UNIVERSIDADE D
COIMBRA

José Songo

Criptografia RSA

Seminário em Matemática Aplicada e Computação no âmbito do Mestrado em Matemática orientada pelo Professor Doutor Pedro Quaresma

Janeiro de 2024

Criptografia RSA

José Diogo Songo

UNIVERSIDADE D
COIMBRA

Seminário de Investigação

Research Seminar

Seminário em Análise Aplicada e Computação

MSc Seminar in Applied Analysis and Computation

Janeiro 2024 / January 2024

Agradecimentos

Agradeço ao professor Pedro Quaresma, que aceitou orientar o meu projeto, revelando uma especial atenção às minhas ideias. Os seus conselhos e sugestões bem como a valorização do trabalho desenvolvido foram determinantes para alcançar este resultado.

Resumo

A criptografia RSA é um sistema de segurança digital que utiliza um par de chaves - uma pública e uma privada - para proteger informações durante a transmissão. Fundamentado na complexidade matemática da fatorização de números de grande dimensão. O RSA é muito utilizado para garantir a confidencialidade e autenticidade de dados em transações online, *e-mails* seguros e na proteção de informações sensíveis. Sua utilidade reside na capacidade de enviar mensagens seguras para qualquer pessoa, usando uma chave pública disponível, enquanto apenas o destinatário autorizado, com acesso à chave privada correspondente, pode decifrar e ler a mensagem. Este método de criptografia eficaz continua a desempenhar um papel essencial na segurança da era digital.

Neste estudo, inicialmente introduzimos os princípios matemáticos essenciais para a compreensão dos sistemas criptográficos. Em seguida, discutimos os sistemas criptográficos, destacando especialmente a criptografia RSA, abordando a cifra RSA assim como a sua criptoanálise (incluindo Métodos de Fermat, Divisões e Euclides). Concluímos com uma análise comparativa dos diversos métodos mencionados anteriormente.

Conteúdo

Capítulo 1

Introdução

A criptografia é uma técnica antiga e importante utilizada para proteger informações confidenciais por meio da codificação e descodificação de mensagens. Ela desempenha um papel essencial na segurança da comunicação, impossibilitando que terceiros não autorizados compreendam ou acessem o conteúdo das mensagens. [2, 3].

Um dos exemplos mais simples de criptografia é a cifra de Júlio César, também conhecida como cifra de deslocamento ou cifra de substituição. Essa técnica foi usada pelo renomado imperador romano Júlio César (100aC – 44aC) para enviar mensagens secretas durante as suas ações militares.

A cifra de Júlio César funciona substituindo cada letra do alfabeto por outra, deslocando um número fixo de posições para a esquerda ou direita no alfabeto. Por exemplo, se usarmos um deslocamento de 3 posições para a direita, a letra “a” seria substituída pela letra “d”, ”b” seria substituída por “e”, e assim por diante. Ao chegar ao fim do alfabeto, volta-se ao início, por exemplo “x” passa a “a”,“y” passa a “b” e “z” passa a “c”.

Essa técnica simples demonstra os princípios fundamentais da criptografia, onde a informação original é escondida de uma maneira específica para ocultar seu significado, e somente aqueles que possuem a chave podem decifrar a mensagem.

No entanto, a cifra de Júlio César é bastante vulnerável a ataques de força bruta, pois há um número limitado de combinações possíveis. Por isso, ela é considerada uma forma trivial de criptografia e não é adequada para proteger informações sensíveis nos dias atuais. Hoje em dia, algoritmos mais complexos e seguros são utilizados para garantir a segurança das comunicações, como AES (Advanced Encryption Standard) e RSA (Rivest-Shamir-Adleman), oferecendo níveis avançados de proteção.

Capítulo 2

Fundamentos Matemáticos

Em seguida são apresentados fundamentos teóricos matemáticos indispensáveis ao estudo da criptografia.

2.1 Funções Unidirecionais

Definição 1 (Função Unidirecional)**.** *Uma função f de um conjunto X para um conjunto Y é dita uma função unidirecional («one-way function») se $f(x)$ é «fácil de calcular» para todo o $x \in X$, mas «essencialmente para todos» os elementos $y \in Im(f)$ é «computacionalmente difícil» achar um $x \in X$ tal que $f(x) = y$.*

Definição 2 (Função Unidirecional com Escapatória)**.** *Uma função unidirecional com escapatória é uma função unidirecional $f : X \longrightarrow Y$ com a propriedade de que dado algum tipo de informação adicional torna-se possível encontrar, para um dado $y \in Im(f)$ e um dado $x \in X$ tal que $f(x) = y$.*

2.2 Resultados da Teoria dos Números

Definição 3 (Divisibilidade)**.** *: Dados $a, b \in \mathbb{Z}$, com $a \neq 0$, diz-se que, a divide b, e escreve-se $a|b$, se existe $q \in \mathbb{Z}$ tal que $b = aq$.*

Convenção: Quando se escreve $a|b$ está implícito que a $a \neq 0$.
Se $a|b$ também se diz que a é um divisor de b, que b é um múltiplo de a ou que b é divisível por a.
Se a não divide b, escreve-se $a \nmid b$.
Para quaisquer $a, b, c \in \mathbb{Z}$ tem-se:

1. $a|0, 1|a$ *e* $a|a$;
2. $a|b \Leftrightarrow a|-b \Leftrightarrow -a|b$;
3. $a|b \wedge b|c \Leftrightarrow a|c$;
4. Para quaisquer $x, y \in \mathbb{Z}, a|b \wedge a|c \Leftrightarrow a|bx + cy$;
5. $a|1 \Leftrightarrow a = \pm b$;

6. $a,b \in \mathbb{N} \wedge a|b \Leftrightarrow a \leq b$;

7. Um inteiro não nulo tem um número finito de divisores.

Teorema 1 (Algoritmo da Divisão Inteira)**.** *Dados $a,b \in \mathbb{Z}$, com $a \neq 0$, existem $q,r \in \mathbb{Z}$, únicos, tais que*

$$b = aq, \text{ com } 0 \leq r < |a|$$

q e r são, respetivamente, o quociente e o resto da divisão inteira de b por a.

Observações:

1. $a|b$ se e só se o resto da divisão inteira de b por a é zero.

2. Em $C/C++$, os operadores "/" e "%" dão-nos o quociente e o resto da divisão inteira (desde que o divisor e o dividendo sejam inteiros).

Definição 4 (Congruência $\mod m$)**.** *Para $m \in \mathbb{N}$ a relação de congruência módulo m é a relação definida em $\mathbb{Z}$ por*

$$a \equiv b(mod\ m) \Leftrightarrow m|a-b,\ a,b \in \mathbb{Z}$$

Se $a \equiv b(\mod m)$ diz-se que a é congruente módulo m com b.

