두개의 S박스를 활용한 ZST 스트림 암호 알고리즘

박미옥° 전문석
숭실대학교 일반대학원 컴퓨터학과
mopark°@cherry.ssu.ac.kr

ZST Stream Cipher using Two S-boxes

Miog Park° Moonseog Jun
Dept. of Computer Science, SoongSil University

요 약

이동통신의 지속적인 발달은 사용자들에게 많은 편리성을 제공해주고 있다. 이와 반면에, 이동통신의 개방성은 무선공격에 심각하게 노출되어 있으며, 안전한 통신을 위해 이동통신망의 보안은 필수적이다.

본 논문에서는 이동통신상에 전송되는 데이터를 보다 안전하게 보호하기 위한 메커니즘으로서 스트림 암호알고리즘에 두개의 S박스를 사용하고, 두 개의 S박스 사용에 따른 메커니즘을 제안한다. 먼저, DES의 각 S박스에 대한 랜덤성을 테스트하여 랜덤특성이 좋은 두개의 S박스를 고찰한다. 두 개의 S박스는 제안하는 메커니즘에 따라 스트림 암호알고리즘에 적용하며, 이 때 두개의 S박스는 비트가 0이면 S박스를 통과하고, 1이면 통과하지 않는 메커니즘을 사용한다. 이에 대한 실험은 기존 모델과의 비교분석을 통해 제안한 모델의 효율성을 증명한다.

1. 서 론

이동통신기술의 지속적인 발달로 인하여 이동통신 사용자들은 언제, 어디서나 원하는 사용자와 통화를 할 수 있을 뿐 아니라 이동단말기를 이용하여 다양한 서비스를 제공받을 수 있다. 그러나, 이러한 이동통신의 편리성에도 불구하고 이동통신의 개방성은 날로 급증하는 무선해킹과 같은 심각한 보안문제에 직면해 있으며, 그에 따른 피해도 날로 급증하고 있다. 유럽의 대표적인 디지털 이동통신은 GSM(Global System for Mobile Communications) 방식이며, 미국의 대표적인 디지털 이동통신은 CDMA(Code Division Mutiple Access) 방식이다. GSM은 CDMA에 비해 훨씬 많은 사용자를 보유하고 있지만, GSM과 CDMA 보안모델 모두 불안전한 것으로 증명되고 있다[1][2][3]. 본 고에서는 무선상의 데이터를 보다 안전하게 전송하기 위해 블록 암호방식에서 주로 사용하는 비선형 함수인 S박스를 변형하여 스트림 암호알고리즘에 적용하는 메커니즘을 제안한다. 본 논문의 구성은 2장에서 스트림 암호의 기본 개념을 설명하고, 3장에서는 기존의 스트림 암호알고리즘에 변형된 형태의 S박스를 적용하기 위한 개념과 구조, 그리고 그에 따른 메커니즘을 설명한다. 4장에서는 DES의 8개 S박스에 관한 고찰과, 기존모델과 제안모델과의 랜덤성과 상관특성 테스트 결과를 제시하고 분석하여 제안모델의 효율성을 검증한다. 마지막으로, 5장에서는 결론을 논하고 본 고를 마친다.

2. 관련 연구

스트림 암호방식은 LFSR(Linear Feedback Shift Register)을 비선형으로 결합한 이진수열 발생기를 근간으로 하는 암호시스템으로서, 키 스트림은 평문과 비트단위로 XOR을 수행하여 암호문을 생성하며, 다음과 같이 나타낼 수 있다.

$$C_i = M_i \oplus K_i \qquad \text{for } i=1,2,3,... \qquad (1)$$

C_i는 암호문의 비트열, M_i는 평문의 비트열을 의미한다. K_i는 일련의 키 스트림을 의미하고, $\oplus$는 XOR 연산을 나타낸다. 전체 키 스트림이 임의적이기 때문에 원타임 패드는 실용적이지 못하다. 즉, 키가 평문 길이만큼 길어야 하므로 키 분배나 관리 문제가 어렵게 된다. 따라서 원타임 패드의 동작 특성에 근사하도록 하면서 실용적으로 적용하기 위해 개발된 것이 의사 난수로서 발생된 키 수열을 사용하는 스트림 암호이다. 스트림 암호시스템의 구성도는 그림 1과 같다[4].

