ROHIT S.R SHIVAPRASAD N.K VINAYAK J

CYPHER XSUITEX

MINI PROJECT REPORT NOV.2018

Contents

Introduction

In cryptography, a **cipher** (or *cypher*) is an algorithm for performing encryption or decryption—a series of well-defined steps that can be followed as a procedure. An alternative, less common term is encipherment. To encipher or encode is to convert information into cipher or code. In common parlance, "cipher" is synonymous with "code", as they are both a set of steps that encrypt a message; however, the concepts are distinct in cryptography, especially classical cryptography.

In non-technical usage, a "(secret) code" typically means a "cipher". Within technical discussions, however, the words "code" and "cipher" refer to two different concepts. Codes work at the level of meaning—that is, words or phrases are converted into something else and this chunking generally shortens the message.

An example of this is the *Commercial Telegraph Code* which was used to shorten long telegraph messages which resulted from entering into commercial contracts using exchanges of Telegrams [1].

Cryptography has been through numerous phases of evolution. Early ciphers in cryptography were designed to allow encryption and decryption to take place by hand, while those which are developed and used today are only possible due to the high computational performance of modern machines.

Cryptography prior to the modern age was effectively synonymous with encryption, the conversion of information from a readable state to apparent nonsense. The originator of an encrypted message shared the decoding technique needed to recover the original information only with intended recipients, thereby precluding unwanted persons from doing the same. Since the development of rotor cipher machines in World War I and the advent of computers in World War II, the methods used to carry out cryptology have become increasingly complex and its application more widespread. [2]

For our project, we are showcasing three of the most historically important cipher algorithms as an implementation in C language. The three ciphers used in our project are the **Caesar's Cipher**, **Atbash Cipher** and the **Vigenère Cipher**. We will also present the modern cryptography algorithms such as the **RSA** and the **Diffie-Hellman** Algorithm.

Caesar Cipher

In cryptography, a **Caesar cipher**, also known as *Caesar's cipher*, the shift cipher, Caesar's code or Caesar shift, is one of the simplest and most widely known encryption techniques. It is a type of substitution cipher in which each letter in the plaintext is replaced by a letter some fixed number of positions down the alphabet. For example, with a left shift of 3, D would be replaced by A, E would become B, and so on. The method is named after Julius Caesar, who used it in his private correspondence

The Caesar cipher is named after Julius Caesar, who, according to Suetonius, used it with a shift of three to protect messages of military significance. While Caesar's was the first recorded use of this scheme, other substitution ciphers are known to have been used earlier

"If he had anything confidential to say, he wrote it in cipher, that is, by so changing the order of the letters of the alphabet, that not a word could be made out. If anyone wishes to decipher these, and get at their meaning, he must substitute the fourth letter of the alphabet, namely D, for A, and so with the others."

— *Suetonius, Life of Julius Caesar 56*

Example

The transformation can be represented by aligning two alphabets; the cipher alphabet is the plain alphabet rotated left or right by some number of positions. For instance, here is a Caesar cipher using a left rotation of three places, equivalent to a right shift of 23 (the shift parameter is used as the key):

Plain: ABCDEFGHIJKLMNOPQRSTUVWXYZ

Cipher: XYZABCDEFGHIJKLMNOPQRSTUVW

When encrypting, a person looks up each letter of the message in the "plain" line and writes down the corresponding letter in the "cipher" line.

Plaintext: THE QUICK BROWN FOX JUMPS OVER THE LAZY DOG

Ciphertext: QEB NRFZH YOLTK CLU GRJMP LSBO QEB IXWV ALD [3]

Vigenère cipher

The Vigenère cipher is a method of encrypting alphabetic text by using a series of interwoven Caesar ciphers, based on the letters of a keyword. It is a form of polyalphabetic substitution.

The cipher is easy to understand and implement, but it resisted all attempts to break it for three centuries, which earned it the description le chiffre indéchiffrable (French for 'the indecipherable cipher'). Many people have tried to implement encryption schemes that are essentially Vigenère ciphers. In 1863, Friedrich Kasiski was the first to publish a general method of deciphering Vigenère ciphers.

The Vigenère cipher was originally described by Giovan Battista Bellaso in his 1553 book *La cifra del. Sig. Giovan Battista Bellaso*, but the scheme was later misattributed to Blaise de Vigenère (1523–1596) in the 19th century and so acquired its present name.

