Cryptanalyzing a Keystream Obtained from the Product of Two LFSRs

D. R. Stinson

In the last supplementary note, we mentioned that we could construct a keystream from two LFSRs by multiplying their outputs. Let's recall how this would be done. Suppose that L_1 is an m-stage LFSR with (maximal) period $2^m - 1$, and L_2 is an n-stage LFSR with (maximal) period $2^n - 1$. Further suppose that $\gcd(m, n) = 1$. L_1 is initialized with a seed $j_1, \ldots j_m$, producing a sequence $\mathbf{j} = j_1, \ldots$; and L_2 is initialized with a seed $k_1, \ldots k_n$, producing a sequence $\mathbf{k} = k_1, \ldots$. Then the keystream is defined to be $\mathbf{z} = z_1, \ldots$, where $z_i = j_i k_i$, $i \geq 1$.

In this note, we describe how $\mathbf{j}$ and $\mathbf{k}$ can be retrieved from $\mathbf{z}$ in a straightforward manner. This means that a stream cipher using a keysytream constructed in this way is completely vulnerable to a known-plaintext attack, as follows. Given plaintext and corresponding ciphertext, we obtain a portion of the keystream $\mathbf{z}$ by exclusive-or-ing them together. Then we reconstruct a portion of $\mathbf{j}$ and $\mathbf{k}$ (as we will shortly describe). Finally, we obtain the linear recurrence relations that were used to construct $\mathbf{j}$ and $\mathbf{k}$ by the method described in *Cryptography Theory and Practice.*

We begin by observing that $j_i = k_i = 1$ whenever $z_i = 1$, so we easily get some keystream bits "for free". However, in order to accomplish the last step of the cryptanalysis described above, we will need $2m - 1$ consecutive bits of $\mathbf{j}$ and $2n - 1$ consecutive bits of $\mathbf{k}$.

Let's look therefore at the reconstruction of $j_1, \ldots, j_{2m-1}$. Consider any particular co-ordinate i $(1 \leq i \leq 2m - 1)$. Then it is the case that

$$j_i = j_{i+2^m-1} = j_{i+2(2^m-1)} = \ldots,$$

since $\mathbf{j}$ has period $2^m - 1$. Hence, if any of the bits

$$z_i, z_{i+2^m-1}, z_{i+2(2^m-1)}, \cdots$$

is a 1, then we can conclude that $j_i = 1$. The question is how many values in this list do we need to consider in order to conclude that $j_i = 0$.

Since $\gcd(m,n) = 1$, it is possible to prove that $\gcd(2^m - 1, 2^n - 1) = 1$ (see any book on number theory!). Consider the following bits in the sequence $\mathbf{k}$:

$$k_i, k_{i+2^m-1}, k_{i+2(2^m-1)}, \ldots, k_{i+(2^n-2)(2^m-1)}.$$

Since $\gcd(2^m - 1, 2^n - 1) = 1$, it follows that these $2^n - 1$ bits in $\mathbf{k}$ are in fact a permutation of the bits

$$k_1, k_2, \ldots, k_{2^n-1}.$$

Hence, there is at least one occurrence of a 1 in these bits of $\mathbf{k}$ (in fact there are many 1's among these bits).

This means that it suffices to consider the $2^n - 1$ bits

$$z_i, z_{i+2^m-1}, z_{i+2(2^m-1)}, \ldots, z_{i+(2^n-2)(2^m-1)}.$$

If any of these values is a 1, then $j_i = 1$ (otherwise, $j_i = 0$).

This is not a very satisfactory approach, however, since it requires a very long portion of the keystream $\mathbf{z}$. So we might ask if we can do better. What we require is a set of ℓ bits of $\mathbf{z}$,

$$z_i, z_{i+2^m-1}, z_{i+2(2^m-1)}, \ldots, z_{i+(\ell-1)(2^m-1)},$$

with the property that at least one of the bits

$$k_i, k_{i+2^m-1}, k_{i+2(2^m-1)}, \ldots, k_{i+(\ell-1)(2^m-1)}$$

is a 1. We have seen that $\ell = 2^n - 1$ will work, but can we get by with some smaller value ℓ?

