Segurança da Informação– GBC083

Prof. Rodrigo Sanches Miani – FACOM/UFU

Aula passada

Segurança da Informação– GBC083

Ataques aos sistemas de criptografia

Existem duas técnicas:

1. Criptoanálise – estudar a natureza do algoritmo, conhecer características do texto claro ou até obter amostras de pares de texto claro-cifrado;
2. Ataque por força bruta – testar todas as chaves possíveis em um texto cifrado.

Ataques aos sistemas de criptografia - Criptoanálise

Definição 1:

- *um esquema de criptografia é* ***incondicionalmente seguro*** *se o texto cifrado gerado por ele não tiver informação suficiente para determinar o texto claro correspondente, não importa quanto texto cifrado esteja à disposição do atacante.*

Ataques aos sistemas de criptografia - Criptoanálise

Definição 2:

- *um esquema de criptografia é considerado* ***computacionalmente seguro*** *se um dos critérios a seguir for atendido:*
  - *i)* O custo para quebrar a cifra ultrapassa o valor da informação cifrada;
  - ii) O tempo exigido para quebrar a cifra supera o tempo de vida útil da informação;

Técnicas clássicas

1. Substituição
2. Transposição
3. OTP
4. Esteganografia
5. Máquinas de rotor

Cifra de Vigenère

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

Exemplo:
Texto claro: ATACARBASESUL
Chave: LIMAOLIMAOLIM
Texto cifrado: LBMCOCJMSSDCX

Técnicas clássicas – *One-time pad*

1. Usar uma chave **aleatória** tão grande quanto a mensagem;
2. A chave deve ser usada para cifrar e decifrar uma mensagem, e depois **descartada**;
3. Cada **nova** mensagem exige uma **nova** chave com o mesmo tamanho.

Dessa forma, o texto cifrado é uma saída aleatória que não possui nenhum relacionamento estatístico com o texto claro, ou seja, não existe um meio de quebrar a criptografia!

Técnicas clássicas – *One-time pad*

texto cifrado:	ANKYODKYUREPFJBYOJDSPLREYIUNOFDOIUERFPLUYTS
chave:	*pxlmvmsydofuyrvzwc tnlebnecvgdupahfzzlmnyih*
texto claro:*	mr mustard with the candlestick in the hall
texto cifrado:	ANKYODKYUREPFJBYOJDSPLREYIUNOFDOIUERFPLUYTS
chave:	*mfugpmiydgaxgoufhkllmhsqdqogtewbqfgyovuhwt*
texto claro:**	miss scarlet with the knife in the library

Dado qualquer texto claro de mesmo tamanho do texto cifrado, existe uma chave que o produz. Portanto, se você fizesse uma busca exaustiva em todas as chaves possíveis, acabaria com muitos textos claros legíveis, sem saber o original! Como o criptonalista sairia dessa situação?

Técnicas clássicas – *One-time pad*

- A segurança do *one-time pad* é inteiramente decorrente da **aleatoriedade** da chave;
- Se o fluxo de caracteres que constitui a chave for verdadeiramente aleatório, então o de caracteres que constitui o texto cifrado também o será;
- Ou seja, achamos o sistema criptográfico perfeito!
- Será?!

Técnicas clássicas – *One-time pad*

Problemas práticos:

1. Como criar grandes quantidades (e qualidade!) de chaves aleatórias regularmente?

2. Como distribuir e proteger uma quantidade absurda de chaves?

- Tais problemas atestam a utilidade **limitada** de tal tipo de cifra.

Tópicos das próximas aulas

Criptografia simétrica – Cifras de bloco

1. Estrutura tradicional
2. DES
3. Projeto de cifra de bloco

Tópicos das próximas aulas

Criptografia simétrica – Cifras de bloco

1. **Estrutura tradicional**
2. DES
3. Projeto de cifra de bloco

Tópicos das próximas aulas

Criptografia simétrica – Cifras de bloco

1. **Estrutura tradicional**
   1. **Cifra de bloco x Cifra de fluxo**
   2. **Motivação para a cifra de Feistel**
   3. **Cifra de Feistel**

