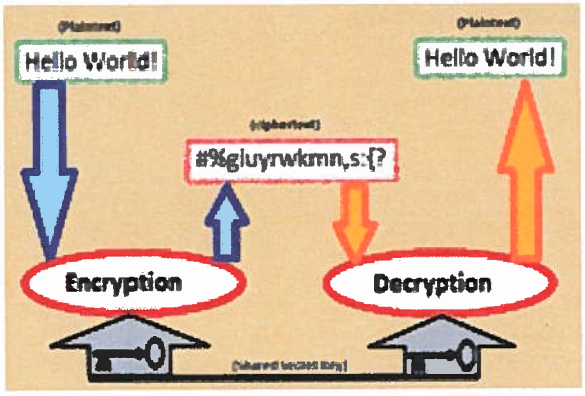
INSTITUTO SUPERIOR DE TECNOLOGIAS AVANÇADAS

Projecto Global



Criptografia e Chaves Públicas (RSA)

Licenciatura:	**Informática - Lisboa - Ano lectivo de 2012/2013**
Nome do Aluno:	**José Pedro Mouta**
Número do Aluno:	**1712**
Coordenador:	**Professor Doutor Nuno Correia**
Orientador Metodológico:	**Professor Doutor Pedro Brandão**
Orientador da Especialidade:	**Professor Doutor Pedro Brandão**

Agradecimentos

Ao meu coordenador, professor Nuno Correia, pela confiança depositada.

Ao meu orientador, professor Pedro Brandão, pelos momentos de apoio e cobrança.

À professora, Cátia Ferreira, pelas suas explicações úteis e pela sua assistência na preparação do manuscrito final.

A todos os professores que me auxiliaram durante a elaboração deste trabalho.

O meu agradecimento especial vai para a minha família, que sempre acreditou na minha capacidade e me apoiou constantemente.

Criptografia e Chaves Públicas (RSA)

Resumo: Este trabalho incide nas origens e evoluçao da criptografia, em geral, e analisa o sistema criptográfico de chave pública RSA, em particular.

O estudo das suas origens e da sua evolução permite descrever os momentos mais relevantes da criptografia e compreender a razão do aparecimento do método RSA. O RSA é um sistema criptográfico de chave pública, criado em 1978 por Ron Rivest, Adi Shamir e Len Adleman.

Define-se criptografia, criptoanálise e outros conceitos que estão relacionados, tais como: criptologia, cifras, chaves e algoritmos. Estudam-se alguns métodos de escrita secreta explicando o seu funcionamento, as suas vantagens e desvantagens. É realizada uma comparação entre as cifras simétricas e assimétricas, analisando os algoritmos em diversos parâmetros, tais como, velocidade de processamento e complexidade computacional. Demonstram-se os principais fundamentos matemáticos que servem de base ao funcionamento do algoritmo RSA. A segurança deste método baseia-se na complexidade dos conceitos matemáticos inseridos na teoria dos números. Abordam-se as etapas do sistema criptográfico RSA, exemplificando o processo de cifrar, decifrar e assinatura digital.

Na parte final deste trabalho, iremos propor as conclusões.

Palavras-chave: Criptografia; Chave Pública; RSA; Criptoanálise; Cifras; Algoritmo; Teoria dos números; Assinatura digital.

Abstract: This paper focuses on the origins and evolution of cryptography in general, and analyzes the cryptosystem RSA public key in particular. The study of its origins and its evolution allows to describe the most relevant moments of cryptography and understand the reason for the appearance of the RSA method. RSA is a public key cryptography system, created in 1978 by Ron Rivest, Adi Shamir and Len Adleman.

Encryption, cryptanalysis and other related concepts are defined, such as cryptology, ciphers, keys and algorithms. The study of some methods of secret writing allow the explanation of their operation, advantages and disadvantages. A comparison between symmetric and asymmetric ciphers is performed by analyzing algorithms in several parameters, such as processing speed and computational complexity. The main mathematical arounds that underpin the operation of the RSA algorithm are demonstrated. The security of this method is based on the complexity of mathematical concepts in theory of numbers. The steps of the RSA cryptosystem are addressed, exemplifying the process of cipher, decipher and digital signature.

Conclusions will be proposed by the end of the paper.

Key word: Cryptography; Public key; RSA; Cryptanalysis; Ciphers; Algorithm; Number theory; Digital signature.

Abreviaturas e Siglas

ACM	*Association for Computing Machinery*
AES	*Advanced Encryption Standard*
ARPA	*Advanced Research Projects Agency*
ARPANET	*Advanced Research Projects Agency Network*
ASCII	*American Standard Code for Information Interchange*
BITS	*Binary Digits* (Dígitos binários)
CBC	*Cipher Block Chaining*
DES	*Data Encryption Standard*
ECB	*Electronic Codebook*
EFF	*Electronic Frontier Foundation*
ENIAC	*Electronic Numerical Integrator and Calculator*
EUA	Estados Unidos da América
IBM	*International Business Machines*
IC	Índice de Coincidência
IK	Índice de Kasiski
MIT	*Massachusetts Institute of Technology*
NBS	*National Bureau of Standards*
NIST	*National Institute of Standards and Technology*
NSA	*National Security Agency*

Índice

Agradecimentos i
Resumo: ii
Abstract: iii
Abreviaturas e Siglas iv
Índice de Figuras vii
Índice de Tabelas viii
Introdução 1
I.Estado da arte 4
1.1 Introdução 4
1.2 Evolução histórica 4
1.3 Conclusão 15
II. Contextualização 17
III. Desenvolvimento 18
3.1 Introdução 18
3.2 Termos e Definições 18
3.3 Criptografia 18
3.4 Tipos de Criptografia 20
3.5 Criptografia Simétrica 20
3.5.1 Substituição 21
3.5.2 Transposição 26
3.5.3 Cifra Simétrica 27
3.5.4 Segurança dos algoritmos simétricos 34
3.6 Criptografia Assimétrica 34
3.7 Criptografia Simétrica vs. Assimétrica 35
3.8 Conclusão 36
IV. Conceitos Matemáticos 37
4.1 Introdução 37
4.2 Divisibilidade 37
4.3 Números Primos 37
4.3.1 Identificação de um número primo 38
4.3.2 Algoritmo para identificar números primos 39
4.4 Potências 40
4.5 Operações com potências 41
4.6 Mínimo Múltiplo Comum (mmc) 42
4.7 Máximo Divisor Comum (mdc) 42
4.7.1 Algoritmo para calcular o Máximo Divisor Comum (mdc) 43
4.8 Algoritmo de Euclides 44
4.9 Inverso Multiplicativo. 47
4.10 Pequeno Teorema de Fermat. 47
4.11 Função de Euler 47
4.12 Conclusão 48
V. Algoritmo RSA 49
5.1 Introdução 49
5.2 Geração das Chaves Pública e Privada no RSA 49
5.3 Cifra e Decifra 50
5.4 Ataques ao RSA 50
5.5 Cuidados necessários em novas aplicações 51
5.6 Assinatura Digital 51

5.7 Conclusão 52
Conclusão 53
Bibliografia 55

Índice de Figuras

Figura 1 - Esquema geral de funcionamento da criptografia 19
Figura 2 - Modelo simplificado de um algoritmo simétrico 21
Figura 3 - Frequência relativa do idioma Inglês 22
Figura 4 - Quadro de Vigenère 24
Figura 5- Estrutura de Feistel para Cifrar 28
Figura 6 - Estrutura de Feistel para Decifrar 29
Figura 7- Fluxograma da cifra DES 30
Figura 8- Fluxograma do sistema criptográfico AES 31
Figura 10 - Transformação *AddRoundKey* da cifra AES 32
Figura 9 - Substituição de bytes para a cifra AES 32
Figura 11 - Tabela de substituição para cifra AES (S-Box) 33
Figura 12 - Transformação *Shift Row* 33
Figura 13 - Transformação *Mix Columns* 33
Figura 14 - Divisibilidade 37
Figura 15 - Tabela de números primos inferiores a 1000 38

Índice de Tabelas

Tabela 1 - Quadro de Cifra Atbash 4
Tabela 2 - Quadro de Políbio 6
Tabela 3 - Arranjos para quebrar Enigma 11
Tabela 4 - Tabela de conversão para chave igual a três 21
Tabela 5 - O método da Força Bruta aplicado à cifra de César 23
Tabela 6 - Cifra de Vigenère 23
Tabela 7 - Quadro para aritmética modular 24
Tabela 8 - Cifra Vigenère pela Aritmética modular 24
Tabela 9 - Método simples para Transposição 27
Tabela 10 - Método mais complexo para Transposição 27
Tabela 11-Comparação de Algoritmos Simétrico e Assimétricos 36
Tabela 12 - Decomposição em factores primos do número 17 38
Tabela 13 - Decomposição em factores primos do número 91 39
Tabela 14 - Resultados de potências definidas por convenção 40
Tabela 15 - Calculo do MDC de 48 e 30 43
Tabela 16 - Calculo do mdc de 195 e 150 44
Tabela 17 - Calculo do mdc de 150 e 195 44
Tabela 18 - Calculo do mdc de 348 e 156 pelo método de Euclides 45
Tabela 19 - Exemplo de calculo de chaves no algoritmo RSA 50
Tabela 20 - Operações do sistema criptográfico RSA 52

Introdução

A segurança da informação sempre foi uma preocupação constante na vida do Homem. Foi no campo militar e diplomático, que o Homem percebeu que tinha de desenvolver uma técnica para enviar mensagens de forma segura. De início era uma arte simples, que utilizava pequenos truques ou fundamentos matemáticos simples.

