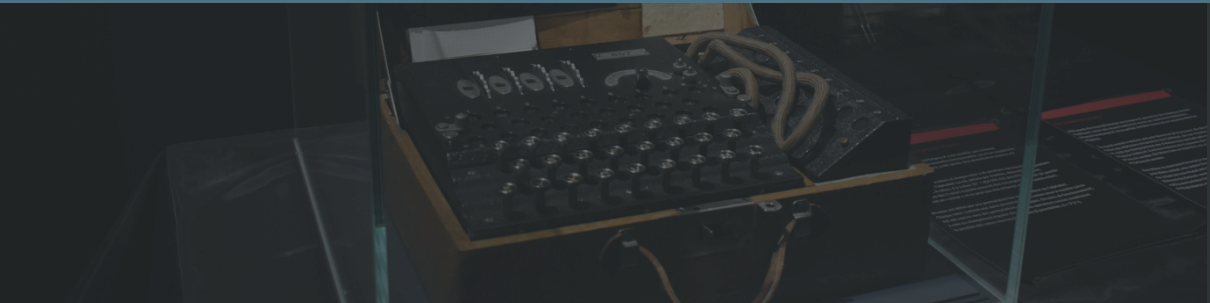
Cryptography, including quantum

Robet von Oliva
naruciakk

2024-11-24



Beginnings

Simple ciphers

- Atbash (aleph-tab-bet-shin) – $\aleph \rightarrow \text{ת}$ etc.
- Caesar's Cipher – $E_n(x) = x + n \mod N(\Sigma)$
- Scytale – using a physical cylinder
- Polybius Square

Figure: Wikicommons, Avi1111 dr. avishai teicher, CC BY 3.0

Figure: Wikicommons, Luringen, CC BY 3.0

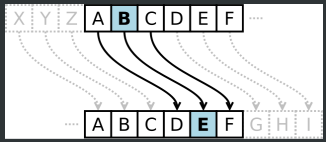


Figure: Wikicommons, Cepheus, domena publiczna

	1	2	3	4	5
1	A	B	Γ	Δ	E
2	Z	H	Θ	I	K
3	Λ	M	N	Ξ	O
4	Π	P	Σ	T	Y
5	Φ	X	Ψ	Ω	

Figure: Wikipedia, CC BY 3.0

1. Atbash - a simple mono-alphabetic representation cipher, involves converting a letter at a distance *r* from the beginning of the alphabet into a letter at a distance *r* from the end of the alphabet. Its name comes from the way it works, i.e. one by one which letters are converted into which: alef into taw, beth into shin, and so on. Its decryption is very simple - just parse the encrypted text again in the same way to get the plaintext. It also occurs in the Bible: Jeremiah 25:26 and Jeremiah 51:41 – Babylon is spelled as Sheshak, Jeremiah 51:1 – Chaldeans are Leb Kamai.
2. Caesar's cipher - each letter of the plaintext is swapped with another letter lying *n* positions away in the alphabet from the original one. Caesar used the shift with key 3 to encrypt private correspondence. It is now used in puzzles and other simple toys.
3. Skytale - a method of encryption used in ancient Sparta. A narrow strip of parchment is wound onto a stick, writing the text on the edges along the stick at the touching edges. With a rod of identical thickness, the plain text can be read. The diameter of the rod is therefore the key.
4. Polybius square - the key is a table assigning each letter a row and column number (encrypted text).

