Scribe: Cryptography and Network Security (Lecture 10)

Kuheli Pratihar

18-Sep-2020

1 Concepts Covered

1. Introduction to Number Theory (Contd.)
   (a) Fermat's Little Theorem
   (b) Homework Questions and Solutions
2. Classical Ciphers and Perfect Ciphers
   (a) Polyalphabetic Ciphers
      i. A Closer Look into the Vigenere Cipher
   (b) Cryptanalysis
      i. Cryptanalysis of the Monoalphabetic Cipher
      ii. Cryptanalysis of the Affine Cipher

2 Introduction to Number Theory(Contd.)

2.1 Fermat's Little Theorem

$$\text{If } gcd(a, m) = 1 \text{ then } a^{\phi(m)} \equiv 1 \, (mod \, m)$$

Proof:

$$R = \{r_1, \cdots, r_{\phi(m)}\} \text{ is a reduced system } (mod \, m)$$

- If $gcd(a, m) = 1$, we see that $\{ar_1, \cdots, ar_{\phi(m)}\}$ is also a reduced system $(mod \, m)$
- It is a permutation of the set R
- Thus, the product of the elements in both the sets are the same. Hence,

$$a^{\phi(m)} r_1, \cdots, r_{\phi(m)} \equiv r_1, \cdots, r_{\phi(m)} \, (mod \, m) \Rightarrow a^{\phi(m)} \equiv 1 \, (mod \, m)$$

Note: We can cancel the residues as they are co-prime with m and hence have multiplicative inverse. These are the two references for this topic[1],[3] and [2].
Explanation: Suppose $m = 10$, then $R_c = \{1, 2, 3, 4, 5, 6, 7, 8, 9\}$ is the complete residue set. Then we consider numbers co-prime to m, to form the set of reduced residual set $R = \{1, 3, 7, 9\}$.
Here, let us consider $\{1, 3, 7, 9\}$ as $\{r_1, r_2, r_3, r_4\}$ and we know that

$$\phi(10) = \phi(5 \times 2) = \phi(5) \times \phi(2) = 4 \times 1 = 4$$

Where $\phi(m)$ is the Euler's function of m which returns the number of positive integers till m which are relatively prime to m. Let us choose a number a such the $gcd(a, m) = 1$ e.g. $a = 3$. Checking the residue of set $\{ar_1, ar_2, ar_3, ar_4\}$ we find that it is another permutation of R (reduced residual set).

$$\{3, 9, 21, 27\} \, mod \, 10 = \{3, 9, 1, 7\}$$

Use this idea we can write the following

$$a \times 1 \times a \times 3 \times a \times 7 \times a \times 9 \, (mod \, 10) \equiv 1 \times 3 \times 7 \times 9 \, (mod \, 10) \Rightarrow a^4 \equiv 1 \, (mod \, m)$$

Therefore, in the general case we shall have

$$a^{\phi(m)} \equiv 1 \, (mod \, m)$$

Example 1: To find the remainder when 72^{1001} is divided by 31

Since, $72 \equiv 10 \, (mod \, 31)$, Hence, $72^{1001} \equiv 10^{1001} \, (mod \, 31)$

- Now, from Fermat's Theorem $10^{30} \equiv 1 \, (mod \, 30)$ (Note: 31 is a prime number, so $\phi(31) = 30$
- Raising both sides to the power of 33

$$10^{990} \equiv 1 \, (mod \, 31)$$

Thus, $10^{1001} = 10^{990+11} = 10^{990}.10^8.10^2.10^1$ and now taking the congruence we get,

$$10^{1001} \equiv 1.(10^2)^4.10^2.10 \, (mod \, 31) \equiv (7)^4.7.10 \, (mod \, 31) \equiv (49^2).7.10 \, (mod \, 31) \equiv (-13)^2.7.10 \, (mod \, 31) \equiv (14.7).10 \, (mod \, 31) \equiv 98.10 \, (mod \, 31) \equiv 5.10 \, (mod \, 31) \equiv 19 \, (mod \, 31)$$

So, the remainder is 19.

