NANYANG TECHNOLOGICAL UNIVERSITY

SEMESTER I EXAMINATION 2017-2018 SUGGESTED SOLUTION

MH4311 – Cryptography

December 2017 TIME ALLOWED: 2 HOURS

INSTRUCTIONS TO CANDIDATES

1. This examination paper contains **FOUR (4)** questions and comprises **FOUR (4)** printed pages.
2. Answer all questions. The marks for each question are indicated at the beginning of each question.
3. Answer each question beginning on a **FRESH** page of the answer book.
4. This is a **RESTRICTED OPEN BOOK** exam. You are allowed to bring into the examination hall **ONE (1)** piece of A4-size paper written or printed on both sides.
5. Candidates may use calculators. However, they should write down systematically the steps in the workings.

Solutions provided by: Brandon Goh – bgoh008@e.ntu.edu.sg

Question 1. Classical Ciphers, Information Theory (40 marks)

(a) A password contains 10 characters. Each character in the password is a lower case letter ('a' to 'z'). Suppose that each character of the password iw independently and randomly generated. What is the entropy of the password? (10 marks)

(b) The ciphertext $UCEMSKIBY$ is obtained by using a Vigenere cipher with the four-letter key $FREE$. Find the plaintext. (10 marks)

(c) In the attack against Vigenere cipher, suppose that we already found that the length of the key is 6. The frequencies of the letters at positions $6i$ in the ciphertext (i is a non-negative integer, the position of the ciphertext is 0) are given below:

Frequency							
A	0.088	B	0.077	C	0.070	D	0.042
E	0.032	F	0.007	G	0.004	H	0.024
I	0.000	J	0.091	K	0.011	L	0.035
M	0.046	N	0.112	O	0.039	P	0.004
Q	0.046	R	0.056	S	0.004	T	0.003
U	0.021	V	0.046	W	0.042	X	0.081
Y	0.021	Z	0.000				

Most likely what is the first letter of the key? (10 marks)

Reference. In English texts, the frequencies of the letters are given below:

Frequency							
A	0.082	B	0.015	C	0.028	D	0.043
E	0.127	F	0.022	G	0.020	H	0.061
I	0.070	J	0.002	K	0.008	L	0.040
M	0.024	N	0.067	O	0.075	P	0.019
Q	0.001	R	0.060	S	0.063	T	0.091
U	0.028	V	0.010	W	0.023	X	0.001
Y	0.020	Z	0.001				

(d) In the attack against Vigenere cipher, suppose that we already found that the length of the key is 6. The frequencies of the letters at the positions $6i + 2$ in the ciphertext (i is a non-negative integer, the position of the first letter of the ciphertext is 0) are given below:

Frequency							
A	0.070	B	0.011	C	0.007	D	0.032
E	0.042	F	0.081	G	0.063	H	0.028
I	0.000	J	0.099	K	0.092	L	0.095
M	0.039	N	0.018	O	0.007	P	0.004
Q	0.007	R	0.000	S	0.081	T	0.007
U	0.042	V	0.021	W	0.095	X	0.011
Y	0.018	Z	0.032				

Most likely what is the third letter of the key? (10 marks)

Reference. In English texts, the frequencies of the letters are given below:

Frequency							
A	0.082	B	0.015	C	0.028	D	0.043
E	0.127	F	0.022	G	0.020	H	0.061
I	0.070	J	0.002	K	0.008	L	0.040
M	0.024	N	0.067	O	0.075	P	0.019
Q	0.001	R	0.060	S	0.063	T	0.091
U	0.028	V	0.010	W	0.023	X	0.001
Y	0.020	Z	0.001				

Answer

(a) Definition of Entropy (of 1 char):

$$H(X) = \sum_{x \in X} P(X) \cdot \log_2 P(X)$$

Since there are 26 lowercase alphabets, the entropy per letter is:

$$\begin{aligned} H(X) &= \sum_{x=1}^{26} -\frac{1}{26}\log_2(\frac{1}{26}) \\ &= -\log_2(\frac{1}{26}) \end{aligned}$$

Each letter is **independent and randomly generated**, hence for the 10 letters, the entropy is of the following value

$$\begin{aligned} H(X) &= 10 \times -\log_2(\frac{1}{26}) \\ &= -10\log_2(\frac{1}{26}) \end{aligned}$$

□

(b) To decrypt a ciphertext that has been encrypted with the Vigenere cipher, we must extend the key to the length of the plaintext. Only then can we perform decryption to calculate the value of the plaintext.