Ocultar as mensagens foi o primeiro método usado na escrita secreta. É o caso da escrita com sumo de limão que permite esconder a mensagem mas não o seu conteúdo. Ao receptor basta aquecer suavemente a mensagem para ficar visível.

O passo seguinte da escrita secreta foi a criação de métodos de substituição e transposição dos caracteres de uma mensagem. Estes consistem em substituir cada letra da mensagem original por outra do mesmo alfabeto, tornando a mensagem sem sentido. Esta técnica não esconde a mensagem, mas sim o seu conteúdo. Esse método deve ser do conhecimento do emissor e do receptor. Esta transformação impede que a mensagem seja lida por pessoas não autorizadas e permite a sua leitura apenas ao destinatário.

Com a evolução, os truques e os algoritmos apoiados em simples conceitos matemáticos ou lógicos, tornaram-se ineficazes e foram descobertos. Hoje em dia, com o incremento cada vez maior do uso das redes computacionais, a qualidade e controlo da informação tornou-se estratégica e importante para os governos, empresas e até para as pessoas. Para salvaguardar a informação, cada vez mais sofisticada, armazenada e transmitida em meios computacionais, foram criados algoritmos cada vez mais elaborados. Vários foram os algoritmos criados e mantidos em segredo, mas praticamente quase todos foram sucumbindo. O desenvolvimento de métodos para ver mensagens secretas não autorizadas é bastante antigo.

O ano de 1976 foi um momento muito importante para a criptografia. Surgiu um artigo *New Direction in Criptography*, de Whitfield Diffie e Martin Hellman, que foi a base de apoio para o novo algoritmo de chaves assimétricas. O conceito das chaves assimétricas, descrito no documento de Diffie e Hellman e estudado por Rivest, Shamir e Adleman, foi muito importante para o aparecimento de um novo algoritmo, que permitia a escrita secreta de forma muito segura, utilizando conceitos matemáticos. RSA foi a designação dada ao novo algoritmo e deve a sua denominação às letras iniciais de cada um dos nomes dos seus criadores.

A metodologia deste trabalho é baseada em pesquisa bibliográfica na internet e livros, relacionados com a criptografia e o método RSA.

Este trabalho tem como objectivo apresentar um estudo sobre a criptografia em geral e, em particular, sobre o método RSA.

Os objectivos específicos deste trabalho são: mostrar a evolução histórica da criptografia desde os tempos antigos até ao aparecimento do sistema criptográfico RSA; demonstrar a importância da criptografia na segurança da informação que circula nos canais inseguros, tal como a internet; destacar na evolução da criptografia a constante disputa entre criptógrafos e criptoanalistas, tornando o seu estudo cada vez mais complexo e essencial; abordar e mostrar o problema da distribuição das chaves secretas nos sistemas criptográficos simétricos; descrever alguns detalhes técnicos sobre os algoritmos DES e AES, muito importantes para a compreensão deste tema; explicar porque razão apareceu o método RSA; explicar as razões pelas quais este método é considerado seguro; comparar o desempenho entre os métodos simétrico e assimétrico. A análise comparativa é realizada através de parâmetros tais como: desempenho, velocidade, tamanho da chave, nível de segurança e distribuição de chaves; identificar e explicar os conceitos matemáticos que serviram de suporte ao sistema criptográfico RSA; codificar e descodificar, através de um exemplo, uma pequena mensagem usando o sistema criptográfico RSA.

O primeiro capítulo documenta a evolução da arte até ao estado actual. Neste capítulo serão abordados alguns métodos e tipos de escrita secreta, explicando o seu funcionamento e os seus pontos fortes e fracos, com apoio a citações, estudos e artigos científicos.

O segundo capítulo contextualiza o assunto do trabalho.

No terceiro capítulo, são abordados os conceitos de criptografia essenciais a este trabalho, como cifras de substituição, simétricas e assimétricas, além das suas formas de criptoanálise.

O capítulo seguinte, o quarto, analisa os conceitos matemáticos que servem de base ao funcionamento do algoritmo RSA. Este capítulo permite compreender a essência e importância do algoritmo e também encontrar argumentos para reflectir nos seus pontos fortes e fracos.

O quinto e último capítulo mostra como funciona o algoritmo RSA. Passo a passo, explica-se como se converte ou codifica uma mensagem e como se descodifica a mesma mensagem.

Finalmente pode-se dizer que a grande motivação para elaborar este trabalho foi de ordem pessoal e profissional. Pessoal por ser um assunto interessante e profissional pois é um algoritmo usado na minha actividade profissional.

I.Estado da arte

1.1 Introdução

A batalha entre codificadores e descodificadores é antiga e em constante evolução. Os codificadores criam códigos cada vez mais seguros e os descodificadores criam métodos para anular a eficácia dos códigos. Qualquer dos ramos se serve da matemática e outras tecnologias para dar apoio a estes estudos. Faz sentido falar da evolução da teoria dos números, pois foi uma área da matemática que despertou do longo período de esquecimento, sendo de muita importância para o aparecimento do algoritmo moderno RSA.

Este capítulo descreve, desde a antiguidade até à actualidade, os factos mais importantes, por ordem cronológica, para o aparecimento do sistema criptográfico RSA e a constante disputa entre codificadores e descodificadores.

O estudo da criptografia através dos tempos é necessário, pois as técnicas empregues nos primeiros algoritmos criptográficos ainda hoje são aplicadas nos algoritmos modernos.

1.2 Evolução histórica

A evolução da criptografia acredita-se ter começado há milhares de anos. Os símbolos usados no Egipto, por volta de 4500 a.C., chamados hieróglifos, podem-se considerar mensagens misteriosas. O mesmo se terá passado na antiga Mesopotâmia com os caracteres cuneiformes (Kahn, 1967; Singh, 2006).

Por volta de 600 a 500 a.C, atribui-se aos hebreus a criação e o uso de cifras de substituição monoalfabéticas. A cifra *Atbash* é um exemplo e usava um método que consistia em trocar a primeira letra do alfabeto o "A", pela última letra do alfabeto o "Z". Esta cifra converte a palavra "OLA" para "LOZ" (Cohen, 1995; Singh, 2006).

Tabela 1 - Quadro de Cifra Atbash
Fonte: (Stallings, 2005)

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

A cifra de transposição foi usada para fins militares pelos Espartanos e o método mais usado para enviar mensagens seguras foi o dispositivo *Scytale*[1] (Kahn, 1967; Mollin, 2007). Na transposição, as letras da mensagem são reorganizadas, gerando um anagrama[2]. Para mensagens, com uma única palavra, este método é muito inseguro, apresenta um número muito limitado para reordenar as letras. Nas mensagens longas o número possível de arranjos é muito grande tornando-se impossível obter o texto original (Singh, 2002, 2006).

Este método basicamente muda a ordem das letras, ou seja, não altera os caracteres originais mas sim a sua posição de acordo com uma regra. A cifra transposição é fácil de reconhecer porque o texto original têm a mesma frequência de letras que o texto cifrado. Em função do número de letras da mensagem esta reorganização pode-se tornar muito complicada. Uma palavra com 35 letras pode ter 50×10^{30} possibilidades de arranjos. Esta reorganização deve ser combinada entre o emissor e o receptor (Singh, 1999, p. 7).

Um exemplo interessante da cifra esteganográfica é o relato no livro *The Histories* onde Herodotus[3] (485-425 a.C.) narrou os conflitos entre a Grécia e a Pérsia ocorridos no século V a.C. Segundo Hérodoto, foi a escrita secreta que salvou a Grécia de ser conquistada por Xerxes[4] (520-465 a.C.), o líder dos persas. O livro aborda várias técnicas para esconder mensagens e o sistema mais analisado é a esteganografia. Esta técnica oferece alguma segurança mas sofre de uma fraqueza quando a mensagem é interceptada e a sua captura torna a mensagem legível. A criptografia não esconde a existência de uma mensagem mas sim o seu conteúdo (Kahn, 1967).

Os gregos na antiguidade usavam a palavra aritmética para estudar os números, as suas propriedades e a relação entre eles. Hoje em dia designa-se por teoria dos números. O interesse pelos números primos cresceu recentemente graças à sua aplicação na criptografia (Mollin, 2007).

Euclides, por volta de 300 a.C., foi dos primeiros a fazer estudos sobre os números primos e demonstrou que existem infinitos números primos. Estes estudos estão publicados na sua obra *The Elements*[5]. Os números primos têm esse nome devido

[1] *Scytale* é um dispositivo para esconder mensagens e consiste num bastão de madeira longa e estreita, ao redor do qual se enrola firmemente uma tira de couro ou pergaminho. Escreve-se a mensagem no sentido do comprimento do bastão, a tira é desenrolada e a mensagem fica cifrada.
[2] Palavras ou frases feita com as letras de outras. http://www.priberam.pt/anagrama.
[3] "As Historias" de Hérodoto.
[4] "Rei dos Reis".
[5] O livro Os Elementos e a Bíblia são os livros mais reproduzidos da história (Mollin, 2007, p. 3).

aos gregos, que dividiam os números em primeiros ou secundários. O termo em latim *primus*[6], significa primeiro. Estes e outros conceitos que se vão abordar servem de suporte ao algoritmo RSA (Mollin, 2007).

