YALE UNIVERSITY
DEPARTMENT OF COMPUTER SCIENCE

CPSC 467b: Cryptography and Computer Security *Handout #10*

*Xueyuan Su* *March 1, 2010*

# Solution to Midterm Examination

## Instructions:

This is a closed book examination. *Answer any 5 of the following 6 questions.* Write the numbers of the **five** questions that you want graded on the cover of your bluebook. All questions count equally. You have 75 minutes. Remember to write your name on your bluebook and to justify your answers. Good Luck!

### Problem 1:

Consider a symmetric cryptosystem with $\mathcal{M} = \mathcal{C} = \{0, 1, 2, 3\}$, where the joint probability distribution over message-ciphertext pairs is described by the table below.

| | | $c$ | | | |
|---|---|---|---|---|---|
| | | 0 | 1 | 2 | 3 |
| $m$ | 0 | .025 | .050 | .075 | .100 |
| | 1 | .050 | .075 | .100 | .025 |
| | 2 | .075 | .100 | .025 | .050 |
| | 3 | .100 | .025 | .050 | .075 |

(a) What is the initial distribution on the message space?

(b) What is the most likely message $m_3$ given that $c = 3$, and what is the conditional probability $P[m = m_3 \mid c = 3]$?

(c) What can you say about the security of this cryptosystem? I.e., what does Eve learn about $m$ when she sees $c$?

(d) Describe a key space $\mathcal{K}$ and an encryption function $c = E_k(m)$ that gives rise to the given joint probability distribution when keys are chosen uniformly from $\mathcal{K}$.

### Solution:

(a) By total probability theorem, for any $0 \leq i \leq 3$, we have

$$P(m_i) = \sum_j P(m = m_i \wedge c = c_j) = .025 + .050 + .075 + .100 = .25$$

(b) Given $c = 3$, the most likely plaintext is $m_3 = 0$. The conditional probability

$$P[m = 0 \mid c = 3] = \frac{P[m = 0 \wedge c = 3]}{P[c = 3]} = \frac{.100}{.25} = .400$$

(c) When Eve sees $c$, she learns information about the distribution of the plaintext $m$ in the form of an *a posteriori* distribution over $\mathcal{M}$. Namely,

$$P[m \mid c] = \begin{cases} .100 & \text{if} \quad m + c \equiv 0 \pmod 4 \\ .200 & \text{if} \quad m + c \equiv 1 \pmod 4 \\ .300 & \text{if} \quad m + c \equiv 2 \pmod 4 \\ .400 & \text{if} \quad m + c \equiv 3 \pmod 4 \end{cases}$$

(d) There are many possible answers. For example, let $\mathcal{K} = \{0, 1, 2, 3, 5, 6, 7, 10, 11, 15\}$ and $E_k(m) = (k - m) \bmod 4$.

## Problem 2: Security of a Symmetric Cryptosystem

Happy Hacker decides to improve on the Caesar cipher by making the key space larger. He keeps $\mathcal{M} = \mathcal{C} = \{0, \ldots, 25\}$ as before, but he lets $\mathcal{K} = \{0, \ldots, 49\}$ and defines

$$E_k(m) = (k + m) \bmod 26.$$

Compare the security of Hacker's system with the ordinary Caesar cipher. Justify your answers.

**Solution:** Hacker's system is *less secure* than ordinary Caesar, despite the larger key space. Because $E_k(m) = E_{k \bmod 26}(m)$ for any $k$, the larger key space does not increase the number of possible encryption functions, but it does alter their distribution. Now, $E_0(), \ldots, E_{23}()$ are twice as likely to be chosen as $E_{24}()$ and $E_{25}()$, so $P[m \mid c]$ is not uniformly distributed, and Eve gets partial information about $m$ from seeing $c$.

## Problem 3: Group Property

Let $\mathcal{M} = \mathcal{C} = \{0, \ldots, 25\}$ as in the Caesar cipher. Let the key space $\mathcal{K}$ be an arbitrary finite set, let $f : \mathcal{K} \to \mathcal{C}$ be an arbitrary function, and let $E_k(m) = (f(k) + m) \bmod 26$.

For each of the following questions, answer whether it holds for *all* $\mathcal{K}$ and $f()$, for *some* $\mathcal{K}$ and $f()$, or for *no* $\mathcal{K}$ and $f()$.

(a) Is there a decryption function $D_k(c)$ such that $D_k(E_k(m)) = m$ for all $m \in \mathcal{M}$? If so, define it. If not, explain why not.

(b) Does the resulting symmetric cryptosystem have the group property? Why or why not?

Justify your answers.

**Solution:**

(a) It is true for all $\mathcal{K}$ and $f()$. The decryption function is

$$D_k(c) = (c - f(k)) \bmod 26$$

(b) It is true for some but not all $(\mathcal{K}, f()$ pairs.

**True for some pairs** If $\mathcal{K} = \mathcal{M}$ and $f()$ is the identity function, then the resulting cryptosystem is the ordinary Caeser cipher, which is known to have the group property.

