스트림 암호 시스템의 설계 및 비도 분석에 관한 연구

°박흥근, 남길현
국 방 대 학 원

A study on Designe and Analysis of Stream Cipher System

Heung-Keun Park, Kil-Hyun Nam
National Defense University

요 약

스트림 암호 시스템의 가장 핵심이 되는 것은 랜덤 수열 발생기이다. 본 연구는 기존의 랜덤 수열 발생기를 연구, 분석하여 선형 복잡도와 랜덤성을 향상시킨 랜덤 수열 발생기를 설계, 구현하여 효율적인 스트림 암호 시스템으로 사용할 수 있는 방안을 제시하였다.

1. 서 론

현대 사회에 있어서 정보란 눈에 보이지는 않지만 유형 자산 못지 않은 중요한 재산으로 여겨지고 있고 이러한 재산에 대한 훼손, 교란, 위협, 침해 등의 행위는 마땅히 제재되어야 한다. 이러한 정보 보호 대책으로서는 정보를 보관하고 있는 시스템 설비면에서의 물리적 안전 대책, 운영 인원에 대한 인적 자원 보안 대책, 법과 제도적 측면에서의 대책, 기술적 측면에서의 대책 등이 있는데 법과 제도적 측면에서의 대책은 기술의 발전 속도에 비해 뒷받침하는 속도가 미약하고 설비면에서의 물리적 대책이나 운용 인원에 대한 대책은 비용 측면이나 효율적인 면에서 비경제적이다. 따라서 기술적 측면으로 보관 자료나 전송 자료를 직접 암호화하여 보호하는 것이 다른 대책보다 효율적이라 판단된다.

본 논문은 정보를 이러한 여러 위협으로부터 보호하기 위한 기술적 측면의 보호 방법인 암호 시스템, 그 중에서 디지탈 암호화 시스템에 관한 연구이다.

스트림 암호 시스템은 1970년대 초반부터 주로 유럽에서 연구, 발전된 것으로써, 선형 쉬프트 레지스터를 이용한 이진 수열 발생기를 사용하는 암호 시스템이다. 이것은 LFSR를 비선형 결합하는 방식으로 구성되어 있다[원동95]. 또한블럭 암호 시스템과는 달리 비교적 수학적 분석이 가능하여 여러가지 중요한 수치 즉, 주기나 선형 복잡도 등에 대해 이론적인 값을 정확히 계산할 수 있고 데이타에 대한 오류 전파 현상을 최소화 할 수 있으며 알고리즘 실현이 비교적 용이하다는 장점 때문에 지금까지 널리 사용되고 있고, 고속을 요구하는 어플리케이션에 적합하다[Lee94]. 따라서 앞으로도 많은 연구와 발전이 있을 것으로 여겨진다.

본 논문에서는 기존의 스트림 암호시스템에 대한 소개와 성능 개선을 목적으로 스트림 암호 시스템의 분석 요건, 설계된 시스템에 대한 평가 항목 을 제시하고 이에 따라 기존 스트림 암호 시스템을 분석하고 이보다 성능이 개선된 스트림 암호 시스템을 제안하고 구현 및 분석하였다.

2. 스트림 암호시스템의 비도 분석

2.1 비도분석 요건

암호알고리즘의 안전성은 암호분석가가 암호알고리즘을 해독할 때의 어려움을 의미하는 것이다.

여기에서 암호알고리즘은 평문을 암호문으로 변환하는 함수들의 집합이며 주어진 함수들의 집합에서 특정한 함수를 선택하는 역할이 키이다. 그런데 주어진 함수를 키에 대한 정보 없이 평문으로 역변환 하는 과정을 알고리즘 해독이라 한다.

암호시스템이 얼마나 안전한가를 측정하는 암호 강도에는 Shannon이 정의한 무조건 안전 (unconditional secure) 과 계산적 안전 (computational secure) 두 가지가 있다. 무조건 안전이란 암호 해독자가 이용할 수 있는 연산 능력은 무한하다고 가정하고, 단지 암호 해독에 필요한 정보의 양이 불충분하여 암호 해독이 불가능한 경우를 의미한다. 계산적 안전이란 암호해독자가 이용할 수 있는 정보의 양이 충분하여 언젠가는 암호를 해독할 수는 있으나 그 해독 과정이 복잡하고 시간과 경비가 많이 요구되어 경제적으로 불합리한 경우의 암호강도를 의미하며 현재 이용되고 있는 암호 알고리즘 대부분이 여기에 속한다. 스트림 암호 시스템의 안전성 측정은 Shannon이 제시한 기준을 사용하는데 Shannon은 1949년 "Communication Theory of Security System" 이란 논문에서 암호 설계자에게는 불리하고 해독자에게는 유리한 조건을 제시하였는데 이것은 다음과 같다.

가정1. 암호 해독자는 키를 제외한 이진 수열 발생기를 포함한 암호 시스템의 모든 사항을 알고 있다.