Example 2: To find the least residue of $72^{973} (mod \, 72)$

We have

$$\begin{aligned} \phi(72) &= \phi(8).\phi(9) \\ &= 4 \times 6 \\ &= 24 \end{aligned}$$

Now,

$$7^{973} = 7^{24\times 40} \cdot 7^{13}$$

Again from Euler's Theorem, we have:

$$7^{24} \equiv 1 \, mod \, 72$$
$$\Rightarrow (7^{24})^{40} \equiv 1 \, mod \, 72$$

Therefore

$$\Rightarrow 7^{973} \equiv 1 \cdot 7^{13} \, (mod \, 72)$$
$$\Rightarrow 7^{973} \equiv 1 \cdot 7^{12} \cdot 7^{1} \, (mod \, 72)$$
$$\Rightarrow 7^{973} \equiv 1 \cdot (-17)^{4} \cdot 7^{1} (mod \, 72)$$
$$\Rightarrow 7^{973} \equiv 1 \cdot (289)^{2} \cdot 7^{1} (mod \, 72)$$
$$\Rightarrow 7^{973} \equiv 1 \cdot 1 \cdot 7^{1} (mod \, 72)$$
$$\Rightarrow 7^{973} \equiv 1 \cdot 1 \cdot 7 (mod \, 72)$$
$$\Rightarrow 7^{973} \equiv 7 (mod \, 72)$$

Hence, we get the least residue of $7^{973} \, (mod \, 72)$ as 7

2.2 Homework Questions and Solutions:

Problem 1:

Let $k|a$ and $k|b$ and $gcd(k, m) = d$
To Prove: $a \equiv b \, (mod \, m) \Rightarrow \frac{a}{k} \equiv \frac{b}{k} \, (mod \, \frac{m}{d})$

Solution: We know that $a \equiv b \, (mod \; m) \Rightarrow (a - b) = cm$ and $a = kc_1$ and $b = kc_2$ where $c, c_1, c_2 \in \mathbb{Z}$

$$(a - b) = cm \Rightarrow \frac{a-b}{d} = c \times \frac{m}{d}$$
$$\Rightarrow (c_1 - c_2) \times \frac{k}{d} = c \times \frac{m}{d}$$
$$\Rightarrow \frac{m}{d} | (\frac{k}{d} \times (c_1 - c_2))$$
$$\Rightarrow \frac{m}{d} | (c_1 - c_2)$$
$$\Rightarrow c_1 \equiv c_2 \, (mod \, \frac{m}{d})$$
$$\Rightarrow \frac{a}{k} \equiv \frac{b}{k} \, (mod \, \frac{m}{d})$$

It is to be noted that $gcd(m, k) = 1 \Rightarrow gcd(\frac{m}{d}, \frac{k}{d}) = 1$.
Problem 2:

Let $a \equiv b \, (mod \, p^k)$ then prove $a^p \equiv b^p \, (mod \, p^{k+1})$
Here, p is prime (Case-I) and p is not prime (Case-II)

Solution: We know from $a \equiv b\,(mod\,p^k) \Rightarrow a - b = c\,p^k$ for some $c \in \mathbb{Z}$. We can prove $a^p \equiv b^p\,(mod\,p^{k+1})$ for general p (prime or composite). Also, we can factorize $(a^p - b^b)$ to the following expression where the second expression has p terms.

$$(a^p - b^p) = (a - b) \times (a^{p-1} + a^{p-2}b + a^{p-3}b^2 + \cdots + ab^{p-2} + b^{p-1}) \quad (1)$$

Using the properties of congruence we can write

$$\begin{gathered} a \equiv b\,(mod\,p^k) \\ \Rightarrow a^{p-1} \equiv b^{p-1}\,(mod\,p^k) \\ \Rightarrow a^{p-1} - b^{p-1} = c_1\,p^k \\ \Rightarrow a^{p-1} = b^{p-1} + c_1\,p^k \end{gathered}$$

We can also represent other terms in the second expression of (1) using the congruence property. For $a^m\,b^n$ with $m + n = p - 1$ and $m, n \geq 1$

$$\begin{gathered} a^m \equiv b^m\,(mod\,p^k) \\ \Rightarrow a^m\,b^n \equiv b^{m+n}\,(mod\,p^k) \\ \Rightarrow a^m\,b^n \equiv b^{p-1}\,(mod\,p^k) \\ \Rightarrow a^m\,b^n = b^{p-1} + c_{m,n}\,p^k \end{gathered}$$

