2018 암호분석경진대회 : 3번 문제 답안

답) KEY : $June%18

1) 알고리즘

공격 방법: LFSR1의 초기 값은 가능한 모든 경우를 전수조사하고, O(2^{28})
각각 그에 따른 LFSR2의 초기 값을 구해 keystream을 비교함으로써 정답을 찾을 것이다.

1. LFSR1 초기 값을 정한다.

2. 전수조사를 통해 주어진 LFSR1의 초기 값에 따른 LFSR2의 초기 값을 구한다.

2-1. LFSR1이 1일때의 LFSR2의 값을 행렬식을 통해 LFSR2의 초기 값 $s_0, s_1, s_2, \cdots, s_{31}$ 으로 구성된 방정식을 만든다.

ex) s_k : $LFSR2$의 $k+1$번 째 값

$$s_{42} = s_1 \bigoplus s_0 \Leftrightarrow s_k = [0\ 1\ 0 \cdots 0\ 1] \cdot \begin{bmatrix} s_0 \\ \vdots \\ s_{31} \end{bmatrix}$$

2-2. 2-1를 32회 반복해 얻어진 32개의 방정식을 합쳐 초기 값에 대한 하나의 32*32행렬식으로 표현한다.

2-3. 가우스 소거법을 통해 행렬식을 풀어 LFSR2의 초기값을 구한다.
(초기값에 곱해진 행렬의 rank가 32라면 역행렬을 구해 식을 풀면 되지만,
대부분의 경우 rank가 28~30 사이였다. 따라서 값이 정해지지 않은 free variable에 대해서는 전수조사를 병행했다.)

3. LFSR1이 1일때의 LFSR2의 값을 주어진 keystream과 비교하여 정답을 판별한다.

시간복잡도 : $2^{28} \times 32^3$

***32^3은 행렬의 곱셈에 대한 복잡도이며, 2^{28}은 LFSR1의 전수조사에 대한 복잡도이다.**

실제 무작위의 LFSR1 초기 값 2^{10}개를 넣어 프로그램을 실행하였을 때 약 0.5초가 걸렸다.
따라서 LFSR1 초기 값에 대한 전수조사의 경우, 약 37시간 정도가 걸릴 것으로 예상하며 컴퓨터 10대에 나누어 실행할 경우 3시간 내외로 답을 얻을 수 있다.

2018 암호분석경진대회 : 3번 문제 답안

2) 풀이방법

1. LFSR1 초기 값을 전수조사한다.

KEY는 아스키 코드로 이루어져 있으므로, LFSR1[k], LFSR2[k]를 각각 LFSR1, LFSR2 초기 값의 k번째 비트라고 하면, LFSR1[x]=0 , LFSR2[x]=0 $(x \in 31,23,15,7)$임을 알 수 있다. 따라서 전수조사의 가짓수는 2^{28}개가 된다.

2. 전수조사를 통해 주어진 LFSR1의 초기 값에 따른 LFSR2의 초기 값을 구한다.

2-1. LFSR1이 1일때의 LFSR2의 값을 LFSR2의 초기값 $s_0, s_1, s_2, \cdots, s_{31}$ 으로 구성된 방정식으로 표현하고, 이를 행렬식으로 구현한다.

32*32행렬 μ를 (1,2),(2,3),....,(31,32)와 (32,23),(32,3),(32,2),(32,1)은 1이고 나머지는 모두 0인 행렬로 두면, LFSR2의 값 $s_i (i = 0,1,2,...)$에 대해

$$\begin{bmatrix} s_1 \\ s_2 \\ s_3 \\ s_4 \\ \vdots \\ s_{29} \\ s_{30} \\ s_{31} \\ s_{32} \end{bmatrix} = \mu \cdot \begin{bmatrix} s_0 \\ s_1 \\ s_2 \\ s_{31} \\ \vdots \\ s_{28} \\ s_{29} \\ s_{30} \\ s_{31} \end{bmatrix}$$

이러한 식이 성립한다. 이를 일반화하면 임의의 자연수 k에 대해

$$\begin{bmatrix} s_{k-31} \\ s_{k-30} \\ s_{k-29} \\ s_{k-28} \\ \vdots \\ s_{k-3} \\ s_{k-2} \\ s_{k-1} \\ s_k \end{bmatrix} = \mu^{k-31} \cdot \begin{bmatrix} s_0 \\ s_1 \\ s_2 \\ s_3 \\ \vdots \\ s_{28} \\ s_{29} \\ s_{30} \\ s_{31} \end{bmatrix}$$

이러한 식이 성립한다.

