Introduction to Symmetric Cryptography

Lars R. Knudsen

June 2014

What is cryptography?

Cryptography is communication in the presence of an adversary
Ron Rivest.

Coding theory

Detection and correction of random errors

Cryptography

Detection and protection of hostile “errors”

What is cryptography about?

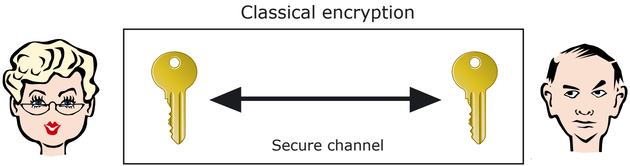
Secrecy (confidentiality)

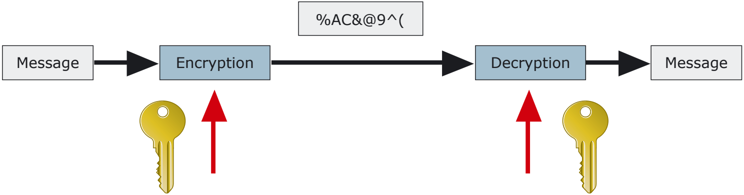
Keeping things secret (data, communication, entity, etc.)

Authentication

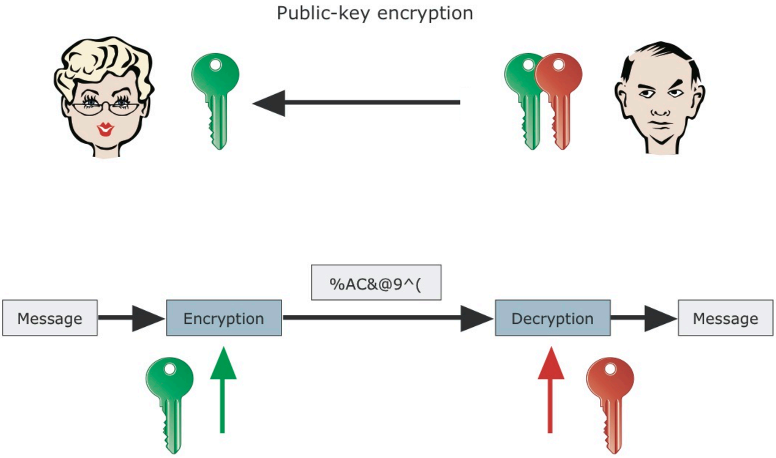
Assurance about authenticity (of data, origin, entity, etc.)

Symmetric encryption





Public-key encryption



Public-key versus symmetric cryptosystems

	Advantages	**Disadvantages**
Symmetric	fast systems	secure key-exchange
Public-key	no secure key-exchange	slow systems

Hybrid encryption

Introduction to symmetric cryptosystems

Cryptosystem $(\mathcal{P}, \mathcal{C}, \mathcal{K}, \mathcal{E}, \mathcal{D})$

$\mathcal{P}$: set of plaintexts

$\mathcal{C}$: set of ciphertexts

$\mathcal{K}$: set of keys

$\mathcal{E}$: for $k \in \mathcal{K}$: $e_k(x)$ encryption rule

$\mathcal{D}$: for $k \in \mathcal{K}$: $d_k(x)$ decryption rule

For every $k \in \mathcal{K}$: it holds for all m that $d_k(e_k(m)) = m$

Symmetric encryption

Kerckhoffs' principle

Everything is known to an attacker except for the value of the secret key.

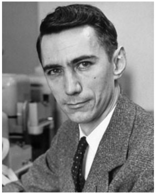
Attack scenarios

- Ciphertext only
- Known plaintext
- Chosen plaintext/ciphertext
- Adaptive chosen plaintext/ciphertext (black-box)

Typical goal

High security even under black-box attack

Claude E. Shannon, 1916-2001



Communication Theory of Secrecy Systems, published in 1949.

Theory

First person to establish a theory for provable security.

Principles

His ideas for building (symmetric) ciphers still used today.

