Cryptography

Fall 2024

R. Sekar

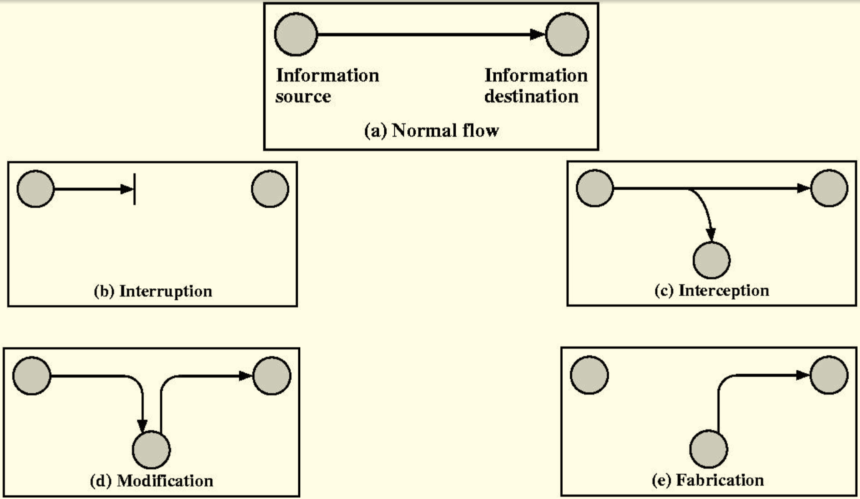
How to achieve security

- Basis is separation
  - Separate adversarial entities
- How to separate adversaries?
  - Physical separation
  - Temporal separation
  - Cryptographic separation
  - Logical separation
- Security vs Functionality
  - Controlled sharing

Communication Vs System Security

- Communication Security
  - security of data channel
  - typical assumption: adversary has access to the physical link over which data is transmitted
  - cryptographic separation is necessary
- System security
  - security at the end points
  - information cannot be encrypted, as it needs to be accessed by applications on the end system
  - logical separation is typically the basis

Communication Security Concerns



Cryptography

- Encode the data in a manner that makes it accessible only to authorized parties
  - Encryption algorithm
  - Encryption key
- Why it is not a good idea to rely on secrecy of algorithm
  - Hard to develop good encryption algorithm
  - Does not scale beyond a few users
  - *Security by obscurity*
- Key point: need to preserve secrecy of key

Terminology

- Plaintext (“unencrypted”)
- Ciphertext (“encrypted”)
- Encryption ($E_k(X)$)
  - Result of encrypting message X using encryption key k
- Decryption ($D_k(X)$)
- Cryptanalysis: Discover k, X or both

Types of Attacks in Cryptography

Type of Attack	Known to Cryptanalyst
Cipher text only	• Encryption algorithm • Cipher text to be decoded
Known plain text	• Encryption algorithm • Cipher text to be decoded • One or more plain text-cipher text pairs formed with the secret key
Chosen plain text	• Encryption algorithm • Cipher text to be decoded • Plain text message chosen by cryptanalyst, together with its corresponding cipher text generated with the secret key
Chosen cipher text	• Encryption algorithm • Cipher text to be decoded • The purported cipher text chosen by cryptanalyst, together with its corresponding decrypted plain text generated with the secret key
Chosen text	• Encryption algorithm • Cipher text to be decoded • Plain text message chosen by cryptanalyst, together with its corresponding cipher text generated with the secret key • The purported cipher text chosen by cryptanalyst, together with its corresponding decrypted plain text generated with the secret key