Letter to number mapping:

Letter	A	B	C	D	E	F	G	H	I	J	K	L	M
Number (mod 25)	0	1	2	3	4	5	6	7	8	9	10	11	12

Letter	N	O	P	Q	R	S	T	U	V	W	X	Y	Z
Number (mod 25)	13	14	15	16	17	18	19	20	21	22	23	24	25

In numbers:

Ciphertext	20	2	4	12	18	10	8	1	24
Key	5	17	4	4	5	17	4	4	5
Plaintext	15	11	0	8	13	19	4	23	19

In letters:

Ciphertext	U	C	E	M	S	K	I	B	Y
Key	F	R	E	E	F	R	E	E	F
Plaintext	p	l	a	i	n	t	e	x	t

$\therefore$ The plaintext is 'plaintext' $\square$

(c) If we map the letter with the highest frequency:

Plaintext		Ciphertext	Key
E	$\rightarrow$	N	9

The key obtained is 9. Using this as a reference for the subsequent letters, we see the following

Plaintext	Frequency		Ciphertext	Frequency
F	0.022	$\rightarrow$	O	0.039
G	0.020	$\rightarrow$	P	0.004
H	0.061	$\rightarrow$	Q	0.046
I	0.070	$\rightarrow$	R	0.056
J	0.002	$\rightarrow$	S	0.004

The difference in frequencies between the plaintext and ciphertext are very minimal and hence the most probable key is 9. $\square$

(d) If we map the letter with the highest frequency:

Plaintext		Ciphertext	Key
E	$\rightarrow$	J	5

Using key 5 as a reference, then

Plaintext	Frequency		Ciphertext	Frequency
F	0.022	$\rightarrow$	K	0.092
G	0.020	$\rightarrow$	L	0.095
H	0.061	$\rightarrow$	M	0.039

From these 3 mappings, we can tell that the frequency fluctuations are very large and so 5 is not a suitable key.

Using the 2^{nd} highest frequency for mapping, we get

Plaintext		Ciphertext	Key
E	$\rightarrow$	W	18

If we do the mappings again, we obtain the following

Plaintext	Frequency		Ciphertext	Frequency
F	0.022	$\rightarrow$	X	0.011
G	0.020	$\rightarrow$	Y	0.018
H	0.061	$\rightarrow$	Z	0.032

We can tell that the frequency variations now are minimal and hence is most probable. $\square$

(Alternatively) We consider mapping from the lower frequencies. Specifically we see that the following two letters have the same distance apart.

Plaintext	Frequency		Ciphertext	Frequency	Key
Q	0.001	$\rightarrow$	I	0.000	18
Z	0.001	$\rightarrow$	R	0.000	18

Taking key 18 as a reference,

Plaintext	Frequency		Ciphertext	Frequency
R	0.060	$\rightarrow$	J	0.099
S	0.063	$\rightarrow$	K	0.092
T	0.091	$\rightarrow$	L	0.095
U	0.024	$\rightarrow$	M	0.039
V	0.067	$\rightarrow$	N	0.018

Extending the mapping to all 26 alphabets, we can tell that the variations in frequency are minimal and hence key 18 is the most probably for letters at positions $6i + 2$. $\square$

Question 2. Hash function (10 marks)

(a) When we use SHA-256 to hash a 200-bit message, how many '0' bits should be padded to the message before the 64-bit message length is padded? (5 marks)

(b) Suppose that you are using SHA-256 to compute the message digest of a 1000-bit message. How many compression function operations are needed? Justify your answer. (5 marks)

Answer

(a) SHA-256: 512-bit block size

$$200 + 1 + X + 64 = \alpha \cdot 512$$

200: Number of bits to hold the message.
1: 1-bit allocated for padding of '1'.
X: X '0's required for padding.
64: 64 bits allocated to hold the length of the message.

We need to solve the above equation to determine the smallest value of α and X.

$$\begin{aligned} \alpha \cdot 512 &= 200 + 1 + X + 64 \\ &= 265 + X \end{aligned}$$

Solving for the smallest values yield that $\alpha = 1$ and $X = 247$. So we need to pad 247 '0's to the message. □

(b) We need to solve the following equation:

$$\begin{aligned} \alpha \cdot 512 &= 1000 + 1 + X + 64 \\ &= 1065 + X \end{aligned}$$

Solving for the least value of α and X yields $\alpha = 3$ and $X = 471$. Therefore, 3 compression functions are required to compute the message digest. □

Question 3. Block Cipher (30 marks)

(a) In DES, the Sbox is non-invertible. In AES, the Sbox is invertible. Can we replace the Sbox of AES with a non-invertible Sbox? Please briefly justify your answer. (10 marks)

(b) A finite field $\mathbf{GF}(2^5)$ is defined by the irreducible polynomial x^5+x^2+1. Find the inverse of 3 in this finite field. (10 marks)

(c) The finite field $\mathbf{GF}(2^8)$ in AES is defined by the irreducible polynomial $x^8+x^4+x^3+x+1$. $a(x) = \{82\}x^3 + x$, $b(x) = \{4\}x^2 + 1$ and $x^4 + 1$ are polynomials with coefficients over $\mathbf{GF}(2^8)$. Here {XY} indicates an integer in hexadecimal format. Compute $a(x) \otimes b(x) = a(x) \bullet b(x)$ mod $(x^4 + 1)$. (10 marks)

Answer

(a) The Sbox is DES is non-invertible as the Feistel Network allows us to retrieve the values at state $i-1 \ \ \forall i > 0$. For AES, a similar structure does not exist and making the Sbox in AES to be non-invertible prevents us from getting the value at state $i-1 \ \ \forall i > 0$. (i.e. Cannot be decrypted if AES Sbox is non-invertible) □

