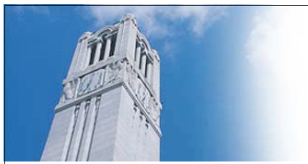


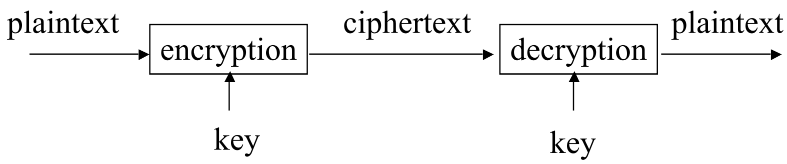
CSC 405
Introduction to Computer Security

Topic 2. Basic Cryptography (Part I)

Cryptography

- Cryptography
  - Original meaning: The art of secret writing
  - Becoming a science that relies on mathematics (number theory, algebra)
  - Process data into unintelligible form, reversible, without data loss
  - Usually one-to-one (not compression)

Encryption/Decryption



- Plaintext: a message in its original form
- Ciphertext: a message in the transformed, unrecognized form
- Encryption: the process that transforms a plaintext into a ciphertext; also known as encode and encipher
- Decryption: the process that transforms a ciphertext to the corresponding plaintext; also known as decode and decipher
- Key: the value used to control encryption/decryption
- Cryptosystem: a system for encryption and decryption

Cryptanalysis

- Ciphertext only:
  - Analyze only with the ciphertext
  - Example: Exhaustive search until “recognizable plaintext”
  - Smarter ways available
- Known plaintext:
  - Secret may be revealed (by spy, time), thus <ciphertext, plaintext> pair is obtained
  - Great for mono-alphabetic ciphers

Cryptanalysis (Cont'd)

- Chosen plaintext:
  - Choose text, get encrypted
  - Useful if limited set of messages
- Chosen ciphertext:
  - Choose ciphertext
  - Get feedback from decryption, etc.

Simple Forms of Encryption

- Substitutions
  - One letter is replaced with another
- Transpositions
  - Also called permutations
  - The order of the letters is rearranged
- Building blocks of modern cryptographic algorithms

Substitution Ciphers

- Monoalphabetic cipher (simple substitution)
  - Use a correspondence table
  - Substitute a character or symbol for each character of the original message
  - Example: Caesar cipher
    - Replace each letter with the one 3 letters later

A	B	C	D	E	F	G	H	I	J	K	L	M
d	e	f	g	h	i	j	k	l	m	n	o	p
N	O	P	Q	R	S	T	U	V	W	X	Y	Z
q	r	s	t	u	v	w	x	y	z	a	b	c

- Exercise
  - E (“COMPUTER SCIENCE”) →
  - D (“qf vwdwh”) →

Caesar Cipher

- Cryptanalysis of Caesar cipher
  - Can be done by guessing
- Clues
  - Break between two words is preserved
    - You can try common letters starting or ending a word
  - Double letters are preserved
  - Always use the same mapping
  - Exercise:
    - `wklv phvvdjh lv qrw wrr kdug wr euhdn`

Other Substitutions

- In general
  - The alphabet is scrambled
  - Each plaintext letter maps to a unique ciphertext letter
  - A substitution table can be defined using a permutation
    - A permutation is a reordering of the elements of a sequence

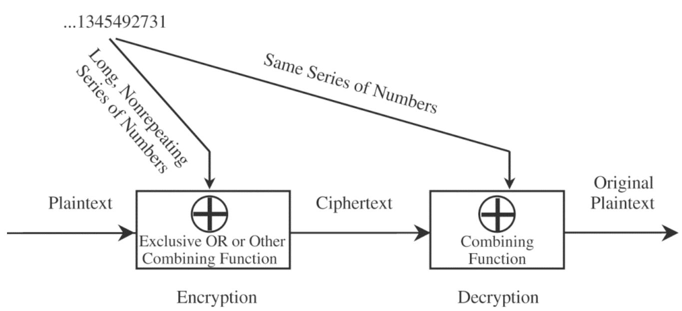
A	B	C	D	E	F	G	H	I	J	K	L	M
d	x	f	t	h	i	w	k	y	m	n	o	p
N	O	P	Q	R	S	T	U	V	W	X	Y	Z
q	r	s	g	u	v	j	e	l	z	a	b	c

