Homework Assignment # 2

Due Date: Wednesday, September 20

For some of the problems in this assignment a program is provided which performs inversion of binary matrices of arbitrary size modulo 2. It is contained in the file "mx_inv_2.c"which you can get from the course web page. You don't have to deal with the source code at all — just compile it and run the executable file. The program will prompt you for the matrix size and the binary entries. There is, however, a brief description in the head of the source code file.

1. What is the pseudorandom sequence generated by the linear shift feedback register (LFSR) characterized by $(c_2 = 1, c_1 = 0, c_0 = 1)$ starting with the initialization vector $(z_2 = 1, z_1 = 0, z_0 = 0)$. What is the sequence generated from the initialization vector $(z_2 = 0, z_1 = 1, z_0 = 1)$. How are the two sequences related?

2. In this problem we will study LFSRs in somewhat more detail. LFSRs come in three flavors

   - LFSRs which generate a maximum-length sequence. These LFSRs are based on *primitive polynomials*.
   - LFSRs which do not generate a maximum-length sequence but whose sequence length is independent of the initial value of the register. These LFSRs are based on *irreducible polynomials* which are not primitive. Note that all primitive polynomials are also irreducible.
   - LFSRs which do not generate a maximum-length sequence and whose sequence length depends on the initial values of the register. These LFSRs are based on *reducible polynomials*.

   We will study examples in the following. Determine *all* sequences generated by

   (a) $x^4 + x + 1$

   (b) $x^4 + x^2 + 1$

   (c) $x^4 + x^3 + x^2 + x + 1$

   Draw the corresponding LFSR for each of the three polynomials. Which of the polynomials is primitive, which is only irreducible, and which one is reducible? Note that the lengths of all sequences generated by a given LFSR should add up to $2^m - 1$ (excluding the all zero sequence.)

   Some additional remarks: Keep in mind that in cryptography one is mainly interested in primitive polynomials and in the following we will only consider those. There are relatively fast primitivity tests available which work as long as the prime factorization

of $2^m - 1$ is known. These algorithms are, however, beyond the scope of this course. Such an algorithm can be found, e.g., in the Handbook of Applied Cryptography. A comprehensive treatment of LFSRs is given in[1] which is available in the Gordon Library.

3. We conduct a known-plain text attack on an LFSR-based stream cipher. We know that the plain text sent was:

   `1001 0010 0110 1101 1001 0010 0110`

   By tapping the channel we observe the following stream:

   `1011 1100 0011 0001 0010 1011 0001`

   (a) What is the degree m of the stream generator?

   (b) What is the initialization vector?

   (c) Determine the feedback coefficients of the LFSR.

   (d) Draw a circuit diagram and verify the output sequence of the LFSR.

4. We want to perform an attack on another LFSR-based stream cipher. In order to process letters, each of the 26 uppercase letters and the numbers 0,1,2,3,4,5 are represented by a five bit vector according to the following mapping:

$$\begin{aligned} A \leftrightarrow 0 &= 00000_2 \\ &\vdots \\ Z \leftrightarrow 25 &= 11001_2 \\ 0 \leftrightarrow 26 &= 11010_2 \\ &\vdots \\ 5 \leftrightarrow 31 &= 11111_2 \end{aligned}$$

   We happen to know the following facts about the system: (1) the degree of the LFSR is $m = 6$; (2) every message starts with the header `WPI`.

   We observe now on the channel the following message (the 4th letter is a zero):

   `J5A0EDJ2B`

   (a) What is the initialization vector?

   (b) What are the feedback coefficients of the LFSR?

[1] Solomon W. Golomb: Shift register sequences. Aegean Park Press

(c) Write a program in C which generates the whole sequence and find the whole plaintext.

(d) Where does the thing after `WPI` live? (1 extra point)

(e) What type of attack did we perform?

5. One important condition for a secure stream cipher is that the period is large. In this problem we look into the consequences of this condition for an application of realistic size.

   The scenario is as follows: We want to encrypt a network link which is part of an ATM (asynchronous transfer mode) network. Data on the link is transmitted at full ATM speed, i.e., at 155 Mbits/sec. Note that 1Mbit=2^{20}bit. As the encryption method of choice, a standard LFSR-based stream cipher is used. What is the minimum degree of the stream cipher in order to ensure that there is no repetition in the key stream within 24 hours? (Of course, single LFSRs are incredible insecure, and we shouldn't use them in actual systems.)

6. One-time pads can easily be generalized to work in alphabets other than the binary. For manual encryption, an especially useful one is a OTP that operates on letters.

   (a) Develop a OTP system which operates with the letters A,B,...,Z, represented by the numbers 0,1,...,25. How does the key (stream) look? What are the encryption and decryption functions?

   (b) Decrypt the following cipher text:

   `BSASPP KKUOSR`

   which was encrypted using the one-time pad:

   `RSIDPY DKAWOA`

   (c) How was the young man murdered? (1 extra point)

7. Assume a one-time pad cipher in which a short key, let's say of only 1000 bit length, is used repeatedly. With which attacks can the cipher be broken efficiently? (It is reported that the Rosenbergs used such a cipher for their communication with the Soviets. As we all know by now, not too successfully.)

8. Assume we store a one-time key on a CD-ROM with 1 Gbyte capacity. Discuss the *real-life* implications of such a system. Address issues such as life cycle of the key, storage of the key during the life cycle/after life cycle, key distribution, generation of key, etc.

9. On the first glance it seems as though a brute force attack against a one-time pad is feasible. This is a paradox since we know that the OTP is unconditionally secure. Describe why a brute force attack does not work. Remark: You have to resolve the paradox! That means answers such as "The OTP is unconditionally secure and therefore a brute force attack does not work" is not a valid one and will get no points.

10. In this problem we will gain a little experience with more advanced stream ciphers that are built from individual LFSRs. Such ciphers *can* be secure.

    We consider the alternating stop-and-go generator introduced in the lecture (see also Fig. 16.10, p. 385, in Schneier's book.) The three LFSRs are specified by the following polynomials and initial vectors:

    LFSR-1 $x^2 + x + 1$; $(z_0 = 1, 0)$

    LFSR-2 $x^3 + x + 1$; $(z_0 = 1, 0, 0)$

    LFSR-3 $x^5 + x^2 + 1$; $(z_0 = 1, 0, 0, 0, 0)$

    (a) Draw the circuit diagram of the stream cipher.

    (b) Compute the first eight output bits.

    (c) It generally holds for stream ciphers build from LFSRs that the sequence length is the product of the sequence lengths of the individual LFSRs if the individual lengths are all relative prime. Is this condition fulfilled for the generator above? What is the length of the sequence generated?