Functionality of the
RSA cipher

CrypTool Team

November 2010

Cryptography and what you need it for

- Sending encrypted messages has always played a major role in the history of humanity. In each era there has been important information which had to be kept secret from other people.

- Especially in today's society, in the age of internet, it is important to be aware of data security.

Data reaches their receiver indirectly by passing between several servers.

At each node, the data can be captured, read and even changed.

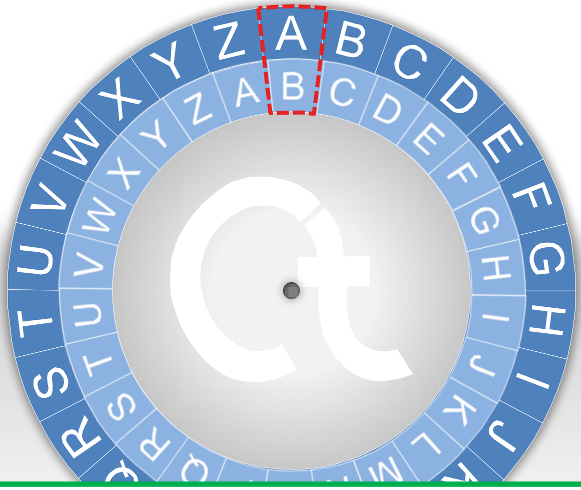
Modern cryptography is about securing this data.

Introductory example: Caesar cipher

- One of the first ways of encrypting a message was the Caesar cipher. The method got its name from the ancient emperor Julius Caesar, who used it 2000 years ago to encrypt secret messages to his generals.
- Here you can see how it works:

Plaintext

This is a secret information!


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

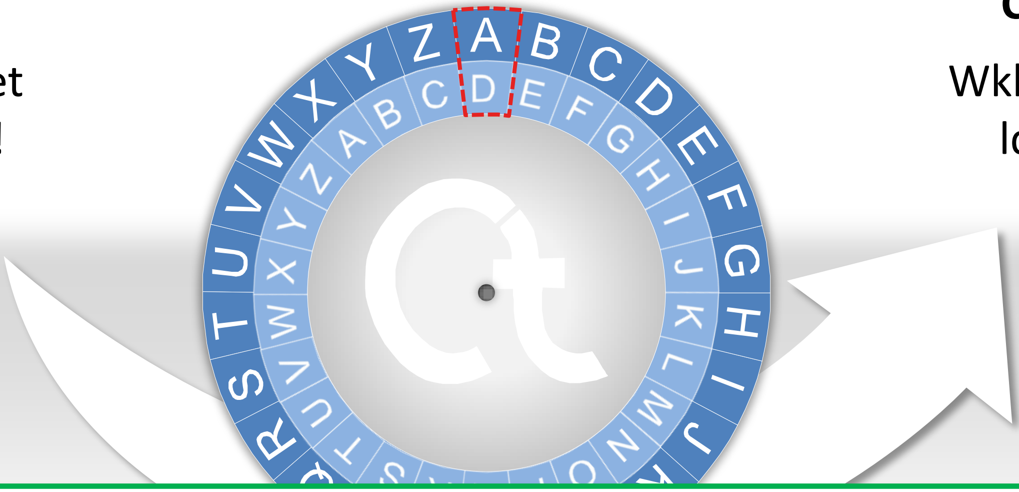
Write out the alphabet twice in two concentric circles. Offset the letters of the inner circle from the outer one by a certain amount.

Introductory example: Caesar cipher

- One of the first ways of encrypting a message was the Caesar cipher. The method got its name from the ancient emperor Julius Caesar, who used it 2000 years ago to encrypt secret messages to his generals.
- Here you can see how it works:

Plaintext

This is a secret information!


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

Chipertext

Wklv lv d vhfuhw lqirupdwlrq!