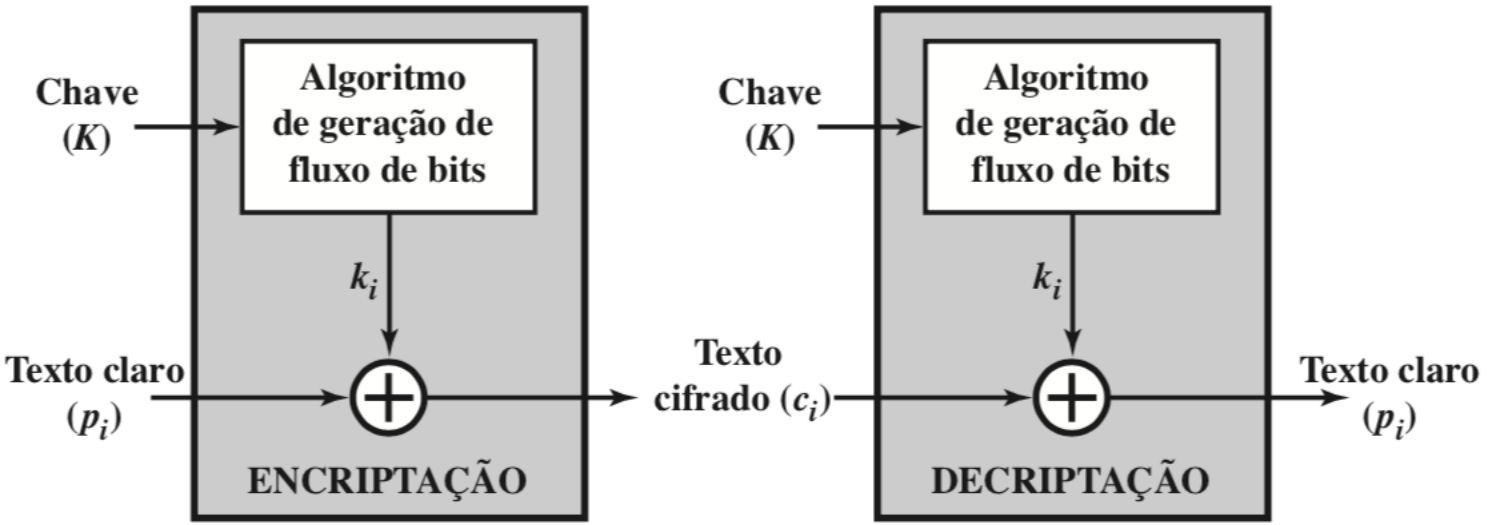
Cifras de bloco x Cifra de fluxo

Segurança da Informação– GBC083

Relembrando...

Cifras de fluxo

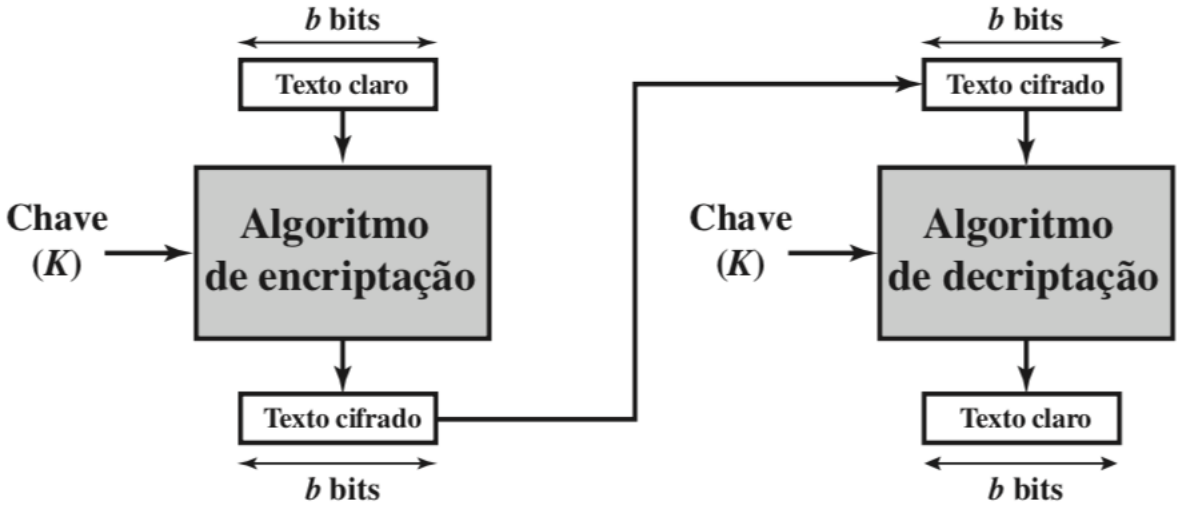
- Cifra um fluxo de dados (bit ou byte, por exemplo) por vez;
  - Cifra de Vigénere, por exemplo.



Relembrando...

Cifras de bloco

- Um bloco de texto claro é tratado como um todo e usado para produzir um de texto cifrado com o mesmo tamanho;
  - Blocos de 128 bits são comumente utilizados.



Cifras de bloco

- Tem sido destinado muito mais esforço para analisar as cifras de bloco;

- Em geral, elas são mais adequadas a uma gama maior de aplicações do que as cifras de fluxo.

Motivação para a cifra de Feistel

Segurança da Informação– GBC083

Cifras de bloco

- Uma cifra de bloco opera sobre um bloco de texto **claro** de *n bits* para produzir um bloco de texto **cifrado** de *n bits*;

- Existem quantos diferentes blocos de texto claro possíveis?

Cifras de bloco

- Uma cifra de bloco opera sobre um bloco de texto **claro** de *n bits* para produzir um bloco de texto **cifrado** de *n bits*;

- Existem quantos diferentes blocos de texto claro possíveis?
- R: 2^n

Transformações

- A transformação de um bloco de texto claro em texto cifrado deve ser reversível;

