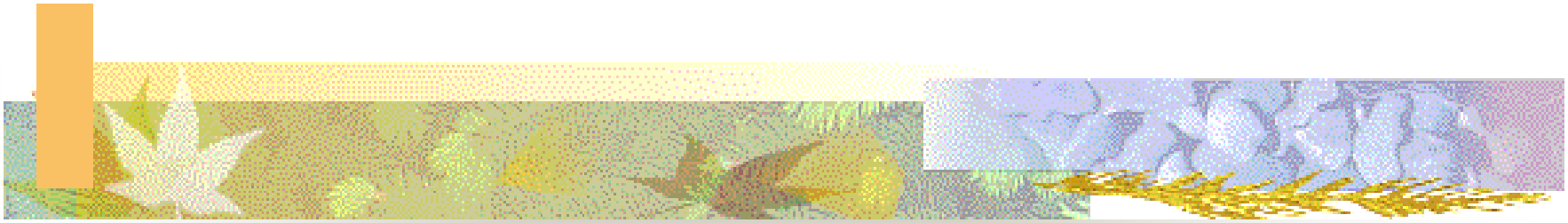
# Introdução à Criptografia



Profa. Yeda

## Aula 1 – Introdução aos Termos e Algoritmos

(**Cap. 2 Stallings**)

# Criptografia - Terminologia

- Texto claro: a mensagem a ser enviada é denominada texto claro. Se for interceptada em uma comunicação poderá ser compreendida pelo interceptador.
- Criptografia: O processo de transformação de uma mensagem que possa ser compreendida em outra não compreensível.
- Texto cifrado: Uma mensagem criptografada é chamada de texto cifrado.

# Algoritmos e Chaves

- Um algoritmo criptográfico é uma função matemática utilizada para a cifragem e decifragem.
- Se a segurança de um algoritmo está baseada na necessidade de se manter o algoritmo secreto, então o algoritmo é dito **restrito**.
- Na criptografia moderna, a *segurança não está na confidencialidade do algoritmo* criptográfico, mas na **chave**
  - O conjunto de todas as possíveis chaves é denominado **espaço de chaves**.

# Sistema Criptográfico

- O conjunto formado pelo **algoritmo** e **chaves** criptográficas é denominado sistema criptográfico.
- Sistemas criptográficos se caracterizam por:
  - Tipo de operação de cifragem usada:
    - **substituição / transposição / produto**
  - Número de chaves usadas
    - Simples ou **Privada** / Dupla ou **Pública**
  - Forma como o texto claro é processado
    - **Bloco** ou em **Fluxo** (Stream)

# Força de um Sistema Criptográfico

- A força de um sistema criptográfico, ou seja, a sua resistência a ataques é função de vários fatores:
  - A confidencialidade da chave.
  - A dificuldade da determinação da chave através da sua adivinhação ou da tentativa de todas as possíveis chaves.
  - A dificuldade em se inverter o algoritmo criptográfico sem o conhecimento da chave criptográfica.

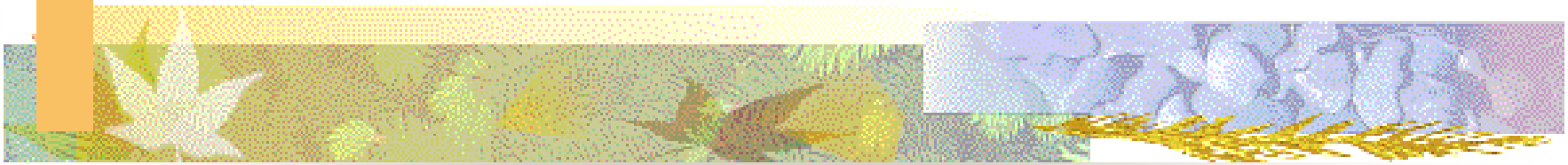
# Força de um Sistema Criptográfico

- A força de um sistema criptográfico (cont...)
  - A inexistência (ou existência) de "portas dos fundos" no sistema, ou outras formas que permitam que um texto cifrado possa ser decodificado de modo simples.
  - A possibilidade de se decodificar todo um texto cifrado dado que se saiba como parte dele é decodificada.
  - O conhecimento de propriedades peculiares da mensagem em texto claro que possam ser utilizadas para sua determinação.

