INTRODUCTION

What is Cryptography?
- Name derived from Greek for (hidden or) secret writing
- The practice and study of techniques for secure communication in the presence of adversaries
- The practical alternative to building secure systems
- Studied in mathematics, computer science, and electrical engineering

Fundamental Concepts
- Confidentiality
  - limitation of access to information to authorized parties
- Integrity
  - prevention of unauthorized alteration of information
- Authentication
  - identity and information origin verification
- Non-repudiation
  - inability to deny previous commitments or actions

Terminology
- plaintext: original unencrypted data
- ciphertext: encrypted (obscured) data
- cipher: (a set of) algorithm(s) for encryption & decryption
  - encryption(plaintext) → ciphertext
  - decryption(ciphertext) → plaintext
- "text" is historical, think bitstream

Evolution of Cryptography
- Basic cryptography principles are well illustrated by the evolution of ciphers and practices
  - transposition ciphers → block ciphers
  - substitution ciphers → one-time pads
  - shift ciphers → algorithms
  - polyalphabetic substitution ciphers → complex algorithms
  - encryption devices, e.g., Scytale, Enigma → computers
**Symmetric Encryption**

- Encryption based on letter reordering
- Cipher text is a permutation of the plaintext
- Transposition patterns can be parametrized
- Often combined with other ciphers
- Foundation for modern (bit-level) block ciphers

**Steganography**

- Secret writing / Invisible ink (lemon juice)
- Text revealed under heat or UV light
- Reels typically operated on text data
- Strength of ciphers relied on secrecy, (lack of) knowledge of mechanisms, and the complexity of cryptanalysis

---

**Route Cipher**

- Basic transposition cipher based on table format
- The quick brown fox jumps over the lazy dog
- Eliminate whitespaces, map plaintext, transpose table

```
THEQU TIOKSTZ
ICKXRA HCMOFHY
OINFOD BRENFEOD
KXJUMP QBFMELO
SOVER UROPRAG
THEIA ZVDOG
```

- Read out message

```
TIOKSTZCHXELYXKNOVEDQBFMELOUROPRAG
```

- Width of table is key

---

**Spartan Scytale**

- Transposition cipher (first mentioned ca 600 BC)
- Plaintext written on strip wrapped around a stick
- Identical sticks exchanged in advance
- Strip (message) sent (without stick)
- Device (stick) is key
- Dimension(s) of stick determines transposition parameter(s)

**Vigenère Cipher**

- Encryption based on a (character) translation table
- Construct a cipher alphabet for translation
  - plain alphabet: ABCDEFGHIJKLMNOPQRSTUVWXYZ
  - cipher alphabet: QWERTYUIOPASDFGHJKLZXCVBNM
- Encrypt (translate) data letter by letter

```
THE QUICK BROWN FOX JUMPS OVER THE LAZY DOG
```

- Cipher alphabet is key
- Foundation for modern (bit-level) one-time pad systems

---

**Historical Perspective**

- Secret communication mechanisms >4000 years
- Secret writing / Invisible ink (lemon juice)
  - Text revealed under heat or UV light
  - Relies on secrecy (knowledge of existence)
- Cryptography >2500 years
- Early ciphers typically operated on text data
- Strength of ciphers relied on secrecy, (lack of) knowledge of mechanisms, and the complexity of cryptanalysis

**Introduction**

- Foundation for modern (bit-level) one-time pad systems
- Strength of ciphers relied on secrecy, (lack of) knowledge of mechanisms, and the complexity of cryptanalysis

---

**Classic Ciphers**

- Spartan Scytale
- Route Cipher
- Substitution Ciphers
- Enigma
- Vigenère Cipher
- Breaking Ciphers
Modern Perspective

**Caesar Cipher**

- Shift (a type of substitution) cipher from ca 50 BC
- Key derived from (rotational) shift of alphabet
  - plain alphabet: ABCDEFGHIJKLMNOPQRSTUVWXYZ
  - cipher alphabet: CDEFGHIJKLMNOPQRSTUVWXYZ
- Encryption translates letter by letter
  - THE QUICK BROWN FOX JUMPS OVER THE LAZY DOG
  - QEB NRFZH YOLTK CLU GRJMP LSBO QEB IXWV ALD
- Algorithm parameter (shift index) is key
  - Caesar +3 = shift 3 steps to the right

**Parametrized Substitution Ciphers**

- Select keyphrase
  - "I AM SO CLEVER"
- Eliminate whitespaces and duplicates
  - "IAMSOCLEVR"
- Pad to get a full cipher alphabet
  - "IAMSOCLEVRXUTQPNKJHGFDB"
- Construct translation table
  - plain alphabet: ABCDEFGHIJKLMNOPQRSTUVWXYZ
  - cipher alphabet: IAMSOCLEVRXUTQPNKJHGFDB
- Encrypt letter by letter
  - THE QUICK BROWN FOX JUMPS OVER THE LAZY DOG
  - QEB NRFZH YOLTK CLU GRJMP LSBO QEB IXWV ALD
- Algorithm parameter (keyphrase) is key