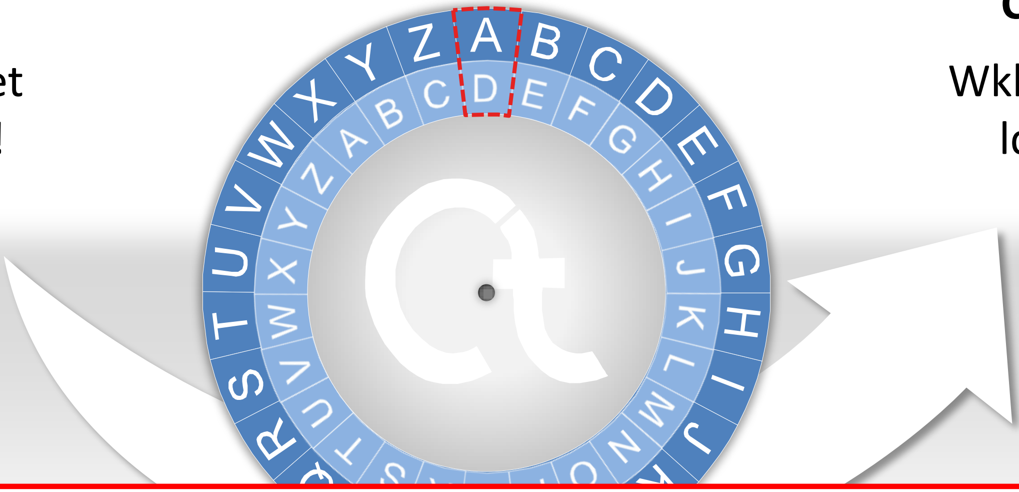
Now each letter in the plaintext will be replaced by its corresponding letter in the inner circle. That’s how you get the ciphertext.

Introductory example: Caesar cipher

- One of the first ways of encrypting a message was the Caesar cipher. The method got its name from the ancient emperor Julius Caesar, who used it 2000 years ago to encrypt secret messages to his generals.
- Here you can see how it works:

Plaintext

This is a secret information!


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

Chipertext

Wklv lv d vhfuhw lqirupdwlrq!

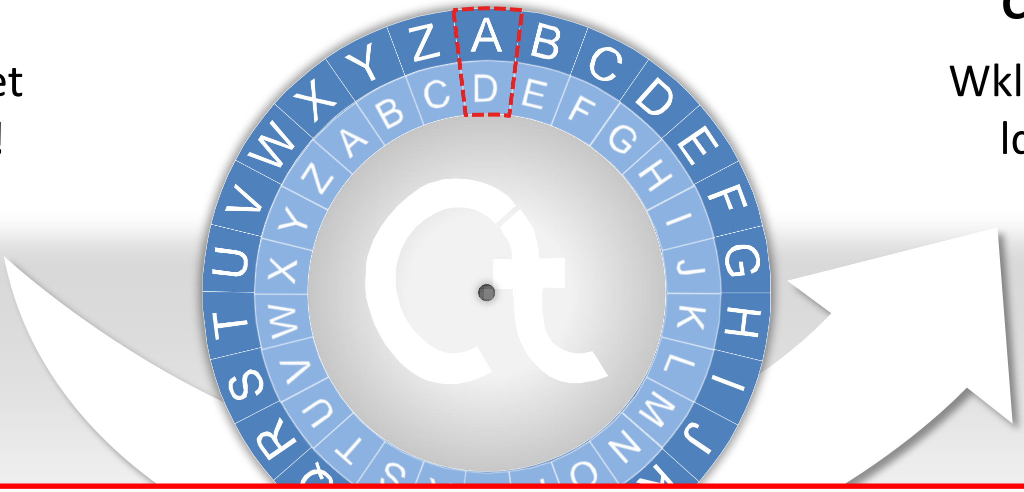
As there are limited possibilities (only 26 possibilities of different chipertexts), this cipher is quite easy to break.

Introductory example: Caesar cipher

- One of the first ways of encrypting a message was the Caesar cipher. The method got its name from the ancient emperor Julius Caesar, who used it 2000 years ago to encrypt secret messages to his generals.
- Here you can see how it works:

Plaintext

This is a secret information!



Chipertext

Wklv lv d vhfuhw lqirupdwlrq!

! **Do you want to try this cipher on your own text?**
You can try it here.

Model of the RSA cipher

- The goal is to achieve a safe means of communication.
  "Safe" in this case means that even if a message is intercepted, it should not be possible for an attacker to read the message.
- How can we realize this security? A modern solution is the **RSA cipher**.
- The idea of the cipher is as follows:

Each participant has a padlock with a matching key.

Model of the RSA cipher

- The goal is to achieve a safe means of communication.
  "Safe" in this case means that even if a message is intercepted, it should not be possible for an attacker to read the message.
- How can we realize this security? A modern solution is the **RSA cipher**.
- The idea of the cipher is as follows:

The main idea is to separate the padlock from the key. You should publicize copies of your padlock, as opposed to your key, which you should keep secret.

Model of the RSA cipher

- The goal is to achieve a safe means of communication.
  "Safe" in this case means that even if a message is intercepted, it should not be possible for an attacker to read the message.
- How can we realize this security? A modern solution is the **RSA cipher**.
- The idea of the cipher is as follows:

Now someone who wants to send you a message is able to encode his or her message with your padlock.