가정2. 암호 해독자는 충분한 양의 암호문과 이에 해당하는 평문도 획득할 수 있다.

이 두 가지 가정에서 스트림 암호 시스템의 안전성 평가 지표들을 도출할 수 있는데 먼저 가정 1에서 키의 길이가 크지 않다면 암호 해독자는 키 소모적 탐색(Exhaustive Search) 방법을 사용하여 스트림 암호 시스템을 공격할 수 있다. 그러므로 키의 길이는 되도록 큰 값을 가져야 한다. 가정 2에서 얻은 충분한 양의 암호문에 통계적 성질이 나타나면 암호 해독자는 암호문의 통계적 성질과 평문의 통계적인 성질을 비교하여 키를 찾으려고 할 것이다. 그러므로 암호문의 통계적 성질을 배제하기 위해 다음과 같은 조건들을 만족해야 한다.

첫째, 주기가 길어야 한다.

둘째, 랜덤성(randomness)을 가져야 한다.

셋째, 선형 복잡도(linear complexity)가 커야 한다.

넷째, 비선형 결합 함수의 상관 면역성(correlation immune)이 커야 한다.

따라서 스트림 암호 시스템의 비도를 결정하는 중요한 변수들은 키 수열의 주기, 키 수열의 불규칙성, 선형 복잡도, 키 관리 문제 등이며 이러한 변수들이 충족되어야 스트림 암호 시스템의 안전성이 있다고 볼 수 있는 것이다.

2.2 스트림 암호 시스템의 비도 평가 항목

2.2.1 주기

어떤 수열 s(t)의 주기는 0보다 큰 모든 t에 대해 s(t + P) = s (t) 인 가장 작은 양의 정수 P를 말한다. 그런데 출력 수열의 한 주기 P를 구할 수 있으면 그 주기에 대한 자기 상관 특성을 이용하여 키 스트림의 키 값을 구할 수 있으므로 주기는 매우 길어야 한다.

2.2.2 랜덤특성

엄밀한 의미에서 주기를 갖는 수열은 랜덤성(*randomness*)을 갖지 않는다. 랜덤성과 주기는 상충되는 의미를 갖는다. 그러나 암호 설계에 있어서 필요한 것은 엄밀한 의미의 랜덤성이 아니라 어느 정도의 불예측성(unpredictability)이다. 즉, 암호 해독자가 어느 정도의 수열을 획득한다고 하여도 획득된 수열이 주기에 비해 아주 짧아서 해독자가 다음 수열을 예측할 수 없다면 아무런 정보도 누출되지 않았다고 본다. Golomb은 주기가 P 인 이진 수열에 대해 다음의 조건을 만족하면 랜덤성이 있다고 제안했다.

첫째, 균형 성질(balance property)이다. 이것은 수열의 한 주기에서 나타나는 0의 갯수와 1의 갯수의 차이가 1이하이어야 한다는 것을 의미한다. 즉, 0과 1이 거의 같은 비율로 나타나야 한다.

둘째, Run 성질이다. 수열에서 '0'이 연속되거나 '1'이 연속된 것을 Run이라 하는데, 예를 들어 011100 1은 하나의 '0'으로 시작되고 세 개의 '1', 두 개의 '0', 하나의 '1'로 구성된 run이라 한다. 이 때 '0'으로 이루어진 run을 'gab' 이라 하고 '1'로 이루어진 run을 'block' 이라 한다. 주기가 P인 수열에서

길이가 1인 run이 P/2개, 길이가 2인 run이 P/4 개, 길이가 3인 run이 P/8 개,... 즉, 각 i에 대하여 최소한 2^{n-i} 인 run이 있고 길이 i인 run은 $P/2^i$ 개 있고, 각 길이에 대한 gab과 block도 같다. 이것은 01 다음에 '0' 이 나타날 확률과 '1' 이 나타날 확률이 같아야 함을 의미한다.

셋째, 자기 상관 특성이 좋아야 한다. 자기 상관 함수는 주기 P인 수열 (s_t)와 여기에서 d만큼 천이된 수열 (s_{t+d})와 비교하여 일치한 위치의 갯수 A와 틀린 갯수 D를 주기로 나눈 값 C(d) = (A - D) / P 이다. 수열의 자기 상관 함수는 편이가 주기의 배수일 때를 제외하고는 매우 작아야 한다. 즉, 어떤 임의의 수열과 그 수열을 일정하게 쉬프트한 수열 간의 일치하는 것(agreement)의 수를 세는 것은 그 수열의 주기만큼 쉬프트 하지 않는 한 그 수열의 주기에 대한 어떠한 정보도 알아낼 수 없음을 의미한다.

