Overview of Cryptography

Hans-Wolfgang Loidl
http://www.macs.hw.ac.uk/~hwloidl

School of Mathematical and Computer Sciences
Heriot-Watt University, Edinburgh

Terminology

The main goal of encryption is *secure* communication between two parties. However, there are more applications, so we need to clarify terminology first.

Cryptography Terminology

cryptography: secret writing with ciphers
cryptanalysis: breaking ciphers
cryptology: both of above
encryption: transforming plain text to cipher text
decryption: recovering plain text from cipher text
cryptosystem: a mechanism for encryption and decryption
(also: **cipher, cryptosystem**)

Goals of cryptography

Cryptography can be used to help ensure these **security properties**:

- **confidentiality:** preventing unauthorised access
- **integrity:** preventing unauthorised modification
- **authentication:** verification of identity. Sometimes split into:
  - entity authentication
  - data origin authentication
- **non-repudiation:** preventing denial of actions

We want to ensure these properties, even when another party may eavesdrop or intercept messages. Carefully designed cryptographic protocols help this.

Attacks against cryptosystems

We assume that the attacker knows the algorithm being used for en-/de-cryption. It is bad design to rely on the confidentiality of the algorithm itself (security by obscurity).
Primary types of attacks are:

- **Ciphertext-only attack**: the attacker knows only one or more encrypted messages (cipher-texts), all using the same key.
- **Known-plaintext attack**: the attacker knows one or more plain-text, cipher-text pairs, all using the same key.
- **Chosen-plaintext attack**: the attacker can choose one or more plain-texts and get the corresponding cipher-text, all using the same key.
- **Chosen-ciphertext attack**: the attacker can choose one or more cipher-texts and get the corresponding plain-text, all using the same key.

Attacks against protocols

- **Known-key attack**: the attacker obtains some keys used previously and then uses this information to determine new keys.
- **Replay**: the attacker records a communication session and replays (part of) the session at some later point in time.
- **Impersonation**: the attacker assumes the identity of one of the legitimate parties in a network (an **interleaving attack** is a more sophisticated variant of this attack).
- **Dictionary**: the attacker takes a list of probable passwords, hashes all entries in this list, and then compares this to the list of true encrypted passwords (a **forward search** is a more sophisticated variant of this attack).

A Taxonomy of Cryptographic Primitives

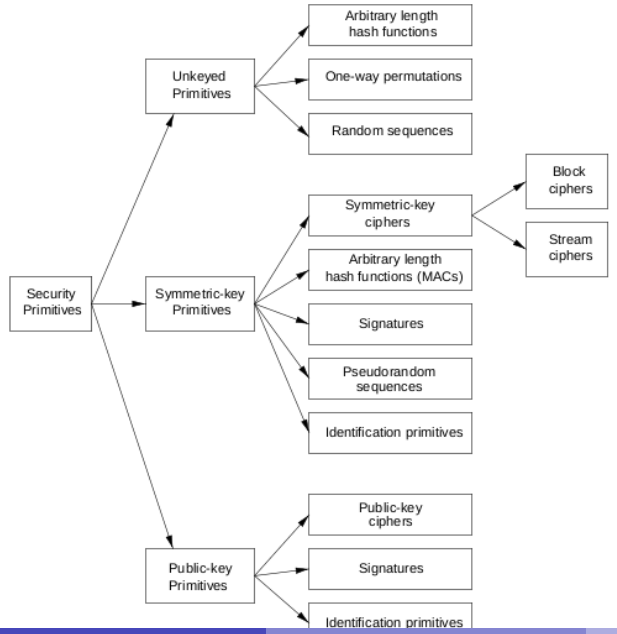

Figure 1.1: *A taxonomy of cryptographic primitives.*

Notation and example applications

- Hash functions $h(m)$
  - **integrity:** "fingerprint" provides tamper evidence
  - **message compression**: hash-then-sign schemes
- Symmetric block ciphers $E_k(m), D_k^{-1}(m)$
  - **bulk encryption**: network comms, data storage
- Public key (asymmetric) ciphers $E_e(m), D_d(m)$
  - **key exchange**: establishing shared keys for symmetric ciphers
- Digital signature schemes $S_A(m), V_A(m, s)$
  - **key signing**: public key infrastructures (PKIs)