**False for some pairs** If $f(k) = 1$ for all $k \in \mathcal{K}$, then the resulting cryptosystem does not have the group property. Choose $k_1 = k_2 \in \mathcal{K}$. Then $E_{k_1,k_2}(m) = E_{k_2}(E_{k_1}(m)) = (m + 2) \bmod 26$. Since $E_{k_3}(m) = (m + 1) \bmod 26$ for all $k_3 \in \mathcal{K}$, there is no $k_3$ for which $E_{k_3} = E_{k_1,k_2}$; hence, the group property does not hold.

## Problem 4: Feistel Network

DES is built from a Feistel network.

(a) Describe how a Feistel network can be used to build a symmetric cryptosystem from an arbitrary "scrambling" function $f()$ of appropriate type.

(b) Describe how encryption and decryption works for a cryptosystem built from a Feistel network.

**Solution:**

(a) A Feistel network consists of some number $t$ of stages. Each stage $i$ uses $f()$ and a subkey $K_i$ to map a pair of $n$-bit words $(L_i, R_i)$ to a new pair $(L_{i+1}, R_{i+1})$. By applying the stages in sequence, a $t$-stage network maps $(L_0, R_0)$ to $(L_t, R_t)$. $(L_0, R_0)$ is the plaintext, and $(L_t, R_t)$ is the corresponding ciphertext.

(b) To encrypt a message $(L_0, R_0)$, each stage works as follows:

$$\begin{aligned} L_{i+1} &= R_i \\ R_{i+1} &= L_i \oplus f(R_i, K_i) \end{aligned}$$

To decrypt a ciphertext $(L_t, R_t)$, each stage works as follows:

$$\begin{aligned} R_i &= L_{i+1} \\ L_i &= R_{i+1} \oplus f(R_i, K_i) \;=\; R_{i+1} \oplus f(L_{i+1}, K_i) \end{aligned}$$

## Problem 5: Message Authentication Codes (MAC)

(a) Define what a Message Authentication Code (MAC) system is and the properties it should have.

(b) Describe a practical use for a MAC.

(c) Describe how to build a MAC system from a given symmetric cryptosystem.

**Solution:**

(a) A MAC system is a cryptographic system that generates a short piece of information used to authenticate a message. A MAC is generated by a function $C_k(m)$ that can be computed by anyone knowing the secret key $k$. It should be hard for an attacker, without knowing $k$ to find any pair $(m, \xi)$ such that $\xi = C_k(m)$. A stronger property requires that the above is true even if the attacker knows a set of valid MAC pairs $\{(m_1, \xi_1), \ldots, (m_t, \xi_t)\}$ as long as $m$ is not a message in this set.

(b) A MAC can be used to authenticate the sender of a message. Given a symmetric cryptosystem, Alice computes the MAC $\xi = C_k(m)$ and sends $(m, \xi)$ to Bob. Upon receipt of the pair, Bob checks whether $\xi = C_k(m)$. If so, the authentication succeeds; otherwise, the authentication fails. By the properties of a MAC, an attacker without knowing $k$ is unlikely to be able to produce a valid pair $(m, \xi)$ that will be accepted by Bob.

(c) A MAC system can be built by using a block cipher in CBC or CFB chaining mode. Let $E_k$ be the resulting cryptosystem, defined for sequences of message blocks $m_1 \dots m_t$, and suppose $c = c_1 \dots c_t = E_k(m_1 \dots m_t)$. We define the MAC $C_k(m) = c_t$. In CBC or CFB mode, each ciphertext block $c_k$ depends on both the current message block $m_k$ and on the previous ciphertext block $c_{k-1}$. Hence, the last ciphertext block $c_t$ depends on the entire message $m_1 \dots m_t$, as desired.

## Problem 6: Congruence Equations

(a) Describe a necessary and sufficient condition for the congruence equation $ax \equiv b \pmod{n}$ to have a solution $x \in \mathbf{Z}_n$.

(b) Solve the congruence equation $45x \equiv 49 \pmod{77}$.

**Solution:**

(a) The congruence equation $ax \equiv b \pmod{n}$ has a solution $x \in \mathbf{Z}_n$ iff $n \,|\, (ax - b)$ for some $x \in \mathbf{Z}$ iff

$$ax + ny = b$$

for some $x, y \in \mathbf{Z}$. This is a linear Diophantine equation. It has a solution iff $\gcd(a, n) \mid b$.

(b) To solve

$$45x \equiv 49 \pmod{77}, \tag{1}$$

we solve the corresponding Diophantine equation

$$45x + 77y = 49. \tag{2}$$

Since $\gcd(45, 77) = 1 \mid 49$, this Diophantine equation has a solution. We use the extended Euclidean algorithm to solve it.

| $i$ | $r_i$ | $u_i$ | $v_i$ | $q_i$ |
|---|---|---|---|---|
| 1 | 45 | 1 | 0 | |
| 2 | 77 | 0 | 1 | 0 |
| 3 | 45 | 1 | 0 | 1 |
| 4 | 32 | -1 | 1 | 1 |
| 5 | 13 | 2 | -1 | 2 |
| 6 | 6 | -5 | 3 | 2 |
| 7 | 1 | 12 | -7 | |

Therefore, $x = 12$ and $y = -7$ satisfies

$$45x + 77y = 1. \tag{3}$$

The solution to (2) is $x = 12 \times 49 = 588$ and $y = -7 \times 49 = -343$. The solution to (1) is $x \bmod 77 = 588 \bmod 77 = 49$.

*(end of exam)*