Data Encryption Algorithms

KR Chowdhary, Professor & Head
Email: kr.chowdhary@gmail.com

Department of Computer Science and Engineering
MBM Engineering College, Jodhpur

Protections

- Security policies
  1. Access policy
  2. Accountability policy
  3. Authentication policy
  4. Availability Policy
  5. Maintenance policy
  6. Violation reporting policy

Protection of users and networks

- Protection of Employees

  a. Shared key encryption schemes

  b. Hash Functions. Hash function H generates $H(m)$ for every variable-length message m. $H(m)$ is message digest

  c. Digital Signature: Binding an information to an originator.

- Protection of Networks:

  a. Firewalls

  b. Access, authorization, and authentication tools

  c. Intrusion detection systems

Cryptography

- Method of sending messages in disguised form, so that only the intended recipients can remove the disguise and read the messages.
- Terms:*plain text, cipher text, cryptosystem*

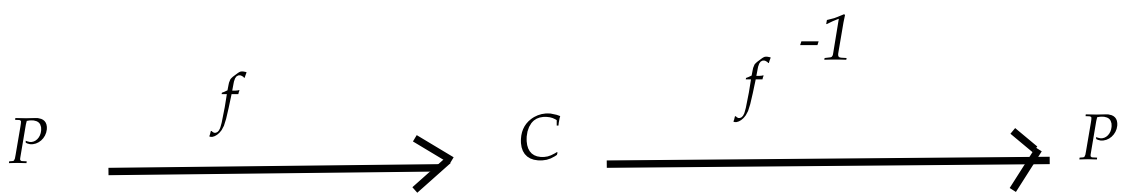


Figure: Cryptosystem

- To facilitate rapid enciphering/deciphering, there should be simple rules.
- For example: $A-Z \Longleftrightarrow 0-25$. Let $p \in \{0,\ldots,25\}$ is plain-text, and f is from $\{0,1,\ldots,25\}$ to itself:

Some Examples of secrecy systems

- Simple Substitution Cipher (Letter of the message is replaced by a fixed substitute) $M = m_1 m_2 \ldots$, where m_1, m_2 are the successive letters, becomes: $E = e_1 e_2 \cdots = f(m_1) f(m_2) \ldots$
- Transposition (Fixed Period d): The message is divided into groups of length d and a permutation applied to the first group, the same permutation to the second group, etc. Thus for d = 5, we might have 2 3 1 5 4 as the permutation. This means that: $m_1 m_2 m_3 m_4 m_5 m_6 m_7 m_8 m_9 m_{10} \ldots$ becomes $m_2 m_3 m_1 m_5 m_4 m_7 m_8 m_6 m_{10} m_9 \ldots$

Data Encryption Algorithm design Criteria

- XOR Function: $X \oplus Y = (x_1 \oplus y_1)(x_2 \oplus y_2)\ldots(x_n \oplus y_n)$
- We use E, D and K to denote encryption algorithm, a decryption algorithm, and secret key, respectively.
- Plaintext M is divided into sequence of blocks: $M_1, M_2, \ldots, M_k$, where each block is $\ell-$bit long, except possibly the last block M_k. If $|M_k| < \ell$, add an 8-bit control code at the end of M_k, once or several times to obtain a new block of size ℓ. This is called *padding*.
- An encryption algorithm which encrypts a block at a time is called block-cipher algorithm.
- When $\ell = 8$ we call it *stream-cipher* algorithm.
- The encryption algorithm encrypts M_i to produce a ciphertext block C_i is:

$$C_i = E_k(M_i) = K \oplus M_i \tag{1}$$

Data Encryption Algorithm design Criteria

- The decryption algorithm decrypts C_i into M_i is as follows:

$$D_k(C_i) = K \oplus C_i = K \oplus (K \oplus M_i) = (K \oplus K) \oplus M_i = 0^\ell \oplus M_i = M_i \tag{2}$$