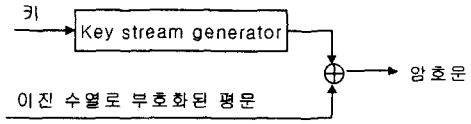

그림 1 스트림 암호시스템의 구성도

3. ZTS 모델

3.1 모델의 원리

본 절에서는 두개의 S박스를 활용한 제안모델의 개념과 동작절차에 대해 설명한다. 제안한 모델의 효율성 비교를 위해 사용하는 기존의 스트림 암호알고리즘으로는 A5를 사용한다. A5는 유럽에서 주로 사용하는 이동통신상의 암호알고리즘으로서 비밀키와 프레임번호를 입력으로 받아들여, 3개의 LFSR(각각 23단, 22단, 19단)의 동작으로 키 수열을 생성한다[2][5].

스트림 암호알고리즘에서는 결과값이 얼마나 좋은 랜덤특성을 나타내는 것이 큰 관건이기 때문에 DES의 8개 S박스 중 상대적으로 더 좋은 랜덤특성을 가지는 여덟 번째 S박스 S8과 두 번째 S박스 S2를 사용한다. 이에 대한 실험결과는 4.1절에서 설명한다. 또한, 제안모델에서는 S8과 S2를 생성된 모든 비트에 사용하는 것이 아니라 그림 2와 같이 기존의 A5의 출력비트가 0인 경우에만 S박스를 통과하는 메커니즘을 사용한다. 만약, 비트가 1인 경우는 S박스를 통과하지 않고 기존의 스트림 알고리즘 방식대로 데이터를 암호화한다. 1과 0의 모든 비트에 대해 S박스를 통과시키지 않는 이유는 모든 비트에 S박스를 통과시킨 후의 출력결과와 한 비트에만 S박스를 통과시킨 후의 출력결과에 나타나는 일정한 패턴을 고려할 경우, 후자의 경우에 일정한 패턴이 덜 나타나기 때문이다. 일정한 패턴이 약화된다는 것은 그만큼 공격에 더 강하게 되어 결국 전체 알고리즘은 공격에 더 강하게 된다. 그래서, 제안모델에서는 0과 1의 모든 비트에 S박스를 통과하지 않고, 0인 경우에만 S박스를 통과시키는 간단한 메커니즘을 사용하여 데이터를 보다 안전하

게 암호화하고자 한다. 본 고에서 제안한 메커니즘은 비트가 0(Zero)일때만 두개(Two)의 S박스(S-box)를 통과시킨다는 의미로서 Zero, Two, S-box의 첫글자를 따라 ZTS 모델이라 칭하기로한다

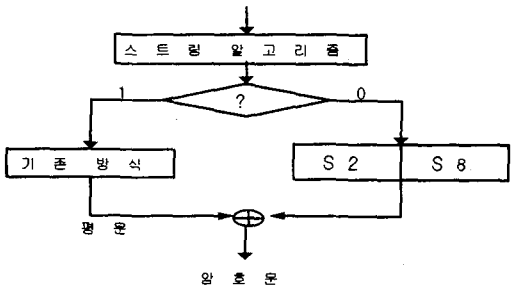

그림 2 제안모델의 구조

그림 2는 제안모델의 동작절차를 나타낸 것이다.

[1단계]비밀키와 프레임 번호를 입력으로 받아 기존의 스트림 암호알고리즘인 A5를 수행하여 키수열을 생성한다.

[2단계]1단계에서 생성된 비트가 1이면, 기존의 스트림 암호방식대로 암호화한 후, 4단계로 이동한다.

[3단계]1단계에서 생성된 비트가 0이면, S2와 S8로 구성된 S박스를 통과한 후, 그 다음 4단계로 이동한다.

[4단계]각 2단계와 3단계에서의 결과값을 평문과 XOR한다.

다음은 [3단계]에서 생성된 비트가 0인 경우, S박스 통과단계로서 S박스를 사용하기 위해 행과 열을 결정하는 S박스 행열 메커니즘을 설명한다. S박스 통과단계는 본 고에서 사용하는 S박스가 DES의 S박스이기 때문에 DES의 S박스 행열 방법에 따라 6개의 입력비트를 사용한다. 또한, DES는 $S(b_0\,b_1\,b_2\,b_3\,b_4\,b_5)$의 여섯 개 입력비트 중 첫 번째 입력비트인 b_0와 여섯 번째 비트인 b_5로 행을 결정하고, 나머지 $b_1\,b_2\,b_3\,b_4$의 4개 비트로 열을 결정한다. 제안모델의 S박스 통과단계에서도 이 S박스 행열 방법에 따라 두개의 S박스 중 하나를 순서대로 통과하게 된다. S박스 행열 메커니즘은 표 1과 같다.