Observe-se que $a \equiv b(mod\ m)$ se e só se a e b têm o mesmo resto quando divididos por m.
Propriedades da congruência

1. $a \equiv a\ (mod\ m)$ (Reflexividade)

2. $a \equiv b\ (mod\ m) \Rightarrow b \equiv a\ (mod\ m)$

3. $a \equiv b\ (mod\ m) \wedge a \equiv b\ (mod\ m) \Rightarrow a \equiv c(mod\ m)$

4. $a \equiv b\ (mod\ m) \wedge c \equiv d(mod\ m) \Rightarrow a+c \equiv b+d\ (mod\ m)$

5. $a \equiv b\ (mod\ m) \wedge c \equiv d(mod\ m) \Rightarrow ac \equiv bd\ (mod\ m)$

6. $a \equiv b\ (mod\ m) \Rightarrow mdc(a.m) = mdc\ (b,m)$

7. $ab \equiv ac\ (mod\ m) \Rightarrow b \equiv c\ (mod \frac{m}{mdc\ (a,m)})$

Das propriedades 1, 2 e 3 resulta que, para $m \in \mathbb{N}$, a relação de congruência módulo m é uma relação de equivalência em $\mathbb{Z}$.

As classes de equivalência desta relação de equivalência chamam-se classes de congruência módulo m.

A classe de congruência módulo m a que pertence $a \in \mathbb{Z}$ é representada por $[a]_m$ ou $\overline{a}$.

Uma vez que $a \in \mathbb{Z}$ é congruente módulo m com o resto da divisão inteira por m, e os m restos possíveis são $0,1,2,\ldots,m-2$ e $m-1$ classes, conclui-se que há m classes de congruência módulo m: $[0]_m,[1]_m,\ldots,[m-1]_m$

Para $m \in \mathbb{N}$, por $\mathbf{Z}_m$ representa-se o conjunto das classes de congruência módulo m, isto é,

$$\mathbb{Z}_m = [0]_m, [1]_m, \ldots, [m-1]_m$$

Uma vez que

$$a \equiv c(mod\ m) \wedge b \equiv d(mod\ m) \Rightarrow a+b \equiv c+d(mod\ m)$$

pode definir-se uma operação em $\mathbb{Z}_m$(adição de classes de congruência) por $[a]_m + [a]_m = [a+b]_m$

$$(\mathbb{Z}_m, +)\ \textit{um grupo abeliano}$$

Neste grupo o elemento neutro é $[0]_m$ e o simétrico de $[a]_m$ é $[-a]_m = [m-a]_m$

Definição 5 (Máximo Divisor Comum)**.** *Sejam $a,b \in \mathbb{Z}$ com $a \neq 0$ ou $b \neq 0$ e considere-se $D = \{c \in \mathbb{Z} : c|a \wedge c|b\}$.*

$D \neq \emptyset$, porque $1 \in D$ e D é finito porque um inteiro não nulo tem um número finito de divisores.

Então D tem um máximo ao qual se chama máximo divisor comum de a e b. Esse máximo é representado por $mdc(a,b)$.

Se $mdc(a,b) = 1$ diz-se que os inteiros a e b são primos entre si ou que a é primo com b.

Propriedades dos máximos divisores comum

Para quaisquer $a,b,c \in \mathbb{Z}\backslash\{0\}$, tem-se:

1. $mdc(a,b) = mdc(b,a) = mdc(a,-b)$;

2. $mdc\left(\frac{a}{mdc(a,b)}, \frac{b}{mdc(a,b)}\right) = 1$

3. $mdc(a,b)$ é o menor elemento positivo de $\{ax+by :\ x,y \in \mathbb{Z}\}$

4. Se $x,y \in \mathbb{Z}$ são tais que $mdc(a.b) = ax+by$ então $mdc(x,y) = 1$;

5. $mdc(a,b)$ é o único divisor comum, positivo, de a e b tal que:

$$x \in \mathbb{Z}\backslash\{0\} \wedge x|a \wedge x|b \Rightarrow x|mdc(a.b)$$

6. $a|bc \wedge mdc(a,b) = 1 \Rightarrow a|c$

Para calcular o máximo divisor comum de $a,b \in \mathbb{Z}\backslash\{0\}$ usa-se o algoritmo de Euclides.

Teorema 2 (Algoritmo de Euclides)**.** *Sejam $a \in \mathbb{N}$ e $b \in \mathbb{Z}$. Aplicando sucessivamente o algoritmo da divisão obtém-se:*

$$
\begin{aligned}
&b = aq_1 + r_1,\ 0 < r_1 < a \\
&a = r_1q_2 + r_2,\ 0 < r_2 < r_1 \\
&r_1 = r_2q_3 + r_3,\ 0 < r_3 < r_2 \\
&\vdots \\
&r_{k-2} = r_{k-1}q_k + r_k,\ 0 < r_k < r_{k-1} \\
&r_{k-1} = r_kq_{k+1}
\end{aligned}
$$

para um dado $k \in \mathbb{N}$ ($r_0 := a$ *e* $r_{-1} := b$)

Então $mdc(a,b) = r_k$.

Proposição 3 (Congruências Lineares)**.** *Sejam* $m \in \mathbb{N}$, $a,b \in \mathbb{Z}$ *com* $mdc(a,m) = 1$. *A congruência* $ax \equiv b(mod\ m)$ *tem solução e o conjunto das soluções é uma classe de congruência módulo m.*

Método de resolução de $ax \equiv b(mod\ m)$ com $mdc(a,m) = 1$:

Usando o algoritmo de Euclides determinam-se $x_0, y_0 \in \mathbb{Z}$ tais que $ax_0 + my_0 = 1$. De $ax_0 \equiv 1(mod\ m)$ resulta que $ax_0 \equiv b(mod\ m)$.

O conjunto das soluções de $ax \equiv b(mod\ m)$ é $[x_0b]_m$.

Proposição 4 (Congruências Lineares—Caso geral)**.** *Sejam* $m \in \mathbb{N}, a,b \in \mathbb{Z}$ *e* $d = mdc(a,m)$. *A congruência* $ax \equiv b(modm)$ *tem solução se e só se* $d|b$.

Se $d|b$ então

$$ax \equiv b(mod\ m) \Leftrightarrow \frac{a}{d}x \equiv \frac{b}{d}\left(mod\frac{m}{d}\right)$$

e o conjunto das soluções é a união de d classes de congruência módulo m.

Proposição 5 (Inverso Multiplicativo)**.** *: Se* $mdc(a,m) = 1$, *então* $[a]_m$ *é invertível em* $\mathbb{Z}_m$ *e , sendo* $x_0, y_0 \in \mathbb{Z}$ *tais que* $ax_0 + my_0 = 1$, *tem-se que:*

$$[a]_m^{-1} = [x_0]_m$$

Notar que se $mdc(a,m) > 1, [a]_m$ não é invertível

Os elementos invertíveis em $(\mathbb{Z}_m, .)$ são os elementos de $\{[a]_m : mdc(a,m) = 1\}$.

Observação: $(\mathbb{Z}_m \backslash \{[0]_m\}, .)$ é um grupo, se e só se m é primo.

Teorema 6 (Grupo multiplicativo)**.** *Seja* $U_m = \{[a]_m : mdc(a,m) = 1\}$. *Consideremos* $m \in \mathbb{N}$. U_m *é um grupo para a multiplicação de classes de congruência módulo m.*

Notações:

1. Quando se trabalha em $\mathbb{Z}_m$ muitas vezes representa-se $[a]_m$ apenas por r, sendo $r \in \{0,1,\ldots,m-1\}$ o resto da divisão inteira por a por m;

2. Sendo $a \in \mathbb{Z}$ e $m \in \mathbb{N}$, é usual representar por *amod m* o resto da divisão inteira de a por m;

3. Se $mdc(a,m) = 1$, por $a^{-1} \mod m$ representa-se o inverso de $[a]_m$ em $\mathbb{Z}_m$

Definição 6 (Menor Múltiplo Comum)**.** *Dados $a,b \in \mathbb{Z}\backslash\{0\}$, um inteiro não nulo c é um múltiplo comum de a e b se $a|c$ e $b|c$.*

Sejam $a,b \in \mathbb{Z}\backslash\{0\}$ e considere-se $M = \{c \in \mathbb{N} : a|c \wedge b|c\}$.

