

Information Security

CS526

Overview of Cryptography

Goals of Cryptography

- The most fundamental problem cryptography addresses: ensure security of communication over insecure medium
- What does secure communication mean?
 - privacy (secrecy, confidentiality)
 - only the intended recipient can see the communication
 - authenticity (integrity)
 - the communication is generated by the alleged sender
- More generally, achieve objectives even when there may be adversaries (bad guys)

Approaches to Secure Communication

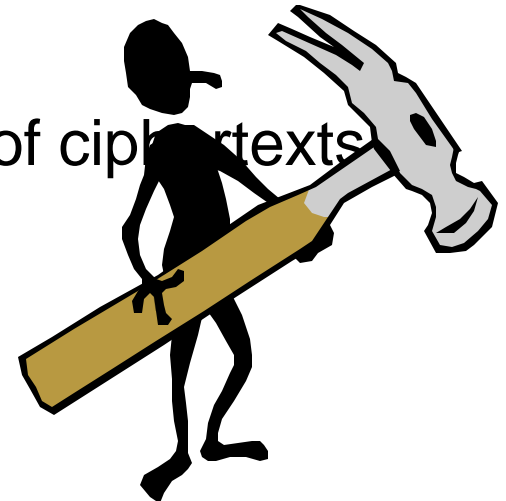
- Steganography
 - “covered writing”
 - hides the existence of a message
- Cryptography
 - “hidden writing”
 - hide the meaning of a message

Symmetric Ciphers

- A Cipher ($K, P, C, \mathbf{K}, \mathbf{E}, \mathbf{D}$)
 - K : the key space
 - P : the plaintext space
 - C : the ciphertext space
 - \mathbf{K} : the key generation function
 - $\mathbf{E}: K \times P \rightarrow C$: the encryption function
 - $\mathbf{D}: K \times C \rightarrow P$: the decryption function
- Sender and receiver must share a common key
- Key needs to be distributed through a secure channel

Adversarial Models for Encryption Schemes

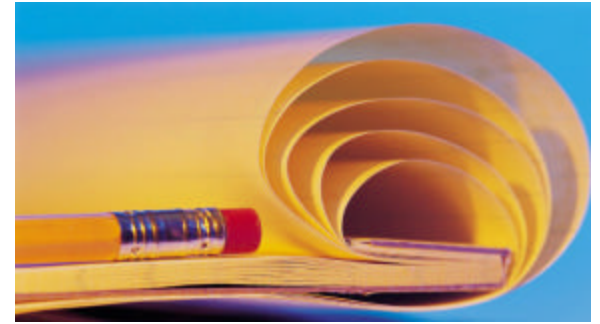
- **Ciphertext-only attack:**
The adversary knows a number of ciphertexts.
- **Known-plaintext attack:**
The adversary knows some random pairs of ciphertext and corresponding plaintext.
- **Chosen-plaintext attack:**
The adversary can choose a number of messages and obtain the ciphertexts for them
- **Chosen-ciphertext attack:**
The adversary can choose a number of ciphertexts and obtain the plaintexts



Common Notions of Security for Encryption

- Shannon (Information Theoretically) Secure, a.k.a. Perfect Secrecy
 - no information about the plaintext in the ciphertext
- Computational Security (consider adversaries that are limited in computation resources):
 - Semantically secure: no information can be learned by the adversary
 - Equivalently, Indistinguishability Secure: attacker cannot distinguish ciphertexts of two messages
 - e.g., IND-CPA, IND-CCA

One-Time Pad



How it works

- A message is encoded as a binary string
- A key is a **random** binary string that is at least as long as the message
- Encryption is by bit-xor

Properties

- Provides perfect secrecy, is informational theoretically secure
- Require key length at least as long as the message to be sent