The Vigenère cipher gained a reputation for being exceptionally strong. Noted author and mathematician Charles Lutwidge Dodgson (Lewis Carroll) called the Vigenère cipher unbreakable in his 1868 piece "The Alphabet Cipher" in a children's magazine. In 1917, Scientific American described the Vigenère cipher as "impossible of translation". That reputation was not deserved. Charles Babbage is known to have broken a variant of the cipher as early as 1854 but failed to publish his work. Kasiski entirely broke the cipher and published the technique in the 19th century, but even earlier, some skilled cryptanalysts could occasionally break the cipher in the 16th century [4]

Example:

Input: Plaintext: GEEKSFORGEEKS
Keyword: AYUSH
Output: Ciphertext: GCYCZFMLYLEIM
For generating key, the given keyword is repeated
in a circular manner until it matches the length of
the plain text.
The keyword "AYUSH" generates the key "AYUSHAYUSHAYU"
The plain text is then encrypted using the process
explained below.

Encryption:

The first letter of the plaintext, G is paired with A, the first letter of the key.
So, use row G and column A of the Vigenère square, namely G.
Similarly, for the second letter of the plaintext, the second letter of the key is used, the letter at row E and column Y is C. The rest of the plaintext is enciphered in a similar fashion.

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

Decryption

Decryption is performed by going to the row in the table corresponding to the key, finding the position of the ciphertext letter in this row, and then using the column's label as the plaintext. For example, in row A (from AYUSH), the ciphertext G appears in column G, which is the first plaintext letter.

Next, we go to row Y (from AYUSH), locate the ciphertext C which is found in column E, thus E is the second plaintext letter [5].

Atbash Cipher:

The Atbash Cipher is a particular monoalphabetic cipher formed by taking the alphabet. And mapping it to its reverse, so that the first letter becomes the last letter, the second letter becomes the second to last letter, and so on. For example, the Latin alphabet would work like this:

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

Due to the fact that there is only one way to perform this, the Atbash cipher provides no communications security, as it lacks any sort of key. If multiple collating orders are available, which one was used in encryption can be used as a key, but this does not provide significantly more security.[6]

RSA Encryption:

RSA (**Rivest–Shamir–Adleman**) is one of the first public-key cryptosystems and is widely used for secure data transmission. In such a cryptosystem, the encryption key is public and it is different from the decryption key which is kept secret (private). In RSA, this asymmetry is based on the practical difficulty of the factorization of the product of two large prime numbers, the "factoring problem". The acronym RSA is made of the initial letters of the surnames of Ron Rivest, Adi Shamir, and Leonard Adleman, who first publicly described the algorithm in 1978. Clifford Cocks, an English mathematician working for the British intelligence agency Government Communications Headquarters (GCHQ), had developed an equivalent system in 1973, but this was not declassified until 1997.[7]

A user of RSA creates and then publishes a public key based on two large prime numbers, along with an auxiliary value. The prime numbers must be kept secret. Anyone can use the public key to encrypt a message, but with currently published methods, and if the public key is large enough, only someone with knowledge of the prime numbers can decode the message feasibly. Breaking RSA encryption is known as the RSA problem. Whether it is as difficult as the factoring problem remains an open question

RSA is an algorithm used by modern computers to encrypt and decrypt messages. It is an asymmetric cryptographic algorithm. Asymmetric means that there are two different keys. This is also called public key cryptography, because one of the keys can be given to anyone. The other key must be kept private. The algorithm is based on the fact that finding the factors of a large composite number is difficult: when the integers are prime numbers, the problem is called prime factorization. It is also a key pair (public and private key) generator. [8]

Operation:

1. Choose two different large random prime numbers p and q.
2. Calculate n = pq. (n is the modulus for p and q)
3. Calculate the totient: Ø(n) = (p-1) (q-1)
4. Choose an integer such that **e** such that 1<**e**<Ø(n) and **e** is the coprime to Ø(n). (e is released as public key exponent).

5. Compute d to satisfy congruence relation.

Encryption:

Consider two entities named Bob and Alice. Alice gives her public key (**n** and **e)** to bob. Now Bob wants to send message **M** to Alice.

First, he turns **M** into a number **m** smaller than **n** by using an agreed-upon reversible protocol known as a **padding scheme**. He then computes the ciphertext **c** corresponding to

$\mathbf{c} = m^e$

This can be done quickly using the method of **exponentiation by squaring**. Bob then sends to Alice.

Decryption:

Alice can recover **m** from **c** by using her private key **d** in the following procedure.

$\mathbf{m} = c^d$

Now she can recover the original distinct prime numbers, applying the **Chinese remainder theorem** to these two congruences yields.

