의사랜덤 수열과 정보통신

송민규, 김정현, 송홍엽
연세대학교

Ⅰ. 랜덤 수열과 의사랜덤 수열

수학에서, 일련의 숫자들이 연속적으로 나열된 형태를 "수열"이라고 부른다. 하나의 수열 내에서 나타나는 서로 다른 숫자들의 개수는 무한할 수도 있고, 유한할 수도 있다. 전자의 경우 자연현상을 설명할 때 많이 사용하는 피보나치 수열이, 후자의 경우는 유한한 길이의 비트스트림이 좋은 예이다. 0과 1, 두 개의 숫자(또는 두 개의 문자)로 이루어진 수열을 "2진 수열"이라 한다. 또한, 이러한 수열 중에서 정해진 길이의 하나의 패턴이 지속적으로 반복되는 구조를 갖는 수열을 일컬어 "주기적 수열"이라 한다.

정보통신의 다양한 분야에서 특정 길이의 2진 수열 혹은 비트스트림이 사용된다. 물론, 비트스트림에 요구되는 특성은 사용 목적만큼이나 다양하다. 이중에서 가장 흔하게 요구되는 특성은 랜덤특성이다. 이는 다음 항이 0인지 혹은 1인지가 확률적으로 결정될 뿐, 예측이 불가능함을 뜻한다. 가장 완벽한 랜덤특성은 동전던지기 시행을 연속적으로 반복하여 얻어질 수 있다. 앞면(0)과 뒷면(1)이 나올 확률이 각각 1/2이므로 다음 항이 무엇일지는 확률로 정의될 뿐, 결정적이지 않다는 특성이 있다. 그렇다면 과연 컴퓨터는 어떻게 동전던지기와 같은 랜덤한 시행을 "쉽게" 답습할 수 있을까?

또 하나의 중요한 둘째 요구 사항은 일반적으로 하나의 랜덤 비트스트림이 한 번 사용된 후에 버려지는 것이 아니라 정확히 동일한 랜덤 비트스트림을 다시 발생시켜서 재사용해야 한다는 것이다. 통신시스템의 송/수신기, 암호시스템의 암/복호화기가 이를 요구하는 좋은 예이다.

두 번째 요구 사항을 위해서는, 컴퓨터가 매우 간단한 연산을 반복함으로써 일정 길이의 비트스트림을 만들어내는 것이 이상적이다. 또한 그 결과로 얻어진 비트스트림이 동전던지기의 결과와 구분이 어렵다면, 첫 번째 요구사항을 만족시키기 충분할 것이다. 이러한 랜덤 비트스트림을 "의사랜덤" 수열이라 부른다. 즉, 만드는 방법은 매우 간단하고 단순하며 일정한 연산법칙을 사용하지만, 그 결과는 랜덤한 혹은 랜덤한 것처럼 보이는 비트스트림인 것이다. 이를 또 다른 표현으로 말하자면, 생성법을 모른다면 랜덤하게 보이지만, 생성법을 안다면 전혀 랜덤하지 않다

만드는 방법은 매우 간단하고 단순하며 일정한 연산법칙을 사용하지만, 그 결과는 랜덤한 것처럼 보이는 비트스트림을 의사랜덤 수열이라 부른다.

...01101001110000101010 01...
(a) 랜덤수열

...00010110001011000101 10...
(b) 의사랜덤수열 (길이7, m-수열)

...01010101010101010101...
(c) 비 랜덤수열

그림 1. 다양한 2진 수열의 예

a \ b	0	1
0	0	1
1	1	0

(a) 2진 덧셈

a \ b	0	1
0	0	0
1	0	1

(b) 2진 곱셈

그림 2. 2진 덧셈과 곱셈 연산

는 뜻이다. 또한, 매우 간단하고 단순한 연산법칙을 사용하므로 두 번째 요구사항인 재생산이 쉽게 가능해진다.

본 논문에서는 의사랜덤 수열을 만드는 다양한 방법들 중에서도 현실적으로 가장 많이 사용되는 "선형 궤환 시프트 레지스터(Linear feedback shift register, LFSR)"에 대해 다룬다. 전개상의 편의를 위해, 본문에서 다루는 수열은 2진 수열로 제한한다. 〈그림 2〉는 0과 1 두 개의 원소로 이루어진 집합에서의 덧셈과 곱셈을 나타내고 있다. 〈그림 2〉가 설명하는 덧셈과 곱셈은 2진 수열의 생성과 분석에 사용되는 기본적인 연산법칙이므로 매우 중요하다.

