CSCE 465 Computer & Network Security

Instructor: Abner Mendoza

Introduction to Cryptography

Roadmap

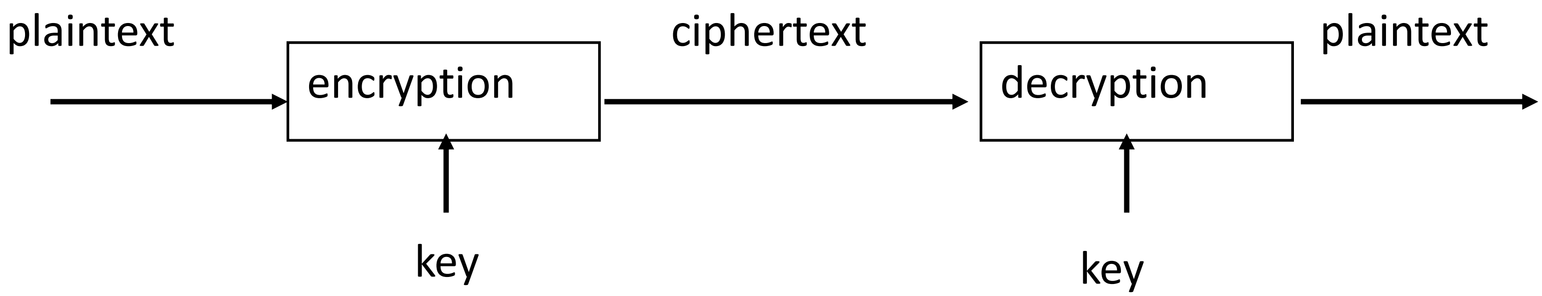
- Basic Crypto Concepts and Definitions
- Some Early (Breakable) Cryptosystems
- “Key” Issues

Basic Concepts and Definitions

Cryptography

- *Cryptography*: the art of secret writing
- Converts data into unintelligible (random-looking) form
  - Must be *reversible* (can recover original data without loss or modification)
- Not the same as compression
  - Usually *n* bits in, *n* bits out
  - Can be combined with compression
    - What's the right order?

Encryption/Decryption



- Plaintext: a message in its original form
- Ciphertext: a message in the transformed, unrecognized form
- Encryption: the process that transforms a plaintext into a ciphertext
- Decryption: the process that transforms a ciphertext to the corresponding plaintext
- Key: the value used to control encryption/decryption.

Cryptanalysis

- “code breaking”, “attacking the cipher”
- Difficulty depends on
  - sophistication of the cipher
  - amount of information available to the code breaker
- Any cipher can be broken by exhaustive trials, but rarely practical
  - When can you recognize if you have succeeded?

Breaking an Encryption Scheme

- Ciphertext only:
  - Exhaustive search until “recognizable plaintext”
  - Need enough ciphertext
- Known plaintext:
  - Secret may be revealed (by spy, time), thus <ciphertext, plaintext> pair is obtained
  - Great for monoalphabetic ciphers
- Chosen plaintext:
  - Choose text, get encrypted
  - Useful if limited set of messages

Brute Force Attacks

- Number of encryption/sec: 1 million to 1 billion/sec
- 56-bit key broken in 1 week with 120,000 processors ($6.7m)
- 56-bit key broken in 1 month with 28,000 processors ($1.6m)
- 64-bit key broken in 1 week with 3.1×10^7 processors ($1.7b)
- 128-bit key broken in 1 week with 5.6×10^{26} processors

The “Weakest Link” in Security

- Cryptography is rarely the weakest link
- Weaker links
  - Implementation of cipher
  - Distribution or protection of keys

Models for Evaluating Security

- Unconditionally Secure (Perfectly Secure)
  - Uncertainty/entropy $H(p)=H(p|c)$
- Provably Secure
  - As difficult to break as solving well-known and *supposedly* difficult problem
- Computationally Secure

Secret Keys v.s. Secret Algorithms

- Security by obscurity
  - We can achieve better security if we keep the algorithms secret
  - Hard to keep secret if used widely
  - Reverse engineering, social engineering