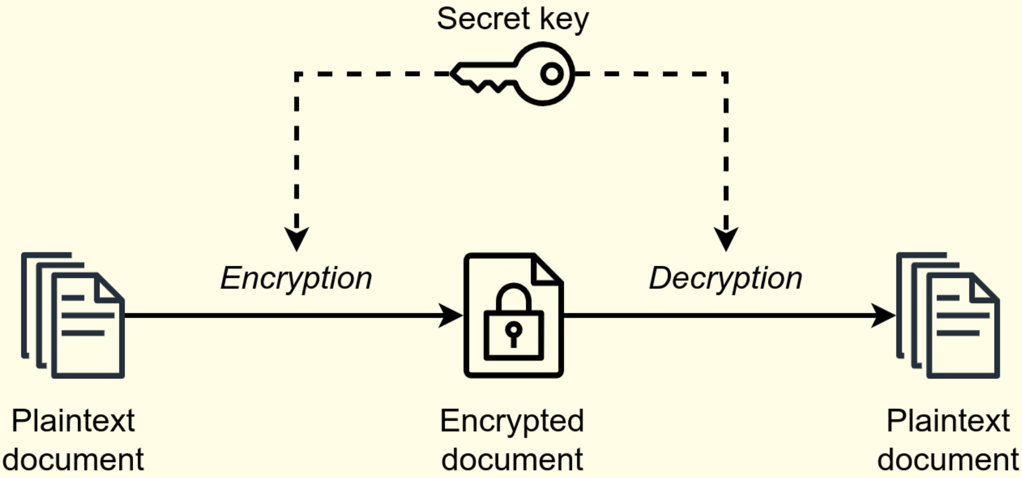
[1]

[1] Source: Applied Cryptography Using Chaos Function for Fast Digital Logic-Based Systems in Ubiquitous Computing — Scientific Figure on ResearchGate.

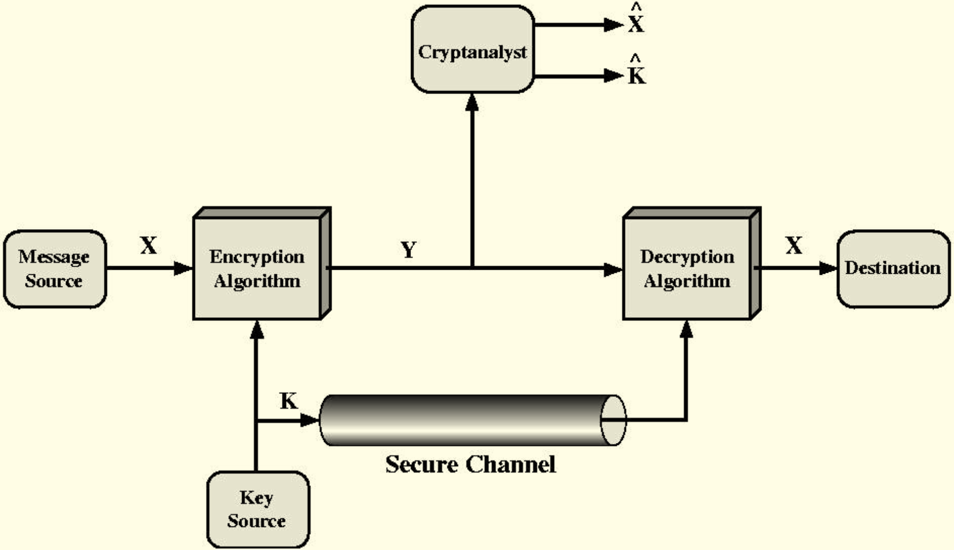
Steganography

- Hiding presence of information
- Use normal-looking messages/pictures that conceal secret data
- Useful if communication is monitored for “suspicious content” by someone
- Also used for copyright protection
  - Watermark: invisible data encoded in messages that is retained in copies, and is robust in the face of typical image transformation operations

Symmetric Cryptography



Model of Symmetric Cryptography



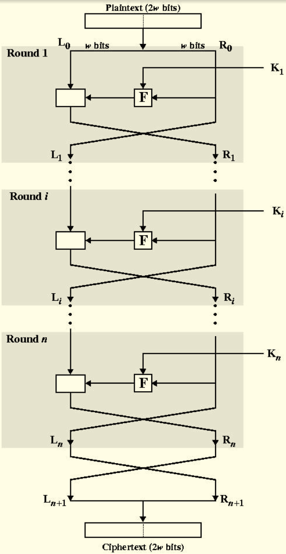
Stream and Block Ciphers

- Stream cipher: used to encrypt digital streams of data, one bit or a byte at a time
  - Provides most flexibility for cryptographic applications
- Block cipher: Operates on data blocks, typicall 128 bits or more
  - Small block sizes are vulnerable to statistical attacks
- Stream ciphers can be constructed from block ciphers
  - But careless applications of a block cipher (e.g., code book) can easily be broken.
  - Use recommended construction, e.g., feedback modes

