CS408
Cryptography & Internet Security

Lecture 2:

Types of attacks,
Models to evaluate security,
Classical cryptosystems
(shift cipher, substitution cipher)

Reza Curtmola
Department of Computer Science / NJIT

Breaking Ciphers…

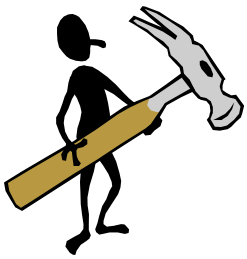
- There are different methods of breaking a cipher, depending on:
  - the type of information available to the attacker
  - the interaction with the cipher machine
  - the computational power available to the attacker

Breaking Ciphers…

- **Ciphertext-only attack**:
  - The cryptanalyst knows **only the ciphertext**.
  - The goal is to find the plaintext and the key.
- NOTE: any encryption scheme vulnerable to this type of attack is considered to be completely insecure.

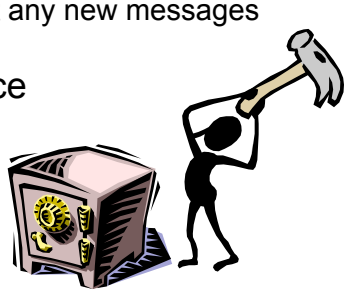
Breaking Ciphers (2)

- **Known-plaintext attack**:
  - The cryptanalyst knows **one or several pairs of ciphertext and the corresponding plaintext.**
  - The goal is to find the key used to encrypt these messages or a way to decrypt any new messages that use that key.
  - How does the cryptanalyst get the pairs of ciphertext and plaintext?



Breaking Ciphers (3)

- **Chosen-plaintext attack**
  - The cryptanalyst **can choose a number of plaintext messages and obtain the ciphertexts for them**
  - The goal is to deduce the key used in the other encrypted messages or decrypt any new messages using that key.
- It can be **adaptive**, the choice of plaintext depends on the ciphertext received from previous requests.



Breaking Ciphers (4)

- **Chosen-ciphertext attack**
  Similar to the chosen-plaintext attack, but the cryptanalyst **can choose a number of ciphertexts and obtain the plaintexts.**
- It can also be **adaptive** The choice of ciphertext may depend on the plaintext received from previous requests.

How Do You Know a Cipher is Secure?

- Show that under the considered attack model, security goals are NOT achieved (break it)
- Show that under the considered attack model, security goals are achieved (evaluate/prove)

Models for Evaluating Security

- **Provable security:**
  - Prove security properties based on assumptions that it is difficult to solve a well-known and supposedly difficult problem (example: computation of discrete logarithms, factoring).
  - Reduce the security of a cryptographic scheme to the difficulty of solving a difficult problem
- **Ad-hoc security:**
  - Variety of convincing arguments that every successful attack requires more resources than the ones available to an attacker.
  - Unforeseen attacks remain a threat.
  - **THIS IS NOT A PROOF**
  - Many symmetric-key ciphers have obscure design principles, and their security is not fully understood
  - Such ciphers are considered secure if they withstand the proof of time!

Provable Security

- **Unconditional (information-theoretic) security**
  - **Assumes that the adversary has unlimited computational resources**.
  - Plaintext and ciphertext modeled by their distribution
  - Analysis is made by using probability theory.
  - For encryption systems: **perfect secrecy**, observation of the ciphertext provides no information to an adversary (besides the length of the plaintext)
  - It is usually expensive and impractical

Provable Security (2)

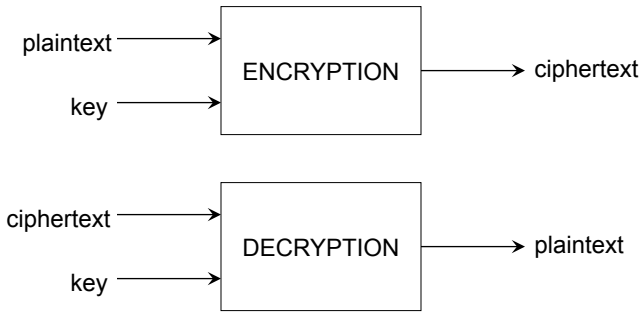
- **Computational security (practical security)**
  - In practice, it is enough to achieve security against adversaries which are computationally bounded
  - Security of a cryptographic scheme is reduced to the difficulty of solving a hard mathematical problem
    - Hard means computationally hard for all adversaries modeled as algorithms that run in polynomial time

Recommended Reading

- Trappe & Washington
  - Chapter 1

Classical cryptosystems

- People have been using secret communication long before the invention of computers
- Encryption scheme (cipher) ensures confidentiality (secrecy) of a message



Symmetric-key cipher

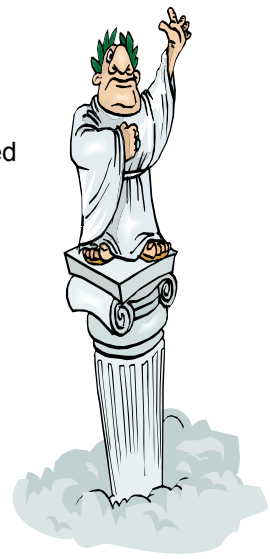
- A symmetric-key encryption scheme is a collection of three algorithms (G, E, D)
  - $\mathcal{K}$ (key space), $\mathcal{P}$ (plaintext space), $\mathcal{C}$ (ciphertext space)
  - Key generation algorithm G generates a key k
  - Encryption algorithm E : $\mathcal{K} \times \mathcal{P} \rightarrow \mathcal{C}$
    we use notation: $c = E_k(p)$
  - Decryption algorithm D : $\mathcal{K} \times \mathcal{C} \rightarrow \mathcal{P}$
    we use notation: $p = D_k(c)$