- Publish the algorithms
  - Security of the algorithms depends on the secrecy of the keys
  - Less unknown vulnerability if all the smart (good) people in the world are examine the algorithms

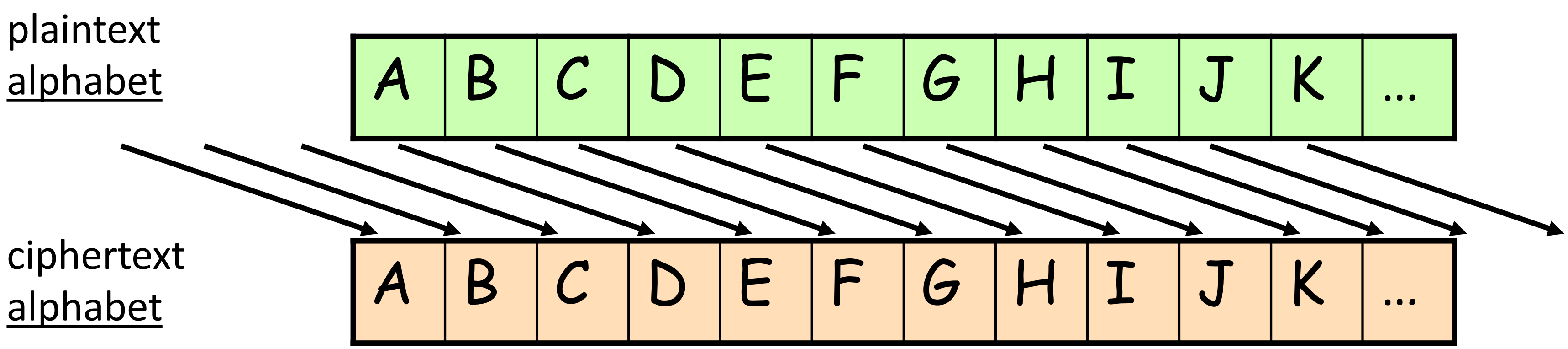
Secret Keys v.s. Secret Algorithms (Cont'd)

- Commercial world
  - Published
  - Wide review, trust
- Military
  - Keep algorithms secret
  - Avoid giving enemy good ideas
  - Military has access to the public domain knowledge anyway.

Some Early Ciphers

Caesar Cipher

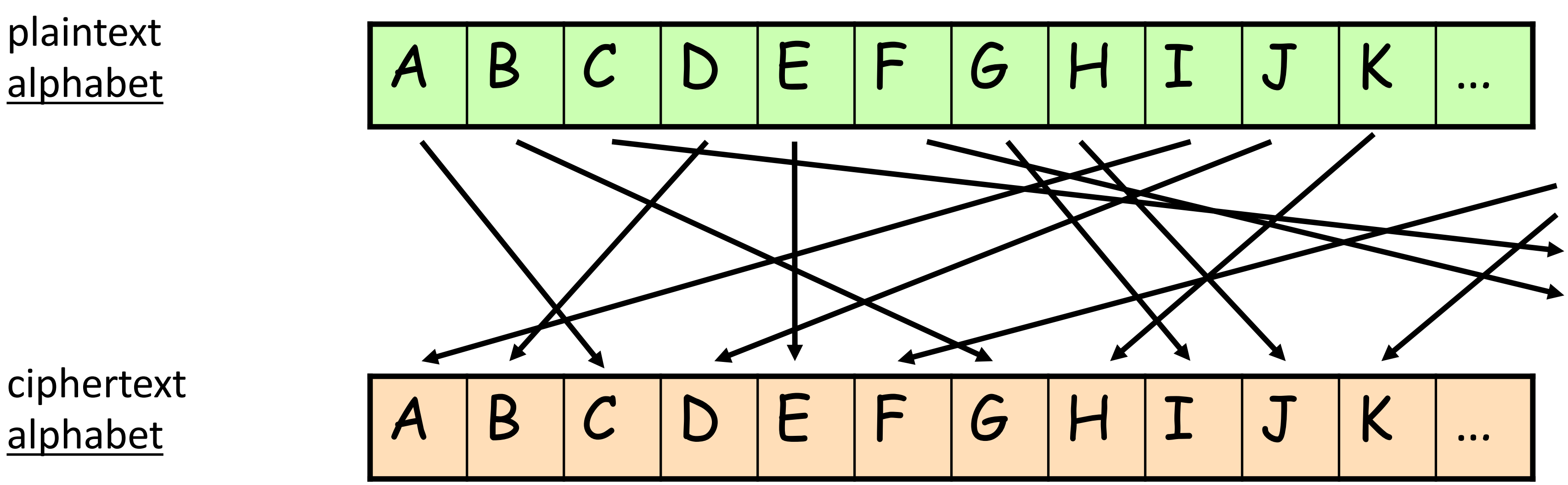
- Replace each letter with the one n (e.g., 3) letters later in the alphabet
  - ex.: plaintext CAT → ciphertext FDW