$\mathbf{m}^{ed} = m \bmod pq.$

Thus,

$\mathbf{c}^d = m \bmod n.$

Padding schemes:

When used in practice, RSA must be combined with some form of padding scheme, so that no values of M result in insecure ciphertexts. RSA used without padding may have some problems:

1. The values $m = 0$ or $m = 1$ always produce ciphertexts equal to 0 or 1 respectively, due to the properties of exponentiation.
2. When encrypting with small encryption exponents (e.g., $e = 3$) and small values of the m, the (non-modular) result of may be strictly less than the modulus n. In this case, ciphertexts may be easily decrypted by taking the eth root of the ciphertext with no regard to the modulus.
3. RSA encryption is a deterministic encryption algorithm. It has no random component. Therefore, an attacker can successfully launch a chosen plaintext attack against the cryptosystem. They can make a dictionary by encrypting likely plaintexts under the public key, and storing the resulting ciphertexts.
4. The attacker can then observe the communication channel. As soon as they see ciphertexts that match the ones in their dictionary, the attackers can then use this dictionary in order to learn the content of the message.

Chinese Remainder Theorem:

The **Chinese remainder theorem** is a theorem from number theory. It is about congruence. The original form was:

> *How many soldiers are there in Han Xin's army? – If you let them parade in rows of 3 soldiers, two soldiers will be left. If you let them parade in rows of 5, 3 will be left, and in rows of 7, 2 will be left.*

The theorem says that there will be a solution to this question if there's no common factor between the row sizes. Using the original example, that is that no number divides both 3 and 7, both 3 and 5, nor both 5 and 7 (except, of course, 1). They're all coprime. [9]

Diffie-Hellman Key Exchange

Diffie–Hellman key exchange (DH) is a method of securely exchanging cryptographic keys over a public channel and was one of the first public-key protocols as originally conceptualized by Ralph Merkle and named after Whitfield Diffie and Martin Hellman. DH is one of the earliest practical examples of public key exchange implemented within the field of cryptography.

Traditionally, secure encrypted communication between two parties required that they first exchange keys by some secure physical channel, such as paper key lists transported by a trusted courier. The Diffie–Hellman key exchange method allows two parties that have no prior knowledge of each other to jointly establish a shared secret key over an insecure channel. This key can then be used to encrypt subsequent communications using a symmetric key cipher.[9]

Elliptic Curve Cryptography (ECC) is an approach to public-key cryptography, based on the algebraic structure of elliptic curves over finite fields. ECC requires a smaller key as compared to non-ECC cryptography to provide equivalent security (a 256-bit ECC security have an equivalent security attained by 3072-bit RSA cryptography).

The Diffie-Hellman algorithm is being used to establish a shared secret that can be used for secret
communications while exchanging data over a public network using the elliptic curve to generate points and get the secret key using the parameters.

- For the sake of simplicity and practical implementation of the algorithm, we will consider only 4 variables one prime P and G (a primitive root of P) and two private values a and b.
- P and G are both publicly available numbers. Users (say Alice and Bob) pick private values a and b and they generate a key and exchange it publicly; the opposite person received the key and from that generates a secret key after which they have the same secret key to encrypt.

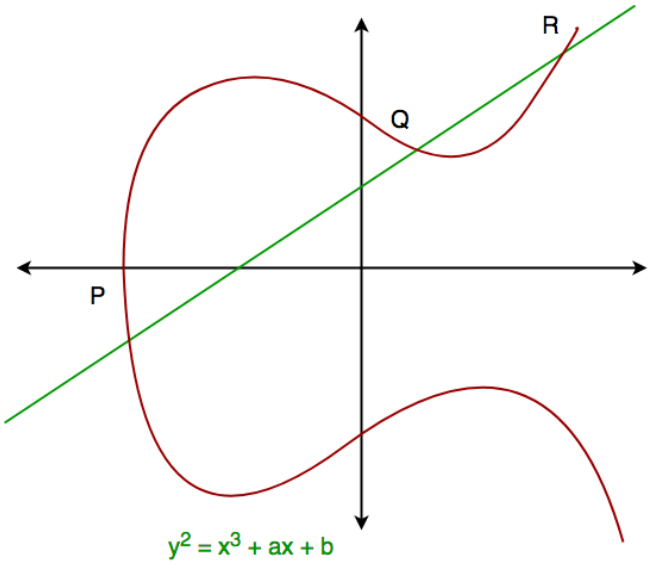
For a better understanding of Elliptic Curve Cryptography, it is very important to understand the basics of Elliptic Curve. An elliptic curve is a planar algebraic curve defined by an equation of the form.