2.2.3 선형 복잡도(linear complexity)

선형 복잡도는 스트림 암호시스템에 있어서 주기와 함께 가장 중요시 되는 보안성 요인이다. LFSR의 출력 수열은 매우 이상적인 통계적 특성을 가지고 있지만, 그 자체가 가지고 있는 선형성으로 인하여 대수적인 기법으로 암호 시스템을 공격할 수가 있기 때문에 스트림 암호화를 위한 수열 생성기로 직접 사용할 수는 없다. 키 스트림 생성기가 이런 종류의 공격에 견딜 수 있는 기준을 제공하는 요소가 바로 선형 복잡도이다. 따라서 키 스트림의 선형 복잡도를 더 증가시키기 위하여 같은 LFSR로부터 여러개의 출력을 조합한다든지 여러개의 LFSR로부터 출력 수열을 조합하는 비선형 결합함수를 사용한다[Mass88].

선형 복잡도는 키 스트림을 생성할 수 있는 가장 짧은 LFSR의 길이를 나타내는 것으로서 생성기가 비선형 함수에 의해 출력을 발생할 경우, 그것에 등가인 선형 생성기를 구하여 선형 복잡도를 나타낸다[Rue92]. 이것은 또한 생성된 출력 수열의 랜덤성(randomness), 불예측성(unpredictability), 등가성(equivalence)을 측정하는 지표를 제공한다.

2.2.4 상관 면역 (correlation immunity)

대다수의 키 스트림 생성기는 LFSR을 그 기본 요소로 갖고 있다. 이러한 LFSR 구조의 키 스트림 생성기는 결국 유한 상태 기계이므로 자체적으로 일정한 주기를 갖는다. 이를 숨기기 위하여 비선형 알고리즘을 도입하여 사용한다. 그러나 일부 비선형 알고리즘의 경우 내부적인 LFSR에 의한 이진 수열과 출력 수열들 사이의 상관성으로 인해 그 취약함을 가지고 있다. 상관 공격(correlation attack)의 가능성은 지겐델러(T. Sigenthaler)가 그 이론을 정립하였다. 지겐델러는 비선형 결합의 경우에도 그 수열을 분석할 수 있는 상관 공격의 개념을 도입하고 Geffe의 생성기, Press의 생성기 등을 예로 들어 이론적 분석 결과와 컴퓨터 실행 분석 결과를 비교함으로써 상관 공격의 타당성을 제시하였다. 키 스트림이 되는 외부 출력 이진 수열은 내부의 LFSR들에 의한 수열과의 비선형 결합이어서 LFSR의 복잡성을 증가시킬수록 내부 출력 수열과 외부 출력 수열간의 종속성으로 인하여 상관 공격에 점점 더 취약해진다.

따라서 키 스트림 발생기의 그 생성기에 의한 출력들이 입력들과 상관 관계가 가급적 작도록 구성하여야 한다. 보통 m 개의 어떤 입력과도 생성기에 의한 출력 사이에는 통계적인 종속성 (statistical dependency)이 존재하지 않을 때 이 이진 수열은 " m 차 상관 면역(correlation immunity)이 있다"고 한다. 지겐델러의 상관 공격에 관한 연구 발표 이후로 암호학에 있어서 상관성에 대한 활발한 연구가 있어왔다[Gie84].

지겐델러는 또한 키 스트림의 분리 정복(divide and conquer)에 의한 상관 공격에 대응할 수 있는 방법을 제안하기도 했다. 상관 면역성을 높여 보다 안전한 암호 시스템을 설계하려 할 때는 비선형 생성기의 복잡성을 감소시킬 수 밖에 없다는 결점이 있어서 일종의 절충 관계(trade off)가 있다. 즉 선형 복잡도를 증가 시키면 상관 면역성이 낮아지고 상관 면역성을 높이면 선형 복잡도가 감소하게 된다. 이것은 키 스트림 발생기에 있어서 구조적인 문제점으로 대두되고 있다. 이와 같은 절충 관계를 없애기 위해 뤼펠(R. A. Rueppel)은 메모리가 있는 키 스트림 생성기를 제안하기도 하였다.

2.2.5 효율성

키 스트림을 생성하기 위해서 목적에 따라 하나의 LFSR 혹은 여러개의 LFSR을 비선형적인 방법으로 결합하여 사용할 수 있다. 임의의 암호시스템에서 LFSR을 몇 개 사용하든지 실질적인 이용 측면에서 볼 때 암호 시스템이 아무리 안전하고 비도가 높다고 해도, 처리 속도가 적절하지 않으면 그 가치가 떨어지게 된다. 따라서 n비트를 생성해 내는데 걸리는 시간을 측정하여 설계된 스트림 암호시스템이 어느 정도의 효율성을 가지고 있는지 비교 분석하여 봄으로써 성능이 개선되었는지의 여부를 알 수 있다.

2.3. 스트림 암호 시스템의 랜덤성 분석 방법

앞에서 키 스트림 생성기의 비도 요인 평가 방법으로는 주기, 랜덤특성, 선형 복잡도, 상관 면역, 효율성 등으로 열거 했으나, 이 중에서 주기와 선형 복잡도는 특성 다항식에 의해 결정되므로 기존 스트림 암호 시스템의 평가 항목으로 랜덤특성과 상관 면역, 효율성 측면에서 분석하고, 특히 이 가운데서 랜덤성 검증에 비중을 두고 분석하고자 한다.