Role of Cryptography

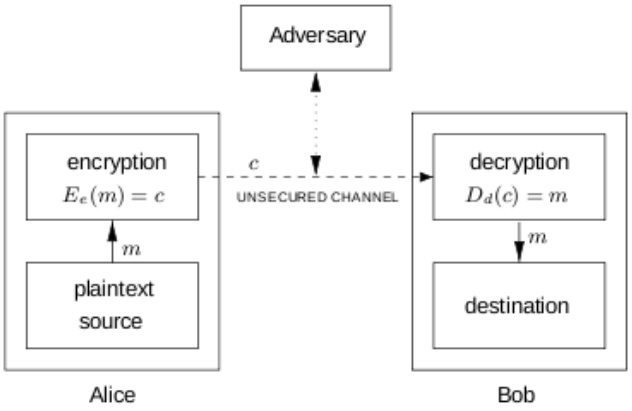

Figure 1.6: *Schematic of a two-party communication using encryption.*

[0]From *Handbook of Applied Cryptography*, CRC Press, 1997

Cryptographic Concepts

- The main goal of encryption is to enable confidential communication over insecure channels, subject to eavesdropping.
- Suppose, Alice wants to send a message M to Bob. M is called the **plaintext**.
- If M is sent unmodified, an eavesdropper Eve can intercept, and read the message.
- To secure the connection, the original message is encrypted, using an **encryption function** E
- The **ciphertext** C is the result of encrypting M i.e. $C = E(M)$
- C is designed in such a way that it is impossible to reconstruct the original text without having secret information, which Bob has.
- Now, C can be sent from Alice to Bob, without the danger of an attacker being able to read the message.

Bijections

- Recall that a **bijection** is a mathematical function which is one-to-one (injective) and onto (surjective).
- In particular, if $f : X \to Y$ is a bijection, then for all $y \in Y$, there is a unique $x \in X$ such that $f(x) = y$. This unique x is given by the inverse function $f^{-1} : Y \to X$.

Bijections are used as the basis of cryptography, for encryption. If f is an encryption transformation, then f^{-1} is the corresponding decryption transformation.

Bijections

Figure 1.3: *A bijection f and its inverse $g = f^{-1}$.*

[0]From *Handbook of Applied Cryptography*, CRC Press, 1997

Message spaces

We assume

- A set $\mathcal{M}$, the message space. $\mathcal{M}$ holds symbol strings, e.g., binary, English. Elements $m \in \mathcal{M}$ are called plaintexts.
- A set $\mathcal{C}$, the ciphertext space. $\mathcal{C}$ also consists of strings of symbols. Elements $c \in \mathcal{C}$ are called ciphertexts.
- Each space is given over some alphabet, a set $\mathcal{A}$. For example, we may consider $\mathcal{A}$ to be the letters of the English alphabet A-Z, or the set of binary digits $\{0, 1\}$.

Cryptography systems

- The encryption transformation is a bijection $E : \mathcal{M} \to \mathcal{C}$ i.e. it is a one-to-one mapping been plaintext and ciphertext.
- The decryption transformation is a bijection $D : \mathcal{C} \to \mathcal{M}$ i.e. it is a one-to-one mapping been ciphertext and plaintext.

Encryption and decryption transformations are indexed using keys.

- The key space $\mathcal{K}$ is a finite set of keys $k \in \mathcal{K}$.
- An encryption scheme consists of two sets indexed by keys
  - a family of encryption functions $\{E_e | e \in \mathcal{K}\}$
  - a family of decryption functions $\{D_d | d \in \mathcal{K}\}$
- such that for each $e \in \mathcal{K}$, there is a unique $d \in \mathcal{K}$ with $D_d = E_e^{-1}$. We call such a pair (e, d) a key pair.
- An encryption scheme is also known as a cryptography system (cryptosystem) or a cipher.