Shannon's Theory

Definition

Perfect secrecy $\iff \Pr_{\mathcal{P}}(x|y) = \Pr_{\mathcal{P}}(x), \forall x \in \mathcal{P}, y \in \mathcal{C}$

Fact

A cryptosystem where $|\mathcal{K}| = |\mathcal{P}| = |\mathcal{C}|$ provides perfect secrecy if and only if

1. $Pr_{\mathcal{K}}(K) = \frac{1}{|\mathcal{K}|}, \forall K \in \mathcal{K}$
2. $\forall x \in \mathcal{P}, y \in \mathcal{C}, \exists$ *unique* K *such that* $e_K(x) = y$

Example

One-time pad: $e_K(x_1, \ldots, x_n) = (x_1 \oplus k_1, \ldots, x_n \oplus k_n)$

- All keys equally likely
- Each key used only once
- Key as long as plaintext and ciphertext

Definition (Redundancy)

R_L: which percentage of a language L is redundant

Example

th weathr is nice 2d.
R_L for English is 75%.

Definition (Unicity distance)

minimum number of ciphertext blocks attacker needs in order to be able to uniquely identify secret key

$$t_0 \simeq \frac{log_2(|\mathcal{K}|)}{R_L log_2(|\mathcal{P}|)}$$

$t_0 = \min_t$: s.t. essentially only one value of the key could have encrypted $c_1, \ldots, c_t$

Unicity distance in known/chosen plaintext attack

Question

What is the unicity distance under a known plaintext attack ??

Assume that we are given t encryptions, that is, the plaintext blocks and the corresponding ciphertext blocks.

Question - again

How big does t have to be, before it is likely that only one value of the key could have encrypted the texts?

$$t_1 = \frac{log_2(|\mathcal{K}|)}{log_2(|\mathcal{P}|)}$$

$t_1 = \min_t$: s.t. essentially only one value of the key could have encrypted m_1 to c_1, m_2 to c_2, ..., m_t to c_t

Definition (Confusion)

The ciphertext statistics should depend on the plaintext statistics in a manner too complicated to be exploited by the cryptanalyst

Definition (Diffusion)

Each digit of the plaintext and each digit of the secret key should influence many digits of the ciphertext

- Substitutions (confusion)
- Permutations (diffusion)
- Product = Substitution $\times$ Permutation

Most popular symmetric ciphers are product ciphers

Shannon's Thoughts

Question

How can we be sure an attacker will require a large amount of work to break a non-perfect system with *every* method???

Hard to achieve! But we can at least

Thoughts/ideas

1. make it secure against all known attacks, and/or
2. make it reducible to some known difficult problem

1. is what is done today in symmetric cryptography
2. is what is done today in public-key cryptography

From classical crypto to modern crypto

looking back..

- (almost) all ciphers before 1920s very weak
- 1920s, rotor machines, mechanical crypto
  - Enigma, Germany
  - Sigaba, USA
  - Typex, UK
- 1949, Shannon's work
- 1970s, computers take over from rotor machines
- ciphers operate on long sequence of bits (bytes)

Symmetric encryption today - two types

Block cipher

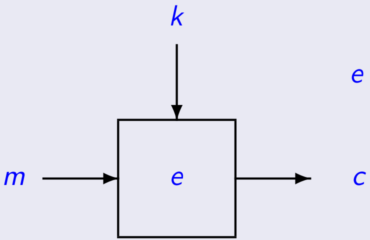
- Operate on from 8 to 16 bytes typically
- No or small internal state

Stream cipher

- Operate on from 1 bit to 4 bytes typically
- Internal state, can be big?

Block ciphers

Input block m, output block c, key k



$$e : \{0,1\}^n \times \{0,1\}^\kappa \to \{0,1\}^n$$

- given k easy to encrypt and decrypt
- given m, c hard to compute k, such that $e_k(m) = c$
- one-way function: $f(k) = e_k(m_0)$ for fixed m_0

Block ciphers

Applications

- block encryption (symmetric)
- stream ciphers
- message authentication codes
- building block in hash functions
- one-way functions

Block ciphers

Block cipher, n-bit blocks, κ-bit key

Family of 2^κ n-bit bijections

How many n-bit bijections are there?