Eratosthenes[7] (284 - 204 a.C.), criou uma das primeiras técnicas para encontrar número primos, o chamado *Sieve of Erathosthenes*[8]. Por exemplo pretende-se encontrar todos os números primos até 30. Faz-se uma lista de todos os números naturais maiores que um e menores que 30. Da lista são retirados todos os números múltiplos de dois superiores a dois e múltiplos de três maiores que três. A esta lista reduzida volta-se a retirar todos os múltiplos de cinco maiores que cinco. Restam os números primos menores de 30. Este sistema funciona bem mas é pouco eficiente (Mollin, 2007, p. 6).

Ao grego Polybios foi-lhe atribuído a criação de um método criptográfico chamado "*Polybios Square*"[9] (203 – 120 a.C). Trata-se de uma cifra de substituição e monoalfabética, em que cada letra pode ser representada por dois números. O quadro de Políbio, na Tabela 2, dá-nos o número da linha e o número da coluna de cada letra do alfabeto inglês. Para exemplificarmos, a conversão da letra "L" é igual ao número 31 e a da letra "H" é 23 (Kahn, 1967).

Tabela 2 - Quadro de Políbio
Fonte: (Kahn, 1967; Mollin, 2007)

	1	2	3	4	5
1	A	B	C	D	E
2	F	G	H	I/J	k
3	L	M	N	O	P
4	Q	R	S	T	U
5	V	W	X	Y	Z

A alternativa à transposição é a substituição e o primeiro uso documentado de uma cifra de substituição para fins militares foi narrado por *Suetonius*.

No livro *The Lives of the Twelve Caesars*, Suetonius fala do método usado pelo imperador romano Júlio César para mascarar as suas mensagens (Kahn, 1967; Stinson, 1995).

[6] http://www.infopedia.pt/primos
[7] Erastóstenes
[8] Crivo de Eratóstenes
[9] Quadrado de Políbio

Este método, conhecido pela cifra de César, consiste em substituir uma letra do texto original por outra n casas à frente. O número n é a chave e representa a distância entre o símbolo original e o cifrado. Para converter é usada uma técnica de substituição monoalfabética com uma chave igual a qualquer número entre um e 26.

Com o colapso do império romano, a Europa mergulhou num período de obscurantismo. As artes e ciências foram esquecidas e a criptografia não escapou à regra. A criptografia entre o ano 500 e 1400 estagnou (Kahn, 1967).

O ano 750 foi a época dourada da civilização islâmica, com vários trabalhos realizados nas artes e nas ciências. Através da matemática, estatística e linguística os árabes inventaram a análise criptológica ou análise de frequências (Kahn, 1967).

Pensava-se que as cifras de substituição monoalfabéticas nunca seriam quebradas, até que os árabes descobriram um método para descodificar estas cifras. Após algum estudo descobriram que algumas letras se repetiam mais que outras. O idioma árabe repete muitas vezes a letra "A" e "L" e o "J" tem uma frequência menor dez vezes. Esta técnica, conhecida por análise de frequências e que parece pouco importante, na realidade foi um passo de gigante na técnica da criptoanálise e permite achar a chave que foi usada na cifra do texto (Kahn, 1967; Singh, 2006):"*Cryptology was born among the Arabs. They were the first to discover and write down the methods of cryptanalysis.*"[10] (Kahn, 1967, p. 76).

O cientista árabe do século IX, Al-Kindi, é o autor do livro mais antigo conhecido em criptologia. Escreveu o livro intitulado *Risalah fi Istikhaj al-Um'amma*[11], que descreve as primeiras técnicas da criptoanálise. Manuscritos antigos descobertos recentemente revelam que a origem da criptologia e os contributos dados pelos árabes são mais antigos e mais extensos do que se pensava (Al-Kadi, 1992).

Durante muitos anos a cifra de substituição monoalfabética tinha sida robusta para garantir o sigilo. O desenvolvimento posterior da análise de frequências, pela primeira vez pelos Árabes, tornou esta cifra frágil. Destruiu a sua segurança e veio mostrar que as técnicas de criptoanálise tinham atingido um nível superior às técnicas de criptografia.

Era a vez dos criptógrafos começarem a estudar e criar uma nova cifra mais forte que pudesse enganar os criptoanalistas. Leon Battista Alberti (1404-1472) lançou as

[10] "A criptologia nasceu entre os Árabes. Eles foram os primeiros a descobrir e a narrar os métodos da criptoanálise".
[11] "Manuscrito para decifrar mensagens criptográficas"

bases da futura cifra polialfabética ao propor o uso de dois ou mais tipos de alfabetos para cifrar as mensagens e assim tornar o texto cifrado mais confuso. Foi considerado o pai da criptografia ocidental (Kahn, 1967; Singh, 2002).

Outro estudo importante para a cifra polialfabética foi a obra *Poligrafhia* de Johannes Tritheminus (1462-1516). Foi ele que pela primeira vez usou a palavra esteganografia (Mollin, 2007, p. 82). É uma técnica criptográfica que permite esconder uma mensagem dentro de outra ou num suporte. Por convenção a esteganografia é considerada um ramo da criptografia.

Vigenère (1523-1596) leu os livros de Tritheminus, Belasco, o manuscrito de Alberti e outros escritores. Este conhecimento ajudou-o a criar uma nova cifra conhecida por cifra Vigenère. O documento "Traicté des chiffres", publicado em 1585, foi um contributo importante para a criptografia (Kahn, 1967; Mollin, 2007; Singh, 2006).

A cifra de Vigenère é forte à análise de frequências e altamente robusta à força bruta quando a chave é grande. A letra "V" é convertida para várias letras "Q", "V", "M" e "C". O criptoanalista também tem dificuldade em decifrar as mensagens pela força bruta, pois as chaves podem ter várias combinações de palavras. De momento uma vantagem clara sobre os criptoanalistas.

Depois de alguns séculos sem qualquer tipo de estudo sobre os números primos, surge Pierre Fermat (1601-1665). O seu teorema, conhecido pelo pequeno teorema de Fermat, é a base de muitos outros trabalhos na teoria dos números e ainda hoje é empregue (Mollin, 2002, 2007).

Vários matemáticos estudaram o sistema binário mas foi documentado de uma forma mais abrangente por Gottfried Leibniz (1646-1716). O sistema de Leibniz utilizou o zero e um, tal como o sistema numérico binário actual (Burnett & Paine, 2002).

Mais de cem anos depois de Fermat, Leonhard Euler (1707-1783) trouxe um novo avanço na teoria dos números. Ele trabalhou o pequeno teorema de Fermat e mostrou outro teorema denominado função de Euler. Este conceito é utilizado na geração da chave pública e privada do algoritmo criptográfico assimétrico RSA (Mollin, 2002, 2007).

A teoria dos números tem papel central na existência do RSA e uma das ferramentas mais importantes na teoria dos números é a aritmética modular, que envolve o conceito de congruência. Uma congruência é a relação entre dois números

que divididos por um terceiro, chamado módulo de congruência, deixam o mesmo resto. Por exemplo, o número nove é congruente ao número dois, modulo sete, pois ambos deixam o resto dois ao serem divididos por sete. A sua representação simbólica é a seguinte: $9 \equiv 2 \bmod 7$. Foi o notável matemático Carl Friedrich Gauss (1777-1855) que estudou este conceito que está relacionado com a divisibilidade e o resto de uma divisão de números inteiros (Burnett & Paine, 2002; Koblitz, 1994).

Surge, em 1854, a álgebra de boole em homenagem a George Boole (1815-1864), que introduziu o cenceito de operação lógica, muito importante para os computadores tomarem decisões. Existem vários operadores lógicos (*AND*, *NAND*, *OR*, *XOR* e *NOT*), mas o mais usado na criptografia é o operador lógico XOR. O "Ou exclusivo" ou *XOR*, é uma operação lógica entre dois operandos que devolve um valor lógico verdadeiro, quando apenas um dos operandos é verdadeiro (Burnett & Paine, 2002).

Desde sempre foi difícil confirmar se os números grandes são primos. Em 1876, François Édouard Lucas (1842-1891) usando o seu método chamado teste de Lucas, demonstrou que 127 era um número primo e foi considerado o maior número primo até 1952. Em 1952, com a era dos computadores, foi possível mostrar que havia mais números primos (Mollin, 2002, 2007).

A cifra de Vigenère foi, durante muitos anos, considerada uma cifra robusta e inquebrável e ficou conhecida por "Cifra indecifrável". Durante o curso da Segunda Guerra Mundial, Babbage deu um grande contributo à criptoanálise ao quebrar a cifra de Vigenère (Singh, 1999). Assim como Babbage, F.W.Kasisski (1805-1881) criou um método para quebrar a cifra de Vigenère (Mollin, 2007; Stinson, 1995).