Structure of Block Ciphers

- Needs to produce a reversible mapping that maps *n*-bit blocks to other n-bit blocks
- Good ciphers are based on Shannon's concepts of "diffusion" and "confusion"
  - Diffusion: disperse bit-patterns within each block of data
  - Confusion: "mix-up" the order of bits within a block. In practice, use permutations specified by a key.
- In principle, good ciphers can be implemented using a table of mappings
  - Encryption key selects which mapping to use
  - Approach impractical for all except smallest block sizes
- Feistel structure: a way to build more complex ciphers from simpler ones

Symmetric Cryptography



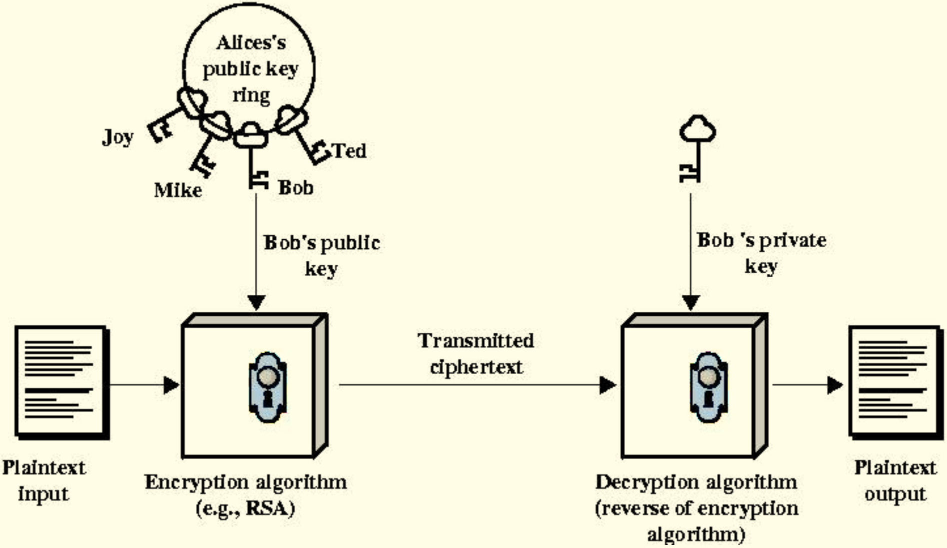
Symmetric Cryptography Algorithms

- DES
  - Not considered very secure (key length of 56 bits)
- Triple DES with two keys (128 bits)
- AES (128 bits)
- IDEA (128 bits)
- Blowfish (up to 448 bits)
- RC5 (up to 2040 bits)
- CAST-128 (40 to 128 bits)
- RC2 (8 to 1024 bits)

