YALE UNIVERSITY
DEPARTMENT OF COMPUTER SCIENCE

Solution to Problem Set 2

In this problem set, we consider a variant of the Caesar cipher which we call the "Happy" cipher (named after the venerable "Happy Hacker" of CPSC 223 fame). Happy (E, D) is defined as follows: Let $X_1 = \{0, \ldots, 12\}$ and $X_2 = \{13, \ldots, 25\}$. Let $\mathcal{M} = \mathcal{C} = \mathcal{K} = X = X_1 \cup X_2$, and let $n = |X| = 26$. Define

$$E_k(m) = \begin{cases} (m+k) \bmod 13 & \text{if } k \in X_1 \wedge m \in X_1 \\ m & \text{if } k \in X_1 \wedge m \in X_2 \\ m & \text{if } k \in X_2 \wedge m \in X_1 \\ ((m+k) \bmod 13) + 13 & \text{if } k \in X_2 \wedge m \in X_2 \end{cases}$$

We also consider Double Happy (E^2, D^2). Here, $\mathcal{K}^2 = \mathcal{K} \times \mathcal{K}$, and $E^2_{(k_1,k_2)} = E_{k_2}(E_{k_1}(m))$.

Problem 1: Happy Encryption (5 points)

Encrypt the plaintext "i am a secret message" using Happy with key $k = 3$. (As usual, we will ignore spaces.)

Solution:

m = "i am a secret message";
k = 3;
c = "l dc d shfrht chssdjh".

Problem 2: Happy Decryption (5 points)

Describe the Happy decryption function $D_k(c)$.

Solution:

$$D_k(c) = \begin{cases} (c-k) \; mod \; 13 & \text{if } k \in X_1 \wedge c \in X_1 \\ c & \text{if } k \in X_1 \wedge c \in X_2 \\ c & \text{if } k \in X_2 \wedge c \in X_1 \\ ((c-k) \; mod \; 13) + 13 & \text{if } k \in X_2 \wedge c \in X_2 \end{cases} \quad (1)$$

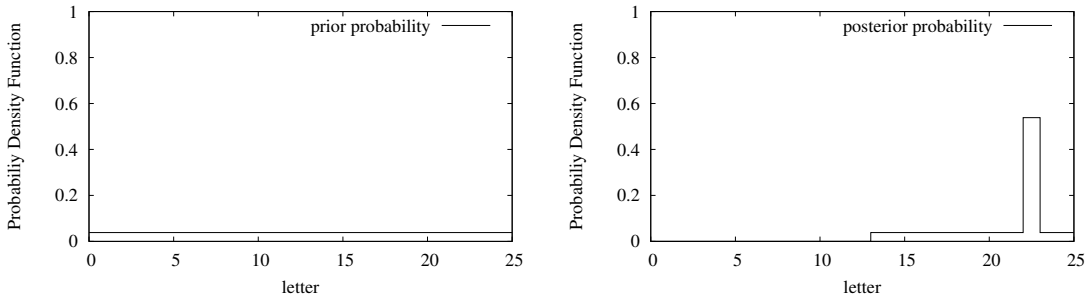
Problem 3: Security (10 points)

(a) Is Happy information-theoretically secure? Why or why not?

(b) Is Double Happy information-theoretically secure? Why or why not?

Solution:

(a) No. Before observing any cipher text, the probability of the plain text m being a specific letter m_0 is $Prob[m = m_0] = \frac{1}{26}$. After observing the cipher text, say $c = m_0$, the probability of the plain text $m = m_0$ becomes $Prob[m = m_0 \mid c = m_0] = \frac{14}{26}$. Assuming $m_0 = 23$, Figure 1



(a) Prior probability (b) Posterior probability

Figure 1: Probability distributions.

illustrates the plain text prior probability distribution and posterior probability distribution after we learn the cipher text $c = m_0 = 23$.

(b) Similar to (a).

Problem 4: Equivalent Key Pairs (10 points)

Suppose $m_0 = c_0 = 4$.

(a) Find all key pairs (k, k') such that $E^2_{(k,k')}(m_0) = c_0$.

(b) Do all such key pairs give rise to the same function $E^2_{(k,k')}$? That is, if $E^2_{(\hat{k},\hat{k'})}(m_0) = E^2_{(k,k')}(m_0) = c_0$, does $E^2_{(\hat{k},\hat{k'})}(m) = E^2_{(k,k')}(m)$ for all $m \in \mathcal{M}$? Why or why not?