$M \neq \emptyset$ porque $|ab| \in M$. Além disso, $M \subseteq \mathbb{N}$. Então M tem um mínimo ao qual se chama menor múltiplo comum de a e b.
Esse mínimo é representado por $mmc(a,b)$.

Para quaisquer $a,b \in \mathbb{Z}\backslash\{0\}$, tem-se:

1. $mmc(a,b)$ é o único múltiplo comum, positivo, de a e b tal que:

$$x \in \mathbb{Z} \wedge a|x \wedge b|x \Rightarrow mmc(a,b)|x$$

2. Se $n \in \mathbb{N}$ é um divisor comum de a e b então

$$mmc\left(\frac{a}{n}, \frac{b}{n}\right)$$

3. $mmc(a,b)mdc(a,b) = |ab|$

O menor múltiplo comum de $a,b \in \mathbb{Z}\backslash\{0\}$ pode ser calculado usando o algoritmo de Euclides e a propriedade 3.

Definição 7 (máximo divisor comum—caso geral)**.** *$n \in \mathbb{N}, n \geq 2, a_1, a_2, \ldots, a_n \in \mathbb{Z}$ não todos nulos. O máximo divisor comum de $a_1, a_2, \ldots, a_n$ é o menor dos divisores comuns positivos de $a_1, a_2, ..., a_n$. Representa-se por $mdc(a_1, a_2, \ldots, a_n)$*

Propriedades:

1. *$mdc(a_1, a_2, \ldots, a_n)$ é o menor inteiro positivo da forma $a_1x_1 + a_2x_2 + \ldots + a_nx_n$, com $x_1, x_2, \ldots, x_n \in \mathbb{Z}$*

2. *$mdc(a_1, a_2, \ldots, a_n)$ é o único divisor comum positivo, de $a_1, a_2, \ldots, a_n$ que é múltiplo de qualquer divisor comum de $a_1, a_2, \ldots, a_n$*

3. *$mdc(a_1, a_2, \ldots, a_n) = mdc(mdc(a_1, a_2, \ldots, a_{n-1}), a_n)$*

Definição 8. *Os inteiros $a_1, a_2, \ldots, a_n$ são primos dois a dois se $mdc(a_i, a_j) = 1$, para $i, j = 1, 2, \ldots, n$, com $i \neq j$.*

$$a_1, a_2, \ldots, a_n \text{ são primos dois a dois}$$

$$\downarrow$$

$$a_1, a_2, \ldots, a_n \text{ são primos entre si}$$

Notar que a implicação reciproca é falsa.

Definição 9. *: $n \in \mathbb{N}, n \geq 2, a_1, a_2, \ldots, a_n \in \mathbb{Z}$ não todos nulos. O menor múltiplo comum de $a_1, a_2, \ldots, a_n$ é o menor dos múltiplos comuns positivos de $a_1, a_2, \ldots, a_n$. Representa-se por $mmc(a_1, a_2, \ldots, a_n)$.*

Propriedades:

1. *$mmc(a_1, a_2, \ldots, a_n)$ é o único múltiplo comum, positivo, de $a_1, a_2, \ldots, a_n$ que divide qualquer múltiplo comum de $a_1, a_2, \ldots, a_n$;*
2. *$mmc(a_1, a_2, \ldots, a_n) = mmc(mmc(a_1, a_2, \ldots, a_{n-1}), a_n)$*

Para $n \leq 3$, em geral,
$mmc(a_1, a_2, \ldots, a_n) = mmc(mmc((a_1, a_2, \ldots, a_{n-1}), a_n) \neq |a_1, a_2, \ldots, a_n|$

Definição 10. *Um inteiro $p > 1$ diz-se um número primo se os únicos divisores positivos de p são 1 e p. Um inteiro diz-se composto se não é primo.*

Teorema 7. *p número primo; $a_1, a_2, \ldots, a_n \in \mathbb{Z}$*

$$p|a_1 a_2 \ldots c a_n \Rightarrow p|a_1 \vee p|a_2 \vee \ldots c \vee p|a_n$$

Teorema 8 (Teorema Fundamental da Aritmética)**.** *Todo o inteiro maior que* 1 *pode ser escrito, de modo único(a menos da ordem dos fatores), como produto de números primos.*

Teorema 9 (Fatorização de Números Primos)**.** *Se $a = p_1^{\alpha_1} p_2^{\alpha_2} \ldots p_k^{\alpha_k}$ e $a = p_1^{\beta_1} p_2^{\beta_2} \ldots p_k^{\beta_k}$ onde $p_1, \ldots p_k$ são números primos dois a dois e $\alpha_1, \ldots, \alpha_k, \beta_1, \ldots, \beta_k \in \mathbb{N}_0$, então*

$$a|b \Leftrightarrow (\alpha_i \leq \beta_i, i = 1, \ldots, k)$$

$$mdc(a,b) = p_1^{min\{\alpha_1, \beta 1\}} p_2^{min\{\alpha_2, \beta 2\}} \ldots p_k^{min\{\alpha_k, \beta_k\}}$$

e

$$mdc(a,b) = p_1^{max\{\alpha_1, \beta 1\}} p_2^{max\{\alpha_2, \beta 2\}} \ldots p_k^{max\{\alpha_k, \beta_k\}}$$

Definição 11 (Função de Euler)**.** *: a função de Euler é a função $\phi : \mathbb{N} \Rightarrow \mathbb{N}$ definida por:*

$$\phi(n) = |a \in \mathbb{N} : a \leq n \text{ e } mdc(a,n) = 1|, \; n \in \mathbb{N}$$

Teorema 10. *A função ϕ é multiplicativa, isto é, se $m, n \in \mathbb{N}$ são tais que $mdc(m,n) = 1$, então*

$$\phi(mn) = \phi(m)\phi(n)$$

Teorema 11. *Sejam, $p_1, p_2, \ldots, p_n$ números primos distintos dois a dois e*
$\alpha_1, \alpha_2, \ldots, \alpha_n \in \mathbb{N}$

$$\phi(p_1^{\alpha_1} \ldots p_k^{\alpha_k}) = (p_1^{\alpha_1} - p_1^{\alpha - 1}) \ldots (p_k^{\alpha_k} - p_k^{\alpha - 1})$$

Teorema 12 (Pequeno Teorema de Fermat)**.** *Se n é um número primo, então* $a^{n-1} \equiv 1(modn)$*, para todo o* $a \in \mathbb{Z}$ *tal que* $mdc(a,n) = 1$

Teorema 13 (Teorema Chinês dos Restos)**.** *Sejam* $m_1, m_2, \ldots, m_k \in \mathbb{Z}$ *primos dois a dois e* $a_1, a_2, \ldots, a_k \in \mathbb{Z}$*.*

O sistema

$$\begin{cases} x \equiv a_1 (mod\ m_1) \\ \vdots \\ x \equiv a_K (mod\ m_k) \end{cases} \tag{2.1}$$

tem solução.

