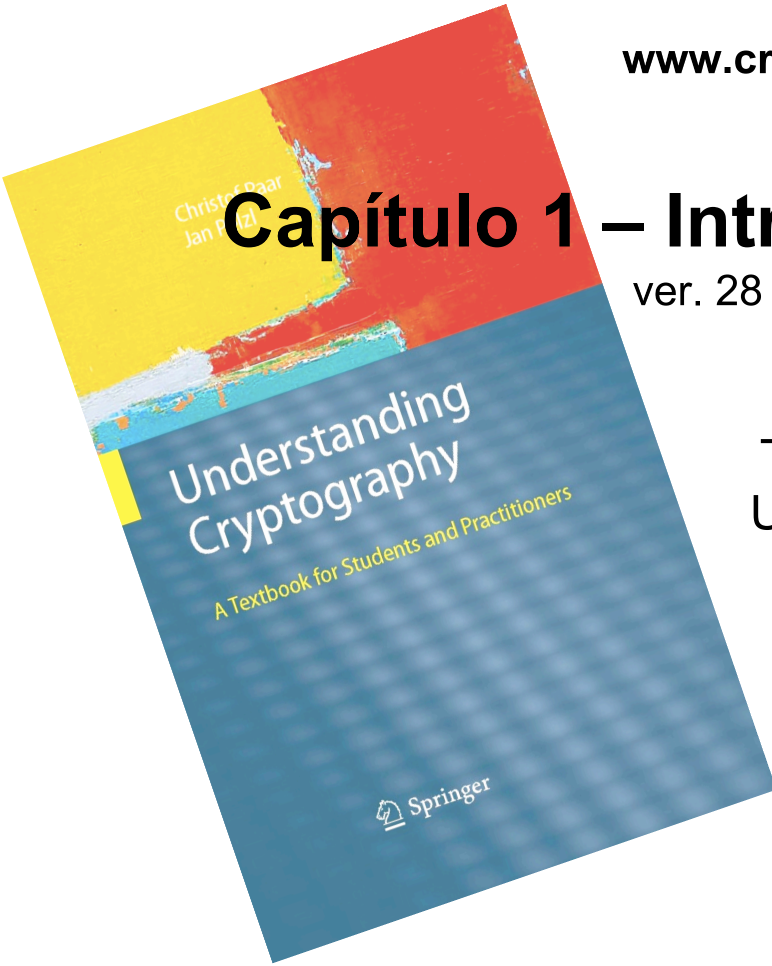
Entendendo Criptografia – Um Livro Texto para Estudantes e Profissionais

por Christof Paar e Jan Pelzl

www.crypto-textbook.com



Capítulo 1 – Introdução a Criptografia

ver. 28 de outubro de 2009

Tradução para Português (Brasil) dos slides:
Understanding Cryptography – A Textbook for Students and Practitioners
by Christof Paar and Jan Pelzl.
Chapter 1 – Introduction to Cryptography.
ver. October 28, 2009

Estes slides foram preparados em inglês por Christof Paar e Jan Pelzl, e traduzidos para o Português por Luiz C. Navarro, Emmanuel F. L. Silva e Ricardo Dahab

Algumas questões legais (desculpem): Condições para uso deste material

- Os slides podem ser usados sem custos. Todos os direitos dos slides permanecem com os autores.
- O título do livro que dá origem aos slides "Understanding Cryptography by Springer" e o nome dos autores devem permanecer em todos os slides.
- Se os slides forem modificados, os créditos aos autores do livro e ao livro devem permanecer nos slides.
- Não é permitida a reprodução de parte ou do todo dos slides em forma impressa sem a permissão expressa por escrito dos autores.

Conteúdo deste Capítulo

- Visão geral do campo da Criptologia
- Conceitos Básicos da Criptografia Simétrica
- Criptoanálise
- Cifra de Substituição
- Aritmética Modular
- Cifra de Deslocamento (ou de César) e Cifra Afim

Conteúdo deste Capítulo

- **Visão geral do campo da Criptologia** A seguir
- Conceitos Básicos da Criptografia Simétrica
- Criptoanálise
- Cifra de Substituição
- Aritmética Modular
- Cifra de Deslocamento (ou de César) e Cifra Afim

Leituras e Informações Adicionais

Leitura adicional ao *Understanding Cryptography*.

- A.Menezes, P. van Oorschot, S. Vanstone, *Handbook of Applied Cryptography*. CRC Press, October 1996.
- H.v.Tilborg (ed.), *Encyclopedia of Cryptography and Security*, Springer, 2005