O método foi desenvolvido por volta de 1863 e quebrou quase todas as cifras da época. Consistia em encontrar repetições de sequência de caracteres no texto cifrado. Procuram-se conjuntos de letras repetidas no texto cifrado e contam-se as distâncias entre eles para obter o comprimento da chave. Ao calcular o maior divisor comum dos valores das distâncias de repetição, permite calcular o comprimento da chave. Estas repetições por vezes ocorrem por puro acaso, o que obriga a várias tentativas para encontrar o verdadeiro comprimento da chave, mas é muito mais eficaz que qualquer outro método (Mollin, 2007).

Em 1883, Kerckhoffs von Nieuwenhof, escreveu no seu livro *La Criptographie militaire*, *"Kerckhoffs' Principle: The security of a crypto-system must not depend on*

keeping secret the cryptoalgorithm, but only on keeping secret the key." (Singh, 2006, p. 2).

Segundo Diffie Whifield, em 1920 foi escrito por William F. Friedman (1891-1969), um dos mais importantes artigos para a criptoanálise do século XX, a monografia "The Index of Coincidence and Its Application in Cryptography"[12] (Mollin, 2007, p. i).

No final do século XIX, de novo a criptografia volta a estar em desvantagem na disputa constante com a criptoanálise. Com a quebra da cifra de Vigenère, por Babbade e Kassiski, os criptográficos voltam a reunir esforços para encontrar novos sistemas de criptografia seguros.

Após a primeira guerra mundial e para substituir os sistemas criptográficos pouco eficazes, foi criado o sistema de cifra chamado Enigma. A máquina criptográfica Enigma foi construída pela empresa *Scherbius & Ritter* e os donos da empresa eram Arthur Scherbius e Richard Ritter. Foi usada pelo exército alemão para codificar e descodificar as suas mensagens militares.

A configuração do engenho consiste em três elementos ligados por fios: um teclado para introdução do texto original, uma unidade central de encriptação que converte cada letra original para uma letra cifrada e várias lâmpadas para mostrar a letra cifrada (Kahn, 1967).

A unidade de encriptação é formada por rodas[13]. Cada cilindro tem 26 pinos de entrada e 26 pinos de saída. No interior de cada cilindro juntam-se os pinos de entrada com os pinos de saída formando um circuito eléctrico. Por exemplo o pino da letra "A", de entrada, pode ser ligado ao pino da letra "X" do lado da saída. Para a criptoanálise uma máquina com um cilindro não apresenta qualquer tipo de dificuldade, corresponde a uma cifra de substituição monoalfabética, mas a complexidade aumenta com o uso de dois ou mais cilindros. Na Tabela 3, podem-se ver os arranjos possíveis em relação ao número de cilindros instalados na máquina Enigma.

Os aliados só conseguiram quebrar os códigos do Enigma, após a captura de uma dessas máquinas (Singh, 2002; Stallings, 2005).

---

[12] O livro "O Índice de Coincidências e a sua Aplicação na Criptografia " descreve o Índice de Coincidências (IC) e como ele pode ser aplicado em criptoanálise polialfabética.

[13] Rodas ou cilindros ou rotores

Tabela 3 - Arranjos para quebrar Enigma
Fonte: (Stallings, 2005)

Enigma - Força Bruta		
Cilindros	Probabilidades	
	Parcelas	Total
2	26 × 26	676
3	26 × 26 × 26	17576
4	26 × 26 × 26 × 26	456976
5	26 × 26 × 26 × 26 × 26	11881376

Durante a Segunda Guerra Mundial foi constante o confronto entre criptoanalistas alemães e britânicos. Os britânicos inventaram um dispositivo chamado *Colossus,* para combater a cifra Lorenz, usada nas comunicações alemãs. Após a guerra, o *Colossus* e tudo em *Bletchley Park*, foi destruído e todos os trabalhadores que participaram na sua elaboração foram proibidos de falar sobre o projecto. *Bletchley Park* era o nome das instalações secretas britânicas, onde se juntaram grandes criptoanalistas para decifrar as mensagens secretas alemãs.

Em 1945, J.P. Eckert e J.W. Mauchly completaram o computador ENIAC (*Electronic Numerial Integrator and Calculator*). A era dos computadores começou a despertar e a eterna batalha entre criptografia e criptoanálise, tem agora mais um instrumento de trabalho muito útil. Os criptógrafos podem empregar a tecnologia dos computadores para criar algoritmos mais complexos e chaves maiores. Por sua vez os criptoanalista podem explorar a velocidade e flexibilidade de programação dos computadores, para pesquisar todas as chaves possíveis até encontrar a correcta.

C.E. Shannon (1916-2001), em 1948 publicou um artigo intitulado "Communication Theory of Secret Systems"[14], onde fala sobre a melhor forma de codificar a informação para enviar. Este documento foi publicado no *Bell System Journal*, em 1949. No ano seguinte, junto com outro grande matemático Warren Weaver (1894-1978) escreveu o livro *A Mathematical Theory of Communication*[15] (Stinson, 1995).

Consciente que no futuro a comunicação entre computadores haveria de ser uma realidade, a IBM nos anos sessenta iniciou um estudo para criar um algoritmo criptográfico que permitisse a protecção de dados. Estes estudos deram origem a um

[14] "A Teoria da Comunicação de Sistemas Secretos"
[15] "Teoria Matemática da Comunicação"

novo algoritmo chamado *Lúcifer*, desenvolvido por Horst Feistel (1915-1990) e o seu grupo (Feistel, 1973).

Em 1972, o governo norte-americano reconheceu a necessidade de criar um algoritmo de criptografia capaz de proteger a informação não confidencial mas sensível.

Os requisitos do projecto eram: Alto nível de segurança, fácil de entender, a sua segurança deve residir na chave e não depender do sigilo do algoritmo, disponível para todos os utilizadores e adaptável a todas as aplicações.

O primeiro concurso foi lançado em 1973, mas nenhuma proposta foi considerada aceitável. A segunda fase decorreu em 1974 e desta vez a IBM entregou uma proposta que foi considerada válida. A nova técnica criptográfica apresentada e baseada no antigo algoritmo *Lúcifer* de Horst Feistel, denominava-se DES (Feistel, 1973).

O algoritmo foi inicialmente muito controverso, com uma pequena chave e suspeitas de um *trap door*[16] da NSA. Martin Helman e Whitfield Diffie, os pioneiros da criptografia assimétrica, criticaram o pequeno tamanho da chave e as misteriosas caixas de substituição[17], como uma evidência de interferência da NSA na feitura do algoritmo (Schneier, 1996). Apesar das críticas, o DES acabou por ser aceite como padrão pelo governo dos EUA e ficou patenteado com a referência FIPS PUB 46. Na década de 1970, em 1977, a disponibilização do algoritmo DES para instalações não governamentais, converteu a criptografia numa arte global. As empresas financeiras, bancos e outras podem a partir de agora, fazer os seus negócios por canais públicos de maneira segura (Schneier, 1996).

O DES surgiu como o primeiro grande sistema criptográfico e por isso concentrou sobre si todos os estudos académicos e privados da criptografia. Nas palavras do especialista em segurança Bruce Schneier, "DES did more to galvanize the field of the cryptanalysis than anything else, now there was a algorithm to study."[18] (Schneier, 1996, p. 267). O DES encripta blocos de 64 bits de entrada e devolve 64 bits à saída. É um algoritmo simétrico, a mesma chave serve para cifrar e decifrar.

O comprimento da chave é de 56 bits. Por cada oito bits, o último, o menos significativo, é ignorado. A estrutura do algoritmo começa por uma permutação inicial e uma final, que são opostas no seu funcionamento. Entre as permutações são executados

[16] Porta dos fundos
[17] S-Box
[18] "O DES fez mais pelo campo da criptoanálise de que qualquer outro, agora há um algoritmo para estudar"

16 estágios ou voltas de processamento repetidos (Menezes, van Oorschot, & Vanstone, 1996; Schneier, 1996).

Com o desenvolvimento tecnológico aparecem computadores capazes de testar milhões de chaves por segundo e a custos relativamente baixos para grandes empresas e governos. Os criptógrafos sabiam que o sistema criptográfico DES não podia durar muito tempo e perante esta ameaça o governo americano reconhece a necessidade de criar um novo algoritmo, o 3DES. É uma simples variação do DES, basicamente é usar o DES original três vezes, com duas ou três chaves distintas. É um algoritmo seguro, porém muito lento para ser considerado um algoritmo padrão (Schneier, 1996).

Uma nova técnica de criptoanálise aparece em 1990, a criptoanálise diferencial, descoberta pelos israelitas Eli Biham e Adi Shamir (Schneier, 1996).