의사랜덤 수열을 만드는 다양한 방법들 중 가장 많이 사용되는 선형 궤환 시스트 레지스터(LFSR)에 대해 다룬다.

2진 선형 궤환 시프트 레지스터 수열에 대한 설명에 앞서, Ⅱ장에서 의사랜덤 수열이 정보통신 시스템에서 사용되는 몇 가지 예를 소개한다. Ⅲ장에서는 정보통신 시스템에서 널리 사용되는 선형 궤환 시프트 레지스터를 이용한 의사랜덤 수열의 생성에 대해 이야기 한다. Ⅳ장에서 선형 궤환 시프트 레지스터 수열의 기본이 되는 m-수열(Maximal length linear feedback shift register sequence, m-sequence)을 소개하고, m-수열의 몇 가지 랜덤특성에 대해 이야기 하며 글을 마무리 한다.

Ⅱ. 의사랜덤 수열의 응용

의사랜덤 수열의 생성에 대해 이야기하기 전에, 정보통신 시스템에서 의사랜덤 수열이 사용되는 예를 통하여 의사랜덤 수열의 중요성을 알아보자. 의사랜덤 수열을 직접적으로 이용하는 대표적인 정보통신 기술로, 정보보호 시스템과 통신 시스템이 있다. 또한, 의사랜덤 수열은 오류정정기술의 한 분야인 오류정정부호와도 밀접한 연관이 있다.

1. 정보보호 시스템

스트림 암호는 의사랜덤 수열을 이용하는 대표적인 정보보호 기법이다[4]. 스트림 암호는

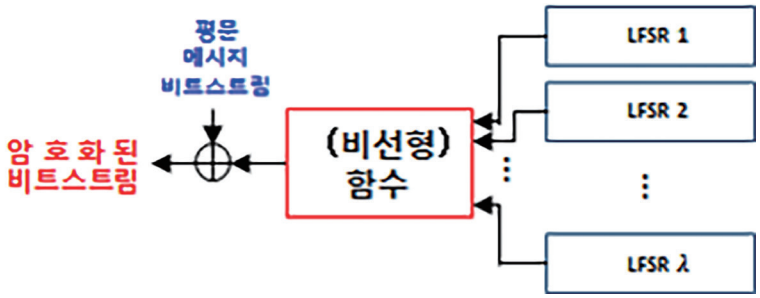

그림 3. 스트림 암호의 일반적인 구성

〈그림 3〉과 같이, 다수의 선형 궤환 시프트 레지스터들의 출력을 가지고 정해진 (비선형) 함수 연산을 수행한 후, 이 결과를 메시지 비트와 2진 덧셈 연산을 수행함으로써 메시지 비트스트림을 암호화 한다. 사용된 선형 궤환 시프트 레지스터들의 궤환함수와 이들을 결합하는 비선형함수가 공개되어도, 초기값에 해당하는 비밀키를 아는 적절한 사용자만이, 암호화 과정에서 메시지 비트스트림과 더해진 수열을 다시 생성하여 메시지 비트스트림을 얻을 수 있다. 이러한 선형 궤환 시프트 레지스터 수열 사용은 선형 궤환 시프트 레지스터 수열이 랜덤 수열과 유사한 특성을 갖는다는 점에 기반한다. 비선형함수의 결합은 암호 키수열(keystream)의 암호학적 강도를 보강한다.