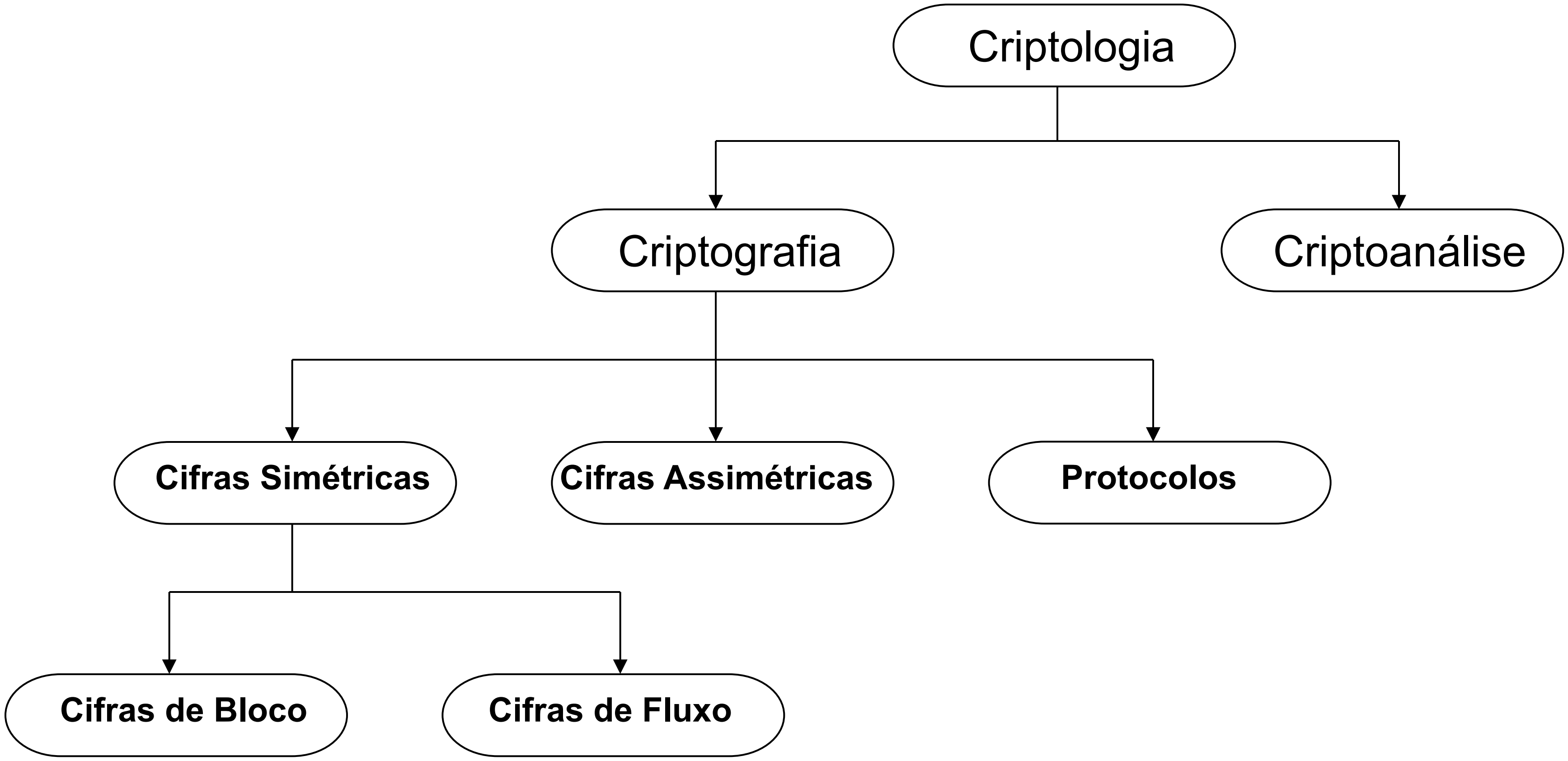
História da Criptografia (bons livros de cabeceira)

- S. Singh, *The Code Book: The Science of Secrecy from Ancient Egypt to Quantum Cryptography,* Anchor, 2000.
- D. Kahn, *The Codebreakers: The Comprehensive History of Secret Communication from Ancient Times to the Internet*. 2nd edition, Scribner, 1996.

Software (demonstração excelente de cifras antigas e modernas)

- *Cryptool, http://www.cryptool.de*

Classificação dos Campos da Criptologia



Alguns Fatos Básicos

- **Criptografia Antiga:** Os primeiros sinais de encriptação são do Egito, de cerca de 2000 A.C.
  Desde então, esquemas de encriptação baseados em letras (p.e., Cifra de César) tornaram-se populares.

- **Cifras Simétricas:** Todos os esquemas de encriptação desde a antiguidade até 1976 eram simétricos.

- **Cifras Assimétricas:** Em 1976, Diffie, Hellman e Merkle propuseram (publicamente) a criptografia de chave pública (ou assimétrica).

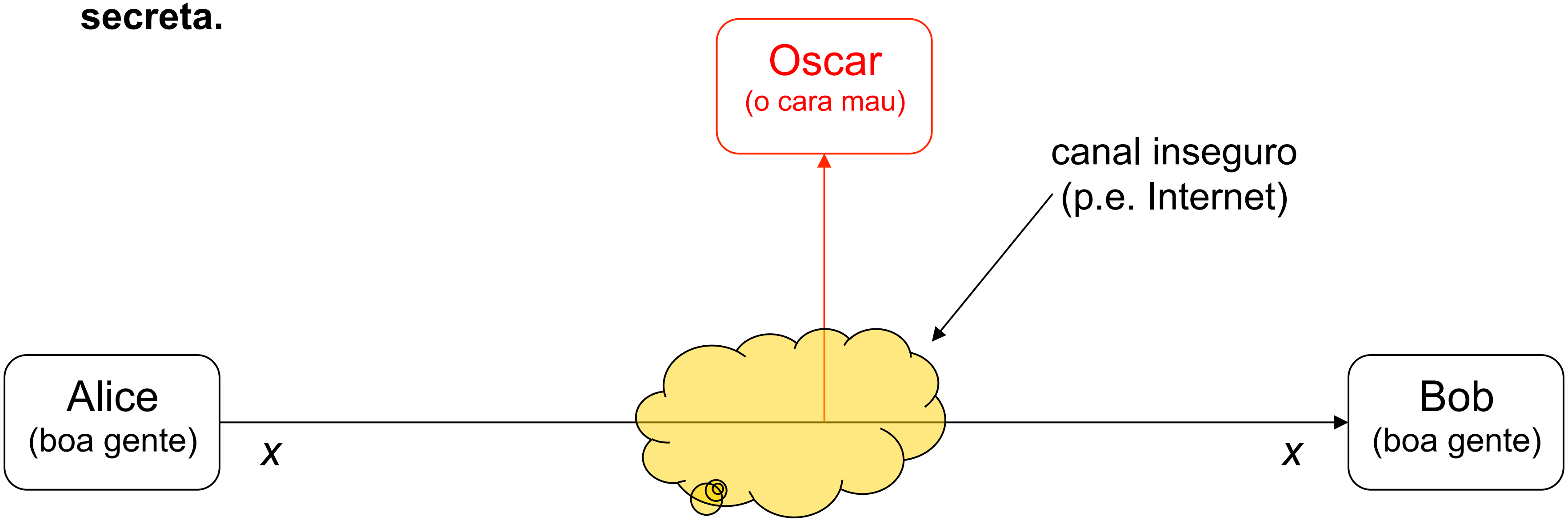
- **Esquemas Híbridos:** A maioria dos protocolos de hoje são esquemas híbridos, ou seja, usam os dois esquemas:
  - cifras simétricas (p. ex. para encriptação e autenticação da mensagem) e
  - cifras assimétricas (p.ex. para a troca de chaves e assinatura digital).

