MPRI 2.13.2 - Error-correcting codes and applications to cryptography

Exercises on stream ciphers - Solutions

Exercise 1 (Irregulary clocked generators).
Let $(c_t)_{t\geq 0}$ and $(z_t)_{t\geq 0}$ denote the output of Device C and of the keystream. Since A has a linear transition function L, its output can be written as $\ell \circ L(x_0)$ where ℓ is the linear extraction function and x_0 the initial state of Device A. With this notation, we have

$$z_t = \ell \circ L^{\varphi(t)}(x_0) \text{ where } \varphi(t) = \sum_{i=1}^{t}(1+c_i) = t + \sum_{i=1}^{t} c_i \ .$$

The following attack then applies :
For each initial state of Device C
Generate $(n_A + d)$ output bits of Device C.
For each t from 0 to $(n_A + d - 1)$, compute $\varphi(t) = t + \sum_{i=1}^{t} c_i$.
For each t from 0 to $(n_A + \varepsilon - 1)$, write the linear equation

$$\ell \circ L^{\varphi(t)}(x_0) = z_t$$

Solve the obtained linear system.
If x_0 is a solution of the system, generate from x_0 the following $(d - \varepsilon)$ keystream bits and check whether they correspond to the known keystream.
The time complexity is then $\mathcal{O}\left(2^{n_C} n_A^3\right)$, while the data complexity is $\mathcal{O}(n_A + d)$. The final test to check whether the recovered value of x_0 is correct has negligible time complexity. It requires the knowledge of d additional keystream bits (besides the $(n_A + \varepsilon)$ needed for solving the system), where d is a small integer depending on the false alarm probability.

Exercise 2 (Statistical test)

1. The sequence has length 100 and contains 34 ones and 66 zeroes. The frequency test then considers

$$x = \frac{(66 - 34)}{\sqrt{100}} = 3.2 \ .$$

The probability that a random sequence satisfies $x \geq 3.2$ is

$$\text{erfc}\left(\frac{3.2}{\sqrt{2}}\right) = 1.3 \cdot 10^{-3} \ .$$

We then conclude that, with a high probability, this sequence is not a truly random sequence.

2. Since the Hamming weight of the 3-variable Boolean function f is 3, $\Pr[z_t = 1] = \frac{3}{8}$. It follows that the keystream is biased and does not pass the frequency test.

3. It can be observed that

$$f(x_1, x_2, x_3) = x_2(1 + x_3 + x_1 x_3) \ .$$

Then, $f(x_1, x_2, x_3) = 1$ implies that $x_2 = 1$. In other words, each time the keystream bit $z_t = 1$, it can be deduced that $s_{t+8} = 1$. The knowledge of 32 keystream bits then provides roughly 12 values of t such that $s_{t+8} = 1$, i.e., roughly 12 bits of the LFSR sequence. The initial state of the LFSR can then be recovered by solving the corresponding linear system (i.e., by expressing each s_{t+8} as a linear function of the LFSR initial state).

Exercise 3 (OFB mode of operation). We focus on known-plaintext attacks, implying that the keystream sequence composed of all $z_t = m_t \oplus c_t$ is supposed to be known to the attacker.

1. A chosen-IV attack can be mounted from the knowledge of the first two words (z_0, z_1) in the keystream sequence generated from IV, and from the first word z'_0 in the keystream generated from $IV' = z_0$. Indeed, for the OFB mode, $z'_0 = z_1$ holds with probability 1 (while it holds with probability 2^{-n} for random sequences, where n is the block size of the involved block cipher).

2. In the random IV-setting, it can be observed that the keystream produced from IV corresponds to a cycle of the permutation $E_K : (z_t)_{t\geq 0} = (E_K^t(IV))_{t\geq 0}$. It follows that, within a keystream of length N, either all z_t are different, or the keystream is a periodic sequence of period smaller than N. This behavior highly differs from the behavior of a random sequence for $N \simeq 2^{n/2}$: indeed, the expected number of collisions in such a random sequence is close to 1. It is worth noticing that, in the OFB mode, the existence of weak IV for which the sequence $(z_t)_{t\geq 0}$ has period smaller than $2^{n/2}$ cannot be avoided, while this phenomenon does not occur for the CTR mode.

Exercise 4 (100 prisoners). The strategy is the following. Each prisoner opens the drawer numbered by his own number. This drawer contains a card, and the prisoner then opens the drawer numbered by the card he just found, and he makes the same until he has opened 50 drawers. The process which assigns to the 100 drawers the 100 cards is a permutation chosen uniformly at random from the set of all permutations of $\{1, \ldots, 100\}$. Then, each prisoner is following a cycle of this permutation beginning with his number. All the 100 prisoners locate their own numbers if the permutation does not have any cycle of length strictly higher than 50.
The number of permutations of N elements having a cycle $\mathcal{C}$ of length exactly $k > \frac{N}{2}$ is $(N)!/k$. Indeed, there are $\binom{N}{k}$ ways to pick the entries in $\mathcal{C}$, $(k-1)!$ to order them, and $(N-k)!$ ways to permute the other elements. Then, the total number of such permutations is :

$$\frac{N!(k-1)!(N-k)!}{k!(N-k)!} = \frac{N!}{k} .$$

Obviously, at most one cycle of length strictly greater than $N/2$ can exist in a given permutation. Then, the total number of permutations having a cycle of length strictly greater than $N/2$ is

$$N!\left(\frac{1}{\frac{N}{2}+1} + \frac{1}{\frac{N}{2}+2} + \ldots + \frac{1}{N}\right) = N!(H_N - H_{N/2}) ,$$

where H_n is the Harmonic number, i.e.,

$$H_n = \sum_{i=1}^{n} \frac{1}{i} \sim \ln n + \gamma .$$

It follows that the probability that a randomly chosen permutation has no cycle of length greater than $N/2$ is

$$1 - H_N + H_{N/2} \simeq 1 - \ln(N) + \ln(N/2) = 1 - \ln(2) \simeq 0.307 .$$

The 100 prisoners will then survive with probability 0.307 (this is an approximation and the exact value for $N = 100$ is a bit higher : 0.312).