# Força de um Sistema Criptográfico

- O objetivo no projeto de um sistema criptográfico:
  - desenvolver um algoritmo que torne muito difícil a reversão do processo de criptografia sem o conhecimento da chave.
  - A dificuldade deve ser no mínimo equivalente ao trabalho requerido para se encontrar a chave criptográfica tentando-se todas as possíveis soluções.

# Técnicas de Algoritmos Simétricos



## Algoritmos Clássicos:
## Substituição e Transposição

# Cifras por Substituição

- Cifras onde letras do texto claro são substituídas por outras letras, números ou símbolos.
- Ou, substitui-se um padrão de bits do texto claro por um padrão de bits de texto cifrado.
- Exemplos:
  - Cifra de Cesar

# Cifra de Cesar

- Matematicamente, dá-se a cada letra um número

| 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 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 | 21 | 22 | 23 | 24 | 25 |

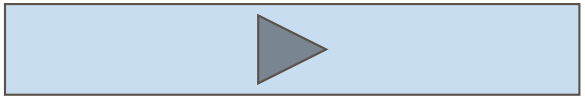
- então, define-se a cifra de Cesar como:

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

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

# Criptoanálise da Cifra de Cesar

- tem somente 26 cifras possíveis: mapas A ... Z
- poderia simplesmente tentar cada um ...
- uma busca por **força bruta**
- dado um texto cifrado, tente todos os deslocamentos possíveis
- é necessário reconhecer quando se tem o texto claro
- Exemplo, quebre o texto cifrado
  - `"PHHW PH DIWHU WKH WRJD SDUWB"`



# SUGESTÃO PARA MELHORAR?

# Cifras Monoalfabéticas

- ao invés de apenas deslocar o alfabeto, poderia trocar as letras arbitrariamente
- assim a chave seria uma sequência de 26 letras

```
Plain:  abcdefghijklmnopqrstuvwxyz
Cipher: DKVQFIBJWPESCXHTMYAUOLRGZN

Plaintext:  ifwewishtoreplaceletters
Ciphertext: WIRFRWAJUHYFTSDVFSFUUFYA
```