Mapeamento reversível	
Texto claro	**Texto cifrado**
00	11
01	10
10	00
11	01

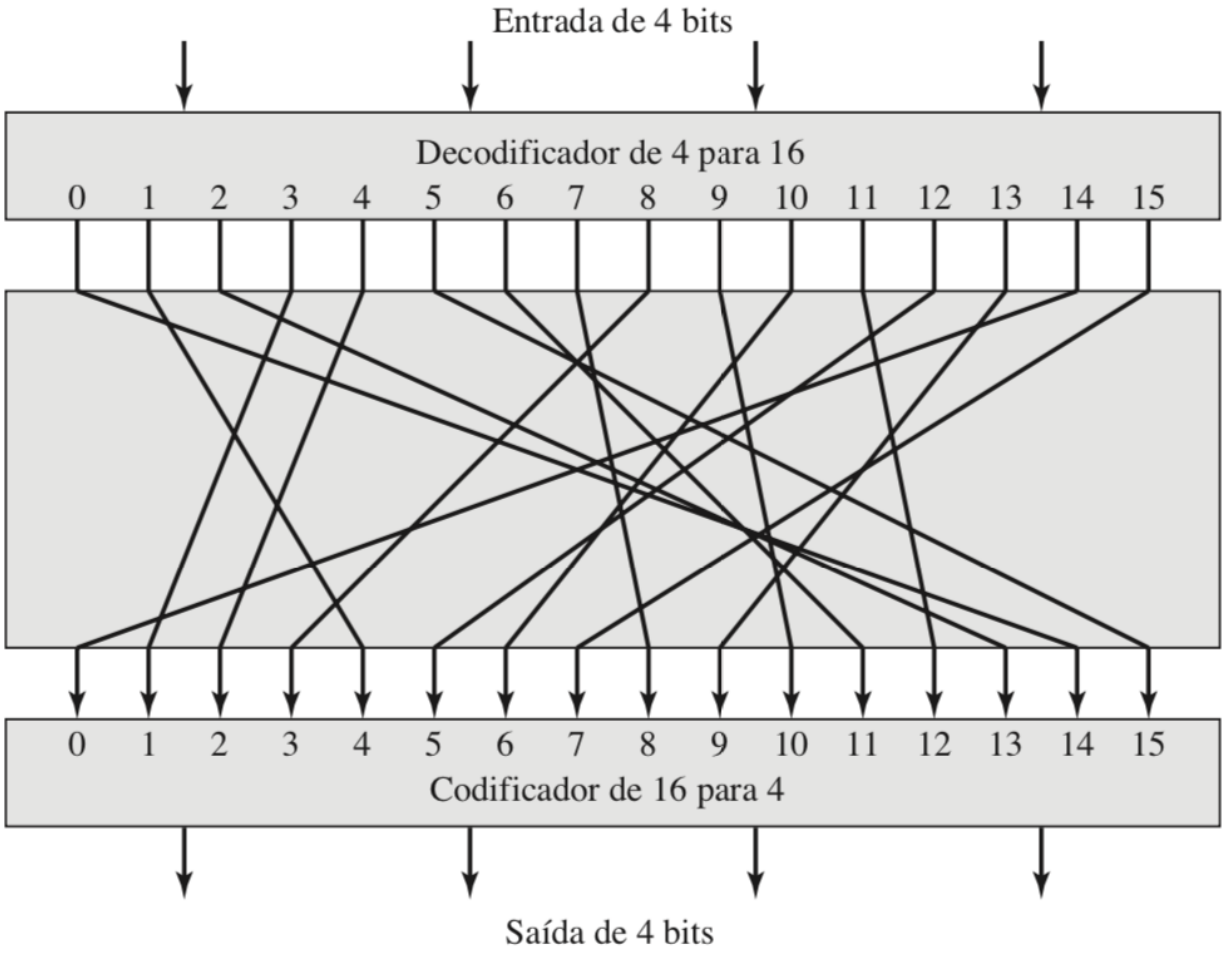
Mapeamento irreversível	
Texto claro	**Texto cifrado**
00	11
01	10
10	01
11	01

Cifra de substituição geral – n=4

Considere a seguinte cifra:

- Uma entrada de 4 bits produz um dos 16 estados de entrada possíveis (números de 0 a 15);
- Tais estados são mapeados pela cifra de substituição para um dos 16 estados de saída possíveis, cada um representado por 4 bits de texto cifrado.

Cifra de substituição geral – n=4



Cifra de substituição geral – n=4

Texto claro	Texto cifrado
0000	1110
0001	0100
0010	1101
0011	0001
0100	0010
0101	1111
0110	1011
0111	1000
1000	0011
1001	1010
1010	0110
1011	1100
1100	0101
1101	1001
1110	0000
1111	0111

Cifra de substituição geral – n=4

Feistel se refere a isso como cifra de bloco ideal, pois permite o número máximo de mapeamentos de cifragem a partir do bloco de texto claro.

Qual o problema prático aqui?

Cifra de substituição geral – Problemas...

- Suponha que tal cifra, com n=4, foi utilizada em um sistema real;

- Como o tamanho do bloco é pequeno (4 bits), tal cifra seria vulnerável a ataques de análise estatística do texto cifrado;
  - Características do texto claro não foram devidamente ocultadas pela cifra!

Cifra de substituição geral – Problemas...

- A ideia, então, seria aumentar o tamanho de n.
- Qual seria o tamanho da nossa chave para o caso com n=4?

Cifra de substituição geral – Problemas...

- A ideia, então, seria aumentar o tamanho de n!
- Qual seria o tamanho da nossa chave para o caso com n=4?
  - Basta multiplicar 4 (bits) x 16 (linhas) = 64 bits (literalmente o tamanho da cadeia de bits da tabela anterior);
- Como ficaria a conta para um tamanho de bloco n "razoável"?

Cifra de substituição geral – Problemas...

- A ideia, então, seria aumentar o tamanho de n!
- Qual seria o tamanho da nossa chave para o caso com n=4?
  - Basta multiplicar 4 (bits) x 2^4 (linhas) = 64 bits (literalmente o tamanho da cadeia de bits da tabela anterior);
- Como ficaria a conta para um tamanho de bloco n "razoável"?
  - Com n = 64 teríamos: 64 x 2^64 = 10^12 bits!!! Isso é prático?

Cifra de substituição geral – Problemas...

- Ou seja, é preciso fazer algum tipo de aproximação entre o modelo ideal e um modelo que funcione na prática...

- É justamente isso que veremos a seguir.

