두개의 S박스를 활용한 이중 S박스 스트림 암호알고리즘

박미옥*, 전문석*
*숭실대학교 컴퓨터학과
e-mail:mopark@cherry.ssu.ac.kr

Double S-box Stream Cipher using Two S-boxes

Mi-Og Park*, Moon-Seog Jun*
*Dept of Computer Science, Soong-Sil University

요 약

이동통신기술의 지속적인 발전은 많은 사용자들에게 시간과 장소에 구애받지 않고 어디서든지 통신 서비스를 받을 수 있도록 제공해주고 있다. 하지만 이동통신기술의 개방성은 항상 심각한 보안 위협에 노출되어 있으며, 보안 안전성이 높은 통신채널 확보를 위해서는 암호화가 반드시 필요하다. 이동통신망의 보안을 위해서 사용하는 가장 일반적인 방법 중의 하나는 스트림 암호알고리즘이다.

본 논문에서는 전송되는 데이터를 보다 안전하게 암호화하기 위해 DES에서 사용하는 S박스중 실험을 통해 랜덤성이 좋은 두 개의 S박스를 조사하고, 비트가 0이면 통과하고 1이면 통과하지 않는 메커니즘을 제안한다. 실험을 통해 제안모델이 기존모델보다 향상된 랜덤성과 상관특성을 나타냄을 증명하며, 비트가 0인 경우에만 S박스를 사용하는 메커니즘과 0과 1의 모든 비트에 S박스를 통과시킨 모델을 비교하여 제안 메커니즘의 효율성을 증명한다.

1. 서론

스트림 암호는 1970년대 초반 유럽에서 발전된 LFSR(Linear Feedback Shift Register)를 이용한 이진수열 발생기를 기본으로 하고 주기, 선형 복잡도 등 비교적 수학적 분석이 가능한 수치를 사용한다는 장점이 있다. 암호화 과정은 평문이진수열을 이진수열 생성기에서 출력된 키 이진수열과 비트별로 XOR시킴으로서 이루어지며, 다음과 같은 수식으로 나타낼 수 있다.

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

여기서, C_i는 암호문의 비트열, M_i는 평문문자의 비트열, K_i는 키 수열, $\oplus$는 XOR 연산을 의미한다[1][2].

본 논문에서는 이동통신상의 데이터를 보다 안전하게 전송하기 위해 블록 암호방식에서 주로 사용하는 비선형 함수인 S박스를 변형하여 스트림 암호알고리즘에 적용하는 메커니즘을 제시한다. 제안모델에서 사용하는 S박스는 DES의 S박스로 DES의 8개 S박스중에 상대적으로 더 좋은 랜덤특성을 가지는 S박스를 조사하고, 이 중 두개의 S박스를 선택하여 본 고에서 제안하는 변형된 메커니즘에 의해 기존의 스트림 암호알고리즘에 적용한다. 또한, 테스트를 통해 변형된 S박스 메커니즘을 적용한 모델이 기존모델보다 향상된 랜덤성과 상관특성을 나타냄을 증명한다.

본 논문의 구성은 2장에서 변형된 형태의 S박스를 적용하기 위한 개념과 구조, 그리고 그에 따른 메커니즘들을 설명하고, 3장에서는 테스트 결과를 제시하고 분석하여 제안모델의 효율성을 증명한다. 마지막으로, 4장에서는 결론을 논하고 본 고를 마친다.

2. 제안 모델

2.1 개념과 동작절차

제안한 모델의 기존모델로는 유럽에서 주로 사용

하는 이동통신상의 암호알고리즘인 A5를 사용한다. 이 알고리즘은 비밀키와 프레임번호를 입력으로 받아, 3개의 LFSR(23단, 22단, 19단)에 의한 동작으로 키 수열을 생성한다[3][4].

제안모델의 동작절차는 그림 1과 같이 기존 알고리즘의 함수중 하나인 getbit()의 결과값에 따라 결과값이 0이면 이중의 S박스를 순서대로 통과하게 되고, 결과값이 1이면 S박스를 통과하지 않고 기존의 알고리즘 방식대로 처리하게 된다. getbit() 함수는 각 LFSR의 출력값을 처리하는 부분이다. 이러한 방법으로 처리된 결과값은 평문과 XOR을 수행하여 최종적인 암호문을 생성한다. S박스 통과단계는 S8을 먼저 통과하고 그 다음에 S2박스를 통과하는 이중의 S박스 통과방법을 사용한다. S8과 S2는 DES의 8개 S박스중 실험을 통해 상대적으로 좋은 랜덤 특성을 가지는 S박스를 의미이다.