Example: Caesar Cipher

How it works:

> *. . . si qua occultius perferenda erant, per notas scripsit, id est sic structo litterarum ordine, ut nullum verbum effici posset; quae si qui investigare et persequi velit, quartam elementorum litteram, id est D pro A et perinde reliquas commutet — Suetonius, Julius Caesar 56*
>
> *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: Caesar Cipher

To summarise:

> *to encrypt a plaintext message M, take every letter in M and shift it by e elements to the right to obtain the encrypted letter; to decrypt a ciphertext, take every letter and shift it by $d = -e$ elements to the left*

As an example, using $e = 3$ as key, the letter *A* is encrypted as a *D*, *B* as an *E* etc.

```
Plain:    ABCDEFGHIJKLMNOPQRSTUVWXYZ
Cipher:   DEFGHIJKLMNOPQRSTUVWXYZABC
```

Encrypting a concrete text, works as follows:

```
Plaintext:  the quick brown fox jumps over the lazy dog
Ciphertext: WKH TXLFN EURZQ IRA MXPSV RYHU WKH ODCB GRJ
```

More formally we have the following functions for en-/de-cryption:

$$E_e(x) = x + e \mod 26$$

$$D_e(x) = x - e \mod 26$$

Characteristics of Caesar's Cipher

Note the following:

- The sets of plain- and cipher-text are only latin characters. We cannot encrypt punctuation symbols etc.
- The en- and de-cryption algorithms are the same. They only differ in the choice of the key.
- The key strength is not tunable: shifting by 4 letters is no more safe than shifting by 3.
- This is an example of a **symmetric or shared-key cryptosystem**.

Question

Is this a strong form of encryption? How can you break this encryption?

Exercise

Implement an en-/de-cryption function based on the Caesar cipher. Implement a function that tries to crack a Caesar cipher, ie. that retrieves plaintext from ciphertext for an unknown key.

Background: Caesar Cipher

An early expert in cryptography, and other areas, was Gaius Julius Caesar (13 July 100 BC – 15 March 44 BC).
Background: Caesar's Life[1]

- Caesar was born into a patrician family, the **gens Julia**
- During his childhood several wars from 91 BC to 82 BC and 82 BC to 80 BC ravaged Rome (under dictator Lucius Cornelius Sulla).
- Caesar, as the nephew of Marius and son-in-law of Cinna, was targeted by Sulla's proscriptions.
- He entered the Roman **cursus honorum** and was elected military tribune, a first step in a political career.
- He was elected quaestor for 69 BC.
- In 63 BC he was elected Pontifex Maximus, chief priest of the Roman state religion.
- He was elected praetor, the second highest position in the cursus honorum.

[1] Source: Wikipedia `http://en.wikipedia.org/wiki/Julius_Caesar`

Background: Caesar's Life (cont'd)

- Then he became governor of Spain, which gave him the chances for military success and to amass a fortune.
- Caesar was elected consul for 59 BC in an election full of (alleged) bribery.
- First Triumvirate between Caesar, Gnaeus Pompeius Magnus, Marcus Licinius Crassus.
- After his consulship, he was appointed to govern Cisalpine Gaul (northern Italy) and Illyricum (southeastern Europe), with Transalpine Gaul (southern France) later added
- In this position he conquered most of Gaul (today's France) from 58 BC to 51 BC.
- In 50 BC, the Senate, led by Pompey, ordered Caesar to disband his army and return to Rome because his term as governor had finished.
- In January 49 BC, Caesar crossed the Rubicon river with only one legion and ignited civil war.
- Caesar defeated Pompey at Pharsalus in 48 BC.

Background: Caesar's Life (cont'd)

- In Rome, Caesar was appointed dictator, then elected consul and stepped down as dictator.
- Late in 48 BC, Caesar was again appointed Dictator, with a term of one year.
- After a victory over the last remaining Pompeyans, he was appointed Dictator for ten years.
- On the Ides of March (15 March) of 44 BC, Caesar was assassinated by a group of republican senators, including Marcus Junius Brutus

Further reading:

