实 验 报 告

课程名称 密码学基础

学　　院 计算机科学与技术学院

专　　业 保密管理

姓　　名 王君可

学　　号 17300240009

开 课 时 间 2019 至 2020 学年第 二 学期

实验项目名　　称	Cracking Vigenere cipher	成绩	

一、实验目的

1. Understanding Vigenère Cipher
2. Understanding Kasiski Test
3. Understanding Friedman's Index of Coincidence Test
4. Understanding Frequency Analysis to crack Caesar Cipher

二、实验内容

In this experiment, we will examine a well-known classical cipher, Vigenère Cipher. Dating back to 16th century, Vigenère Cipher is a polyalphabetic substitution cipher.

We will use some techniques and write computer program to cryptanalyze (to Crack!) the Vigenère Cipher.

An overview of the process is as follows. Inputting ciphertext encrypted by Vigenère Cipher and using Kasiski examination + Friedman test to determine the keyword length. Then using the ciphertext separated by the length of the keyword as a group, and crack it by analyzing the Caesar cipher. Finally, outputting the keyword and the corresponding plaintext.

三、实验步骤

1. Computing the chance for the false positive rate of Kasiski test

Read the paper Kasiski's Test: Couldn't the Repetitions be by Accident? *and answer the. question below:*

The likelihood of two same trigraphs not being from the same plaintext fragment?

2. Implementing Vigenère Cipher

a. func1: vigenere_encrypt (plaintext, key, alphabet)
b. func2: vigenere_decrypt (ciphertext, key, alphabet

3. Determining the Key size

a. func1: kasiski_test(ciphertext)
b. func2: index_of_coincidence(ciphertext, alpha)

4. Determining the keyword

func: find_likelyletters (coset, alpha, eng_freq)

四、实验结果及分析

1. the likelihood of two same trigraphs not being from the same plaintext. fragment?

From Theorem 1 in the given paper, we know that the probability that a repetition of substring with length t occurs in a string with length r is:

$K(N,r) = 1 - Q(N,r)$

Where

$Q(N,r) = \frac{N*(N-1)*(N-2)....(N-r+1)}{N^r}$, and $N = 26^t$

Therefore, if we set N to 26^3, r is the length of plaintext, the result is $\boldsymbol{K(26^3, r)} = 1 - \frac{N*(N-1)*(N-2)....(N-r+1)}{N^r}, N = 26^3$.

which means the probability that two same trigraphs are not from the same plaintext is $1 - \frac{26^{3r}*(26^{3r}-1)....(26^{3r}-r+1)}{26^{3r}}$. From table1, we can know that when r=100, the probability is 0.246. using the given upper boundary formula, we can know that for random strings of length r a repetition of any substring of length t is rather unlikely (< 50%) as long as t<=1.413 logr, when t=3, r should satisfy r <= 132.

2. the cracked key length:

Cracked key length by Kasiski test	**7**
Cracked key length by Friedman test	**4**

3. Using 7 as key length, I find the letters in the key are: BITCOIN

4. the cracked plaintexts:

COMMERCEONTHEINTERNETHASCOMETORELYALMOSTEXCLUSIVELYONFINANCIALINSTITUTIONSSERVIN GASTRUSTEDTHIRDPARTIESTOPROCESSELECTRONICPAYMENTSWHILETHESYSTEMWORKSWELLENOUG HFORMOSTTRANSACTIONSITSTILLSUFFERSFROMTHEINHERENTWEAKNESSESOFTHETRUSTBASEDMODE LCOMPLETELYNONREVERSIBLETRANSACTIONSARENOTREALLYPOSSIBLESINCEFINANCIALINSTITUTIONS CANNOTAVOIDMEDIATINGDISPUTESTHECOSTOFMEDIATIONINCREASESTRANSACTIONCOSTSLIMITINGTHE MINIMUMPRACTICALTRANSACTIONSIZEANDCUTTINGOFFTHEPOSSIBILITYFORSMALLCASUALTRANSACTIO NSANDTHEREISABROADERCOSTINTHELOSSOFABILITYTOMAKENONREVERSIBLEPAYMENTSFORNONREV ERSIBLESERVICESWITHTHEPOSSIBILITYOFREVERSALTHENEEDFORTRUSTSPREADSMERCHANTSMUSTBE WARYOFTHEIRCUSTOMERSHASSLINGTHEMFORMOREINFORMATIONTHANTHEYWOULDOTHERWISENEEDA CERTAINPERCENTAGEOFFRAUDISACCEPTEDASUNAVOIDABLETHESECOSTSANDPAYMENTUNCERTAINTIE SCANBEAVOIDEDINPERSONBYUSINGPHYSICALCURRENCYBUTNOMECHANISMEXISTSTOMAKEPAYMENTS OVERACOMMUNICATIONSCHANNELWITHOUTATRUSTEDPARTY.

5. time cost:

Experiment	Time(s)
Vigenere encryption	0.00114083
Vigenere decryption	0.00115180
Kasiski test	0.02843308
Friedman test	8.60691070e-5
Cracking	5.41879034

五、实验总结

1. The time cost of Vigenere encryption and decryption is relatively small.

2. The time cost of Friedman test is shorter than that of Kasiski test. From the code, there are only one for loop in the Friedman test while there two in the Kasiski test, therefore, Friedman is O(N) while Kasiski is O(N^2).

3. When the length of the ciphertext is small, the statistical pattern of the frequency of letters is less obvious, therefore, Kasiski test is more reliable than Friedman test.

4. Most time is spent on the process of finding the cipher letter for every coset.

5. Generally, the Vigenere is easy to be deciphered with small time costs, which makes it less popular in model encryption system.