A brief Introduction to Classical cipher

Li Zihao

Class 29, Grade 2017, School of Computing, Zhuhai College of Jilin University

Summary: This article mainly introduces some brief introductions to classical cipher, and gives the principles of substitution, such as Caesar Cipher, Vigenere Cipher, and Transposition Cipher, the Cryptanalysis, and the deciphering methods. Finally, the implementation program of Caesar Cipher and Transposition Cipher and cracking methods in the above cryptosystem are given

Key Words: classical cipher; Substitution Cipher; Transposition Cipher

1 The basic purpose of cryptography

The fundamental objective of cryptography is to enable two people, usually referred to as alice and bob, to communicate over an insecure channel in such a way that an opponent, Oscar, cannot understand what is being said. This channel could be a telephone line or computer network, for example, The information that Alice wants to send to Bob, which we call"plaintext", can be English text, numerical data, or anything at all-its structure is completely arbitrary. Alice encrypts the plaintext, using a predetermined key, and sends the resulting ciphertext over the channel. Oscar, upon seeing the ciphertext in the channel by eavesdropping cannot determine what the plaintext was; but Bob, who knows the encryption key can decrypt the ciphertext and reconstruct the plaintext. [1]

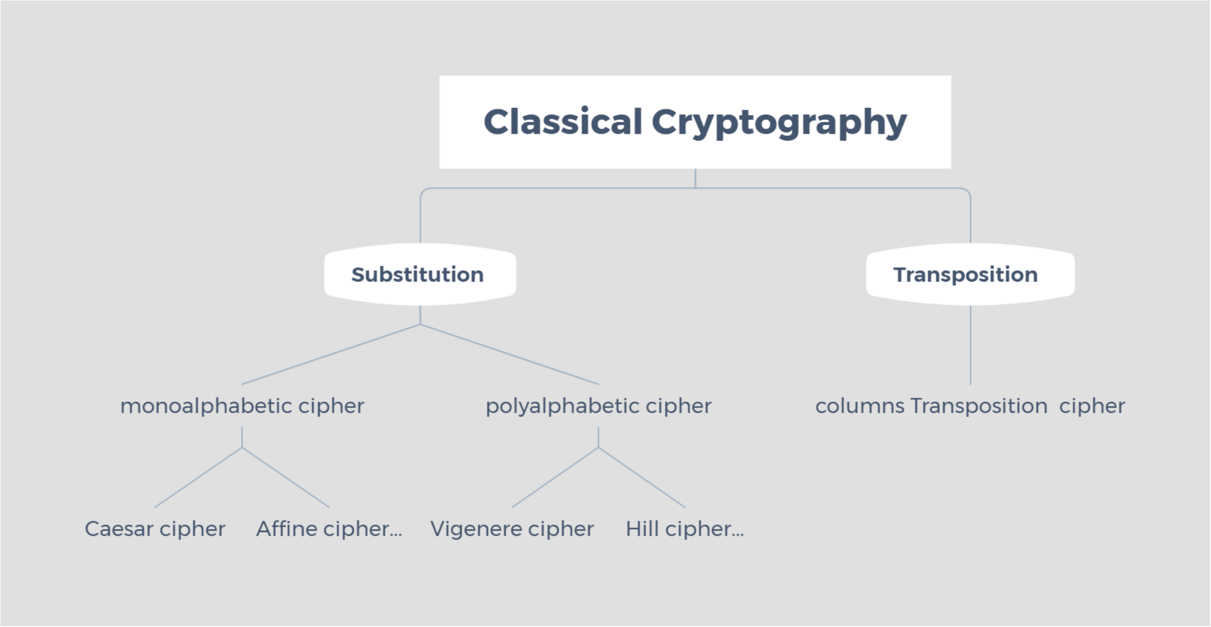
2 Encoding Method of Classical cipher

There are two main types of classical cipher coding methods, namely Substitution ciphers and Transposition ciphers.

Rearrange the letters in the plain text, as their positions have changed with the letters themselves remain the same place. The password thus compiled is called a Transposition password. the simplest way to Transposition Cipher is to reverse the alphabetical order of the plaintext and cut it into a set of fixed-length letters as the ciphertext.