- The XOR is the simplest encryption algorithm. For example, $\ell = 16$, $K = 1001101010011011$, then E encrypts 'FUN' as follows:

plaintext	F	U	N	(padding)
ASCII	01000110	01010101	01001110	00001010
secret key:$\oplus$	10011010	10011011	10011010	10011011
ciphertext:	11011100	11001110	11010100	10010001

- The XOR algorithm is simple and fast, but resulting level of security is low. The eavesdroppers can easily can easily calculate the secret key K from a plain text-cipher text pair as follows:

$$M_i \oplus C_i = M_i \oplus (M_i \oplus K) = K. \tag{3}$$

Data Encryption Algorithm design Criteria

- Attacks like this derive secret keys using a small number of samples of cipher-text data and corresponding plain text, are called *known-plain text attacks.*
- Therefore, the key must be frequently changed. If key is used once only, the XOR provides the best security. but, this requires large number of keys.
- Think of other problems of one-time pad?

Criteria for Data Encryption

Efficiency:

- Must be easy to implement on hardware and software.
- The time and space complexity of both encryption and decryption must be small and constant factor of input.
- only those operations must be employed, which are easy to implement on computer.
- Following operations are common: OR, AND, substitution, ex-OR, shift-R/L, permutations, unary operations, etc.

Criteria for Data Encryption

Resistance to statistical analysis:

- The encryption algorithm must destroy any statistical structure in the plain text data. This is satified by the *diffusion* and *confusion* properties of the algorithm.
- *diffusion:* Changing a single bit in plain text will cause a number of bits in cipher text to be changed. These should be distributed.
- *confusion:* change of single bit in key will cause number of bits in the cipher text to be changed, and evenly distributed.
- Diffusion is achieved by repeatedly executing fixed sequence of operations for a fixed rounds on strings generated from the previous round.
- Confusion may be achieved by generating number of sub-keys from the key K. And, use the subkeys to operate the plaintext in different rounds.

Criteria for Data Encryption

Resistance to Brute-Force Attacks:

- If encryption key is ℓ bit long. An eavesdropper could brute force to decipher C by calculating $M' = D_{K'}(C)$ for each ℓ bit string. There are 2^{ℓ} different ℓ-bit strings.
- The ℓ must be large. It is common belief that $\ell = 128$ bit will take many years to break.

Criteria for Data Encryption

Resistance to other attacks:

- Encryption algorithm must resists other attacks. These include chosen plain text attacks and mathematical attacks.
- *Chosen Plain text attacks:* The attacker choses a plain text message to lure the opponents to encrypt it as a bait. It contains useful information for the attacker.
- *mathematical attacks:* Attacker uses the mathematical methods to decipher the cipher text. These are: differential cryptanalysis, linear cryptanalysis, algebraic cryptanalysis.

Symmetric Encryption

- Sender and receiver share the same secret key. Algorithm is called secret key or symmetric key algorithm.
- The encryption assumes that it is computationally unfeasible to decrypt a message given the ciphertext and knowledge of encryption algorithm.
- Secret-key encryption features: Message space M, key space K, cipher space C, and two maps (encryption and decryption).

$$\phi : M \times K \rightarrow C \tag{4}$$

$$\gamma : C \times K \rightarrow M \tag{5}$$

- Both functions must satisfy:

$$\phi(\gamma(c,k),k) = c \text{ and } \gamma(\phi(m,k),k) = m$$

$\phi(-,k) : M \rightarrow C$; One way function for all k
$\phi(m,-) : K \rightarrow C$; One way function for all m

One way function

- One way functions: It is easy to compute (compute in polynomial complexity). Infeasible: not computable in polynomial time.
- Secret key must be communicated prior to use, through a secure channel.
- Disadv: Key need to be communicated securely. Hence, requires key distribution.
- Example of secret key: Data Encryption Standard (DES), block based technique, 64-bit long fixed block text, transform it into 64-bit ciphertext. (56-bits are encrypted and decrypted, 8-bits for parity check),
- Total keys: $2^{56} = 7.2 * 10^{16}$.