Em 1997, o governo americano, através do NIST, lançou um concurso para criar um novo algoritmo que pudesse substituir o DES. O AES surge porque havia necessidade de escolher um algoritmo mais robusto, que corrigisse as debilidades do DES. Foi o tamanho pequeno da sua chave e as várias propostas de máquinas para quebrar o DES, que originaram a sua substituição. Em Setembro de 1997, publicaram-se os requisitos necessários para o novo algoritmo: divulgação pública, algoritmo simétrico (chave privada), cifra em bloco de 128 bits, aceitar chaves com 128, 192 e 256 bits de comprimento e maior rapidez em relação ao DES e 3DES. O concurso decorreu entre 1997 e 2000 e teve duas fases. Na primeira conferência apresentaram-se 15 candidatos e apenas cinco continuaram a disputa no segundo congresso. A luta final foi entre os algoritmos criptográficos *MARCS*, *RC6*, *Rijndael*, *Serpent* e *Twofish*. Em 2000, após análise rigorosas de especialistas criptográficos, é conhecido o vencedor: *Rijndael*. O algoritmo é uma melhoria da cifra *square* e os seus criadores foram os Belgas Vincent Rijmen e Joan Daemen (Dent & Mitchell, 2005; Singh, 2002).

A entrega das chaves aos seus destinatários sempre foi um processo pouco prático e tem atormentado os criptógrafos. A maneira mais segura é entregá-la na mão e a menos segura mas mais prática é através de mensageiro. Por estas vias, o fornecimento regular das chaves a todos os receptores válidos de um emissor, é um grande problema logístico, principalmente para destinos muito distantes.

A distribuição das Chaves pode parecer fácil mas trata-se de um problema primordial para os criptógrafos da pós-guerra. Em meados da década de 1970, surgiu um grupo de cientistas com uma ideia brilhante para resolver o problema de distribuição

de chaves. De facto esta descoberta é considerada a maior conquista da criptografia do século XX (Singh, 2002).

De volta ao ano de 1960, o departamento de defesa dos EUA começou a financiar uma organização chamada ARPA e um dos projectos era encontrar um sistema para ligar os vários computadores militares à distância.

Em 1982, a ARPANET, que criou a internet em finais de 1980, disponibilizou esta tecnologia para entidades não governamentais (Singh, 2002).

Whitfield Diffie previu que as pessoas comuns um dia teriam os seus próprios computadores e que estariam interligadas entre si através de linhas telefónicas. Esta configuração permitiria aos seus utilizadores trocar entre si mensagens e tinham o direito de o fazer através de mensagens seguras. Para garantir a privacidade, a criptografia necessitava de um meio ou mecanismo para troca segura de chaves. Em 1974, Whitfield Diffie, Martin Hellman e Ralph Merkle juntam-se para estudar o problema da distribuição de chaves. Para troca de chaves segura este grupo de trabalho imaginou esta situação. Como enviar uma mensagem secreta entre o emissor Alice e o receptor Bob, sem haver troca de chave (Singh, 2002).

Alice coloca a sua mensagem numa caixa de ferro com cadeado e envia para Bob. O Bob acrescenta mais um cadeado à caixa e devolve à Alice. Alice recebe a caixa com dois cadeados. Remove o seu cadeado, deixa ficar o de Bob e envia de novo a caixa para o Bob. Ao receber a caixa, Bob pode abri-la com a sua chave e ler a mensagem da Alice (Singh, 2002).

Este exemplo demonstrou que era possível o envio de mensagens secretas sem a troca de chaves. Porém, este processo tem um senão que é a ordem das cifras e decifras, ou seja, a última cifra deve ser a primeira a ser decifrada. No nosso exemplo o Bob foi o último a aplicar a cifra e deve ser o primeiro aplicar a decifra. O problema da distribuição da chave continuava por resolver, mas este estudo foi muito importante para a fase seguinte. Em 1975, Diffie volta a ter outra ideia brilhante e inventa uma nova cifra, que assenta numa chave assimétrica. Neste sistema de chave assimétrica, as chaves para cifrar e decifrar não são idênticas. Este sistema não foi aplicado, contudo foi revolucionário (Singh, 2002).

Em 1977, Rivest, Shamir e Adleman, resolveram estudar o processo para criar uma cifra assimétrica. Assim nasceu a cifra assimétrica RSA. Este algoritmo trabalha com uma chave pública, do domínio público e uma chave privada, apenas do conhecimento do emissor (Singh, 2002).

O algoritmo RSA é publicado, em 1978, na ACM (*Association for Computing Machinery*), num dos melhores meios de divulgação de pesquisas científicas. Os detalhes desta organização podem ser lidos no link http://www.acm.org. O sistema criptográfico assimétrico também pode ser usado para verificar a originalidade das mensagens. Pode-se ao enviar uma mensagem querer também assiná-la digitalmente, tal como se assina um cheque ou um documento. O sistema criptográfico RSA foi o primeiro sistema criptográfico a usar este método (Mollin, 2002).

"*One of the most significant contributions provided by public-key cryptography is the digital signature*"[19] (Menezes, van Oorschot, & Vanstone, 1996, p. 2). A assinatura digital é o resultado da encriptação de dados ou documento, pela utilização da chave privada de um sistema criptográfico assimétrico. A sua validação é feita decifrando o documento com a chave pública correspondente. Se o resultado for válido, a assinatura é autêntica, uma vez que apenas o detentor da chave privada, par da chave pública utilizada, poderá ter gerado aquela assinatura (Rivest, Shamir, & Adleman, 1978).

Em 1998, o hardware chamado *Deep Crack* construído pela EEF, por cerca de $250,000, conseguiu quebrar o DES em 56 horas[20] (Dent & Mitchell, 2005; Foundation, 1998; Singh, 2002).

Um grupo da Web em todo o mundo, organizado pela Distributed.net[21] e a EFF em 1999, juntou as suas tecnologias para quebrar o DES em 22 horas e 15 minutos.

Em 2001 o DES foi oficialmente substituído pelo AES. Após um longo concurso o NIST seleccionou o *RIjndael*. Em 2006, o AES, anunciado pelo NIST, tornou-se um padrão oficial de criptografia com a referência FIPS PUB 197 (Burnett & Paine, 2002; Dent & Mitchell, 2005; Singh, 2002).

1.3 Conclusão

Neste capítulo analisou-se a evolução da criptografia desde as suas origens até ao aparecimento do sistema criptográfico RSA, que representa, para este trabalho, o estado actual da arte criptográfica.

[19] " Uma das contribuições mais importantes prestada pela criptografia de chave pública é a assinatura digital"
[20] http://www.eff.org/descracker.html
[21] http://www.distributed.net

Durante este período, vários foram os momentos importantes que contribuiram para o seu desenvolvimento. Contudo, considera-se que os dois marcos mais relevantes para criptografia pública tiveram lugar em 1976.

O primeiro foi a publicação do DES pelo governo american, um algoritmo aberto de criptografia simétrica, selecionado pelo NIST num concurso onde foi escolhida uma variante do algoritmo Lucifer, proposto pela IBM. O DES foi o primeiro algoritmo de criptografia disponibilizado abertamente ao público. A partir desse momento houve um acrescimo de estudos académicos públicos sobre criptografia.

Pela primeira vez o estado da arte da criptografia é praticado e estudado pelo meio público, fora do controlo exclusivo das entidades militares e governamentais. É possível empregar as técnicas de segurança que nos protegem dos nossos adversários.

O segundo foi a publicação do artigo *New Direction in Criptography*, de Whitfield Diffie e Martin Hellman, que introduziu o conceito de chave pública e um novo e engenhoso processo de troca de chaves. Inspirado no documento do Diffie-Hellman e apoioados em fundamentos matemáticos, Rivest, Shamir e Adleman estudaram e criaram uma nova cifra assimétrica, o RSA.

No capítulo seguinte falar-se-à sobre a contextualização da criptografia e do sistema criptográfico assimétrico denominado RSA.

II. Contextualização

Vive-se numa sociedade onde o fluxo da informação é intenso e nesse contexto, surgiu a necessidade de protegê-la de modo a permitir o acesso ao seu conteudo apenas às pessoas autorizadas. Foi essa necessidade que gerou o desenvolvimento de métodos para mascarar as mensagens. Por essa razão o estudo e o aperfeiçoamento da criptografia vêm-se tornando cada vez mais complexos e essenciais.

A criptografia teve a sua origem nos tempos antigos e com a sua evolução converteu-se numa poderosa ferramenta de segurança da informação que circula em redes inseguras, tal como a internet. A criptografia significa escrita secreta ou escrita oculta, havendo vários algoritmos para ocultar as mensagens.

Embora a criptografia seja de importância fundamental para o meio militar, os sistemas criptográficos são amplamente utilizados no meio civil, em aplicações comerciais e nas simples trocas de mensagens que são efectuadas pela internet.

Um dos marcos mais importantes para a criptografia foi a descoberta do sistema criptografico assimétrico. O método assimétrico mais utilizado actualmente é o RSA, usado em transacções bancárias, correio electrónico, troca de mensagens e também utilizado, por exemplo, no Netscape, o mais popular dos aplicativos de navegação na internet. Este sistema foi apresentado pela primeira vez em 1978, por Rivest, Shamir e Adelman e faz parte de um conjunto de algoritmos criptográficos de chave pública, onde a chave de codificação é diferente da chave de descodificação.

Pela sua ampla utilização, o RSA tem sido testado em diversos aspectos a fim de garantir a sua segurança. O método não só garante a transmissão de informações confidenciais através de canais inseguros como também pode garantir a autenticação do remetente. O RSA é um sistema criptográfico extremamente útil e de grande importância nos dias de hoje.