The Substitution is to replace the plain text characters with some other characters according to a certain rule, and the replaced letters retain their original positions.

According to different processing methods, the specific classification is shown in the following figure:



3 Substitution Cipher

The substitution Cipher is to replace the plain text characters with some other characters according to a certain rule, and the replaced letters retain their original positions. Password substitution can be divided into monoalphabetic cipher and polyalphabetic cipher.

3.1 Monoalphabetic Cipher

Monoalphabetic cipher is a substitution cipher in which for a given key, the cipher alphabet for each plain alphabet is fixed throughout the encryption process. For example, if 'A' is encrypted as 'D', for any number of occurrence in that plaintext, 'A' will always get encrypted to 'D'. [2]

A classic example of Monoalphabetic cipher is the Caesar Cipher. which is based on modular arithmetic.

3.1.1 Caesar Cipher

Caesar Cipher is very simple, and its general expression is as follows:

$$e_{k(x)} = (x + k) \bmod 26 \text{（Encryption）}$$

$$d_k(y) = (y - k) \bmod 26 \text{（Decryption）}$$

FOR，$0 \leq k \leq 25$

We would use the Caesar Cipher to encrypt ordinary English text by setting up a correspondence between alphabetic characters and residues modulo 26 as follows: A ⇔ 0,B ⇔1,…,Z ⇔25.

A	B	C	D	E	F	G	H	I	J	K	L	M
0	1	2	3	4	5	6	7	8	9	10	11	12
N	O	P	Q	R	S	T	U	V	W	X	Y	Z
13	14	15	16	17	18	19	20	21	22	23	24	25

A small example will illustrate.

Suppose the key for a Caesar Cipher is K=11,and the plaintext is:

wewillmeetatmidnight

We first convert the plaintext to a sequence of integers using the specified correspondence, obtaining the following:

22	4	22	8	11	11	12	4	4	19	0	19	12	8	3	13	8	6	7	19

Next, we add 11 to each value, reducing each sum modulo 26:

7	15	7	19	22	22	23	15	15	4	11	4	23	19	14	24	19	17	18	4

Finally, we convert the sequence of integers to alphabetic characters, obtaining the ciphertext:

HPHTWWXPPELEXTOYTRSE

To decrypt the ciphertext, you only need to first convert the ciphertext to a sequence of integers, then subtract 11 from each value (reducing modulo 26), and finally convert the sequence of integers to alphabetic characters.

3.1.2 The flaw of Caesar Cipher

Obviously, Caesar Ciphers are not secure, because the key space is too small, there are only 26 possible situations, and all possible keys can be exhausted to get the meaningful plain text we want.

Here is an example: deciphering the ciphertext ” GUVF VF ZL FRPERG ZRFFNTR .”

```
Key #0 : GUVF VF ZL FRPERG ZRFFNTR.
Key #1 : FTUE UE YK EQODQF YQEEMSQ.
Key #2 : ESTD TD XJ DPNCPE XPDDLRP.
Key #3 : DRSC SC WI C0MB0D W0CCKQ0.
Key #4 : CQRB RB VH BNLANC VNBBJPN.
Key #5 : BPQA QA UG AMKZMB UMAAIOM.
Key #6 : AOPZ PZ TF ZLJYLA TLZZHNL.
Key #7 : ZNOY OY SE YKIXKZ SKYYGMK.
Key #8 : YMNX NX RD XJHWJY RJXXFLJ.
Key #9 : XLMW MW QC WIGVIX QIWWEKI.
Key #10: WKLV LV PB VHFUHW PHVVDJH.
Key #11: VJKU KU OA UGETGV OGUUCIG.
Key #12: UIJT JT NZ TFDSFU NFTTBHF.
Key #13: THIS IS MY SECRET MESSAGE.
Key #14: SGHR HR LX RDBQDS LDRRZFD.
Key #15: RFGQ GQ KW QCAPCR KCQQYEC.
Key #16: QEFP FP JV PBZOBQ JBPPXDB.
Key #17: PDEO EO IU OAYNAP IAOOWCA.
Key #18: OCDN DN HT NZXMZO HZNNVBZ.
Key #19: NBCM CM GS MYWLYN GYMMUAY.
Key #20: MABL BL FR LXVKXM FXLLTZX.
Key #21: LZAK AK EQ KWUJWL EWKKSYW.
Key #22: KYZJ ZJ DP JVTIVK DVJJRXV.
Key #23: JXYI YI CO IUSHUJ CUIIQWU.
Key #24: IWXH XH BN HTRGTI BTHHPVT.
Key #25: HVWG WG AM GSQFSH ASGGOUS.
```

The decrypted output of the key 13 is ordinary English, so the original encryption key must be 13.