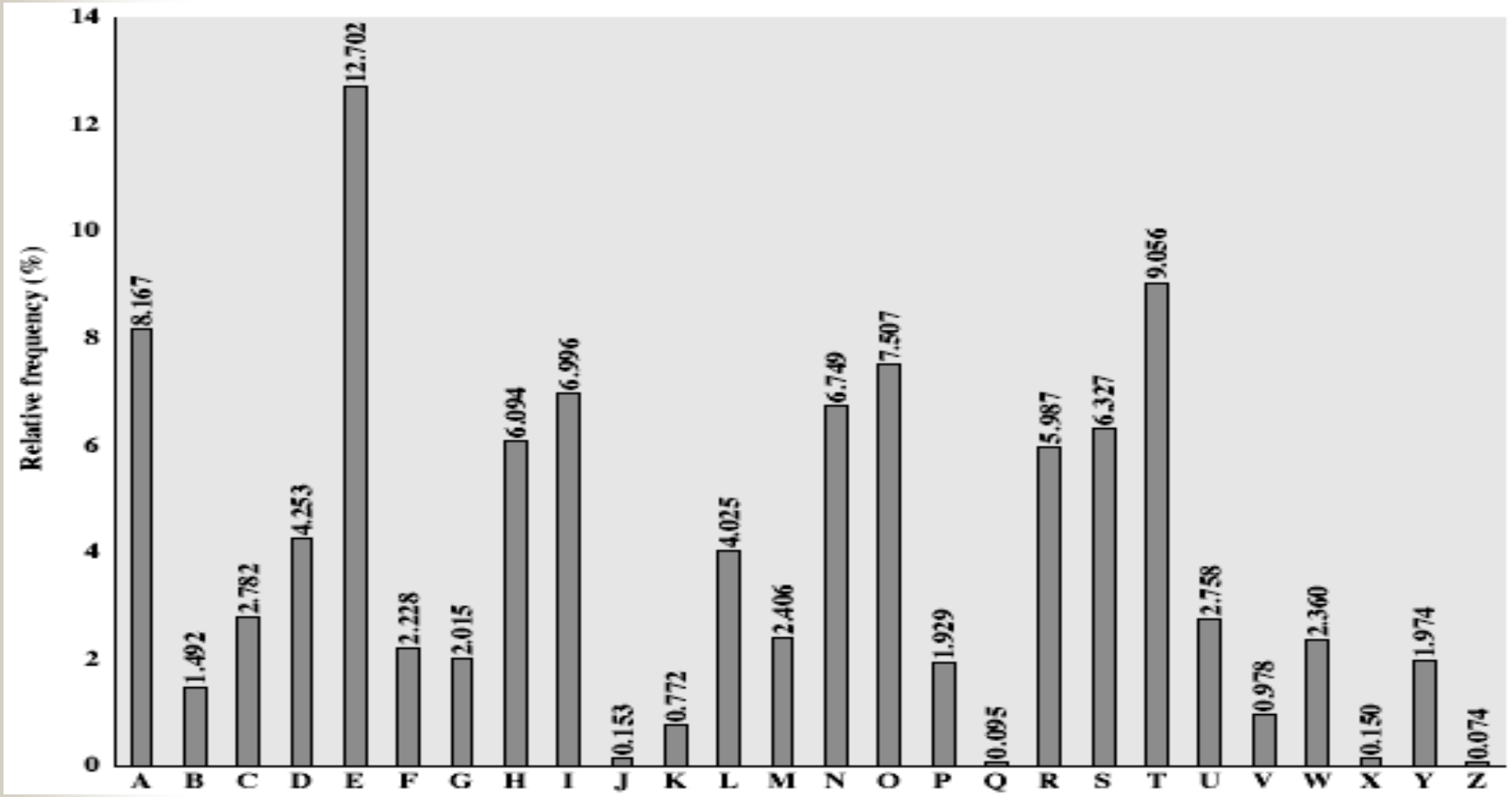
# Segurança de Cifras Monoalfabéticas

- agora temos um total de $26! = 4 \times 10^{26}$ chaves
- com tantas chaves, poderia-se pensar que é seguro?
- **!!!ERRADO!!!**
- o problema está na característica da linguagem.

# Redundância e Criptoanálise da Linguagem

- linguagem humana é **redundante**
- por exemplo, "*td q prcs m prt d txt*" (complete!)
- letras não são usadas com mesma frequência
- em inglês E é de longe a mais utilizada
  - seguida por T,R,N,I,O,A,S
- outras letras como Z,J,K,Q,X são raras
- há tabelas para frequências de letras simples, duplas e triplas para vários idiomas.

# English Letter Frequencies



# Uso em Criptoanálise

- Conceito chave – cifradores de substituição monoalfabética não mudam a frequência relativa das letras
- Foi descoberto por um cientista no século IX
- Calcula-se a frequência das letras para o texto cifrado e compara-se com as frequências conhecidas
- Por exemplo, se olhássemos um texto inglês, cifrado com a cifra de Cesar, veríamos:
  - picos em: simples A-E-T, dupla NO, tripla RST
  - baixa em: JK, X-Z

# Exemplo de Criptoanálise

- Dado o texto cifrado:

  UZQSOVUOHXMOPVGPOZPEVSGZWSZOPFPESXUDBMETSXAIZ
  VUEPHZHMDZSHZOWSFPAPPDTSVPQUZWYMXUZUHSX
  EPYEPOPDZSZUFPOMBZWPFUPZHMDJUDTMOHMQ