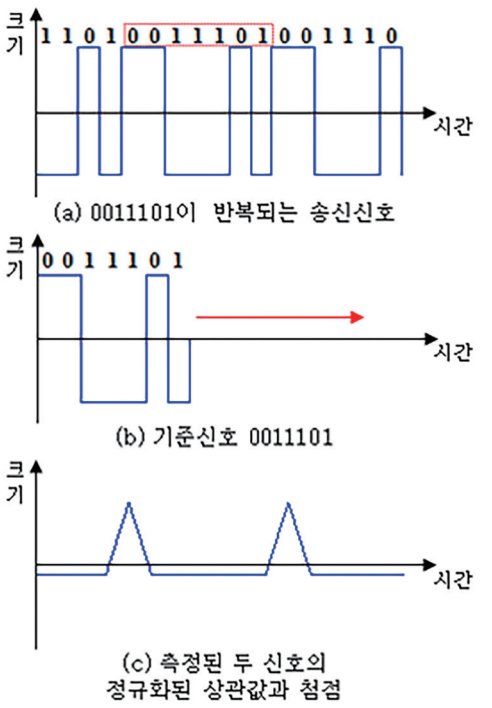

그림 4. 길이 7인 m-수열을 이용한 송수신기의 동기화

2. 통신 시스템

통신 시스템에서 의사랜덤 수열을 사용하는 방법은 정보보호 시스템과는 조금 다르다. 통신 시스템에서는 의사랜덤 수열의 상관 특성을 주로 이용한다. 〈그림 4〉는 통신 시스템에서 의사랜덤 수열을 이용하는 하나의 예이다. 송신기가 사각형 펄스(Rectangular pulse)를 이용하여 주기가 7인 의사랜덤 수열 0011101을 반복적으로 전송하는 시스템을 가정해 보자. 본 예에서 사용한 의사랜덤 수열은 m-수열로, 후에 Ⅳ장에서 자세히 다루어진다. 전송을 위해서 송신기는 비트 0을 +A, 비트 1을 -A의 전압을 갖는 사각형 펄스로 변환한다. 〈그림 4(a)〉는 사각형 펄스로 변환된 송신신호이다. 수신기에서는 송신기가 전송한 신호 0011101을 사각형 펄스로 변화한 "기준신호"(〈그림 4(b)〉 참고)의 시작시간을 바꿔가며 상관도를 측정한다. 측정된 상관도는 m-수열의 자기 상관 특성에 따라 〈그림 4(c)〉과 같이 나타나고, 송신기가 전송한 m-수열 0011101과 기준신호의 시작 시간이 정확히 일치할 때, 상관값이 최대가 된다. 따라서, 수신기는 상관값이 최대가 되는 지점, 즉 첨점(Peak)를 확인하여 수신한 신호에서 m-수열이 시작되는 시간을 알 수 있다. 이처럼, 서로 약속된 신호를 송수신함으로써 송수신기간의 시간동기를 일치 시킬 수 있는데, 이때 사용되는 신호를 파일럿 신호(Pilot Signal)라 한다. 〈그림 4〉는 파일럿 신호 0011101을 사용한 예이다. 파일럿 신호는 시간동기를 맞추는 작업 외에도, 무선 전송 채널의 상태를 추정하는 등의 다양한 용도로 사용된다. 중화권의 지상 방송 표준인 DTMB(Digital Terrestrial Multimedia Broadcast)는 주기가 127 인 m-수열을 파일럿 신호로 채택하고 있다.

코드 분할 다중 접속(CDMA)에서는 상관도가 낮은 다수의 의사랜덤 수열을 각 사용자에게 할당하여 동시에 송수신을 가능하게 한다.

코드 분할 다중 접속(Code division multiple access, CDMA)는 의사랜덤 수열을 이용하는 통신 시스템의 또 다른 좋은 예이다[3][4]. 상호간의 상관도가 낮은 다수의 의사랜덤 수열을 각 사용자에게 할당하고, 이를 이용하여 동시에 다수 가입자의 송수신을 가능케 하는 이 기법은 다양한 통신 시스템에서 사용되고 있다.

일상생활에서 위치기반 통신을 위한 차량용 네비게이션은 GPS(Global positioning system) 신호를 사용하는데 바로 이 신호는 긴 길이와 짧은 길이의 m-수열을 사용한다. 또한, 군통신에

0000000	1111111
1110100	0001011
1101001	0010110
1010011	0101100
0100111	1011000
1001110	0110001
0011101	1100010
0111010	1000101

그림 5. [7,4,3] 해밍부호

서는 주파수도약 통신시스템이 주로 사용되는데, 이때 주파수 도약 순서를 결정하는 비2진 수열생성에도 LFSR이 유용하게 사용된다.

3. 오류정정부호와의 관계

마지막으로, 의사랜덤 수열은 디지털 통신 시스템의 오류를 제어하기 위한 기술인 "오류정정부호"와도 밀접한 관련이 있다. 예를 들어, 이전 절에서 설명한 통신 시스템 예에서 파일럿 신호로 사용된 0011101은 의사랜덤 수열임과 동시에, 가장 기본적인 오류정정부호로 알려져 있는 길이 7의 해밍부호(Hamming code)의 부호어이기도 하다.

구체적으로 길이 7인 해밍부호는 0011101와 이것의 모든 순환부호, 1111111과의 합인 1100010과 이것의 모든 순환부호로 이루어져있다. 〈그림 5〉는 16개의 [7,4,3] 해밍부호어 전체를 표시한다.

