SE 4C03 Winter 2001

07. Overview of Cryptography

Instructor: W. M. Farmer

Revised: 25 February 2001

What is Cryptography?

- **Cryptography** is a collection of mathematical techniques for:
  - Protecting data privacy
  - Protecting data integrity
  - Verifying the identity of objects
  - Verifying the identity of subjects
  - Producing random objects
- Principal techniques:
  - Cryptographic hashing
  - Conventional encryption
  - Public key encryption
  - One-way encryption
  - Random number generation

Hashing

- Given an object as input, a **hash function** returns an identification code (called a **hash code**) for the object
- A hash function has the following properties:
  - The output has a fixed size, much smaller than the size of the input
  - The function is many-to-one (so **collisions** are possible)
  - The function is deterministic and easy to compute
- Hash functions are used to:
  - Build rapidly accessible data storage structures called **hash tables**
  - Produce **checksums** for checking data integrity

Cryptographic Hashing

- A **cryptographic hash function** is a hash function whose purpose is to produce a "fingerprint" (called a **message digest**, **cryptographic hash code**, or **cryptographic checksum**) of an input object
- A cryptographic hash function h has the following properties:
  - Given a hash code c, it is mathematically infeasible to find an object x such that $f(x) = c$ (**one-way property**)
  - Given an object x, it is mathematically infeasible to find another object y such that $f(x) = f(y)$ (**weak collision property**)
  - It is mathematically infeasible to find two objects x and y such that $f(x) = f(y)$ (**strong collision property**)

Conventional Encryption

- A single **key** is required that is kept secret
- **Encryption**: plaintext, key $\xrightarrow{f}$ ciphertext
- **Decryption**: ciphertext, key $\xrightarrow{f^{-1}}$ plaintext
- f and f^{-1} are the encryption and decryption algorithms, respectively
- **Key assumption**: Computation of the plaintext from the ciphertext is mathematically infeasible without the key
- In practice, the security of the process depends primarily on maintaining the secrecy of the secret key

Ciphers

- A **cipher** is a encryption/decryption method
- Monoalphabetic ciphers (letter-for-letter substitution)
  - Caesar (rotation) ciphers (25 possible keys)
  - Shuffle ciphers (26! possible keys)
- Cipher techniques
  - Substitution
  - Transposition
  - Stream translation
  - Block translation

Cryptanalysis

- **Cryptanalysis** is the process of discovering how to decrypt ciphertext without the secret key
- Approaches:
  - Brute force: try all possible keys
  - Exploit known plaintext
  - Exploit chosen plaintext
  - Analyze encryption and decryption algorithms
- Criteria for measuring the effectiveness of a cipher:
  - Cost of breaking the cipher vs.
    Value of the encrypted information
  - Time required to break the cipher vs.
    The useful lifetime of the encrypted information

Data Encryption Standard (DES)

- Most widely used conventional encryption algorithm
- Developed by IBM in the late 1960s
- Adopted by the USA National Institute of Standards and Technology (NIST) in 1977
- Process
  - Encryption is performed in 64-bit blocks
  - Change of single input bit changes almost all output bits
  - Key is 56 bits long (as requested by USA NSA)
- Security concerns:
  - Key length (brute force attacks work)
  - Internal algorithm structure (details of which are classified)

International Data Encryption Algorithm (IDEA)

- Developed by Xuejia Lai and James Massey of Swiss Federal Institute of Technology and published in 1990
- 128-bit key is used to encrypt data in 64-bit blocks
- Included in the Pretty Good Privacy (PGP) package

Blowfish

- Developed by Bruce Schneier around 1993
- Fast, compact, easy to implement
- Variable-length key (up to 448 bits long) is used to encrypt 64-bit blocks
  - Higher speed and higher security can be traded off
- Considered to be an extremely strong algorithm