Devido às razões apresentadas acima, o método RSA foi o escolhido para a elaboração deste trabalho.

Neste capítulo abordou-se a contextualização da criptografia e do sistema criptográfico assimétrico, chamado RSA. No próximo capítulo serão analisados os principais algoritmos criptográficos e serão discutidas as suas vantagens e desvantagens.

III. Desenvolvimento

3.1 Introdução

Neste capítulo reflecte-se sobre os assuntos que são importantes para compreender a escrita secreta: as grandes ciências da escrita secreta, termos e definições úteis para auxiliar na sua leitura e os métodos usados pela escrita secreta com exemplos para melhor compreensão dos mesmos. Abordar-se-ão também os pontos fortes e fracos de cada um desses métodos.

3.2 Termos e Definições

Alguns termos e definições específicos serão usados durante este trabalho. O propósito desta secção é auxiliar na leitura do mesmo e apresentar a sua explicação.

A comunicação é um processo pelo qual o ser humano pode trocar informação e o objecto da comunicação é a mensagem. Há várias formas de comunicação, oral, escrita e electrónica. Os principais intervenientes da comunicação são o emissor e receptor. O emissor ou remetente envia a mensagem e o receptor ou destinatário recebe e lê a mensagem. O texto da mensagem pode ser original, ou seja, legível para muitas pessoas, ou cifrado, apenas entendível pelo emissor e receptor.

O processo que permite a transformação de texto original para cifrado, dá pelo nome de cifra, codificação ou encriptação, o processo oposto designa-se pelos termos decifra, descodificação ou decriptação (Dent & Mitchell, 2005; Kahn, 1967; Stallings, 2005).

3.3 Criptografia

A Criptografia (do grego *Kryptos* "secreto", + "*Gráphé* " escrever" + -ia)[22] permite converter informação através de um protocolo acordado entre o emissor e receptor. Este método permite enviar mensagens irreconhecíveis para quem desconhece o protocolo. Criptógrafos são as pessoas que se dedicam a estudar a criptografia. Os objectivos principais da criptografia são: confidencialidade, integridade, autenticação e não-repúdio. A confidencialidade assegura que só o receptor autorizado deve conseguir decifrar e ler o conteúdo da mensagem, a integridade garante que a mensagem não foi

[22] www.infopédia.pt/lingua-portuguesa/criptografia

alterada, a autenticação permite ao receptor validar se a mensagem foi enviada pelo emissor correcto e, por fim, o não-repúdio é um processo que permite validar a autoria da mensagem.

Nem todos os algoritmos criptográficos são usados para atingir todos os objectivos, há algoritmos específicos para cada um dos objectivos. A sua principal função é manter o texto original e a chave secreta fora do conhecimento dos seus opositores, adversários, atacantes ou piratas.

Os atacantes, para recuperar o texto original podem usar dois processos, a criptoanálise ou a força bruta. A criptoanálise é um processo ilegal, realizado por alguém que não têm conhecimento do processo de decifra e pretende fazê-lo através de técnicas de encriptação que são do seu conhecimento. "While cryptographer develops new methods of secret writing, it is the cryptanalyst who struggles to find weaknesses in these methods in order to break into secret messages."[23] (Singh, 1999, p. 15). O outro método é a Força Bruta que é um processo de recuperação que consiste em executar várias tentativas, por alguém sem grandes conhecimentos de criptoanálise, até obter algo com sentido e a partir daí deduzir regras para converter o texto cifrado no texto original.

A esteganografia (Do grego *steganós*, "oculto" + *gráphein*, "escrever" +-ia)[24] é um método que permite ocultar uma mensagem. Não converte a mensagem, apenas esconde o seu conteúdo, se a mensagem é descoberta o seu conteúdo é revelado.

O nome da escrita secreta que junta estas duas artes, a criptografia e a criptoanálise, intitula-se Criptologia (Do grego *Kryptos*, "oculto" + *logos*, "tratado" +-ia)[25] e os seus praticantes são os criptólogos (Schneier, 1996; Stallings, 2005).

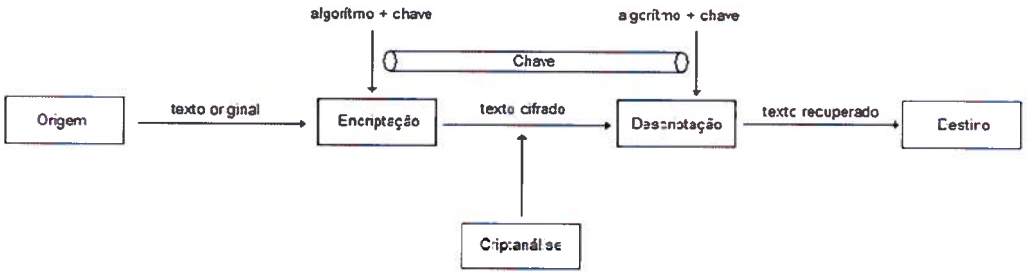


Figura 1 - Esquema geral de funcionamento da criptografia
Fonte: (Stinson, 1995)

[23] "Enquanto o criptógrafo desenvolve novos métodos de escrita secreta, é o criptoanalista que luta para encontrar fraquezas nesses métodos, de modo a quebrar a mensagem secreta."
[24] www.infopedia.pt/lingua-portuguesa/esteganografia
[25] www.infopedia.pt/lingua-portuguesa/criptologia

3.4 Tipos de Criptografia

Na criptografia existem dois grandes métodos, cada um com as suas vantagens e desvantagens, a cifra simétrica e a cifra assimétrica. A cifra simétrica usa apenas uma chave no processo de encriptação e a chave é partilhada pelo emissor e receptor. A cifra assimétrica trabalha com duas chaves. Uma das chaves serve para cifrar e a outra para decifrar. Cada um dos interlocutores possui duas chaves, uma chave privada como emissor ou secreta e uma chave pública como receptor.

Para efectuar os processos de codificação e descodificação, também podemos dizer cifra e decifra, é necessário um algoritmo e uma chave criptográfica. O algoritmo e a chave formam um sistema criptográfico e a sua robustez depende da capacidade em resistir aos processos de quebra. O sistema criptográfico diz-se seguro se a decifra não ocorre em tempo útil, ou seja, o tempo para quebrar a cifra é superior ao prazo de validade da mensagem e os custos dos recursos envolvidos para quebrar a cifra são superiores aos ganhos obtidos com a informação contida na mensagem.

Para aumentar a segurança do algoritmo criptográfico pode-se aumentar o tamanho da chave, enviar a chave por um canal seguro diferente da mensagem e aumentar a complexidade do algoritmo. O comprimento da chave é muito importante para tornar o ataque pela via da força bruta, muito difícil e complicado (Stallings, 2005).

Os algoritmos podem ser divididos em duas grandes classes quanto ao modo de converter a mensagem na cifra e decifra. Quando a mensagem é processada em pequenos pedaços de texto, bit a bit ou carácter a carácter, chama-se *stream cipher*[26]. A *block cipher*[27] trata grandes blocos de texto, de tal maneira que uma pequena alteração no bloco de entrada provoca uma alteração muito grande no bloco de saída (Stallings, 2005).

3.5 Criptografia Simétrica

Os métodos simples ou clássicos de substituição e transposição serviram de suporte ao aparecimento da cifra moderna, que dá pelo nome de cifra simétrica ou de chave única. O emissor e o receptor partilham a mesma chave.

Embora a criptografia moderna utilize as mesmas ideias básicas da substituição e transposição clássicas, o seu foco actual é criar um algoritmo complexo e emaranhado

[26] Cifra de fluxo
[27] Cifra de bloco

para que o intruso não seja capaz de obter qualquer sentido ou padrão da mensagem (Dent & Mitchell, 2005; Stallings, 2005).

Na Figura 2 mostra-se um modelo simplificado de um algoritmo simétrico.

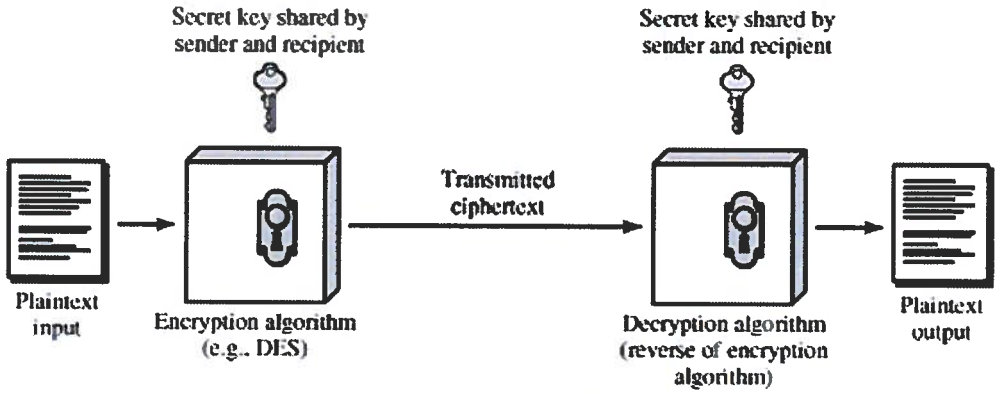


Figura 2 - Modelo simplificado de um algoritmo simétrico
Fonte: (Stallings, 2005)

3.5.1 Substituição

Trata-se de uma das técnicas mais antigas e que consiste em substituir cada letra do texto original por outra do mesmo alfabeto, tornando a mensagem sem sentido.