일반적으로 m-수열은 항상 같은 길이의 해밍부호의 부호어(Codeword)이다. 더 나아가, 모든 주기가 있는 2진 수열은 오류정정부호의 한 종류인 순환부호(Cyclic Code)의 부호어가 된다. 특히, m-수열이 부호어로 사용되면 자기상관값이 우수하다는 특성으로 인해 부호어들의 최소거리를 최대화시키는데 유용하다. 더 자세한 내용은 [5]를 참고하기 바란다.

모든 주기가 있는 2진 수열은 오류정정부호의 한 종류인 순환부호의 부호어가 된다.

Ⅲ. 선형 궤환 시프트 레지스터 수열

Ⅱ장에서 알아본 바와 같이, 의사랜덤 수열은 여러 정보통신 시스템에 성공적으로 적용되어 다양한 역할을 수행하고 있다. 본 장에서는 여러가지 많은 방법 중에서 선형 궤환 시프트 레지스터 수열에 대해 알아보고자 한다.

매 인가된 클럭(Clock)마다 기존의 값을 출력하고 새로운 값을 입력 받는 다수의 메모리를 선입선처리(FIFO, First Input First Output) 형태로 구성한 디지털 논리회로의 기본적인 요소를 시프트 레지스터라 한다. 〈그림 6〉과 같이 시프트 레지스터의 각 메모리가 가지는 값으로 선형 궤환함수의 값을 계산하고, 이를 시프트 레지스터의 입력에 되먹임(궤환) 함수로써 선형 궤환 시

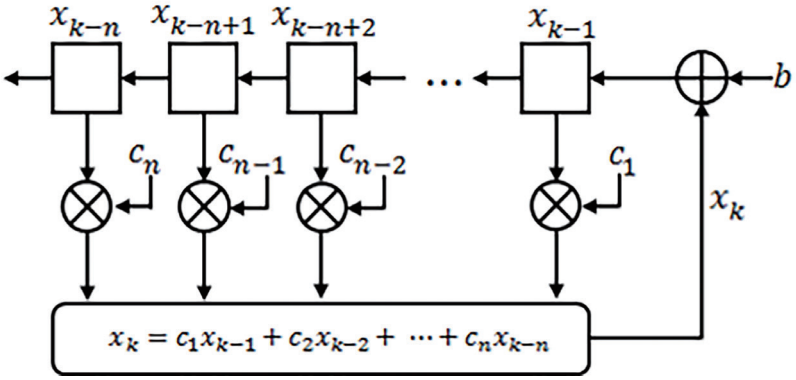

그림 6. 선형 궤환 시프트 레지스터의 구조

선형 궤환 시프트 레지스터에서 궤환 값을 계산하기 위해서 n개의 메모리가 배열되어 사용되면 단수가 n이라고 한다.

프트 레지스터를 구성한다.

〈그림 6〉에 나타난 선형 궤환 시프트 레지스터는 궤환 값을 계산하기 위해서 연속적으로 배열된 n개의 메모리가 사용되기에, 보통 n단 선형 궤환 시프트 레지스터라 불린다. 〈그림 6〉에서 왼쪽 끝에 위치한 메모리의 출력을 선형 궤환 시프트 레지스터의 출력이라 하고, 오른쪽 끝에 있는 b를 선형 궤환 시프트 레지스터의 입력이다 한다. 이때 선형 궤환 시프트 레지스터는 항상 하나의 출력이 존재하나, 입력은 존재하지 않을 수 있다. 입력이 존재하지 않으면 동차(Homogeneous)라고 부르며, 의사랜덤 수열을 생성하는 방법으로는 동차 LFSR이 주로 사용되므로, 앞으로의 전개에서 등장하는 선형 궤환 시프트 레지스터는 동차로 제한한다.

