Hacettepe University
Department of Computer Engineering

BBM203 EXPERIMENT 4:
DATA STRUCTURES

Advisors	R.A. Hüseyin Temuçin
Subject	Data Structures, Polynomial Operations
Submission Date	08.12.2014
Due Date	26.12.2014
Programming Language	ANSI C
Host Platform	Windows, MinGW

1. Introduction

In this experiment you are expected to implement a simplified Shamir Secret Sharing Scheme [6] using C programming language. In the experiment, you will use and implement polynomial operations adding and multiplication as a part of Shamir Secret Sharing Scheme.

1.1 Background Information

1.1.1 Information Systems and Security

Organizations (military, education, business etc.) works on information. They store their information (data), process them depend on their purpose and store processed data again. This kind of data operations are handled by information systems. An information system is an integrated set of components for collecting, storing, processing, and communicating information [6]. In today world, the computers constitute the information systems.

Because always there are attackers who wants to get into system without authorization and obtain information that is not permitted, the security of information systems are crucial. The protection of systems and information it includes is the work of information security. **Information security** means protecting information and information systems from unauthorized access, use, disclosure, disruption, modification or destruction [2].

1.1.2 Cryptography

Information security services must transform data to unrecognizable form to protect from unauthorized users, especially the data communication networks. Information security uses cryptography to transform usable information into a form that renders it unusable by anyone other than an authorized user [2]. Cryptography is a mathematical approach to protect information (data). The cryptography provides encryption/decryption algorithms and ways to systems.

1.1.2.1 Basic Terminology in Cryptography[5]

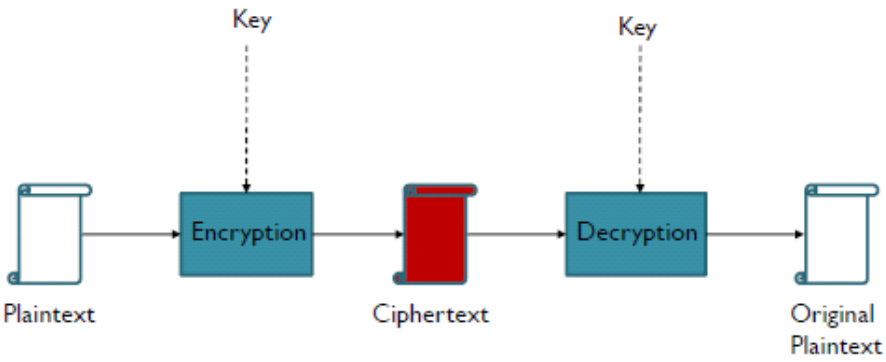
Plaintext: the data to be concealed.

Key: the special knowledge shared between communicating parties

Encryption (encipherment): the process of transforming information (plaintext) using an algorithm (cipher) to make it unreadable to anyone except those possessing special knowledge

Ciphertext: the result of encryption on the plaintext

Decryption (decipherment): the process of making the encrypted information readable again



1.1.3 Secret Sharing Scheme (Threshold)

Some systems cannot be controlled individually. For example, a bank safe is controlled by many authorized personnel, like five managers and it can only be opened with any three of managers' keys. Or a bank account that holds to a company with three trading partner can operated only with all partners keys.

Secret Sharing Schemes handle these kinds of multi-party sharing problems. Secret sharing distributes a secret to a group of party and the secret can only be constructed with a sub-group of parties. If a secret is shared in n-parties and it can be constructed with t-parties, it called as A (t, n) threshold.

More information can be found by searching references about these terms.

1.2 Shamir Secret Sharing Scheme

Shamir Secret Sharing Scheme (SSS) algorithm is an approach to secret sharing problem. It is found by Adil Shamir [4] and widely used in modern cryptography. This approach uses polynomial operations to share a secret in parties.

In Shamir Secret Sharing Scheme, the parties use a polynomial with variable degrees. The polynomial is created randomly. The constant of polynomial (the term with degree 0) is the main key of system.

All parties' keys are the polynomial's result for party id. For example if the polynomial is $p(x) = 39 + 2x + 3x^2$, the key of 1st party's key will be:

*p(1) = 39 + 2*1 + 3 = 44*

The polynomial, so the main key is the secret information of system. All parties know only their keys.