Less simple ciphers

ROT13 – Caesar's cipher with the rotation of the half of the alphabet

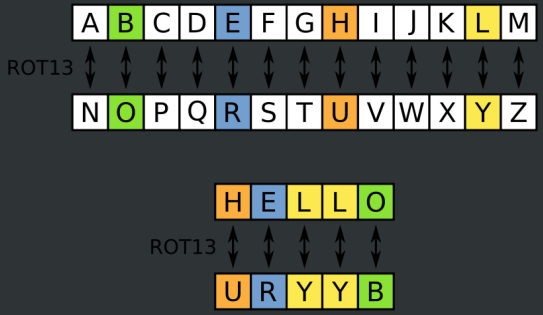


Figure: Wikicommons, domena publiczna

VIC cipher

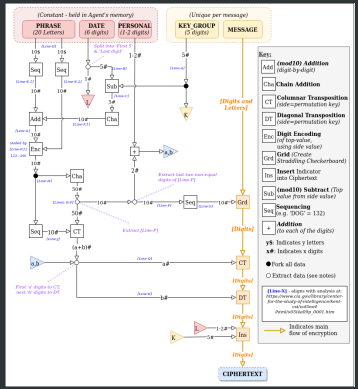


Figure: Wikicommons, R. Fallon, CC BY-SA 4.0

1. ROT13 - Caesar's cipher with an offset of 13 (half the length of the Latin alphabet). It is its own inverse (we encrypt and decrypt with the same function). Used sometimes on internet forums/usenet to hide spoilers and other sensitive messages where the author wanted to avoid accidental reading.
2. VIC - a cipher used in the 1950s by Soviet spies in the US, named after the spy Reino Häyhänen, alias 'VICTOR'. It is one of the most complex 'paper' ciphers (counted using a piece of paper and a pen). It is based on the Nihilist cipher, a basic cipher with a double key, similar to Polybius' square. The encryption is divided into two stages: we create the keys and, based on them, a Polybius square with a bifurcation and two arrays where the cipher digits are shuffled. Four fixed elements are needed: a group of six digits (e.g. a date), a sequence of 20 letters, a key word and one short number, and a variable element - a five-digit number invented once for each plaintext. Stage two is encryption: using a checkerboard, the plaintext is converted into a sequence of digits, which is then shuffled by rearranging the columns. The sequence of digits is then divided into five-digit groups and a variable element is added to the sequence.

Enigma

Portable electromechanical cipher machine:

- mechanical part - keyboard, rotors on a common axis and a mechanism that rotates one or more rotors;
- electrical circuit - encoding is done electrically;
- single rotor - simple substitution cipher;
- plugboard - additionally swapping some letters;
- initial setting (which is the key) - the order of rotors, initial position of rotors, rotors setting, plugboard setting

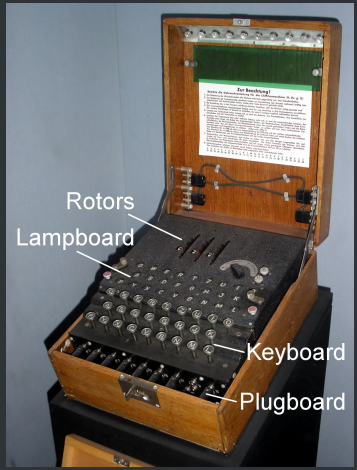


Figure: Wikicommons, domena publiczna

1. Operating procedure - set the rotors according to the daily setting (codebook), select a random combination of rotor settings (individual message key) and broadcast it at the beginning of the transmission.
2. The difficulty is that each letter of the plaintext is encrypted using a different permutation of the alphabet, each letter having a different substitution cipher.
3. The objective is to obtain the daily setting.
4. The bomb (bomba kryptologiczna) – an electric-mechanical machine for automatically breaking the daily cipher. Initially designed by cryptologists from the Polish Cipher Bureau for an early (smaller number of cipher drums and different encryption procedure) version of Enigma. The concept was improved by Alan Turing (the approach of looking for certain words, such as 'Wetter' (weather) or a Nazi greeting - thus ruling out a lot of possible settings).

One-time pad

It is not possible to crack this method!

- the key is single-use,
- the key is as random as possible (discrete uniform probability distribution),
- the key is at least the length of the plaintext,
- the key cannot be reused even partially,
- the key must be secret for the third-party.