Conteúdo deste Capítulo

- Visão geral do campo da Criptologia
- **Conceitos Básicos da Criptografia Simétrica** A seguir
- Criptoanálise
- Cifra de Substituição
- Aritmética Modular
- Cifra de Deslocamento (ou de César) e Cifra Afim

Criptografia Simétrica

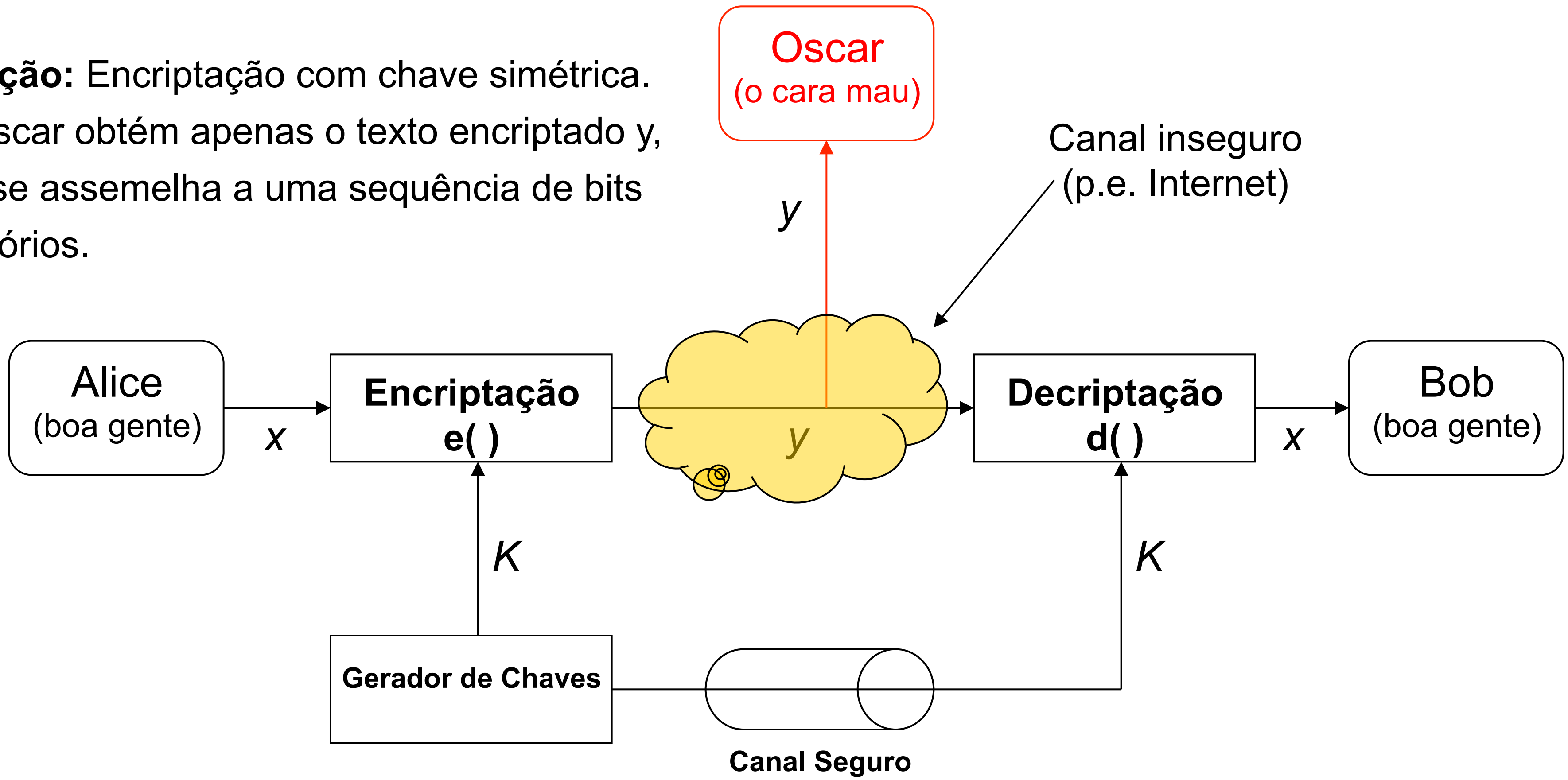
- Nomes alternativos: criptografia de **chave-privada**, **chave-única** ou **chave-secreta.**



- **Definição do Problema:**
  1) Alice e Bob gostariam de comunicar-se por meio de um canal não seguro (por exemplo uma WLAN ou a Internet).
  2) Um terceiro malicioso Oscar (o cara mau) tem acesso à informação que trafega no canal mas não deve ser capaz de entender a comunicação.

