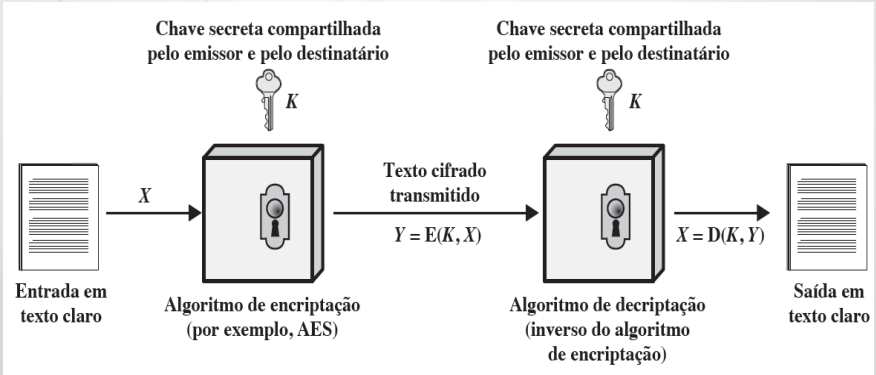
Técnicas clássicas de encriptação

Objetivos de aprendizagem

- Apresentar uma visão geral dos principais conceitos de criptografia simétrica.
- Explicar a diferença entre criptoanálise e ataque por força bruta.
- Entender a operação de uma cifra de substituição monoalfabética.
- Entender a operação de uma cifra polialfabética.
- Apresentar uma visão geral da cifra de Hill.
- Descrever a operação de uma máquina de rotor.

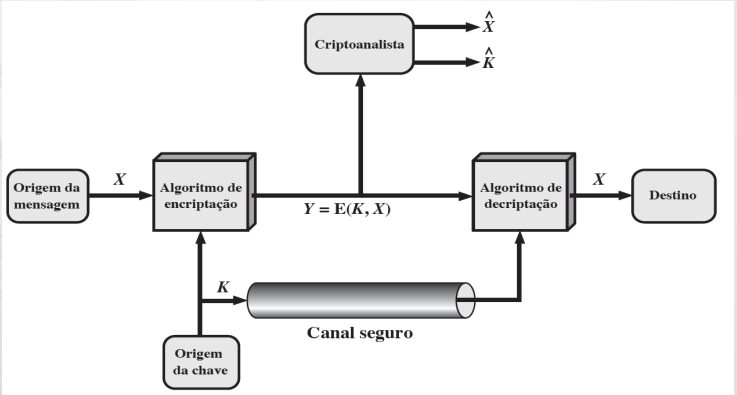
Modelo de cifra simétrica

- Modelo simplificado da encriptação simétrica:



Modelo de cifra simétrica

- Modelo de criptossistema simétrico:



Criptografia

- Os sistemas criptográficos são caracterizados ao longo de três dimensões independentes:

1. O tipo das operações usadas para transformar texto claro em texto cifrado.
2. O número de chaves usadas.
3. O modo em que o texto claro é processado.

- Existem duas técnicas gerais para o ataque a um esquema de encriptação convencional:

1. Criptoanálise
2. Ataque por força bruta

Criptografia

- Tipos de ataque sobre mensagens encriptadas:

TIPO DE ATAQUE	CONHECIDO AO CRIPTOANALISTA
Apenas texto cifrado	■ Algoritmo de encriptação ■ Texto cifrado
Texto claro conhecido	■ Algoritmo de encriptação ■ Texto cifrado ■ Um ou mais pares de texto claro-texto cifrado produzidos pela chave secreta
Texto claro escolhido	■ Algoritmo de encriptação ■ Texto cifrado ■ Mensagem de texto claro escolhida pelo criptoanalista, com seu respectivo texto cifrado gerado com a chave secreta
Texto cifrado escolhido	■ Algoritmo de encriptação ■ Texto cifrado ■ Texto cifrado escolhido pelo criptoanalista, com seu respectivo texto claro decriptado produzido pela chave secreta
Texto escolhido	■ Algoritmo de encriptação ■ Texto cifrado ■ Mensagem de texto claro escolhida pelo criptoanalista, com seu respectivo texto cifrado produzido pela chave secreta ■ Texto cifrado escolhido pelo criptoanalista, com seu respectivo texto claro decriptado produzido pela chave secreta

Técnicas de substituição

- O uso mais antigo que conhecemos de uma cifra de substituição, e o mais simples, foi feito por Júlio César.

