학과(hjan@pknu.ac.kr)
접수일자 : 2010. 06. 29 심사(수정)일자 : 2010. 07. 09 게재확정일자 : 2010. 08. 05

shift)를 계산하는 방법을 연구하였다[3-4].

최근 Sabater 등은 수축수열을 삽입수열로 해석하여 수축생성기에 의해 생성된 수열 중 일부를 알 때 90/150 셀룰라 오토마타를 이용하여 알지 못하는 새로운 비트스트림을 재구성하는 방법을 제안하였다[5].

본 논문에서는 수축생성기에 의해 생성된 수축수열을 삽입수열로 해석하고, 제어레지스터에서 생성되는 수열의 성질을 이용하여 출력된 수축수열의 위상이동차를 분석하여 가로챈 일부 수열로부터 원래 수열을 복원해내는 방법을 제안한다.

II. 배경지식 및 기존연구

키스트림에 비선형성을 내재시키기 위한 시각제어 생성기(clock-controlled generator)는 두 개의 레지스터로 구성된 생성기이다. 시각제어생성기는 규칙적으로 클럭이 주어지는 제어 레지스터(control register)와 제어 레지스터의 상태 값에 따라 클럭 수가 결정되는 생성 레지스터(generator register)로 구성되었다. 두 개의 LFSR R_1과 R_2를 이용하는 키스트림 생성기인 수축생성기는 제어 레지스터 R_1이 생성하는 수열을 이용하여 생성 레지스터 R_2가 생성하는 수열을 일부 선택하여 키스트림을 생성한다. 따라서 수축생성기에 의해 생성되는 키스트림은 R_2에 의해 생성되는 키스트림의 수축된 키스트림이다. 수축된 이 수열를 수축수열이라 한다. 그림 1은 수축생성기의 구조이다.

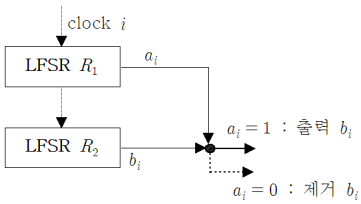

그림 1. 수축생성기의 구조
Fig. 1 The structure of shrinking generator

R_1과 R_2에 클럭신호가 주어질 때마다 R_1의 출력이 1이면 R_2의 출력이 키스트림에 포함되고, R_1의 출력이 0이면 R_2의 출력은 배제된다. 이 수축생성기는 생성된 키스트림의 비트의 위치가 고정되지 않는다는 점이 장점이고, 긴 주기, 높은 선형복잡도(linear complexity), 우수한 통계적 특성 등 암호학적으로 좋은 성질을 가지고 있다. LFSR R_1과 R_2의 특성다항식이 각각 L_1차, L_2차의 원시다항식일 때, $\gcd(L_1, L_2) = 1$인 경우 주기는 $(2^{L_2}-1)\cdot 2^{L_1-1}$이고, 선형복잡도 LC는 식 (1)과 같다[2].

$$L_2 \cdot 2^{L_1-2} \le LC \le L_2 \cdot 2^{L_1-1} \quad \cdots\cdots (1)$$

a_i: 000100110101111000100110101111...
b_i: 0000101011101100011111001101001 0...
c_i: 0101011011010100...

그림 2. 수축수열의 생성과정
Fig. 2 The process of generation of a shrunken sequence

R_1의 길이가 4, 특성다항식이 x^4+x+1, 초기수열벡터가 {0001}이면 R_1에 의해 생성되는 수열 $\{a_i\}$는 최대주기수열인 000100110101111...이고 주기는 15이다. R_2의 길이가 5, 특성다항식이 x^5+x^3+1, 초기수열벡터가 {00001}이면 R_2에 의해 생성되는 수열 $\{b_i\}$는 최대주기수열인 00001010111011000111111001 101001...이고 주기는 31이다. 이때 수축생성기에 의해 생성되는 수축수열 $\{c_i\}$는 주기가 $2^{4-1}\times 31 = 248$인 수열로 그림 2는 수축수열의 생성과정을 보여준다.