The steps of Shamir's threshold process are like below[1]:

Shamir's (t, n)-threshold Scheme

- Preparing and distributing the keys:
- The dealer chooses prime p such that $p \geq n+1$, $K \varepsilon Z_p$;
- generates distinct, random, non-zero $x_i \varepsilon Z_p$, i=1,...,n;
- generates random $a_i \varepsilon Z_p$, i=1, 2,..., t – 1;
- $a_0 = K$, the secret;
- $f(x) = \Sigma_{i=0 \text{ to } t-1} a_i x_i \bmod p = a_0 + a_1 x + ... + a_{t-2} x_{t-2} + a_{t-1} x_{t-1} \bmod p$
- ith person's share is (xi, f(xi)).

Combining t keys and reconstructing the secret K

- $I_i(x) = \prod_{j=1 \text{ to } t,\, j \neq i} (x - xj) / (xi - xj) \bmod p$
- $f(x) = \Sigma_{i=0 \text{ to } t} I_i(x)\, y_i \bmod p$
- $f(0) = K$

[1]The theoretical information about Shamir Secret Sharing is taken from Ahmet Burak Can's Computer and Network Security Lecture Notes [5]

Example: Shamir's (3, 6)-threshold Scheme[1]

- n=6, t=3, K=1234,
- We randomly obtain 2 numbers: a_1=166, a_2=94
- a_0 = K = 1234
- $f(x) = 1234 + 166x + 94x^2$
- We construct six points:
- To reconstruct the key any 3 points will be enough. Assume that we have these keys:

(1,1494);(2,1942);(3, 2578); (4, 3402); (5, 4414); (6, 5614)

- To reconstruct the key any 3 points will be enough. Assume that we have these keys:

$(x_0, y_0) = (2, 1942); (x_1, y_1) = (4, 3402); (x_2, y_2) = (5, 4414);$

- From these 3 keys, we compute I_i values:

$$\ell_0 = \frac{x-x_1}{x_0-x_1}\cdot\frac{x-x_2}{x_0-x_2} = \frac{x-4}{2-4}\cdot\frac{x-5}{2-5} = \frac{1}{6}x^2 - 1\frac{1}{2}x + 3\frac{1}{3}$$

$$\ell_1 = \frac{x-x_0}{x_1-x_0}\cdot\frac{x-x_2}{x_1-x_2} = \frac{x-2}{4-2}\cdot\frac{x-5}{4-5} = -\frac{1}{2}x^2 + 3\frac{1}{2}x - 5$$

$$\ell_2 = \frac{x-x_0}{x_2-x_0}\cdot\frac{x-x_1}{x_2-x_1} = \frac{x-2}{5-2}\cdot\frac{x-4}{5-4} = \frac{1}{3}x^2 - 2x + 2\frac{2}{3}$$

- Then, we compute f(x):

$$\begin{aligned} f(x) &= \sum_{j=0}^{2} y_j \cdot \ell_j(x) \\ &= 1942\cdot\left(\frac{1}{6}x^2 - 1\frac{1}{2}x + 3\frac{1}{3}\right) + 3402\cdot\left(-\frac{1}{2}x^2 + 3\frac{1}{2}x - 5\right) + 4414\cdot\left(\frac{1}{3}x^2 - 2x + 2\frac{2}{3}\right) \\ &= 1234 + 166x + 94x^2 \end{aligned}$$

[1]Example is taken from Ahmet Burak Can's Computer and Network Security Lecture Notes [5]

2. Problem

This experiment consist a simplified Shamir's Threshold encryption. Your implementation will be named as s*hamir* and must operate two functions:

1. It will calculate the main key for a given party and crypt a plain text with main key. This function will run with parameter "–e"
2. It will find sub-polynomials of parties and will write them to text file. This function will run with parameter "-l"

2.1 Encryption

In encryption process the program will take three argument text file names: [*keys_file*], [*plaintext_file*] and [*cipher_text_file*]. The program will be run from command line for encryption mode in format below:

shamir -e [*keys_file*] [*plaintext_file*] [*cipher_text_file*]