**Breaking Classic Ciphers**

- Direct letter-level translations preserve basic statistical properties of the original (plaintext) message
- Statistical properties (e.g., word lengths and letter frequencies in the English language) provide hints that can be used in deciphering messages
- Analysis compares statistical properties of ciphertext to known properties of the language of the plaintext
  - provides direct clues for translation
  - provides heuristics for brute force attacks

**Breaking Ciphers**

- Vulnerabilities
  - exploit knowledge of structure or content of plaintext (or ciphertext)
  - e.g., word length, frequency analysis
- Exploits
  - attack (potentially forced) misuse
- Exhaustive parameter space search (brute force)
  - with known plaintext-ciphertext (e.g., find cipher parameters)
  - without known plaintext-ciphertext (e.g., search for syntax patterns or known values)

**English Letter Frequencies**

- Graph showing letter frequencies in the English language.
Symmetric Encryption

Lessons to Learn

Steganography

Codes

Enigma

Vigenère Cipher

Polyalphabetic substitution cipher that aims to disguise letter frequency of plaintext messages

- Uses a (repeated) key to switch alphabets in encryption (each letter encrypted using a different alphabet)
- Shift-based translation table (Vigenère square)
- Cipher index constructed from (repeated) keyphrase
- Plaintext and keyphrase letters used as indices in table
- Encryption: plaintext horizontal, keyphrase vertical
- Decryption: keyphrase vertical, find ciphertext character
- Encrypt letter by letter using indexed translation table

- THE QUICK BROWN FOX JUMPS OVER THE LAZY DOG
- BHQ TINNOUP YFEP EFLQKW FDED LVG WEUC UWG

- The most frequently occurring letter is 'E'
- The most frequently occurring 3-letter word is 'the'

- The Enigma machine
- Lampboard - output
- Keyboard - input
- Plugboard - remaps letters prior to input
- Reflector - remaps character and returns the signal
- A key is pressed, signal is mapped on the plugboard, passes through rotors, returned from the reflector, back through the rotors, back through plug board, and a key is lit on the lampboard
- Rotors turn to a new position on each key press
- Each character is remapped at least 7 times (minimum 3 rotors + reflector)
Enigma Message Analysis

- Hypothesize plaintext based on knowledge of patterns
  - military jargon and operator communication habits
    (improper Enigma use provided clues to configurations)
- Manual investigation deduced plugboard connections
  - the Enigma would never translate a letter to itself
  - any configuration that produced a letter self-mapping could be eliminated
- Candidate parameter sets evaluated by Bombes
- Decryption setting results were fed into a replica Enigma to see if they produced German language

Bombe

- WWII Enigma decryption machines initially developed by the Polish, adapted and developed by the allies
- Enigma message decryption requires Enigma parameters to be determined (settings changed daily)
  - selection of rotors: 60 (3 of 5) - 336 (3 of 8) permutations
  - plugboard connections: ca. $1.5 \times 10^{9}$ permutations
  - initial rotor configuration: ca. $1.7 \times 10^{8}$ permutations
- Each Bombe emulated 36 Enigmas for parallel investigation of rotor settings
  - reduced the rotor order and position search space to a manageable size for further (manual) analysis

Code Examples (Code Talkers)

- In WWII native languages were used in combination with codes (Apache, Comanche, Navajo, Basque, Welsh)
- Navajo - no written form, complex syntax and tonal forms
- Letters and everyday words with replaced semantics
  - A = WOL-LA-CHEE (ant), BE-LA-SANA (apple), TSE-NILL (axe)
  - January = ATSAM-BE-YAZ (small eagle)
  - tank = CHAY-DA-GAHI (tortoise)
- Code never broke, even confused a native speaker not trained in the code
- Native dialects differ from non-native dialects
  (robust form of impersonation detection)
Modern Perspective

Codes vs Cryptography

- Codes are expensive to develop and maintain
- Ciphers are easier to operate
- Ciphers scale to greater scale
- Combinations can be used to improve security

Lessons to Learn

- Even the most complex ciphers have lifetimes
- Strength of ciphers rely on
  - the integrity of the algorithm
  - the complexity of the algorithm
  - the strength (size) of the key
- Strength of security system rely on
  - cipher systems
  - correct application of ciphers
  - protocols
  - user awareness, understanding and training

Kerckhoff's Design Principles

- A set of design principles for ciphers formulated in 1883
  1. The system must be practically, if not mathematically, indecipherable;
  2. It should not require secrecy, and it should not be a problem if it falls into enemy hands;
  3. It must be possible to communicate and remember the key without using written notes, and correspondents must be able to change or modify it at will;
  4. It must be applicable to telegraph communications;
  5. It must be portable, and should not require several persons to handle or operate;
  6. Lastly, given the circumstances in which it is to be used, the system must be easy to use and should not be stressful to use or require its users to know and comply with a long list of rules.