$y^2=x^3+ax+b$

where ‘a’ is the co-efficient of x and ‘b’ is the constant of the equation.

The curve is non-singular; that is its graph has no cusps or self-intersections (when the characteristic of the
co-efficient field is equal to 2 or 3).

In general, an elliptic curve looks like as shown below. Elliptic curves could intersect atmost 3 points when a straight line is drawn intersecting the curve. As we can see that elliptic curve is symmetric about the x-axis, this property plays a key role in the algorithm.



This approach uses six tuple {P, a, b, G, n, h}

P = Field that the curve is define over

G = Generator point

a, b = Values define the curve

h = Co- factor

n = Prime order of G

Step by Step Explanation:

ALICE	BOB
Public Keys available = P, G	Public Keys available = P, G
Private Key Selected = a	Private Key Selected = b
Key generated = $x = G^a mod P$	Key generated = $y = G^b mod P$
Exchange of generated keys takes place	
Key received = y	key received = x
Generated Secret Key = $k_a = y^a mod P$	Generated Secret Key = $k_b = x^b mod P$
Algebraically it can be shown that $k_a = k_b$	
Users now have a symmetric secret key to encrypt	

Example:

Step 1: Alice and Bob get public numbers P = 23, G = 9

Step 2: Alice selected a private key a = 4 and
Bob selected a private key b = 3

Step 3: Alice and Bob compute public values
Alice: $x = (9^4 \bmod 23) = (6561 \bmod 23) = 6$
Bob: $y = (9^3 \bmod 23) = (729 \bmod 23) = 16$

Step 4: Alice and Bob exchange public numbers

Step 5: Alice receives public key y =16 and
Bob receives public key x = 6

Step 6: Alice and Bob compute symmetric keys
Alice: $ka = y^a \bmod p = 65536 \bmod 23 = 9$
Bob: $kb = x^b \bmod p = 216 \bmod 23 = 9$

Step 7: 9 is the shared secret.

Methodology:

In our project we have used several methodologies.

For the implementation of Caesar Cipher, Atbash Cipher and Vigenère Cipher, we used **string.h** for string manipulation functions and arrays as well.

For modern cryptographic functions, we used **math.h**.

Algorithm for Caesar Cipher:

Input:

1. A String of lower case letters, called Text.
2. An Integer between 0-25 denoting the required shift.

Procedure:

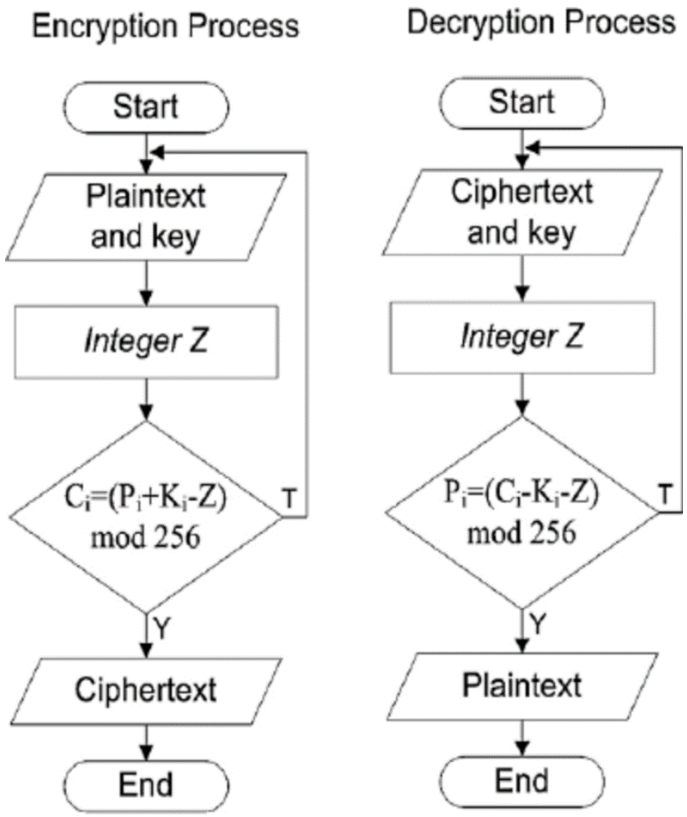
1. Traverse the given text one character at a time.
2. For each character, transform the given character as per the rule, depending on whether we're encrypting or decrypting the text.
3. Return the new string generated.