Criptografia Simétrica

Solução: Encriptação com chave simétrica.
⇒ Oscar obtém apenas o texto encriptado y, que se assemelha a uma sequência de bits aleatórios.



- x é o **texto em claro**
- y é o **texto encriptado**
- *K* é a **chave**
- O conjunto de todas as chaves {*K*1, *K*2, ...,*Kn*} é o **espaço de chaves**

Chapter 1 of *Understanding Cryptography* by Christof Paar and Jan Pelzl

Criptografia Simétrica

- Equação de Encriptação **$y = e_K(x)$**
- Equação de Decriptação **$x = d_K(y)$**

- Encriptação e decriptação são operações inversas se a mesma chave K é usada em ambos os lados.

$$d_K(y) = d_K(e_K(x)) = x$$

- Importante: A chave deve ser transmitida via um **canal seguro** entre Alice e Bob.
- O canal seguro pode ser implementado, p.ex., pela instalação manual da chave no protocolo Wi-Fi Protected Access (WPA), ou por um mensageiro humano.
- Entretanto, o sistema só é seguro se um atacante não consegue descobrir a chave K !

⇒ **Assim o problema da comunicação segura fica reduzido apenas à transmissão e armazenamento seguros da chave K.**

Conteúdo deste Capítulo

- Visão geral do campo da Criptologia
- Conceitos Básicos da Criptografia Simétrica
- **Criptoanálise** A seguir
- Cifra de Substituição
- Aritmética Modular
- Cifra de Deslocamento (ou de César) e Cifra Afim

Porque precisamos da Criptoanálise?

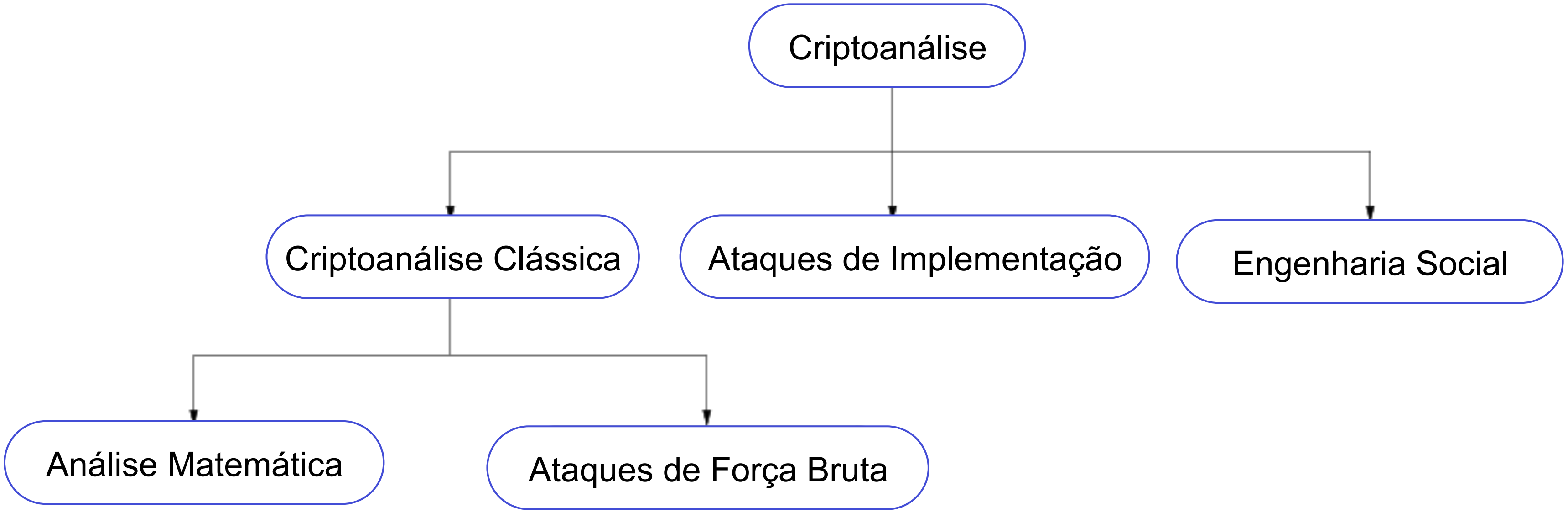
- Não há *nenhuma prova matemática de segurança* para qualquer das cifras práticas hoje em uso.
- Tentar quebrá-las (e não conseguir) é o único caminho de assegurar-se de que uma cifra é segura !

O Princípio de Kerckhoff é um princípio fundamental na criptografia moderna:

> Um sistema de criptografia deve ser seguro mesmo que o adversário (Oscar) conhecer todos os detalhes do sistema, com exceção da chave secreta.

- De forma a atingir o Princípio de Kerckhoff na prática: s**omente use cifras bem conhecidas, e que já tenham sido criptoanalisadas por vários anos por bons criptoanalistas !** (*Understanding Cryptography* só trata dessas cifras).
- **Observação:** É tentador presumir que uma cifra é "mais segura" se os seus detalhes são mantidos em segredo. No entanto, a história tem mostrado repetidamente que cifras secretas podem quase sempre ser quebradas, uma vez que se consegue fazer sua engenharia reversa. (Exemplo: o Sistema de Embaralhamento de Conteúdo (CSS) usado para proteção do conteúdo de DVDs.)

