Cryptography Homework 1

陈赉(12212231)

Problem 1. The following is an encryption of English text using a shift cipher. Find the key and decrypt the ciphertext.

"O QFMDHCGMGHSA GVCIZR PS GSQIFS SJSB WT SJSFMHVWBU OPCIH HVS GMGHSA SLQSDH HVS YSM WG DIPZWQ YBCKZSRUS"

Solution. The shift cipher has a key space of barely 26. We can try all possible keys to decrypt the ciphertext. And by the observation that, the word with single letter could either be A or I, we can test these two possible result through the first letter to obtain the key. Using the tools provided by http://www.xarg.org/tools/caesar-cipher/ we can decrypt the ciphertext, and the key is 12.

"A CRYPTOSYSTEM SHOULD BE SECURE EVEN IF EVERYTHING ABOUT THE SYSTEM EXCEPT THE KEY IS PUBLIC KNOWLEDGE"

Problem 2. Prove that Definition 1.6 and Definition 1.7 on slides of Lecture 02 are equivalent.

Solution. Suppose that the plaintext space is $\mathcal{M} = \{x, x'\}$. For arbitrary key $k \leftarrow_R \{0,1\}^n$, let the cipher text $c = Enc_k(x), c' = Enc_k(x')$. Prove: Definition 1.6 $\to$ Definition 1.7

Suppose

$$\exists x, x' \in \mathcal{M},\ Enc_{U_n}(x) \not\equiv Enc_{U_n}(x')$$

which means there also exists a string y_0 such that

$$\Pr[Y_x = y_0] > \Pr[Y_{x'} = y_0]$$

Then we can construct an attacker Eve such that

$$Eve(y) = \begin{cases} x, & \text{if } y = y_0 \\ x_b \leftarrow_R \{x, x'\}, & \text{if } y \neq y_0 \end{cases}$$

Thus, Eve has chances larger than $1/2$ to obtain the correct plaintext. And this gives that

$$\neg \text{ Definition 1.7} \to \neg \text{ Definition 1.6} \equiv \text{Definition 1.6} \to \text{Definition 1.7}$$

Prove: Definition 1.7 $\to$ Definition 1.6. Suppose we have

$$Enc_{U_n}(x) \equiv Enc_{U_n}(x')$$

then Eve cannot distinguish them after seeing the ciphertext, and can only guess the plaintext with probability at most $1/2$.

Problem 3. Let n be a positive integer. The affine cipher modulo n is defined as follows. A key (a, b) consists of an element $a \in \mathbb{Z}_n^*$ and an element $b \in \mathbb{Z}_n$. For a message $m \in \mathbb{Z}_n$, the ciphertext is $C = Enc_{(a,b)}(m) = (am + b) \bmod n$. If we randomly choose a key (a, b) for each message m to be encrypted, is this affine cipher perfectly secure? Explain your answer.

Solution. Perfectly secure. Since for every $m \in \mathbb{Z}_n$ and $c \in \mathbb{Z}_n$, we have

$$\Pr\left[Enc_{(a,b)}(m) = c\right] = \frac{|\{a \in \mathbb{Z}_n^*, b \in \mathbb{Z}_n : (a \times m + b) \bmod n = c\}|}{|\mathbb{Z}_n^*| \times |\mathbb{Z}_n|} = \frac{1}{n}$$

over a random choice of (a, b), so the distrubition is equivalent. For detailed explanation, the numerator is always $n - 1$, since for any $a \in \mathbb{Z}_n^*$ and $m \in \mathbb{Z}_n$ we can consider $(a \times m) \bmod n$ as a base element and b as the offset of the group

$$\{b \in \mathbb{Z}_n : (a \times m + b) \bmod n\}$$

which is a permutation of $\mathbb{Z}_n$, and where the cipher c appears always. Thus, c appears $|\mathbb{Z}_n^*|$ times in the numerator. Therefore, affine cipher is perfectly secure.