Solution:

(a) Note that $m_0, k_0 \in X_1$. There are four different cases:

- 1) $\{(k, k^{'}) \mid k, k^{'} \in X_2\}$
- 2) $\{(k, k^{'}) \mid k = 0, k^{'} \in X_2\}$
- 3) $\{(k, k^{'}) \mid k^{'} = 0, k \in X_2\}$
- 4) $\{(k, k^{'}) \mid k, k^{'} \in X_1, (k + k^{'}) \text{ mod} 13 = 0\}$

(b) No. For example, if $m = 13$, then $E^2_{13,14}(13) = 14$ (case 1), but $E^2_{1,12}(13) = 13$ (case 4).

Problem 5: Group Property (10 points)

Is Happy a group? Why or why not?

Solution:

No. group property requires that for any key pairs $k_1, k_2 \in X$, there is always a single key $k \in X$, such that $E_{k_1,k_2}(m) = E_k(m)$ for any $m \in X$. Taking $k_1 = 1, k_2 = 14$ for example, now we will show that $\forall k$, there $\exists m$ such that $E_{k_1,k_2}(m) \neq E_k(m)$.

- $k \in X_1$: $E_{1,14}(13) = 14$, but $E_k(13) = 13$

- $k \in X_2$: $E_{1,14}(0) = 1$, but $E_k(0) = 0$

The following problems ask you to compute probabilities. You may do so either analytically (if you're facile with combinatorial counting techniques) or experimentally by writing a program to simulate 1000 random trials and reporting the fraction of times that the desired result is obtained. Either way, you should show your work – analytic derivation, or program and simulation results.

Problem 6: Birthday Problem (20 points)

Suppose $u_1, \ldots, u_6$ and $v_1, \ldots, v_6$ are chosen uniformly and independently at random from X (so duplicates are possible. Find the probability that $\{u_1, \ldots, u_6\} \cap \{v_1, \ldots, v_6\} \neq \emptyset$. (Note that $6 = \lceil \sqrt{n} \rceil$.)

Solution:

We first calculate the probability that sets $\{u_i\}$ and $\{v_i\}$ do not have intersection and analyze this situation case by case as follows:

1) 6 members of set $\{u_i\}$ are identical, (i.e., the form 6), $\{v_i\}$ choose from the remaining 25 letters:

$$p_1 = [1 \times (\frac{1}{26})^5] \times [(\frac{25}{26})^6] \tag{2}$$

2) 5 members of set $\{u_i\}$ are identical, (i.e., the form 5:1), $\{v_i\}$ choose from the remaining 24 letters:

$$p_2 = \binom{6}{5} \times [1 \times (\frac{1}{26})^4 \times \frac{25}{26}] \times [(\frac{24}{26})^6] \tag{3}$$

3) 4 members of set $\{u_i\}$ are identical, the rest 2 members themselves are identical, (i.e., the form 4:2), $\{v_i\}$ choose from the remaining 24 letters:

$$p_3 = \binom{6}{4} \times [1 \times (\frac{1}{26})^3 \times \frac{25}{26} \times \frac{1}{26}] \times [(\frac{24}{26})^6] \tag{4}$$

4) 4 members of set $\{u_i\}$ are identical, the rest 2 members themselves are different, (i.e., the form 4:1:1), $\{v_i\}$ choose from the remaining 23 letters:

$$p_4 = \binom{6}{4} \times [1 \times (\frac{1}{26})^3 \times \frac{25}{26} \times \frac{24}{26}] \times [(\frac{23}{26})^6] \tag{5}$$

5) 3 members of set $\{u_i\}$ are identical, the rest 3 members themselves are identical, (i.e., the form 3:3), $\{v_i\}$ choose from the remaining 24 letters:

$$p_5 = \binom{6}{3} \times \frac{1}{2!} \times [1 \times (\frac{1}{26})^2 \times \frac{25}{26} \times (\frac{1}{26})^2] \times [(\frac{24}{26})^6] \tag{6}$$

6) 3 members of set $\{u_i\}$ are identical, other 2 members themselves are identical, the last one is different (i.e., the form 3:2:1), $\{v_i\}$ choose from the remaining 23 letters:

$$p_6 = \binom{6}{3} \times \binom{3}{2} \times [1 \times (\frac{1}{26})^2 \times \frac{25}{26} \times \frac{1}{26} \times \frac{24}{26}] \times [(\frac{23}{26})^6] \tag{7}$$