Criptoanálise: Atacando Sistemas Criptográficos



- **Ataques Clássicos**
  - Análise Matemática
  - Ataque de Força Bruta
- **Ataques de Implementação**: Tentativa de extrair a chave por meio de engenharia reversa ou ataques por canais colaterais, tais como medidas de consumo de potência.
- **Engenharia Social**: p.ex., induzir um usuário a fornecer a sua senha.

Ataque de Força Bruta (ou de busca exaustiva da chave) contra cifras simétricas

- Trata o algoritmo como uma caixa preta.
- Requer (no mínimo) um par de texto em claro e seu correspondente encriptado (x_0, y_0).
- Verifique todas as chaves possíveis até que a condição abaixo seja satisfeita:

$$\mathbf{d_K(y_0) \stackrel{?}{=} x_0}$$

- De quantas chaves você precisa?

Tamanho da chave em bits	Tamanho do espaço de chaves	Horizonte de segurança (supondo que o ataque de força bruta seja o melhor ataque possível)
64	2^{64}	**Curto prazo** (poucos dias ou menos)
128	2^{128}	**Longo prazo** (várias décadas, sem o uso de computadores quânticos)
256	2^{256}	**Longo prazo** (resistente também ao ataque com computadores quânticos – notar que tais computadores no momento ainda não existem e talvez nunca existam)

Importante: Um adversário só necessita ter sucesso em **um** ataque. Assim, um espaço de chaves extenso não ajuda se outros ataques (p.e. Engenharia social) forem possíveis.

Conteúdo deste Capítulo

- Visão geral do campo da Criptologia
- Conceitos Básicos da Criptografia Simétrica
- Criptoanálise
- **Cifra de Substituição** A seguir
- Aritmética Modular
- Cifra de Deslocamento (ou de César) e Cifra Afim

Cifra de Substituição

- Cifra histórica
- Ótima ferramenta para entender ataques de força bruta vs. ataques analíticos
- Encripta letras em vez de bits (como todas as cifras até a 2.a Guerra Mundial)

Ideia: Troque cada letra do texto em claro por uma outra letra fixa.

Texto em Claro		Texto Encriptado
A	→	k
B	→	d
C	→	w
	....	

Por exemplo, `ABBA` seria encriptado como `kddk`

- Exemplo (texto encriptado):

  `iq ifcc vqqr fb rdq vfllcq na rdq cfjwhwz hr bnnb hcc hwwhbsqvqbre hwq vhlq`

- Quão segura é a Cifra de Substituição? Vamos examinar os ataques…

Ataques contra Cifras de Substituição

1. Ataque: Busca exaustiva da chave (Ataque de Força Bruta)

- Simplesmente tente todas as tabelas de substituição possíveis até que um texto em claro inteligível apareça (note que cada tabela de substituição é uma chave)
- Quantas tabelas de substituição (= chaves) existem?

$$26 \times 25 \times \ldots \times 3 \times 2 \times 1 = 26! \approx 2^{88}$$

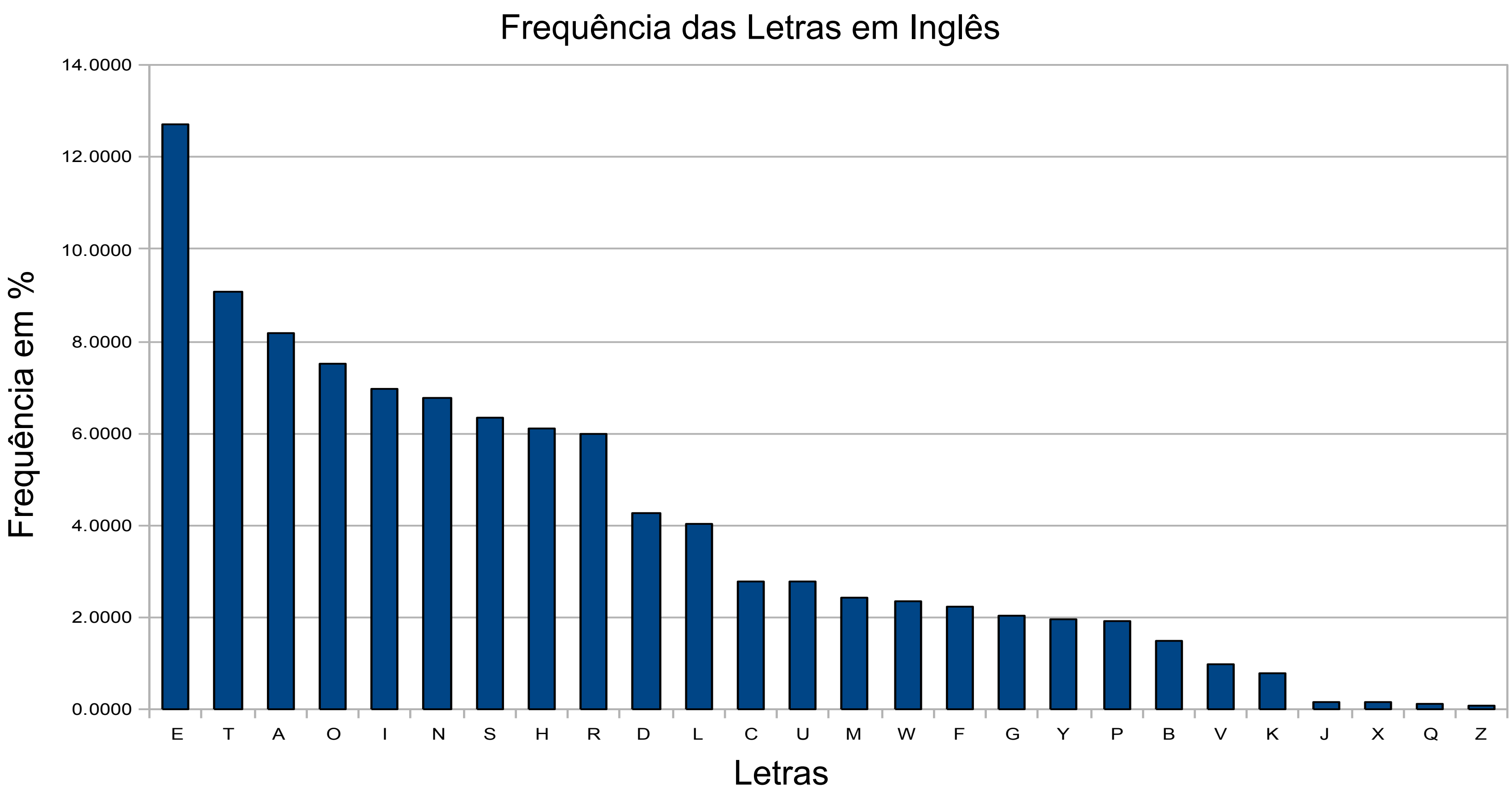
Pesquisar 2^{88} chaves é completamente inviável com os computadores de hoje! (conforme tabela anterior de comprimentos de chave)