Sabater 등은 수축수열을 삽입수열로 해석하여 수축생성기에 의해 생성된 수열 중 일부를 알 때 90/150 셀룰라 오토마타를 이용하여 알지 못하는 새로운 비트스트림을 재구성하는 방법을 제안하였다[5]. 이 방법은 주어진 생성기로부터 출력된 키스트림 중 일부를 알고 있을 때 이를 이용하여 알지 못하는 새로운 비트들을 계산하는 방법으로 여기서 90/150 CA를 합성하고 이를 이용하여 키스트림을 이용하여 부분삼각형을 만들어 다시 유한체를 이용하여 새로운 비트를 재구성한다. 이 방법은 계산이 복잡하고 무엇보다 출력수열을 일부만 알아내는 것이지 모두 알아내기는 어렵다는 문제점을 가지고 있다. 본 논문에서는 이러한 문제점을 극복하고 제어레지스터에 의해 생성되는 수열의 특성과 수축생성기에 의해 생성된 수축수열의 위상이동차를 계산하여 가로챈 키스트림의 일부를 이용하여 출력수열 전체를 알 수 있는 방

법을 제안한다.

III. 위상이동차를 이용한 수축수열의 분석

m차 원시다항식(primitive polynomial)에 의해서 생성되는 0이 아닌 수열을 주기 2^m-1인 *PN수열(Pseudo-Noise sequence)*이라 한다. PN수열에서 가장 주목해야 할 점은 *런분포 성질(run-distribution property)*이다. 이진 수열에서 0 또는 1이 연이어 나타나는 부분을 주어진 수열의 *런(run)*이라 하고, 특히 0만으로 이루어진 런과 1만으로 이루어진 런을 각각 이 수열의 *갭(gap), 블록(block)*이라고 한다.

<예제 1> 초기수열벡터가 {0001}일 때 원시다항식 x^5+x^2+1에 의해 생성된 PN수열 00001001011001111100011011101 01에는 길이가 4인 갭이 1개, 길이가 3인 갭이 1개, 길이가 2인 갭이 2개, 길이가 1인 갭이 4개 있고, 길이가 5인 블록이 1개, 길이가 3인 블록이 1개, 길이가 2인 블록이 2개, 길이가 1인 블록이 4개 있다.

위 예제를 일반화 하면 2^m-1인 PN수열의 런분포는 표 1과 같다.

표 1. PN수열의 런분포
Table 1. The run distribution of PN sequence

길이	0-runs	1-runs
1	2^{m-3}	2^{m-3}
2	2^{m-4}	2^{m-4}
⋮	⋮	⋮
r	2^{m-r-2}	2^{m-r-2}
⋮	⋮	⋮
$m-2$	1	1
$m-1$	1	0
m	0	1
합계	2^{m-2}	2^{m-2}

수축생성기에서 R_1의 길이가 4, 특성다항식이 x^4+x+1, 초기수열벡터가 0001이고, R_2의 길이가 5, 특성다항식이 x^5+x^2+1, 초기수열벡터가 00001일 때, 생성된 수축수열을 R_1에 의해 생성된 길이 2^4-1인 PN수열의 한 주기에 1의 개수가 8개이므로 수축수열을 8비트씩 행으로 나열하면 그림 3과 같은 31×8행렬로 표현되는 삽입수열을 얻는다.

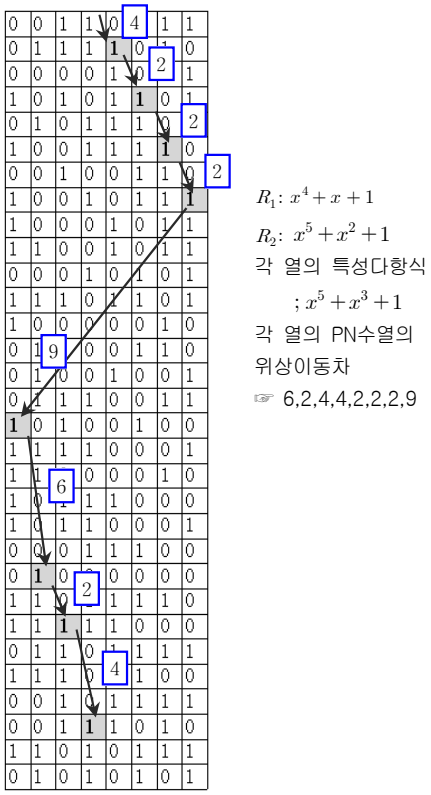

그림 3. 삽입수열로 표현된 수축수열 ($L_1=4$, $L_2=5$)
Fig. 3 The shrunken sequence represented with the interleaved sequence($L_1=4$, $L_2=5$)