- A cifra de César envolve substituir cada letra do alfabeto por aquela que fica três posições adiante. Por exemplo,

```
claro: meet me after the toga party
cifra: PHHW PH DIWHU WKH WRJD SDUWB
```

- Podemos definir a transformação listando todas as alternativas, da seguinte forma:

```
claro: 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
cifra: 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
```


Técnicas de substituição

- Vamos atribuir um equivalente numérico a cada letra:

a	b	c	d	e	f	g	h	i	j	k	l	m
0	1	2	3	4	5	6	7	8	9	10	11	12

n	o	p	q	r	s	t	u	v	w	x	y	z
13	14	15	16	17	18	19	20	21	22	23	24	25

Técnicas de substituição

- Então, o algoritmo pode ser expresso da forma a seguir.
- Para cada letra em texto claro *p*, substitua-a pela letra do texto cifrado *C*:

$$C = \mathrm{E}(3, p) = (p + 3) \bmod 26$$

- Um deslocamento pode ser de qualquer magnitude, de modo que o algoritmo de César geral é

$$C = \mathrm{E}(k, p) = (p + k) \bmod 26$$

- onde *k* assume um valor no intervalo de 1 a 25. O algoritmo de decriptação é simplesmente

$$p = \mathrm{D}(k, C) = (C - k) \bmod 26$$

Cifras monoalfabéticas

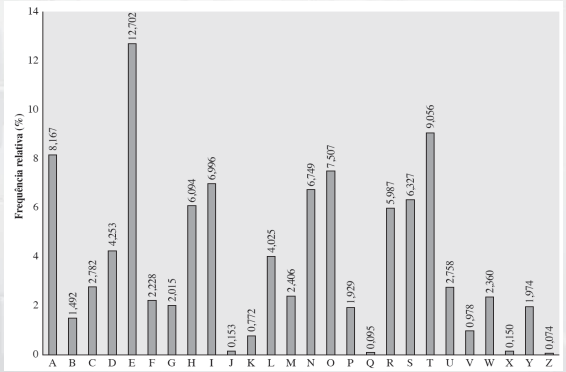
- Com apenas 25 chaves possíveis, a cifra de César está longe de ser segura.
- Uma **permutação** é um conjunto finito de elementos *S* em uma sequência ordenada de todos os elementos de *S*, com cada um aparecendo exatamente uma vez.

As frequências relativas das letras no texto cifrado são as seguintes:

P 13,33	H 5,83	F 3,33	B 1,67	C 0,00
Z 11,67	D 5,00	W 3,33	G 1,67	K 0,00
S 8,33	E 5,00	Q 2,50	Y 1,67	L 0,00
U 8,33	V 4,17	T 2,50	I 0,83	N 0,00
O 7,50	X 4,17	A 1,67	J 0,83	R 0,00
M 6,67				

Cifras monoalfabéticas

- Frequência relativa de letras no texto em inglês:



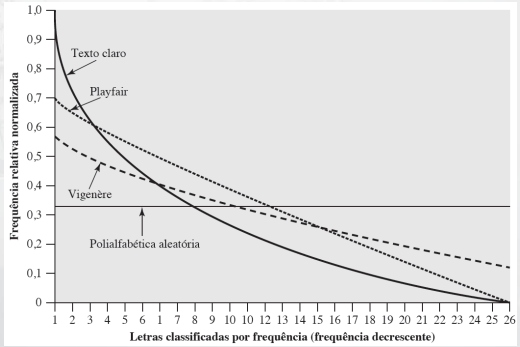
Cifra Playfair

- O algoritmo Playfair é baseado no uso de uma matriz 5 × 5 de letras construídas usando uma palavra-chave.
- Aqui está um exemplo, solucionado por Lord Peter Wimsey em *Have His Carcase*, de Dorothy Sayers:

M	O	N	A	R
C	H	Y	B	D
E	F	G	I/J	K
L	P	Q	S	T
U	V	W	X	Z

Cifra Playfair

- Um modo de revelar a eficácia da Playfair e outras cifras aparece abaixo:



Cifra de Hill

- O algoritmo de Hill utiliza m letras de texto claro sucessivas e as substitui por *m* letras de texto cifrado.
- A substituição é determinada por m equações lineares, em que cada caractere recebe um valor numérico (a = 0, b = 1, ..., z = 25).
- Para *m* = 3, o sistema pode ser descrito da seguinte forma:

$$c_1 = (k_{11}p_1 + k_{21}p_2 + k_{31}p_3) \bmod 26$$
$$c_2 = (k_{12}p_1 + k_{22}p_2 + k_{32}p_3) \bmod 26$$
$$c_3 = (k_{13}p_1 + k_{23}p_2 + k_{33}p_3) \bmod 26$$

Cifra de Hill

- Isso pode ser expresso em termos de vetores de linhas e matrizes:

$$(c_1\ c_2\ c_3) = (p_1\ p_2\ p_3)\begin{pmatrix} k_{11} & k_{12} & k_{13} \\ k_{21} & k_{22} & k_{23} \\ k_{31} & k_{32} & k_{33} \end{pmatrix} \bmod 26$$

- ou

$$\mathbf{C} = \mathbf{PK} \bmod 26$$

- onde **C** e **P** são vetores de coluna de tamanho 3, representando o texto claro e o texto cifrado, e **K** é uma matriz 3 × 3, indicando a chave de encriptação. As operações são realizadas com mod 26.

Cifras polialfabéticas

- Outra forma de melhorar a técnica monoalfabética simples é usar diferentes substituições monoalfabéticas enquanto se prossegue pela mensagem de texto claro.
- O nome geral para essa técnica é **cifra por substituição polialfabética**. Características em comum:

1. Um conjunto de regras de substituição monoalfabéticas é utilizado.
2. Uma chave define qual regra em particular é escolhida para determinada transformação.

Cifras polialfabéticas

Cifra de Vigenère

- Considere uma sequência de letras em texto claro $P = p_0, p_1, p_2, \ldots, p_{n-1}$ e uma chave consistindo na sequência de letras $K = k_0, k_1, k_2, \ldots, k_{m-1}$, onde normalmente $m < n$.
- A sequência de letras em texto cifrado $C = C_0, C_1, C_2, \ldots, C_{n-1}$ é calculada da seguinte forma:

$$C = C_0, C_1, C_2, \ldots, C_{n-1} = \mathrm{E}(K, P) = \mathrm{E}[(k_0, k_1, k_2, \ldots, k_{m-1}), (p_0, p_1, p_2, \ldots, p_{n-1})]$$
$$= (p_0 + k_0) \bmod 26, (p_1 + k_1) \bmod 26, \ldots, (p_{m-1} + k_{m-1}) \bmod 26,$$
$$(p_m + k_0) \bmod 26, (p_{m+1} + k_1) \bmod 26, \ldots, (p_{2m-1} + k_{m-1}) \bmod 26, \ldots$$

Cifras polialfabéticas

- Uma equação geral do processo de encriptação é

$$C_i = (p_i + k_{i \bmod m}) \bmod 26$$

- De modo semelhante, a decriptação é $p_i = (C_i - k_{i \bmod m}) \bmod 26$
- Para encriptar uma mensagem, é preciso que haja uma chave tão longa quanto ela. Normalmente, a chave é uma palavra-chave repetida.

```
chave:            deceptivedeceptivedeceptive
texto claro:      wearediscoveredsaveyourself
texto cifrado:    ZICVTWQNGRZGVTWAVZHCQYGLMGJ
```