- Pergunta: Podemos concluir que a cifra de substituição é segura, já que o ataque de força bruta não é viável?
- Resposta: Não! Nós temos que estar protegidos contra **todos** os possíveis ataques.

Ataques contra Cifras de Substituição

2. Ataque: Análise de frequência das letras (Ataque analítico)

- As letras na língua inglesa tem frequências de uso muito diferentes
- Além disso: o texto encriptado preserva a frequência das letras do texto em claro.
- Por exemplo, “e” é a letra mais comum em inglês; quase 13% de todas as letras em um texto típico em inglês são “e”s. A letra seguinte mais comum é o “t” com aproximadamente 9%.



Chapter 1 of *Understanding Cryptography* by Christof Paar and Jan Pelzl

Quebrando a Cifra de Substituição com a Análise de Frequência das Letras

- Vamos voltar ao exemplo e identificar a letra mais frequente:

  `iq ifcc vqqr fb rdq vfllcq na rdq cfjwhwz hr bnnb hcc hwwhbsqvqbre hwq vhlq`

- Trocamos a letra `q` por `E` e obtemos:

  `iE ifcc vEEr fb rdE vfllcE na rdE cfjwhwz hr bnnb hcc hwwhbsEvEbre hwE vhlE`

- Por um processo de tentativa e erro, baseado na frequência das letras restantes do texto encriptado, nós obtemos o texto em claro:

  `WE WILL MEET IN THE MIDDLE OF THE LIBRARY AT NOON ALL ARRANGEMENTS ARE MADE`

Quebrando a Cifra de Substituição com o Ataque de Frequência das Letras

- Na prática, não só a frequência das letras individuais, mas também a frequência de pares de letras (p.e., “th” é muito comum no inglês), triplas de letras, etc, podem ser usadas.
- No Problema 1.1 do *Understanding Cryptography* você deve tentar quebrar um texto encriptado mais longo!

Lição Importante: Ainda que a cifra de substituição tenha um espaço de chaves suficientemente grande, de aprox. 2^{88}, ele pode ser facilmente derrotado com métodos analíticos. Este é um exemplo excelente de que um esquema de encriptação deve suportar todos os tipos de ataques.

Conteúdo deste Capítulo

- Visão geral do campo da Criptologia
- Conceitos Básicos da Criptografia Simétrica
- Criptoanálise
- Cifra de Substituição
- **Aritmética Modular** A seguir
- Cifra de Deslocamento (ou de César) e Cifra Afim

Uma breve introdução à aritmética modular

Porque precisamos estudar aritmética modular?

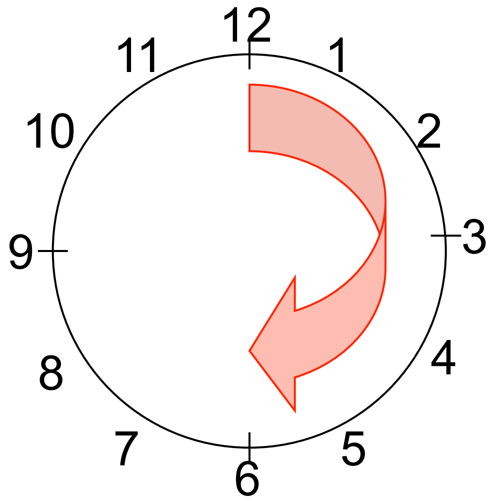
- Extremamente importante para a criptografia assimétrica (RSA, curvas elípticas, etc.)
- Algumas cifras históricas podem ser descritas elegantemente usando aritmética modular (conforme descrito a seguir, para a Cifra de César e a Cifra Afim).

Uma breve introdução à aritmética modular

Genericamente falando, a maioria dos sistemas criptograficos são baseados em **conjuntos de números** que são:

1. **discretos** (conjuntos com números inteiros são particularmente úteis)
2. **finitos** (i.e., computados com um conjunto finito de números)

Parece muito abstrato? --- Vamos olhar para um conjunto de números discretos que nos é muito familiar: um relógio.