이렇게 얻은 삽입수열의 8개 열은 x^5+x^2+1에 의해 생성된 PN수열 중에서 x^4+x+1에 의해 생성된 PN수열의 1이 있는 위치에서 15칸씩 건너뛰면서 만들어진 PN수열로서, x^5+x^2+1의 상반다항식(reciprocal polynomial)에 의해 생성된 PN수열들이다. 삽입수열의 각 PN수열들은 다른 PN수열들의 위치를 이동함으로써 얻을 수 있다. 8개의 PN수열의 위상이동차는 R_1에 의해 생성된 PN수열로부터 분석이 가능하다.

표 1에 의하면 x^4+x+1에 의해 생성된 주기 2^4-1인 PN수열 000100110101111에 길이가 4인 블록이 1개이므로 위상이동차 2가 연이어 3번 나타나고,

길이가 2인 블록이 1개이므로 위상이동차 2가 한 번 더 나타난다. 한편 길이가 1인 갭이 2개 존재하므로 위상이동차 4가 한번 나타나고 길이가 2인 갭이 1개 존재하므로 위상이동차 6이 한번 나타난다. PN수열의 주기가 31이므로 위상이동차 31-(6+2+4+4+2+2+2)=9가 한번 나타난다.

[정리 1] 차수가 L_1인 원시다항식으로 구성된 LFSR R_1과 차수가 L_2인 원시다항식으로 구성된 LFSR R_2로 이루어진 수축생성기에 의해 생성된 수축수열을 $(2^{L_2}-1)\times 2^{L_1-1}$ 행렬로 표현되는 삽입수열로 나타내었을 때, R_1에 의해 생성된 PN수열에서 1이 처음 발생한 위치에서 발생한 PN수열의 위상이동차 n은 식 (2)를 만족한다.

$$(i+1)+(2^{L_1}-1)n \equiv i \pmod{2^{L_2}-1} \quad \cdots\cdots (2)$$

그림 3의 예제에서 $L_1=4, L_2=5$이다. x^4+x+1에 의해 생성된 주기 2^4-1인 PN수열 000100110101111에서 $2^{4-1}=8$개의 1중에서 1이 연이어 나타나는 곳이 두 번째와 세 번째이므로 삽입수열의 2열과 3열의 PN수열의 위상이동차 n은 $1+15\cdot n=31$을 풀면 2임을 알 수 있다.

[정리 2] 차수가 L_1인 원시다항식으로 구성된 LFSR R_1과 차수가 L_2인 원시다항식으로 구성된 LFSR R_2로 이루어진 수축생성기에 의해 생성된 수축수열을 $(2^{L_2}-1)\times 2^{L_1-1}$ 행렬로 표현되는 삽입수열로 나타내었을 때, R_1에 의해 생성된 PN수열에서 1이 처음 발생한 위치에서 발생한 PN수열을 삽입수열의 1열이라 하자. R_1에 의해 생성된 PN수열에서 i번째 1과 $i+1$번째 1사이에 0이 k번 나타난다면 삽입수열의 i열과 $i+1$열의 PN 수열의 위상이동차 n은 식 (3)을 만족한다.