Steganography

- Hiding messages inside other messages
- Historical versions: invisible ink, microdots
- Modern versions: digital watermarking
  - amplitude modulation, e.g., alter (least significant bit) pixels values in images
  - spread-spectrum, e.g., add (noise) values in the frequency domain representation of signals
  - quantization, e.g., modify bit representations of signals
- Often used in conjunction with cryptography techniques
  - encoded messages hidden in plain sight

Cryptographic Design

- Kerckhoff's principle
  - "A cryptosystem should be secure even if everything about the system, except the key, is public knowledge"
  - "the strength of a crypto-system must not depend on the secrecy of the cipher but only on the secrecy of the key"
- Shannon's maxim
  - "the enemy knows the system, i.e., one ought to design systems under the assumption that the enemy will immediately gain full familiarity with them"
- If attackers cannot break an algorithm with detailed knowledge of the algorithm mechanics, they can certainly not do so without that knowledge
- Err on the side of caution / assume worst case scenario

Modern Perspective

- Computers facilitate new types of security systems
  - more advanced ciphers
  - more advanced security attacks
- Mathematics of cryptography (both ciphers and attacks) are now both better understood and more well spread
- …and so are brute force attack tools and resources
Symmetric Encryption

- Commonly referred to as private key encryption
- Data encrypted and decrypted using the same key
- Anyone with access to the key can decrypt the data
- Comparatively fast
- Suffers from the key distribution problem

Block Cipher

- Deterministic (fixed complexity) algorithm that operates on a fixed length group of bits (a block) to encrypt data using a symmetric key
- Typically implemented using iterated product ciphers
  - repeated combined application of simple operations, e.g., substitutions and permutations, that obscure data
- Common design: substitution-permutation networks
  - graphs of key diffusion (rotation and derivation) schemes, (Substitution) boxes, and (Permutation) boxes
- Aims for high avalanche effect
  - small changes in input bits yield large changes in output

Substitution-Permutation Network

- S(ubstitution)-boxes, and P(ermutation) boxes: e.g., substitutions and permutations, that obscure data
- n-bit substitution (S-box) for each bit of the input
- Key: key components and round keys
- Security: non-repeating, repeated combined application of simple operations, graphs of key diffusion (rotation and derivation) schemes, (Substitution) boxes, and (Permutation) boxes

Block Cipher Modes of Operation

- Electronic Code Book (ECB)
  - blocks encrypted independently
- Cipher-Block Chaining (CBC)
  - blocks XORed to the previous encrypted block
  - Initialization Vector (IV) used for the first block
- Output Feedback (OFB)
  - repeated encryption of the IV yields keystream
- Cipher Feedback (CFB)
  - self-synchronizing stream ciphers
  - one of the output bits is added to the next message bit
- Counter Encryption (CTR)
  - key \( k_{n} = (IV + n) \mod \text{key length} \)
  - similar to OFB
  - less costly and less sensitive to birthday attacks
**Electronic Codebook (ECB)**

- **Encryption**
  - Key: Plaintext
  - Ciphertext: Ciphertext
- **Decryption**
  - Key: Ciphertext
  - Plaintext: Plaintext

**Cipher-Block Chaining (CBC)**

- **Encryption**
  - Key: Plaintext
  - Ciphertext: Ciphertext
- **Decryption**
  - Key: Ciphertext
  - Plaintext: Plaintext

**Message Authentication Code (MAC)**

- Use a block cipher with Cipher-Block Chaining
- Discard all blocks but the last
- Provides integrity and authenticity

**Stream Cipher**

- A symmetric encryption algorithm that obscures data by combining plaintext digits with those of a keystream
- Keystreams (streams of pseudorandom values) derived from key during encryption
- Typically implemented using bit-level XOR
- Very fast
- Low hardware requirements (compared to block ciphers)
- Limited security
  - Keystream must have long periods
  - Keystreams (or initialization vectors) must not be reused
- Keystreams can be generated from block ciphers (using e.g., OFB or CTR modes)
Modern Perspective

Symmetric Encryption

- Advanced Encryption Standard (AES), Rijndael
- Data Encryption Standard (DES), Lucifer (obsolete)
- Triple DES (3DES), 3x DES with CBC
- International Data Encryption Algorithm (IDEA)
- Blowfish
- RC4

One-Way Encryption

- Data encrypted using secret keys
- Data can not be decrypted
- Cipher texts (to a high probability) uniquely mapped to data content
- Cipher texts used instead of data in situations where data must be kept secret (e.g., passwords)
- Closely related to hashcodes and Message Authentication Codes (MACs)

Challenge-Response

- Basic assumption: shared secret $S$
- Client authentication
  - server generates and sends a random challenge $C$
  - client computes and sends hash($C$ & $S$)
  - server computes the same hash to verify client secret
- Mutual client-server authentication
  - server generates and sends a random challenge $C_s$
  - client generates a random challenge $C_c$ and sends $C_c$ & hash($C_s$ & secret)
  - server verifies client hash and sends hash($C_c$ & secret)
  - client verifies server hash
- To increase security, add
  - key derivation functions
  - time stamps
- Session key exchange
  - authenticate and exchange encrypted session key