Stream Ciphers

- Idea: approximate one-time pad by replacing **random** by **pseudo random**
 - Use Pseudo Random Number Generator (PRNG)
 - Secret key is the seed
 - $E_{\text{seed}}[M] = M \oplus \text{PRNG}(\text{seed})$
 - $D_{\text{seed}}[C] = C \oplus \text{PRNG}(\text{seed})$
 - PRNG: $\{0,1\}^s \rightarrow \{0,1\}^n$
 - expand a short random seed into a long bit string that “looks random”

Example Stream Ciphers

- RC4 (widely used)
 - a byte oriented stream cipher
- Linear Feedback Shift Register:
 - vulnerable to known-plaintext attack
 - a m -stage LFSR can be completely broken given $2m$ bits outputs

Properties of Stream Ciphers

- Security depends on strength of PRNG
- PRNG must be “unpredictable”
 - Don’t use UNIX rand for crypto!
- Never reuse any part of an output stream
- Typical stream ciphers are very fast
- Widely used:
 - SSL (RC4), Cell phones, DVD (LFSR)

Block Ciphers

- A block cipher algorithm operates on a block
 - DES uses 64-bit blocks, and 56-bit key
 - AES uses 128-bit blocks, key can be 128, 192, 256 bits
- Security of block ciphers
 - When a random key is picked, the cipher should behave like a random permutation

Encryption Modes

- How to encrypt a message
 - A message is divided into blocks
 - Different encryption modes may be used
- Electronic Code Book (ECB)
- Cipher Block Chaining (CBC)
- Cipher Feedback (CFB)
- Output feedback (OFB)
- Counter Mode (CTR)

Block Cipher Encryption Modes: ECB

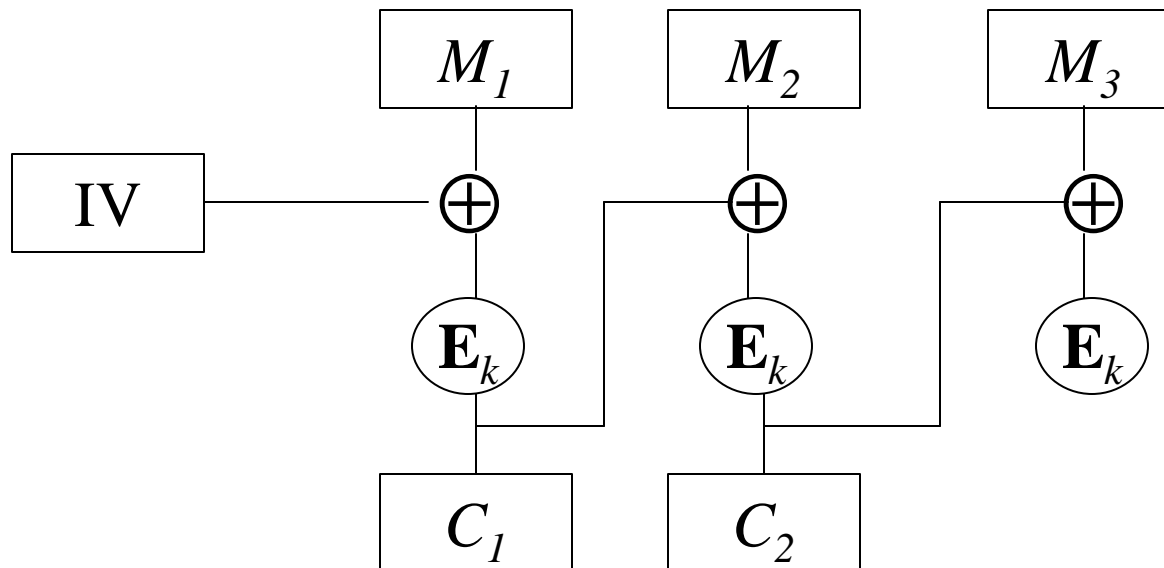
- **Electronic Codebook Book (ECB)**: each block encrypted separately.
- Not secure: it is deterministic, the same data gets encrypted the same way, **vulnerable if data repeats**, reordering ciphertext determines reordered plaintext.
- Errors in one block do not propagate.