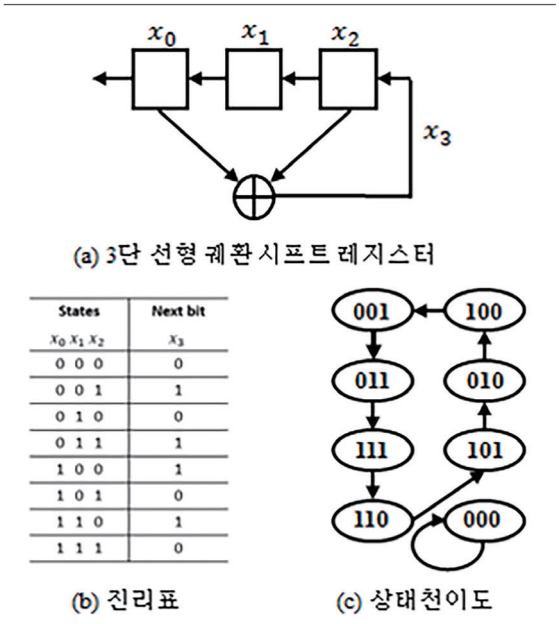

그림 7. 진리표와 상태천이도를 이용한 3단 LFSR의 해석

〈그림 7(a)〉에 나타난 3단 선형 궤환 시프트 레지스터를 이용하여 선형 궤환 시프트 레지스터의 동작에 대해 자세히 알아보자. 선형 궤환 시프트 레지스터가 동작을 시작한 이래로 k 번째 클럭이 인가되었을 때, 〈그림 7(a)〉의 3단 선형 궤환 시프트 레지스터에서 각각의 메모리가 보유하는 값들을 $(x_{k-3}, x_{k-2}, x_{k-1})$와 같이 표현할 수 있다. 이를 선형 궤환 시프트 레지스터의 상태라 한다. 자명하게도, 〈그림 7(a)〉의 2진 선형 궤환 시프트 레지스터는 $2^3 = 8$ 개의 서로 다른 상태를 가질 수 있다. 〈그림 6〉과 비교하면, 그림 7 (a)에서 $c_1 = 1$, $c_2 = 0$, $c_3 = 1$이기 때문에, k 번째 클럭에서 선형 궤환 시프트 레지스터는

$$x_k = x_{k-1} + x_{k-3} \tag{1}$$

의 궤환함수 연산을 통하여 궤환 값을 계산하고, 이 궤환 값을 이용해서 선형 궤환 시프트 레지스터가 (x_{k-2}, x_{k-1}, x_k)의 상태로 바뀐다. 자명하게도, k가 0일 때, 즉 동작 후 한 번도 클럭이 인가되지 않았을 때에도, 각각의 메모리들은 일정한 값을 가져야 한다. 이 값을 초기값(Seed)라 한다.

따라서, 단 하나의 동차 선형 궤환 시프트 레지스터의 출력수열은 다음의 값들로부터 결정된다.

① 선형 궤환함수

② 초기값

식 (1)의 궤환함수로부터 출력수열이 다음을 만족함을 알 수 있다.

$$x_k + x_{k-1} + x_{k-3} = 0, \quad \text{for all } k = 3,4,\ldots$$

위의 수열이 만족하는 관계를 선형점화식이라고 부르며, 이 수열의 일반항을 표현하는 핵심은 이 점화식의 특성다항식(Characteristic polynomial)이다.

$$x^3 + x^2 + 1 \tag{2}$$

이를 또한 최소차 연결다항식(Minimum Degree Connection polynomial)이라고도 한다.

선형 궤환 시프트 레지스터 역시 디지털 논리 회로이기에, 진리표와 상태천이도를 이용하여 그 동작을 예측할 수 있다. 〈그림 7〉은 예시로 주어진 3단 선형 궤환 시프트 레지스터의 진리표와

상태천이도를 이용한 해석을 보여주고 있다. 〈그림 7(c)〉를 통해서 알 수 있는 선형 궤환 시프트 레지스터의 특성이 있다.

① 모든 상태들은 오직 하나의 이전 상태와 이후 상태를 갖는다.

② (000)상태는 항상 스스로를 이전/이후 상태로 갖는다.

이를 통해서, 선형 궤환 시프트 레지스터 수열은 주기적 수열이고, 더 나아가 그 주기는 $(0, 0, \ldots, 0)$을 제외한 모든 상태의 개수보다 작거나 같다는 사실을 유추할 수 있다. 〈그림 7(a)〉의 3단 선형 궤환 시프트 레지스터의 경우, 상태의 변화는 일정한 주기를 갖고 있으며, (000)이 아닌 모든 초기값에 대해 변화 주기 내에서 (000)을 제외한 모든 상태가 정확히 한 번씩 나타난다. 선형 궤환 시프트 레지스터의 출력은 왼쪽에 있는 상태 값과 같기에, 결과적으로 〈그림 7(a)〉의 선형 궤환 시프트 레지스터는 주기가 7인 수열을 생성한다. 이는 3단 선형 궤환 시프트 레지스터 수열이 가질 수 있는 최대 주기이다. 일반적으로 단수가 n일 때, 2진 선형 궤환 시프트 레지스터 수열이 가질 수 있는 최대주기는 $2^n - 1$이 되며, 이러한 최대 주기를 가지는 출력 수열을 m-수열이라 정의한다.