Continuing (1) we get

$$\begin{aligned} (a^p - b^p) &= (a-b) \times [b^{p-1} + c_1\,p^k + b^{p-1} + c_{p-2,1}\,p^k + \cdots + b^{p-1} + c_{1,p-2}\,p^k + b^{p-1}] \\ &= c\,p^k \times [p \times b^{p-1} + p^k\{c_1 + c_{p-2,1} + \cdots + c_{1,p-2} + c_{1,p-2}\}] \\ &= c\,p^{k+1}[b^{p-1} + c_0 \times p^{k-1}] \\ &\Rightarrow a^p \equiv b^p\,(mod\,p^{k+1}) \end{aligned}$$

Taking all the constants $(c_1 + c_{p-2,1} + \cdots + c_{1,p-2} + c_{1,p-2})$ as a single constant c_0 and using $a \equiv b\,(mod\,p^k) \Rightarrow a - b = c\,p^k$ we complete the above proof.

3 Classical Ciphers and Perfect Ciphers

3.1 Polyalphabetic Ciphers

A Polyalphabetic Cipher is a substitution cipher in which a particular character in the plaintext can be mapped into more than one possible character in the Ciphertext. This makes the cryptanalysis of polyalphabetic Ciphers more difficult when compared to the monoalphabetic ciphers. Some popular examples of polyalphabetic ciphers include the Vigenere Cipher, Hill Cipher etc

3.1.1 A Closer Look into the Vigenere Cipher

The Vigenere Cipher was devised by Blaise de Vigenère in the 16th century. It is a type of polyalphabetic cipher that uses two or more cipher alphabets for data encryption. In other words, the letters in the Vignere Cipher are shifted by different amounts, which is normally done using a keyword or a phrase.

Example:

Consider the plaintext shown in Figure 1.

t	h	i	s	c	r	y	p	t	o	s	y	s	t	e	m	i	s	n	o	t	s	e	c	u	r	e

Figure 1: Sample Plaintext

Let m be the number of characters in the key. Here, we shall consider the key as (2, 8, 15, 7, 4, 17) i.e. m = 6.

First, we shall first convert the given plaintext into its residue modulo 26. Then we write them in groups of 6 (in this case since key size is 6) and add the keyword. From this, we obtain the corresponding cipher text. Here we assume a = 0, b = 1,.... , z = 25. This method is illustrated in Figure 2

Plaintext	t	h	i	s	c	r	y	p	t	o	s	y	s	t	e	m	i	s	n	o	t	s	e	c	u	r	e
Value	19	7	8	18	2	17	24	15	19	14	18	24	18	19	4	12	8	18	13	14	19	18	4	2	20	17	4
Key	2	8	15	7	4	17	2	8	15	7	4	17	2	8	15	7	4	17	2	8	15	7	4	17	2	8	15
Residue %26	21	15	23	25	6	8	0	23	8	21	22	15	20	1	19	19	12	9	15	22	8	25	8	19	22	25	19
Ciphertext	v	p	x	z	g	i	a	x	i	v	w	p	u	a	t	t	m	j	p	w	i	z	i	t	w	z	t

Figure 2: Vigenere Cipher Mapping

As shown in Figure 2, the letter 't' in the plaintext maps to more than one letter in the ciphertext. Similarly, the letter 's' also maps to more than one character in the ciphertext and so on.

Key Size: Let us assume that the length of the keyword is 'm'. Therefore, there are 26^m possible keys. Now, if m = 5, then 26^5 = 1.1 x 10^7 possible keys, which is **large enough** to **preclude exhaustive key search** by hand. However, we shall later see in the next lecture that there will be a systematic methodology to break the Vigenere Cipher.

3.2 Cryptanalysis

Cryptanalysis is the art of decrypting or analysing codes, ciphers or encrypted text. In this section, we shall discuss some techniques of Cryptanalysis. To do so, we shall first differentiate between different attack models on cryptosystems. The *attack models* specify the information available to the adversary when he launches the attack. The different types of attack models discussed are as follows:

1. **ciphertext only:** The opponent possesses a string of ciphertext, $\mathbf{y}$.
2. **known plaintext:** The opponent posses a string of plaintext, $\mathbf{x}$, and the corresponding ciphertext, $\mathbf{y}$.
3. **chosen plaintext:** The opponent has temporary access to the plaintext function. Hence, he can choose a plaintext, $\mathbf{x}$, and can obtain the corresponding ciphertext string, $\mathbf{y}$.
4. **chosen ciphertext:** The opponent has temporary access to the decryption function. Hence, he can choose a ciphertext string, $\mathbf{y}$, and construct the corresponding plaintext string, $\mathbf{x}$.