Cryptanalysis of Substitution Ciphers

- Ad hoc clues
  - Short words, words with repeated patterns, common initial and final letters
- Language specific knowledge
  - Frequency of letters
    - E, T, O, and A occur far more often than J, Q, X, and Z
  - Letter patterns
    - th, er, en, ss, st, …

One-Time Pads

- Encrypt plaintext with a large, non-repeating set of keys
  - Absolute synchronization between sender and receiver
  - Unlimited number of keys



Vernam cipher

Book Cipher

- Use book, piece of music, or other object with which structure can be analyzed
  - Both sender and receiver need access to identical objects
  - Example: book cipher with Vigenère tableau
    - Key: I am, I exist, that is certain.
    - Plaintext: MACHINES CANNOT THINK

```
iamie xistt hatis cert      → column
MACHI NESCA NNOTT HINK      → row
-------------------------
Uaopm kmkvt unbhl jmed
```

	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
A	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
B	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	a
C	c	d	e	f	g	h	i	j	k	l	m	n	o	p	q	r	s	t	u	v	w	x	y	z	a	b
D	d	e	f	g	h	i	j	k	l	m	n	o	p	q	r	s	t	u	v	w	x	y	z	a	b	c
E	e	f	g	h	i	j	k	l	m	n	o	p	q	r	s	t	u	v	w	x	y	z	a	b	c	d
F	f	g	h	i	j	k	l	m	n	o	p	q	r	s	t	u	v	w	x	y	z	a	b	c	d	e
G	g	h	i	j	k	l	m	n	o	p	q	r	s	t	u	v	w	x	y	z	a	b	c	d	e	f
H	h	i	j	k	l	m	n	o	p	q	r	s	t	u	v	w	x	y	z	a	b	c	d	e	f	g
I	i	j	k	l	m	n	o	p	q	r	s	t	u	v	w	x	y	z	a	b	c	d	e	f	g	h
J	j	k	l	m	n	o	p	q	r	s	t	u	v	w	x	y	z	a	b	c	d	e	f	g	h	i
K	k	l	m	n	o	p	q	r	s	t	u	v	w	x	y	z	a	b	c	d	e	f	g	h	i	j
L	l	m	n	o	p	q	r	s	t	u	v	w	x	y	z	a	b	c	d	e	f	g	h	i	j	k
M	m	n	o	p	q	r	s	t	u	v	w	x	y	z	a	b	c	d	e	f	g	h	i	j	k	l
…																										

Vigenère Tableau

Cryptanalysis of Book Cipher

- Flaw of book cipher
  - Distributions of both key and message cluster around high frequency letters
  - Example
    - A, E, O, T, N, I account for 50% of all letters
    - Probability of both key and plaintext letters are one of them: 0.25
- Cryptanalysis
  - Look for intersections of the above six letters
  - For each cipher text letter, identify the possible plaint text letter from those intersections

```
uaopm kmkvt unbhl jmed
?AA?E ?E??A ?ANN? ?EA?
  O I   I  T  NTT   IE
    T   T           T
```

Correct prediction underlined

Transpositions (Permutations)

- Letters of the message are rearranged
  - Aim to break established patterns
- Confusion and diffusion
  - Confusion
    - Make it difficult to determine how message and key are transformed into cipher text
    - Complex relationship between plaintext, key, and ciphertext
    - Done through substitution
  - Diffusion
    - Widely spread the information from the message or key across the cipher text
    - Done through transposition (permutation)

Columnar Transpositions

- Rearrange characters of the plain text into columns

Key:	4	3	1	2	5	6	7
Plaintext:	A	T	T	A	C	K	P
	O	S	T	P	O	N	E
	D	U	N	T	I	L	T
	W	O	A	M	X	Y	Z

Cipher text: ________________________________

Cryptanalysis of Transpositions

- Diagram analysis
  - Frequent diagram
    - Patterns of pairs of adjacent letters
    - RE, EN, ER, NT, …
  - Frequent trigrams
    - Groups of three letters
    - ENT, ION, AND, ING, …
  - Infrequent diagrams and trigrams
    - VK and QP

Cryptanalysis by Diagram Analysis

- Confirms it is a transposition
  - Compute the letter frequencies
- Find adjacent columns
  - Try different column sizes
  - Look for common diagrams
  - Verify possible matches for different positions
- Rely heavily on a human's judgment of what "looks right"