Seja $m = m_1 m_2 \ldots\ m_k$. Para $i = 1, 2, \ldots, m_k$ seja $b_i \in \mathbb{Z}$ tal que $\frac{m}{m_i} b_i \equiv 1 (mod\ m_i)$ e considere-se

$$x_0 = \sum_{i=1}^{k} \frac{m}{m_i} a_i b_i$$

O conjunto das soluções é $[x_0]_m$.

Capítulo 3

Criptografia

O surgimento da criptografia (do grego:Kryptós,oculto + graph, r.de graphein, escrever) remota a milhares de anos, sendo uma das técnicas mais antigas na transmissão de informação de forma secreta. No ano 400 a.C., os Espartanos conceberam um método intrigante: eles gravavam mensagens em uma tira de couro enrolada em um bastão. Ao desenrolar a tira do bastão, a mensagem era cifrada, requerendo que fosse enrolada novamente em um bastão de diâmetro similar para decifrá-la [3].

Um sistema criptográfico é então um conjunto de procedimentos que possibilitam tornar uma mensagem ilegível para qualquer pessoa que não seja o destinatário autorizado, garantindo que somente o destinatário legítimo possa decodificá-la e acessar o conteúdo original [3].

Portanto, a criptografia tem os seguintes objetivos:

Confidencialidade Manter o conteúdo da informação confidencial para todos exceto o destinatário;

Integridade da informação Certificar-se de que não há adulteração da informação por partes não autorizadas;

Autenticação

- das entidades que comunicam entre si;
- da informação (origem, conteúdo, data de envio,...)

Não repudiação o criador da informação não pode negar a autoria.

3.1 Sistemas Criptográficos Simétricos

Os primeiros sistemas criptográficos desenvolvidos foram baseados na criptografia simétrica, conhecida como sistemas de chave secreta. Esses sistemas utilizam uma única chave para cifrar e decifrar informações, onde os processos de encriptação e desencriptação são idênticos.

Mas, esses sistemas enfrentam dois problemas que limitam a sua eficácia na proteção das informações:

A chave de cifração deve ser compartilhada por todos os elementos da organização “amiga” e, ao mesmo tempo, deve ser mantida em segredo absoluto de organizações consideradas como adversárias. Quanto mais complexa for a estrutura da organização “amiga”, mais complicado será garantir esta condição [3].

3.2 Sistemas Criptográficos Assimétricos

Aparecem então os sistemas de criptografia assimétrica, ou de chave pública. Nestes sistemas, a cifragem utiliza uma chave diferente, conhecida como chave privada.

Este tipo de sistema resolve os dois problemas mencionados anteriormente:

Os algoritmos desenvolvidos são significativamente mais complexos para serem quebrados em comparação com os sistemas anteriores.

+ A chave privada é conhecida por apenas uma única entidade, o destinatário da mensagem. Manter essa chave secreta torna-se assim consideravelmente mais simples;

- os algoritmos desenvolvidos são menos eficientes do que as atuais cifras do tipo simétrico.

Para avaliar as ferramentas criptográficas utilizam-se os seguintes parâmetros:

Nível de Segurança número de operações solicitadas pelo melhor método conceituado para quebrar o código;

Funcionalidade quais são as primitivas mais eficientes para uma dada finalidade;

Método de Operações o procedimento de cada primitiva depende da maneira como são aplicadas e de quais os valores que lhe são dados;

Performance as ferramentas têm de ser eficientes em termos de tempo e espaço;

Facilidade de Implementação implementar uma dada ferramenta num dado sistema operacional de forma simples.

Para criar um esquema de encriptação vamos precisar selecionar:

- um alfabeto(finito) definição;
- um espaço de mensagens
- um espaço de chaves;
- um conjunto de transformações de encriptação;
- um correspondente conjunto de transformações de desencriptação

Para tal vão ser introduzidas , formalmente, as seguintes definições:

Definição 12 (Alfabeto de Definição)**.** *$\mathscr{A}$ denota de um conjunto finito de símbolos designado por alfabeto de definição.*

Definição 13 (Espaço das Mensagens)**.** *$\mathscr{M}$ denota um conjunto designado por espaço das mensagens. $\mathscr{M} = \mathscr{A}^*$ consiste de sequências de elementos de um alfabeto de definição. Um elemento de $\mathscr{M} = \mathscr{A}^*$ é designado por mensagem de texto claro (não cifrado).*

Definição 14 (Espaço das Mensagens Cifradas)**.** *$\mathscr{C}$ denota um conjunto designado por espaço das mensagens cifradas. $\mathscr{C}$ consiste de sequências de elementos de um dado alfabeto de definição, o qual pode diferir do usado em $\mathscr{M}$. Um elemento de $\mathscr{C}$ é designado por texto cifrado.*

Definição 15 (Definição das chave)**.** *$\mathscr{K}$ denota um conjunto designado por espaço das chaves. Um elemento de $\mathscr{K}$ é designado por chave.*

Definição 16 (Função de Encriptação)**.** *Cada elemento $e \in \mathscr{K}$ determina, de forma única, uma bijeção de $\mathscr{M}$ para $\mathscr{C}$, designada por $\mathscr{E}_e$ é designada por função de encriptação, ou transformação de encriptação.*

Definição 17 (Função de Desencriptação)**.** *para cada $d \in \mathscr{K}$, $\mathscr{D}_d$ denota a bijeção de $\mathscr{C}$ para $\mathscr{M}$. $\mathscr{D}_d$ é designada por função de desencriptação, ou transformações de desencriptação.*

As funções $\mathscr{D}_d e \mathscr{E}$ devem ser tais que se verifica:

$$\mathscr{D}_d(E_e(m)) = m$$

Definição 18 (Encriptaçao)**.** *o processo de aplicar a transformação E_e a mensagem $m \in M$ é usualmente designado por encriptar m, ou a encriptação de m, onde M é o espaço das mensagens.*

Definição 19 (Desencriptação)**.** *O processo de aplicar a transformação D_d ao texto cifrado $c \in C$ é usualmente designado por desencriptar c, ou a desencriptação de c, onde C é o espaço das mensagens cifradas.*

Definição 20 (Par de Chaves)**.** *As chaves e e d na definição anterior são designadas por par de chaves, e usualmente denotadas por (e,d). Note-se que as chaves podem ser iguais.*

Definição 21 (Esquema de Encriptação—Cifra)**.** *Um esquema de encriptação consiste de um conjunto $\{E_e : e \in K\}$ de transformações de encriptação e um conjunto correspondente $\{D_d : d \in K\}$ de transformações de desencriptação com a propriedade de que para todo o $e \in K$ existe uma chave única $d \in K$ tal que $D_d = E_d^{-1}$, isto é, $D_d(E_e(m)) = m$ para todo o $m \in M$, onde K é o espaço de chaves K, $\{E_e : e \in K\}$ um conjunto de transformações de encriptação e $\{D_d : d \in K\}$ de transformações de desencriptação.*

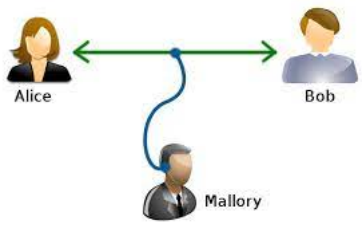

Figura 3.1 Bob e Alice

Exemplificando, tem-se as seguintes etapas na utilização de uma cifra:

Utilização de uma cifra de chaves simétricas

1. Alice e Bob escolhem (secretamente) um par de chaves;
2. Bob decide enviar uma mensagem, $m \in \mathscr{M}$, a Alice. Calcula $c = E_e(m)$ e envia o texto resultante;
3. Ao receber a mensagem a Alice calcula $D_d(c) = m$ recuperando deste modo a mensagem original.