É interessante notar que, embora os números sejam incrementados a cada hora, nunca deixam o conjunto dos números inteiros:

1, 2, 3, … 11, 12, 1, 2, 3, … 11, 12, 1, 2, 3, …:

Uma breve introdução à aritmética modular

- Desenvolveremos agora um sistema aritmético que nos permitirá calcula em um conjunto finito de números inteiros similar aos 12 inteiros que temos em um relógio (1,2,3, … ,12).
- É crucial ter uma operação que “mantenha os números dentro dos limites“, i.e., depois da adição e multiplicação, os resultados nunca devem sair do conjunto (i.e., nunca devem ser maiores que 12).

Definição: Operação Módulo

Sejam a, r, m inteiros e $m > 0$. Escreve-se

$$a \equiv r \bmod m$$

se $(r-a)$ *é divisível por* m.

- “m” é chamado de **módulo**
- “r” é chamado de **resto**

Exemplos de redução modular

- Sejam a= 12 e m= 9: $12 \equiv 3 \bmod 9$
- Sejam a= 34 e m= 9: $34 \equiv 7 \bmod 9$
- Sejam a= -7 e m= 9: $-7 \equiv 2 \bmod 9$

(você deve verificar se a condição “m divide (r-a)” é satisfeita em cada um desses 3 casos)

Propriedades da aritmética modular (1)

- **O resto não é único**

É surpreendente que, para cada par módulo “m” e número “a” dados, existam infinitos restos válidos.

Exemplos:

- $12 \equiv 3 \bmod 9$ → 3 é um resto válido pois 9 divide (3-12)
- $12 \equiv 21 \bmod 9$ → 21 é um resto válido pois 9 divide (21-12)
- $12 \equiv -6 \bmod 9$ → -6 é um resto válido pois 9 divide (-6-12)

Propriedades da aritmética modular (2)

- **Qual resto devemos escolher?**

Por convenção, nós estipulamos que **o menor inteiro positivo "r"** seja o resto. Este inteiro pode ser calculado como:

$$a = \overbrace{q}^{\text{quociente}}\ m + \overbrace{r}^{\text{resto}} \qquad \text{onde } 0 \leq r \leq m\text{-}1$$

- Exemplo: a=12 e m= 9

$$12 = 1 \text{ x } 9 + 3 \qquad \rightarrow \; r = 3$$

Observação: Está é apenas uma convenção. Num algoritmos, somos livres para escolher qualquer outro valor válido como resto para calcular nossas funções criptográficas.

Propriedades da aritmética modular (3)

- **Como fazemos a divisão modular?**

Primeiro, perceba que, em vez de fazer a divisão, preferimos multiplicar pelo inverso. Ex.:

$$b / a \equiv b \times a^{-1} \bmod m$$

O inverso de um número (a^{-1}) é definido de tal forma que:

$$a\, a^{-1} \equiv 1 \bmod m$$

Ex.:O que é 5 / 7 mod 9 ?

O inverso de 7 mod 9 é 4 pois $7 \times 4 \equiv 28 \equiv 1 \bmod 9$, assim:

$$5 / 7 \equiv 5 \times 4 = 20 \equiv 2 \bmod 9$$

- **Como o inverso é calculado?**

O inverso de um número *a* mod *m* existe se e somente se :

$$\text{mdc}\,(a, m) = 1$$

(note que no exemplo acima, mdc(5, 9) = 1; então, o inverso de 5 módulo 9 existe.

Por enquanto, a melhor forma de calcular o inverso é usar busca exaustiva. No capítulo 6 do livro *Understanding Cryptography* vamos aprender o poderoso algoritmo de Euclides que, na verdade, calcula o inverso de um dado número e um módulo.

Propriedades da aritmética modular (4)

- **A redução modular pode ser feita em qualquer ponto durante um cálculo**

Vamos examinar primeiro um exemplo. Queremos calcular 3^8 mod 7 (veremos que a exponenciação é extremamente importante na criptografia de chave pública).

1. Abordagem: Exponenciação seguida de redução modular

$$3^8 = 6561 \equiv \mathbf{2} \bmod 7$$

Note que nós temos o resultado intermediário 6561, mesmo sabendo que o resultado final não pode ser maior que 6.

2. Abordagem: Exponenciação com redução modular intermediária

$$3^8 = 3^4\, 3^4 = 81 \times 81$$

Neste ponto reduzimos o resultado intermediário 81 módulo 7:

$$3^8 = 81 \times 81 \equiv 4 \times 4 \bmod 7, \text{ e}$$

$$4 \times 4 = 16 \equiv \mathbf{2} \bmod 7.$$

Note que podemos fazer todas essas multiplicações sem uma calculadora de bolso; entretanto, calcular mentalmente $3^8 = 6561$ é um desafio para a maioria de nós.

Regra geral: Para a maioria dos algoritmos, é vantajoso reduzir resultados intermediários o mais cedo possível.

Uma visão algébrica da aritmética modular: O Anel Z_m (1)

Podemos ver a aritmética modular em termos de conjuntos e operações sobre conjuntos. Ao fazer aritmética módulo *m*, obtemos o **Anel Inteiro Z_m** com as seguintes propriedades:

- **Fechamento**: Podemos somar e multiplicar dois números inteiros quaisquer; o resultado é sempre um elemento do anel.
- Adição e multiplicação são **associativas**, i.e., para todo $a,b,c \,\varepsilon\, Z_m$

  $a + (b + c) = (a + b) + c$

  $a \times (b \times c) = (a \times b) \times c$

  E a adição é **comutativa**: $a + b = b + a$
- A **Lei distributiva** assegura que $a\times(b+c) = (a\times b)+(a\times c)$ para todos $a,b,c \,\varepsilon\, Z_m$
- Existe o **elemento neutro, *0*, da adição**, i.e., para todo $a \,\varepsilon\, Z_m$

  $a + 0 \equiv a \bmod m$
- Para todo $a \,\varepsilon\, Z_m$, sempre há **o elemento inverso aditivo $-a$** tal que

  $a + (-a) \equiv 0 \bmod m$
- Existe o **elemento neutro, 1, da multiplicação**, i.e., para todo $a \,\varepsilon\, Z_m$

  $a \times 1 \equiv a \bmod m$
- O **inverso multiplicativo a^{-1}**

  $a \times a^{-1} \equiv 1 \bmod m$

  somente existe para alguns, mas não todos os elementos em Z_m.