표 1 S박스 행·열 메커니즘

```
S_ROW = gb[0]*2 + gb[5];
S_COL = gb[1]*8 + gb[2]*4 + gb[3]*2 + gb[4];
S_OUT = S_BOX[S_index][S_ROW*S_COL];
```

*S_ROW*는 행을 결정하기위해 사용되는 변수로서, 첫 번째와 여섯 번째 비트인 gb[0]와 gb[5]를 계산한 값을 저장한다. *S_COL*은 열을 결정하기 위한 값을 저장하는 변수로서 두 번째에서 다섯 번째 비트를 나타내는 gb[1], gb[2], gb[3], gb[4]을 계산한 값이다. *S_index*는 두개의 S박스 사용순서로서, 사용순서는 순차적으로 사용한다. 처음에 S2가 사용되면 그 다음 사용순서는 S8이다. *S_BOX*는 사용되는 S박스 배열이름을 의미한다. 선택된 S박스의 행과 열은 앞에서 계산된 *S_ROW*와 *S_COL* 변수값을 곱셈연산하여 S박스안의 임의의 값을 출력하게 된다. S박스안의 임의의 출력값을 저장하는 변수는 *S_OUT*이다. 제안한 모델의 S박스 통과과정은 S박스의 입력을 고려하여 DES의 S박스 입력비트인 6비트로하여 S박스를 통과하고 이를 통과한 출력은 평문과 XOR를 수행한다. 이러한 방법으로 모든 동작절차를 통과한 출력비트는 다음과 같은 간단한 수식으로 나타낼 수 있다.

□비트가 0일 경우 $S_i[\,Out_{A5_{i-6}}]$ (2)

□비트가 1일 경우 Out_{A5_i} (3)

비트가 0일 경우 수식 $S_i[\,Out_{A5_{i-6}}]$에서 i-6 의 의미는 S박스를 통과하기 위해서 다음에 연속적으로 생성된 6개 비트를 모두 사용한다는 의미이다. 수식 (3)은 비트가 1인 경우 S박스를 통과하지 않고 한 비트씩 처리하기 때문에 i로 나타낸다.

4. 실험 결과 및 분석

본 절에서는 제안모델이 기존모델보다 더 높은 랜덤성을 실제로 제공하는지를 테스트한다. 실험환경은 UltraSPAC-II 400MHz(두개)의 CPU와 2048M의 메모리, 디스크는 8G(7개)인 Sun Enterprise 3500에서 실험하였고, 사용한 언어는 C 언어이다. 각 출력 수열의 랜덤성 테스트는 Ent 의사난수 테스트 프로그램(Pseudorandom Number Sequence Test Program)을 사용하였고[7], 각 실험에 사용한 비트는 약 30500비트이다.

4.1 S박스

본 절에서는 DES의 8개 S박스에 대한 실험결과를 함께 보임으로써 제안모델에서 사용하는 두개의 S박스가 상대적으로 더 좋은 랜덤성을 가진다는 것을 증명한다.

그림 3은 DES의 8개 S박스에 대한 랜덤성을 실험한 결과이다. 그림의 랜덤값은 arithmetic mean 테스트 결과를 의미하며, 파일안의 모든 바이트를 합하여 파일 길이로 나눈 결과로서 127.5에 가까울수록 더 좋은 랜덤특성을 가진다는 의미이다. 그래서, 더 낮은 값을 출력한 S박스들은 *arithmetic mean*의 정의에 따라 127.5에 더 멀기 때문에 랜덤성이 더 나쁘다는 것을 알 수 있고, 반대로 다른 S박스들에 비해 가장 높은 값을 출력한 S8 박스는 상대적으로 더 좋은 랜덤성을 가진다는 것을 의미한다. 두 박스에 대한 횟수의 비교는 127번의 각 실험에서 S8이 S4보다 112번 더 높은 값을 출력하고, S8이 S2보다는 126번 더 높은 값을 출력하여 각 횟수의 비교에서 S8이 가장 좋은 랜덤특성을 가진다. S4와 S2의 비교에서는 S2가 S4보다 67번 더 좋은 랜덤특성을 나타내어 큰 차이없이 비슷한 랜덤특성을 나타내었다. 실험결과, 8개 S박스 중 상대적으로 더 좋은 랜덤성을 가지는 S8과 S2를 사용하여, 제안모델의 S박스 통과단계에서 사용했기 때문에 더 나쁜 랜덤성을 가지는 다른 S박스를 사용한 것보다 제안모델에 더 좋은 랜덤성을 제공한다는 것을 알 수 있으며, 본 고에서 다른 S박스가 아닌 S8과 S2의 사용근거를 제시한다.