Algorithm for Atbash cipher

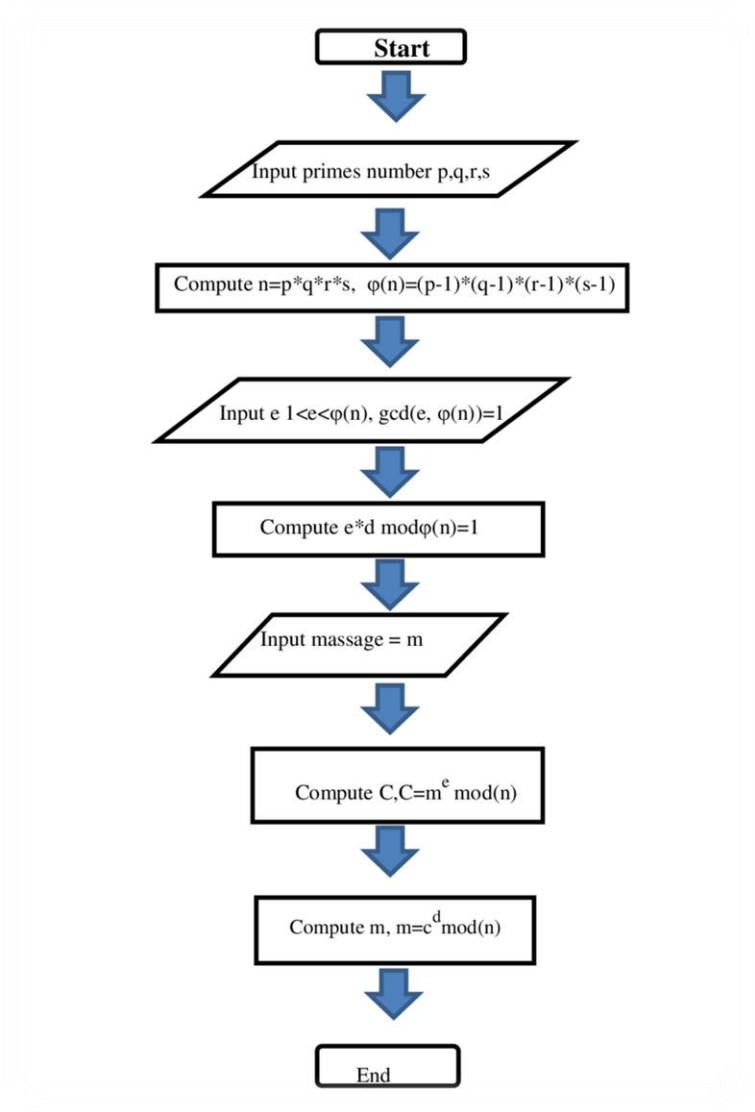
The Atbash cipher is essentially a substitution cipher with a fixed key, if you know the cipher is Atbash, then no additional information is needed to decrypt the message. The substitution key is:

ABCDEFGHIJKLMNOPQRSTUVWXYZ
ZYXWVUTSRQPONMLKJIHGFEDCBA

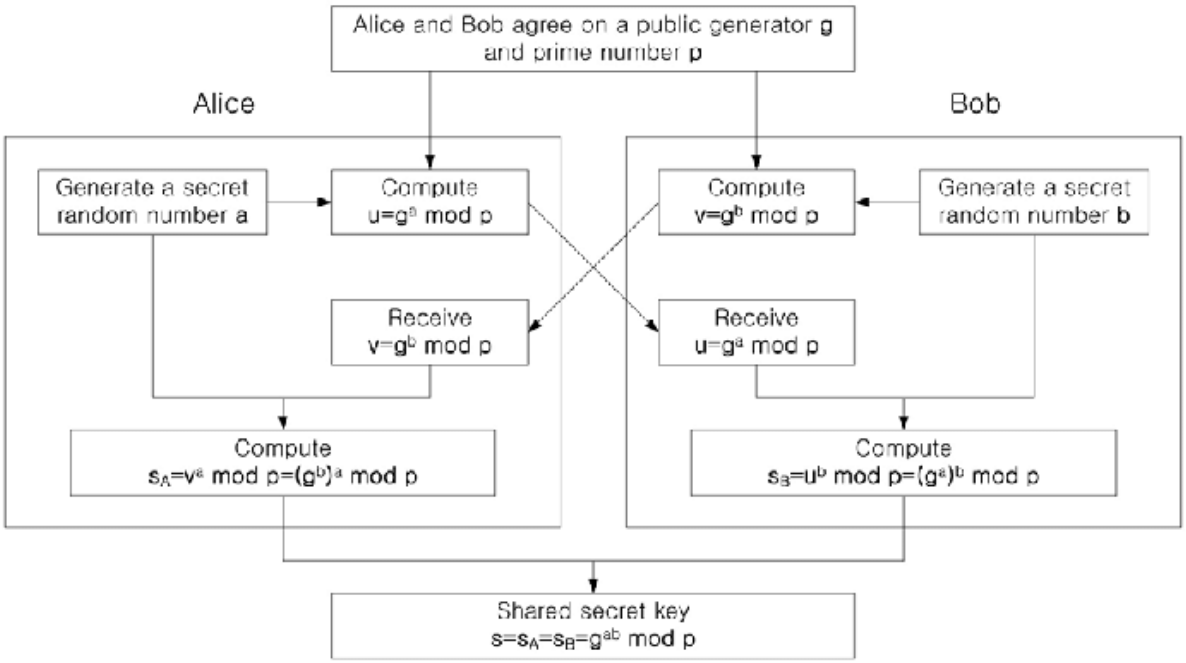
Flowchart for Vigenère Cipher:



RSA Flowchart:



Diffie-Hellman Key Exchange Flowchart:



Applications

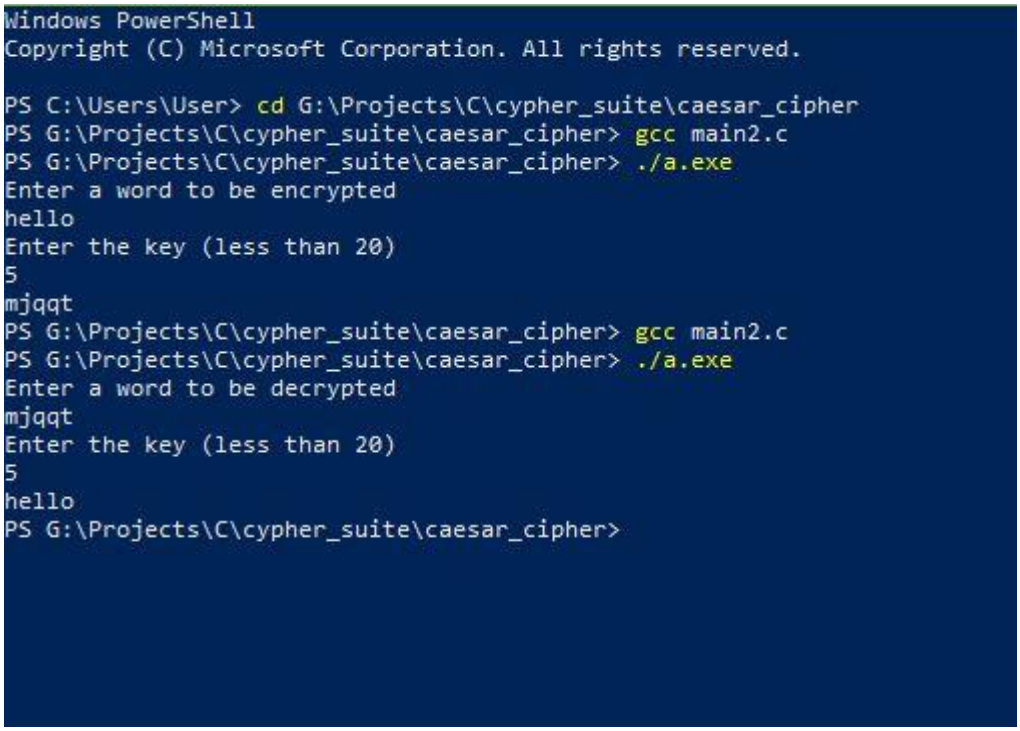
The most obvious use of cryptography, and the one that all of us use frequently, is encrypting communications between us and another system. This is most commonly used for communicating between a client program and a server. Examples are a web browser and web server, or email client and email server. When the internet was developed it was a small academic and government community, and misuse was rare. Most systems communicated in the clear (without encryption), so anyone who intercepted network traffic could capture communications and passwords. Modern switched networks make interception harder, but some cases – for example, public WIFI – still allow it. To make the internet more secure, most communication protocols have adopted encryption. Many older protocols have been dropped in favour of newer, encrypted replacements.