- conte a frequência relativa das letras

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| 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 | 2,5 | I | 0,83 | N | 0,00 |
| O | 7,50 | X | 4,17 | A | 1,67 | J | 0,83 | R | 0,00 |
| M | 6,67 | | | | | | | | |

# Exemplo de Criptoanálise

- Dado o texto cifrado:

```
UZQSOVUOHXMOPVGPOZPEVSGZWSZOPFPESXUDBMETSXAIZ
 t a         e  e te  a that e e a       a
VUEPHZHMDZSHZOWSFPAPPDTSVPQUZWYMXUZUHSX
  e t   ta t ha e ee  a e  th    t  a
EPYEPOPDZSZUFPOMBZWPFUPZHMDJUDTMOHMQ
 e  e e tat  e   the   t
```

- suponha que ***P*** & ***Z*** são ***e*** e ***t*** (maior freq.)
- suponha que ***S,U,O,M*** e ***H*** são ***a,h,i,n,o,r*** e ***s***
- suponha que ***ZW*** é ***th*** e assim ***ZWP*** é ***the***

# Exemplo de Criptoanálise

- siga com a tentativa e erro e obtenha:

```
it was disclosed yesterday that several informal but
direct contacts have been made with political
representatives of the viet cong in moscow
```

# SUGESTÃO PARA MELHORAR?

# Cifrador Playfair

- um grande número de chaves não foi suficiente para fornecer segurança para cifradores monoalfabéticos
- uma abordagem para melhorar a segurança foi criptografar múltiplas letras
- o **Playfair Cipher** é um exemplo
- inventado por Charles Wheatstone em 1854

# Matriz de Chave Playfair

- Uma matriz de letras 5X5 baseada em palavra chave
- Preencha as letras da palavra chave (sem duplicatas)
- Preencha o restante da matriz com as outras letras
- ex. Usando a palavra chave MONARCHY

| | | | | |
|---|---|---|---|---|
| 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 |

# Cripto e Decriptografia

| | | | | |
|---|---|---|---|---|
| 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 |

- O texto plano é cifrado duas letras por vez

1. se um par é uma letra repetida, inserir um preenchedor, tal como ‘x’

Ex.: balloon ⟹ ba lx lo on

2. se ambas as letras estão na mesma linha da tabela, troque cada qual com a letra a direita (linha cíclica)

Ex.: ar (*plano*) ⟹ rm (*cifrado*)

# Cripto e Decriptografia

| | | | | |
|---|---|---|---|---|
| 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 |

- O texto plano é cifrado duas letras por vez

3. se ambas as letras estão na mesma coluna, troca cada qual com a letra abaixo dela (coluna cíclica)

Ex.: mu (*plano*) ⟹ cm (*cifrado*)