On average, using the above method, you only need to try 26/2 times to get the plaintext.

The above example shows that the key space of a cryptosystem is too small to resist exhaustive password search attacks (violent attacks). The common solution is to increase the space for the keys. Although increasing the key space is not a sufficient condition to ensure the security of the cryptosystem.

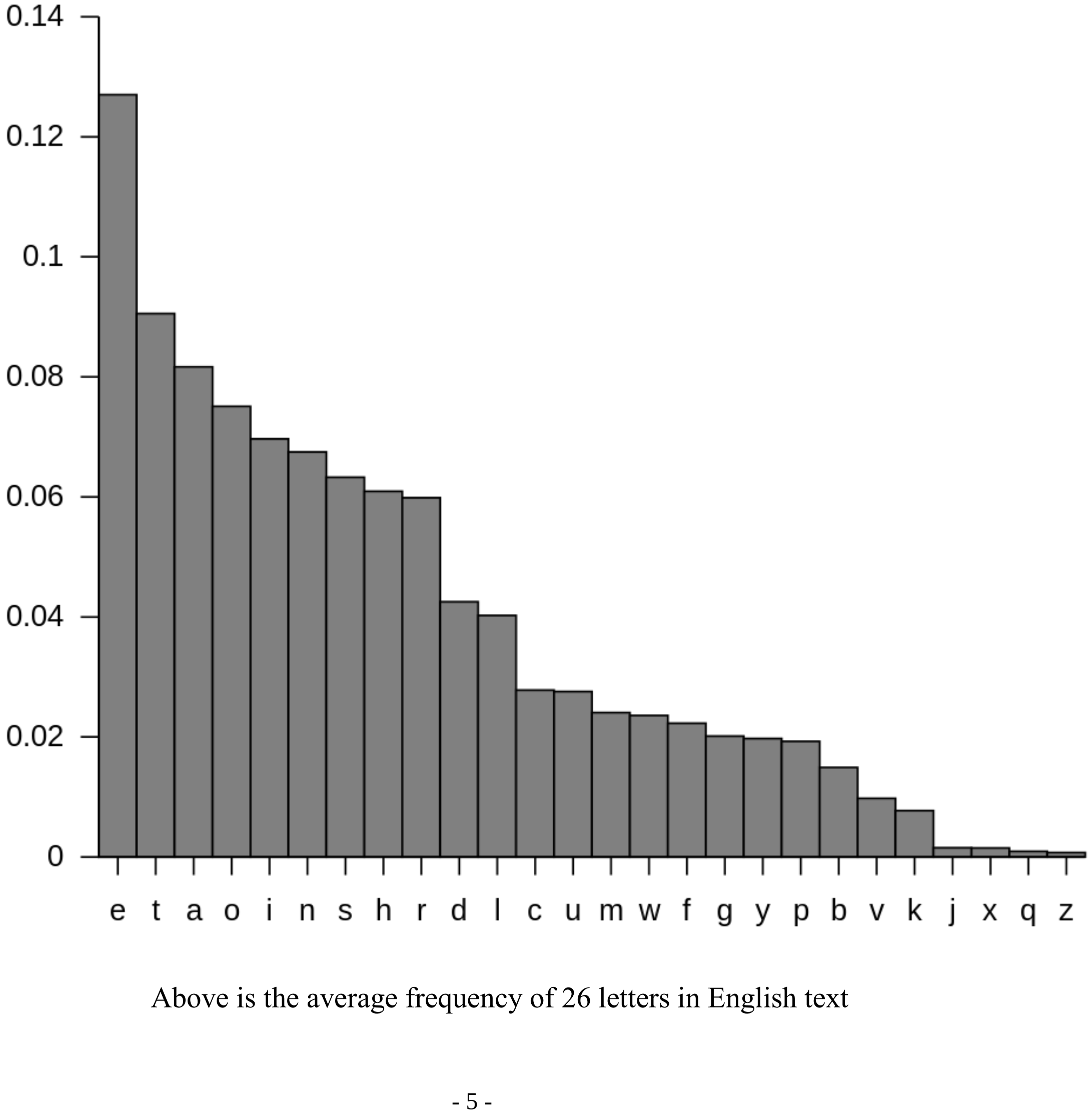
After the shortcomings of Caesar Cipher were exposed, someone improved it: replace the normal alphabet with a random alphabet. This simple replacement method increases the key space to 26! This means that exhaustive deciphering has been temporarily disabled.