Estas cifras de substituição podem ser tratadas individualmente, letra a letra, ou em grupos de letras, em blocos.

Quando a substituição converte uma letra do texto original, numa outra letra do texto cifrado, estamos em presença de uma substituição monoalfabética. A conversão diz-se polialfabética quando mais de um carácter do texto original, é utilizado para converter num carácter do texto cifrado. Dois grandes exemplos de cifras que utilizam estas técnicas são a cifra de Júlio César e a cifra de Vigenère.

A cifra de César é uma técnica conhecida por cifra de deslocamento de César, que consiste em substituir uma letra do texto original por outra k casas à frente. O número k é a chave e representa a distância entre o carácter original e o carácter cifrado. Como exemplo vamos criar a
Tabela 4 de conversão, para a chave igual a três (k=3).

Tabela 4 - Tabela de conversão para chave igual a três
Fonte: (Kahn, 1967)

Letra Original	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
Letra cifrada	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

Se a mensagem a cifrar for "criptografia", ao executar a cifra, o resultado é "FULSWRJUDILD".

Para converter foi usada uma técnica de substituição monoalfabética e com uma chave igual a três. Significa que cada carácter é trocado por outro que é sempre o mesmo e três é a deslocação entre o carácter original e o cifrado. Durante anos este foi o algoritmo usado na escrita secreta. Pensava-se que a cifra nunca seria quebrada, até que os árabes descobriram um método para quebrar esta cifra. Após algum estudo descobriram que algumas letras se repetiam mais que outras. Por exemplo no idioma árabe as letras "A" e "L" são as mais usadas, enquanto a letra "J" apresenta uma frequência menor dez vezes. Esta técnica, conhecida por análise de frequências, permite achar a chave que foi usada na cifra do texto (Mollin, 2007).

As mensagens mais longas permitem uma melhor decifra do que os textos pequenos. As mensagens curtas têm um desvio maior em relação às frequências padrão. Para executar esta técnica é necessário conhecer o idioma da mensagem. Obtêm-se uma mensagem não codificada do mesmo idioma, suficientemente longa e analisa-se a frequência com que cada letra se repete no texto. Depois, faz-se a mesma análise sobre o texto cifrado. O processo seguinte é trocar o símbolo com maior frequência do texto cifrado pela letra com maior frequência do texto do idioma e assim por diante (Kahn, 1967; Stallings, 2005).

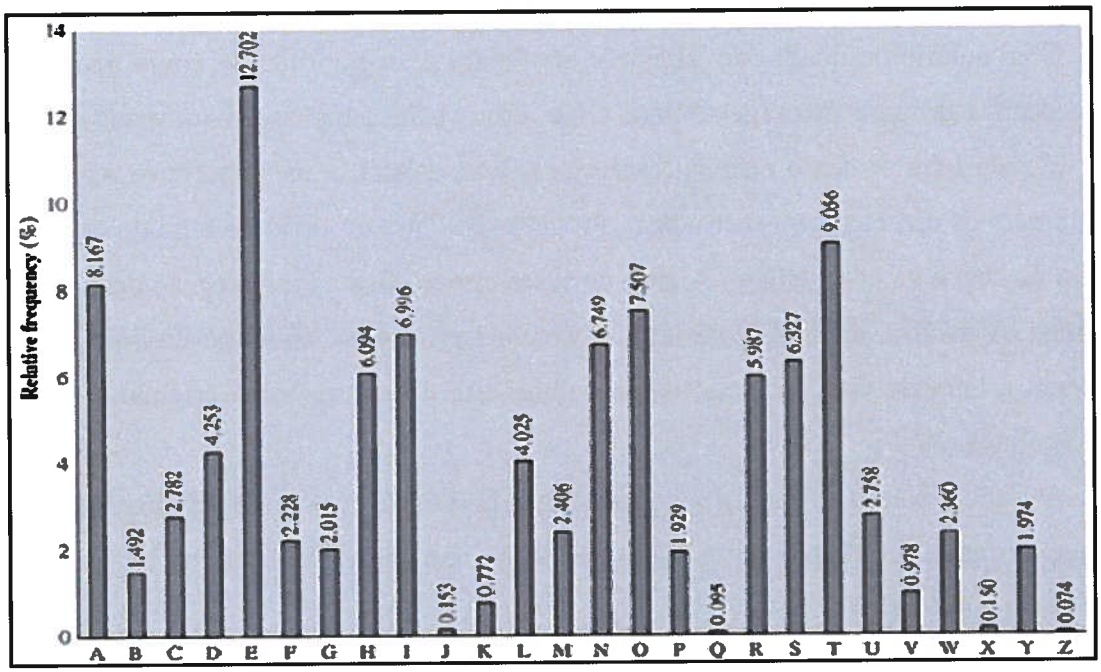


Figura 3 - Frequência relativa do idioma Inglês
Fonte: (Stallings, 2005)

Outra forma de quebrar este algoritmo é usando a técnica da Força Bruta. Normalmente executada por alguém com poucos conhecimentos técnicos, consiste em fazer várias tentativas até obter algo com sentido e a partir daí criar deduções e regras para cifrar a mensagem. Basta testar as vinte e cinco possibilidades da chave. A Tabela 5 mostra a aplicação desta técnica e a mensagem é decifrada à terceira tentativa.

Tabela 5 - O método da Força Bruta aplicado à cifra de César
Fonte: (Stinson, 1995)

Texto Cifrado	F	U	L	S	W	R	J	U	D	I	L	D
Chave 1	E	T	K	R	V	Q	I	T	C	H	K	C
Chave 2	D	S	J	Q	U	P	H	S	B	G	J	B
Chave 3	**C**	**R**	**I**	**P**	**T**	**O**	**G**	**R**	**A**	**F**	**I**	**A**
Chave 4	B	Q	H	O	S	N	F	Q	Z	E	H	Z
...												

Para evitar os ataques pela análise de frequências, a cifra de Vigenère usa uma cifra polialfabética. Para cifrar, é usada uma tabela de 26 linhas e 26 colunas. Por cada linha é escrito o alfabeto deslocado ciclicamente uma posição para a direita. Na primeira linha podemos ver o alfabeto de "A" a "Z", na segunda linha de "B" a "A" e assim sucessivamente até à linha 26 que começa no "Z" e termina no "Y". Antes de iniciar o processo de cifra deve-se escolher uma chave, conhecida apenas pelo emissor e receptor, de preferência sem letras repetidas (Mollin, 2007; Stallings, 2005).

Com auxílio do quadro de Vigenère, na Figura 4, exemplifica-se como ocorre a cifra e decifra de uma mensagem. Para cifrar escreve-se a mensagem original e por baixo de cada letra do texto original escreve-se, letra a letra, a palavra-chave repetida até acabarem os caracteres da mensagem. Procura-se a letra do texto no topo do quadro e a letra da chave na parte lateral. A letra do texto cifrado fica na intercepção da coluna e da linha. A decifra, inicia-se pela letra da chave e procura-se ao longo da linha uma letra igual à letra da cifra, no topo dessa coluna está a letra do texto original (Kahn, 1967; Stallings, 2005).

A Tabela 6 mostra qual é a mensagem cifrada que resulta da combinação da mensagem original "TREINO DE CRIPTOGRAFIA" e a chave "SEGREDO".

Tabela 6 - Cifra de Vigenère
Fonte: (Stinson, 1995)

Texto original	T	R	E	I	N	O	D	E	C	R	I	P	T	O	G	R	A	F	I	A
Palavra-chave	S	E	G	R	E	D	O	S	E	G	R	E	D	O	S	E	G	R	E	D
Texto cifrado	L	V	K	Z	R	R	R	W	G	X	Z	T	W	C	Y	V	G	W	M	D

A Cifra de Vigenère pode ser vista na forma algébrica. Enumeraram-se as letras de "A" a "Z" com os números inteiros de zero a 25.

Tabela 7 - Quadro para aritmética modular
Fonte: (Stinson, 1995)

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

Se for usado o operador mod 26, que é o resto da divisão por 26, ou seja, o número de letras do nosso alfabeto, pode-se representar a cifra e a decifra conforme o exemplo representado na Tabela 8 (Shannon, 1949).