Secret-Key Distribution

- Problem: Often too many secret keys are needed to deliver them all physically
- Key distribution scheme
  - A **key distribution center (KDC)** holds a unique **master key** for each end system
  - Communication between end systems is encrypted using a temporary key called a **session key**
    - One end system A requests a session key from KDC to communicate with another end system B
    - The KDC sends A back a message encrypted with A's master key containing the session key and a message for B encrypted with B's master key
    - The latter message, which contains the session key and A's identity, is sent to B by A
- The whole system fails if the KDC is compromised

Application: Link Encryption

- Data transmitted on a communication link is encrypted
- Every pair of routers that share a link need to share a unique secret key
- The entire data area of a frame is encrypted
- The data area of the frame must be decrypted when it arrives at a router
  - The message is exposed to intermediate routers

Application: End-to-End Encryption

- Data is encrypted by the sender and decrypted by the receiver
- Only the data part of a packet is encrypted
- Can be performed at different TCP/IP layers
  1. Application layer (e.g., telnet, e-mail)
     – Only parts of the TCP/UDP data area are encrypted
     – IP, TCP, and UDP software need not be modified
  2. Transport layer (TCP, UDP)
     – Entire TCP/UDP data area is encrypted
     – TCP/UDP layer software must be modified
  3. Internet layer (IP)
     – Entire IP data area is encrypted
     – IP layer software must be modified

Application: IP Tunneled Through IP

- Encrypted IP datagram is encapsulated in another IP datagram
- The Internet is treated as an SPN
- Used to create Virtual Private Networks (VPNs)

Public Key Encryption

- Discovery
  – Discovered but held secret by USA NSA and UK Communications-Electronic Security Group in mid to late 1960s
  – Discovered and publicized by Whitfield Diffie and Martin Hellman at Stanford University in 1976
- Motivation
  – Difficulty of secret-key distribution: secrecy must be shared
  – Need for digital signatures that can be verified by arbitrary parties

Public Key Encryption: Basic Process

- Each end system has two keys:
  – **Private key** that is kept secret
  – **Public key** that is made public
- **Encryption**: plaintext, public key $\xrightarrow{f}$ ciphertext
- **Decryption**: ciphertext, private key $\xrightarrow{f}$ plaintext
- **Signature writing**: plaintext, private key $\xrightarrow{f}$ ciphertext
- **Signature reading**: ciphertext, public key $\xrightarrow{f}$ plaintext
- The same algorithm is used for both encryption and decryption
- It is mathematically infeasible to derive the private key from the public key

Public Key Encryption Applications

1. Privacy
   - The sender encrypts the plaintext message with the receiver's public key
   - The receiver decrypts the ciphertext message with its private key
2. Integrity, digital signature, and nonrepudiation
   - The sender encrypts the plaintext digest of the message with its private key
   - The receiver decrypts the ciphertext digest of the message with sender's public key

Public Key Encryption Applications (cont.)

3. Privacy and integrity
   - The sender encrypts the plaintext message with its private key
   - The sender encrypts the ciphertext message with the receiver's public key
   - The receiver decrypts the ciphertext message with its private key
   - The receiver decrypts the ciphertext message with the sender's public key
4. Secret-key exchange

Diffie-Hellman Key Exchange Algorithm

- Appeared in original 1976 Diffie-Hellman paper
- Used only for secret key exchange

RSA Algorithm

- Developed by Ron Rivest, Adi Shamir, and Len Adleman at MIT in 1977
- Supports privacy, digital signature, and secret-key exchange
- Most widely used public key algorithm
- The keys are generated from two large prime numbers p and q
  - p and q are private
  - The product of p and q is public

Public Key Management

- Problem: How are public key forgeries prevented?
- Public key distribution
  - Public announcement
  - Public key directory
  - Public key authority secured using cryptographic measures
  - Public key certificates provided by a certificate authority

Conventional vs. Public Key Encryption

- Conventional encryption is more efficient than public key encryption
- Public key encryption is very versatile
- Public/private key pairs are easily changed or revoked