Trivial to break

Mono-Alphabetic Ciphers

- Generalized substitution cipher: an arbitrary (but fixed) mapping of one letter to another
  - 26! ($\approx 4.0*10^{26} \approx 2^{88}$) possibilities
- The key must specify which permutation; how many bits does that take?



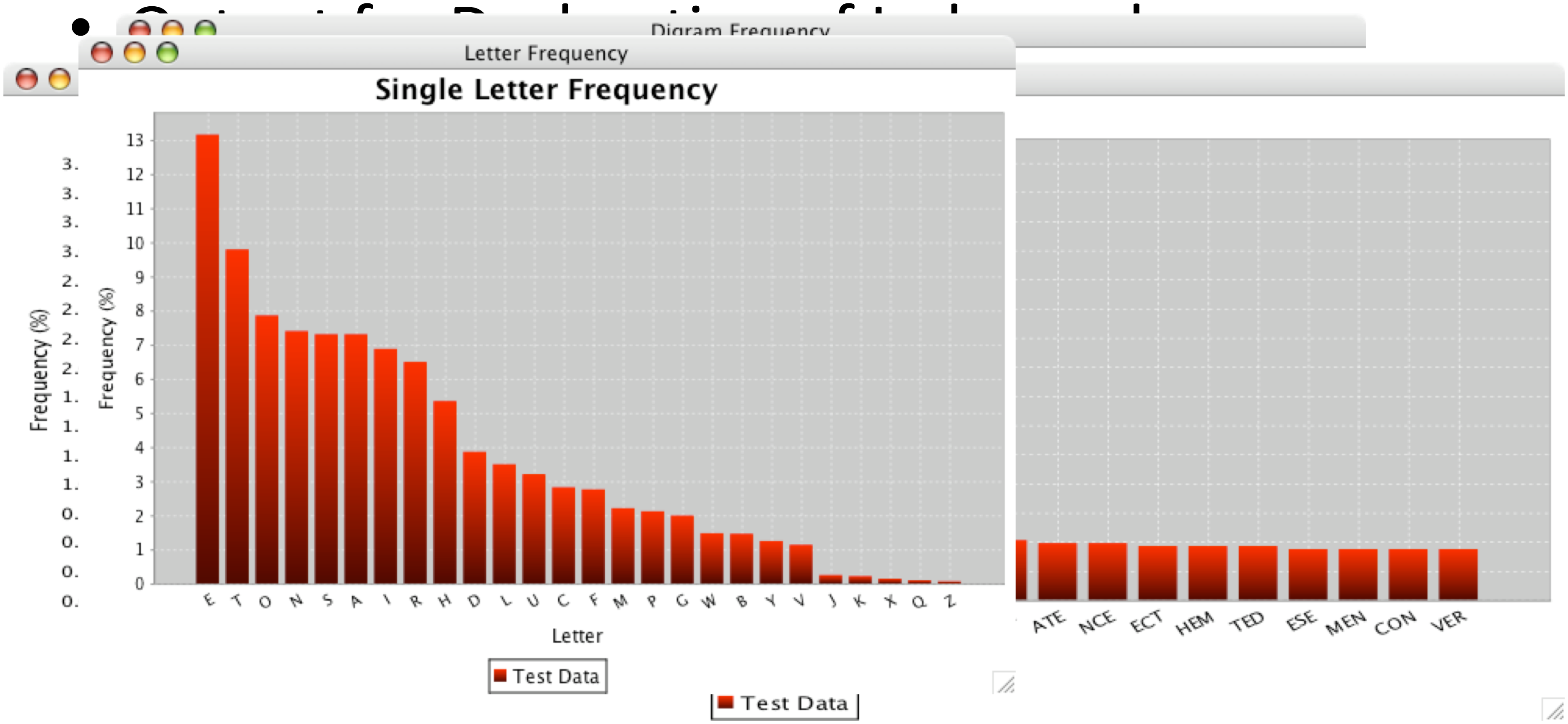
Attacking Mono-Alphabetic Ciphers

- Broken by statistical analysis of letter, word, and phrase frequencies of the language
- Frequency of single letters in English language, taken from a large corpus of text:

A ≈ 8.2%	H ≈ 6.1%	O ≈ 7.5%	V ≈ 1.0%
B ≈ 1.5%	I ≈ 7.0%	P ≈ 1.9%	W ≈ 2.4%