Model of the RSA cipher

- The goal is to achieve a safe means of communication.
  "Safe" in this case means that even if a message is intercepted, it should not be possible for an attacker to read the message.
- How can we realize this security? A modern solution is the **RSA cipher**.
- The idea of the cipher is as follows:

Then the message can be sent in public, as only the right recipient will be able to open the padlock with the appropriate key.

The essential problem

- The RSA cipher is the electronic implementation of the model described before.
- The cipher got its name from its inventors: **R**ivest, **S**hamir and **A**dleman.
- The algorithm is based on an underlying mathematical problem. Specifically, it is the problem of factoring a given large number into prime numbers.
- When you have a number that is a product of large prime numbers, it is quite hard to find its decomposition. Still today no one has found a fast and effective way of finding the factors. The security of RSA is based on this difficulty.

3347807169895689878604416984821269081770479498371376856891
2431388982883793878002287614711652531743087737814467999489

* *3674604366679959042824463379962795263227915816434308764267*
6032283815739666511279233373417143396810270092798736308917

= *1230186684530117755130494958384962720772853569595334792197*
3224521517264005072636575187452021997864693899564749427740
6384592519255732630345373154826850791702612214291346167042
9214311602221240479274737794080665351419597459856902143413

Bit length: 768 Decimal length: 232

! **Current PCs can quickly factor numbers with about 80 digits.**
Therefore, practical RSA implementations must use moduli with at least 300 digits to achieve sufficient security.

How does the RSA cipher work?

To understand how RSA cipher works you need some basic mathematical concepts. We will explain this in the next slides.

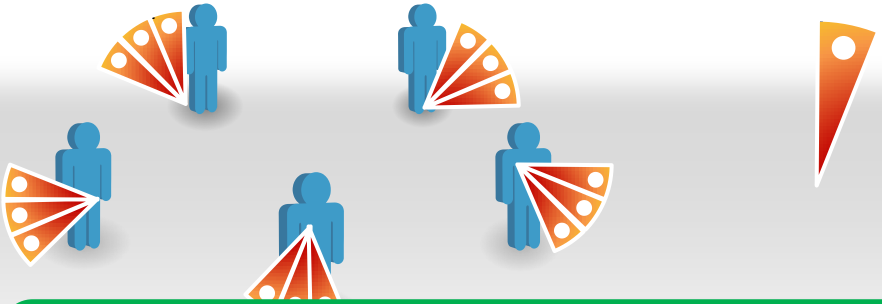
1. The modulo operator
2. Euler‘s totient function
3. Euler-Fermat theorem

Mathematical basics - 1

The modulo operator

- This sign is the modulo operator. With the modulo operation you are interested in the remainder left over from division with an integer number.
- To get a better idea, take a look at the following:

≡



$$16 \equiv 1 \, mod \, 5$$

Five people want to share a cake which is already cut into 16 pieces.
Each of them can get three pieces of cake, but one will be left over.
The modulo operator calculates precisely this remainder.

Mathematical basics - 1

The modulo operator

- This sign is the modulo operator. With the modulo operation you are interested in the remainder left over from division with an integer number.
- To get a better idea, take a look at the following:

$$\equiv$$

Mathematical definition

$$a \equiv b \; mod \; N$$

means that there exists an integer number k such that a can be represented as

$$a = k * N + b$$

with the condition that: $0 \leq b \leq N - 1$

! We are not interested in the value of k. The important part is its existence.

An example

The modulo operator is commutative with the basic arithmetic operations. For example it does not matter whether you **first** multiply

$$18 * 13 = 234 \equiv 4 \; mod \; 10$$

or first calculate the modulus **and then** multiply:

$$18 * 13 \equiv 8 * 3 \; mod \; 10$$
$$= 24 \; mod \; 10 \equiv 4 \; mod \; 10$$

? Further information can be found in the CrypTool Script (chap. 4.4).

Mathematical basics - 2

Euler’s totient function

- Euler’s totient function φ of an integer N counts how many whole numbers are both coprime to N and smaller than$N$.
- Here how the formula looks:

$$\varphi(N) = \#\{a \epsilon \mathbb{N} \mid gcd(a, N) = 1 \;\; and \;\; 1 \leq a < N\}$$

?