Cifra de Feistel

Segurança da Informação– GBC083

Cifra de Feistel

- Horst Feistel foi um funcionário da IBM na década de 70 que trabalhava com sistemas criptográficos;

- Usando a ideia da cifra de bloco ideal ele propôs o seguinte:
  - Podemos aproximar a cifra de bloco ideal utilizando o conceito de uma **cifra de produto**, que é a execução de **duas** ou **mais** cifras simples em sequência, de tal forma que o resultado seja criptograficamente mais forte do que qualquer uma das cifras componentes.

Cifra de Feistel

- A essência da técnica é desenvolver uma cifra de bloco com um tamanho de chave de ***k*** bits e de bloco de ***n*** bits;

- Feistel propôs o uso de uma cifra que alterna substituições e permutações:
  - **Substituição (confusão):** cada elemento de texto claro ou grupo de elementos é substituído exclusivamente por um elemento ou grupo de elementos de texto cifrado correspondente.
  - **Permutação (difusão):** uma sequência de elementos de texto claro é substituída por uma permutação dessa sequência.

Conceito de difusão e confusão

- Considere que o atacante tem algum conhecimento das características estatísticas do texto claro;

- Por exemplo, uma mensagem legível pode ter uma distribuição de frequência conhecida ou talvez haja palavras ou frases que provavelmente aparecem na linguagem.

Conceito de difusão e confusão

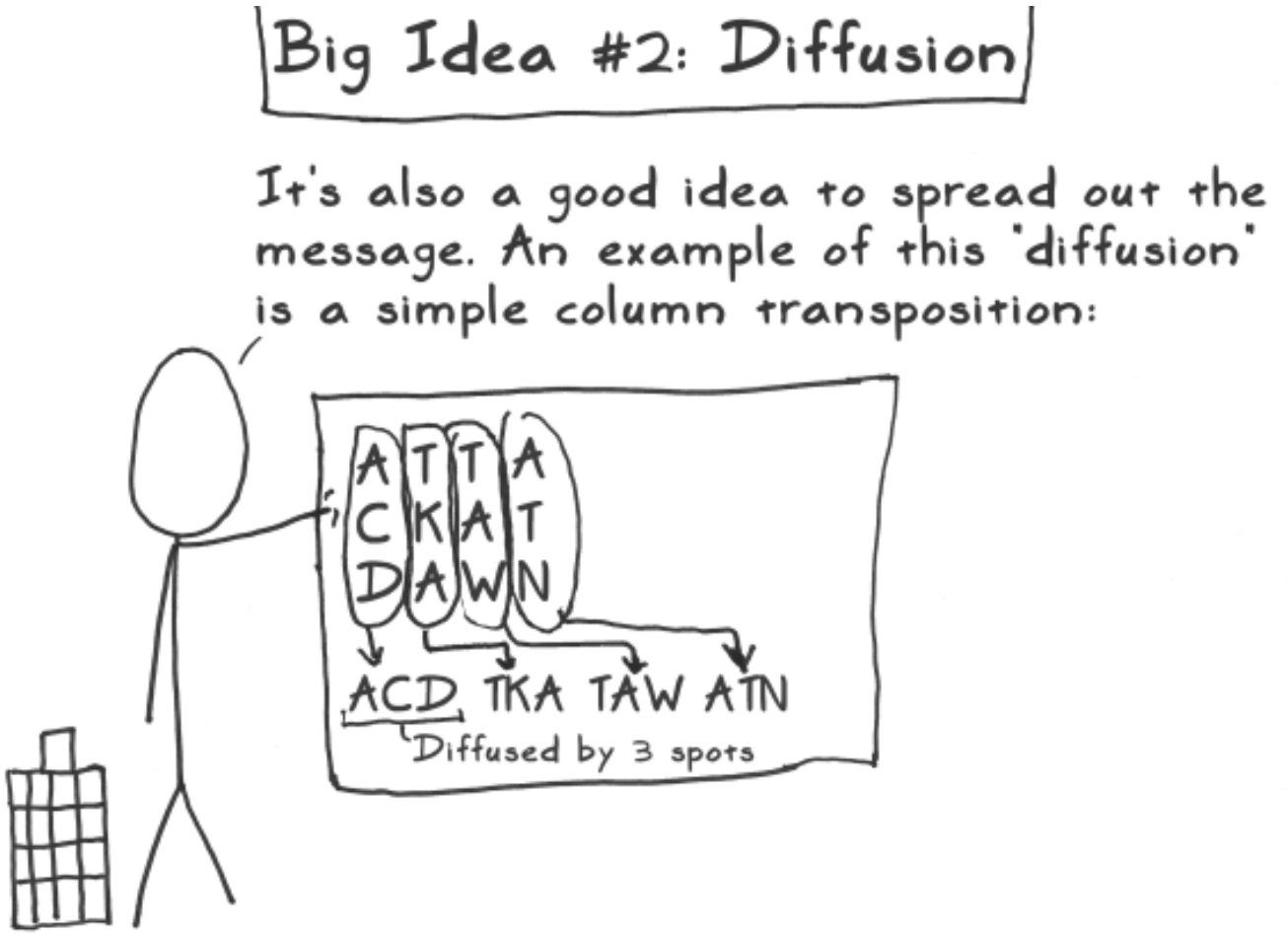
Um pesquisador chamado Claude Shannon estudou esse problema e sugeriu dois métodos para minimizar os problemas da criptoanálise estatística:

- **Difusão**
- **Confusão**

Difusão

- **Difusão** – modificar a estrutura estatística do texto claro;
  - Cada dígito do texto claro deve afetar o valor de muitos do texto cifrado;
  - Solução: aplicar diversas permutações seguida de alguma função.