단수가 n일 때, 2진 선형 궤환 시프트 레지스터 수열이 가질 수 있는 최대주기인 2^n-1을 가지는 출력 수열을 m-수열이라고 한다.

Ⅳ. m-수열의 몇가지 랜덤 특성

선형 궤환 시프트 레지스터 수열 중에서, 주어진 단수에 대해 생성 가능한 최대 주기의 수열을 일컬어 m-수열이라고 정의했다. 주기 $L = 2^n - 1$인 m-수열은 최대 주기를 갖는다는 특성 외에도 몇 가지 상당히 중요한 랜덤 특성을 갖는다. 이러한 연유로 m-수열은 의사랜덤수열이며 다양한 분야에서 널리 사용되고 있다. 본 장에서는 주기가 L인 2진 m-수열이 갖는 몇가지 랜덤 특성에 대해 이야기 하고자 한다. m-수열의 특성에 대한 보다 자세한 정보는 [3]을 참고하기 바란다.

1. 균형 특성(Balanced property)

균형 특성은 한 주기 내부에 0과 1의 개수가 최대한 균형을 맞춘다는 특성이다. 주기가 홀수이므로 그 차이가 1일 때 최대한 균형이 맞는다. 주기가 $L = 2^n - 1$인 2진 m-수열에서 0은 $(L-1)/2$번 나타나며, 1은 $(L+1)/2$번 나타난다. 이를 m-수열의 균형특성이라 부른다.

2. 연속마디 분포 특성(Run property)

r-연속마디(Run)란, 하나의 문자가 연속적으로 r번 나타나고, 그 양쪽 끝에 다른 문자가 위치하는 것을 의미한다. 예를 들어

…01110…

은 1의 3-연속마디이다. m-수열의 연속마디 분포 특성은 짧은 길이의 연속마디가 보다 긴 길이의 연속마디보다 자주 나타나고, 1의 연속마디 개수와 0의 연속마디 개수가 같음을 의미한다. m-수열에서 각 연속마디들이 나타날 확률은, 2진 랜덤 수열에서 같은 연속마디들이 나타날 확률과 매우 유사하다.

3. 이상적 자기상관 특성

주기가 L인 임의의 2진 수열 $s(k)$에 대해, 정규화된 주기적 자기상관 함수 $R(\tau)$를

$$R(\tau) = \frac{1}{L}\sum_{k=0}^{L-1}(-1)^{s(k)-s(k+\tau)} \qquad (3)$$

와 같이 정의하자. 여기서, τ는 임의의 정수이다. 위의 수식 계산에서, $k+\tau$는 모듈로-L로 연산한다. 주기 $L=2^n-1$을 갖는 m-수열의 자기상관 함수는

$$R(\tau)=\begin{cases} 1\ , & \tau\equiv 0\ (\mathrm{mod}\ L) \\ -1/L, & otherwise \end{cases} \tag{4}$$

이다. 이는 주기가 홀수인 2진 수열이 가질 수 있는 최적의 자기상관이다. 참고로, 길이 L인 2진 랜덤 수열이 갖는 정규화된 주기적 자기 상관의 평균값은 $\tau \neq 0\ (\mathrm{mod}\, L)$일 때 항상 0이다. m-수열은 주기가 $L=2^n-1$인 모든 2진 수열 중에서 자기상관 특성이 2진 랜덤 수열과 가장 유사한 수열이다.

이 밖에도, m-수열은 스팬 특성(Span property)과 순환-덧셈 특성(Cycle-and-add property) 같은 다양한 랜덤 특성을 갖는다[3].

Acknowledgement

이 논문은 2015년도 정부(교육부)의 재원으로 한국연구재단의 지원을 받아 수행된 기초연구사업임(No.2013R1A1A2062061).

참 고 문 헌

[1] S. W. Golomb, Digital Communications with Space Applications, Prentice-Hall, 1964.

[2] M. K. Simon, J. K. Omura, R. A. Scholtz, and B. K. Levitt, Spread Spectrum Communications Handbook, McGraw-Hill, 1994.

[3] S. W. Golomb, Shift Register Sequences, Aegean Park Press, 1982.

[4] A. J. Menezes, P. C. van Oorschot, and S. A. Vanstone, Handbook of Applied Cryptography, CRC Press, 1996.

[5] 양경철, "오류제어기술," 정보와 통신 열린강좌, 33권, 6호, pp. 18-26, 2016년 6월.