Utilização de uma cifra de chaves assimétricas

1. O Bob escolhe um par de chaves: torna pública a chave de encriptação e, mantém secreta a chave de desencriptação d.
2. A alice decide enviar uma mensagem, $m \in \mathscr{M}$, ao Bob. Obtém a chave pública do Bob e e calcula $c = E_e(m)$. Depois envia o texto resultante.
3. Ao receber a mensagem o Bob calcula $D_d(c) = m$ recuperando deste modo a mensagem original.

Assim, Bob e Alice conseguem comunicar secretamente, sem conceder ao Mallory acesso ao conteúdo da conversa.

Naturalmente todo o código têm duas etapas: uma para codificar a mensagem e outra para decodificar a mensagem. Mas para decifrar uma mensagem, isto é, ler um mensagem codificada sem ser o verdadeiro destinatário dela é preciso "quebrar" o código. Portanto é necessário introduzir a seguinte definição:

Definição 22 (Criptoanálise)**.** *É o estudo dos procedimentos necessários para tentar comprometer as técnicas criptográficas, e mais genericamente, os serviços de segurança da informação.*

Definição 23 (Cifra (parcialmente) Quebrada)**.** *É o ataque sistemático a uma cifra com a finalidade principal de obter texto claro a partir de texto cifrado. Em caso de sucesso, diz-se que a cifra for parcialmente quebrada.*

Definição 24 (Cifra (formalmente) Quebrada)**.** *Tem por objetivo obter a chave privada de uma entidade, nesse caso a cifra é completamente e formalmente quebrada.*

Relativamente as classes de ataques aos esquemas de encriptação, tem-se as seguintes definições:

Definição 25 (Ataque Passivo)**.** *É um ataque em que o adversário apenas monitoriza o canal de comunicação. Um atacante passivo apenas ameaça a confidencialidade da informação.*

Definição 26 (Ataque Ativo)**.** *É um ataque em que o adversário tenta apagar, acrescentar, ou de alguma forma modificar a informação. Um atacante ativo ameaça a integridade da informação assim como a sua confidencialidade.*

Em relação ao tipo de ataque, tem-se as seguintes definições:

Definição 27 (Ataque de texto cifrado)**.** *neste tipo de ataque o criptoanalista tenta deduzir a chave de decifração ou o texto claro por observação unicamente do texto cifrado. Uma cifra que seja vulnerável a este tipo de ataque considera-se completamente inseguro.*

Definição 28 (Ataque de texto claro conhecido)**.** *É um ataque aonde o criptoánlista consegue obter um excerto de um texto claro com base num seu corresponde teste cifrado.*

Definição 29 (Ataque de texto claro escolhido)**.** *É um ataque aonde o adversário escolhe o texto em claro obtendo de seguida o corresponde texto cifrado. Toda a informação daí deduzida é posteriormente usada em outros textos cifrados.*

Definição 30 (Ataque de de força bruta)**.** *É um ataque por procura exaustiva no espaço das chaves. Uma cifra sujeita a este tipo de ataque é designado por cifra fraca.*

Por exemplo, suponhamos que o Mallory, extremamente perspicaz, conseguiu encontrar uma forma de comprometer os métodos de encriptação e desencriptação utilizados pelo Bob e a Alice. Primeiramente, efetuou a monitorização das conversas trocadas pelo Bob e Alice tendo conseguido chegar ao texto cifrado a partir de um certo texto original que conseguiu reunir antes do Bob enviar para a Alice. Apercebendo-se do quanto o conteúdo da informação era preciosa, Mallory quis impedir Alice de ter acesso ao resto do informação. Para isso encontrou a chave certa testando todas as chaves disponíveis. E com isso conseguiu,facilmente, decifrar o conteúdo dos textos cifrados do Bob e passou a enviar falsas correspondências, em nome do Bob, a alice.

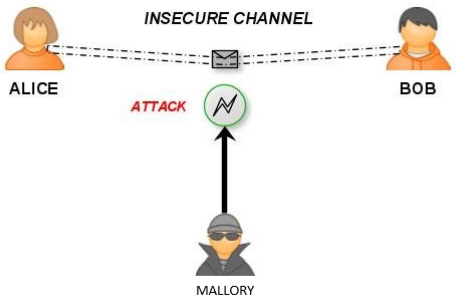

Figura 3.2 Bob, Alice e Mallory

Observação: Num sistema de chave pública é preciso saber a chave pública do destinatário de modo a encriptar uma mensagem a ele destinada. Apenas o destinatário é capaz de decifrar a mensagem.

Nota: Cada entidade tem um par (chave pública, chave privada). De forma a evitar ataques as chaves públicas disponibilizadas pelas entidades, tais como, a substituição da chave pública (este tipo de ataque é designado por personificação) foi criado um repositório de chaves que é capaz de garantir a manutenção das chaves [3].

Capítulo 4

Criptografia Clássica

Designa-se usualmente por *criptografia clássica* as cifras pré-computacionais, desenvolvidas e utilizadas tendo por base processos mecânicos ou manuais. O mais simples deste tipo de criptografia consiste em trocar uma letra pela seguinte. Um código similar foi usado por Júlio César, cuja a chave era estabelecida pelo deslocamento, de três posições, nas letras do alfabeto [1].

Definição 31 (Cifra Deslocamento)**.** *Seja* $\mathscr{M} = \mathscr{C} = \mathbb{Z}_{26}^*$, $\mathscr{K} = \mathbb{Z}_{26}$. *7 Para* $0 \leq K \leq |\ \mathbb{Z}_{26}\ |$*=26, define-se:*

$$e_k(x) = (x+K)\ mod\ 26$$

e

$$d_k(y) = (y-K)\ mod\ 26$$

para todo o $x, y \in \mathbb{Z}_{26}$

Nota: para $\mathbb{Z}_n$ tem-se $n = 26$ uma vez que o alfabeto adotado tem 26 caracteres.

Teorema 14 (Cifra Deslocamento Simples)**.** *As funções* e_k *e* d_k *constituem uma cifra.*

Demonstração

$$\begin{aligned} d_k(e_k(x) &= d_k(x+k) = \\ &= (x+k)-k = \\ &= x\ mod\ 26 \end{aligned}$$

Então, por definição,a cifra de deslocamento é uma cifra. □

Definição 32 (Cifra Deslocamento Linear)**.** *Seja* $\mathscr{M} = \mathscr{C} = \mathbb{Z}_{26}^*$, *e seja:* $K = \{(a,b) \in \mathbb{Z} \times \mathbb{Z} : mdc(a,26) = 1\}$.