Difusão



- http://www.moserware.com/2009/09/stick-figure-guide-to-advanced.html

Difusão - Exemplo

A ideia aqui é tentar ocultar relacionamentos estatísticos entre o texto claro e o texto cifrado.

Considere o texto claro M = “Só sei que nada sei”.

1. É possível notar algum padrão no texto?
2. Como aplicar a difusão?

Difusão - Exemplo

A ideia aqui é tentar ocultar relacionamentos estatísticos entre o texto claro e o texto cifrado.

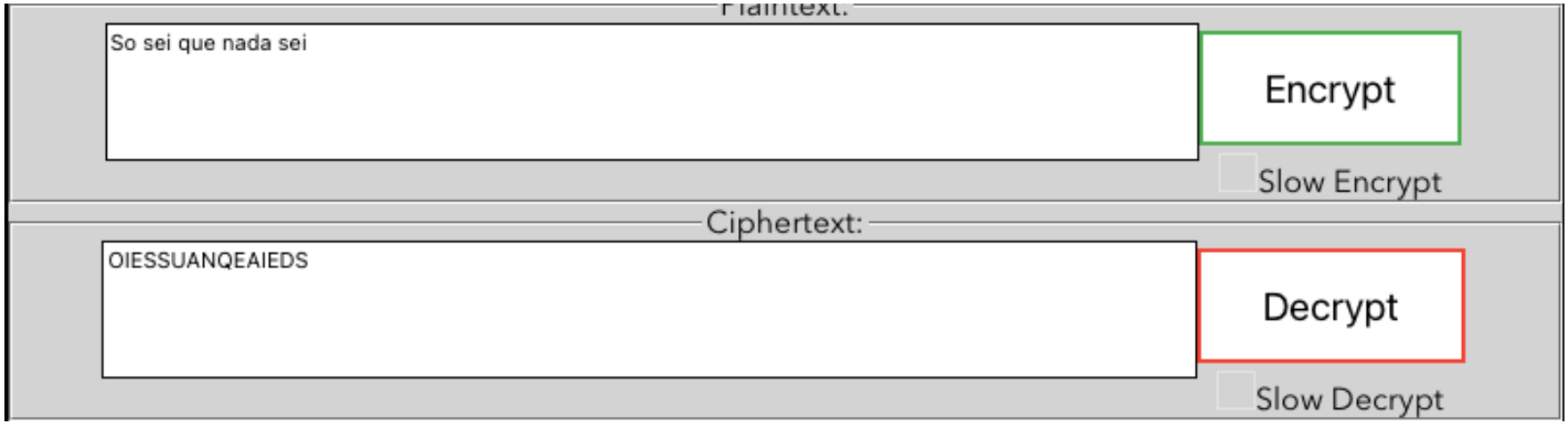
Considere o texto claro M = "Só sei que nada sei".

1. É possível notar algum padrão no texto?
2. Como aplicar a difusão?

Vamos usar um algoritmo simples de permutação e ver se esse padrão é mitigado.