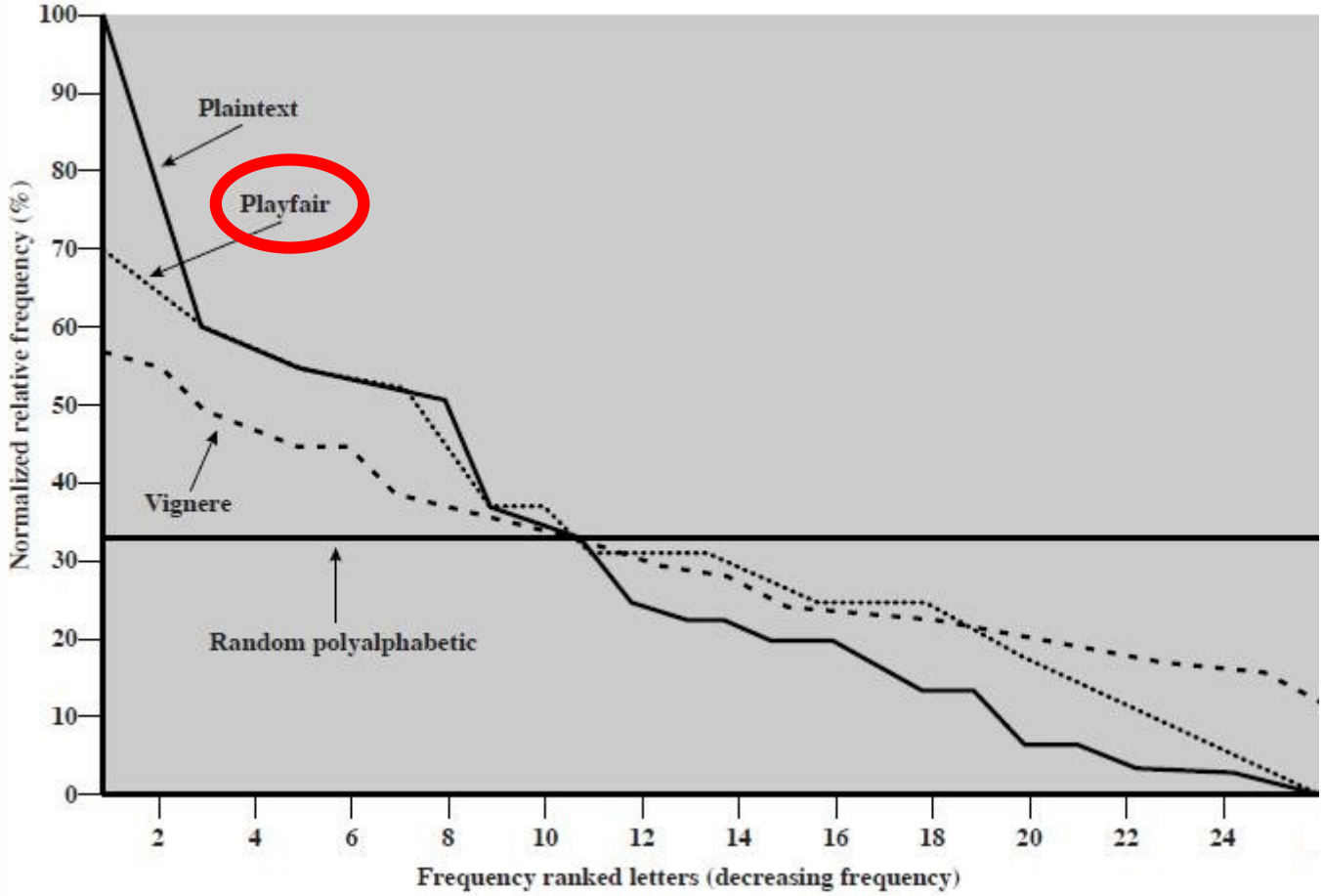
4. caso contrário, cada letra é trocada pela letra na mesma linha e coluna da outra letra do par.

Ex.: hs (*plano*) ⟹ bp (cifrado)

# Segurança do Cifrador Playfair

- Grande melhoria na segurança em relação às cifras monoalfabéticas,
- uma vez que tem 26 x 26 = 676 digramas.
- Necessitaria de 676 tabelas de frequências de entradas para analisar (versos 26 para uma monoalfabética) ,
- assim como mais textos cifrados.
- Foi utilizada por muitos anos
  - Ex.: pelos militares US & Britânicos na 2ª guerra
- Pode ser quebrado, dado algumas centenas de letras, pois ainda mantém certa estrutura do texto plano.

# Frequência Ordenada de Letras



# Cifradores Polialfabéticos

- **Cifradores polialfabéticos de substituição**
- melhora a segurança usando múltiplos cifradores alfabéticos
- torna a criptoanálise mais difícil com mais alfabetos para desvendar e uma distribuição de frequência mais plana,
- usa uma chave para selecionar qual alfabeto é usado para cada letra da mensagem,
- repete do início após o fim da chave ser atingido.

# Cifradores de Vigenère

- Cifrador de substituição polialfabético mais simples.
- efetivamente um cifrador de César múltiplo,
- chave com tamanho de múltiplas letras
  - $K = k_1\, k_2 \ldots k_d$
- $i^{th}$ letra especifica o $i^{th}$ alfabeto a usar,
- repete do início após *d* letras da mensagem
- decriptografia faz o reverso.