Product Cipher

- Product cipher
  - Combination of two ciphers
  - Modern ciphers: interleaved substitutions and transpositions (permutations)
  - S→P→S→P→…
- But
  - How about S→P→S→S→P→…
  - How about S→P→P→S→…

"Good" Encryption algorithms

- What does it mean for a cipher to be "good"?
  - Meaning of "good" depends on intended use of the cipher
  - Commercial applications
  - Military applications

Characteristics of "Good" Ciphers

- Shannon's principles
  - The amount of secrecy needed should determine the amount of labor appropriate for the encryption and decryption
  - The set of keys and the enciphering algorithm should be free from complexity
    - No restrictions on keys or plain text; keys should be short
  - The implementation of the process should be as simple as possible
    - Formulated with hand encryption in mind
    - Implementation on a computer need not be simple, as long as the time complexity is tolerable

Characteristics of "Good" Ciphers

- Shannon's Principles (Cont'd)
  - Errors in ciphering should not propagate and cause corruption of further information in the message
    - No error propagation
  - The size of the enciphered text should be no larger than the text of the original message
    - Dramatic cipher expansion in size does not carry more information, but
    - It gives the cryptanalyst more data to infer patterns

Security of An Encryption Algorithm

- Unconditionally secure
  - It is impossible to decrypt the ciphertext
  - One-time pad (the key is as long as the plaintext)

$$C_i = P_i \oplus K_i$$

- Computationally secure
  - The cost of breaking the cipher exceeds the value of the encrypted information
  - The time required to break the cipher exceeds the useful lifetime of the information

Secret Keys v.s. Secret Algorithms

- Security by obscurity
  - We can achieve better security if we keep the algorithms secret
  - Hard to keep secret if used widely
  - Reverse engineering, social engineering
- Publish the algorithms
  - Security of the algorithms depends on the secrecy of the keys
  - Less unknown vulnerability if all the smart (good) people in the world are examine the algorithms

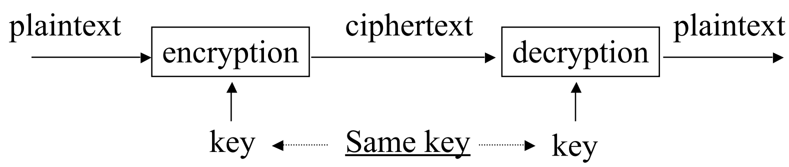
Secret Keys v.s. Secret Algorithms (Cont'd)

- Commercial world
  - Published
  - Wide review, trust
- Military
  - Keep algorithms secret
  - Avoid giving enemy good ideas
  - Military has access to the public domain knowledge anyway.

Types of Cryptography

- Number of keys
  - Hash functions: no key
  - Secret key cryptography: one key
  - Public key cryptography: two keys - public, private
- The way in which the plaintext is processed
  - Block cipher: divides input elements into blocks
  - Stream cipher: process one element (e.g., bit) a time

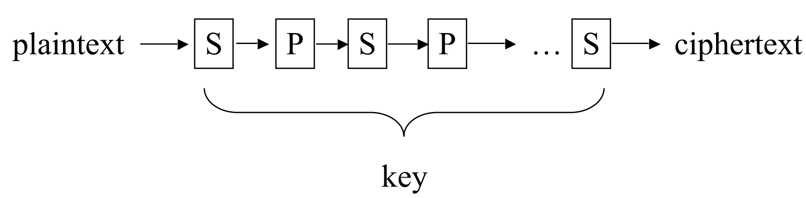
Secret Key Cryptography



- Same key is used for encryption and decryption
- Also known as
  - Symmetric cryptography
  - Conventional cryptography

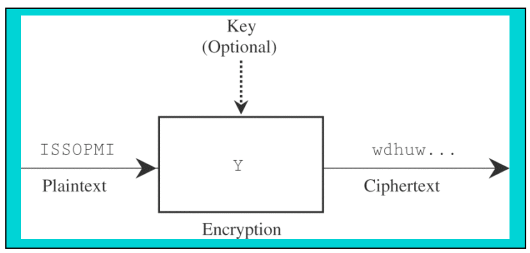
Secret Key Cryptography (Cont'd)

- Basic technique
  - Product cipher
  - Multiple applications of interleaved substitutions and permutations
- Cipher text approximately the same length as plaintext