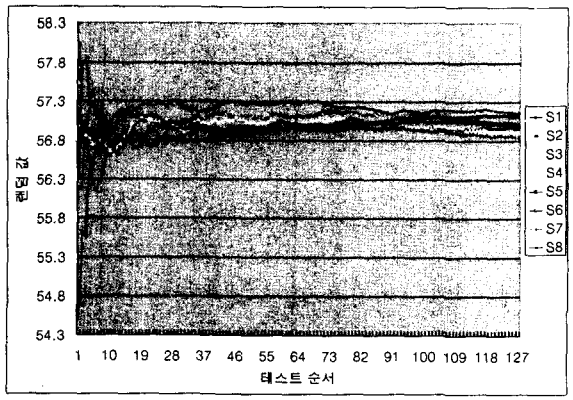

그림 3 각 S박스에 대한 랜덤성 비교

4.2 세 모델의 비교

본 절에서는 기존모델과 제안모델, 그리고 비트가 0과 1인 모든 경우에 S박스를 통과시키는 S박스 모델에 대한 실험결과를 비교·분석한다. 그림 4에서 보는 것처럼, 기존모델은 56.대의 랜덤값을 출력하면서 비트가 증가할수록 조금씩 더 낮은 값으로 이동되는데 반해, 제안모델은 68.대의 랜덤값을 출력하고 사용비트가 증가할수록 조금씩 더 높은 값을 출력하는 것을 알 수 있다. 모든 비트에 대해 S박스를 통과하도록 사용한 S박스 모델은 기존모델보다 약간 더 높은 59.대의 랜덤값을 출력하였다. 횟수비교에서는, 127번의 모든 실험에서 제안모델이 기존모델보다 더 높은 랜덤값을 출력하기 때문에 arithmetic mean의 정의에 따라 제안모델이 기존모델보다 더 좋은 랜덤특성을 가지는 것을 알 수 있다. 그 다음으로는 S박스 모델이 기존모델보다 더 좋은 랜덤특성을 가지지만 랜덤값에 많은 변화가 없음을 알 수 있고, 결과적으로, 비트가 0인 경우에만 S박스를 통과시키는 메커니즘이 전체적인 알고리즘의 랜덤성을 향상시키는 것으로 나타나 제안모델의 효율성을 증명한다. 좋은 랜덤특성을 가진다는 것은 스트림 암호알고리즘에서 큰 관건이 되는 중요한 기준중의 하나로서 제안모델이 기존모델보다 더 향상된 랜덤특성을 가진다는 것은 출력결과값에 랜덤성이 향상된 만큼 더 좋은 랜덤특성이 나타난다는 것을 의미하기 때문에 출력결과가 그만큼 난수에 더 가깝다는 것이고, 출력결과로부터 입력이나 키 값을 유추하기가 어렵다는 것이다. 그래서, 제안모델은 기존모델보다 더 좋은 랜덤특성을 나타냄으로써 전송되는 데이터를 더 안전하게 보호한다고 말할 수 있다.

