MAS 433 Tutorial 1

Wang Xueou (087199E16)

August 23, 2011

Remark For Question(4) and Question(6), we use matlab to help decrypt the cipher. The matlab code is attached.

Question (1) Solution: Key letter is "D ", this means $K = 3$. We have:

A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z
D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C

Thus we know "Tanjung Pagar "is encrypted into "WDQMXQJ SDJDU".

Question (2) Solution:
Using $K = 1, 2, 3, 4, 5, 6$, we can get the following decryption:

	Q	F	X	Y	Y	W	F	N	S	T	Z	Y	L	J	Y	X	W	T	D	F	Q	X	J	S	I	T	K	K
$K = 1$	p	e	w	x	w	v	e	m	r	s	y	x	k	i	x	w	v	s	c	e	p	w	i	r	h	s	j	j
$K = 2$	o	d	v	w	w	u	d	l	q	r	x	w	j	h	w	v	u	r	b	d	o	v	h	q	g	r	i	i
$K = 3$	n	c	u	v	v	t	c	k	p	q	w	v	i	g	v	u	t	q	a	c	n	u	g	p	f	q	h	h
$K = 4$	m	b	t	u	u	s	b	j	o	p	v	u	h	f	u	t	s	p	z	b	m	t	f	o	e	p	g	g
$K = 5$	l	a	s	t	t	r	a	i	n	o	u	t	g	e	t	s	r	o	y	a	l	s	e	n	d	o	f	f
$K = 6$	$\cdots$																											

So far, we have got the plaintext: Last train out gets royal send off.

Question (3) Solution:
3.1. Use the substitution table, "smartphone "is enctypted into ESYUIODRNB.
3.2. Use the substitution table, we get the following decryption:

A	I	E	M	A	X	X	A	C	J	T	I	I	R	E	W	A	I	C	D
i	t	s	d	i	f	f	i	c	u	l	t	t	o	s	w	i	t	c	h

Thus we get the plaintext: It's difficult to switch.

Question (4) Solution:
4.1. We attack the substitution cipher by frequency analysis. The encryption of substitution cipher does not randomize the frequency of occurrence of letter properly. We can calculate the frequency of occurence of letters in ciphertext, and then compare those frquency with the frequency of occurrence of letters in the language to determine the substition table. During the frequency analysis, we usually compare the frequency of letters, digram and trigram.

4.2. First, we analyse the frequency for each letter (use lettersubfreq.m) and find Y has the highest frequency, so we correspond Y to e. Then we analyse the frequency of digram of Y (use disubfreq.m) and find YW and WY both occur with very high frequency, so we assign W to r. Similarly do so for other letters and futher some guessing, we get the following substitution table:

a	b	c	d	E	f	g	h	i	j	k	l	m	n	o	p	q	r	s	t	u	v	w	x	y	z
K	F	O	C	Y	I	D	X	Z	P	B	N	R	M	Y	E	H	W	J	A	S	P	T	G	V	L

And we get the first sentence: Scientists at Rice University and HP are reporting this week that they can overcome a fundamental barrier to the continued rapid miniaturization of computer memory that has been the basis for the consumer electronics revolution in recent years.

Question (5) Solution:
By using the key "many ", we get the following decryption:

Ciphertext	O	I	C	F	Q	R	G	C	J	T
Key	M	A	N	Y	M	A	N	Y	M	A
Plaintext	c	i	p	h	e	r	t	e	x	t

Question (6) Solution:
6.1. To attack Vigenere cipher,we need first find the length of the key, and then for each letter in the key, the problem becomes a simple shift cipher. To find the length of the key, we can use 2 methods:Kasiski test and the index of coincidence.
For Kasisti test, we observe that two identical segments of plaintext will be encrypted to the same ciphertext if their distance is the multiple of the key. Thus we search for pairs of identical segments of length at least 3. Then we record distances between the two segments: $\Delta_1, \Delta_2, \cdots$. The length of the key is a divisor of $gcd(\Delta_1, \Delta_2, \cdots)$.
For the index of coincidence, if we guess the length of the key correctly, then the distribution of the ciphertext letters would be close to the distribution of the English letters; if we guess it wrongly, then the distribution of the ciphertext letters would be random. We can calculate the probability that two random elements of X are identical. If X is a string of Engilish language, the probability is 0.065, while if it is a random string, it is 0.038.

6.2. The key is $jksnfe$, and the first sentence is "A computer virus is a computer program that can copy itself and infect a computer. "
By using index of coincidence (decrypt.m), we try the key length of $1, 2, 3, 4, 5, 6$ and find that the key length is 6. Then, for each letter in the key, the problem becomes a simple shift cipher. Thus, we do frequency analysis for each key letter (test.m), and find that the cipher corresponding to each key letter has N, O, J, W, J, I respectively as highest frequency. We make them to correspond to e and get the key $jkdsfe$. However, this key is not the real key. After analysing the text we have got using this key, we find the third and fourth letter in the key is not correct. Thus, we try the second highest-frequency letter in the third and fourth letter. For the third one, we have W, L of equal frequency as second highest ones. Try them and we find W should correspond to e but not L. Similarly for the fourth letter, we find H should correspond to e. Now we can get the real key: $jksdfe$.

