EE 720: An Introduction to Number Theory and Cryptography (Spring 2019)
Instructor: Saravanan Vijayakumaran
Indian Institute of Technology Bombay

Midsem Exam : 25 points — February 23, 2019

1. (5 points) Suppose we try to define perfect secrecy for the encryption of two messages *using the same key* in the following manner. Let the message space be $\mathcal{M}$. Let M_1, M_2 be the random variables denoting the first and the second message, respectively. Given a pair of messages $(m_1, m_2) \in \mathcal{M} \times \mathcal{M}$ and key $k \in \mathcal{K}$, the ciphertext is $(c_1, c_2) \in \mathcal{C} \times \mathcal{C}$ where $c_1 \leftarrow \texttt{Enc}_k(m_1)$ and $c_2 \leftarrow \texttt{Enc}_k(m_2)$. Here $\mathcal{C}$ is the ciphertext space. Let C_1, C_2 be the random variables denoting the first and second ciphertexts, respectively.

   We say that $\Pi = (\texttt{Gen}, \texttt{Enc}, \texttt{Dec})$ is **perfectly secret for two messages** if for all distributions over $\mathcal{M} \times \mathcal{M}$, all $m_1, m_2 \in \mathcal{M}$, and all ciphertexts $c_1, c_2 \in \mathcal{C}$ with $\Pr[C_1 = c_1 \cap C_2 = c_2] > 0$, we have

   $$\Pr[M_1 = m_1 \cap M_2 = m_2 \mid C_1 = c_1 \cap C_2 = c_2] = \Pr[M_1 = m_1 \cap M_2 = m_2].$$

   Prove that **no encryption scheme** can satisfy this definition. Note that $\mathcal{K}$ can be larger than $\mathcal{M} \times \mathcal{M}$. Also note that Enc can be a probabilistic algorithm but Dec is a deterministic algorithm.

2. (5 points) Consider a linear feedback shift register (LFSR) which has n registers. Let the initial state of the LFSR be $s = (s_1, s_2, \ldots, s_n)$ where each $s_i \in \{0,1\}$. Let the feedback equation be given by

   $$s_{j+n+1} = \bigoplus_{i=1}^{n} a_i s_{j+i}$$

   where $a_i \in \{0,1\}$ and $j \geq 0$. Let $G : \{0,1\}^n \mapsto \{0,1\}^m$ be the output of the LFSR when restricted to m bits where $m > n$. So $G(s) = (s_1, s_2, \ldots, s_m)$.

   Prove that G is **not a pseudorandom generator** irrespective of how the values of a_i are chosen.

3. (5 points) Let F be a length-preserving pseudorandom permutation having key length, input length, and output length all equal to n bits. Suppose a fixed-length private key encryption scheme $\Pi = (\texttt{Gen}, \texttt{Enc}, \texttt{Dec})$ is defined as follows:

   - Gen: Key k is chosen uniformly from $\{0,1\}^n$.
   - Enc: The message space $\mathcal{M} = \{0,1\}^{n/2}$. A string r is chosen uniformly from $\{0,1\}^{n/2}$ and the ciphertext $c \in \{0,1\}^n$ corresponding to $m \in \{0,1\}^{n/2}$ is given by

     $$c := F_k(r \| m).$$

     Here $\|$ is the string concatenation operator.
   - Dec: Given key k and ciphertext $c \in \{0,1\}^n$, the message m is obtained by taking the last $n/2$ bits of $F_k^{-1}(c)$.

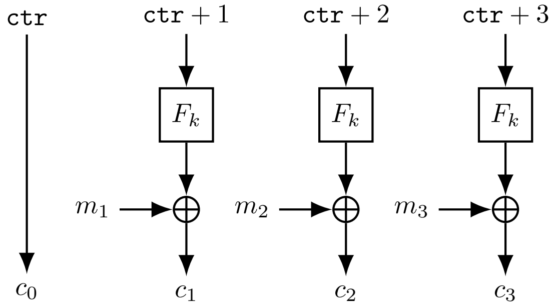
   Prove that Π is CPA-secure for messages of length $n/2$.

4. (5 points) Consider the basic CBC-MAC construction where the receiver only accepts messages of length $3n$ bits for authentication. So given a message-tag pair (m, t) the receiver will output $\texttt{Vrfy}_k(m, t) = 0$ if the length of the message m is not $3n$ bits. If $|m| = 3n$, then the receiver calculates $t' = \texttt{CBC-MAC}_k(m)$ and outputs $\texttt{Vrfy}_k(m, t) = 1$ if $t' = t$.

   Suppose the sender authenticates messages of lengths n, $2n$, or $3n$. Show that a adversary who can query the sender for tags of some messages in a query set $\mathcal{Q}$ **can forge a valid tag** on a new message. By new message, we mean a message which is not in the query set $\mathcal{Q}$. Note that this new message should have length $3n$ bits (otherwise the receiver will reject the tag).

5. Recall that the PKCS #5 padding scheme is used to pad a message x having length some integral number of bytes into a *encoded data* m having length jL bytes where L is the block length in bytes. The number of bytes which are appended to x to get m is b where $1 \leq b \leq L$. Each of these padding bytes is equal to the byte representation of the integer b. Assume that $L < 256$.

   Suppose the encoded data m has length $3L$ bytes, i.e. $m = (m_1, m_2, m_3)$ where $|m_i| = L$ bytes for $i = 1, 2, 3$. Now suppose the encoded data is encrypted using CTR mode where F is a length-preserving pseudorandom function as shown below. The input and output lengths of F_k are both equal to $n = 8L$ bits. Here the value ctr is uniformly chosen from $\{0,1\}^n$.



Suppose an adversary has access to a padding oracle. On input some ciphertext block $c' = (c'_0, c'_1, c'_2, c'_3)$, the padding oracle only returns a message from the set $\{\texttt{ok}, \texttt{padding\_error}\}$. The `ok` is returned when there is no padding error in the encoded data m' obtained from c'.

(a) (1 point) Describe a procedure by which the adversary can recover the **length** b of the padding in the encoded data m.

(b) (1 point) Describe a procedure by which the adversary can recover the **first** byte in the encoded data block m_2. By first byte, we mean the most significant byte. For example, if $L = 3$ and $m_2 =$ `0x01 0x07 0x20`, then `0x01` is the first byte of m_2.

(c) (1 point) Describe a procedure by which the adversary can recover the **last** byte in the encoded data block m_2. In the above example, `0x20` is the last byte of m_2.

(d) (2 points) What is the **maximum number** of padding oracle queries required by the adversary to recover all the bytes in the encoded data m? By all the bytes, we mean all the bytes of m_1, all the bytes of m_2, and all the bytes of m_3. Describe the procedure used by the adversary which results in this maximum number.