Uma visão algébrica da aritmética modular: O Anel Z_m (2)

Grosso modo, um anel é uma estrutura na qual sempre podemos adicionar, subtrair e multiplicar, mas só podemos dividir por certos elementos (ou seja, por aqueles para os quais existe inverso multiplicativo).

- Lembramos das explicações anteriores que um elemento $a \varepsilon Z_m$ tem seu inverso multiplicativo se e somente se:

  mdc $(a, m) = 1$

  Nós dizemos que "a" e "m" são **coprimos** ou **primos entre si**.

- Ex: Considere o Anel $Z_9 = \{0,1,2,3,4,5,6,7,8\}$

  Os elementos 0, 3, e 6 não tem inversos pois eles não são coprimos de 9.

  Os inversos multiplicativos dos elementos 1, 2, 4, 5, 7, e 8 são:

  $1^{-1} \equiv 1 \bmod 9$ $\quad$ $2^{-1} \equiv 5 \bmod 9$ $\quad$ $4^{-1} \equiv 7 \bmod 9$

  $5^{-1} \equiv 2 \bmod 9$ $\quad$ $7^{-1} \equiv 4 \bmod 9$ $\quad$ $8^{-1} \equiv 8 \bmod 9$

Conteúdo deste Capítulo

- Visão geral do campo da Criptologia
- Conceitos Básicos da Criptografia Simétrica
- Criptoanálise
- Cifra de Substituição
- Aritmética Modular
- **Cifra de Deslocamento (ou de César) e Cifra Afim**

A seguir

Cifra de Deslocamento (ou de César) (1)

- Cifra antiga, supostamente usada por Júlio César.
- Substitui cada letra do texto em claro por uma outra letra.
- Regra de substituição é bem simples: Pegue a letra que está 3 posições à frente no alfabeto.
- É preciso mapear letras para números:

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

- Exemplo para $k = 7$

Texto em claro = `ATTACK` = 0, 19, 19, 0, 2, 10

Texto encriptado = `haahjr` = 7, 0, 0, 7, 9, 17

Note que as letras continuam circularmente do final para o início do alfabeto, o que pode ser expresso matematicamente pela redução módulo 26, p. ex., $19 + 7 = 26 \equiv 0 \bmod 26$

Cifra de Deslocamento (ou de César) (2)

- Descrição matematicamente elegante da cifra.

Sejam k, x, y ε {0,1, …, 25}

- Encriptação: $y = e_k(x) \equiv x + k \bmod 26$
- Decriptação: $x = d_k(x) \equiv y - k \bmod 26$

- Pergunta; A cifra de deslocamento é segura?
- Resposta: Não! vários ataques são possíveis, incluindo:
  - Busca exaustiva de chaves (o tamanho do espaço de chaves é somente **26 (!!)**)
  - Análise de frequência das ketras, similar ao ataque contra a cifra de substituição

Cifra Afim (1)

- Extensão da cifra de deslocamento: em vez de apenas adicionarmos a chave ao texto em claro, nós também multiplicamos pela chave.
- Usamos então uma chave consistindo de 2 partes: $k = (a, b)$

Sejam $k, x, y \in \{0, 1, \ldots, 25\}$

- Encriptação: $y = e_k(x) \equiv a\,x + b \bmod 26$
- Decriptação: $x = d_k(x) \equiv a^{-1}(y - b) \bmod 26$

- Como é necessário o inverso de *a* para a decriptação, podemos somente usar valores de a para os quais:

$$\text{mdc}(a, 26) = 1$$

  Existem 12 valores de *a* que satisfazem a essa condição.
- Dessa forma, segue que o tamanho do espaço de chaves é somente 12 x 26 = 312 (conforme: Sec 1.4 do *Understanding Cryptography*)
- Novamente, vários ataques são possíveis, incluindo:
  - Busca exaustiva de chaves e análise de frequência das letras, similarmente aos ataques contra a cifra de substituição.

Lições aprendidas

- Nunca, jamais, desenvolva o seu próprio algoritmo de criptografia, a menos que você tenha uma equipe de experientes criptoanalistas verificando o seu projeto.
- Não utilize algoritmos de criptografia não comprovados ou protocolos não comprovados.
- Os atacantes vão sempre olhar para o ponto mais fraco de um sistema de criptografia. Por exemplo, um grande espaço de chaves por si só não é garantia de uma cifra segura; a cifra ainda pode estar vulnerável contra ataques analíticos.
- Comprimentos de chave de algoritmos simétricos, a fim de frustrar ataques de busca exaustiva da chave:
  - 64 bits: inseguro exceto para dados cujo uso se dará num prazo muito curto.
  - 128 bits: segurança de longo prazo para várias décadas, a menos que os computadores quânticos se tornem disponíveis (computadores quânticos não existem e talvez nunca existam).
  - 256 bits: como acima, mas provavelmente seguros até contra ataques por computadores quânticos.
- Aritmética modular é uma ferramenta para exprimir os esquemas de encriptação históricos, tais como a Cifra Afim, de uma maneira matematicamente elegante.