Important properties

Given a number which is product of two factors a and b :

$$\varphi(a * b) = \varphi(a) * \varphi(b)$$

Given a prime number p:

$$\varphi(p) = p - 1$$

Therefore , given a number composed of two primes, $N = p * q$:

$$\varphi(N) = \varphi(p * q) = \varphi(p) * \varphi(q) = (p - 1)(q - 1)$$

Example

Suppose we want to calculate $\varphi(10)$.
First we find the factor of 10:

$$10 = 5 * 2$$

Becauseo 5 and 2 are primes, we can use the formula given to the left:

$$\varphi(10) = \varphi(5) * \varphi(2) = 4 * 1 = 4$$

$$\varphi(5) = \#\{1, 2, 3, 4\} = 4 \qquad \varphi(2) = \#\{1\} = 1$$
$$\varphi(10) = \#\{1, 3, 7, 9\} = 4$$

Näheres siehe CrypTool-Skript, Kap 4.8.2

Mathematical basics - 3

The Euler-Fermat theorem

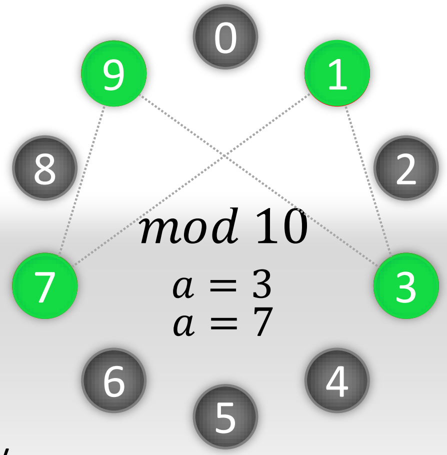
- The last basic equation is the Euler-Fermat theorem.

Given two coprime numbers a and N:

$$a^{\varphi(N)} \equiv 1 \, mod \, N$$

Modulus calculations operate in the **finite** set $\{0, 1, \ldots, N-1\}$. A function is called **cyclic** if, after repeated application, the results repeat themselves within this set.

For example, one such cyclic function is multiplication with a fixed base. We will choose the numbers $a = 3$ and $a = 7$ as the fixed bases. We can multiply each number by itself until we reach it again. In our example $N = 10$ with $\varphi(N) = 4$.



3 ▸ 9 ▸ 7 ▸ 1 ▸ 3

7 ▸ 9 ▸ 3 ▸ 1 ▸ 7

The cycles generated by this operation both have length 4, which is exactly $\varphi(N)$.

If you multiply a number a by itself, you will, with absolute certainty, reach a again in at most $\varphi(N)$ steps. You can verify this by multiplying both sides of the formula above by a.

With these basic equations we can start looking at the actual cipher.

Step 1: Generate the keys

- We separate the RSA cipher algorithm into three different steps which will be explained on the following slides.
- First we have to generate our RSA keys. This step must be done only once as an initial step.

Formal

1 Choose two primes p and q with $p \neq q$

2 Calculate their product: $N = p * q$

3 Calculate the value of Euler’s totient function of N

$$\varphi(N) = \varphi(p * q) = (p - 1)(q - 1)$$

4 Choose a number e between 1 and $N - 1$ which is coprime to $\varphi(N)$

5 Find another number d where

$$d * e \equiv 1 \; mod \; \varphi(N)$$

Example

1 Suppose we select $p = 13$ and $q = 7$

2 Thus: $N = 13 * 7 = 91$

3 $\varphi(91) = \varphi(13 * 7) = (13 - 1)(7 - 1) = 72$

4 Suppose we choose $e = 5$, because: $gcd(5, 72) = 1$

5 We will select $d = 29$ as thus:

$$d * e = 145 = 2 * 72 + 1 \equiv 1 \; mod \; 72$$

? Here you can get more information on how to find an appropiate number d
(by means of the extended Euclidean algorithm)

(e, N) **is the *public RSA key.***
(d, N) **is the *private key.***

You can find further details in the CrypTool script, chap. 4.10.3

Step 2: Encrypt messages

- Now we have the requirements to encrypt and decrypt messages.
- First we must convert the letters into numbers to be able to use them in our calculations.

For example you can use the following substitution:

A	B	C	D	…	Z
01	02	03	04	…	26

Formal

To encrypt a message we have to calculate

$$C \equiv K^e \; mod \; N$$

Here K is the converted message and C is the encoded text, the ciphertext. The numbers e and N are taken from the public RSA key.