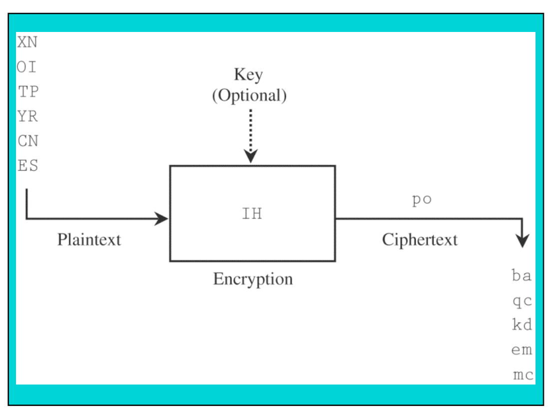
Stream and Block Ciphers

- Stream ciphers
  - Convert one symbol of plaintext immediately into a symbol of ciphertext
    - A symbol: a character, a bit
  - Examples
    - Substitution ciphers discussed earlier
    - Modern example: RC4



Stream and Block Ciphers (Cont'd)

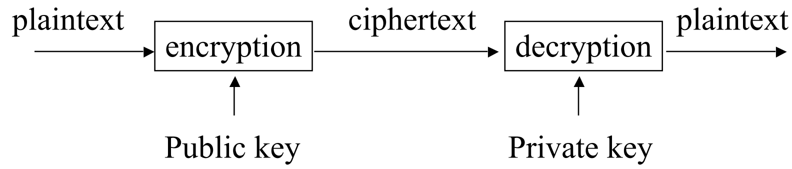
- Block cipher
  - Encrypt a group of plaintext symbols as on block
  - Examples
    - Columnar transposition
    - Modern examples: DES, AES



Applications of Secret Key Cryptography

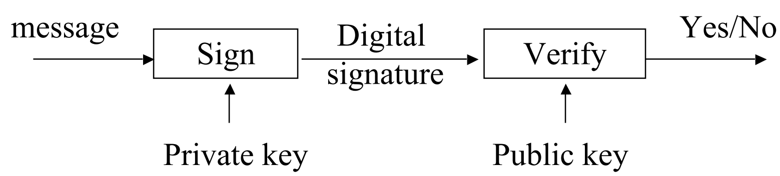
- Transmitting over an insecure channel
  - Challenge: How to share the key?
- Secure Storage on insecure media
- Authentication
  - Challenge-response
  - To prove the other party knows the secret key
  - Must be secure against chosen plaintext attack
- Integrity check
  - Message Integrity Code (MIC)
  - Also called Message Authentication Code (MAC)

Public Key Cryptography



- Invented/published in 1975
- A public/private key pair is used
  - Public key can be publicly known
  - Private key is kept secret by the owner of the key
- Much slower than secret key cryptography
- Also known as
  - Asymmetric cryptography

Public Key Cryptography (Cont'd)



- Another mode: digital signature
  - Only the party with the private key can create a digital signature.
  - The digital signature is verifiable by anyone who knows the public key.
  - The signer cannot deny that he/she has done so.

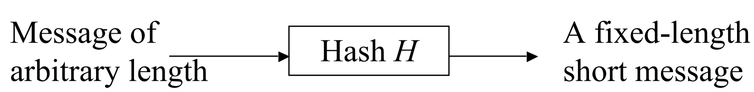
Applications of Public Key Cryptography

- Data transmission
  - Alice encrypts m_a using Bob's public key e_B, Bob decrypts m_a using his private key d_B.
- Storage
  - Can create a safety copy: using public key of trusted person.
- Authentication
  - No need to store secrets, only need public keys.
  - Secret key cryptography: need to share secret key for every person to communicate with.

Applications of Public Key Cryptography (Cont'd)

- Digital signatures
  - Sign hash $H(m)$ with the private key
    - Authorship
    - Integrity
    - Non-repudiation: can't do with secret key cryptography
- Key exchange
  - Establish a common session key between two parties