Cifras polialfabéticas

- Expresso numericamente, temos o seguinte resultado:

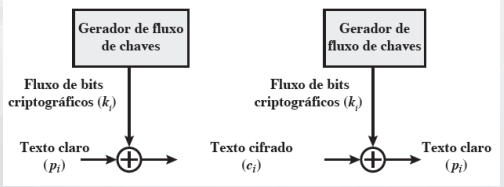
chave	3	4	2	4	15	19	8	21	4	3	4	2	4	15
texto claro	22	4	0	17	4	3	8	18	2	14	21	4	17	4
texto cifrado	25	8	2	21	19	22	16	13	6	17	25	6	21	19

chave	19	8	21	4	3	4	2	4	15	19	8	21	4
texto claro	3	18	0	21	4	24	14	20	17	18	4	11	5
texto cifrado	22	0	21	25	7	2	16	24	6	11	12	6	9

Cifras polialfabéticas

Cifra de Vernam

- A principal defesa contra a técnica criptoanalítica descrita é escolher uma palavra-chave que seja tão longa quanto o texto claro e que não possua relacionamento estatístico com ele.



Cifras polialfabéticas

- O sistema pode ser expresso de forma sucinta da seguinte forma

$$c_i = p_i \oplus k_i$$

- Assim, o texto cifrado é gerado realizando-se o XOR (operação lógica ou-exclusivo) bit a bit entre texto claro e a chave.
- Por conta das propriedades do XOR, a decriptação simplesmente envolve a mesma operação bit a bit:

$$p_i = c_i \oplus k_i$$

Cifras polialfabéticas

One-time pad

- Um oficial do Exército, Joseph Mauborgne, propôs uma melhoria na cifra de Vernam, que gera o máximo em segurança.
- Mauborgne sugeriu o uso de uma chave aleatória que fosse tão grande quanto a mensagem, de modo que a chave não precisasse ser repetida.
- Além disso, a chave deve ser empregada para encriptar e decriptar uma única mensagem, e depois descartada.

Cifras polialfabéticas

- Cada nova mensagem exige uma nova chave com o mesmo tamanho.
- Esse esquema, conhecido como ***one-time pad***, é inquebrável.
- Ele produz saída aleatória que não possui qualquer relacionamento estatístico com o texto claro.
- Como o texto cifrado não contém qualquer informação sobre o texto claro, simplesmente não existe um meio de quebrar o código.

Técnicas de transposição

- Uma espécie bem diferente de mapeamento é obtida realizando-se algum tipo de permutação nas letras do texto claro.
- Essa técnica é referenciada como uma **cifra de transposição**.
- A cifra mais simples desse tipo é a técnica de **cerca de trilho**, em que o texto claro é escrito como uma sequência de diagonais, e depois lido como uma sequência de linhas.

Técnicas de transposição

- A cifra de transposição pode se tornar muito mais segura realizando-se mais de um estágio de transposição.
- O resultado é uma permutação mais complexa, que não é facilmente reconstruída.
- Assim, se a mensagem anterior for reencriptada usando o mesmo algoritmo,

```
Chave:     4 3 1 2 5 6 7
Entrada:   t t n a a p t
           m t s u o a o
           d w c o i x k
           n l y p e t z
Saída:     NSCYAUOPTTWLTMDNAOIEPAXTTOKZ
```


Técnicas de transposição

- Para visualizar o resultado dessa dupla transposição, designe as letras na mensagem de texto claro original pelos números que indicam a sua posição. Assim, com 28 letras na mensagem, a sequência original das letras é

```
01 02 03 04 05 06 07 08 09 10 11 12 13 14
15 16 17 18 19 20 21 22 23 24 25 26 27 28
```