- The following should always hold true (correctness):
  - $D_k(E_k(p)) = p$, for all k and p

Shift Cipher

- A substitution cipher
- The Key Space:
  - [0 .. 25]
- Encryption given a key K:
  - each letter in the plaintext P is replaced with the K'th letter following the corresponding number (shift right) :
    x = x + K (mod 26)
- Decryption given K:
  - shift left: x = x - K (mod 26)

History: K = 3, Caesar's cipher



Shift Cipher: An Example

A B C D E F G H I J K L M N O P Q R S T U V W X Y Z
0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25

P = CRYPTOGRAPHYISFUN
K = 11
C = NCJAVZRCLASJTDQFY

C → 2; 2+11 mod 26 = 13 → N
R → 17; 17+11 mod 26 = 2 → C
…
N → 13; 13+11 mod 26 = 24 → Y

Shift Cipher: Cryptanalysis

- Can an attacker find K?
  - YES: exhaustive search, key space is small (<= 26 possible keys).
- Once K is found, very easy to decrypt

General Mono-alphabetic Substitution Cipher

- The key space: all permutations of alphabet Σ = {A, B, C, …, Z}
- Encryption given a key π:
  - each letter X in the plaintext P is replaced with $\pi(X)$
- Decryption given a key π:
  - each letter Y in the cipherext P is replaced with $\pi^{-1}(Y)$

Example:

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z
π=	B	A	D	C	Z	H	W	Y	G	O	Q	X	L	V	T	R	N	M	S	K	J	I	P	F	E	U

BECAUSE → AZDBJSZ

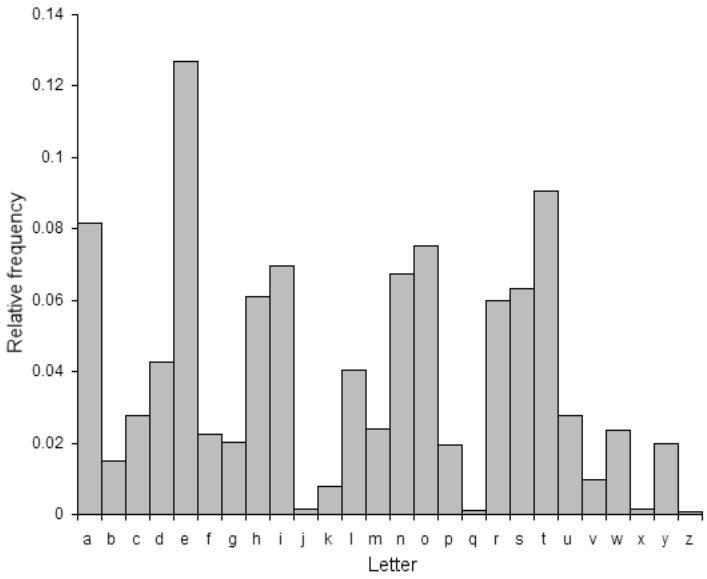
Strength of the General Substitution Cipher

- Exhaustive search is infeasible
  - key space size is 26! ≈ $4*10^{26}$
- Dominates the art of secret writing throughout the first millennium A.D.
- Thought to be unbreakable by many back then

Cryptanalysis of Substitution Ciphers: Frequency Analysis

- Basic ideas:
  - Each language has certain features:
    - Frequency of individual letters, or
    - Frequency of groups of two or more letters
  - Substitution ciphers preserve the language features.
  - Substitution ciphers are vulnerable to frequency analysis attacks.

Frequency of Letters in English



Other Frequency Features of English

- Vowels, which constitute 40% of plaintext, are often separated by consonants.
- Letter A is often found in the beginning of a word or second from last.
- Letter I is often third from the end of a word.
- Letter Q is followed only by U
- And more …

Substitution Ciphers: Cryptanalysis

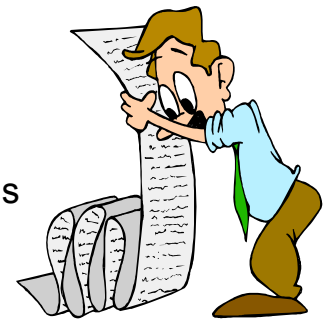
- The number of different ciphertext characters or combinations are counted to determine the frequency of usage.
- The cipher text is examined for patterns, repeated series, and common combinations.
- Replace ciphertext characters with possible plaintext equivalents using known language characteristics.

Frequency Analysis History

- Discovered by the Arabs
  - earliest known description of frequency analysis is in a book by the ninth-century scientist Al-Kindi
- Widely used in Europe during the Renaissance period (14th-17th centuries)
- Frequency analysis made substitution cipher insecure

Summary

• Shift ciphers are easy to break using brute force attacks, they have small key space.
• Substitution ciphers preserve language features and are vulnerable to frequency analysis attacks.



Recommended Reading

- Chapter 2.1, 2.3, 2.4