Para $K = (a,b) \in \mathscr{K}$*, define-se:*

$$e_k(x) = (ax+b)\ mod\ 26$$

e

$$d_k(y) = a^{-1}(y-b)\ mod\ 26$$

para todo o $x, y \in \mathbb{Z}_{26}$

Teorema 15 (Cifra deslocamento linear)**.** *As funções* e_k *e* d_k *constituem uma cifra.*

Demonstração

$$\begin{aligned} d_k(e_k(x)) &= d_k(ax+b) \\ &= a^{-1}(ax+b-b) \quad \text{por definição } a \text{ é invertível em } \mathbb{Z}_{26} \\ &= a^{-1}ax \\ &= x \end{aligned}$$

Por definição, verifica-se que o resultado apresentado anteriormente é verdadeiro. □

Relativamente à criptoanálise, as cifras clássicas são cifras muito fracas ou completamente inseguras pelo que estão suscetíveis a ataques por procura exaustiva. Também é possível quebrar esta cifra por ataques baseada na frequência relativa das letras, diagramas,trigramas, letras iniciais e finais das palavras.

Por exemplo, se todas as todas as ocorrências da letra *a* são substituídas pela letra *x*, uma mensagem cifrada contendo muitas instâncias da letra *x*, iria sugerir ao criptoanalista, que a letra *x* representa a letra *a*.

De facto, para uma dada linguagem verifica-se que cada letra aparece de acordo com uma frequência própria. No caso da língua portuguesa tem-se a seguinte tabela de frequência: [3]:

a	12,71%	b	0.81%	c	4,16%	d	5,52%	e	11,99%
f	1,43%	g	1,32%	h	0,74%	i	7,18%	j	0,21%
k	0,00%	l	3,23%	m	4,48%	n	5,24%	o	11,32%
p	3,07%	q	1,41%	r	6,47%	s	7,99%	t	5,31%
u	3,44%	v	1,36%	w	0.02%	x	0,28%	y	0,02%
z	0,37%								

Com estes dados, agrupa-se estes valores em grupos, das letras com maior frequência para as menos frequentes:

Primeiro grupo	Segundo grupo	Terceiro grupo	Quarto grupo	Quinto grupo
a,e,o	s,r,i	n,d,m,u,t,c	l,p,v,g,h,q,b,f	z,j,x,k,w,y

Notar que este processo reduz o número de tentativas e erro a realizar antes de se conseguir quebrar o código tornando-o mais eficiente.

Logo para quebrar a cifra utilizada para obter o texto encriptado “a fkdyh whp gh vhu pdqwlgd’, sabendo que foi utilizado um sistema de cifração de deslocamento simples, precisamos de calcular as frequências relativas para o texto cifrado,tem-se:

d	17,9%	f	7,1%	g	7,1%	h	21,4%
k	3,6%	l	3,6%	p	7,1%	q	3,6%
u	7,1%	v	7,1%	w	10,7%	y	3,6%

Por análise das frequências relativas e consultado o primeiro grupo, onde as letras tem maior frequência relativa, uma vez que “d”, com $17,9\%$, e o “h”, com $21,4\%$ então podemos fazer “d” corresponder a “a”, por exemplo. Neste caso, a nossa chave é 3.

Assim, obtemos o texto claro “a chave tem de ser mantida secreta”. Logo a cifra foi quebrada. [3]

Com recurso do computador podemos facilmente realizar todo o processo anterior rapidamente pelo que este tipo de cifra é muito fraca.

Capítulo 5

Criptografia RSA

Metódo de criptografia de chave pública criado em 1978 por R.C. **R**ivest, A. **S**hamir e L.**A**dlemar(RSA), que na época trabalhavam no massachusetts Institute of Tecnology(MIT). Presentemente, é o código de chave pública mais utilizado em aplicações comerciais [1].

Para implementar a criptografia RSA vamos precisar de:

- escolher dois números primos de grande dimensão,[1] p e q;
- conhecer $n = pq$ para codificar a mensagem;
- saber p e q para decodificar a mensagem;

A segurança do metódo advêm da dificuldade que é fatorizar números primos, uma vez que para descobrir a chave de encriptação é preciso fatorizar n de forma a encontra p e q.

Seja $\mathscr{C}_p = (e,n)$ Chave Pública, onde $1 < e < \phi(n)$. Tem-se, pela função de Euler e *teorema 10*, $mdc(e,\phi(n)) = mdc(e,(p-1)(q-1)) = 1$

Seja $\mathscr{C}_s = (d,n)$ Chave Privada, onde d é o inverso multiplicativo de e(Proposição 2), módulo $\phi(n)$.

O algoritmo de encriptação, $\mathscr{A}_{C_p} : M \to \mathscr{A}_{C_p}(M)$, é:

$$C = M^e (mod\ n)$$

O algoritmo de desencriptação, $\mathscr{A}_{C_s} : C \to \mathscr{A}_{C_s}(C) = M$, é:

$$M = C^d (mod\ n)$$

Para que o algoritmo RSA possa ser considerado uma cifra o procedimento tem de ser invertível, isto é,

$$A_{C_s}(A_{C_P}(M)) = A_{C_P}(A_{C_s}(M)) = M^{ed} (mod\ n) = M$$

[1] Size considerations for public and private keys, https://www.ibm.com/docs/en/zos/2.3.0?topic=certificates-size-considerations-public-private-keys

5.1 Teorema RSA

Para provar este resultado vamos precisar da congruência mod m e da consequência que daí se tira que se $a \equiv b(mod\ n)$ então $a = b + kn$, para um dado $k \in \mathbb{Z}$.

Para o desenvolvimento da demonstração são necessários alguns resultados auxiliares, nomeadamente, congruência mod m, *pequeno teorema de Fermat* e *teorema chinês dos restos*.

Teorema 16 (Cifra RSA)**.** *Sendo (e,n) e (d,n) as chaves públicas e privada respetivamente do sistema de Criptografia RSA verifica-se então que:*

$$(m^e)^d (mod\ n) = m$$

para qualquer inteiro m, com $0 \leq m < n$.

Demonstração Da definição de e e d tira-se que $ed \equiv 1(mod\ \phi(n))$ existe então um $k \in \mathbb{Z}$ tal que $ed = 1 + k\phi(n)$, ou seja:

$$ed = 1 + k(p-1)(q-1),\ K \in \mathbb{Z}$$

donde

$$(m^e)^d = m^{ed} = m^{1+k(p-1)(q-1)} = m(m^{(p-1)(q-1)})^k$$

segue-se que

$$(m^e)^d \equiv m(m^{(p-1)})^{(q-1)k} \equiv m(mod\, p)$$

Se p não é um divisor de m esta congruência é uma consequência do *pequeno teorema de fermat*. Caso contrário a asserção é trivial dado que ambos os membros da equação são congruentes com $0\ mod\ p$.

De forma análoga ter-se-ia que:

$$(m^e)^d \equiv m\ (mod\ q)$$

Dado que p e q são números primos distintos pode-se aplicar o teorema chinês dos restos e dado que se assume que $0 \leq m < n$, obtêm-se [3]:

$$(m^e)^d \equiv m\ (mod\ pq) \equiv (mod\ n) = m$$

□

Porque o RSA é seguro?

Seja p e q os parâmetros que estamos a utilizar. Sendo o RSA um método de chaves públicas então a chave de codificação corresponde a chave pública do sistema, ou seja, o par (n,e) é acessível a qualquer utlizador. Logo a segurança do RSA provém da dificuldade do cálculo de d sabendo n e e.

Efetivamente, apenas sabemos calcular d por aplicação do algoritmo de Euclides a $\phi(n)$ e e. Por outro lado, apenas consegue-se calcular $\phi(n)$ se formos capazes de fatorizar n de forma a obter p e q, isto é, apenas é possível quebrar o código se conseguirmos fatorizar n.

