

DEPARTMENT OF COMPUTER SCIENCE

The Mango Messenger

DIPLOMA PROJECT

Poznan, 2021

Contents

1	Project Assumptions	2
1.1	Project description	2
1.2	Project objectives	2
2	Impelentation	4
2.1	Project tasks	4
2.2	Project implementation	5
2.3	Project outcomes	5
2.4	Usefulness of project	5
2.5	Project self-evaluation	5
2.6	Material and bibliography used to carry out the project	6
2.7	List of annexes	6
	Bibliography	7
A	RSA Algorithm comments	8
A.1	Details	8
A.1.1	One way functions	8
A.1.2	Euler's totient theorem	8
A.2	RSA Encryption algorithm	9

Partner Details

Mentor's details

First name and surname	Szymon Murawski
Degree	Dr. Inz.
Date and signature	

Team members' details

First name and surname	Petro Kolosov
Course of study	Computer Science
Type of study program	Daytime
Date and signature	

First name and surname	Serhii Holishevskyi
Course of study	Computer Science
Type of study program	Daytime
Date and signature	

First name and surname	Illia Zubachov
Course of study	Computer Science
Type of study program	Daytime
Date and signature	

First name and surname	Arslanbek Temirbekov
Course of study	Computer Science
Type of study program	Daytime
Date and signature	

Chapter 1

Project Assumptions

1.1 Project description

Please provide an abbreviated description of the project, in line with the following structure. The description should not be more than 10,000 characters long (including spaces). Please use Times New Roman font, 12 pts, 1.5 spacing.

1. Research project
2. Justification for selecting the subject
3. Project's scope in terms of subject matter, object of research, time and space
4. Work methodology research (method and techniques)

Nowadays, instant messaging systems achieve a great success and became the main mean of communication between people via an internet. Thanks to the simplicity and quickness of the message exchanging more and more people over the world start to use instant messengers on daily basis. However, such a great attention forces us to discuss another aspect of these systems, an aspect of the information security and user privacy. The main aim of this thesis is to design and implement an instant messaging system that copes with the required functionalities and satisfies the defined security requirements.

1.2 Project objectives

1. To provide the system requirements for instant messaging system, both functional and non-functional.
2. To analyze and propose mitigations for security and user privacy vulnerabilities of the instant messaging system.
3. To propose web service (API's) architecture that fits the requirements.

4. To discuss an authorization mechanism that fits the requirements.
5. To discuss and apply E2E Encryption to the system, if necessary.
6. To implement web service.
7. To implement web client.
8. To implement mobile client.
9. To implement desktop client.

Chapter 2

Impelentation

2.1 Project tasks

Task 1.

Task name	
Entities involved to solve the task	
Task completion outcomes	
Star date of task execution	
End date of task execution	

Task 2.

Task name	
Entities involved to solve the task	
Task completion outcomes	
Star date of task execution	
End date of task execution	

Task 3.

Task name	
Entities involved to solve the task	
Task completion outcomes	
Star date of task execution	
End date of task execution	

Task 4.

Task name	
Entities involved to solve the task	
Task completion outcomes	
Star date of task execution	
End date of task execution	

2.2 Project implementation

[Please develop the theoretical assumptions of the project, including notes; present the empirical part of the project – research results and conclusions, as well as subject-matter description, etc. Please present computations and calculations, if any, in annexes. Please do not change the names of the points below. There is no predefined structure within individual points. Also, additional parts, forming individual points, can be enumerated according to your own concept. The theoretical and empirical part should not exceed 50,000 characters. Please use Times New Roman font, 12 pts, 1.5 spacing.]

1. Theoretical assumptions
2. Description of facts
3. Empirical research

2.3 Project outcomes

Please describe the achieved outcomes of the project. If possible, please provide figures showing the described outcomes. Please confront them with the objectives of the project. This part should be between 2000 and 10,000 characters long. Please use Times New Roman font, 12 pts, 1.5 spacing. Full description of solutions that were worked out and project outcomes, if any, should be presented in annexes.

2.4 Usefulness of project

Please justify how this project is useful (how the project can be used in practice). The description should not exceed 6000 characters. Please use Times New Roman font, 12 pts, 1.5 spacing.