일반적으로 랜덤 수열은 결정적 방법에 의해 만들어지기 때문에 의사 랜덤 수열로 부르며 이것이 랜덤하다는 것을 보장하기 위한 검증 방법이 필요한데 키 스트림 생성기에서 생성되는 수열의 랜덤성을 검증하는 이론적 검증은 한 주기동안 발생하는 모든 수열에 대하여 랜덤성을 검증한다. 즉 0의 갯수와 1의 갯수가 비슷한가, 자기 상관 함수가 적은 값인가등을 검증한다. 수학적으로 완전한 랜덤 수열(Truly random sequence)이란 확률변수열 { X_n }이 독립이고 같은 분포를 가질 때 $X_n = x_n$인 실현치의 수열 { x_n }을 의미한다. 특히,P($X_n = 0$)= P($X_n = 1$)= 1/2이고 i.i.d(*independent and identically distributed*) 인 확률변수열(Xn)을 완전한 이진 랜덤 수열이라 한다.

통계적 검증의 원리로부터 한 수열이 어떤 통계적 검증을 통과하였다 하는 것은 랜덤하지 않다고는 말할 수 없는 소극적 긍정이지 그 수열이 랜덤하다고 단정하는 적극적인 긍정은 아니다. 그것은 한 수열이 통계적 검정 T_1, T_2, ..., T_4 을 통과하였다고 해서 다른 통계적 검정 T_{n+1}을 통과한다는 보장은 없으며 T_{n+1}까지 통과할 경우에 그 수열의 랜덤성에 더 많은 신뢰성을 부여할 수 있다는 것을 의미한다.

확률 및 통계 이론으로부터 랜덤성에 대한 많은 검정 방법이 알려져 있으나 본 절에서는 기존의 스트림 암호 시스템에서 생성된 이진 수열의 랜덤성 검증을 위해 frequency test, serial test, poker test, autocorrelation test, run test 등을 사용하였다.

2.3.1 Frequency Test

키 스트림을 테스트하는 방법은, 먼저 키 스트림으로부터 충분한 양의 이진 수열 (S_t)를 얻어 0의 갯수를 n_0, 1의 갯수를 n_1이라 한다. 이상적인 랜덤 수열이라면 어떤 한 비트가 0이 될 확률은 1/2 이므로 키 스트림 내에서의 0의 기대치는 n/2이고 1의 기대치도 n/2이다. 그러므로 0의 측정치 n_0와 1의 측정치 n_1으로부터 다음과 같은 검정 통계량을 갖는 χ^2 분포를 생각할 수 있다.

$$\chi^2 = \frac{(n_0 - n_1)^2}{n}$$

χ^2 분포로부터 유의수준 5%의 값은 3.84이므로 통계량 χ^2 의 값이 3.84보다 큰 키 스트림은 frequency test 에 대해 랜덤성이 없다고 판단되어 기각한다.

2.3.2 Serial Test

이것은 이진 수열 (S_t)에서 한 비트가 다음 비트로 가는 천이 확률을 테스트하는 것으로서 한개의 비트 즉, '0' 비트나 '1' 비트가 주어졌을 때 그 다음 비트가 '0' 인 경우와 '1' 인 경우를 테스트한다. 테스트하는 방법은 우선 이진 수열 (S_t)에서 테스트할 데이타를 얻은 다음, 이 수열을 연속된 2비트로 분리하여, '00' 의 갯수를 n_{00}, '01' 의 갯수를 n_{01}, '10' 의 갯수를 n_{10}, '11' 의 갯수를 n_{11} 이라 하고 전체 N 비트 중 '0' 의 갯수를 n_0, '1' 의 갯수를 n_1이라 하자. 그러면 다음과 같은 식을 얻는다.

$$n_{00} + n_{01} = n_0 \quad \text{혹은} \quad n_0 - 1$$

$n_{10} + n_{11} = n_1$ 혹은 $n_1 - 1$

$n_{00} + n_{01} + n_{10} + n_{11} = N-1$

$n_0 + n_1 = N$

n_{ij}의 기대치는 (N-1)/4 이므로 다음의 통계량은 근사적으로 자유도가 2인 χ^2 분포를 따른다.

$$T = \sum_{i,j=0}^{1} \frac{n_{ij} - (N-1)/4}{(N-1)/4} - \sum_{i=0}^{1} \frac{(n_i - N/2)^2}{N/2}$$

이것을 다른 식으로 표현하면 다음과 같다.

$$T = \frac{4}{N-1}(n_{00}^2 + n_{01}^2 + n_{10}^2 + n_{11}^2) - \frac{2}{N}(n_0^2 + n_1^2) + 1$$

위의 T는 자유도 2인 χ^2 분포를 뜻하며, χ^2 분포에서의 유의수준 5%, 자유도 2의 값은 5.99이므로 검정 통계량의 값이 5.99보다 큰 키 스트림은 serial test에 대해 랜덤성을 갖지 못하므로 기각한다.