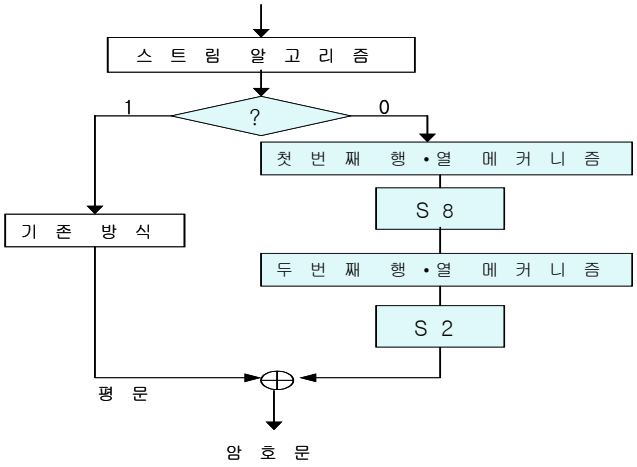

그림 1 제안 모델의 구조

다음은 제안한 모델의 동작절차이다.

[1단계] 비밀키와 프레임번호를 입력으로 받아 기존의 스트림 암호알고리즘인 A5를 수행하여 키 수열을 생성한다.

[2단계] 생성된 비트가 1이면, 기존의 스트림 암호 방식대로 연산처리한 후, 4단계로 이동한다.

[3단계] 1단계에서 생성된 비트가 0이면, 첫번째 S박스 행・열 메커니즘에 따라 S8을 통과한 후, 두 번째 행・열 메커니즘에 의해 S2박스를 통과한다. 그 다음은 4단계로 이동한다.

[4단계] 각 2단계와 3단계에서의 결과값을 평문과 함께 XOR한다.

2.2 S박스 행・열 메커니즘

본 절에서는 이중의 S박스 통과단계에서 사용하는 S박스 행・열 메커니즘에 대해 설명한다. S박스 통과과정은 DES의 S박스를 사용하기 때문에 입력 비트를 DES의 S박스 원리에 따라 6개비트중 첫번째 비트인 b_0와 여섯번째 비트인 b_5의 두 개 비트로 행을 결정하고, $b_1\ b_2\ b_3\ b_4$의 4개비트로 열을 결정한다. 첫번째 S박스 통과단계에서의 S박스 행・열 메커니즘은 이 방식을 그대로 사용하며, 이 때 사용되는 S박스는 S8이다.

표 2 첫 번째 S박스 행・열 메커니즘

```
S1_ROW = b[0]*2 + b[5];
S1_COL  = b[1]*8+ b[2]*4 + b[3]*2 + b[4];
S1_OUT = S1[S1_ROW][S1_COL];
```

표 2의 *S1_ROW*는 행을 결정하기 위해 계산된 결과값을 저장하는 변수이고, *S1_COL*은 열을 결정하기 위한 결과값을 저장하는 변수이다. *b*[]로 표시된 부분은 여섯개의 비트중 몇 번째 비트인가를 나타낸다. 예를들면, *b*[0]은 첫번째 비트, *b*[1]는 두번째 비트를 의미한다. *S1*은 첫번째 S박스의 배열이름, *S1_OUT*은 첫번째 S박스를 통과하고 나서 출력되는 값을 의미한다.

두 번째 S박스 통과단계에서의 행・열 메커니즘은 첫 번째 S박스 통과단계의 출력값을 이용하여 S2박스의 행・열을 결정하는 메커니즘으로서 표 3과 같다.