Caesar Cipher, in addition to the shortcomings of the key space, have another drawback.

Once the key is selected, the number corresponding to each letter is encrypted and transformed into a unique number. If the original plaintext structure remains the same, this encryption method is vulnerable to frequency analysis. That is, by looking at how often certain characters or character sequences appear, you can find them without having to perform a full brute force attack.

3.1.3 Frequency Analysis

There are 26 letters in the English alphabet, but they don’t each appear an equal amount of the time in English text. It can be determined by counting a large number of English texts that some letters are used more than others. For example, the letters E, T, A and O occur very frequently in English words. But the letters J, X, Q, and Z are rarely found in English text. This technique is called frequency analysis.



Above is the average frequency of 26 letters in English text

When using simple Substitution Cipher such as Caesar Cipher, certain letters in the alphabet are always replaced by the same letter, the imbalance in the frequency also appear in the letters of the cipher text, with the decipherer maybe Knowing more about the ciphertext sender, it is still not difficult to determine the information contained in the cipher text.

The polyalphabetic Cipher can solve the problem that the English text is easy to be analyzed by frequency.

3.2 Polyalphabetic Cipher

Another encryption system corresponding to monoalphabetic Cipher is Polyalphabetic Cipher. Polyalphabetic Cipher is a substitution cipher in which for a given key, the cipher alphabet for each plain alphabet is not fixed throughout the encryption process.

The Vignere Cipher is one of the Polyalphabetic Cipher. It is a combination of multiple Caesar Cipher. Every other bit is replaced with a Caesar cipher secret key. The cycle begins until the end. Its key sequence is expressed as

K=k0,K1,K2,…Km−1，

It can be seen that the Caesar Cipher is a special case when the length of the Vigenere Cipher key is 1.The specific encryption and decryption formula is as follows:

yi = (xi + ki mod m) mod 26 （Encryption）

xi = (yi - ki mod m) mod 26 （Decryption）

Using the correspondence A ⇔ 0, B ⇔1,…,Z ⇔25 described earlier, we can associate each key with an alphabetic string of length m, called a keyword. The Vignere Cipher encrypts m alphabetic characters at a time: each plaintext element is equivalent to m alphabetic characters.

A	B	C	D	E	F	G	H	I	J	K	L	M
0	1	2	3	4	5	6	7	8	9	10	11	12
N	O	P	Q	R	S	T	U	V	W	X	Y	Z
13	14	15	16	17	18	19	20	21	22	23	24	25

Suppose m = 6 and the keyword is CIPHER， This corresponds to the numerical equivalent K=(2,8,15,7,4,17).Suppose the plaintext is the String:

thiscryptosystemisnotsecure

We convert the plaintext elements to residues modulo 26,write them in groups of six, and then ”add” the keyword modulo 26, as follows:

plaintext	I	L	O	V	E	Y	O	Y
corresponds	8	11	14	21	4	24	14	24
keywords	M	A	T	H	M	A	T	H
corresponds	12	0	19	7	12	0	19	7
modulo 26	20	11	7	2	16	24	7	5
ciphertext	U	L	H	C	Q	Y	H	F

he alphabetic equivalent of the ciphertext string would thus be:

YLHCQYHF

When decrypting, use the same key word and perform the inverse operation.

It can be seen that the size of the key space of the Vignere Cipher is 26^m, so even if the value of m is small, it takes a long time to use the exhaustive key search method. But because it still retains many characteristics of frequency distribution, Kasiski Examination method and frequency analysis can be easily cracked in the case of a small key.