In fact, it suffices to take $\ell = n$ (see the discussion below). So we obtain the following algorithm to construct the desired $2m - 1$ bits of $\mathbf{j}$ from $\mathbf{z}$:

for $i \leftarrow 1$ **to** $2m - 1$ **do**
    $j_i \leftarrow 0$
    for $h \leftarrow 0$ **to** $n - 1$ **do**
        if $z_{i+h(2^m-1)} = 1$ **then**
            $j_i \leftarrow 1$

Of course, we can construct $\mathbf{k}$ from $\mathbf{z}$ by the same technique, i.e., by considering m bits in $\mathbf{z}$ that are $2^n - 1$ co-ordinates apart.

Let's look an example. Suppose that $m = 6$ and $n = 5$, so $\mathbf{j}$ has period 63 and $\mathbf{k}$ has period 31. Suppose we have a portion of $\mathbf{z}$. Then we can compute any j_i from

$$z_i, z_{i+63}, z_{i+126}, z_{i+189}, z_{i+252};$$

and we can compute any k_i from

$$z_i, z_{i+31}, z_{i+62}, z_{i+93}, z_{i+124}, z_{i+155}.$$

Here is the beginning of a keystream $\mathbf{z}$ produced by this method:

00000000000100001101010001000001001011001101100010
00100101000000100000100010000101100000000000000001
00111010001101010101000000000000011000000000111010
00000000001000011100000100100001000000000100001100

Using the algorithm described above, we would obtain $\mathbf{j}$ to be

$$00100001100\ldots.$$

We leave the determination of $\mathbf{k}$ as an exercise for the reader.

We now discuss the reason why there must be at least one occurrence of a 1 among the bits

$$k_i, k_{i+2^m-1}, k_{i+2(2^m-1)}, \ldots, k_{i+(n-1)(2^m-1)}.$$

(Warning: this discussion assumes some knowledge of finite field theory.)

The sequence $k_1, k_2, \ldots$ is generated by a linear recurrence relation of degree n

$$k_{n+i} = \sum_{h=0}^{n-1} c_h k_{h+i},$$

where $c_h = 0, 1$ for $0 \leq h \leq n-1$. The fact that the recurrence has maximum period $2^n - 1$ is equivalent to saying that the polynomial

$$f(x) = x^n + \sum_{h=0}^{n-1} c_h x^h$$

is irreducible in $\mathbb{Z}_2[x]$, and that x is a primitive element in the field $GF(2^n) = \mathbb{Z}_2[x]/(f(x))$.

Now, let d be a positive integer, and consider the subsequence

$$k_i, k_{i+d}, k_{i+2d}, \ldots.$$

Let the minimal polynomial of x^d be

$$m_{x^d}(x) = x^a + \sum_{h=0}^{a-1} d_h x^h,$$

where $c_h = 0, 1$ for $0 \leq h \leq n-1$. (The *minimal polynomial* of an element y is the polynomial $m_y(x) \in \mathbb{Z}_2[x]$ of smallest degree for which y is a root.)

Then it can be shown that this subsequence can be generated by the linear recurrence relation

$$k_{ad+i} = \sum_{h=0}^{a-1} d_h k_{hd+i}.$$

Since the minimal polynomial of any element of $GF(2^n)$ has degree at most n, it follows that $a \leq n$, and this recurrence has degree at most n.

As an example, consider the recurrence $x_{i+4} = x_{i+1} + x_i$. The polynomial $x^4 + x + 1$ is irreducible, and x is a primitive element of $GF(16)$. Suppose we take $d = 3$. It can be shown that the minimal polynomial of x^3 is $x^4 + x^3 + x^2 + x + 1$. So the recurrence relation

$$k_{i+12} = k_{i+9} + k_{i+6} + k_{i+3} + k_i$$

can be used to generate a subsequence $k_i, k_{i+3}, k_{i+6}, \ldots$, given initial values k_i, k_{i+3}, k_{i+6} and k_{i+9}.

Now, we are interested in the subsequence

$$k_i, k_{i+2^m-1}, k_{i+2(2^m-1)}, \ldots,$$

i.e., we have $d = 2^m - 1$. We observed above that this sequence contains some bits equal to 1. On the other hand, we have just stated that this sequence is generated by a linear recurrence relation of degree at most n. This means that the sequence cannot begin with n consecutive zeroes, which is exactly the result we wanted to establish.