https://crypto.interactive-maths.com/permutation-cipher.html

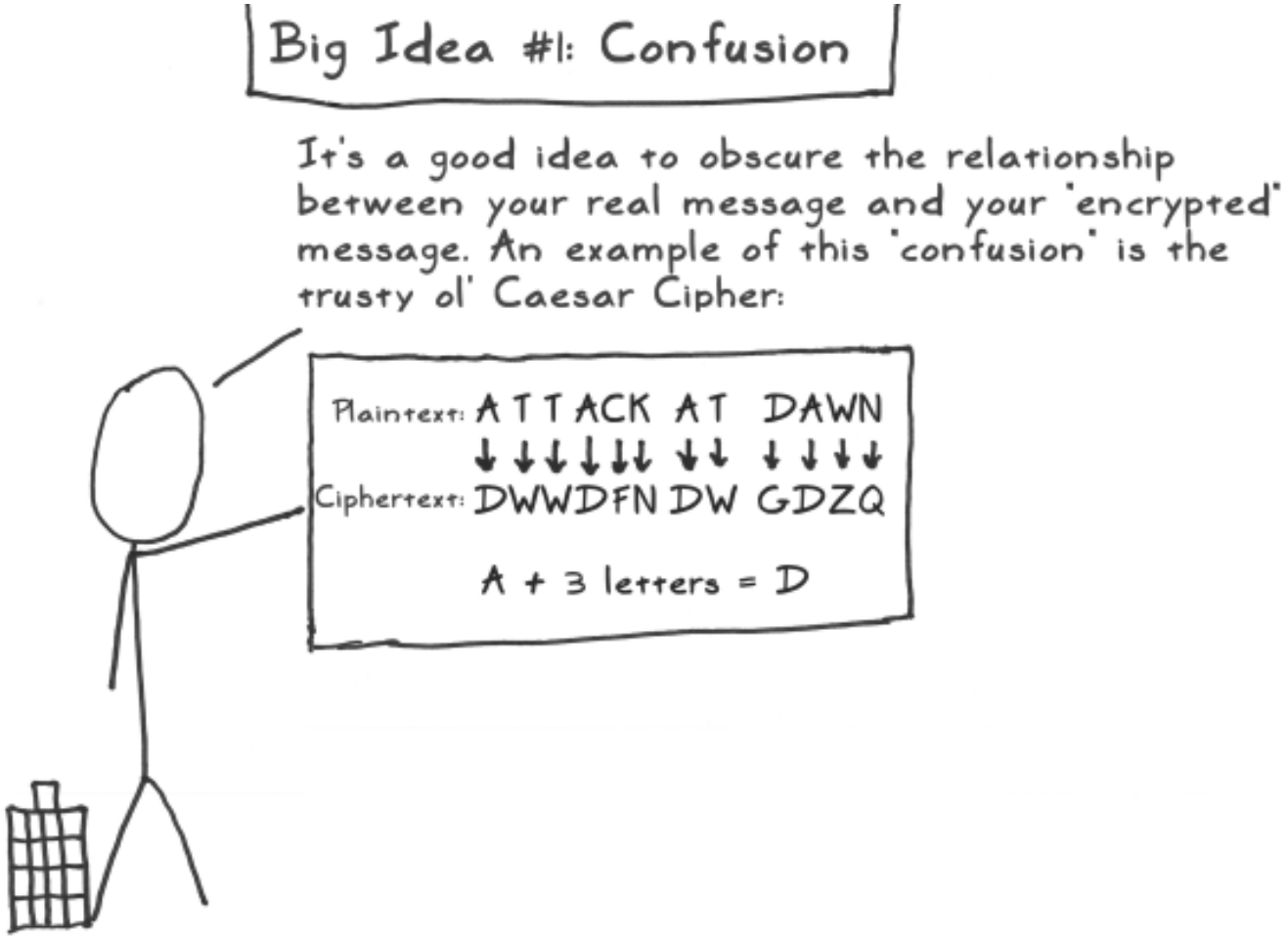
Difusão - Exemplo



Confusão

- **Confusão** – tornar o relacionamento entre as estatísticas do **texto cifrado** e da **chave** o mais complexo possível;
  - Ou seja, mesmo que o atacante possa ter alguma ideia das estatísticas do texto cifrado, o modo pelo qual a chave foi usada para produzir esse texto cifrado é tão complexo que torna difícil deduzi-la;
  - Portanto, deve ser difícil encontrar a chave à partir do texto cifrado;
  - Solução: aplicar diversas **substituições** no texto claro.

Confusão



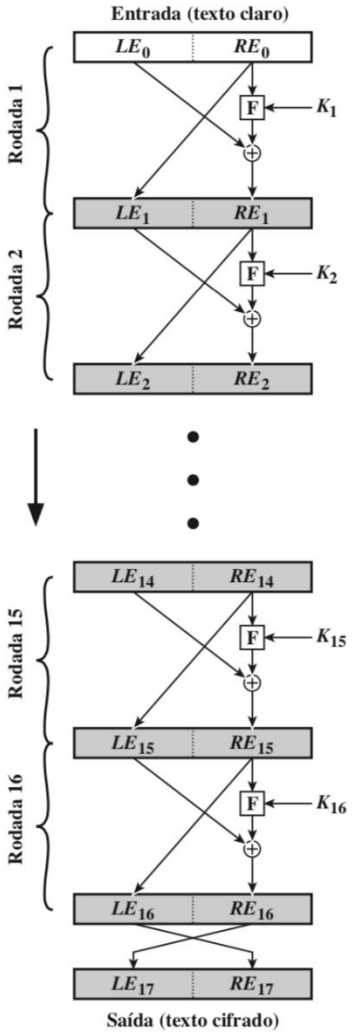
- http://www.moserware.com/2009/09/stick-figure-guide-to-advanced.html

Confusão - Exemplo

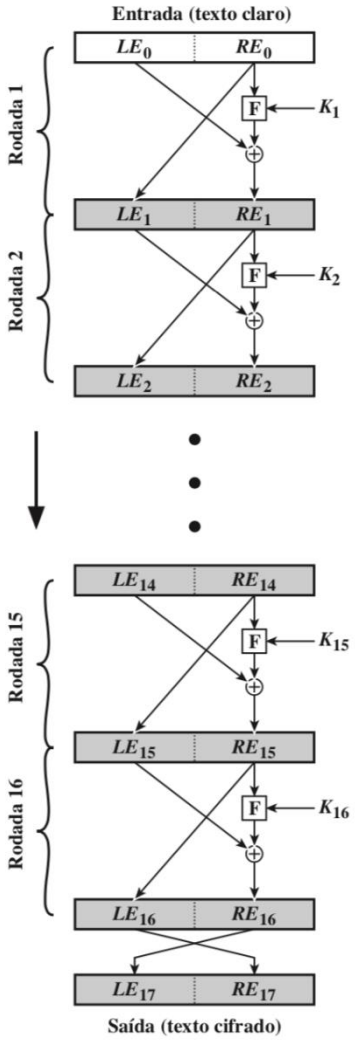
As **substituições** possuem esse papel. Considere uma Cifra de Vigenère.

1. Olhando somente para o texto cifrado, não é trivial encontrar a chave, mesmo com as limitações das cifras alfabéticas;
2. Por exemplo, cifrar o texto em claro M = "Só sei que nada sei" usando a chave K = "cripto", gera o texto cifrado C = "uf atb ewv vpwo uvq".
3. As cifras modernas usam funções de substituição (substitution boxes) que terão efeitos parecidos mas com um detalhes muito importante: *a alteração de um bit na chave afeta a muitos bits no texto cifrado*.