Problem 4. Prove that ans encryption scheme (Gen, Enc, Dec) with message space $\mathcal{M}$ is perfectly secure if and only if

$$\Pr[Enc_K(m) = c] = \Pr[Enc_K(m') = c]$$

holds for every two $m, m' \in \mathcal{M}$ and every $c \in \mathcal{C}$

Solution. "If" part: Suppose we have a message $m \in \mathcal{M}$ and a ciphertext c for which $\Pr[C = c] > 0$. If $\Pr[M = m] = 0$, then trivally $\Pr[M = m \mid C = c] = \Pr[M = m] = 0$. So, considering the case $\Pr[M = m] > 0$, we have firstly

$$\Pr[C = c|M = m] = \Pr[Enc_K(M) = c \mid M = m] = \Pr[Enc_K(m) = c]$$

which denotes as δ_c. From the assumption we have for every $m' \in \mathcal{M}$

$$\Pr[Enc_K(m') = c] = \Pr[C = c|\ M = m'] = \delta_c$$

Using Bayes' theorem, we have

$$\begin{aligned}
\Pr[M = m|C = c] &= \frac{\Pr[C = c|M = m] \cdot \Pr[M = m]}{\Pr[C = c]} \\
&= \frac{\Pr[C = c|M = m] \cdot \Pr[M = m]}{\sum_{m' \in \mathcal{M}} \Pr[C = c|M = m'] \cdot \Pr[M = m']} \\
&= \frac{\delta_c \cdot \Pr[M = m]}{\sum_{m' \in \mathcal{M}} \delta_c \cdot \Pr[M = m']} \\
&= \frac{\Pr[M = m]}{\sum_{m' \in \mathcal{M}} \Pr[M = m']} = \Pr[M = m]
\end{aligned}$$

Thus, we have $\Pr[M = m|\ C = c] = \Pr[M = m]$, which means the encryption scheme is perfectly secure. For the "only if" part, suppose that we have two messages m, m' and a ciphertext c with nonzero probability. Then by Definition 1.6, we have

$$\Pr[M = m|C = c] = \Pr[M = m'|C = c] = \frac{1}{2}$$

This is equivalent to

$$\Pr[M = m|C = c] = \frac{\Pr[C = c|M = m] \cdot \Pr[M = m]}{\Pr[C = c]}$$
$$= \frac{\Pr[C = c|M = m]}{2 \cdot \Pr[C = c]}$$

which gives

$$\Pr[C = c|M = m] = \Pr[Enc_k(m) = c] = \Pr[C = c]$$

Similarly, we have $\Pr[C = c|M = m'] = \Pr[Enc_k(m') = c] = \Pr[C = c]$ which gives that the two sides are equivalent. Thus, the "only if" part is proved.

Problem 5. For an encryption scheme (Gen, Enc, Dec), consider the following game:

- Eve chooses $m_1, m_2, m_3 \in \{0, 1\}^l$.
- Alice selects $k \leftarrow_R \{0, 1\}, i \leftarrow_R \{1, 2, 3\}$ and gives Eve $c = E_{k(m_i)}$
- Eve sends a number $j \in \{1, 2, 3\}$

Eve wins if $i = j$. Prove that (Gen, Enc, Dec) is perfectly secure if and only if Eve can guess i with probability at most $\frac{1}{3}$.

Solution. "If" part: Prove by contropositive. Suppose the scheme is not perfectly secure, which means that there exists a strategy for Eve to guess m_i from $\mathcal{M}$ with probability larger than $1/|\mathcal{M}|$. Then, w.l.o.g., we assume $x_1 = 0^l$ and $x_2 \leftarrow_R \mathcal{M}$. Then, for random key k, we have

$$\Pr[Eve(Enc_k(x_2)) = x_2] > \frac{1}{|\mathcal{M}|}$$

But for every k, the decrypted message $Eve(Enc_k(x_1))$ is independent of x_1, so we have

$$\Pr[Eve(Enc_k(x_1)) = x_2] \leq \frac{1}{\mathcal{M}} < \Pr[Eve(Enc_k(x_2)) = x_2]$$

So for Eve's strategy, we have

$$Eve'(c) = \begin{cases} x_2, & \text{if } Eve(c) = x_2 \\ x_i \leftarrow_R \{m_1, m_2, m_3\}, & \text{otherwise} \end{cases}$$

which gives that Eve can guess i with probability larger than $1/3$.
"Only if" part: Suppose that the scheme is perfectly secure. Then by definition, we have

$$\Pr[M = m|C = c] = \Pr[M = m]$$

which means Eve gains no information about the plaintext and can only guess i with probability at most $1/3$.