표 3 두 번째 S박스 행・열 메커니즘

```
S2_ROW = b[n1]*2 + b[n2];
S2_COL = b[n1]*8+b[n2]*4+b[n3]*2+b[n4];
S2_OUT = S2[S2_ROW][S2_COL];
```

표 3의 *S2_ROW*는 행을 계산한 결과값을 저장하는 변수이고, *S2_COL*은 열을 결정하기 위해 계산된 결과값을 저장하는 변수이다. *b*[]로 표시된 부분은 표 2와 동일한 의미로서, []안의 n1, n2, n3, n4는 첫번째 S박스를 통과한 후의 출력값인 4개의 비트를 나타낸다. 예를 들어, n1은 S8박스의 출력값중 첫번째 비트, n2는 두 번째 비트, n3는 세 번째, n4는 네 번째 비트를 의미한다. 두 번째 S박스 행・열 메커니즘은 첫 번째 S박스 통과단계에서 출력된 이 4개비트를 사용해서 두 번째 S박스의 행・열을 결정하기 위해 사용된다. *S2*는 두 번째 S박스 통과단계에서 사용하는 S박스로서, 그림 1에서 S2로 표기된 부분을 의미한다. *S2_OUT*은 두 번째 S박스를 통과한 후의 출력값을 저장하는 변수를 나타낸다. 두 번째 행・열 메커니즘은 행을 결정하기 위해서 첫 번

째 S박스의 출력값중 첫 번째(n1)와 두 번째(n2) 비트를 사용하고, 열을 결정하기 위한 값으로는 첫 번째 S박스의 출력값인 4개비트를 모두 사용한다. 두 개의 S박스 행·열 메커니즘에 의해 [3단계]의 S박스 통과단계와 비트가 1인 경우 S박스를 통과하지 않고 [2단계]의 기존방식대로 암호화한 후의 각각의 출력비트는 평문과 XOR을 수행한다.

3. 실험 결과 및 분석

실험환경은 UltraSPAC-II 400MHz(두개)의 CPU와 2048M의 메모리, 디스크는 8G(7개)인 Sun Enterprise 3500에서 실험하였고, 사용한 언어는 C 언어이다. 각 출력수열의 랜덤성 여부를 테스트하기 위해 사용한 프로그램은 John Walker에 의해 작성된 Ent(Pseudorandom Number Sequence Test Program) 프로그램이다[5]. 실험에 사용한 비트는 약 30500 비트이며, 127번으로 나누어 실험하였다.

3.1 DES의 S박스에 관한 고찰

본 절에서는 DES의 8개 S박스에 대한 실험결과를 함께 보임으로써 제안모델에서 사용하는 두개의 S박스가 상대적으로 더 좋은 랜덤성을 가진다는 것을 증명한다. 그림 2의 랜덤값은 arithmetic mean 테스트 결과를 의미하며, 파일안의 모든 바이트를 합하여 파일 길이로 나눈 결과로서 127.5에 가까울수록 더 좋은 랜덤특성을 가진다는 의미이다. 127번의 각 실험에서 S8이 S4보다 112번 더 높은 값을 출력하고, S8이 S2보다는 126번 더 높은 값을 출력하여 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 있다.