Hash Algorithms



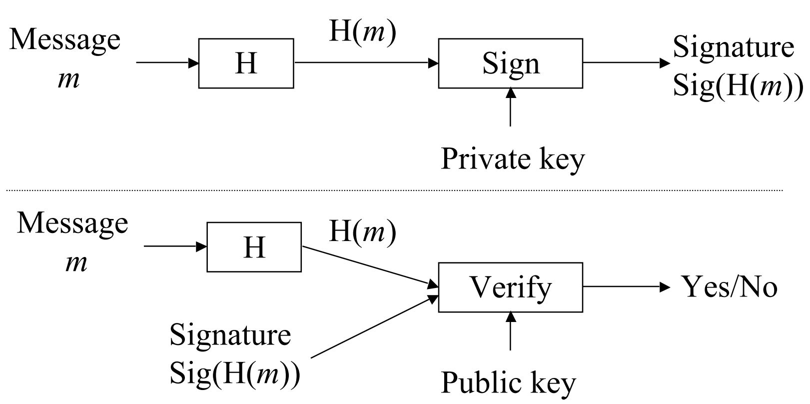
- Also known as
  - Message digests
  - One-way transformations
  - One-way functions
  - Hash functions
- Length of $H(m)$ much shorter then length of m
- Usually fixed lengths: 128 or 160 bits

Hash Algorithms (Cont'd)

- Desirable properties of hash functions
  - Performance: Easy to compute $H(m)$
  - One-way property: Given $H(m)$ but not m, it's difficult to find m
  - Weak collision free: Given $H(m)$, it's difficult to find m' such that $H(m') = H(m)$.
  - Strong collision free: Computationally infeasible to find m_1, m_2 such that $H(m_1) = H(m_2)$

Applications of Hash Functions

- Primary application
  - Generate/verify digital signature

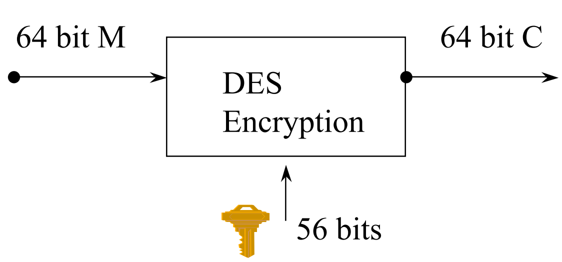


Applications of Hash Functions (Cont'd)

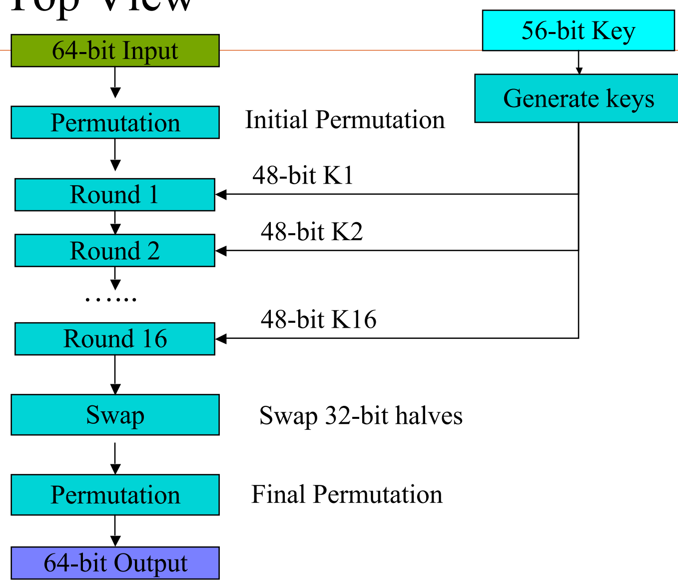
- Password hashing
  - Doesn't need to know password to verify it
  - Store *H*(*password+salt*) and salt, and compare it with the user-entered password
  - Salt makes dictionary attack more difficult
- Message integrity
  - Agree on a secrete key *k*
  - Compute *H*(*m*|*k*) and send with *m*
  - Doesn't require encryption algorithm, so the technology is exportable

DES (Data Encryption Standard)

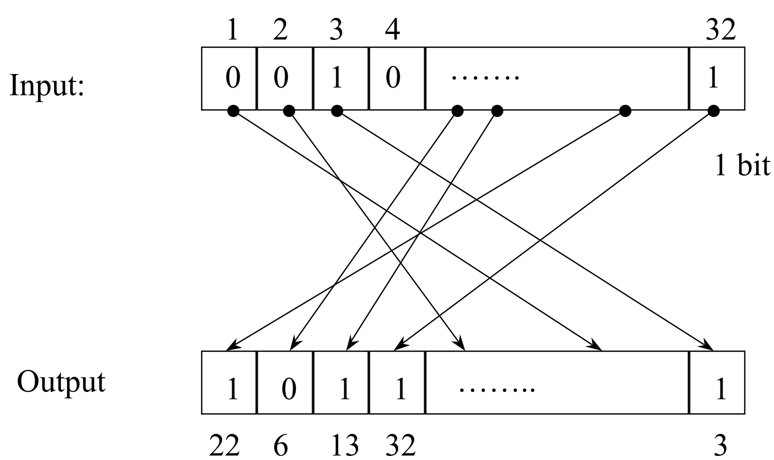
- Officially adopted in 1976
- Expired in 1998
- Key: 64 bit quantity=8-bit parity+56-bit key
  - Every 8th bit is a parity bit.
- 64 bit input, 64 bit output.