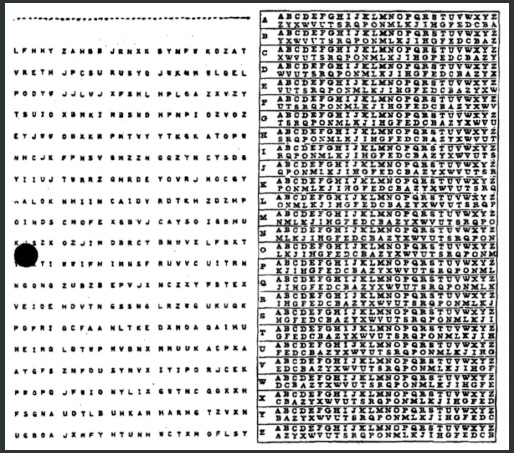


Figure: Wikicommons, domena publiczna

1. Each key is used only once. It is created completely at random, and once encrypted, the messages must be destroyed.
2. Using, for example, a polyalphabetic cipher, it is possible to achieve complete message security, since for each pair *T* and *S* (plaintext and encrypted text) there is a matching *K* key, i.e. knowing only *S* we can match any *M* of the same length and calculate the matching key.
3. Even up until the 1970s, the National Security Agency (NSA) was creating single-use keys used for general-purpose but also special-purpose codes, such as nuclear codes.

The dawn of real cryptography

A Mathematical Theory of Cryptography

In the 1940s, Claude Shannon described the basics of information theory:

- *A Mathematical Theory of Communication* – quantification of the level of uncertainty (Shannon entropy): $H = -\sum_{i=1}^{n} p_i \log p_i$,
- *A Mathematical Theory of Communication* – examination methods for cryptographic systems (the probability of receiving certain values),
- *A Mathematical Theory of Communication* – two types of systems: *unconditional secrecy* and *conditional secrecy*,
- *Communication Theory of Secrecy Systems* – all unbreakable ciphers have the limitations of a cipher with a one-time key.

Figure: Wikicommons, Konrad Jacobs, CC BY-SA 2.0 de

1. Entropy is used to measure the amount of information available in a message. It measures the amount of uncertainty, the more information is contained in a message, the smaller it is. Entropy is subject to standard mathematical transformations, so values such as conditional entropy can be determined ($H(Y|X) = \sum_{x \in X} p(x) \sum_{y \in Y} p(y|x) \log \frac{1}{p(y|x)} = \sum_{x \in X} \sum_{y \in Y} p(x,y) \log \frac{p(x)}{p(x,y)}$, where $p(y|x)$ is the probability of the result y if and only if x is true, $p(x,y)$ is the joint probability) or mutual information – the amount of information shared by two systems ($I(X;Y) = H(X) - H(X|Y) = \sum_{y \in Y} \sum_{x \in X} p(x,y) \log \frac{p(x,y)}{p(x)p(y)}$).
2. When examining a language from the cryptographic side, all we care about are the statistical properties of a language: the distribution and frequency of characters, binary characters, words, phrases and so on.

Modern techniques

Symmetric-key cryptography

The first type, more similar to the initial systems:

- we use the same key to encode and decode,
- the key is as random as possible (discrete uniform probability distribution),
- a secure key exchange protocol must be established,
- there are two types: block ciphers and stream ciphers.

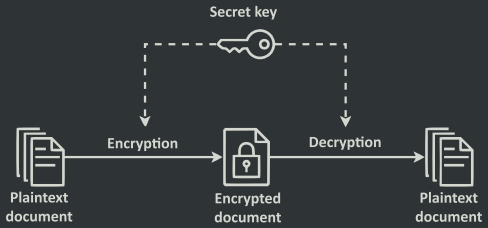


Figure: Wikicommons, Michel Bakni, CC BY-SA 4.0

1. The keys do not necessarily have to be the same, and a simple (i.e. performed in polynomial time) transformation between two keys can also be used. Effectively, this means that systems of this type require a shared secret and the longer the secret, the better.
2. This secure key exchange protocol is often a simple key transfer procedure using asymmetric encryption. Although this is not intuitive at first glance (if we already have a secure channel for exchanging information, why create a second one?), but symmetric algorithms are very efficient (see AES) and doing it this way leads to significant savings in computing power/time. There are, however, other classical key exchange protocols which are not the subject of this presentation.
3. Block ciphers, as the name implies, divide the plaintext (in either byte or letter form) into blocks and encrypt the individual blocks as a whole. Stream ciphers encrypt each character (or byte) one at a time - to simplify, we can say that stream ciphers are block ciphers with a block size of 1.