- Wikipedia entry on Julius Caesar: `http://en.wikipedia.org/wiki/Julius_Caesar`
- Suetonius: The Life of Twelve Caesars: `http://penelope.uchicago.edu/Thayer/E/Roman/Texts/Suetonius/12Caesars/home.html`
- *"De Bello Gallico"*, Gaius Julius Caesar `http://wiki.dickinson.edu/index.php/Caesar_Gallic_War`

Symmetric and Asymmetric Cryptosystems

- **symmetric cryptography**:
  - the keys for en-/de-cryption (e and d) are (essentially) the same
  - aka secret-key, shared-key, single-key, conventional
- **asymmetric cryptography**:
  - Given e, it is (computationally) infeasible to find d.
  - aka public-key (PK), since e can be made public.
- There are more differences between symmetric and asymmetric cryptosystems than the key-pair relation, but this is the characteristic one.
- In later lectures we will refine the meaning of "essentially" and "computationally infeasible."

Symmetric Cryptosystems

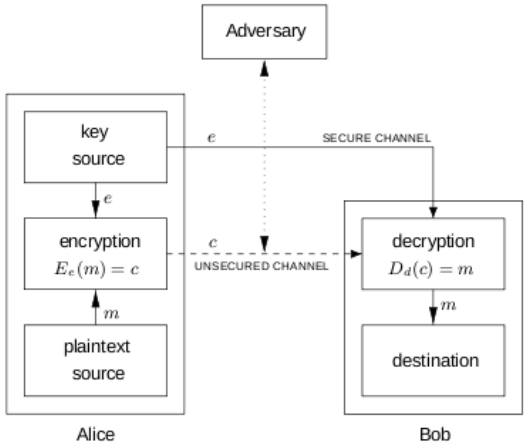

Figure 1.7: *Two-party communication using encryption, with a secure channel for key exchange. The decryption key d can be efficiently computed from the encryption key e.*

[1] From: "Handbook of Applied Cryptography"

Foundations of Asymmetric Cryptosystems

A function $f : X \to Y$ is called a one-way function if

- it is feasible to compute $f(x)$ for all $x \in X$, but
- it is infeasible to find any x in the pre-image of f, such that $f(x) = y$, for a randomly chosen $y \in Im(f)$. (If f is bijective, this means it is infeasible to compute $f^{-1}(y)$).

Question

Is a one-way function useful for encryption?

Note: The definition above is vague: to be exact, we should give precise notions of feasible and infeasible. This is possible, but so far no-one has proved the existence of a true one-way function. Some functions used in modern ciphers are properly called candidate one-way functions, which means that there is a body of belief that they are one-way.

Trapdoor one-way functions

A trapdoor one-way function is a one-way function f that has a *trapdoor*: given some additional information, it is feasible to compute an x such that $f(x) = y$, for any $y \in Im(f)$.

Question

Is a trapdoor one-way function useful for encryption?

- Again, we know candidates, but no function has yet been proved to be a trapdoor one-way function.
- In principle, there is a possibility of breaking crypto systems by new algorithms based on advances in mathematics and cryptanalysis.
- It's unlikely that one-way functions do not exist; some hash functions are as secure as NP-complete problems.
- Catastrophic failure for present functions is less common than gradual failure due to advances in computation power and (non-revolutionary but clever) algorithms or cryptanalysis, bringing some attacks closer to feasibility.

Asymmetric Cryptosystems

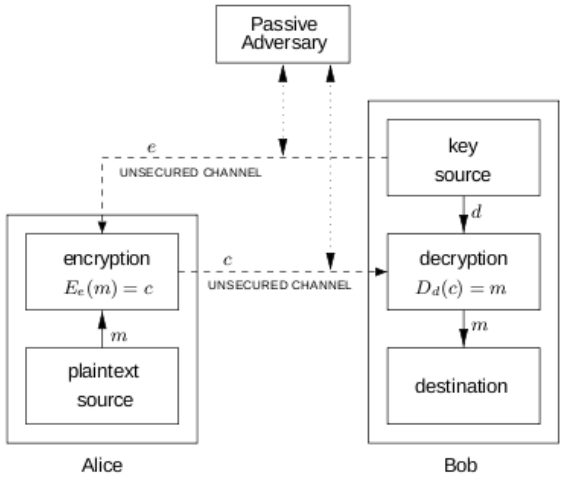

Figure 1.11: *Encryption using public-key techniques.*

[1] From: "Handbook of Applied Cryptography"