2.3.3 Poker Test

8 비트 ASCII 코드의 종류는 2^8=256가지인데, 길이 n인 키 스트림을 취하면, F= $\frac{n}{8}$ 개의 8 비트 ASCII 문자가 생긴다. (엄밀히 말하면 F는 $\frac{n}{8}$ 을 넘지 않는 가장 큰 정수이다.) F개의 8 비트 이진 코드 중에서 *00000000*의 갯수를 f_0, *00000001*의 갯수를 f_1, ... , *11111111*의 갯수를 f_{255} 라 하자. 만일 키 스트림이 균일하게 산포되어 있다면, 256 가지의 8 비트 ASCII 코드 문자는 각각 동일한 1/256 이 되어 그에 따른 기대치는 F/256 이 된다.

이러한 f_i 에 대해 다음의 검정 통계량을 갖는 χ^2 을 생각할 수 있다.

$$T = \frac{256}{F}\sum_{i=1}^{255}(f_i - F)^2$$

위의 T는 자유도 255인 χ^2 분포를 하며, χ^2 분포에서 유의수준 5%, 자유도 255의 값은 292.84이므로 이 값보다 큰 키 스트림은 poker test에 대해 랜덤성을 갖지 못하므로 기각한다.

2.3.4 Autocorrelation Test

이 테스트는 이진 수열 (S_t)와 이로부터 양의 정수 n만큼 천이시킨 수열 (S_{t+n})와의 상관관계를 조사하는 것이다.임의의 n개의 비트를 $a_1a_2a_3...a_n$ 이라 하면, 각 autocorrelation 값은 다음과 같이 주어진다.

$$A(d) = \sum_{i=1}^{n-d} a_i a_{i+d} \qquad (0 \le d < n-1)$$

여기에서

$$A(0) = \sum_{i=1}^{n-d} (a_i)^2 = n_1$$

즉, 1의 갯수가 됨을 알 수 있다.

만일 키 스트림으로부터 n 개의 수열이 n_0개의 '0' 을 갖고 1 개의 '1' 을 가지고 있으며 균일하게 분포되어 있다면, 수열의 autocorrelation의 이상적인 값은 다음과 같다.

$$u(d) \simeq \frac{n_1^2 (n-d)}{n^2} \qquad (d = 0)$$

이 때, 측정치 A(d)와 기대치 u(d) 사이의 차이를 x(d)= A(d) - u(d), ($1 \le d < n-1$)라 하고, x(1), x(2), ..., x(n-1)은 정규분포를 이룬다고 하며, 전체 키 스트림의 x(1), x(2), ..., x(n-1)의 모평균 u는 0이라고 하자.

각 x(d) 값들의 표준 평균을 X, 표본 분산을 S라 하면, t검정에서의 자유도 n-2인 검정 통계량은 다음과 같다.

$$t = \frac{X-u}{S/\sqrt{n-1}}$$

그러나 자유도가 큰 t분포 (>>30)는 표준정규분포를 이루게 된다. 이를 이용하면 유의 수준 5%로 하는 한계치는 1.96 이므로 이 값보다 크면 autocorrelation이 랜덤성을 가지지 못하는 것으로 판단하여 기각한다.

2.3.5 Run Test

이것은 이진 수열(S_t)내에 0이나 1이 연속해서 나오는 것을 말하며, 알고리즘 자체의 어떤 취약성에 의해 키 스트림이 이러한 특성을 갖게 되는지를 테스트하는데 필요하다.

이것을 테스트하는 방법은 다음과 같다.

출력 수열을 '0' 의 나열인 gab과 '1' 의 나열인 block으로 나눈다. r_{0i}는 길이 i인 gab의 갯수이고 r_{1i}는 길이 i인 block의 갯수라면, gab의 전체의 갯수와 block의 전체의 갯수는 다음과 같이 주어진다.

$$r_0 = \sum_{i=1}^{17} r_{0i} \qquad (r_0;\ \text{전체의 gab 갯수})$$

$$r_1 = \sum_{i=1}^{17} r_{1i} \qquad (r_1;\ \text{전체의 block 갯수})$$

이 때, 키 스트림의 수열이 충분히 크고 frequency test를 통과하였다면, '0' 의 run 수와 '1' 의 run 수는 거의 같으므로 두 개의 run을 구별할 필요는 없다.

$n_i = n_{0i} + n_{1i}$, $n = r_0 + r_1$이라 하면, 어떤 run이 길이 i가 될 이상적인 확률은 $1/2^i$이므로 기대치는 $n/2^i$가 된다. 그러므로 실제 측정치인 n_i와 추정치인 $n/2^i$ 사이의 χ^2 test를 생각할 수 있다.

$$\chi^2 = \frac{1}{n}\sum_{i=1}^{17} 2^i (n_i - n)^2$$

χ^2 test로부터 자유도 16인 χ^2 분포의 5% 유의 수준에 의한 값은 26.300 이므로 χ^2 의 값이 이것보다 크면 run test에 대하여 랜덤성을 갖지 못하므로 기각한다.