$$\{(i+1)+k\}+(2^{L_1}-1)n \equiv i \pmod{2^{L_2}-1} \quad \cdots\cdots (3)$$

그림 3의 예제에서 x^4+x+1에 의해 생성된 주기 2^4-1인 PN수열 000100110101111에서 $2^{4-1}=8$개의 1중에서 첫 번째 1과 두 번째 1사이에 0이 두 번 나타나므로 삽입수열의 1열과 2열의 PN수열의 위상이동차 n은 $3+15\cdot n=93$을 풀면 6임을 알 수 있다. 같은 방법을 이용하여 2열부터 8열까지의 위상이동차를 계산하면 6, 2, 4, 4, 2, 2, 2이고 8열과 1열사이의 위상이동차는 31-(6+2+4+4+2+2+2)=9임을 알 수 있다. 그러므로 R_1에 의해 생성되는 PN수열의 1의 위치를 알면 수축수열을 삽입수열로 표현했을 때, 각 삽입수열의 위상이동차를 알 수 있다. 이러한 특성은 출력수열 중 일부만을 알 때 위상이동차를 이용하여 나머지 수열 전체를 알 수 있음을 보여주고 있다.

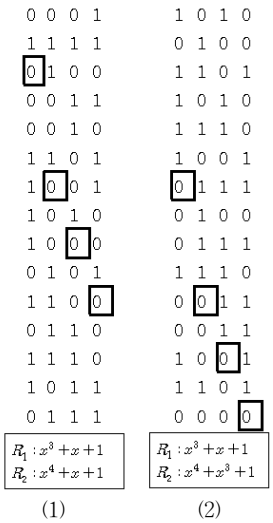

그림 4. R_1이 같고 R_2가 다른 수축생성기
Fig. 4 shrinking generators with the same R_1 and different R_2

그림 4는 R_1이 같고 R_2는 길이만 같고 그 특성다항식이 다른 경우 두 수축생성기에서 출력된 키스트림을 삽입수열로 해석한 경우이다. 각 생성기에서 출력된 두 수열은 그 특성다항식이 서로 다르지만 주어진 수열을 2^{L_1-1}개의 열로 나열하여 삽입수열로 해석하였을 때 두 생성기의 수열은 달라도 각 열 사이의 위상이동차가 4, 2, 2, 7로 동일하다. 이러한 사실은 수축생성기의 구성요소인 R_1, R_2중 R_1의 정보가 더

중요함을 알 수 있다. 즉 R_1의 정보에 대해 취약하다는 것이다. 그림 3의 예제는 [5]에서 분석한 수축생성기와 동일한 주기의 수열을 생성하는 수축생성기이다. [5]에서 분석한 자료와 동일한 조건으로 가로챈 출력수열의 크기를 24비트라 하고 이 수열이 {110010110001010111101101}라 하자. R_1의 특성다항식이 x^4+x+1이므로 초기벡터가 {0001}일 때 생성되는 수열의 주기는 {000100110101111}이므로 R_2의 길이가 5이므로 출력수열의 위상이동차는 8, 6, 2, 4, 4, 2, 2, 2가 된다. 가로챈 수열을 2^{L_1-1}개씩 잘라 행으로 나열하면 그림 5와 같다.

1	0	0	0	1	0	1	1
1	1	0	0	1	0	1	1
0	0	0	1	0	1	0	1

그림 5. 가로챈수열
Fig. 5 Intercepted sequence

PN수열의 특성 중 4차 PN수열은 반드시 런의 길이가 4인 경우가 1개 있으므로 이는 위상이동차가 2인 경우가 연이어 3번 나온다는 것을 의미한다. 그림 5에서 위상이동차가 2가 될 가능성이 있는 부분은 4열과 5열부터 6열과 7열 사이일 수 있으며 또는 5열과 6열부터 7열과 8열 사이가 될 수 있다. 그런데 주어진 R_1의 특성다항식에 의해 만들어지는 수열을 정확히 알고 있으므로 연이은 위상이동차 2, 2, 2 앞서 위상이동차가 2인 부분을 확인하면 그림 6의 윗부분에 있는 각 열의 위상이동차가 되며 첫 째 열의 위상이동차 9는 마지막 열과 첫 째 열사이의 위상이동차로 31에서 위상이동차 7개의 값을 모두 더하여 뺀 값으로 계산된다. 그림 6은 각 열에 대한 위상이동차를 바탕으로 각 열사이의 관계를 통하여 점차적으로 새롭게 알게 되는 출력비트들을 찾아가는 과정이다. 마지막 열에서 알게 된 비트수가 9개 이므로 이를 이용하여 출력된 수축수열의 특성다항식을 구할 수 있게 되므로 나머지 부분까지 모두 알 수 있다.