Model of secret-key cryptosystem

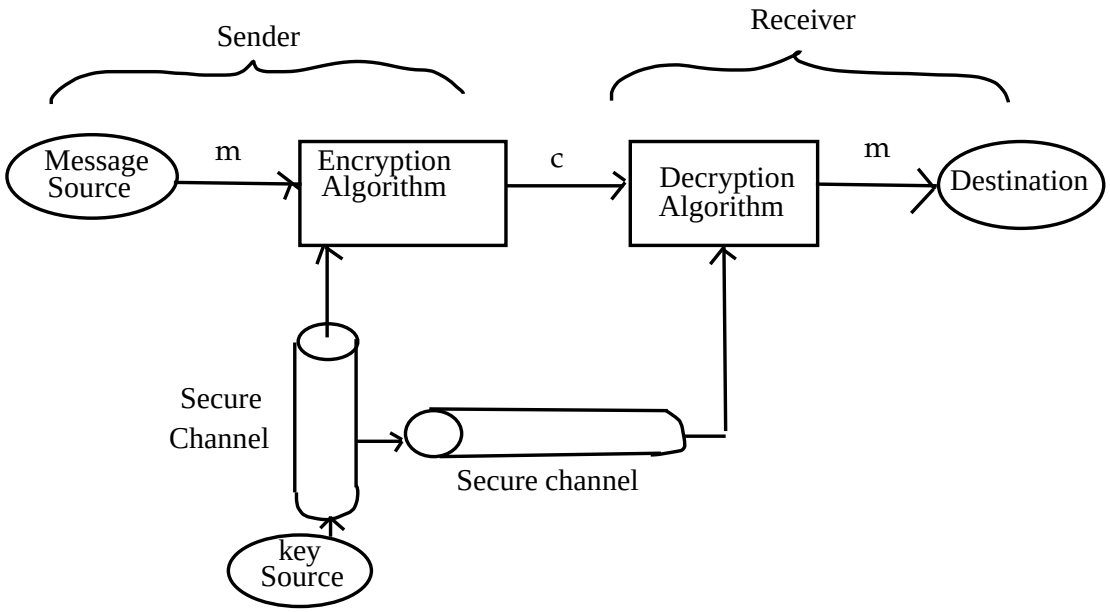


Figure: Secret-key Cryptosystem

Table: ciphertext obtained by XOR of m and k.

Plaintext:	s e c u r i t y o f e - s e r v i c e s
Key:	x k l p o r w x r u g e s m y a w f h v
Ciphertext:	k n o e 2 c a - 2 s - b e - n s w 3 e b e

Figure: DES encription.

DES Algorithm

- Algo: Series of permutations and substitutions. A block (56-bits) subjected to an initial permutations (IP), followed with complex, key dependent cycles, involving 16-sub keys and functions. Finally, through the permutation IP^{-1}.
- $c_i = m_i \oplus k_i$, for all $i \leq \text{length}(m)$, where $\oplus$ is exlusive -OR, i is ith binary bit of plaintext m, k_i is ith bit of key k, c_i is ith binary bit of cipher text c.
- Intermediate halves of ciphertext are comparable to intermediate keys. Also, it provides *non-linearity*.

$$L_i = R_{i-1} \tag{6}$$

$$R_i = L_{i-1} \oplus F(R_{i-1}, K_i). \tag{7}$$

DES Algorithm

- Figure shows the main operations executed at each cycle. Each right half expanded from 32 to 48 bits.