7) 3 members of set $\{u_i\}$ are identical, the rest 3 members themselves are all different, (i.e., the form 3:1:1:1), $\{v_i\}$ choose from the remaining 22 letters:

$$p_7 = \binom{6}{3} \times [1 \times (\frac{1}{26})^2 \times \frac{25}{26} \times \frac{24}{26} \times \frac{23}{26}] \times [(\frac{22}{26})^6] \tag{8}$$

8) $\{u_i\}$ are in the form of 2:2:2, $\{v_i\}$ choose from the remaining 23 letters:

$$p_8 = \binom{6}{2} \times \binom{4}{2} \times \frac{1}{3!} \times [1 \times \frac{1}{26} \times \frac{25}{26} \times \frac{1}{26} \times \frac{24}{26} \times \frac{1}{26}] \times [(\frac{23}{26})^6] \tag{9}$$

9) $\{u_i\}$ are in the form of 2:2:1:1, $\{v_i\}$ choose from the remaining 22 letters:

$$p_9 = \binom{6}{2} \times \binom{4}{2} \times \frac{1}{2!} \times [1 \times \frac{1}{26} \times \frac{25}{26} \times \frac{1}{26} \times \frac{24}{26} \times \frac{23}{26}] \times [(\frac{22}{26})^6] \tag{10}$$

10) $\{u_i\}$ are in the form of 2:1:1:1:1, $\{v_i\}$ choose from the remaining 21 letters:

$$p_{10} = \binom{6}{2} \times [1 \times \frac{1}{26} \times \frac{25}{26} \times \frac{24}{26} \times \frac{23}{26} \times \frac{22}{26}] \times [(\frac{21}{26})^6] \tag{11}$$

11) All members of $\{u_i\}$ are different, (i.e., the form of 1:1:1:1:1:1), $\{v_i\}$ choose from the remaining 20 letters:

$$p_{11} = [1 \times \frac{25}{26} \times \frac{24}{26} \times \frac{23}{26} \times \frac{22}{26} \times \frac{21}{26}] \times [(\frac{20}{26})^6] \tag{12}$$

Then the probability of non-empty intersection between u_i and v_i is:

$$p = 1 - \sum_{j=1}^{11} p_j \approx 0.752486 \tag{13}$$

The above calculation is pretty tedious, but it helps to understand the detailed analysis. We can also use Monte Carlo simulation to obtain the success probabilities.

We can make use of the randRange() function we constructed in PS1 to generate a 6 random numbers u_i and another 6 random numbers v_i. Then we check each pair of (u_i, v_i) one by one to see whether there is a match. If there is such match, the counter increases by 1. In each experiment we run $1,000,000$ trials and calculates the non-empty intersection probability. Then we run the experiments for 100 times and calculate the probability. The result is shown in Figure 2. The average probability is about 75.25%.

Problem 7: Birthday Attack on Double Happy (40 points)

Assume Alice chooses a random key pair (k_0, k_0') and a random message m and computes $c = E^2_{(k_0,k_0')}(m)$ using Double Happy. Eve learns the plaintext-ciphertext pair (m, c) and then carries out the *Birthday Attack* for $m \in \mathcal{M}$ and $c \in \mathcal{C}$ as follows:

- She chooses $k_1, \ldots, k_6$ uniformly at random from $\mathcal{K}$ and computes $u_i = E_{k_i}(m)$ for $i = 1, \ldots, 6$.
- She chooses $k_1', \ldots, k_6'$ uniformly at random from $\mathcal{K}$ and computes $v_j = D_{k_j'}(c)$ for $j = 1, \ldots, 6$.

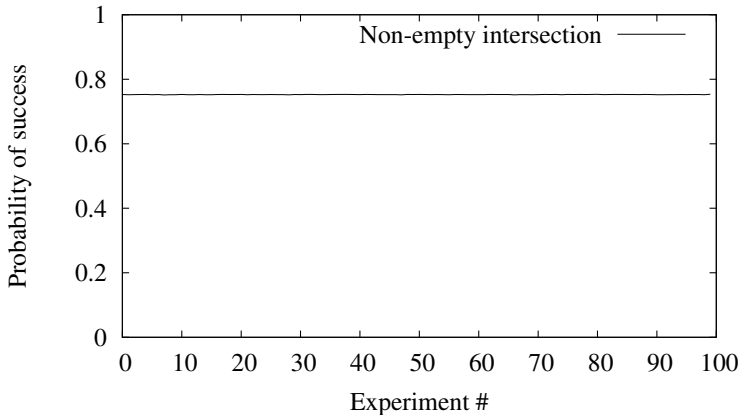


Figure 2: Probability of non-empty intersection.