! **The presented cipher is simplified. Further information is provided in the next slides.**

Example

We shall continue our example by encoding the word "SECRET":

Letters	S	E	C	R	E	T
Numbers	19	05	03	18	05	20

Now we take the first letter S = 19 and encrypt it by using the public key: $(5, 91)$

$K^e = 19^5 = 19 * (19^2)^2 = 19 * (361)^2$

$\equiv 19 * (88)^2 \equiv 19 * 9 = 171 = 80 \; mod \; 91$

Following this pattern, "SECRET" is encrypted as follows:

80	31	61	44	31	76

Step 3: Decrypt Messages

- The receiver gets the message now in its encrypted form only.

Formal

To decipher the original message the receiver needs to calculate the following:

$$K \equiv C^d \ mod \ N$$

Here K will produce the plaintext. The values d and N are saved in the receiver's private key (d, N).

Example

The encrypted message is as follows:

30	31	61	44	31	76

According to the formula given to left, he or she can decipher by using his or her private key $(29, 91)$:

$$C^d = 30^{29} = \cdots \equiv 19 \ mod \ 91$$

The complete plaintext is obtained by calculating accordingly for each value.

Nnumbers	19	05	03	18	05	20
Letters	S	E	C	R	E	T

?

Why do you get the plaintext by using these formulas?
You can learn the answer on the following slides.

Explanation of the formulas

- The following formulas explain why the receiver will obtain the plaintext from the encrypted text.
- First we should examine the process of decryption more precisely. Since $C = K^e$,

$$C^d = (K^e)^d = K^{e*d}$$

- Thus $d * e \equiv 1 \; mod \; \varphi(N)$, which is equivalent to $d * e = 1 + l * \varphi(N)$, where l is an arbitrary integer number.
- We can then derive the following sequence of equations:

$$K^{e*d} = K^{1+l*\varphi(N)} = K * K^{l*\varphi(N)} = K * (K^{\varphi(N)})^l$$

- By means of Euler-Fermat theorem, $K^{\varphi(N)} \equiv 1 \; mod \; N$, we get:

$$K * (K^{\varphi(N)})^l \equiv K \; mod \; N$$

- All in all we get the following:

$$C^d \equiv K \; mod \; N$$

By raising the ciphertext to a higher power, we reobtain the plaintext.

Security of the cipher

- The given example was simplified to make the explanation clearer. If you were to use the cipher as it was just explained, communication would be insecure.

S	E	C	R	E	T
19	05	03	18	05	20

30	31	61	44	31	76

.	E	.	.	E	.

By encoding each letter to one number, the resulting encryption will be a one-to-one mapping: for each letter there is just one corresponding number in the ciphertext.
So an easy way of attacking the ciphertext is by using a frequency analysis. The idea is that there exists an unequal distribution of letters in each language. In English the most frequent letter is the letter “E”, so you can try to replace the most frequent number in the ciphertext with “E”.

- To avoid this problem, a possible solution is to combine several numbers into a block. In our example we could unite it as follows an then encode it again in another manner:

SEC	RET
190503	180520

! **By combining several numbers to a block, we have to pay attention in choosing our module N. It has to be bigger than the largest possible number in the block.**

- In practice, RSA is not used to encrypt text blocks, but rather combined with a symmetric cipher. In this case, RSA is only used to encrypt the key of the symmetric cipher (Hybrid cryptosystem).

The factorization problem and RSA

- You may be asking yourself why all of this is based on the problem of factoring large numbers.
- We will explain this with the help of our example, as it is easy to find the factors of the number N:

$$N = 91 = 13 * 7 = p * q$$

- As soon as you have the factorization, you can calculate $\varphi(N)$. By means of e and the connection of d and e with the formula $d * e \equiv 1 \, mod \, \varphi(N)$, you can easily find the number d, which is – together with N – the private key. Once you have the private key you can decrypt the entire ciphertext.
- No one has yet found a way to:
  - calculate d with the help of e without knowing the factorization of the number N.
  - calculate the plaintext from the cipher without knowing the private key d.

!

In fact, by knowing the factorization and the public key (e, N) it is possible to generate the private key. Therefore, the attacker could repeat the first step of the process, the generation of the keys.

Further information and references

- http://www.cryptool.org
  An open-source software tool for learning cryptographic ciphers and cryptanalysis
- http://cryptool.org/download/CrypToolScript-en.pdf
  A thorough script with more information about the mathematical aspects of cryptography
- http://en.wikipedia.org/wiki/Cryptography
  Wikipedia article about cryptography in general
- http://en.wikipedia.org/wiki/RSA_cipher
  Wikipedia article about the RSA cipher
- http://www.gax.nl/wiskundePO/
  Online RSA encryption application (Dutch)