(b) **NEVER attempt GF/AES calculation questions in decimal form. Always convert to polynomial form first!**

3 in the field $GF(2^5)$ is expressed as the polynomial $x+1$. Calculate the inverse as follows:

$$x^5 + x^2 + 1 \mod x + 1 = (x^4 + x^3 + x^2)(x + 1) + 1$$

$\therefore (x+1)^{-1} = (x^4 + x^3 + x^2)$.
$x^4 + x^3 + x^2$ expressed in hex(adecimal) form is {1C}. □

(c) *Note that finite field for polynomials is separate from finite field for coefficients. For simplicity, we will use y as a placeholder for the polynomial form of **coefficients** to prevent confusion.

$$\begin{aligned} a(x) \otimes b(x) &= a(x) \bullet b(x) \mod (x^4+1) \\ &\equiv (\{82\}x^3 + x) \bullet (\{4\}x^2 + 1) \\ &\equiv (\{82\} \bullet \{4\})x^5 + (\{82\} + \{4\})x^3 + x \\ &\equiv (\{82\} \bullet \{4\})x^5 + \{86\}x^3 + x \\ &\equiv \{86\}x^3 + (\{82\} \bullet \{4\} + \{1\})x \end{aligned}$$

We need to determine the value of $\{82\} \bullet \{4\}$ and have to do so by making use of the irreducible polynomial.

$$\begin{aligned} \{82\} &= y^7 + y \\ \{4\} &= y^2 \\ \{1\} &= 1 \\ \{82\} \bullet \{4\} + 1 &= (y^7 + y)(y^2) + 1 \\ &= y^9 + y^3 + 1 \end{aligned}$$

The coefficient is subject to the irreducible polynomial $y^8+y^4+y^3+y+1$.

$$\begin{aligned} y^9 + y^3 + 1 \mod (y^8 + y^4 + y^3 + y + 1) &= (y^5 + y^4 + y^2 + y) + y^3 + 1 \\ &= y^5 + y^4 + y^3 + y^2 + y + 1 \\ &= 111111 \text{ (Binary form)} \\ &= \{3F\} \text{ (Hexadecimal form)} \end{aligned}$$

Going back to our previous equation, we have obtained the solution to the coefficient $\{82\} \bullet \{4\} + \{1\}$ and can finish up with the full answer.

$$\begin{aligned} a(x) \otimes b(x) &= \{86\}x^3 + (\{82\} \bullet \{4\} + \{1\})x \\ &= \{86\}x^3 + \{3F\}x \end{aligned}$$

□

Question 4. Modes of Block Cipher (20 marks)
Suppose that there is a semiconductor chip with AES being implemented, so we can perform 64 AES encryptions and 64 AES decryptions in parallel (i.e., encrypt and decrypt 64 message blocks at the same time). Assume that each AES encryption takes 10 clock cycles and each AES decryption also takes 10 clock cycles. You do not need to consider padding in this question.

(a) Suppose that you need to encrypt a 1024-byte plaintext. We ignore the cost of one XOR operation comparing to the AES operation. What is the minimum number of clock cycles that are needed to encrypt the plaintext if CBC mode is used? (5 marks)

(b) Suppose that you need to decrypt a 1024-byte ciphertext which was encrypted using CBC mode. We ignore the cost of one XOR operation comparing to the AES operation. What is the minimum number of clock cycles that are needed to decrypt the ciphertext? (5 marks)

(c) Suppose that you need to encrypt a 1024-byte plaintext. We ignore the cost of one XOR operation comparing to the AES operation, and ignore the cost of updating the counter. What is the minimum number of clock cycles that are needed to encrypt the whole message if CTR mode is used? (5 marks)

(d) Suppose that you need to decrypt a 1024-byte ciphertext which was encrypted using OFB mode. We ignore the cost of one XOR operation comparing to the AES operation. What is the minimum number of clock cycles that are needed to decrypt the ciphertext? (5 marks)

Answer

*For this question, refer to the lecture nodes on encryption modes when unsure. There are 5 types of encryption modes, ECB, CBC, OFB, CFB and CTR. Depending on the encryption/decryption mode used, the process may be parallelised (i.e. depending on the encryption/decryption mode used, the time taken can be reduced by making more efficient use of system resources).

(a) One 1024-byte plaintext is split into 64-byte message blocks. So, we will need $\frac{1024}{64} = 16$ message blocks in total.

CBC encryption of plaintext i $\forall i > 0$ (i starts from 0) requires the ciphertext obtained from state $i-1$. Therefore, we will need to perform 16 encryptions in succession and hence 160 clock cycles are required. □

(b) CBC decryption of ciphertext i $\forall i > 1$ does not depend on state $i-1$. Hence, the number of clock cycles required for decryption is 10. (Note that the number of cycles required is still 160, but it is done in parallel so the number of **clock** cycles is 10.) □

(c) Both CTR encryption and decryption can be performed in parallel since the process is isolated within each block. Therefore only 10 clock cycles are needed. □

(d) Each block requires the criphertext obtained from state $i - 1$. As such, the decryption cannot be performed in parallel and requires 160 clock cycles. □

END OF PAPER