# Exemplo de Cifra de Vigenère

- escreva o texto plano,
- escreva a palavra chave repetidas vezes sobre ele,
- use cada letra da chave como chave para o cifrador de Cezar,
- criptografe a correspondente letra do texto plano,
- ex.: usando a *deceptive* palavra chave

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

# Segurança da Cifra de Vigenère

- tem múltiplas letras de texto cifrado para cada letra de texto plano,
- assim a frequência das letras são obscurecidas,
- mas não totalmente perdidas,
- inicie com a frequência das letras,
  - veja se parece com monoalfabético ou não.
- se não, então precisa determinar o número de alfabetos, assim pode atacar cada um deles.

# Método de Kasiski

- Método desenvolvido por Babbage/Kasiski:
- repetições em texto cifrado dá a dica para o período.
- Ache para um mesmo texto plano um período exato,
- obtenha o mesmo resultado no texto cifrado,
- é claro, poderia também ser um acaso aleatório,
- Ex. repetidos “VTW” no exemplo anterior sugere tamanho de 3 ou 9,
- então ataque cada cifrador monoalfabético individualmente usando a mesma técnica anterior.

# Método de Kasiski

- Texto plano e cifrado:
  - WEAREDISCOVEREDSAVEYOURSELF
  - ZICVTWQNGRZGVTWAVZHCQYGLMGJ

- distância de 9 letras, então a chave provavelmente tem tamanho 3 ou 9.

# Cifrador de Autokey

- Deseja-se uma chave tão longa quanto a mensagem.
- Vigenère propôs o cifrador de auto-chave,
- a palavra-chave é prefixada à mensagem como chave,
- conhecendo a palavra-chave pode-se recuperar as primeiras poucas letras,
- usa-se estas letras para obter o restante da mensagem.

# Cifrador de Autokey

- Ex. dado a chave *deceptive*
  - key: *deceptivewearediscoveredsav*
  - plaintext: wearediscoveredsaveyourself
  - ciphertext:

ZICVTWQNGKZEIIGASXSTSLVVWLA

- Mas ainda tem características para o ataque de frequência, pois a frequência das letras no texto plano se repetirá no alfabeto sendo usado para cifrar.

# One-Time Pad

- "*Se uma chave totalmente aleatória e do tamanho da mensagem for usada, o cifrador será seguro.*"
- Essa técnica é chamada One-Time pad.
- É inquebrável, uma vez que o texto cifrado não produz uma relação estatística com o texto plano.
- Assim, para **qualquer texto plano** & **qualquer texto cifrado** existe um mapeamento de chave para o outro.
- Mas a chave pode ser utilizada uma **única vez**.

# One-Time Pad

## Que problema vocês percebem nisso?

- Problema na geração e segura distribuição de chaves

# Cifradores de Transposição

- Agora considere os cifradores clássicos de **transposição** ou **permutação.**
- Estes escondem a mensagem por rearranjar a ordem das letras.
  - A ideia é rearranjar a ordem das unidades básicas (letras/bytes/símbolos) sem alterar seus valores atuais.

# Cifrador de Rail Fence

- Escreva as letras da mensagem diagonalmente sobre um número de linhas,
- então leia do cifrador linha por linha
- Ex.: "meet me after the toga party“, profundidade 2

```
m e m a t r h t g p r y
  e t e f e t e o a a t
```

- resultando no texto cifrado

```
MEMATRHTGPRYETEFETEOAAT
```

# Cifrador de Linha de Transposição

- Uma transposição mais complexa.
- Escreve letras da mensagem em linhas sobre um número especificado de colunas,
- então reordene as colunas de acordo com alguma chave antes de permutar as colunas

```
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
Ciphertext: TTNAAPTMTSUOAODWCOIXKNLYPETZ
```