C ≈ 2.8%	J ≈ 0.2%	Q ≈ 0.1%	X ≈ 0.2%
D ≈ 4.3%	K ≈ 0.8%	R ≈ 6.0%	Y ≈ 2.0%
E ≈ 12.7%	L ≈ 4.0%	S ≈ 6.3%	Z ≈ 0.1%
F ≈ 2.2%	M ≈ 2.4%	T ≈ 9.1%	
G ≈ 2.0%	N ≈ 6.7%	U ≈ 2.8%	

(Tip: Counting Letter Frequencies)

- Program **`letter`**, written by TJ O'Connor



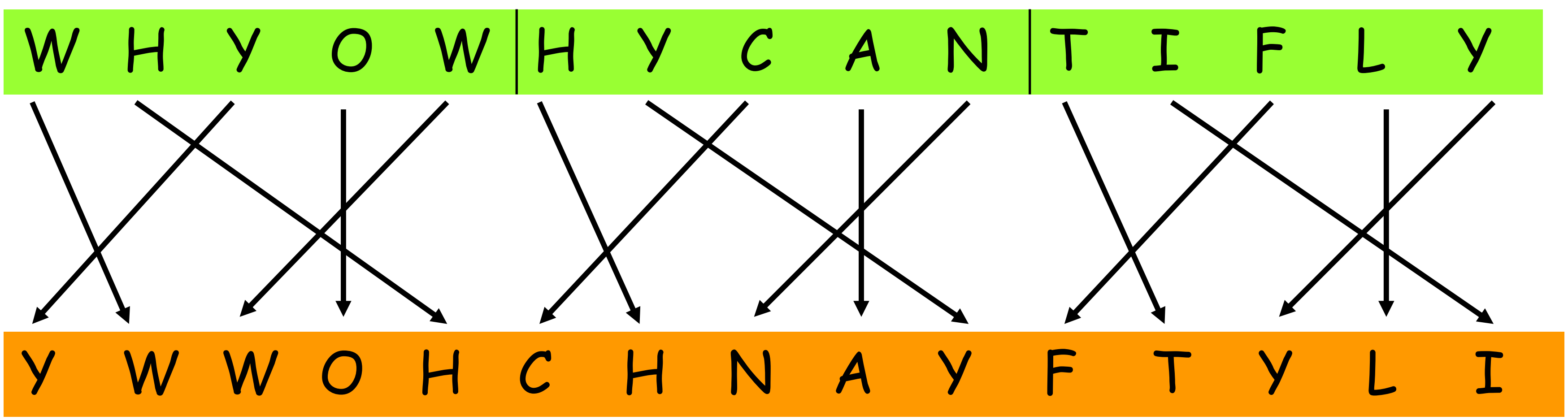
Permutation Ciphers

- The previous codes are all based on substituting one symbol in the alphabet for another symbol in the alphabet
- Permutation cipher: permute (rearrange, transpose) the letters in the message
  - the permutation can be fixed, or can change over the length of the message