- Depois da primeira transposição, temos

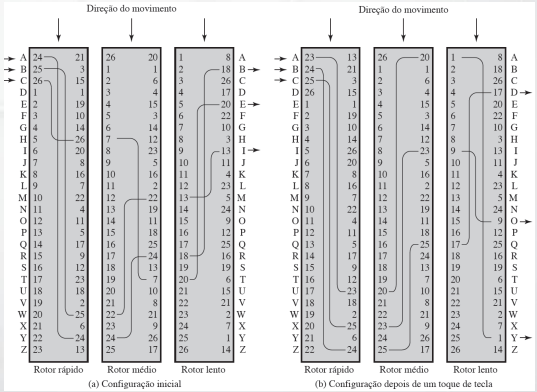
```
03 10 17 24 04 11 18 25 02 09 16 23 01 08
15 22 05 12 19 26 06 13 20 27 07 14 21 28
```

- Mas, depois da segunda transposição, temos

```
17 09 05 27 24 16 12 07 10 02 22 20 03 25
15 13 04 23 19 14 11 01 26 21 18 08 06 28
```


Máquinas de rotor

Máquina de três rotores com fiação representada por contatos numerados:



Esteganografia

- Uma forma simples de esteganografia, mas demorada de se construir, é aquela em que um arranjo de palavras e letras dentro de um texto aparentemente inofensivo soletra a mensagem real.
- Por exemplo, a sequência de primeiras letras de cada palavra da mensagem geral soletra a mensagem escondida.
- Diversas outras técnicas têm sido usadas historicamente, e alguns exemplos são [MYER91]:

Esteganografia

- **Marcação de caractere**: letras selecionadas do texto impresso ou datilografado são escritas com lápis por cima.
- **Tinta invisível**: diversas substâncias podem ser usadas para a escrita sem deixar rastros visíveis.
- **Perfurações**: pequenos furos em letras selecionadas normalmente não são visíveis.
- **Fita corretiva de máquina de escrever**: usada entre as linhas digitadas com uma fita preta, os resultados de digitar com a fita corretiva são visíveis apenas sob uma luz forte.

Cifras de bloco e o data encryption standard

Objetivos de aprendizagem

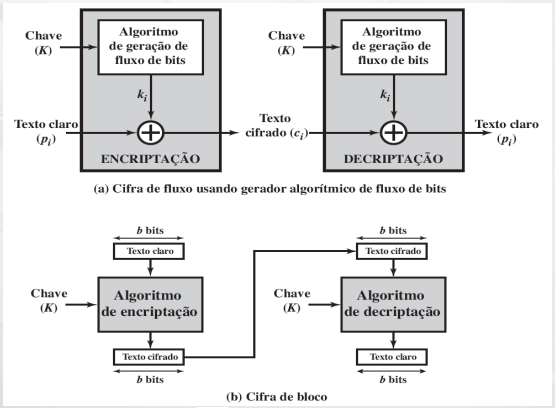
- Entender a distinção entre cifras de fluxo e cifras de bloco.
- Apresentar uma visão geral da cifra de Feistel e explicar como a decriptação é o inverso da encriptação.
- Apresentar uma visão geral do data encryption standard (DES).
- Explicar o conceito do efeito avalanche.
- Discutir a força criptográfica do DES.
- Resumir os princípios mais importantes do projeto de uma cifra de bloco.

Estrutura tradicional de cifra de bloco

- Uma **cifra de fluxo** é aquela que encripta um fluxo de dados digital um bit ou um byte por vez.
- Uma **cifra de bloco** é aquela em que um bloco de texto claro é tratado como um todo e usado para produzir um de texto cifrado com o mesmo tamanho.
- Uma cifra de bloco opera sobre um bloco de texto claro de n bits para produzir um bloco de texto cifrado de n bits.
- A figura a seguir ilustra a lógica de uma cifra de substituição geral para $n = 4$.