이상을 요약하여 < 표 1 > 에 분석 기준을 정리하였다.

< 표 1 > 각 Test 별 임계 값

Test 종류	Frequency Test	Serial Test	Poker Test	Autocor-relation Test	Run Test
임계 값	3.480	5.990	292.840	1.960	26.300

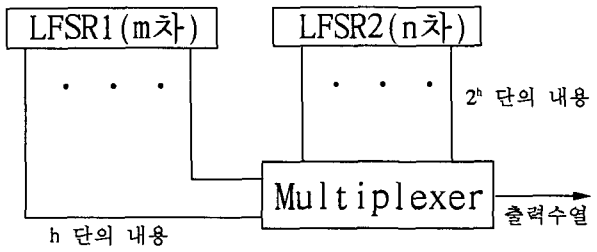
3. 기존 스트림 암호 시스템의 분석

본 절에서는 위에서 소개된 스트림 암호 시스템의 분석 방법을 이용하여 기존의 스트임 암호 시스템을 분석하여 보았다. 기존의 스트림 암호 시스템에는 J-K-flip-flops, Pless 시스템, 게프(Geffe system) 시스템, 상호대칭 시스템, 멀티플렉싱 시스템 (Multiplexing system), BRM 시스템 등이 있는데, 이 중 임의성이 양호하고 보안성도 우수하다고 알려진 멀티플렉싱 시스템, BRM 시스템에 대해서 분석해 보고, 프로그램으로 구현하여 스트림 암호 시스템의 분석 기준에 따른 적합성 여부를 조사하였다.

3.1 멀티플렉싱 시스템 (Multiplexing system)

멀티플렉싱 시스템은 간단히 MUX 시스템이라고 하며, 멀티플렉서 2개와 LFSR로 구성되는데 이것을 도식하면 < 그림 1 > 과 같다[한국91].

LFSR1과 LFSR2의 차수가 각각 m, n이고 LFSR1의 단이 A_0, A_1, ..., A_{m-1}, LFSR2의 단이 B_0, B_1, ..., B_{n-1}, 그리고 LFSR 1의 출력 수열을 (a_n), LFSR 2의 출력 수열을 (b_n)이라고 하자. 먼저 $1 < h < m$ 인 정수 h를 선택한 후, LFSR 2의 n 단 중에서 2^h 단을 선택한다. 시간이 t일 때 MUX 시스템의 출력 U_t는 LFSR 1의 h 단의 내용에 의해 LFSR 2의 2^h 단 중 한 단의 내용으로 결정된다.



< 그림 1 > 멀티플렉싱 시스템

MUX 시스템을 구현하기 위해 다음과 같이 각 차수가 서로소인 특성 다항식을 무작위 추출하여 사용하였다.

LFSR 1 : $X^{33} + X^5 + X + 1$

LFSR 2 : $X^{35} + X^4 + X^2 + 1$

이 시스템의 주기는 $(2^{33} -1)(2^{35} -1) = 2^{68}$이고, h=4인 경우를 택하였으므로 선형 복잡도는 $35(1 + {}_{33}C_1 + {}_{33}C_2 + {}_{33}C_3 + {}_{33}C_4) = 2^{21}$이다.

MUX 시스템의 랜덤성 테스트 분석 결과는 < 표 2 > 와 같다.

< 표 2 > 에서 알 수 있듯이 MUX 시스템은 Frequency Test는 모두 무사히 통과하였으나 Serial Test , Poker Test, Autocorrelation Test, Run Test 에서는 기각 영역에 속하는 부분들이 있어 시스템이 불안정 하다는 것을 알 수 있다.

< 표 2 > MUX 시스템 분석 결과

	Frequency Test	Serial Test	Poker Test	Autocor-relation Test	Run Test	Time (sec)
8000 bit	0.288	9.0496	348.096	2.441	25.999	0.327
16000 bit	2.704	5.3092	296.928	1.990	29.087	0.654
24000 bit	1.290	5.1667	280.454	2.302	23.564	0.983
32000 bit	0.072	5.3942	290.276	1.843	22.372	1.449

3.2 BRM 시스템

BRM 시스템은 2개의 LFSR과 BRM(Binary Rated Multiplexer)으로 구성되며 이것을 도식하면 < 그림 2 > 와 같다 [Jen84].

BRM 시스템을 구현하기 위해 다음과 같이 각 차수가 서로소인 특성 다항식을 무작위 추출하여 사용하

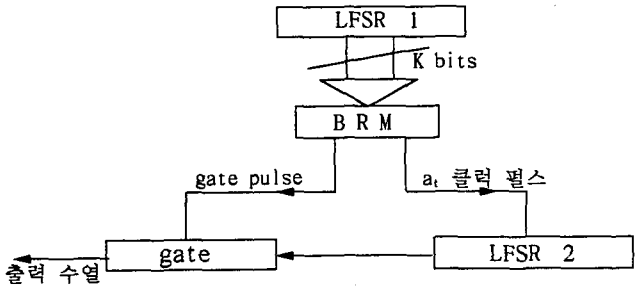
였다.