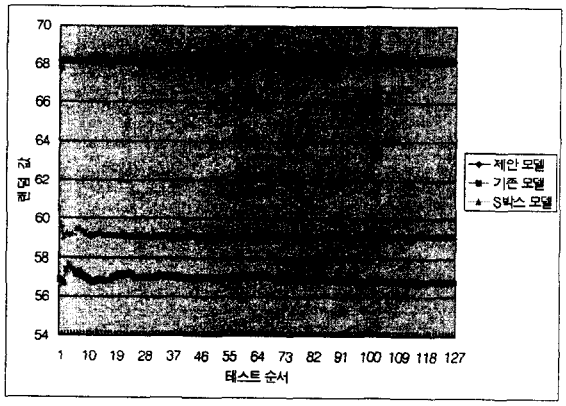

그림 4 Randomness 비교

Serial correlation은 파일안의 각 바이트와 이전 바이트와의 의존도를 나타내는 것으로서, autocorrelation을 의미한다. 결과값은 양수나 음수가 될 수 있고, 0에 가까울수록 더 좋은 상관특성을 가진다는 의미이다. 그림 5는 세 모델에 대한 serial correlation 특성을 실험한 결과로서, 기존모델은 비트가 증가할수록 처음보다는 0에 더 가까워지지만, 제안모델과 비교하면 0과 많이 떨어져 있는 것을 확인할 수 있다. 기존모델은 처음의 최고 0.12대의 값에서 0.001대의 값의 변동으로 0.074대의 값까지 0에 가까워지고, 제안모델은 0.059값이 0에서 가장 멀리 떨어진 값이고, 대부분 0.03대의 값을 출력하였다. S박스 모델은 0.34에서 0.33대의 값을 출력함으로써, 기존모델과 제안모델에 비해 0과 훨씬 멀어져 있는 것을 알 수 있다. 결과적으로, serial correlation 정의에 따라 제안모델이 기존모델보다 0에 훨씬 더 가깝기때문에 제안모델이 기존모델보다 더 좋은 상관특성을 가진다고 말할 수 있으며, S박스 모델이 가장 나쁜 상관특성을 나타내는 것을 알 수 있다. 좋은 상관특성은 상관특성의 정의에 따라 그 만큼 각 바이트들간의 관계성이 적기 때문에 상관공격에 더 강하다는 것을 의미하여, 상관공격에 가장 좋은 모델은 제안모델이라는 사실을 증명하였다.

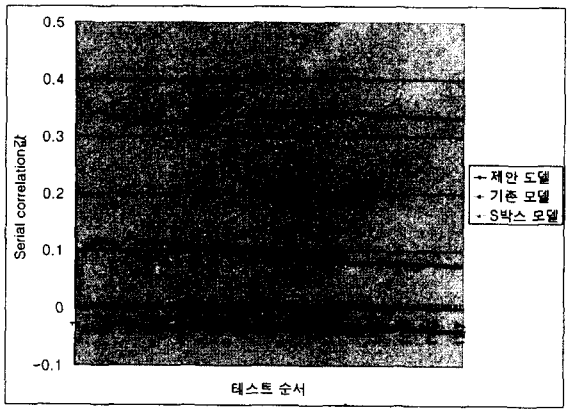

그림 5 Serial Correlation 비교

5. 결 론

본 고에서는 불안전한 기존의 스트림 암호알고리즘을 보다 안전하게 구성하기위한 메커니즘으로서 블록 암호방식에서 주로 사용한 비선형함수인 S박스의 사용을 제안하였다. 제안모델에서 사용하는 S박스는 8개 S박스중 랜덤성 테스트를 거쳐 상대적으로 좋은 랜덤성을 가지는 두 개의 S박스만을 사용하여 이동통신상의 특징인 작은 메모리와 낮은 계산율에 의한 연산의 효율성을 이루고자하였다. 또한, 이 S박스의 사용은 비트가 0인 경우에만 S박스를 사용하고 1인 경우는 S박스를 사용하지 않고 기존방식대로 암호화하는 간단한 메커니즘에 의해 전체 알고리즘의 랜덤 특성과 상관 특성의 향상을 이루었다.

제안모델의 효율성에 대한 검증은 4장의 그림 4와 5를 통해 기존모델, 제안모델, S박스 모델을 비교분석함으로써 제안모델이 가장 좋은 모델이라는 것을 알 수 있었고, S박스 모델과 제안모델의 비교를 통해서는 비트가 0인 경우에만 S박스를 통과시키는 간단한 메커니즘의 사용이 전체 알고리즘의 랜덤특성과 상관특성에 많은 영향을 미치는 것을 확인하여 비트가 0인 경우에만 S박스를 사용하는 메커니즘의 효율성을 증명하였다. 결과적으로, 제안모델은 비트가 0인 경우에만 두 개의 S박스를 순서대로 통과시키는 간단한 메커니즘에 의해서 이동통신상에 전송되는 데이터를 보다 안전하게 보호할 수 있다고 말할 수 있다.

참 고 문 헌

[1] Jovan D. Golic, "Cryptanalysis of Alleged A5 Stream Cipher", Advances in Cryptology- EUROCRYPT '97, 1977, pp.239-255, May. 1997.

[2] Alex Biryukov, Adi Shamir, David Wager, "Real Time Cryptanalysis of A5/1 on a PC," Fast Software Encryption Workshop 2000, Vol.40, pp.71-79, Apr. 2000.

[3] David Wanger. Bruce Schneier, John Kelsey, "Cryptanalysis of the Cellular Message Encryption Algorithm", Draft-2000, Counterpane Labs. March 1997.

[4] 강주성, 김재현, 박상우, "현대 암호학", pp93-135, ETRI부설 국가보안기술연구소, 영문사, 2000.

[5] Eli Biham, Orr Dunkelman, "Cryptanalysis of the A5/1 GSM Stream Cipher", Progress in Cryptology - INDOCRYPT 2000, LNCS 1977, pp.43-51, Dec. 2000.

[6] John Walker, "ENT A Pseudorandom Number Sequence Test Program", http://www.fourmilab.ch/random/