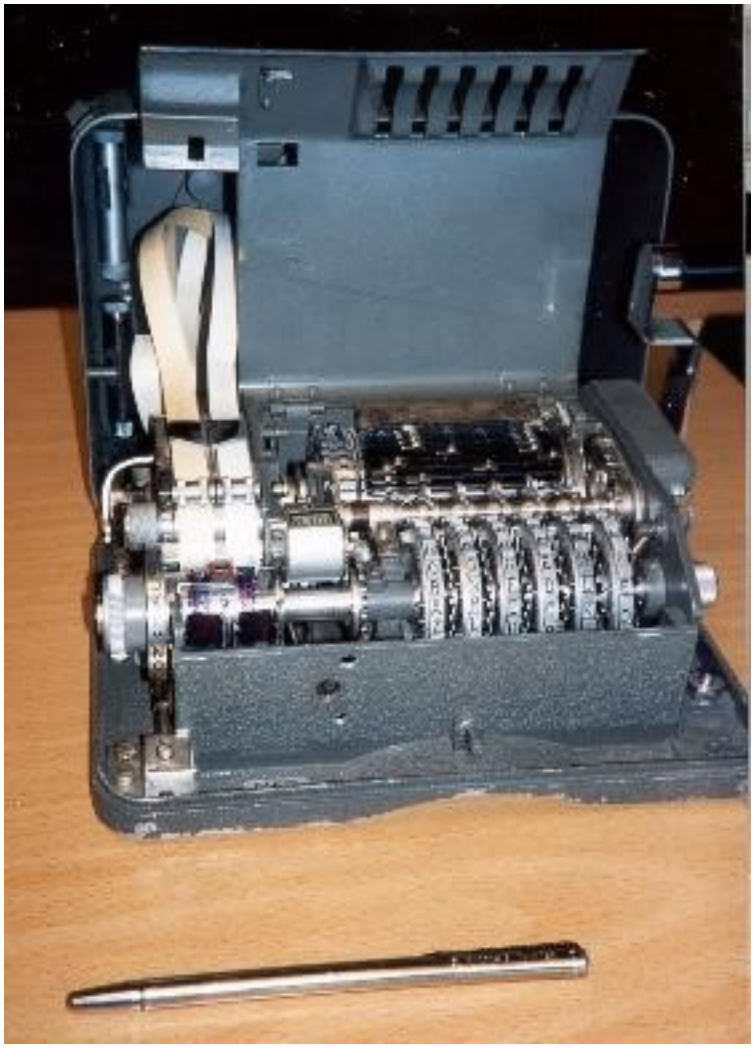
# Cifradores de Produto

- cifradores de substituição e transposição não são suficientemente seguros devido às características da linguagem,
- assim considere o uso de vários cifradores em sucessão para torná-lo mais forte:
  - duas substituições torna a substituição mais complexa,
  - duas transposições torna a transposição mais complexa,
  - mas uma substituição seguida por uma transposição torna o novo cifrador muito mais forte.
- esta é a ponte dos modernos cifradores.

# Máquina de Rotores

- Antes dos cifradores modernos, máquina de rotores foram os cifradores mais complexos em uso,
- foram amplamente usados na 2a guerra
  - **Enigma Alemão**, Allied Hagelin, Japanese Purple
- Usava uma série de cilindros, cada um fazendo uma substituição, o qual rotacionava e mudava após cada letra ser cifrada,
- com 3 cilindros tem $26^3=17576$ alfabetos

# Hagelin Rotor Machine



# Steganografia

- Uma criptografia alternativa
- Esconde a existência da mensagem
  - usando somente um subconjunto de letras/palavras em uma mensagem mais longa marcada de alguma forma,
  - usando tinta invisível,
  - escondendo no bit LSB em imagens gráficas ou arquivo de som.
- Contras: alta sobrecarga para esconder poucos bits

# Resumo

- Terminologia e técnicas de cifradores clássicos,
- Cifradores monoalfabéticos de substituição,
- Criptoanálise usando frequência das letras,
- Cifrador Playfair,
- Cifradores polialfabéticos,
- Cifradores de transposição,
- Cifradores de produto e máquina de rotores,
- Esteganografia.