MAS433 Cryptography: Tutorial 2
One-Time Pad, Information Theory
02.09.2010

Instructions.

1. Submission of tutorial solution is compulsory.
2. Submission deadline: 01 September 2011, 6PM
3. Please submit your solution by sending email to wuhj@ntu.edu.sg (If your solution is handwritten, you can scan and convert it into digital format, such as .pdf document.)
4. Tutorial solution will not be provided after the tutorial class.

Question 1. One-time Pad

1.1 For the bit-wise one-time pad, the encryption is performed as: $C_i = K_i \oplus P_i = (K_i + P_i) \bmod 2$. Now an encryption system operates as: $C_i = K_i + P_i$. How to attack this modified one-time pad?

1.2 Show that the above modified one-time pad encryption scheme is not perfectly secure.

Question 2. A 1000-bit key used in one-time pad is not randomly generated.

2.1 Suppose that the values of the first 5 bits are 0, and the other 995 bits are randomly generated and uniformly distributed (each bit with value 0 and 1 with probability 0.5), what is the entropy of the key? What is the risk of using this key in one-time pad?

2.2 Suppose that each bit of the key is randomly generated but with value 0 with probability 0.54. What is the entropy of the key? What is the risk of using this key in one-time pad? (Hint: show that the entropy of the key is the sum of the entropy of individual key bits if each key bit is generated independently.)

Question 3. A 128-bit key is randomly generated.

3.1 What is the entropy of the key? Given an unknown key, how many guesses are needed to guess the value of the key correctly on average, suppose that each guessed value is different from the values guessed already?

3.2 A 128-bit key takes a fixed value "010101...0101" half the time, and is one of 2^{127} possible values half the time (it takes each of those 2^{127} values with equal probability).

3.2.1 What is the entropy of the key?

3.2.2 Given n keys, how many guesses are needed to guess half of the keys correctly?

3.2.3 Given a key, how many guesses are needed to guess its value correctly on average?

Question 4. Information Theory, Entropy
Let the plaintext space $\mathbf{P} = \{p_1, p_2\}$ with $\mathbf{Pr}[P = p_1] = 1/4$, $\mathbf{Pr}[P = p_2] = 3/4$. Let $\mathbf{K} = \{k_1, k_2, k_3\}$ with $\mathbf{Pr}[K = k_1] = 1/2$, $\mathbf{Pr}[K = k_2] = \mathbf{Pr}[K = k_3] = 1/4$. The encryption is performed as follows:

$$E_{k_1}(p_1) = c_1,\ E_{k_1}(p_2) = c_2,$$
$$E_{k_2}(p_1) = c_2,\ E_{k_2}(p_2) = c_3,$$
$$E_{k_3}(p_1) = c_3,\ E_{k_3}(p_2) = c_4,$$

4.1 Compute the probabilities $\mathbf{Pr}[C = c_i]$ for $i = 1, 2, 3, 4$.

4.2 Compute the entropy of $\mathbf{P}$, $\mathbf{K}$ and $\mathbf{C}$.

4.3 Compute the conditional probabilities $\mathbf{Pr}[p_i|c_j]$ for $i = 1, 2$, $1 \leq j \leq 4$.

4.4 Compute the entropy of $\mathbf{P}$ if the ciphertext is given as c_i $(1 \leq i \leq 4)$. Are these results different from the entropy of $\mathbf{P}$? Why?

Question 5. Information Theory, Unicity Distance
A substitution cipher over a plaintext space of size n has $|\mathbf{K}| = n!$ (i.e., the key space size is $n!$). Let m be a positive integer. The m-gram substitution cipher is the substitution cipher where the plaintext (and ciphertext) spaces consist of all 26^m m-grams. Let $n! \approx \sqrt{2\pi n}(\frac{n}{e})^n$. Compute the unicity distance of the m-gram substitution cipher if $R_L = 0.75$ for $m = 1, 2, 3, 4, 5$.

Question 6. (Bonus Question) Prove that if only a single character is encrypted, then the shift cipher is perfectly secure.

Question 7. In the attack against Vigenere cipher,

7.1 Develop a new method to determine the key length using correlation instead of the index of coincidence.

7.2 Develop a new method to determine the key length using entropy instead of the index of coincidence.