Além disso, é impossível alguém criar uma forma de descobrir d sem ter que fatorizar. De facto, a partir de $n = pq$ e $\phi(n) = (p-1)(q-1)$ conhecidos tem-se

$$\phi(n) = (p-1)(q-1) = pq - (p+q) + 1 = n - (p+q) + 1$$

tal que $p+q = n - \phi(n) + 1$ é conhecido. Mas

$$(p+q)^2 - 4n0(p^2+q^2+2pq) - 4pq = (p-q)^2$$

Logo

$$p - q = \sqrt{(p+q)^2 - 4n}$$

Assim sabendo $p+q$ e $p-q$ calcula-se facilmente p e q [1].

5.2 Criptoanálise

Ao longo dos anos, vários métodos de criptoanálise têm sido desenvolvidos para tentar encontrar vulnerabilidades no algoritmo RSA e em outros sistemas criptográficos. Estes incluem ataques de fatoração de inteiros, ataques de timing, entre outros, cada um com vista a encontrar falhas na segurança para comprometer a proteção dos dados [1, 3].

A contínua evolução da criptoanálise e a descoberta de nova formas de ataque salientam a importância de manter algoritmos como o RSA em constante revisão e atualização, de modo a garantir assim a segurança dos dados em um ambiente digital em constante mudança e cada vez mais propenso a ameaças cibernéticas.

Em seguida são apresentados alguns dos métodos de criptoanálise, que em casos específicos, conseguem quebrar a cifra RSA.

5.2.1 Método de Fermat

Proposto Pierre Fermat para encontrar dois inteiros a e b que permitam representar o número natural n como diferença de dois quadrados:

$$n = a^2 - b^2 \leftrightarrow n = (a - b)(a + b)$$

Teorema 17. *qualquer inteiro n ímpar maior que 1 pode ser escrito como a diferença de dois quadrados.*

Demonstração Seja $n = pq$, com $q \leq p$(no caso de n ser primo considera-se $n = n \times 1$). Por hipótese n é ímpar, então p e q também o são, logo: $\frac{p+q}{2}$ e $\frac{p-q}{2}$ são inteiros, ma então temos:

$$\begin{aligned}\left(\frac{p+q}{2}\right)^2 - \left(\frac{p-q}{2}\right)^2 &= \frac{p^2+2pq+q^2}{4} - \frac{p^2-2pq+q^2}{4} \\ &= \frac{p^2+2pq+q^2-p^2+2pq-q^2}{4} \\ &= \frac{4pq}{4} \\ &= pq \\ &= n\end{aligned}$$

□

Para determinar os inteiros a e b de forma que $n = a^2 - b^2$, o processo pode ser conduzido da seguinte maneira:

- Dado um inteiro n ímpar começamos por $a = \lfloor \sqrt{n} \rfloor + 1$
- Se $b = \sqrt{a^2 - n}$ é um inteiro, obtém-se o pretendido
- Caso contrário, incrementamos a de uma unidade até que b seja um inteiro;

Por exemplo, para $n = 2027651281$temos
$a = \lfloor \sqrt{n} \rfloor + 1 = 45030$
Para b obteve-se a seguinte tabela de tentativas [3]:

a	b	a	b
1º45030	$\sqrt{45030^2 - 2027651281^2} = 222,75$	7º45036	$\sqrt{45036^2 - 2027651281} = 768,12$
2º45031	$\sqrt{45031^2 - 2027651281^2} = 373,73$	8º45037	$\sqrt{45037^2 - 2027651281^2} = 824,67$
3º45032	$\sqrt{45032^2 - 2027651281} = 479,31$	9º45038	$\sqrt{45038^2 - 2027651281} = 877,58$
4º45033	$\sqrt{45033^2 - 2027651281} = 565,51$	10º45049	$\sqrt{45039^2 - 2027651281} = 927,49$
5º45034	$\sqrt{45034^2 - 2027651281} = 640,21$	45040	$\sqrt{45040^2 - 2027651281} = 974,84$
6º45035	$\sqrt{45035^2 - 2027651281} = 707,06$	12º45041	$\sqrt{45041^2 - 2027651281} = 1020$

De

$$\begin{cases} \frac{p+q}{2} = 46041 \\ \frac{p-q}{2} = 1020 \end{cases} \tag{5.1}$$

obtém-se $p = 47061$ e $q = 45021$

Assim $n = 45041^2 - 1020^2$, $n = 46061 \times 44021$. Relativamente ao algoritmo de fermat prova-se que,quanto maior for a diferença entre p e q, maior é o número de tentativas que vão ser precisas obter um primeiro valor inteiro para a raiz [3].

5.2.2 Método das Divisões

Neste método, a abordagem consiste em empregar a técnica da fatoração por tentativa, que requer a divisão iterativa por todos os números primos até chegar ao valor de $\lfloor \sqrt{n} \rfloor$ ou até que a solução seja identificada [3].

Para isso precisamos de gerar uma lista de números primos até ao limite pretendido. Presentemente, ainda não é conhecida nenhuma formula para gerar números primos. Assim, é necessário adotar um procedimento que apresente a enumeração exaustiva de todos os números primos até o limite especificado, tal como o algoritmo do Crivo de Eratóstenes.

5.2.3 Crivo de Aristóteles

O método do Crivo de Eratóstenes, criado pelo matemático grego Eratóstenes, representa um algoritmo prático e direto para determinar números primos até um certo limite.

Para exemplificar, vamos listar os números primos de 1 a 30.

- Primeiramente, precisamos determinar $\lfloor\sqrt{30}\rfloor$, o número limite a ser verificado;
- gerar uma lista de inteiros de 2 a 30;
- obter o primeiro número da lista, neste caso é o 2;
- eliminar todos os números da lista, excepto o 2;
- O número sucessor na lista após o primo anterior também é primo.

Obviamente, se repetirmos esse raciocínio até ao final da lista anteriormente gerada, os elementos que não forem eliminados pelas sucessivas aplicações do crivos são números primos de 1 a 30.

Contudo a utilização deste algoritmo adiciona um elevado peso tanto em termos de tempo como de espaço uma vez que é preciso gerar o lista de inteiros de 2 a n e estar a aplicar constantemente o crivo a lista [3].

Alternativamente, existem outros métodos capazes de produzir uma sequência de números primos, assim como números que não são primos com ganhos temporal e espacial comparativamente ao crivo de Aristóteles. Mas são, obviamente, menos eficientes a sua utilização [3].

5.2.4 Método de Euclides

Este metódo ganha o seu nome da utilização do algoritmo de Euclides para o cálculo do máximo divisor comum de dois inteiros. Este algoritmo é muito eficiente e pode ajudar-nos a obter um dos fatores primos de n, de forma a obter o factor primo desejado. Basta multiplicar todos os números primos entre 2 e $\lfloor\sqrt{n}\rfloor$., calcular de seguida o m.d.c entre esse produto e n, de forma a obter o factor primo desejado.

Este método recebe seu nome da utilização do algoritmo de Euclides para calcular o máximo divisor comum de dois números inteiros. Este algoritmo é muito eficaz e pode auxiliar na obtenção de um dos fatores primos de n. É suficiente multiplicar todos os números primos no intervalo de 2 a $\lfloor\sqrt{n}\rfloor$, em seguida, calcular o máximo divisor comum entre esse produto e n para obter o fator primo pretendido.