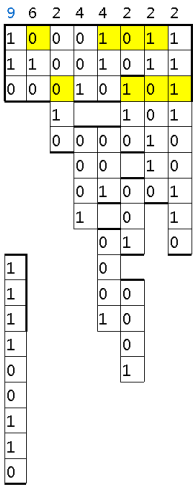

그림 6. 위상이동차를 이용한 출력수열의 복원과정
Fig. 6 The reconstruction process of output sequence using the phase shifts

Ⅳ. 결 론

본 논문에서는 수축생성기에 의해 생성되는 수축수열을 삽입수열로 해석하여 위상이동차를 분석하여 일부 수열로부터 원래 수열을 복원해내는 새로운 공격방법을 제안하였다. 이 결과는 Sabater 등이 수축생성기에 의해 생성된 수열 중 일부를 알 때 90/150 셀룰라 오토마타를 이용하여 알지 못하는 새로운 부분 비트스트림을 재구성하는 방법을 제안한 방법과 달리 출력된 수축수열의 위상이동차를 분석하여 가로챈 일부 수열로부터 원래 수열 전체를 복원해내는 방법을 제안하였다.

참고 문헌

[1] D. Coppersmith, H. Krawczyk, Y. Mansour, “The shrinking generator,” LNCS, Vol. 773, pp. 22-39, 1994.
[2] G. Gong, “Theory and applications of q-ary interleaved sequences,” IEEE Trans. Inform. Theory, Vol. 41(2), pp. 400-411, 1995.
[3] S.J. Cho, U.S. Choi, Y.H. Hwang, Y.S. Pyo, H.D. Kim and S.H. Heo, “Computing phase shifts of maximum-length 90/150 cellular

automata," LNCS, Vol. 3305, pp.31-39, 2004.
[4] S.J. Cho et al., "Phase Shifts of LFSM as Pseudorandom Number Generators for BIST for VLSI," Logic and Theory of Algorithms, Proc. CiE 2008, pp.77-86, 2008.
[5] A. Fuster-Sabater, P. Caballero-Gil, "Concatenated automata in cryptanalysis of stream ciphers," LNCS, Vol. 4173, pp. 611-616, 2006.
[6] R.J. McEliece, Finite fields for computer scientists and engineers, Kluwer Academic Publishers, 1987.

저자 소개

최언숙(Un-sook Choi)

1992년 2월 성균관대학교 산업공학과 졸업 (공학사)
2000년 2월 부경대학교 대학원 응용수학과 졸업(이학석사)
2004년 2월 부경대학교 대학원 응용수학과 졸업(이학박사)
2009년 8월 부경대학교 대학원 정보보호학과 졸업(공학박사)
동명대학교 미디어공학과 전임강사
※ 주관심분야 : 셀룰라 오토마타론, 정보보호, 암호이론

조성진(Sung-jin Cho)

1979년 2월강원대학교 수학교육과 졸업 (이학사)
1981년 2월 고려대학교 대학원 수학과 졸업(이학석사)
1988년 2월 고려대학교 대학원 수학과 졸업(이학박사)
부경대학교 응용수학과 교수
※ 주관심분야 : 셀룰라 오토마타론, 정보보호, 부호이론, 컴퓨터 구조론

김한두(Han-doo Kim)

1982년 2월 고려대학교 수학과 졸업 (이학사)
1984년 2월 고려대학교 대학원 수학과 졸업(이학석사)
1988년 2월 고려대학교 대학원 수학과 졸업(이학박사)
1989년~현재: 인제대학교 컴퓨터응용과학부 교수
※ 주관심분야 : 셀룰라 오토마타론, 전산수학

안현주(Hyun-joo Kim)

2010년 3월~현재 부경대학교 대학원 응용수학과 석사과정
※ 주관심분야 : 셀룰라 오토마타론, 정보보호