Estrutura tradicional de cifra de bloco

- Cifra de fluxo e cifra de bloco:

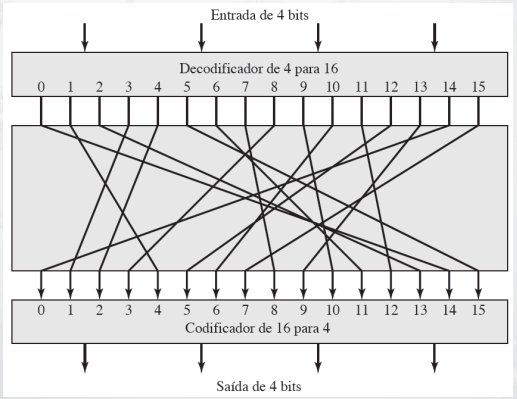


(a) Cifra de fluxo usando gerador algorítmico de fluxo de bits

(b) Cifra de bloco

Estrutura tradicional de cifra de bloco

- Substituição de bloco geral de *n* bits para *n* bits (mostrados com $n = 4$):



Cifra de feistel

- Em particular, Feistel propôs o uso de uma cifra que alterna substituições e permutações:

- **Substituição**: cada elemento de texto claro ou grupo de elementos é substituído exclusivamente por um elemento ou grupo de elementos de texto cifrado correspondente.

- **Permutação**: uma sequência de elementos de texto claro é substituída por uma permutação dessa sequência. Ou seja, nenhum elemento é acrescentado, removido ou substituído na sequência, mas a ordem em que os elementos aparecem nela é mudada.

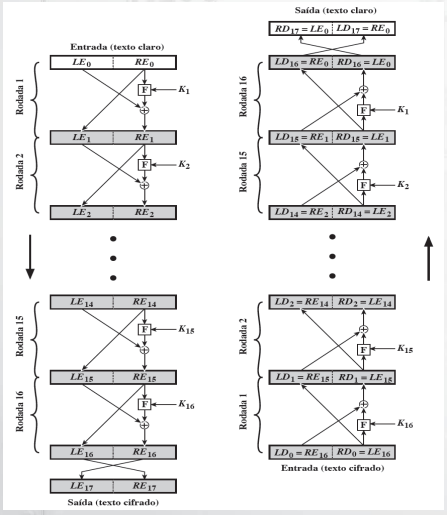
Cifra de feistel

- Os termos **difusão** e **confusão** foram introduzidos por Claude Shannon para abranger os dois ingredientes básicos para a montagem de qualquer sistema criptográfico [SHAN49].

- Na **difusão**, a estrutura estatística do texto claro é dissipada em estatísticas de longa duração do texto cifrado.

- A **confusão** procura estabelecer o relacionamento entre as estatísticas do texto cifrado e o valor da chave de encriptação o mais complexo possível, novamente para frustrar tentativas de descobrir a chave.

Cifra de feistel



Encriptação e decriptação de Feistel (16 rodadas):

Cifra de feistel

- A execução exata de uma rede de Feistel depende da escolha dos seguintes parâmetros e recursos de projeto:
- **Tamanho de bloco**: tamanhos de bloco maiores significam maior segurança (mantendo as outras coisas iguais), mas velocidade de encriptação/decriptação reduzida para determinado algoritmo.
- **Tamanho de chave**: tamanho de chave maior significa maior segurança, mas pode diminuir a velocidade de encriptação/decriptação.

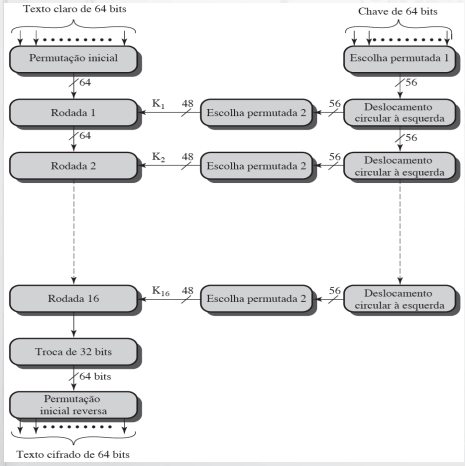
Cifra de feistel

- **Número de rodadas**: a essência da cifra de Feistel é que uma única rodada oferece segurança inadequada, mas várias proporcionam maior segurança. Um tamanho típico é de 16 rodadas.
- **Algoritmo de geração de subchave**: maior complexidade nesse algoritmo deverá levar a maior dificuldade de criptoanálise.
- **Função F**: novamente, maior complexidade geralmente significa maior resistência à criptoanálise.