DES Top View



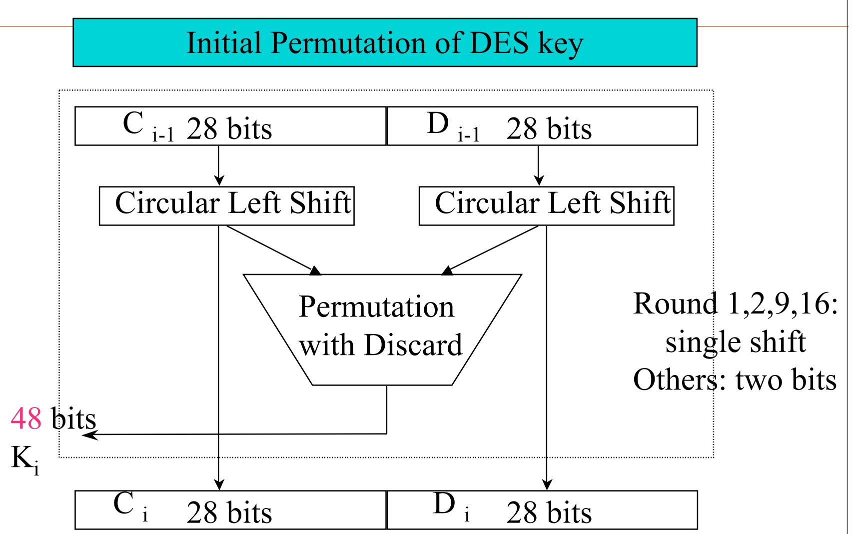
Bit Permutation (1-to-1)



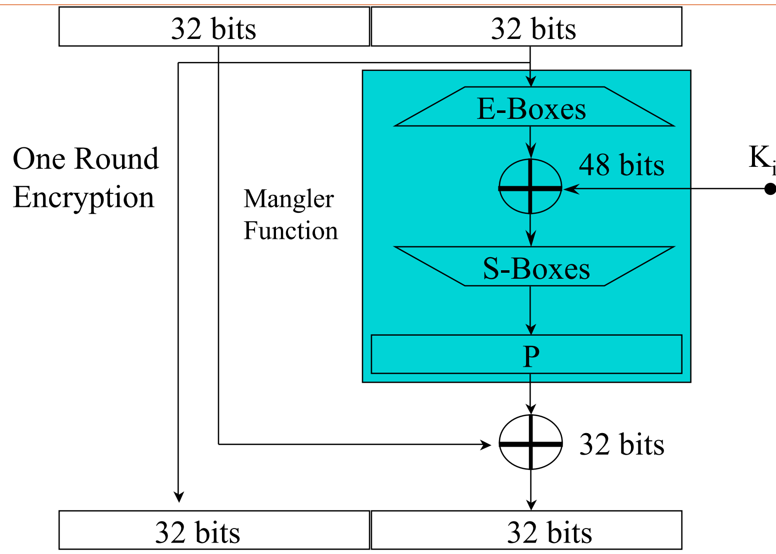
Initial and Final Permutations

- Initial permutation (IP)
- View the input as M: 8 X 8 bit matrix
- Transform M into M’ in two steps
  - Transpose row x into column (9-x), $0<x<9$
  - Apply permutation on the rows:
    - For even row y, it becomes row y/2
    - For odd row y, it becomes row (5+y/2)
- Final permutation FP = IP^{-1}