In each case, the objective is to obtain the key. Arranging them in the increasing order of their strength, we have:
Ciphertext only < Known plaintext <Chosen Plaintext < Chosen Ciphertext

Several techniques of Cryptanalysis use statistical properties of the English language. These statistical properties are:

1. Probabilities of occurrences of 26 letters:
   (a) E has a probability of about 0.12 (12%)
   (b) T, A, O, I, N, S, H, R each between 0.06 and 0.09
   (c) D, L each around 0.04
   (d) C, U, M, W, F, G, Y, P, B each between 0.015 and 0.028
   (e) V, K, J, X, Q, Z each less than 0.01
2. 30 common diagrams
   (a) A diagram is a sequence of two consecutive letters.
   (b) TH, HE, IN, ER, AN, RE.... (in decreasing order)
3. 12 common triagrams
   (a) A triagram is a sequence of three consecutive letters.
   (b) THE, ING, AND, HER, ERE... (in decreasing order)

3.2.1 Cryptanalysis of Monoalphabetic Ciphers

Cryptanalysis of Monoalphabetic Ciphers is done using the **ciphertext only** attack model i.e. by using the letter frequencies of the English language.

Let's consider a Shift Cipher with Shift = 2. Then A, B, C, D, E....Z maps to C, D, E, F, G.....B respectively. Knowing that the letter 'E' is the most occurring alphabet, the adversary shall take the ciphertext and start analyzing it to find that 'G' has the maximum number of occurrences (since 'E' gets mapped to 'G' and 'E' has the highest probability of occurrence), thereby tentatively concluding the shift value used.

3.2.2 Cryptanalysis of Affine Cipher

We shall explain this with the help of an example covered in the class.

Example: Consider the ciphertext obtained from an Affine Cipher: [2]
FMXVEDKAPHFERBNDKRXRSREFMORUDSDKDVSHVUFEDKAPRKDLYEVLRHHRH

Step 1: We shall perform a frequency analysis of the ciphertext.
We observe that there are 57 characters of ciphertext. The most frequent ciphertext characters are R (8 occurrences), D (7 occurrences), E, H, K (5 occurrences each), and F, S, V (4 occurrences each).
Step 2: First Guess: We hypothesize that R is the encryption of 'e' and D is the encryption of 't', since 'e' and 't' are the two most common letters used
Step 3: Expressing numerically, we have $e_K(4) = 17$ and $e_K(19) = 3$.
This gives us two linear equations in two unknowns:

$$4a + b = 17$$
$$19a + b = 3$$

This system has a unique solution with a = 6 and b = 19 in Z_{26}. But, this is an illegal key, since gcd(a, 26) = 2, which is greater than 1. Therefore our first guess is incorrect.
Step 4: Our next guess is R is the encryption of 'e' and E is the encryption of 't'. Proceeding as the previous step, we get, a = 13, which is again not possible.
Step 5: We again take a guess with R is the encryption of 'e' and H is the encryption of 't'. This yields, a = 8, which is clearly impossible.
Step 6: We suppose R is the encryption of 'e' and K is the encryption of 't'. This produces a = 3, b = 5. This is a legal key.
Step 7: On performing the required operations, we obtain $d_K(y) = 9y - 19$ and the given ciphertext decrypts to

algorithmsarequitegeneraldefinitionsofarithmeticprocesses

4 Conclusion

In this lecture, we successfully discussed Fermat's Little Theorem with suitable examples, Classical Cryptosystems with a closer look into the Vignere Cipher

and also discussed the strategies used for the Cryptanalysis of Monoalphabetic Ciphers as well as the Affine Cipher.

References

[1] Johannes Buchmann. *Introduction to cryptography*. Springer Science & Business Media, 2013.

[2] Douglas Robert Stinson and Maura Paterson. *Cryptography: theory and practice*. CRC press, 2018.

[3] Sadanand G Telang and MG Nadkarni. *Number theory*. Tata McGraw-Hill, 1996.