“Kasiski Examination” is a process used to determine how long the Vignere key used to encrypt a ciphertext was. After this is determined, frequency analysis can be used to break each of the subkeys.

4 Transposition Cipher

Transposition cipher, the letters are just moved around. The letters or words of the plaintext are reordered in some way, fixed by a given rule (the key).

As mentioned above, the Vignere Cipher can be cracked by Kasiski Examination and frequency analysis.

“Kasiski Examination” is a process used to determine how long the Vignere key used to encrypt a ciphertext was.” The distance between the occurrence of the same substring in the ciphertext may be a multiple of t. Find out the distance of all the same substrings, especially the number of occurrences, t is the greatest common divisor of these distances.”

It can be seen here that by changing the position of characters in the ciphertext, the Kasiski Examination can be disabled.

One of the Transposition Cipher which called Column Transposition Cipher is described below:

Suppose m = 4,and m is a positive integer.

The plaintext is :

lizihaotebieshuai

First construct a table with 4 columns and several rows, and then fill the table from left to right and top to bottom, as shown below:

l	i	z	i
h	a	o	t
e	b	i	e
s	h	u	a
i			

Then read the characters in the table from top to bottom, and from left to right(skip the space). to construct the ciphertext.

lhesiiabhzoiuitea

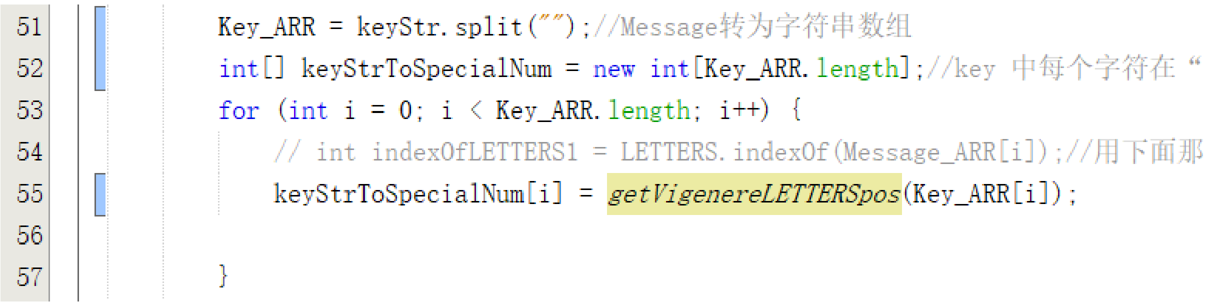
The decryption process is the reverse of the encryption process.

5 Program key code analysis

5.1 Vignere Cipher implementation

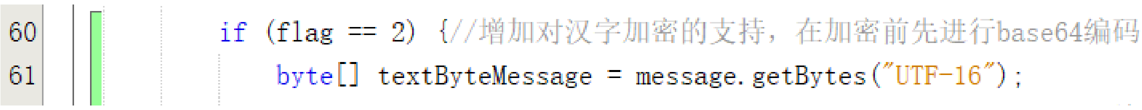
Vignere encryption

(1) Generate a corresponding key array based on the key. The specific generation method is that each character in the key is stored in the array keyStrToSpecialNum in the order corresponding to "VigenereLETTERS".



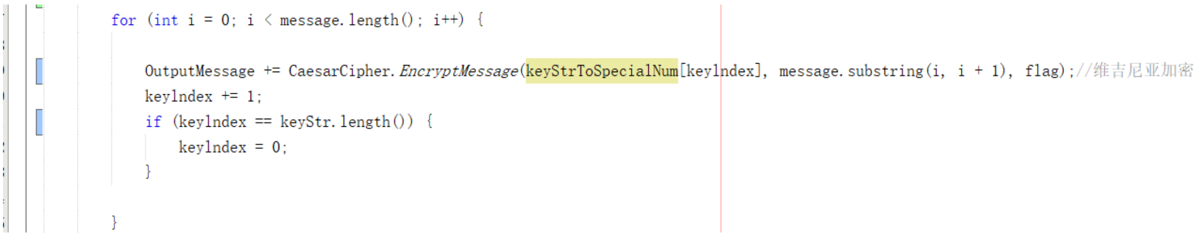
```
Key_ARR = keyStr.split("");//Message转为字符串数组
int[] keyStrToSpecialNum = new int[Key_ARR.length];//key 中每个字符在 “
for (int i = 0; i < Key_ARR.length; i++) {
    // int indexOfLETTERS1 = LETTERS.indexOf(Message_ARR[i]);//用下面那
    keyStrToSpecialNum[i] = getVigenereLETTERSpos(Key_ARR[i]);

}
```