Per-Round Key Generation



A DES Round

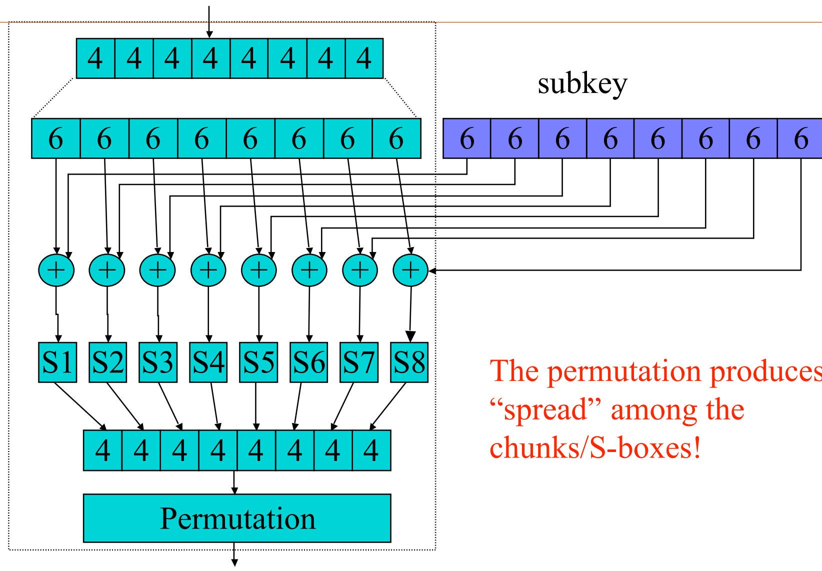


E Box of DES (Expansion Permutation)

- How is the E box defined
  - Each row expands from 4 bits to 6 bits

32	1	2	3	4	5
4	5	6	7	8	9
8	9	10	11	12	13
12	13	14	15	16	17
16	17	18	19	20	21
20	21	22	23	24	25
24	25	26	27	28	29
28	29	30	31	32	1

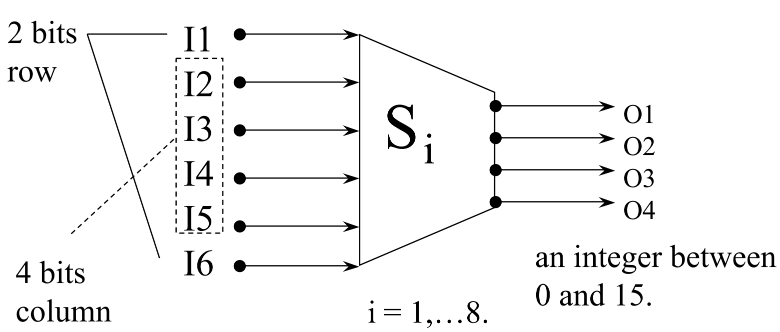
Another View of the Mangler Function



The permutation produces "spread" among the chunks/S-boxes!

S-Box (Substitute and Shrink)

- 48 bits ==> 32 bits. (8*6 ==> 8 *4)
- 2 bits used to select amongst 4 permutations for the rest of the 4-bit quantity



The First S Box S1

Each row and column contain different numbers.

	0	1	2	3	4	5	6	…	15
0	14	4	13	1	2	15	11		
1	0	15	7	4	14	2	13		
2	4	1	14	8	13	6	2		
3	15	12	8	2	4	9	1		

Example: input: 100110 output: ???

DES Standard

- Cipher Iterative Action
  - Input: 64 bits
  - Key: 48 bits
  - Output: 64 bits

- Key Generation Box
  - Input: 56 bits
  - Output: 48 bits

One round (Total 16 rounds)

Avalanche Effect

- A small change in either the plaintext or the key should produce a significant change in the ciphertext
- DES has a strong avalanche effect
- Example
  - Plaintexts: 0X0000000000000000 and 0X8000000000000000
  - Same key: 0X016B24621C181C32
  - 34 bits difference in cipher-texts
  - Similar result with same plaintext and slightly different keys

Concerns about DES

- Key space problem: 56 bit key (2^{56})
  - DESCHALL recovered RSA challenge I key on June 17, 1997 (6 month into the contest)
  - $.25m (total cost), July 15, 1998, RSA DES challenge II key recovered in 56 hours
- Cryptanalysis
  - Sixteen weak and semi-weak keys:
  - Differential cryptanalysis require less tries using chosen plaintext/ciphertext [Biham, 1993]
    - Effective up to 15 rounds
    - DES is well designed to defeat differential analysis
  - Linear cryptanalysis requires only known plaintext/ciphertext [Matsui, 1993]