

Introduction: What is crypto?

Crypto can be understood as *cryptography* or *cryptology*; the two are distinct notions, but this won't matter to us in this course.

*Crypto is the science of securing data
against adversaries.*

This is not just encryption, not just data transit (*e.g.* managing data), not all information security. Security is a lot broader than crypto: it contains crypto but also social engineering, finding implementation bugs, risk management, system security.

This first chapter will be more informal, to give a high-level overview of crypto.

1 Some basic examples.

1.1 How to check that two distant files are the same?

Instead of comparing the two files byte by byte, we can compare *cryptographic hashes*. They are a lot smaller in size compared to the original data. Some famous examples of cryptographic hash functions are SHA-256, SHA-2, SHA-3 or MD-5 (this last one is outdated).

1.2 Secure communication over the Internet.

HTTPS is a secure protocol for browsing the Internet. The “S” in HTTPS means *secure*. The underlying protocol is called TLS/SSL. When browsing, for example, Google, we want that our communication is secured against *eavesdropper*; usually this means encrypting the communications. This communication goes in multiple parts.

1. **Authentication.** Alice asks Bob some kind of identification, and Bob will send Alice a certificate of authentication.
2. **Handshake protocol.** Alice and Bob then create a shared secret key.
3. **Encrypted communication.** Finally, Alice and Bob can communicate without risking being eavesdropped. This is done with symmetric encryption.

All of these steps can be done while being eavesdropped, and this does not compromise the communication between Alice and Bob. This course will cover all the ingredients for this protocol:

1. Digital signature (for step 1)
2. Public-key encryption (for step 2)
3. Symmetric encryption (for step 3).

We will see all of this in the reverse order.

2 Summary.

The *core* of crypto is the three following ideas.

Authenticity. Am I certain to whom I am talking?

Integrity. Am I certain what I am saying is what the other person hears?

Confidentiality. Am I certain that our communication is private?

Encryption, what most people have in mind when talking about crypto, sits in the third step “Confidentiality.” However, without the other two, there is no point to encryption for secure communication.

As we can see from SSL/TSL, defining a secure protocol is complicated. We should define: functionality, security properties, who is the adversary, what the adversary is capable of. This last question can be divided in two subquestions:

- ▷ What is its computational resources?

Crypto experts try to define protocols that require at least 2^{128} operations for the adversary to access your secured communication. Even with a limited computational power on the user side, the adversary would require exponentially more computational power.

- ▷ What kind of access does it have to your data?

Having access to someone's phone or not having access to it makes a lot of difference in terms of security.

2.1 TL;DR.

Protocols are complex: we break them into simpler pieces called “cryptographic primitives” (*e.g.* digital signatures, hash functions, symmetric encryption, *etc*). The security of a complex system can be reduced to the security of its building blocks. One of the goal of this course is to define all these primitives and prove that they are secure.

3 Some more applications of crypto.

VPNs or the TOR network are more common examples of applications of crypto. One important part is to *define* what you are trying to hide. For example, your internet provider can still see, even while using a VPN, that you are watching videos as a lot of data is flowing (the ISP does not know the exact videos, but still know you are watching something).

Another very common example is cryptocurrencies, for example the famous BitCoin. Hashes and digital signatures are used a lot with these cryptocurrencies. E-cash is a little different, but with the same idea. How can you check that someone did not spend some amount of money twice? Guaranteeing that you cannot duplicate money.

Then, crypto is also used by E-voting. An E-voting system needs to make sure of a few properties: assure that there is no ballot

stuffing, verifiability at the end of the protocol, anonymity of the vote, authentication.

Also, the French app *TousAntiCovid* from the Covid pandemic used crypto. Here, functionality is important: in itself the protocol is not sure, and can be easily manipulated to leak its data, no matter how perfectly encrypted the data is.

Disk encryption is also using crypto. Is there a way to compute over encrypted data? This can be used as, for examples, LLMs such as ChatGPT are manipulating lots of data, and some of it can be confidential.

4 Conclusion.

Cryptography is a science that can be split in five steps:

1. Define functionality;
2. Define the goal of the attacker;
3. Define its capabilities;
4. Propose a realisation;
5. Prove the security of the realisation.

This last step is usually done by proving that, if an attack succeeds in breaking the protocol, then some unrealistic (computational) conditions are true. This is done via reduction, *e.g.* prove that attacking implies being able to solve famous intractable problem, for example the FACTORING problem. **NP**-complete problems, like SAT, are not always the answer as it is difficult to get a SAT instance that is hard to solve; to get a hard-to-solve instance for FACTORING, you only need to pick two large prime numbers which is a lot easier to do.

Crypto is related to complexity theory and the intractable problems we rely on are usually algebraic problems. Beautiful math is involved.

Crypto is not an art! Non-public protocols must be considered insecure. Protocols should be public and sources of security should be clearly identified and concentrated. If the protocol is not open-source, then it usually means that something sketchy is going on. The implementation can be private as a lot of hard work can be done in how to implement this protocol effectively, but the protocol should be public.