The party keys file will include party's (*x, f(x)*) couples. The *x* value will be party's id and *f(x)* value will be the party's key. So every line in this text file will consist a party's id and party's key in "*party id: party key"* format. For given example the party list text file will be like this:

```
1 : 1494

2 : 1942

3 : 2578

4 : 3402

5 : 4414

6 : 5614
```

The party key file will include just enough point count to calculate the polynomial. If the file includes six points, our polynomial will be 5^{th} degree. For A (n, t) threshold, the text file will include t parties. Then the program will calculate I_i (sub-polynomial) for each group as given in example, and the main key using $\Sigma_{i=0 \text{ to } t} I_i$ formula.

[*plaintext_file*] file will include plain text that will be encrypted. The encryption process will be simply xor (exclusive-OR) of each character's integer value with key. The result of encryption will be another character. For example if we want to encrypt 'A' character and if the key is 44:

'A' = 65

'A' xor 44 = 65 xor 44 = 109 = 'm'

So the encrypted value will be 'm'. In C programming language you can use simply "^" bitwise XOR operator for encryption. The output "cipher" text file will include key and encrypted text.

For example if our plain.txt contains:

```
Hacettepe University Computer Science and Engineering
```

If the key is 44, the cipher.txt will contain:

```
Key : 44

dMOIXXI\IyBEZI^_EXUoCA\YXI^□ÐEIBOIMBHiBKEBII^EBK&Ó
```

The encrypted text will be written to text file named with 3rd argument. The program can be run from command line with encryption mode as :

The program's output will be [*cipher_text_file*], which contains encrypted (XORed) form of plain.txt

2.2 Listing sub-polynomials

In listing process, the program will take two argument text file names. [*keys_file*] and [*sub_polynomials_file*]. The program will be run from command line for listing sub-polynomials mode in format below:

shamir -l [*keys_file*] [*subpolynomials_file*]

In listing mode the program will calculate sub-polynomial of each group and will write sub-polynomials to text file named as 2nd argument in order with *party id: party polynomial*. For example, for the polynomial $1234 + 166x + 94x^{2,}$ if the keys file is:

```
1 : 1494

2 : 1942

3 : 2578
```

The sub-polynomials file will be:

```
1 : 0.17 x^2 - 3.50 x + 3.33

2 : -0.50 x^2 + 3.50 x - 5

3 : 0.33 x^2 - 2 x + 2.67
```

Float coefficient will be written with two precision in output file

3. Restrictions

In this implementation some features of Shamir threshold are restricted:

1. The key length is limited with 8bit. So the key value cannot be bigger than 255
2. There is no level mechanism between users. All users have same authorizations

4. Evaluation

4.1. Required Files

You should create and submit a ZIP archive in the following structure for evaluation. An invalid structured archive will cause you partial or full score loss.

Directory	**Files**	**Description**
source	*.c, *.h	Program source/header files
report	*.pdf	Your report (Only pdf format is accepted)

4.2 Reports

The report must adhere to the Hacettepe University Computer Science Department Report Writing Guidelines. Submissions with poorly written software can't be expected to receive a high score for the report.

5. Notes

- Save all your work until the experiment is graded.
- The assignment must be original, INDIVIDUAL work. Shared source codes will be considered as cheating. Also the students who share their works will be punished in the same way.
- You can ask your questions via course's news group (*dersler.bbm203* on news.cs.hacettepe.edu.tr). Discussion of the solution on the newsgroups will be considered "group work", which means "cheating".
- You are expected to follow the course's new group and you will be held responsible for the announcements made there.
- Cheaters will not be tolerated, and anyone caught will be reported to university authorities.

6.References

1. http://www.britannica.com/EBchecked/topic/287895/information-system
2. http://en.wikipedia.org/wiki/Information_security
3. http://www.wisegeek.com/what-is-information-security.htm
4. http://en.wikipedia.org/wiki/Adi_Shamir
5. http://web.cs.hacettepe.edu.tr/~abc/teaching/bil438/index.php
6. http://en.wikipedia.org/wiki/Shamir%27s_Secret_Sharing