Permutation… (Cont'd)

- Permutation cipher ex. #1:
  - Permute each successive block of 5 letters in the message according to position offset <+1,+3,-2,0,-2>

plaintext message



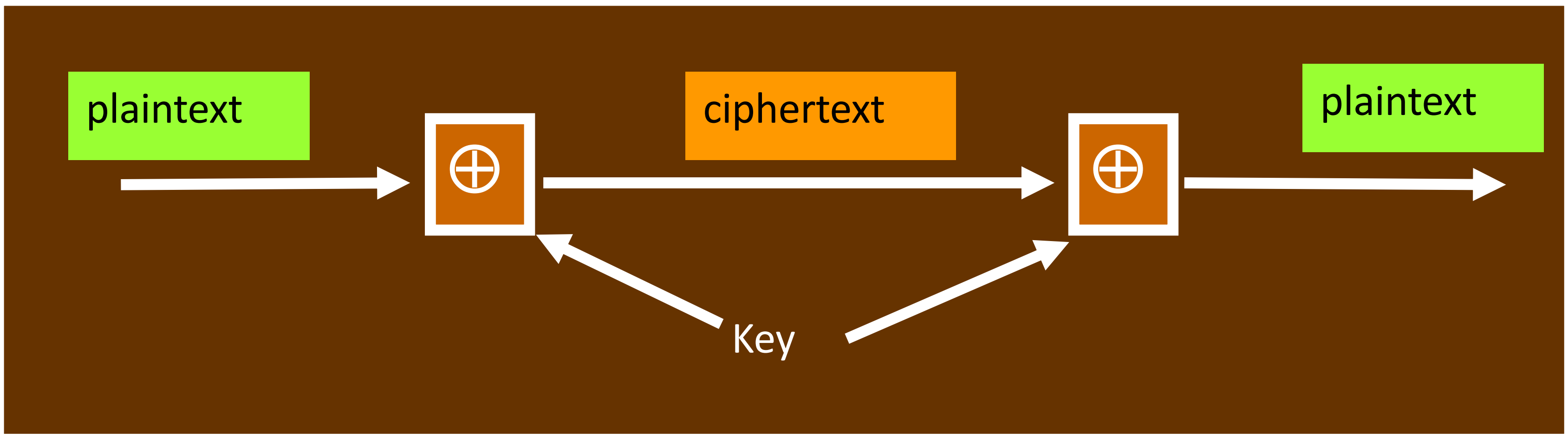
ciphertext message

A Perfectly Secure Cipher: One-Time Pads

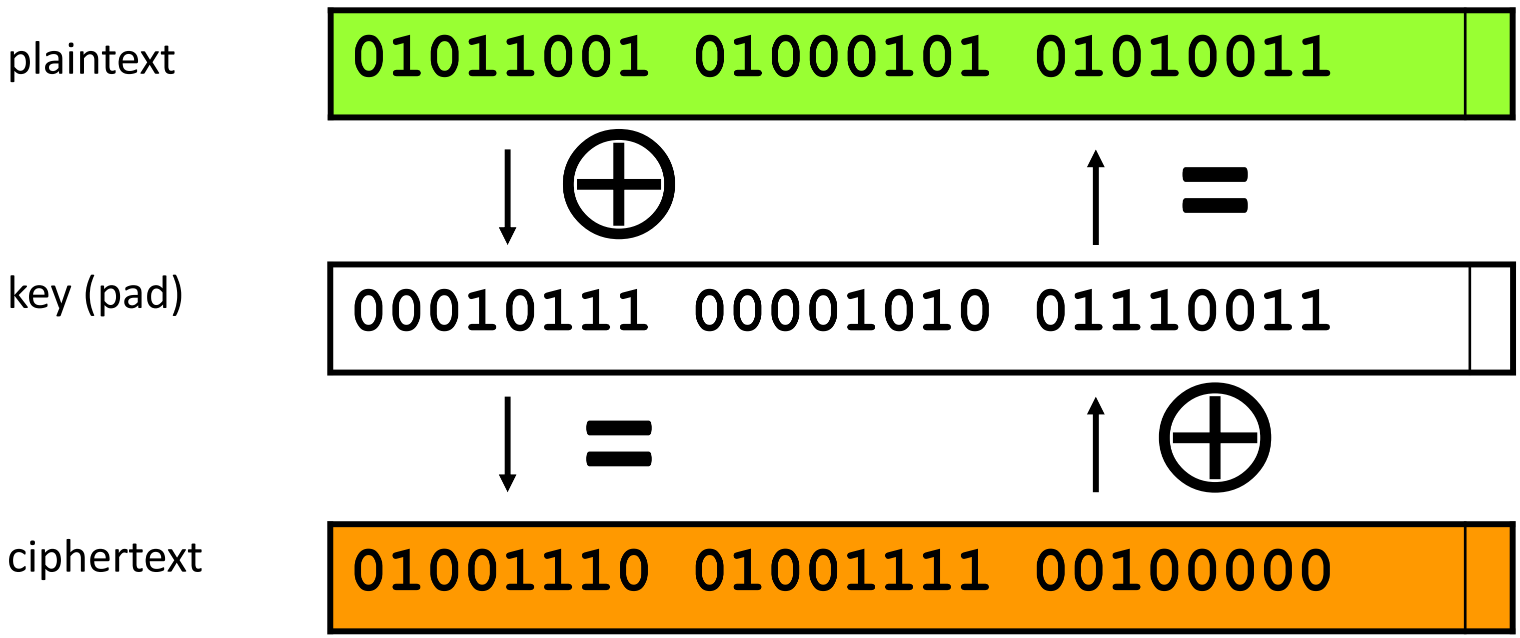
- According to a theorem by Shannon, a perfectly secure cipher requires:
  - a key length at least as long as the message to be encrypted