Com este procedimento, obviamente, ainda vamos continuar com o mesmo problema de elevado peso temporal e espacial em consequência da exigência de criar um registo de todos os números primos até um limite específico.

Por outro lado, vamos ter um problema de representação computacional decorrente dos número obtidos do produto de números primos uma vez que rapidamente o resultado excede a capacidade de representação da maioria das linguagens de programação disponíveis. Para evitar esse problema final, subdividi-se a operação de multiplicação em múltiplas multiplicações menores.

Para tal, os passos seguintes são adotados:

- Começa-se por definir os conjuntos auxiliares:

  $R = r_1, r_2, \ldots, r_n$, representando r_i um limite inferior $(r_1 = 1,\ r_i < r_{i+1})$;

  $S = s_1, s_2, \ldots, s_n$, representando s_i um limite superior$(s_i < si+1, s_{m-1} < \lfloor\sqrt{n}\rfloor < s_m)$;
- Para cada par r_i e s_i, multiplicam-se todos os números primos entre estes dois limites, $P_i = \prod_{r_i \leq p_i \leq s_i} p_i$;
- Para cada um dos P_i calcula-se o $mdc(P-i,n) = a_i$;
- Se $a_i \neq 1$, então a_i é o factor primo de n que se pretende obter [3].

Para exemplificar, seja $n = 1223$. Tem-se $\lfloor\sqrt{1223}\rfloor = 34$, e realizemos a adição de forma iterativa, agrupando de 10 em 10.

$$R = \{1, 11, 21, 31\}\ \ S = \{10, 20, 30, 40\}$$

$P_1 = \prod_{1 \leq 10 \leq s_1} p_1 = 2 \times 3 \times 5 \times 7 = 210,\ mdc(210, 1223) = 1$
$P_2 = \prod_{11 \leq p_2 \leq 20} p_2 = 11 \times 13 \times 17 \times 19 = 46189,\ mdc(46189, 1223) = 1$
$P_3 = \prod_{21 \leq p_3 \leq 30} p_i = 23 \times 29 = 667,\ mdc(667, 1223) = 1$
$P_4 = \prod_{31 \leq p_3 \leq 30} p_i = 31 \times 37 = 1147,\ mdc(1147, 1223) = 31$

Como foi dito anteriormente, embora este método seja muito eficiente, mesmo considerando o problema inicial anteriormente descrito, vai prontamente resultar em problemas de representação devido à capacidade limitada dos tipos de dados disponíveis na maioria das linguagens de programação [3].

5.2.5 Estudo Comparativo dos Vários Métodos

Vimos que do ponto de vista teórico existem abordagens construtivas de decomposição em números primos que podem ser empregadas para alcançar o objetivo. Entretanto, é importante avaliar se existem ferramentas computacionais viáveis para realizar essa quebra de maneira eficiente e prática.

Para analisar isso, uma vez que a análise da complexidade dos algoritmos apresentados está além do âmbito deste texto, simplesmente vamos exibir os resultados de um estudo comparativo dos diferentes métodos com base em um conjunto de testes de execução.

Cada dado foi adquirido de testes executados em condições computacionais uniformes:sistema windowns 10 home,Intel(R) Core(TM) i3-7020U CPU 2.30GHz, RAM 4 GB.

Podemos afirmar que, com a escolha correta dos fatores primos, a cifra RSA permanece protegida. De facto:

> Os métodos de Divisão e Euclides não apenas observam um aumento significativo no tempo de execução à medida que os fatores primos crescem de maneira substancial, mas também perdem a capacidade de resolver o problema a partir de valores relativamente pequenos de

Tabela 5.1 Estudo comparativo dos métodos

n	Fatores	Divisão	Euclides	Fermat
1457	$p = 31, q = 47$	$0,000$s	$0s$	$0s$
13199	$p = 67, q = 197$	$0,002s$	$0.002s$	$0s$
281161	$p = 79, q = 3559$	0.136s	$0.145s$	$0s$
701123	$p = 3559, q = 197$	0.521s	$0.482s$	$0.003s$
23420707	$p = 41017, q = 571$	$71.124s$	$64.86s$	$0s$
488754769	$p = 110503, q = 4423$	–	–	$0.005s$
2027651281	$p = 46061, q = 41017$	–	–	$0s$
103955963689	$p = 47188363, q = 2203$	–	–	$1.75s$
210528952589	$p = 95564663, q = 2203$	–	–	$3.635s$
2746662891777043	$p = 47188363, q = 58206361$	–	–	$0.024s$
4509540007616669	$p = 47188363, q = 95564663$	–	–	$0.299s$

$n = p \times q$. Essas limitações surgem da necessidade de gerar números primos até $\lfloor \sqrt{n} \rfloor$ e, no caso do método de Euclides, da multiplicação de números de grande escala.

O método de Fermat observa um aumento muito acentuado em seus tempos de execução com o crescimento da dimensão dos fatores primos; no entanto, uma análise mais detalhada revela que quando os fatores primos estão em proximidade, o método de Fermat demonstra ser altamente eficaz.

Podemos inferir que, para garantir a segurança da criptografia RSA, os fatores primos devem estar consideravelmente distantes um do outro, e o valor de "n"deve ser maior que 20 dígitos[3].

Efetivamente, o tamanho de n deve ser consideravelmente maior. Devido a algoritmos mais eficazes do que os previamente mencionados, a criptografia RSA foi violada utilizando valores de n contendo 129, 155 e até mesmo 576 dígitos. Atualmente, a cifra RSA emprega valores de n com 1024 dígitos ou mais[3].

Capítulo 6

Conclusões

A cifra clássica representa um ponto de partida fundamental na história da criptografia, servindo como base para o desenvolvimento de métodos mais complexos e sofisticados de proteção de informações. Por meio da Cifra de César, cifra de deslocamento simples, um dos primeiros métodos conhecidos para codificar mensagens, demonstrou-se a importância de ocultar informações sensíveis para protegê-las de terceiros.

No entanto, vimos que essas cifras são, nos dias de hoje, completamente inseguras o que levou a criptografia a evoluir para algoritmos mais complexos tal como o RSA.

O método de RSA é hoje um dos métodos de base nos sistemas critográficos.

Como verificamos no secção 5.2, os métodos aí estudados são incapazes de quebrar o método RSA para chaves com um comprimento superior a 30bits. As chaves seguras actuais para o método RSA estão actualmente na ordem do 2048 bits.[1]

Para a tentativa de quebrar a cifra RSA os métodos actuais são os métodos do crivo quadrático, assim como aproximações usandos computadores quânticos (algoritmo de Shor).

No semestre seguinte, dedicaremos uma análise mais aprofundada à criptoanálise do método RSA, com especial ênfase no crivo quadrático

[1] Ver página: *Size considerations for public and private keys*, https://www.ibm.com/docs/en/zos/2.3.0?topic=certificates-size-considerations-public-private-keys

Bibliografia

[1] Coutinho, S. (2005). *Números Inteiros e Criptografia RSA*. IMPA.

[2] Quaresma, P. and Lopes, E. (2008). Criptografia. *Gazeta de Matemática*, 154:7 – 11.

[3] Quaresma, P. and Pinho, A. (2009). Criptoanálise. *Gazeta de Matemática*, 157:22–31.