Advanced Encryption Standard (Rijndael)

The algorithm that won the 1998 competition for a new standard for encryption run by NIST (National Institute of Standards and Technology):

- the most popular symmetric-key algorithm;
- it uses a byte matrix of size 4×4:
  1. each byte is added with the corresponding byte from the key (*XOR*),
  2. we conduct 9/11/13 rounds: substitution + row shifting + column mixing + addition of key byte,
  3. the final round without mixing columns;
- there are three types: 128 bits of the key, 192 bits albo 256 bits;
- the algorithm is (currently) secure, $2^{126.2}$ operations for 128 key bits ($4.36 * 10^{17}$ seconds since the Big Bang).

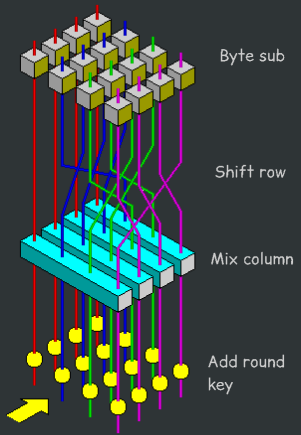


Figure: Wikicommons, public domain (CC0)

1. AES is a sub-family of the Rijndael family of algorithms, created by 2 Belgian cryptographers Vincent Rijmen and Joan Daemen, for a competition organised by NIST.
2. The number of rounds depends on the key length: 128 bits gives 10 rounds (including the last one), 192 is 12 rounds and 256 is 14 rounds.
3. The substitution cipher is non-linear and based on a certain matrix, the shifting of rows and the mixing of columns are done within a 4×4 matrix (these are the transposition steps).

Asymmetric cryptography

- We use two different keys for encoding (*public key*) and decoding (*private key*);
- the public key can be made available to anyone without the loss of security;
- it is also possible to use this technique for digital signing;
- for RSA, a security condition is that it is not possible to factorise numbers composed of two large prime numbers quickly.

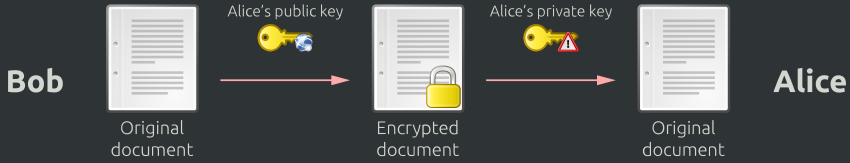


Figure: Wikicommons, public domain

1. Asymmetric encryption is sometimes called the public key encryption.
2. The key is that it should impossible to obtain the private key by just knowing the public key - we rely here on the properties of one-way functions, easy to compute one way, but the inverse function is difficult - e.g. multiplication of two large prime numbers is easy, but factorisation is extremely computationally difficult.
3. The security of the RSA algorithm is symmetric - knowing the public key we cannot know the private key and vice versa. Let us denote the public key as (e, n) and the private key as (d, n). Let p and q be two large and close in length prime numbers generated as randomly as possible. We obtain d and e from the equation $de \equiv 1(\mod (p-1)(q-1))$, where p and q are then treated together as n, and either of the numbers, either d or e, is drawn randomly.
4. A digital signature system can be achieved, for example, by swapping the RSA public and private keys in place. The digest (hash) of the message is encrypted with the public key (in the application of encryption) and verified with the private key. It is therefore sufficient to provide a private key instead of the public key and we can sign messages digitally.

Quantum computing

The basic unit of information

1. Classical – bit ($b = 0 \vee b = 1$)
2. Quantum – qubit
   ($|q\rangle = |0\rangle \vee |q\rangle = |1\rangle \vee \ldots$)

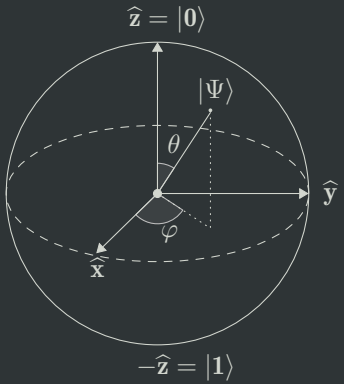


Figure: Glosser.ca, Wikicommons, CC BY-SA 3.0

Information processed in classical computers is stored as numbers in the binary system. In the case of quantum computers, on the other hand, information is processed using quantum physical objects described in a two-dimensional Hilbert space called qubits. The physical implementation is possible, for example, as the polarisation of a single photon or an elementary particle with spin $1/2$.