2.5 Project self-evaluation

Each of the project's Authors describes his or her skills and competencies that were developed while working on the project and identifies issues encountered while working on the project. If during the work on the project the team had not completed any tasks planned earlier, or omitted them altogether, please specify what were these tasks and why they had not been completed. This part should not exceed 6000 characters. Please use Times New Roman font, 12 pts, 1.5 spacing.

2.6 Material and bibliography used to carry out the project

Please enumerate sources used by the team during the work on the project (as per the applying layout, Times New Roman font, 12 pts., 1.5 spacing).

2.7 List of annexes

In this place you should list all additional documents, e.g. preprinted forms, data sets, financial statements, survey templates, diagrams, concepts, strategies, studies, analyses, procedures, regulations, technical documents, plans, models, etc. which significantly contributed to the project. Please prepare all annexes in accordance with the template in place. Please use Times New Roman font, 12 pts, 1.5 spacing. All annexes form an integral part of the project.

Bibliography

- Cocks, Clifford C (1973). “A note on non-secret encryption”. In: *CESG Memo*.
- Rivest, Ronald L, Adi Shamir, and Leonard Adleman (1978). “A method for obtaining digital signatures and public-key cryptosystems”. In: *Communications of the ACM* 21.2, pp. 120–126.

Appendix A

RSA Algorithm comments

A.1 Details

A.1.1 One way functions

One way function – is a function that is easy to compute on every input, but hard to invert given the image of a random input. For instance, the function

$$f(m) = m^e \bmod N \equiv C$$

where e, N are public constants is one-way function, because it is easy to compute C given m , however it is hard to compute m given C .

A.1.2 Euler's totient theorem

Given a number N and its prime factorization $p_1^{e_1} \cdot p_2^{e_2} \cdots p_k^{e_k}$, then Euler's totient function $\phi(N)$ is defined as

$$\phi(N) = (p_1^{e_1} - p_1^{e_1-1}) \cdot (p_2^{e_2} - p_2^{e_2-1}) \cdots (p_k^{e_k} - p_k^{e_k-1})$$

In particular, for positive number M such that its factorization is $p_1 \cdot p_2$, the $\phi(M)$ is

$$\phi(M) = (p_1 - 1) \cdot (p_2 - 1)$$

Euler's theorem relates the modular division and exponent as follows, given number m , then

$$m^{\phi(N)} = 1 \bmod N$$

It means that remainder of division $m^{\phi(N)}$ by N is always 1. By the equality $1^K = 1$

$$M^{K \cdot \phi(N)} = 1 \bmod N$$

Bibliography

If we multiply both parts by M , we get

$$M \cdot M^{K \cdot \phi(N)} = M^{K \cdot \phi(N) + 1} = M \bmod N$$

A.2 RSA Encryption algorithm

The RSA algorithm is named after Ron Rivest, Adi Shamir and Len Adleman, who invented it in 1977 [Rivest, Shamir, and Adleman, 1978]. The basic technique was first discovered in 1973 by Clifford Cocks [Cocks, 1973] of CESG (part of the British GCHQ) but this was a secret until 1997. The patent taken out by RSA Labs has expired.

Historically, the process of encryption is considered to be symmetric one. That means that prior the communication, the sides conclude on the common key to be used in encryption. This process is similar to the first sharing keys and only after that the locked chest with the message. Such approach is highly cost since it requires to share the defined keys between each actor if the number of actors is greater than 2. Much more simpler is to think about secured communication channel that in terms of asymmetric encryption. The real life example would be if Alice shares with all actors an opened lock having key. So that Bob receives an opened lock, writes letter to Alice, puts letter to the chest, locks this chest with received from Alice lock. This way, only Alice will be able to open the chest and to read the letter. This is an idea of the asymmetric encryption. However, such a simple from first glance idea requires complex number theory approach. A concept of opened lock may be interpreted in terms of one-way functions. One way function – is a function that is easy to compute on every input, but hard to invert given the image of a random input. Thus, it is much simpler to close the lock without key, but very difficult to open lock trying the combinations of the key. For instance, the function

$$f(m) = m^e \bmod N = C$$

where e, N are public constants is one-way function, because it is easy to compute C given m , however it is hard to compute m given C . So, assume that Alice defines two positive integer constants e, N and sends it to Bob. Bob encrypts the secret message m using $f(m)$

$$f(m) = m^e \bmod N = C$$

Then Bob sends encrypted message C to the Alice. Given C Alice must fetch the Bob's message m . In order to decrypt C , Alice has to compute