Email is one area where encryption is not widely in use. When email moves from server to server, and from server to you, it is encrypted. On the mail server and on your system, however, an administrator can read it. There are options to implement "end-to-end" encryption for email but email systems are complex and these options are complex. Truly secure messaging systems – where only the sender and receiver can read the message – are those where encryption has been built in from the start. WhatsApp is good; Signal is better.

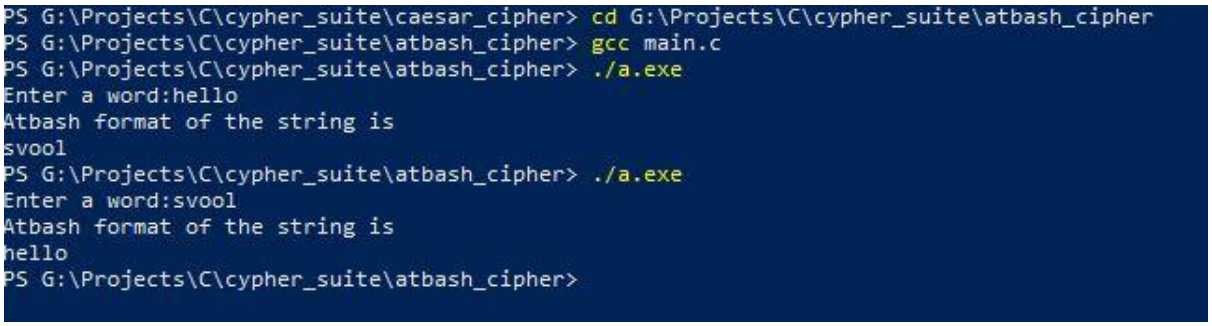
A more notable use of encryption is to encrypt the entire drive, and require correct credentials to access it. UCL has recently implemented Microsoft's Bit locker on Desktop machines, and this means that without the user logging in the data on the drive is completely opaque. If someone took the drive and tried to read it, they would not be able to access any data. This has the occasional side effect of locking the system, so some UCL readers may have had to request the recovery key.

Results

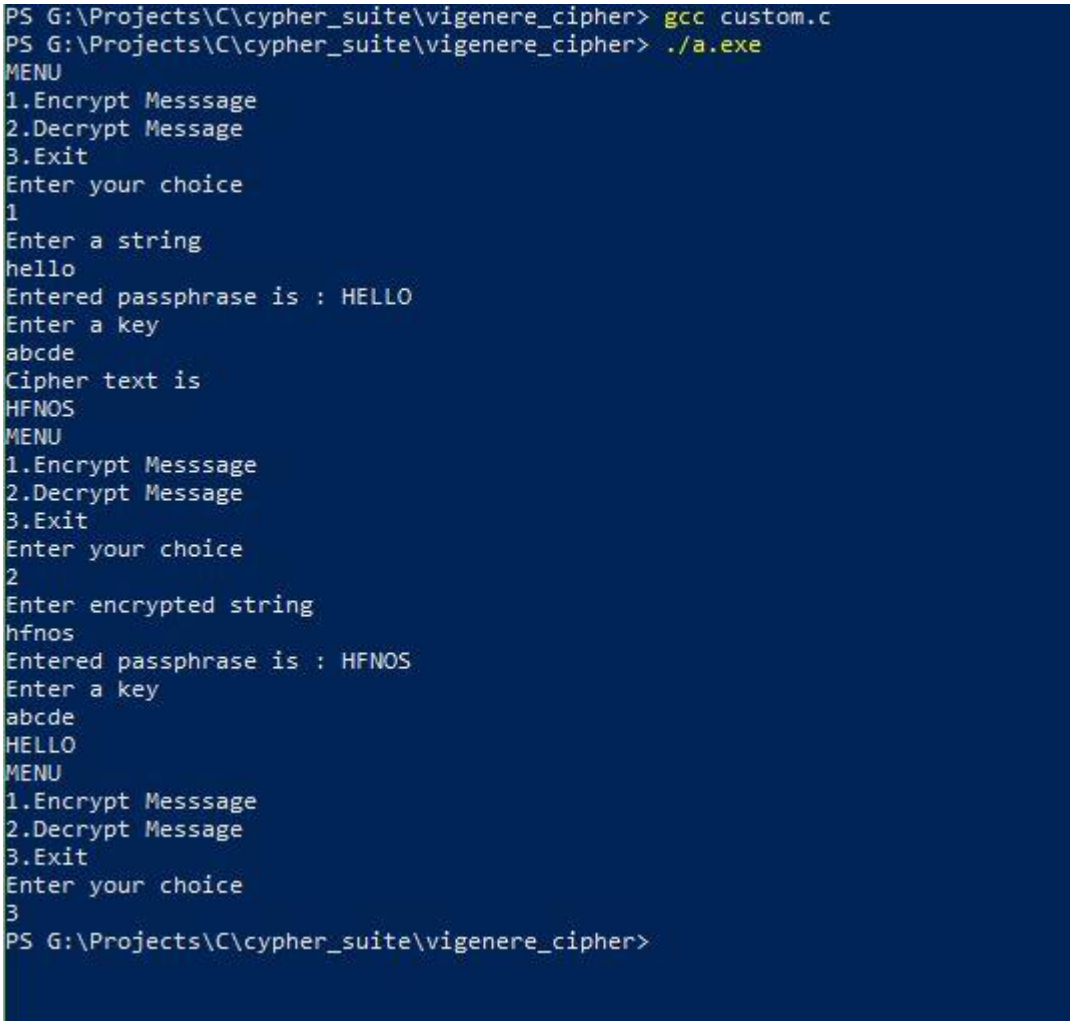
Output produced for Caesar Cipher:



Output produced for Atbash Cipher:



Output produced for Vigenère Cipher



Output produced for RSA

```
PS G:\Projects\C\cypher_suite\vigenere_cipher> cd G:\Projects\C\cypher_suite\rsa
PS G:\Projects\C\cypher_suite\rsa> gcc rsa.c
PS G:\Projects\C\cypher_suite\rsa> ./a.exe
Message data = 12.000000
p = 3.000000
q = 7.000000
n = pq = 21.000000
totient = 12.000000
e = 5.000000
d = 5.000000
Encrypted data = 3.000000
Original Message Sent = 12.000000
PS G:\Projects\C\cypher_suite\rsa>
```

Output produced for Diffie-Hellman Key Exchange Algorithm.

```
PS G:\Projects\C\cypher_suite\rsa> cd G:\Projects\C\cypher_suite\diffie-hellman
PS G:\Projects\C\cypher_suite\diffie-hellman> gcc df.c
PS G:\Projects\C\cypher_suite\diffie-hellman> ./a.exe
The value of P : 23
The value of G : 9

The private key a for Alice : 4
The private key b for Bob : 3

Secret key for the Alice is : 9
Secret Key for the Bob is : 9
PS G:\Projects\C\cypher_suite\diffie-hellman>
```

Software and Hardware Specifications

- **Compiler**:
  GCC (GNU Compiler Collection) v8.2

- **Operating System**:
  Windows XP/7/8/10, Fedora 29, Ubuntu 18.04(Bionic Beaver).

- **Processor**:
  Intel i3 or higher (Recommended for faster execution).

- **RAM**:
  2 GB or higher recommended.

- **Hard Disk**:
  250 GB or higher.

- **Code Editors used**:
  Atom editor, Vi, Nano.

Conclusion

Security in the Internet is improving. The increasing use of the Internet for commerce is improving the deployed technology to protect the financial transactions. Extension of the basic technologies to protect multicast communications is possible and can be expected to be deployed as multicast becomes more widespread.

Control over routing remains the basic tool for controlling access to streams. Implementing particular policies will be possible as multicast routing protocols improve. Cryptography is a tool which may alleviate many of the perceived problems of using the Internet for communications. However, cryptography requires the safe implementation of complex mathematical equations and protocols, and there are always worries about bad implementations. A further worry is that users are integral to securing communications, since they must provide appropriate keys. As the founders of First Virtual point out

a safe application of cryptographic technology will pay close attention to how public keys are associated with user identities, how stolen keys are detected and revoked and how long a stolen key is useful to a criminal.

Cryptography may be groovy technology, but since security is a human issue, cryptography is only as good as the practices of the people who use it. Users leave keys lying around, choose easily remembered keys, don't change keys for years. The complexity of cryptography effectively puts it outside the understanding of most people and so motivation for the practices of cryptographic security is not available

References

[1] https://en.wikipedia.org/wiki/Cipher

[2] https://en.wikipedia.org/wiki/Cryptography

[3] https://en.wikipedia.org/wiki/Caesar_cipher

[4] https://en.wikipedia.org/wiki/Vigen%C3%A8re_cipher

[5] https://www.geeksforgeeks.org/vigenere-cipher/

[6] https://en.wikipedia.org/wiki/Atbash

[7] https://en.wikipedia.org/wiki/RSA_(cryptosystem)

[8] https://simple.wikipedia.org/wiki/RSA_algorithm