Superposition

$$|\rho\rangle = c_0 |0\rangle + c_1 |1\rangle = \begin{bmatrix} c_0 \\ c_1 \end{bmatrix}$$

where $c_0, c_1 \in \mathbb{C}$ oraz ${c_0}^2 + {c_1}^2 = 1$.
Quantum states can be added and new quantum states can be created from them, when measuring we get base states, e.g. 0 or 1.

Entanglement

The system is better described than its components

$$|\rho\rangle_{12} = c_{00}\,|0\rangle_1|0\rangle_2 + c_{01}\,|0\rangle_1|1\rangle_2 + c_{10}\,|1\rangle_1|0\rangle_2 + c_{11}\,|1\rangle_1|1\rangle_2$$

for comparison

$$|\psi\rangle_1 \otimes |\phi\rangle_2 = (a|0\rangle_1 + b|1\rangle_1) \otimes (c|0\rangle_2 + d|1\rangle_2),$$

$$|\psi\rangle_1 \otimes |\phi\rangle_2 = ac\,|0\rangle_1|0\rangle_2 + ad\,|0\rangle_1|1\rangle_2 + bc\,|1\rangle_1|0\rangle_2 + bd\,|1\rangle_1|1\rangle_2$$

is not equal to the first one!

Additionally ${c_{00}}^2 + {c_{01}}^2 + {c_{10}}^2 + {c_{11}}^2 = 1$.

We are not able to describe a system of particles using the states of its constituent particles

Constraints

In the case of quantum computing, there are certain constraints on the construction of solutions, stemming directly from the limitations of quantum mechanics.

1. Reversibility – quantum operations are reversible, any gates provide the opportunity to return to the information contained in the input.
2. No-cloning theorem – that it is impossible to create an independent and identical copy of an arbitrary unknown quantum state.

The issue

Shor's Algorithm

Take a certain natural number N. The goal is to find the divisor of this number (in practice, the decomposition into prime factors of the product of two large prime numbers) Classical computing: there are no polynomial solutions to this problem: $O(e^{\sqrt[3]{\frac{64}{9}b(\log b)^2}})$.

The Shor algorithm - why is it an issue?

The difficulty in factorising the product of two large prime numbers is the basis of the RSA asymmetric algorithm used for key distribution. So we have the ability to encrypt using a symmetric key, but key exchange is a problem.

(or at least we won't have until such an algorithm can for sufficiently large numbers be executed on an existing quantum computer, but shhh)

Quantum cryptography

What is that exactly?

Using quantum mechanics methods for cryptographic purposes:

1. cryptanalysis, undermining the security of classical ciphers;
2. quantum methods of information encryption;
3. quantum key distribution methods.

1. Some classical ciphers, for example, are susceptible to facilitating problems that are computationally complex for classical computers (e.g. factorisation of prime numbers or computation of the discrete logarithm), but there are proven proposals for quantum algorithms that, when executed on suitable quantum computers, could allow fast computation of such problems. This problem applies to RSA, for example.
2. Quantum objects and their properties (e.g. superposition) can be used to encode information. For example, we can use the polarisation of photons for this purpose. Let us impose that the horizontal polarisation is 0 and the vertical polarisation is 1. We can also shift this arrangement by 45 degrees. So in this case, we have two possible types of encoding - normal and rotated by 45 degrees. Our key, then, is to know how we should position the polarisers in turn. If we set them up incorrectly, we will get either 0 or 1 with equal probability.
3. Since we are using symmetric methods such as RSA to exchange the symmetric key, maybe we can use a quantum method to accomplish this task?

Quantum key distribution

Since we use symmetric methods such as RSA for key distribution, maybe we can use some quantum method to accomplish this task?