Question (7) Solution:
7.1.
Step 1: The playfair key table is

N	A	Y	G	B
C	D	E	F	H
I	K	L	M	O
P	Q	R	S	T
U	V	W	X	Z

Step2: Break the message into digrams and encrypt the message according to the table:

fo	lx	lo	ws	as	tr	on	gl	ys	ci	en	ti	fi	ca	px	pr	oa	ch
HM	MW	MI	XR	GQ	PS	IB	YM	GR	IP	CY	PO	CM	DN	SU	QS	KB	DC

7.2. According to the playfair table above, we get the following decryption:

PX	AB	CL	NZ	MI	EZ	HR	KY	AC	WL	RF	QW	FW
su	ng	ei	bu	lo	hw	et	la	nd	re	se	rv	ex

Thus we get the plaintext: Sungei Buloh wetland reserve.

Question (8) Solution:
8.1. The π^{-1} is:

x	1	2	3	4	5	6	7	8
$\pi^{-1}(x)$	4	1	6	2	7	3	8	5

8.2. We partition the ciphertext into groups of 8 letters and rearrange them according to π^{-1}

E	T	E	G	E	N	L	M	D	N	T	N	E	O	O	R	D	A	H	A	T	E	C	O	E	S	A	H	L	R	M	I
g	e	n	t	l	e	m	e	n	d	o	n	o	t	r	e	a	d	e	a	c	h	o	t	h	e	r	s	m	a	i	l

Thus we get the plaintext: Gentlemen don't read each other's mail.

Question (9) Solution:
9.1. We claim the composite of 2 substition ciphers is still a substitution cipher. Thus, we can attack this cipher by frequency analysis.

Proof. Define $S = S_2 \circ S_1$. We now show that S is invertible. Suppose $c_i = S_1(p_i), q_i = S_2(c_i)$, then we have $q_i = S(p_i) = S_2(S_1(p_i))$. As S_1 and S_2 are both invertible, we get: $S^{-1}(q_i) = (S_2 \circ S_1)^{-1}(q_i) = S_1^{-1}(S_2^{-1}(q_i)) = S_1^{-1}(c_i) = p_i$, i.e., $S^{-1}(q_i) = p_i$. Thus S is invertible, and the composite of 2 substituion cipher is just another substitution cipher. □

9.2. Let $K_1 = k_1k_2\cdots k_m$, $K_2 = h_1h_2\cdots h_n$, $t = lcm(m,n)$. Now for $1 \le i \le t$, define a new key $L = l_1l_2\cdots l_t$ such that:

$$l_i = k_{i \mod m} + h_{i \mod n} \mod 26$$

By applying V_{k_1} and V_{k_2} to P, we get

$$c_i = ((p_i + k_{i \mod m}) \mod 26) + h_{i \mod n} \mod 26 = (p_i + k_{i \mod m} + h_{i \mod n}) \mod 26$$

By applying the new key L, we get

$$c_i = (p_i + l_i) \mod 26 = (p_i + (k_{i \mod m} + h_{i \mod n} \mod 26)) \mod 26 = (p_i + k_{i \mod m} + h_{i \mod n}) \mod 26$$

Thus, applying $C = V_{K_2}(V_{K_1}(P))$ is just the same as applying the new Vigenere cipher V_L. We can attack this Vigenere cipher by analysing the key length with Kasiski and index of coincidence approaches.
The security of using more than two Vigenere cipher increases as it is equivalent to using a new Vigenere cipher with a longer length. Applying Kasiski and index of coincidence approaches will take a much longer time. However, it is still a Vigenere cipher, so the methods of attacking still apply to it.

Question (10) Solution:
Shift cipher. As the attacker has already known the corresponding cipher to the piece of plaintext, the key is already known. He can just shift back the cipher to get the plain text.
Substitution cipher. In this case, the attacker has already known part of the substitution table. He can first use the substitution table to decrypt part of the cipher. The remained plaintext can be gotten by further analysing the frequency of the letters, digrams and trigrams, and he may also be able to guess out some of the remained plaintext.
Vigenere cipher. First, the attacker can do Kasiski test or index of coincidence method to find the length of the key. If the key length if shorter than the length of the known plaintext, then the key is easy to find out. If the length of the key is longer than the plaintext, he knows part of the key. Thus he can use this part of the key to decrypt part of the cipher. The remained cipher can be decryted using frequency analysis andd some guessing.
Playfair By comparing the plaintext and the corresponding cipher, the attacker can work out part of the key table. He can use the incomplete table to decrypt part of the cipher. Then he can try to fill the remained entries in the table. May be it is easy to fill it from Z and backwards. Dropping the duplicated words that already appeared and at the same time use it to decrypt. This may involve error and trial method.
Composition of Vigenere ciphers. We have analysed that the composite of Vigenere cipher is also a Vigenere cipher, so the break can be done as aforementioned Vigenere cipher in this question.