2018 암호분석경진대회 : 3번 문제 답안

따라서, 32 이상의 임의의 자연수 k에 대해 s_k가 초기값으로 어떻게 만들어졌는지는, μ^{k-31}의 마지막 행을 보면 알 수 있다. 예를 들어, k=42일 때, 계산한 행렬의 마지막 행이 (0 1 0 0 …0 1) 이면,

$$s_{42} = s_1 \bigoplus s_{31} \Leftrightarrow s_k = [0\ 1\ 0 \cdots\ 0\ 1] \cdot \begin{bmatrix} s_0 \\ \vdots \\ s_{31} \end{bmatrix}$$

인 것이다. 이 과정은 LFSR1[k]의 값이 1일때만 시행한다.

2-2. 2-1을 32회 반복해 얻어진 32개의 식을 초기값에 대한 하나의 32*32행렬식으로 표현한다.

문제에서 주어진 keystream의 k번째 값을 keystream[k]라고 하고, LFSR1이 1인 값들 중 m번째를 LFSR1[k_m]이라고 하면, LFSR2[k_m]=keystream[m]이 성립한다.

위 2-1 과정을 32번 반복하면서 계산한 $\mu^{k_m-31}(0 \le m \le 31 \text{의 정수})$의 마지막 행을 32X32행렬 M에 차례로

한 행씩 넣는다. 즉, μ^{k_m-31} 의 마지막 행을 μ_m이라고 하면

$$M = \begin{bmatrix} \mu_0 \\ \mu_1 \\ \vdots \\ \mu_{30} \\ \mu_{31} \end{bmatrix}$$

이 되는 것이다. 따라서

$$\begin{bmatrix} keystream[0] \\ keystream[1] \\ keystream[2] \\ \vdots \\ keystream[29] \\ keystream[30] \\ keystream[31] \end{bmatrix} = \begin{bmatrix} LFSR2[k_0] \\ LFSR2[k_1] \\ LFSR2[k_2] \\ \vdots \\ LFSR2[k_{29}] \\ LFSR2[k_{30}] \\ LFSR2[k_{31}] \end{bmatrix} = M \cdot \begin{bmatrix} LFSR2[0] \\ LFSR2[1] \\ LFSR2[2] \\ \vdots \\ LFSR2[29] \\ LFSR2[30] \\ LFSR2[31] \end{bmatrix}$$

이러한 식이 만들어진다. keystream값은 문제에서 주어졌으므로, M의 역행렬을 양변에 곱하면 LFSR2의 초기 값을 알 수 있다.

하지만 시행 결과. 대부분의 경우 M의 RANK는 28에서 30 사이였고, 따라서 역행렬이 존재하지 않았다. 따라서 M의 역행렬을 구하지 않고, 가우스 소거법을 통해 M을 upper triangle 꼴로 만들고, 그때 leading 1이 없는 행의 초기 값들은 전수조사를 통해 정답을 구해내었다.

4*4행렬로 예를 들면,

$$M = \begin{bmatrix} 1\,0\,1\,0 \\ 0\,1\,0\,0 \\ 1\,0\,1\,0 \\ 0\,0\,0\,1 \end{bmatrix}$$

이고, 그때의 keytstream값이 차례로 (1 0 1 0) 이었다면, 가우스 소거법 이후에는

2018 암호분석경진대회 : 3번 문제 답안

$$M = \begin{bmatrix} 1\,0\,1\,0 \\ 0\,1\,0\,0 \\ 0\,0\,0\,0 \\ 0\,0\,0\,1 \end{bmatrix}$$

이 되고 keystream의 값은 (1 0 0 0)이 된다. 이때에는 3행의 leading 1 이 없으므로 s_2의 값에 0과 1을 모두 대입해본다.

3. LFSR1이 1일때의 LFSR2의 값을 주어진 keystream과 비교하여 정답을 판별한다.

초기 값이 모두 구해졌다면 정답 유무를 확인하는 것은 직접 LFSR1의 값과 LFSR2의 값을 차례로 계산하여 문제에 주어진 알고리즘에 따라 keystream 값을 구해, 주어진 keystream과 비교해본다. 이때 그 값이 200개의 keystream값과 모두 일치한다면 정답으로 출력하였다.