- If $\{u_1, \ldots, u_6\} \cap \{v_1, \ldots, v_6\} \neq \emptyset$, we say the Birthday Attack *succeeds in producing a candidate key pair*. In that case, Eve obtains the candidate key pair $(k, k') = (k_i, k'_j)$, where (i, j) is the lexicographically smallest pair such that $u_i = v_j$.
- If a candidate key pair (k, k') is produced and (k, k') can be used to decrypt any message Alice might send using her key, that is, if $D^2_{(k,k')}(E^2_{(k_0,k'_0)}(m)) = m$ for all $m \in \mathcal{M}$, then we say the Birthday Attack *succeeds in breaking Double Happy*.

(a) Find the probability that the Birthday Attack succeeds in producing a candidate key pair, and compare your result with your answer to problem 6.

(b) Find the probability that the Birthday Attack succeeds in breaking Double Happy.

Solution:

(a) To successfully produce a candidate key pair, it means that the Birthday Attack succeeds in finding a key pair that decrypts a known plaintext-ciphertext pair (m, c).

We still make use of the randRange() function to generate a random plain letter m and a key pair $(k_0, k_0^{'})$. We use this key pair to encrypt m and get c. Now we need to generate 6 random keys k_i and another 6 random keys $k_i^{'}$. Use k_i to encrypt m we get the set of $\{u_i\}$ and use $k_i^{'}$ to decrypt c we get the set of $\{v_i\}$. Then if there is a match between any pair of (u_i, v_i), the counter increases by 1. In each experiment we run $1,000,000$ trials and calculates the probability of succeeding in producing a candidate key pair. Then we run the experiments for 100 times and calculate the probability. The result is shown in Figure 3. The average probability is about 75.29%.

(b) To successfully break Double Happy, it means that the candidate key can decrypts all the cipertext produced from any $m \in X$ with the key pair that Alice owns. Due to the special property of Double Happy, it is suffice to show that if the candidate key pair can decrypts one plaintext-ciphertext pair $(m1, c1) \in X_1$ and another plaintext-ciphertext pair $(m2, c2) \in X_2$, it can break the entire encryption system.

Therefore, we make use of the randRange() function to generate two random plain letter $m_1 \in X_1$ and $m_2 \in X_2$, and a key pair $(k_0, k_0^{'})$. We use this key pair to encrypt m_1, m_2 and get c_1, c_2. Now we need to generate 6 random keys k_i and another 6 random keys $k_i^{'}$. Use k_i to encrypt m we get the set of $\{u_i\}$ and use $k_i^{'}$ to decrypt c we get the set of $\{v_i\}$. By finding the match between any pair of (u_i, v_i), we find the candidate key pair that can decrypts (m_1, c_1). If we also succeed in

decrypting (m_2, c_2) with the same candidate key pair, we have succeeded in producing a candidate key pair that breaks Double Happy. In each experiment we run $1,000,000$ trials and calculates the probability of succeeding in breaking Double Happy. Then we run the experiments for 100 times and calculate the probability. The result is shown in Figure 3. The average probability is about 13.53%.

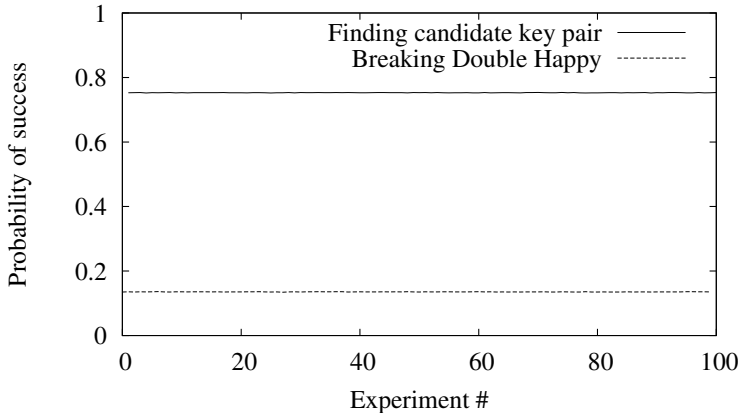


Figure 3: Birthday attack on Double Happy.