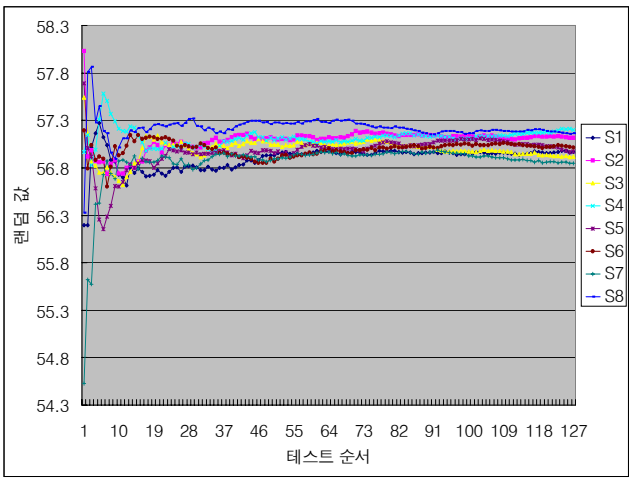

그림 2 각 S박스에 대한 랜덤성 비교

3.2 세가지 모델에 관한 고찰

본 절에서는 기존모델과 제안모델, 그리고 S박스 모델에 대한 실험을 통해 제안모델의 효율성을 증명한다. S박스 모델이란 0과 1인 모든 비트에 이중 S박스를 통과시킨 모델을 지칭한 것이다. 이러한 S박스 모델과 제안모델을 비교하는 것은 비트가 0인 경우에만 이중 S박스를 통과시키는 메커니즘의 효율성을 알아보기 위한 것이다. 그림 3은 세 모델에 대한 랜덤검정 결과를 나타낸 것으로, 기존모델은 56.대와 57.대의 값을 계속 번갈아가면서 출력되었고 비트가 증가할수록 56.대로 값이 조금씩 감소하는 것을 볼 수 있다. 제안모델은 1번째의 68.9344값에서 시작하여 3번째의 가장 낮은값인 68.1967을 출력한 후, 그 이후부터는 68.373에서 0.1씩 값이 계속 증가하여 마지막에서는 68.9431을 출력하였다. 기존모델은 4번째에서 가장 높은값인 57.6107을 출력하고, 나머지는 모두 이보다 작은값으로서 57.대와 56.대의 값을 번갈아가면서 출력하였고, 비트가 증가할수록 56.76대로 값이 조금씩 낮아져서 마지막 값은 56.766을 출력하였다. 비트가 증가할수록 값이 지속적으로 낮아진다는 것은 전송되는 데이터양이 많아질수록 더 나쁜 랜덤성을 나타낸다는 의미이기 때문에 기존모델은 랜덤성이 감소한다고 말할 수 있다. 반면에, 제안 모델은 처음 68.9344 값에서 출력하여, 3번째에서 가장 낮은 값인 68.1967을 출력하고, 그 다음부터는 68.373에서 0.1씩 값이 증가하거나 0.01

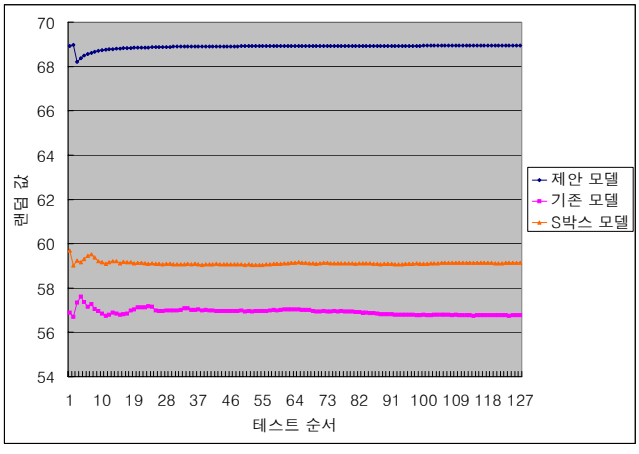

그림 3 세 모델의 랜덤성 비교

씩 값이 증가하는 추세를 보이면서 마지막으로 68.9431 값을 출력하였다. 제안모델의 값이 지속적으로 증가하는 경향을 보인다는 것은 사용하는 비트가 많아질수록 암호화되어 전송되는 데이터에 랜덤 특성이 증가한다는 것을 의미하는 것이다. S박스 모델은 59.5886에서 시작하여 59.0대와 59.1대의 값을 출력하면서 감소와 증가를 반복하였고, 그림에서처럼 처음부분에 가장 높은값들이 출력되는 경향을 보였다. S박스 모델은 제안모델에서 사용하는 비트가 0일 경우에만 이중 S박스를 통과시키는 메커니즘의 효율성을 보이기 위해 사용한 모델로서, 실험결과에서 보는 것처럼 1과 0의 모든 비트에 이중 S박스 통과단계를 사용하는 S박스 모델보다 비트가 0인 경우에만 이중 S박스를 사용하는 메커니즘의 단순함에 비해, 랜덤성이 크게 향상됨을 알 수 있다. 이것은 비트가 0인 경우에만 이중 S박스를 통과시켜 출력결과에 일정한 패턴을 약화시키고자 한 목적을 이루었다고 말할 수 있다. 실험횟수의 모든 경우에 arithmetic mean의 정의에 따라 127.5에 더 가까운 제안모델, S박스 모델, 기존 모델순으로 좋은 랜덤성을 나타내는 것을 알 수 있다. Serial correlation은 파일안의 각 바이트와 이전 바이트와의 의존도를 나타내는 것으로서, 양수나 음수값을 가질 수 있으며 0에 가까울수록 더 좋은 상관특성을 의미한다. 그림 4에서 기존모델은 대부분 0.1대의 값을 출력하였고, 비트가 증가할수록 0.09, 0.08, 0.07대까지 값이 낮아진다. 제안모델은 3번째에서 가장 높은값인 -.236695을 출력하는 것을 제외하고는 0에 훨씬 가까운 값을 출력하였고, 비트가 증가할수록 0.0001대의 값에서 0.001씩 값이 증가하면서 마지막에서는 0.010966값을 출력하였다. S박스 모델은