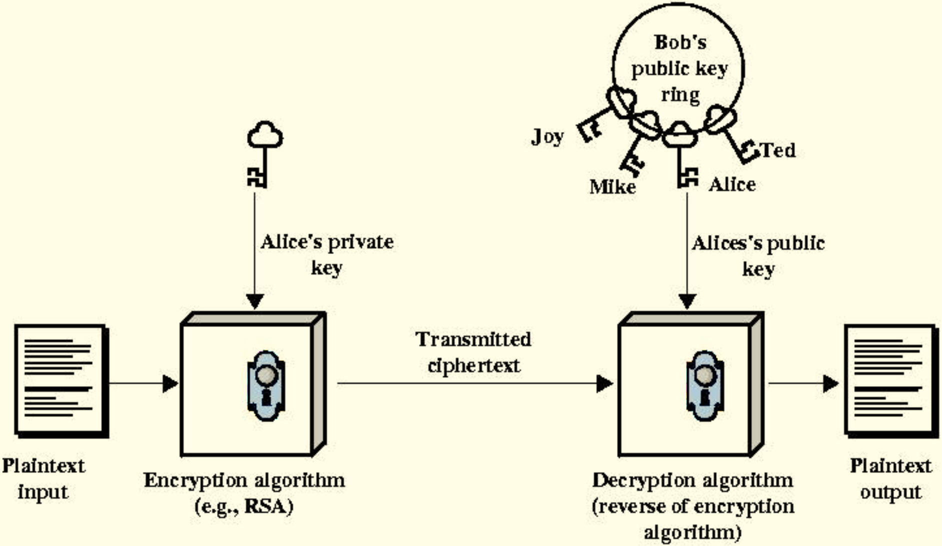
Public Key (Asymmetric) Cryptography

- Uses one key for encryption and another one for decryption
  - Requires that it be computationally infeasible to compute one of the keys based on the other
- One of the two keys is private to a principal; the other key can be freely distributed to any one
  - Each principal generates his/her own pair of public/private keys, and the private key need not be revealed to any one.
- Some public key algorithms (e.g., RSA) permit both keys to be used for encryption and decryption
  - What is encrypted with one key can be decrypted with the other

Encryption in Public Key Crypto



Authentication in Public Key Crypto



Encryption Vs Signing

- When the encoding operation is performed using someone's public key, the results are accessible only to that person
  - This operation can be used to ensure confidentiality of data — hence called "encryption"
- When the encoding operation is done using someone's private key, the results are accessible to every one.
  - But one can be sure that the message came only from the person whose public key is used for decoding — hence called "signing"

RSA Algorithm

- Alphabet = $\{0, \ldots, n-1\}$
- Encryption: $C = M^e \; mod \; n$
- Decryption: $M = C^d \; mod \; n = M^{ed} \; mod \; n$
- Need: $M^{ed} \equiv M \; (mod \; n)$
- Both sender and receiver know n.
- Sender knows e, while only the receiver knows d.
- $KU = (e, n)$, $KR = (d, n)$

RSA Algorithm Requirements

- It is possible to find d, e and n such that $\forall M \;\; M^{ed} \equiv M \; (mod \; n)$
- M^e and C^d can be efficiently computed
- It is infeasible to determine d from e

RSA Key generation

- Select two large prime numbers p and q
- Calculate $n = p \times q$
- Calculate $\phi(n)$. For the n we have chosen, it will be $(p-1)(q-1)$
- Select an $e < \phi(n)$ that is relatively prime to $\phi(n)$
- Calculate $d = e^{-1} \ (mod \ \phi(n))$
- Set $KU = (e, n)$, $KR = (d, n)$

Miller-Rabin Test for Primality

- Pick an odd number n. Note $n-1 = 2^k q$, where q is odd.
- Pick a number $1 < a < n-1$, compute $a^q, a^{2q}, \ldots, a^{2^k q}$.
- If n is prime, by Fermat's theorem:

$$a^{2^k q} = a^{n-1} = 1 \ \ (mod\ n)$$

Hence, for some $0 \leq j \leq k$, $a^{2^j q} \mod n = 1$

  - **Case 1:** $j = 0$: This means $a^q \mod n = 1$
  - **Case 2:** for some $j > 0$, $a^{2^{(j-1)} q} \mod n \neq 1$ but $a^{2^j q} \mod n = 1$:
    i.e., $(a^{2^{(j-1)} q} - 1)(a^{2^{(j-1)} q} + 1) \mod n = 0$
    Since the first factor is nonzero, we have
    $(a^{2^{(j-1)} q} + 1) \mod n = 0$, or $a^{2^{(j-1)} q} \mod n = n-1$

Miller-Rabin Test

- The algorithm tests for case 1 or case 2.
- If the test fails, that means n is composite
- If it succeeds, n is not *guaranteed* to be prime
  - but the probability of success for a nonprime is less than 0.25
  - repeat the test for r different a's to get a prime with probability 0.75^r

Conventional Vs Public Key Crypto

- Conventional crypto is fast
  - Fast enough that we don't think twice about encrypting all internet traffic
- Public key crypto is much slower
  - At least 3 orders of magnitude slower
- Key distribution is easier with public keys
  - Need to ensure authenticity of public keys
  - For conventional keys, confidentiality is needed
- Solution
  - Use conventional crypto for encrypting bulk data
  - Use public key crypto to set up keys for such encryption.
    - conventional keys are generated by one party and sent to the other, encrypted using public keys.
  - Use certificates and certification authorities (CAs) to establish authenticity of public keys