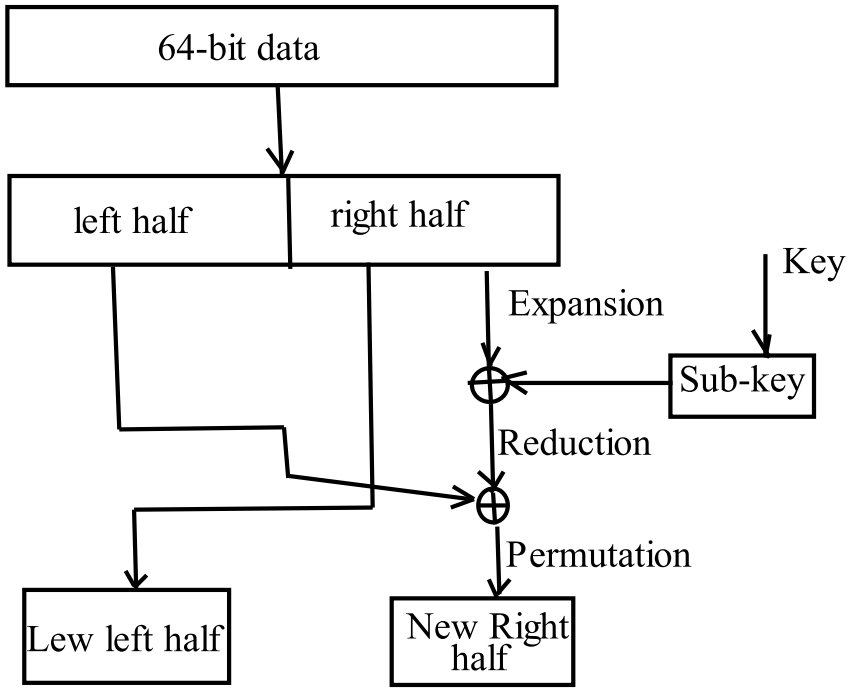


Figure: A DES Cycle.

DES Algorithm

- DES Weaknesses: Two messages m and $\underline{m}$ are called complement if $m \oplus \underline{m} = 1111\ldots$ hence if m is encrypted with a k, the complement of the resulting cipher will be the encryption of the complement $\underline{m}$, using the complement key $\underline{k}$
- Semi-key waekness: there are specific pair of keys having identical decryption, implying that two keys decrypt a message encrypted by a single key.
- Better algorithms are Triple DES (56*3=168 length), and AES.

Public-key cryptosystem

- User U_b wants to send a secret message to U_a. K_b is public, K_a is secret (in possession of U_a). There is a function:

$$\alpha : C \to C \text{ such that}$$
$$\gamma(\phi(m, \alpha(k)), k) = m, \text{ and}$$
$$\phi(\gamma(c, k), \alpha(k)) = c$$

- $(\alpha(k), k)$ are pair of private and public keys, and α is link between private and public keys.

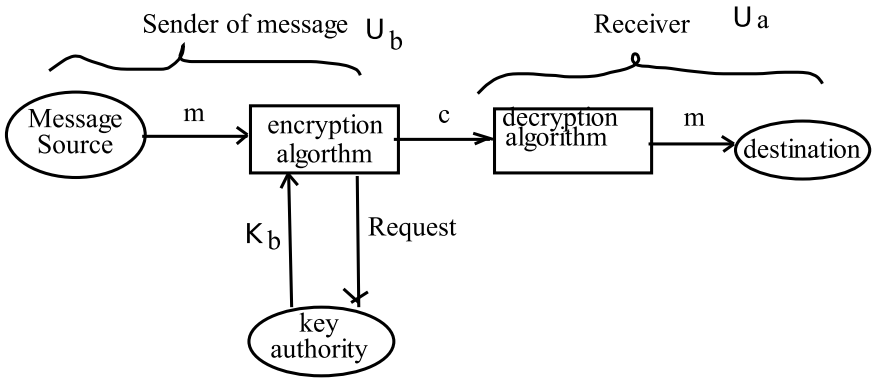


Figure: Public-key Cryptosystem

Public-key main algorithm

RSA Algorithm:

- 1978: secure and widely accepted, difficulty to determine prime factor of large integers.
- Block cypher; Plain-text, cipher-text are 0 to $n-1$.
- following requirement must be met:

  It is possible to find integers n, e and d such that:

  $m^{ed} = mmodn$, for all $0 \leq m < n$. It is relatively easy to compute $m^d modn$ and $c^e modn$ for $m < n$ but computationally infeasible to compute d given e and n.
- Two interchangeable keys (n, d) (public) and (n, e) (private-key).

Public-key main algorithm

RSA Algorithm: continued..