  - the key can only be used once (i.e., for each message we need a new key)
- Very limited use due to need to negotiate and distribute long, random keys for every message

OTP… (Cont'd)

- Idea
  - generate a random bit string (the key) as long as the plaintext, and share with the other communicating party
  - encryption: XOR this key with plaintext to get ciphertext
  - decrypt: XOR same key with ciphertext to get plaintext



OTP… (Cont’d)



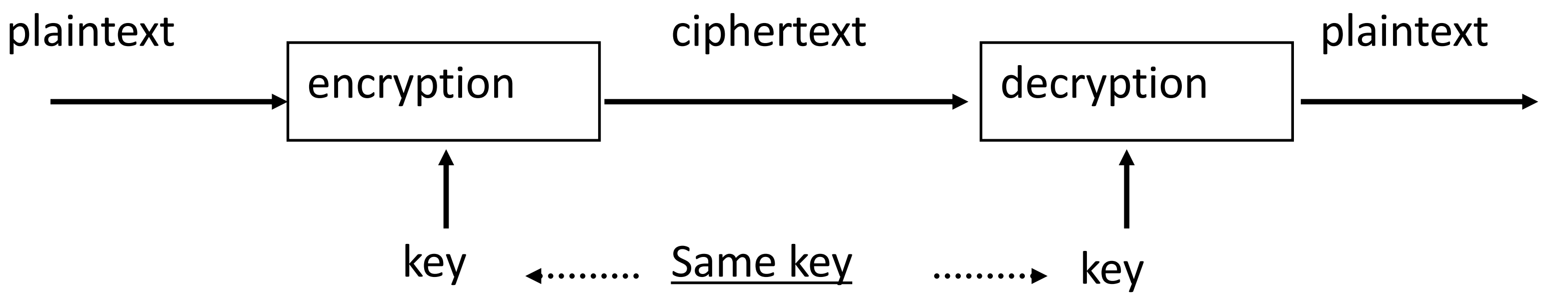
- Why can’t the key be reused?