Use of Random Numbers

- Nonces (to protect against replay attacks)
- Session key generation
- RSA key generation
- Need cryptographically strong random number generator
  - Not enough if we had "random" numbers in a statistical sense
  - Need unpredictability

Pseudorandom number generators

- Linear congruential method
  - $X_{n+1} = (aX_n + b)\ mod\ m$
  - Not good for crypto applications, as it is predictable

Natural Random Noise

- Best source is natural randomness in real world
  - Radiation counters
  - Radio noise
  - Keystroke intervals
  - Network packet arrival characteristics
  - Especially at the nanosecond timescale, these times are likely to be truly random.
- Built into OSes now
  - `/dev/random` is a source of cryptographically secure random numbers
  - But there is a limited amount, so you can't read it too often
    - `read` will block if the random pool is exhausted

Digital Signatures

- Required properties
  - receiver can verify who sends
  - sender can not repudiate
  - receiver can not generate
- Conventional crypto is not very useful
  - Sender and recipient share key, so nonrepudiability is a problem

Digital Signatures

- Public-key signature
  - Originator simply encrypts the message using private key
  - When the receiver gets the message decrypted using the originator's public key, then we can be sure about who sent the message
- Note that the encrypted message can be produced only by the originator, so all of the above properties are satisfied.

Message Digests

- Encrypting the whole message for signature purposes is impractical (too inefficient)
- Solution
  - use one-way hash functions: compute a fixed-size (e.g., 128-bit) hash on the message
  - Encrypt the hash using private key
- One-way hash code:
  - Given P, it is easy to compute $H(P)$
  - Given $H(P)$, it is impossible find P
  - No one can generate two messages that have the same message digest
    - *Strong collision resistance*

Message Digests

- Common hash functions
  - MD5 (128 bits)
  - SHA-1 (160 bits)
  - SHA-3 (224 to 512 bits)
  - SHA-256
  - RIPEMD-160
- MD5 and SHA-1 are considered weak now.

Digital Certificates

- Certificates are issued by a CA
- Every one knows the public keys of the CA
- A certificate for a principal A is simply A's public key that is encrypted with CA's private key
  - Only the CA could have produced such a message, so the recipient of the certificate knows that the CA vouches for A's public key
  - If the recipient trusts CA, then the certificate provides a simple way to authenticate the public key of A.

Public-Key Certificates

- certificates allow key exchange without real-time access to public-key authority
- a certificate binds identity to public key
  - usually with other info such as period of validity, rights of use etc
- with all contents signed by a trusted Public-Key or Certificate Authority (CA)
- can be verified by anyone who knows the public-key authorities public-key

X.509 Certificates

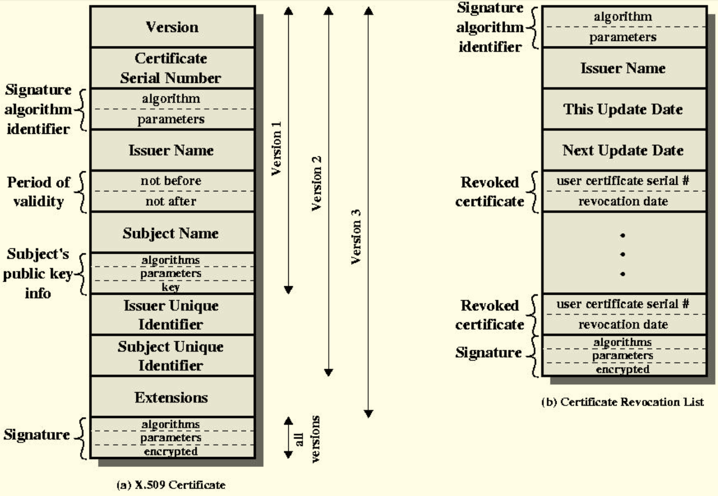


Figure 11.3 X.509 Formats