Data encryption standard

- DES foi adotado em 1977 pelo National Bureau of Standards.
- O algoritmo é conhecido como data encryption algorithm (DEA).
- Para DEA, os dados são encriptados em blocos de 64 bits usando uma chave de 56 bits.
- O algoritmo transforma a entrada de 64 bits em uma série de etapas para uma saída de 64 bits. As mesmas etapas, com a mesma chave, são empregadas para reverter a encriptação.

Encriptação DES



Representação geral do algoritmo de encriptação DES:

Um exemplo do DES

Rodada	K_i	L_i	R_i
IP		5a005a00	3cf03c0f
1	1e030f03080d2930	3cf03c0f	bad22845
2	0a31293432242318	bad22845	99e9b723
3	23072318201d0c1d	99e9b723	0bae3b9e
4	05261d3824311a20	0bae3b9e	42415649
5	3325340136002c25	42415649	18b3fa41
6	123a2d0d04262a1c	18b3fa41	9616fe23
7	021f120b1c130611	9616fe23	67117cf2
8	1c10372a2832002b	67117cf2	c11bfc09
9	04292a380c341f03	c11bfc09	887fbc6c
10	270321260728040 3	887fbc6c	600f7e8b
11	2826390c31261504	600f7e8b	f596506e
12	12071c241a0a0f08	f596506e	738538b8
13	300935393c0d100b	738538b8	c6a62c4e
14	311e09231321182a	c6a62c4e	56b0bd75
15	283d3e0227072528	56b0bd75	75e8fd8f
16	2921080b13143025	75e8fd8f	25896490
IP^{-1}		da02ce3a	89ecac3b

Efeito avalanche

- Efeito avalanche no DES: mudança no texto claro:

Rodada		δ
	02468aceeca86420 12468aceeca86420	1
1	3cf03c0fbad22845 3cf03c0fbad32845	1
2	bad2284599e9b723 bad3284539a9b7a3	5
3	99e9b7230bae3b9e 39a9b7a3171cb8b3	18
4	0bae3b9e42415649 171cb8b3ccaca55e	34
5	4241564918b3fa41 ccaca55ed16c3653	37
6	18b3fa419616fe23 d16c3653cf402c68	33
7	9616fe2367117cf2 cf402c682b2cefbc	32
8	67117cf2c11bfc09 2b2cefbc99f91153	33

Rodada		δ
9	c11bfc09887fbc6c 99f911532eed7d94	32
10	887fbc6c600f7e8b 2eed7d94d0f23094	34
11	600f7e8bf596506e d0f23094455da9c4	37
12	f596506e738538b8 455da9c47f6e3cf3	31
13	738538b8c6a62c4e 7f6e3cf34bc1a8d9	29
14	c6a62c4e56b0bd75 4bc1a8d91e07d409	33
15	56b0bd7575e8fd8f 1e07d4091ce2e6dc	31
16	75e8fd8f25896490 1ce2e6dc365e5f59	32
IP^{-1}	da02ce3a89ecac3b 057cde97d7683f2a	32

Efeito avalanche

- Efeito avalanche no DES: mudança na chave:

Rodada		δ
	02468aceeca86420 02468aceeca86420	0
1	3cf03c0fbad22845 3cf03c0f9ad628c5	3
2	bad2284599e9b723 9ad628c59939136b	11
3	99e9b7230bae3b9e 9939136b768067b7	25
4	0bae3b9e42415649 768067b75a8807c5	29
5	4241564918b3fa41 5a8807c5488dbe94	26
6	18b3fa419616fe23 488dbe94aba7fe53	26
7	9616fe2367117cf2 aba7fe53177d21e4	27
8	67117cf2c11bfc09 177d21e4548f1de4	32