Asymmetric Cryptosystems

The idea of **asymmetric** cryptography (public-key cryptography):

- Every participant in a communication has two keys
  - a **private key**, which is kept secret
  - a **public key**, which can be published e.g. on the web
- For safe communication,
  - the sender uses the recipient's **public key** to **encrypt** the data,
  - the recipient uses his **private key** to **decrypt**.
- Since the private key is kept secret, only the recipient can read the message.
- This technology relies on the fact, that the private key cannot be efficiently computed from just knowing the public key (and the crypto algorithm).
- This idea was a **major research breakthrough** in the area of cryptography.

Symmetric vs Asymmetric Cryptosystems

Advantages of **symmetric** cryptosystems:

- **Performance:** Symmetric crypto-algorithms are typically much faster than asymmetric ones, and achieve higher throughput.
- **Compositionality:** Ciphers can be combined to achieve stronger encryption.
- **Key sizes:** Due to sophisticated ciphers, keys can be shorter
- **Pragmatics:** Long history

Disadvantages of **symmetric** cryptosystems:

- **Management:** Communication key must remain secret
- **Scalability:** For secure communication a separate key is needed for every pair of communication
- **Stability**: Keys should be changed frequently

Question

How many symmetric keys are needed for secure communication among n participants?

Symmetric vs Asymmetric Cryptosystems (cont'd)

Advantages of **asymmetric** cryptosystems:

- **Easier to manage:** Only 1 key of the key-pair must be kept secret.
- **Better scalability:** For communication among *n* participants only their *n* public keys are needed.
- **Flexibility:** Encryption technology can be naturally used for electronic signatures.
- **Stability**: Keys can remain unchanged over a long period of time

Disadvantages of **asymmetric** cryptosystems:

- **Performance:** Public-key algorithms are significantly slower
- **Key sizes:** Keys are typically much longer
- **Foundations:** No public-key scheme has been proven to be secure
- **Pragmatics:** Shorter history.

In order to combine the advantages of both, often strong, asymmetric encryption is only used on a **session key**, which is the symmetric key used to encrypt the bulk of the message.

Man-in-the-middle Attacks

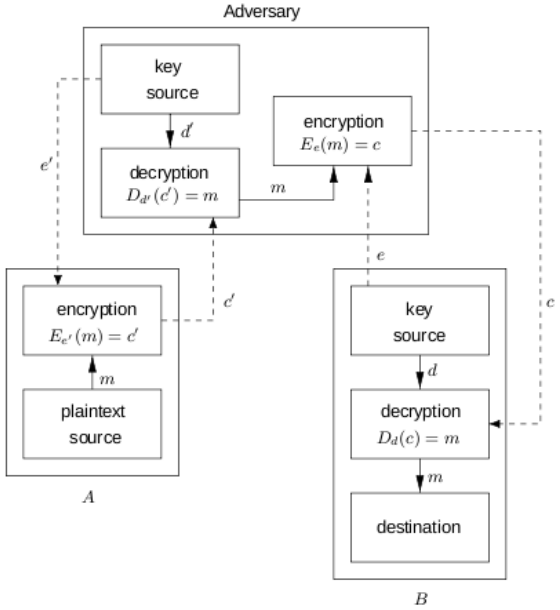

Figure 1.13: *An impersonation attack on a two-party communication.*

[1] From: "Handbook of Applied Cryptography"

Further Reading

- Michael T. Goodrich and Roberto Tamassia *"Introduction to Computer Security"*, Addison Wesley, 2011. ISBN: 0-32-151294-4
- Alfred J. Menezes, Paul C. van Oorschot and Scott A. Vanstone, *"Handbook of Applied Cryptography"*, CRC Press, 2001. ISBN 0-8493-8523-7. On-line: `http://www.cacr.math.uwaterloo.ca/hac/`
- Nigel Smart, *"Cryptography: An Introduction"*, On-line: `http://www.cs.bris.ac.uk/~nigel/Crypto_Book/`