$$C^d \bmod N = m^{ed} \bmod N \equiv m,$$

Bibliography

where e for encryption and d for decryption. Now the problem is to define such d that it is hard to the listener to fetch it. In order to define the secret d , Alice chooses two enough big prime numbers: P , Q , let's say around 150 digits both. Then Alice multiplies these two prime numbers in order to get N

$$N = P \cdot Q$$

The N is around 300 digits. Now Alice can share N with anyone, since it takes decades to find its prime factorization by the fundamental problem of prime factorization. Next, it is very important to know such a function, which depends on the knowledge of factorization of N . Such function is an Euler's totient function. Given a number N and its prime factorization $p_1^{e_1} \cdot p_2^{e_2} \cdots p_k^{e_k}$, the Euler's totient function $\phi(N)$ is defined as

$$\phi(N) = (p_1^{e_1} - p_1^{e_1-1}) \cdot (p_2^{e_2} - p_2^{e_2-1}) \cdots (p_k^{e_k} - p_k^{e_k-1})$$

In particular, for positive number M such that its factorization is $p_1 \cdot p_2$, the $\phi(M)$ is

$$\phi(M) = (p_1 - 1) \cdot (p_2 - 1)$$

Euler's theorem relates the modular division and exponent as follows, given number m , then

$$m^{\phi(N)} = 1 \bmod N$$

It means that reminder of division $m^{\phi(N)}$ by N is always 1. By the equality $1^K = 1$

$$M^{K \cdot \phi(N)} = 1 \bmod N$$

If we multiply both parts by M , we get

$$M \cdot M^{K \cdot \phi(N)} = M^{K \cdot \phi(N) + 1} = M \bmod N$$

It follows that Alice is able to define the secret d as follows

$$\begin{aligned} e \cdot d &= K \cdot \phi(N) + 1 \\ d &= \frac{K \cdot \phi(N) + 1}{e} \end{aligned}$$

The following image demonstrates the concept of RSA approach

Bibliography

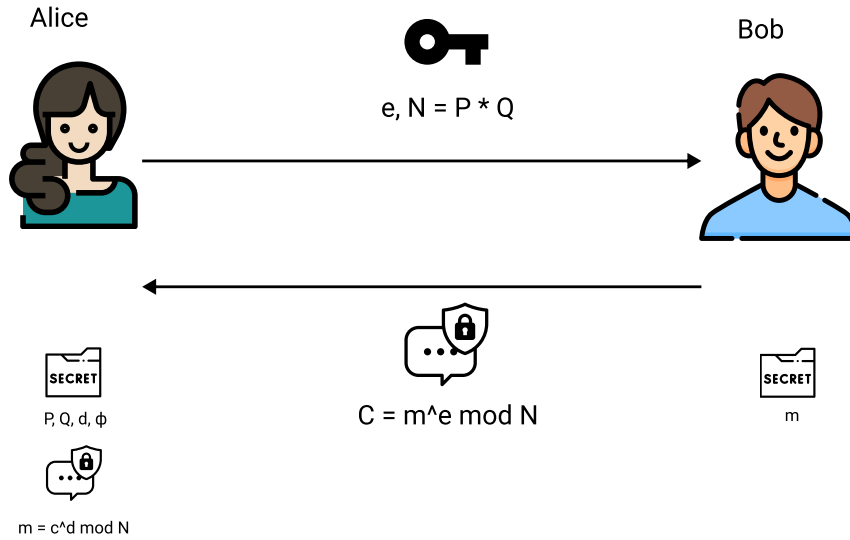


FIGURE A.1: Secret chat encryption concept diagram. Source:

To summarize, the process by the steps is as follows

- Alice defines the large secret prime numbers P, Q .
- Alice computes $N = P \cdot Q$ and $\phi = (P - 1)(Q - 1)$
- Alice chooses an integer e , $1 < e < \phi$ such that $\gcd(e, \phi) = 1$.
- Alice computes secret exponent d , $1 < d < \phi$ such that $ed \equiv 1 \bmod \phi$.
- Alice shares public key (N, e) with Bob and keeps private key (d, p, q) is secret.
- Bob defines the message m , encrypts it as $C = m^e \bmod N$.
- Bob sends C to Alice.
- Alice decrypts C using her secret d , so she gets m

$$m = C^d \bmod N$$

Security of the RSA approach is based on the complexity of fundamental problem of prime factorization, which takes decades to solve having enough large number.