(2) Base64 encoding the ciphertext. This step is to increase the support for non-English character encryption, which can be ignored.



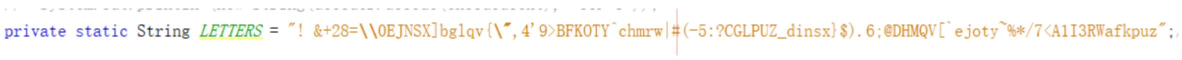
```
if (flag == 2) {//增加对汉字加密的支持, 在加密前先进行base64编码
    byte[] textByteMessage = message.getBytes("UTF-16");
```

(3) Use the number corresponding to the keyStrToSpecialNum array as the key to perform Caesar encryption on each character of the ciphertext, respectively.



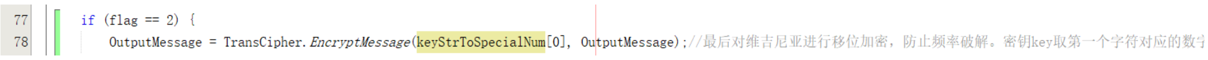
```
for (int i = 0; i < message.length(); i++) {

    OutputMessage += CaesarCipher.EncryptMessage(keyStrToSpecialNum[keylndex], message.substring(i, i + 1), flag);//维吉尼亚加密
    keylndex += 1;
    if (keylndex == keyStr.length()) {
        keylndex = 0;
    }

}
```

Among them, in order to increase the key space, the order of "encryption roulette" is randomly shuffled. For crackers, the keyspace of such Caesar encrypted has been increased from 26 to 26!



```
private static String LETTERS = "! &+28=\\0EJNSX]bglqv{\",4'9>BFKOTY^chmrw|#(-5:?CGLPUZ_dinsx}$).6;@DHMQV[`ejoty~%*/7<A1I3RWafkpuz";
```

(4) In order to prevent frequency analysis, the ciphertext is finally Transposition Cipher encrypted again.



```
if (flag == 2) {
    OutputMessage = TransCipher.EncryptMessage(keyStrToSpecialNum[0], OutputMessage);//最后对维吉尼亚进行移位加密，防止频率破解。密钥key取第一个字符对应的数字
```

The Vignere encryption process in turn is the decryption process, which is skipped here.

6 Conclusion

Classical cipher are usually easy to crack. Many cipher of Classical cipher can be cracked by ciphertext alone, so they are vulnerable to ciphertext-only attacks. Some Classical cipher (such as Caesar Cipher) have a limited number of keys, so this type of Cipher can be bruteforced to try all the keys.Substitution Cipher have a large number of keys, but they are easy to analyze by frequency, because each password letter represents a plaintext letter. polyalphabetic Cipher, such as the Virginia Cipher, use multiple substitutions to prevent simple frequency analysis, however, more advanced techniques such as the Kasiski Examination can be used to crack such ciphers.

Classical cipher Although seem to be very simple when it compared to the Modern cryptography, but Classical cipher are still worth learning for us on the principles of constructing passwords and some methods of solving problems. Therefore, most are not used directly for security applications(Some techniques from classical ciphers can be used to strengthen modern ciphers. For example, the MixColumns step in AES is a Hill cipher.), it can still be introduced as an introduction to cryptography.[3]

References

[1] Douglas Robert Stinson ,Cryptography: Theory and Practice, Third Edition:1

[2] https://www.tutorialspoint.com/cryptography/traditional_ciphers.htm; （accessed 2019/11/28）

[3] https://en.wikipedia.org/wiki/Classical_cipher（accessed 2019/11/29）