Problem 6. Prove that the statistical distance $\Delta(X, Y)$ is a metric.

Solution. Firstly, by the definition of statistical distance, we have

$$\Delta(X, X) = \max_{T \subseteq \{0,1\}^n} |\Pr[X \in T] - \Pr[X \in T]| = 0$$

Then, for the symmetry, we have

$$\begin{aligned}\Delta(X,Y) &= \max_{T\subseteq\{0,1\}^n} |\Pr[X\in T]-\Pr[Y\in T]| \\ &= \max_{T\subseteq\{0,1\}^n} |\Pr[Y\in T]-\Pr[X\in T]| \\ &= \Delta(Y,X)\end{aligned}$$

For the transitivity, by Lemma 2.3 from , we have

$$\Delta(X,Y) = \frac{1}{2}\sum_{\omega\in Supp(X)\cup Supp(Y)} |\Pr[X=\omega]-\Pr[Y=\omega]|$$

and therefore,

$$\begin{aligned}\Delta(X,Y)+\Delta(Y,Z) &= \frac{1}{2}\sum_{\omega\in Supp(X)\cup Supp(Y)} |\Pr[X=\omega]-\Pr[Y=\omega]| \\ &+\frac{1}{2}\sum_{\omega\in Supp(Y)\cup Supp(Z)} |\Pr[Y=\omega]-\Pr[Z=\omega]| \\ &\geq \frac{1}{2}\sum_{\omega\in Supp(X)\cup Supp(Y)\cup Supp(Z)} |\Pr[X=\omega]-\Pr[Z=\omega]| \\ &= \frac{1}{2}\sum_{\omega\in Supp(X)\cup Supp(Z)} |\Pr[X=\omega]-\Pr[Z=\omega]| \\ &= \Delta(X,Z)\end{aligned}$$

Thus, the statistical distance is a metric.

Problem 7. Let $\{X_n\}, \{Y_n\}$ be sequences of distributions with X_n and Y_n ranging over $\{0,1\}^{p(n)}$ for some polynomial $p(n)$ in n. $\{X_n\}$ and $\{Y_n\}$ are computationally indistinguishable $(X_n \approx Y_n)$ if for every polynomial-time algorithm A there is a negligible function ε such that

$$|\Pr[A(X_n)=1]-\Pr[A(Y_n)=1]| \leq \varepsilon(n)$$

Prove that the computationally indistinguishable relation is an equivalence relation.

Solution. Equivalence relation contains symmetry, reflexivity and transitivity. For symmetry, trivally, we have

$$|\Pr[A(X_n)=1]-\Pr[A(Y_n)=1]| = |\Pr[A(Y_n)=1]-\Pr[A(X_n)=1]|$$

For reflexivity, by definition, we have

$$|\Pr[A(X_n)=1]-\Pr[A(X_n)=1]| = 0 \leq \varepsilon(n)$$

For transitivity, we have firstly

$$\begin{aligned}&|\Pr[A(X_n)=1]-\Pr[A(Z_n)=1]| \\ &= |\Pr[A(X_n)=1]-\Pr[A(Y_n)=1]+\Pr[A(Y_n)=1]-\Pr[A(Z_n)=1]| \\ &\leq |\Pr[A(X_n)=1]-\Pr[A(Y_n)=1]|+|\Pr[A(Y_n)=1]-\Pr[A(Z_n)=1]| \leq 2\varepsilon(n)\end{aligned}$$

Since $\varepsilon(n)$ is negligible, we have

$$|\Pr[A(X_n) = 1] - \Pr[A(Z_n) = 1]| \leq \varepsilon(n)$$

which is $X_n \approx Z_n$ by definition if $X_n \approx Y_n$ and $Y_n \approx Z_n$. Thus, the computationally indistinguishable relation is an equivalence relation.