- Plain-text block m having value less than n is encrypted/decrypted as:
  $c = m^d \bmod n$, and $m = c^e \bmod n$.
- e is selected such that: $(m^e)^d \bmod n = m$, $e.d = k\phi(n) + 1$, $\phi(n) = (p-1)(q-1)$ is Euler Function.

  n=p.q, where p, q are primes, d is relative prime to (p-1)(q-1), i.e. d has no factors common with (p-1)(q-1)
- let p=17, q=13, $\therefore n = pq$=221,$\phi(n) = 192, d = 5, e = 77$,
- Let $m = 6, \therefore, c = 6^5 \bmod 221 = 41$. is encryption, and $m = 41^{77} \bmod 221 = 6$ is decryption.

Valuation of Cryptosystem

- Amount of Secrecy
- Size of Key
- Complexity of Enciphering and Deciphering Operations
- Propagation of Errors
- Expansion of Message

Lattice Model for secure information flow

- Lattice: Finite Set + Partial Ordering relation, such that for every pair there is least upper bound and greatest lower bound.
- $(S, \leq)$ is POSET.

Linear ordered lattice

$$S = \{A_1, A_2, \ldots, A_n\}$$

$$A_i \rightarrow A_j \text{iff } A_i \leq A_j$$

$$A_i \oplus A_j = A_{max}(i, j)$$

$$A_i \otimes A_j = A_{min}(i, j)$$

$$L = A_1; H = A_n.$$

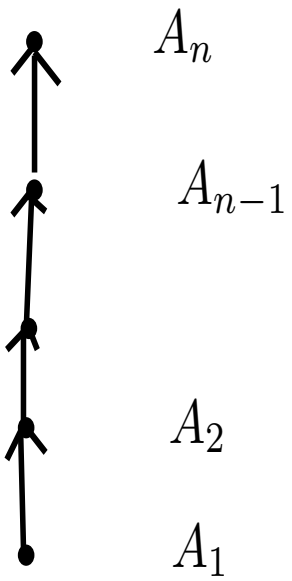


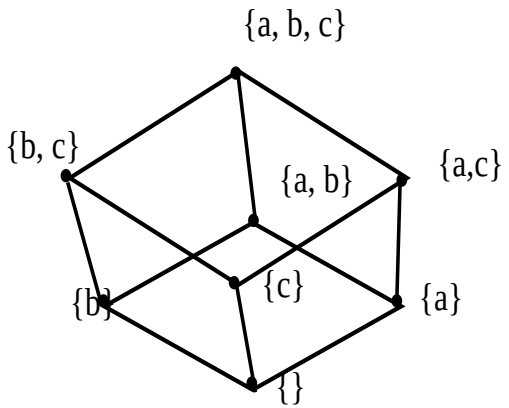
Figure: Linear ordered lattice

- Classes are linearly ordered.

Nonlinear lattice model

- It is richer structure; classes are hierarchically ordered.
- Suitable for providing security at different levels (Govt. & Military) or objects in programs, and OS.
- Information may be: unclassified, confidential, secret, top secret.

$$A \to B, (iff) A \subseteq B$$
$$A \oplus B = A \cup B$$
$$A \otimes B = A \cap B$$
$$L = \phi; H = S$$



Lattice Structure for (P(S), <=)
S={a, b, c}

Access Matrix Model

- Three principle components: *(1) Set of passive objects*, *(2) set of active subjects* (which manipulate other objects), and *(3) set of rules governing manipulation of objects by subjects*.
- Objects: files, terminals, devices. Subjects: Processes (every subject may be also an object, hence it may be manipulated by other subjects).
- Entry for a row-column reflects the mode of access and ownership for subject-object

Final Words

- No security protection mechanism is perfect yet.
- It is claimed that quantum cryptography, which is based on the quantum state of particles, will be a strong cryptosystem, and shall be impossible to break.
- On the other side, as the processing power of CPUs are increasing constantly, and 100 core processors are not far off, it will be easier to break the encryption easier using multiple desktops using these systems.