LFSR 1 : $X^{33} + X^5 + X + 1$

LFSR 2 : $X^{35} + X^4 + X^2 + 1$

이 시스템의 주기는 ($2^{33}-1$)($2^{35}-1$)=2^{68}이고 선형 복잡도는 35($2^{33}-1$)=2^{38}이다.

BRM 시스템의 랜덤성 테스트 분석 결과는 < 표 3 > 과 같다.



< 그림 2 > BRM 시스템

< 표 3 > 에서 보는 바와 같이 BRM 시스템은 랜덤성 테스트에서 32000bit가지는 모두 통과하였다. 그러나 비트수가 증가할수록 임계치를 초과하는 것을 알 수 있다.

< 표 3 > BRM 시스템 분석 결과

	Frequency Test	Serial Test	Poker Test	Autocor-relation Test	Run Test	Time (sec)
8000bit	0.024	1.2765	240.064	0.948	20.428	0.988
16000bit	0.009	1.2923	234.624	1.132	21.032	2.075
24000bit	0.160	1.5142	222.699	1.325	23.602	3.087
32000bit	1.058	1.4157	228.224	1.453	23.650	3.972
64000bit	2.324	2.350	256.984	1.785	24.897	7.546
120000bit	3.500	4.210	295.438	1.959	27.200	13.962

4. 개선된 암호 시스템의 제안 및 분석

안전한 키 스트림이 되기 위해서는 67차 이상의 높은 차수의 원시 다항식을 사용하여야 하는데 앞에서 분석한 기존의 스트림 암호 시스템은 랜덤 수열 발생시 33차 원시 다항식과 35차 원시 다항식을 사용해야 하기 때문에, 원시 다항식 판별 소요 시간이 과다하여 원시 다항식을 결정하는데 있어서 융통성이 결여되기 쉽고, 하드웨어 실현의 어려움이 있어 수열 생산의 효율이 떨어지게 된다. 또한 BRM 시스템의 경우 암호화 할 TEXT가 긴 경우 랜덤 테스트에 대한 임계치를 초과할 우려가 있다. 따라서 5차 이하의 낮은 차수의 특성 다항식을 사용하여 주기에는 큰 변화가 없으나 선형 복잡도를 증가시킬 수 있는 < 그림 3 > 과 같은 개선된 스트림 암호 시스템을 설계하였다.

각 시스템의 특성 다항식은 다음과 같다.

LFSR : $x^{35}+x^4+x^2+1$

BRM1 : x^3+x+1, x^5+x^2+1

BRM2 : x^4+x+1, x^4+x^2+1

BRM3 : x^5+x^3+x, x^3+x+1

BRM4 : $x^4+x^3+x^2$, x^4+x^3+x

각 특성 다항식은 서로소인 임의의 차수로 무작위 축출하였으며, 설계한 시스템의 작동 원리는 다음과 같다.

우선 BRM 시스템 4개를 하나의 LFSR로 구성하여 각 BRM 시스템의 출력을 하나의 단위 비트로 저장한다. 저장된 결과를 4 비트 이진 수열로 출력시켜 LFSR 의 임의의 16개 비트 (2^h)의 내용 중 BRM 출력 결과에 따른 위치의 비트를 추출하여(MUX 시스템의 출력) 키 스트림으로 발생 시킨다.

본 시스템의 주기는 각 시스템의 주기를 곱한 것과 같다.

BRM 1의 주기 : (2^3-1)(2^5-1)

BRM 2의 주기 : (2^4-1)^2

BRM 3의 주기 : (2^5-1)(2^3-1)

BRM 4의 주기 : (2^4-1)^2

LFSR의 주기 : ($x^{35}-1$)

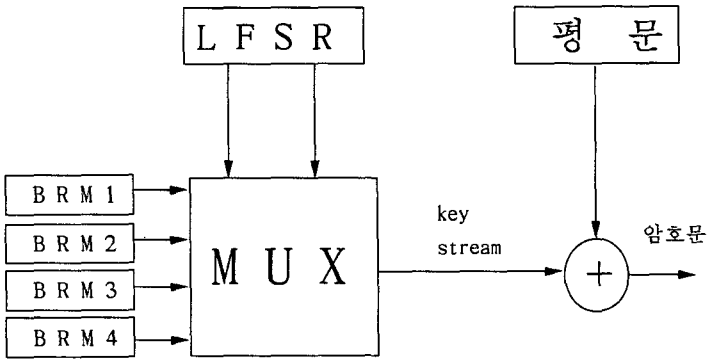
이것을 계산하면 제안한 시스템의 주기는 2^{67}이다.