Cifra de feistel – Cifragem (*encryption*)

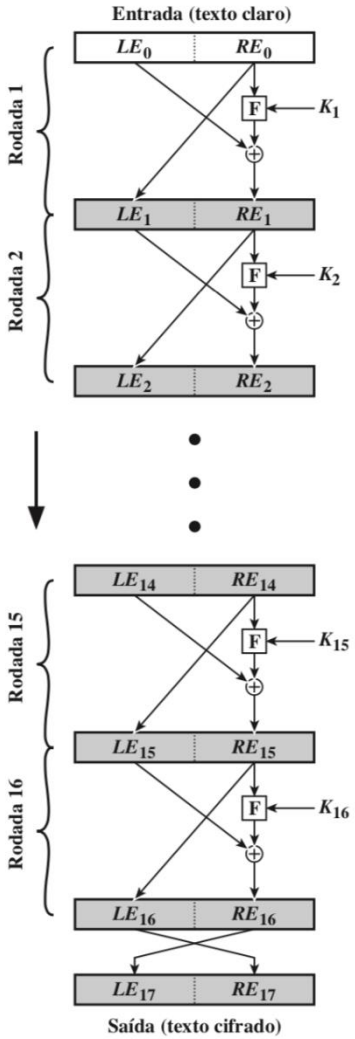


Cifra de feistel – Cifragem (*encryption*)



- Entradas do algoritmo:
  1. Um bloco de texto claro de tamanho 2w bits;
  2. Uma chave K.
- Funcionamento básico:
  - O bloco do texto claro é dividido em duas metades L_0 e R_0 (cada uma com w bits);
  - As duas metades dos dados passam por n rodadas de processamento e depois se combinam para produzir o bloco do texto cifrado;
  - Cada rodada i possui como entradas L_{i-1} e R_{i-1}, derivadas da rodada anterior, assim como uma subchave K_i derivada de K;

Cifra de feistel – Cifragem (*encryption*)



Perguntas:

- Onde está a substituição?
- Onde está a permutação?
- O algoritmo está aberto! Não seria possível fazer alguma engenharia reversa em um dado texto cifrado?

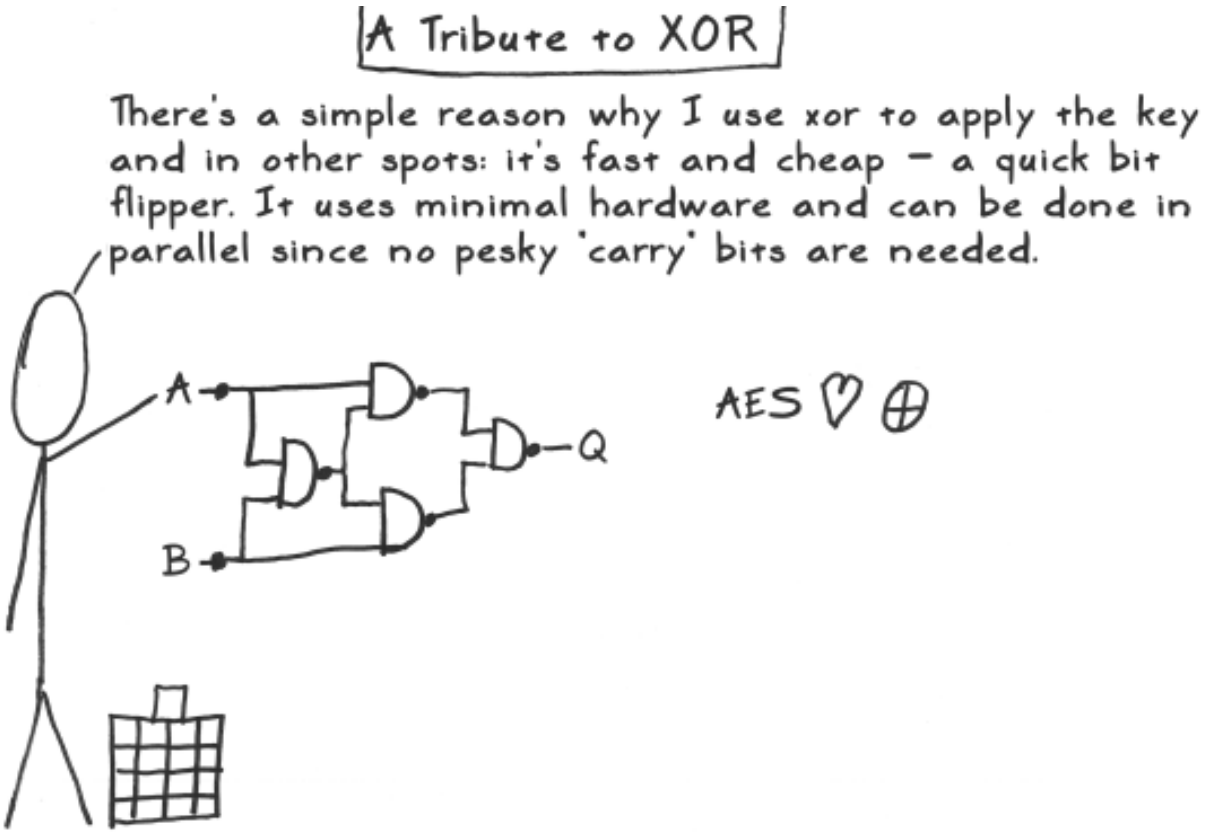
Rede clássica de Feistel – Rodada

Todas as rodadas têm a mesma estrutura:

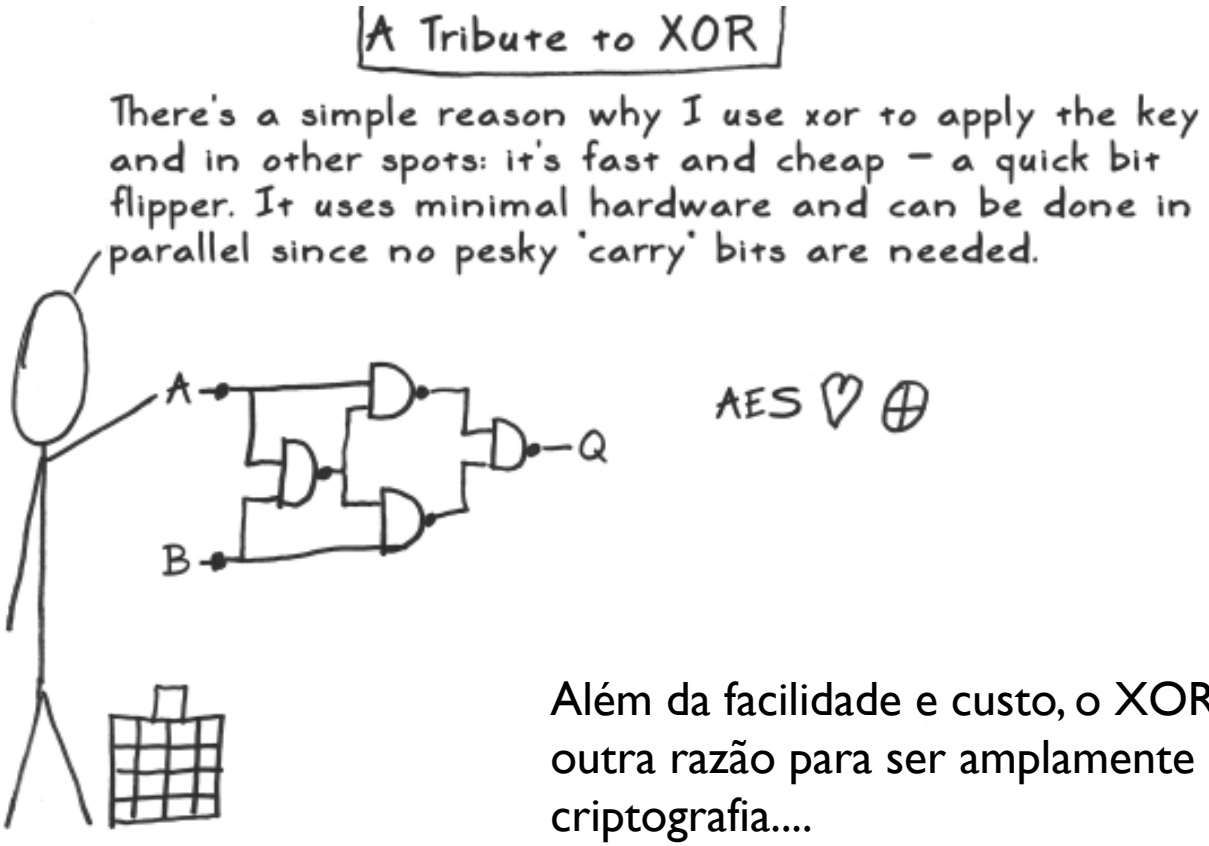
- Uma **substituição** é realizada na metade esquerda dos dados aplicando uma função F (rodada) à metade direita dos dados e depois aplicando **XOR** entre a saída dessa função e a metade esquerda dos dados.
- Depois **permuta-se** as duas metades dos dados.
- Pergunta: qual a razão para o uso da operação XOR?

A	B	A **XOR** B
0	0	0
0	1	1
1	0	1
1	1	0

Rede clássica de Feistel – Rodada



Rede clássica de Feistel – Rodada



Além da facilidade e custo, o XOR tem uma outra razão para ser amplamente utilizado em criptografia....

Rede clássica de Feistel – Rodada

Quais as semelhanças/diferenças entre as operações?

AND		
A	B	Result
0	0	0
0	1	0
1	0	0
1	1	1

OR		
A	B	Result
0	0	0
0	1	1
1	0	1
1	1	1

XOR		
A	B	Result
0	0	0
0	1	1
1	0	1
1	1	0

Rede clássica de Feistel – Rodada

Quais as semelhanças/diferenças entre as operações?

AND		
A	B	Result
0	0	0
0	1	0
1	0	0
1	1	1

OR		
A	B	Result
0	0	0
0	1	1
1	0	1
1	1	1

XOR		
A	B	Result
0	0	0
0	1	1
1	0	1
1	1	0

Ao usar o XOR eu sempre terei 50% de chance de "flipar" um bit...

Isso está diretamente relacionado a **perda de informação**... Considere **A = 1010**, **B = 1100**. Se eu fizer A OR B = C eu consigo recuperar o A sabendo somente C e B? O que acontece se eu fizer o XOR ao invés do OR?

Rede clássica de Feistel – Parâmetros

1. **Tamanho de bloco** – maior → mais seguro e maior overhead computacional;
2. **Tamanho de chave** – maior → mais seguro e maior overhead computacional;
3. **Número de rodadas** – 16 rodadas é o tamanho típico;
4. **Algoritmo de geração de subchave** – maior complexidade → maior segurança;
5. **Função F** – maior complexidade → maior segurança.

Rede clássica de Feistel – Exercício

- Como funciona o processo de decifragem em uma rede de Feistel? Qual deve ser o texto de entrada? O que muda quando comparado ao processo de cifragem?

- Estudar páginas 53 e 54 do livro do Stallings.

Próximas aulas

- Exemplo de um algoritmo baseado na cifra/rede de Feistel – DES.

Roteiro de estudos

1. Leitura da seção 3.1 do livro "Criptografia e segurança de redes. Princípios e práticas". William Stallings;
2. Estudo da vídeo-aula referente ao tópico 5;