$2^n! \simeq (2^{n-1})^{2^n}$

Design dream/aim

2^κ bijections chosen uniformly at random from all $2^n!$ bijections

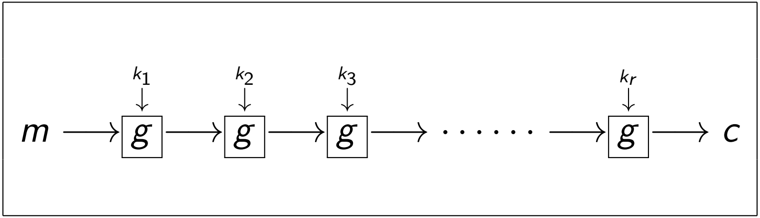
Famous block ciphers

	block size, n	key size, κ	year
DES	64	56	1977
Kasumi	64	128	1999
AES	128	128, 192, 256	2000
Present	64	80, 128	2007

Ciphers pick only a tiny fraction of all possible n-bit bijections

Unicity distance, known-plaintext attack?

Iterated block ciphers (DES, AES, ...)



- plaintext m, ciphertext c, key k
- key-schedule: user-selected key $k \rightarrow k_0, \ldots, k_r$
- round function, g, weak by itself
- idea: g^r, strong for "large" r

DES

Data Encryption Standard

- blocks: 64 bits, keys: 56 bits
- iterated cipher, 16 rounds
- developed in early 70's by IBM using 17 man years
- evaluation by National Security Agency (US)
- 1977: publication of FIPS 46 (DES)
- 1991: differential cryptanalysis, 2^{47} chosen plaintexts
- 1993: linear cryptanalysis, 2^{45} known plaintexts
- 1999: world-wide effort to find one DES-key: 22 hours

AES

Advanced Encryption Standard

- blocks: 128 bits
- keys: choice of 128-bit, 192-bit, and 256-bit keys
- iterated cipher, 10, 12 or 14 iterations depending on key
- FIPS (US governmental) encryption standard
- open (world) competition announced January 97
- October 2000: AES=Rijndael

Cryptanalysis

Assumption

Assume cryptanalyst has access to black-box implementing the cipher with secret key k

Aims of cryptanalyst

- find key k, or
- find (m, c) such that $e_k(m) = c$ for unknown k, or
- show non-random behaviour of the cipher

Generic attacks. Block size n, key size κ

Exhaustive key search

- try all keys, one by one
- $\lceil \kappa/n \rceil$ texts, time 2^{κ}, storage small

Table attack

- store $e_k(m_0)$ for all k
- storage 2^{κ}, time (of attack) small

Trade-offs

- Hellman tradeoff, $2^{2\kappa/3}$ time, $2^{2\kappa/3}$ memory

Generic attacks (continued)

Dictionary and birthday attacks on block ciphers

- known plaintexts: Collect pairs (m, c)
- ciphertext-only: Collect ciphertexts, look for matches $c_i = c_j$.

Example (CBC mode)

1. Collect $2^{n/2}$ ciphertext blocks
2. With 2 equal ciphertext blocks
   $c_i = c_j \Rightarrow e_k(m_i \oplus c_{i-1}) = e_k(m_j \oplus c_{j-1})$
   $\Rightarrow m_i \oplus m_j = c_{i-1} \oplus c_{j-1}$

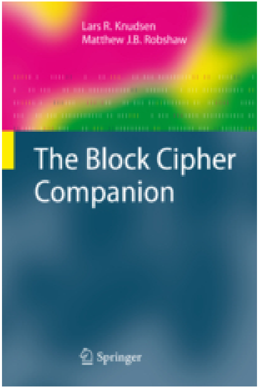
(similar attacks for ECB and CFB)

Short-cut attacks

Success dependent on intrinsic properties of $e(\cdot)$

- Differential cryptanalysis
- Linear cryptanalysis
- Higher-order differentials. Truncated differentials. Boomerang attack. Rectangle attack
- Integral attack. Related key attack. Interpolation attack
- Multiple linear cryptanalysis. Zero-correlation attack
- Side-channel cryptanalysis

By Lars R. Knudsen and Matt Robshaw.



Available online for free via Springer,
hard copies also available from Springer, Amazon etc.