Some “Key” Issues

Types of Cryptography

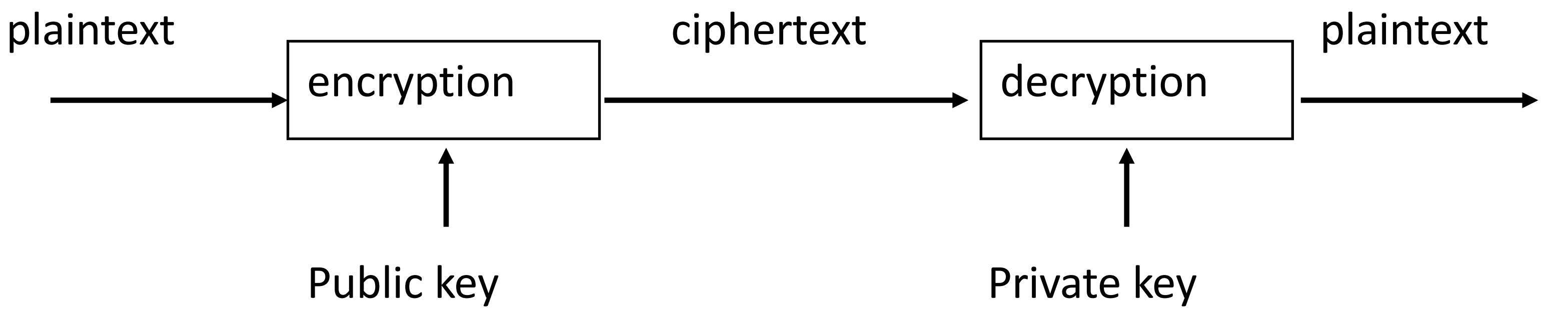
- Number of keys
  - Hash functions: no key
  - Secret key cryptography: one key
  - Public key cryptography: two keys - public, private
- The way in which the plaintext is processed
  - Stream cipher: encrypt input message one symbol at a time
  - Block cipher: divide input message into blocks of symbols, and processes the blocks in sequence
    - May require padding

Secret Key Cryptography



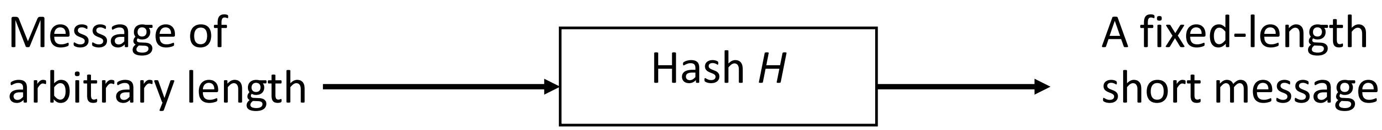
- Same key is used for encryption and decryption
- Also known as
  - Symmetric cryptography
  - Conventional cryptography
- Ciphertext approximately the same length as plaintext
- Examples
  - Stream Cipher: RC4
  - Block Cipher: DES, IDEA, AES

Public Key Cryptography



- Invented/published in 1975
- A public/private key pair is used
  - Public key can be publicly known
  - Private key is kept secret by the owner of the key
- Much slower than secret key cryptography
- Also known as
  - Asymmetric cryptography

Hash Algorithms



- Also known as
  - Message digests
  - One-way transformations
  - One-way functions
  - Hash functions
- Length of $H(m)$ much shorter then length of m
- Usually fixed lengths: 128 or 160 bits

Summary

- Cryptography is a fundamental, and most carefully studied, component of security
  - not usually the "weak link"
- "Perfectly secure" ciphers are possible, but too expensive in practice
- Early ciphers aren't nearly strong enough
- Key distribution and management is a challenge for any cipher