1. the method does not necessarily allow for a consistent exchange of information;
2. two types of protocols: prepare and measure or entanglement-based;
3. prepare-and-measure methods exploit the property of superposition in an unknown (to the eavesdropper) quantum state;
4. entanglement methods exploit the properties of quantum entanglement - a measurement on one qubit will cause changes in the state of the other one.

The key here is to obtain an identical bit sequence (e.g. 128 bits) between the two parties to the conversation in a way that makes it impossible for a third party to intercept this. As we will see in example BB84, this property actually occurs despite the loss of a large portion of the initially generated bits of the sifted key.

Quantum key distribution - BB84

- It has been published in 1984 by Charles H. Bennett and Gilles Brassard.
- Alice and Bob are connected by a quantum communication channel, allowing for the transmission of quantum states (e.g. optical fibre).
- In addition - a public, classic, unsecured (but authenticated) channel.
- We assume that the eavesdropper (Eve) can interfere with, tap into etc. the quantum channel.
- We encode information using non-orthogonal states.

Quantum key distribution – BB84

1. Alice generates one bit (either 0 or 1), randomly chooses the base in which to encode this value, and then creates the corresponding qubit.
2. Transmission of the qubit follows – eavesdropping by Eve may occur.
3. Bob randomly selects a base in which to measure, records that base and the read bit.
4. Bob then announces over the public channel which base he used for the readout.
5. Alice checks which bases match and rejects where they do not match.

There are various ways in which this protocol can be attacked. The first one worth mentioning is the intercept-and-resend attack, which, as the name suggests, involves intercepting a qubit with encoded information and trying to read it. With two possible non-orthogonal bases, we have a 50% chance of succeeding. Such an attack, however, causes a corresponding quantum error that can be detected by the authentic communication parties. Another type of attack is photon number splitting. Weak lasers are usually used to create the photons used in communication, so just one photon per bit is prepared and sent away. Unfortunately, very often this is not possible and two or three photons are actually prepared. In this case, these 'excess' photons are set up in the same way as our photon intended for communication. So the eavesdropper can simply seize one of these redundant photons and store it, and measure it when the bases are announced. This problem can be combated by a suitable laser constitution, or by using a different protocol operating on the same method but with different announced elements, etc., e.g. SARG04, which was developed to counteract the effects of this very attack.

Quantum key distribution – BB84

Quantum transmission

Alice's random bits	0	1	1	0	1	1	0	0	1	0	1	1	0	0
Random sending bases	×	+	×	+	+	+	+	+	×	×	+	×	×	×
Photons Alice sends	↗	↑	↘	→	↑	↑	→	→	↘	↗	↑	↘	↗	↗
Photons receiving bases	+	×	×	+	+	×	×	+	×	+	×	×	×	×
Bits as received by Bob	1		1		1	0	0	0		1	1	1		0

Public discussion

Bob's reported bases	+		×		+	×	×	+		+	×	×		×
Alice's feedback			✓		✓			✓				✓		✓
Initial key			1		1			0				1		0
Bob reveals randomly					1									0
Alice's confirmation					✓									✓

Outcome

Sifted key			1					0				1		

Sources and further reading

1. IBM Quantum Learning, accessed on 2024-11-23
2. Menezes A., van Oorschot P., Vanstone S., Handbook of Applied Cryptography, CRC Press, Waterloo 1996, ISBN 0-8493-8523-7
3. Nielsen M., Chuang I., Quantum Computation and Quantum Information, Cambridge University Press, Cambridge 2000, ISBN 978-1-107-00217-3
4. A Short Introduction to Quantum Information and Quantum Computation, Michel Le Bellac, Cambridge University Press, 2006, ISBN 978-0-52-186056-7
5. Shannon C., A mathematical theory of communication, The Bell System Technical Journal, 27(3):379–423, 7 1948
6. Bennett, C. H.; Brassard, G., Quantum cryptography: Public key distribution and coin tossing, Proceedings of the International Conference on Computers, Systems & Signal Processing, Bangalore, India. Vol. 1. New York: IEEE. pp. 175–179, DOI:10.1016/j.tcs.2014.05.025, arXiV:2003.06557