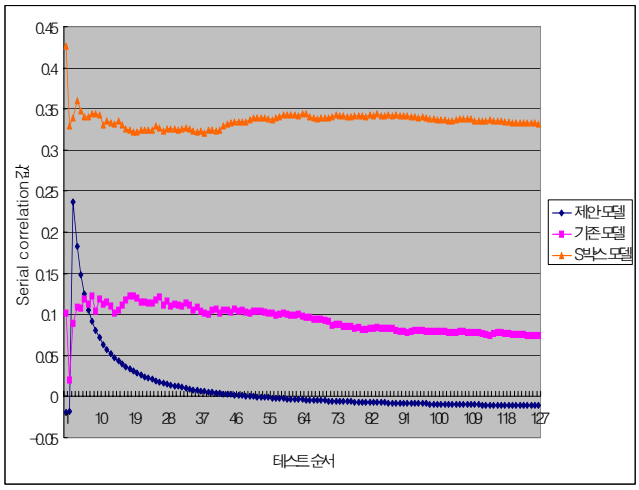

그림 4 Serial correlation 비교

0.426675에서 시작하여 2번째에서부터는 0.32대와 0.34대의 값을 출력하면서 마지막으로 0.331824값을 출력하였다. S박스 모델은 세가지 모델중 가장 높은 값을 출력하기 때문에 serial correlation값의 정의에 따라 제안모델, 기존모델, S박스 모델의 순으로 상관특성이 좋다는 것을 알 수 있다. 상관특성이 좋다는 것은 각 바이트와의 의존도가 더 낮아진다는 의미로서 세 모델중 제안모델이 상관공격에 가장 강하여 이동통신상의 데이터를 가장 안 전하게 보호한다고 말할 수 있다. 결과적으로, 그림 3과 4의 세가지 모델에 대한 실험을 통해 제안모델은 기존모델보다 더 좋은 랜덤특성과 상관특성을 나타냄으로써 제안된 메커니즘의 효율성을 증명한다고 말할 수 있다.

4. 결론

본 고에서는 이동통신상의 데이터를 보다 안전하게 보호하기 위한 메커니즘으로서 S박스의 사용과 그에 대한 새로운 메커니즘을 제시하였다. 이때, S박스의 사용은 기존모델의 getbit()함수의 출력값이 0인 경우에만 이중으로 연결된 S박스를 통과하도록 하고, 두 개의 S박스를 통과할 때 각 S박스 행·열 메커니즘에 따라 행과 열을 결정할 수 있도록 하는 메커니즘을 사용하였다. 3장의 세가지 모델에 대한 실험을 통해 제안모델의 효율성을 증명하였다. 또한, DES의 S박스 사용은 연산의 효율성과 전체 알고리즘의 비도를 향상시키기 위해 사용하는 일반적인 S박스의 사용목적과 일치하여 기존의 불안전한 알고리즘에 의해 암호화되어 전송되는 데이터를 보다 안전하게 보호하기 위한 제안목적을 증명하였다.

참고문헌

[1] A. Ruppel, "Analysis and Design of Stream Ciphers", p.5-16, Springer-Verlag, 1986.

[2]이민섭, 현대암호학, pp.118-155, 교우사, 2000.

[3] Alex B., A야 S. David W., "Real Time Cryptanalysis of A5/1 on a PC", Fast Software Encryption Workshop 2000, Vol.40, pp.71-79, Apr. 2000.

[4] Eli B., Orr D., "Cryptanalysis of A5/1 GSM Stream Cipher", Progress in Cryptology - INDOCRYPT 2000, LNCS 1977, pp.43-51, Dec. 2000.

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