Rodada		δ
9	c11bfc09887fbc6c 548f1de471f64dfd	34
10	887fbc6c600f7e8b 71f64dfd4279876c	36
11	600f7e8bf596506e 4279876c399fdc0d	32
12	f596506e738538b8 399fdc0d6d208dbb	28
13	738538b8c6a62c4e 6d208dbbb9bdeeaa	33
14	c6a62c4e56b0bd75 b9bdeeaad2c3a56f	30
15	56b0bd7575e8fd8f d2c3a56f2765c1fb	33
16	75e8fd8f25896490 2765c1fb01263dc4	30
IP^{-1}	da02ce3a89ecac3b ee92b50606b62b0b	30

A força do DES

- Com um tamanho de chave de 56 bits, existem 2^{56} chaves possíveis, o que é aproximadamente $7{,}2 \times 10^{16}$ chaves.
- Assim, um ataque de força bruta parece ser impraticável.
- Supondo que, em média, metade do espaço de chave tenha que ser pesquisado, uma única máquina realizando uma encriptação DES por microssegundo levaria mais de mil anos para quebrar a cifra.
- A tabela a seguir mostra quanto tempo é necessário a um ataque de força bruta para diversos tamanhos de chave.

A força do DES

- Tempo médio exigido para uma busca exaustiva no espaço de chaves:

Tamanho de chave (bits)	Cifra	Número de chaves alternativas	Tempo exigido a 10^9 decriptações/s	Tempo exigido a 10^{13} decriptações/s
56	DES	$2^{56} \approx 7{,}2 \times 10^{16}$	2^{55} ns = 1,125 ano	1 hora
128	AES	$2^{128} \approx 3{,}4 \times 10^{38}$	2^{127} ns = $5{,}3 \times 10^{21}$ anos	$5{,}3 \times 10^{17}$ anos
168	Triple DES	$2^{168} \approx 3{,}7 \times 10^{50}$	2^{167} ns = $5{,}8 \times 10^{33}$ anos	$5{,}8 \times 10^{29}$ anos
192	AES	$2^{192} \approx 6{,}3 \times 10^{57}$	2^{191} ns = $9{,}8 \times 10^{40}$ anos	$9{,}8 \times 10^{36}$ anos
256	AES	$2^{256} \approx 1{,}2 \times 10^{77}$	2^{255} ns = $1{,}8 \times 10^{60}$ anos	$1{,}8 \times 10^{56}$ ano
26 caracteres (permutação)	Monoalfabético	$2! = 4 \times 10^{26}$	2×10^{26} ns = $6{,}3 \times 10^{9}$ anos	$6{,}3 \times 10^{6}$ anos

Princípios de projeto de cifra de bloco

- Quanto maior o número de rodadas, mais difícil é realizar a criptoanálise, mesmo para uma função F relativamente fraca.
- Em geral, o critério deverá ser de que o número de rodadas seja escolhido de modo que os esforços criptoanalíticos conhecidos exijam maior ação do que um ataque de busca de chave por força bruta.
- Quanto menos linear for F, mais difícil será qualquer tipo de criptoanálise.

Princípios de projeto de cifra de bloco

- Em termos gerais, quanto mais difícil for aproximar F de um conjunto de equações lineares, mais não linear será F.
- **Critério de avalanche estrito**: afirma que qualquer bit de saída *j* de uma S-boxes deverá mudar com probabilidade 1/2 quando qualquer bit de entrada isolado *i* for invertido para todo *i*, *j*.
- **Critério de independência de bit**: afirma que os bits de saída *j* e *k* devem mudar independentemente quando qualquer bit de entrada isolado *i* for invertido, para todo *i*, *j* e *k*.