BRM 1,BRM 2,BRM 3,BRM 4의 선형 복잡도는 (2^3-1)×5, (2^4-1)×4, (2^5-1)×5, (2^4-1)×4 이므로 이들의 곱은 2^{25}이고, 이것을 M이라 놓으면 제안한 시스템의 선형 복잡도는 다음과 같다.

이 시스템은 MUX 시스템에서 h=4인 경우이므로,

$$LC = n(1+\sum_{i=1}^{h} {}_{M}C_i)$$

에 의해서 선형 복잡도는 $35(1 + {}_{M}C_1 + {}_{M}C_2 + {}_{M}C_3 + {}_{M}C_4)=2^{105}$이다.



< 그림 3 > 개선된 암호 시스템

이것은 기존 시스템에 비해 주기가 같으면서도 선형 복잡도가 크게 증가한 것이다.

제안한 시스템을 구현하여 스트림 암호 시스템의 분석 기준에 의해 랜덤성 테스트 결과는 < 표 4 > 와 같다.

표에서 나타낸 바와 같이 랜덤성 테스트를 모두 통과 하였고 MUX 시스템 이나 BRM 시스템보다는 우수하지만, BRM 시스템보다 효율성 측면에서 난수를 생성하는데 걸리는 시간이 지체되었음을 알 수 있다. 그 이유는 이 시스템이 BRM 시스템을 4개 사용하여 구현되었으므로 BRM 시스템의 출력이 나와야만 난수를 발생시킬 수 있기 때문에 생기는 현상이다. 하지만 각 BRM 시스템을 하드웨어적으로 구성하여 제안한 시스템의 입력비트 발생을 병렬로 처리한다면 난수 발생 시간이 더욱 단축될 것이다.

< 표 4 > 제안한 시스템의 분석 결과

	Frequency Test	Serial Test	Poker Test	Autocor-relation Test	Run Test	Time (sec)
8000bit	1.624	1.475	241.024	0.940	20.036	3.066
16000bit	0.100	1.240	228.480	1.088	20.124	6.134
24000bit	0.146	1.121	212.197	1.266	22.803	9.246
32000bit	0.128	1.170	220.184	1.344	23.114	12.389
64000bit	1.325	2.423	224.564	1.756	25.449	23.254
120000bit	2.341	1.989	221.654	1.698	23.453	45.675

5. 결 론

현재 사용되고 있는 암호 시스템으로는 블럭 암호 시스템인 DES 와 궤환 레지스터를 이용한 스트림 암호 시스템이 있다. 이 중 스트림 암호 시스템은 통계적인 특성과 구현의 용이함, 특히 수학적으로 보안성의 검증이 가능하기 때문에 LFSR 을 이용하여 구현되고 있다. 그러나 LFSR 자체의 선형성으로 인해 그 자체만으로는 고도의 안전성과 보안성을 요구하는 암호 시스템으로 사용하기에는 부적합하다. 따라서, 본 논문에서는 스트림 암호 시스템으로 제안되어 있는 MUX 시스템, BRM 시스템 등의 알고리즘을 연구해보고, 구현하여 통계적 방법을 통해 랜덤성을 검증해 보았으며, 기존의 암호 시스템이 가지고 있는 높은 원시 다항식의 차수를 사용하여야 한다는 문제점을 보완하여 이들 두 시스템의 특성을 이용하여 5차 이하의 원시 다항식을 사용한 새로운 암호 시스템을 제안하고 구현하였다.

본 논문에서 제안한 새로운 시스템은 암호 강도 이론에 있어서 우수함이 분석을 통해 증명 되었고, 차수가 낮은 원시 다항식을 적용한 LFSR을 사용하여 동일한 주기에서도 선형복잡도를 증가시킬 수 있었다. 이것을 하드웨어적으로 구현하게 되면 빠른 데이타의 처리를 요구하는 무선 데이타 통신에 이용할 수 있을 것으로 생각된다.

앞으로는 스트림 암호 시스템을 하드웨어적으로 구현할 수 있는 연구가 진행되어야 하겠다.

참고 문헌

[원동95] 원동호,"스트림 암호알고리즘의 안전성 평가 연구", 성균관 대학교,1995

[한국91] 한국전자통신연구소, "현대암호학 ",한국전자통신연구소,1991

[Jen84] S. M. Jennings, " Linear Equivalence of Certain BRM Shift Register Sequence ", IEEE Electronics Letters Vol.20 NO.24 1984.11.22

[Lee94] Man Young Lee, "Cryptograph and Secure Communications",McGraw-Hill Series on Computer Communications, 1994

[Mas88] G. Xiao, J. L. Massey, " A Special Characterization of Correlation Immune Combining Functions ",IEEE Trans.Inform.Theory, Vol.34,No.3, 1988

[Rue92] R.A.Rueppel,"Security Models and Notions for Stream Ciphers",Oxford University Press, 1992

[Sie84] T.Siegenthaler,"Correlation-Immunity of Nonlinear Combining Functions for Cryptographic Applications ",IEEE Trans,Information Theory, Vol.30 No.5, Sep. 1984