Block Cipher Encryption Modes: CBC

- Cipher Block Chaining (CBC): next input depends of previous output

Encryption: $C_i = E_k (M_i \oplus C_{i-1})$, with $C_0 = IV$

Decryption: $M_i = C_{i-1} \oplus D_k (C_i)$, with $C_0 = IV$



Properties of CBC

- Errors in one block propagate, but only two blocks away
- Randomized encryption: repeated text gets mapped to different encrypted data.
 - can be proven to be semantically secure
- When IV is randomized, can be proven to be semantically secure, assuming that the block cipher is Pseudo Random Permutation (PRP)
- Sequential encryption, cannot use parallel hardware

Towards Public Key Encryption?

- Key distribution is a very difficult problem in symmetric encryption
- Basic Idea:
 - Encryption key and decryption key are different
 - Given encryption key, one cannot find out the decryption key
 - Encryption key can be made public, only decryption key needs to be kept private
 - Every one has a pair of public/private key

Common Public Key Encryption Algorithm

- RSA
 - security related to factoring large numbers
 - not semantically secure
- El Gamal
 - security related to discrete logarithm
 - randomized encryption

Achieving Integrity/Authenticity

- Unkeyed approach: using cryptographic hash functions
 - send the hash value of a message securely
- Symmetric approach: Using Message Authentication Code (MAC)
- Asymmetric approach: Using digital signature

Requirements for Cryptographic Hash Functions

Given a function $h: X \rightarrow Y$, then we say that h is:

- **preimage resistant (one-way):**
if given $y \in Y$ it is computationally infeasible to find a value $x \in X$ s.t. $h(x) = y$
- **2-nd preimage resistant (weak collision resistant):**
if given $x \in X$ it is computationally infeasible to find a value $x' \in X$, s.t. $x' \neq x$ and $h(x') = h(x)$
- **collision resistant (strong collision resistant):**
if it is computationally infeasible to find two distinct values $x', x \in X$, s.t. $h(x') = h(x)$

Hash Family

- A hash family is a four-tuple (X, Y, K, H) , where
 - X is a set of possible messages
 - Y is a finite set of possible message digests
 - K is the keyspace
 - For each $K \in K$, there is a hash function $h_K \in H$. Each $h_K: X \rightarrow Y$
- Alternatively, one can think of H as a function $K \times X \rightarrow Y$

Message Authentication Code

- A MAC scheme is a hash family, used for message authentication
- $MAC = C_K(M)$
- The sender and the receiver share K
- The sender sends $(M, C_K(M))$
- The receiver receives (X, Y) and verifies that $C_K(X)=Y$, if so, then accepts the message as from the sender
- To be secure, an adversary shouldn't be able to come up with (X, Y) such that $C_K(X)=Y$.

MAC Construction

- Based on hash functions
 - HMAC (widely used, provable secure)
- Based on block ciphers
 - CBC-MAC

Authentication with public keys: Digital Signature

- Each party has a private key and a corresponding public key
 - using the private key, one can compute an digital signature for any message
 - everyone can use the public key to verify that the digital signature is authentic
- Digital signature algorithms
 - RSA
 - DSA, Schnorr, etc

Summary

	Symmetric	Asymmetric
Confidentiality (Secrecy)	<ul style="list-style-type: none">•stream ciphers•block ciphers + modes	<ul style="list-style-type: none">•public key encryption
Integrity (Authenticity)	<ul style="list-style-type: none">•MAC	<ul style="list-style-type: none">•digital signature

Notions of Security for Encryption

	Key Recovery	Plaintext Recovery	Ciphertexts Distinguishing
Ciphertext Only			
Known Plaintext			IND-KPA
Chosen plaintext			IND-CPA
Chosen ciphertext			IND-CCA

Other Objectives in Cryptography

- Zero-knowledge proof
- Commitment schemes
- Oblivious transfer
- Secure Multi-party Computation (Secure Function Evaluation)