Tabela 8 - Cifra Vigenère pela Aritmética modular
Fonte: (Shannon, 1949)

Cifra Vigenère		
	Cifra	Decifra
Chave	C = P + K (mod 26)	P = C – K+26 (mod 26)
Exemplo	C = ?, P= "T",K="S"	P=?, C="L", K="S"
	C=?, P=19, K=18	P=?, C=11, K=18
	C=19+18(mod 26	P=11-18+26 (mod 26)
	C=37 (mod 26)	P=19 (mod 26)
	C= 11	P= 19
	C= Letra "L"	P= Letra "T"
Legenda	P= Letra original, C= Letra cifrada, K = Chave (de zero a 25)	

	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

Figura 4 - Quadro de Vigenère
Fonte: http://pt.wikipedia.org

Durante muitos anos esta cifra foi considerada indecifrável pela via da criptoanálise. A análise de frequência é um método de criptoanálise que não é eficaz a quebrar sistemas criptográficos de substituição polialfabéticos. Para quebrar as cifras de substituição polialfabéticas, como por exemplo a cifra Vigenère, existem alguns métodos para estimar o comprimento da chave usada na encriptação de uma mensagem. As técnicas mais usadas são o método de ataque chamado Kasiski/Kerckoff e o índice de Coincidências. A principal ideia associada a estas técnicas, baseia-se no facto de que repetidas porções de texto original encriptados com a mesma chave originam textos cifrados com padrões de letras iguais. As chaves curtas e repetidas tornam o algoritmo fraco. Para evitar este tipo de análise a chave deve ser de um tamanho muito próximo do texto original. Contudo, chaves muito longas impedem a sua memorização e implica uma grande probabilidade de cometer um erro na sua escrita (Mollin, 2007).

O teste de Kasiski foi descrito pela primeira vez em 1863 por Friedrick Kasiski (1805-1881), em *Die Geheimschriften und die Dechiffrierkunst*[28]. A ideia geral do método é encontrar duas sequências de letras iguais no texto cifrado e a distância entre elas na mensagem. Os passos para executar este método são: identificar padrões repetidos de um conjunto de letras (três ou quatro), para cada sequência anotar as posições iniciais de cada ocorrência dela no texto cifrado, calcular as diferenças entre as posições iniciais de sequências idênticas e determinar todos os factores das diferenças. O tamanho da chave é determinado pelo máximo divisor comum das diferenças encontradas. Em resumo, procura padrões de letras repetidas e a distância entre elas ao longo da mensagem cifrada. Por exemplo, é muito provável que a palavra "QUE", bastante comum na língua portuguesa, seja substituída constantemente pelo mesmo anagrama na cifra polialfabética. Assim, pode-se dizer que o grupo de letras "HNS", que se repetem ao longo do texto cifrado nas posições 16, 94, 256 tem as distâncias 78, 240 e 162 (94-16=78, 256-16=240, 256-94=162).

Repetem-se os cálculos para outros conjuntos de letras, por exemplo, "UHR", "WEO", etc.

As distâncias assim calculadas são múltiplos de dois e seis. Se o comprimento da chave fosse dois de comprimento, devia haver também múltiplos de quatro, logo pode-se concluir que o comprimento possível da chave é seis (Kahn, 1967; Stallings, 2005).

[28] "Escrita Secreta e a Arte da Decifra"

O algoritmo de Vigenère é uma cifra de substituição, logo não altera a frequência das letras.

O índice de coincidência, com o nome abreviado IC, é um valor estatístico que indica a probabilidade de dois símbolos, tomados ao acaso, corresponderem à mesma letra. Esta técnica tanto pode ser aplicada em textos originais como cifrados. O valor do IC varia de dialecto para dialecto, pois depende do número de vezes que cada símbolo do alfabeto é utilizado. Para exemplificar o IC inglês é 0,0677 ou 6,677% e o português é de 0,0762 ou 7,62%. O IC está intimamente relacionado com o alfabeto e a língua em questão e designa-se por IC_L. A indicação IC_R, que representa o valor estatístico para uma distribuição regular de letras, tem o valor 0,0385. O IC do texto m, designado por IC(m), é definido como a probabilidade de dois caracteres aleatórios de serem iguais. A formula $IC(m) = {}^{z}\sum_{A} n_i \times (n_i - 1)/n \times (n - 1)$, em que o n_i é a frequência da letra i em m e n é o tamanho da mensagem m, permite calcular a ocorrência de cada letra do alfabeto do texto.

Para melhor explicar este conceito vamos exemplificar com a letra "A" que ocorre 11 vezes num texto cifrado de 108 caracteres. Assim, tem-se $11 \times (11 - 1)/108 \times (108 - 1) = 110/11556 = 0{,}0095$. Repete-se os cálculos para as restantes letras e depois basta somar os valores achados para se obter o IC da mensagem.

Friedman, em 1920, propôs ainda uma forma de estimar o tamanho da chave de um texto cifrado. O IC de uma mensagem, encriptada por uma cifra polialfabética, deve variar entre o IC_L e o IC_R da linguagem em questão, dependendo do tamanho da chave. Pode-se concluir que o tamanho da chave T pode ser determinado, em função do IC do texto cifrado m, pela formula: $T \approx n \times (IC_L - IC_R)/(n - 1) \times IC(m) - n \times IC_R + IC_L$ (Friedman, 1920; Kahn, 1967; Mollin, 2007; Stallings, 2005).

3.5.2 Transposição

O método de transposição, basicamente muda a ordem das letras, ou seja, não altera o carácter original mas sim a sua posição de acordo com uma regra ou função. As letras são reorganizadas gerando uma mensagem cifrada. Em função do número de letras da mensagem esta reorganização pode-se tornar muito complicada. Uma palavra com 35 letras pode ter 50×10^{30} possibilidades de arranjos. Esta reorganização deve ser combinada entre o emissor e o receptor (Singh, 1999).

A cifra transposição é fácil de reconhecer porque o texto original têm a mesma frequência de letras que o texto cifrado.

Existem métodos simples e mais complexos. Descrevem-se a seguir os dois modelos sobre este método de substituição.

A Tabela 9 explica como funciona o processo simples.

Imaginem-se duas linhas e escreve-se a mensagem na diagonal.

Tabela 9 - Método simples para Transposição
Fonte: (Stallings, 2005)

Texto original:	encontrohoje										
Texto escrito:	e		c		n		r		h		j	
		n		o		t		o		o		e
Texto cifrado:	ecnrhjnotooe										

Juntam-se as letras da primeira linha com as letras da segunda linha.

Um método mais complexo é escrever a mensagem em várias linhas e com uma chave numérica que indica a sequência.

O método escolhido para a mensagem "estou a vigiar o meu amigo Ricardo", foi a chave de sete dígitos com a ordem definida na Tabela 10.

Tabela 10 - Método mais complexo para Transposição
Fonte: (Stallings, 2005)

Chave:	4	3	1	2	5	6	7
Texto Original:	e	s	t	o	u	a	v
	i	g	i	a	r	o	m
	e	u	a	m	i	g	o
	r	i	c	a	r	d	o
Texto cifrado:	TIACOAMASGUIEIERURIRAOGDVMOO						

O texto cifrado é a junção das colunas por ordem da chave, primeiro a coluna um depois a coluna dois e assim por diante. Para melhor entender com se forma o texto cifrado colocou-se a negro a coluna um e a sublinhado a coluna dois (Stallings, 2005).

3.5.3 Cifra Simétrica

Os algoritmos simétricos, de chave única ou chave secreta usam apenas uma chave no processo de cifra e decifra e a chave é partilhada pelo emissor e receptor. Quanto à sua maneira de processar podem ser divididos em cifra de fluxo ou de bloco. A cifra de fluxo trata a informação bit a bit, enquanto a cifra de bloco converte uma

sequência de bits como se fosse uma única unidade. Das cifras deste grupo destacam-se o ***Data Encrypt Standard*** (DES) e o ***Advanced Encryption Standard*** (AES).

O DES, foi o primeiro algoritmo criptográfico simétrico da época moderna e foi utilizado em larga escala internacionalmente. O DES cifra blocos de 64 bits de texto original em blocos de texto cifrado com 64 bits, usando uma chave de 64 bits mas apenas 56 bits são utilizados. Foi criado pela IBM nos anos 70 e foi baseado no algoritmo de encriptação designado por *Lúcifer*. Esta técnica foi desenvolvida por Horst Feistel, com apoio da IBM nos anos 60. Foi o primeiro algoritmo usado pelas empresas comerciais e financeiras. O seu processo usa blocos de informação que são cifrados pelo método de substituição e transposição. O sistema criptográfico, no processo de cifra executa dois passos. No primeiro cria as 16 sub-chaves de 48 bits e no segundo converte cada um dos blocos de dados com 64 bits. A cifra e decifra são processos similares a única diferença reside na aplicação das sub-chaves de modo inverso. As permutações do DES são de três tipos: *straigh permutation,*[29] os bits da mensagem são apenas reordenados; *expanded permutation,*[30]alguns bits são duplicados e então reorganizados, aumentando assim o número de bits na saída; *permuted choice*,[31] partes dos bits são descartados e os restantes são reordenados, o que provoca uma diminuição no bloco de saída (Foundation, 1998; Stallings, 2005).

O algoritmo que gera as chaves, transforma a chave inicial de 64 bits em 56 bits, aplicando uma tabela de permutação. Os restantes oito bits são descartados ou utilizados como bits de paridade. Os 56 bits são então divididos em dois blocos de 28 bits cada e tratados separadamente. Em voltas sucessivas, 16 ao todo, as duas metades são deslocadas à esquerda em um ou dois bits, especificado em cada volta por uma tabela. A formatação final das 16 sub-chaves, obtêm-se aplicando a segunda tabela de permutação

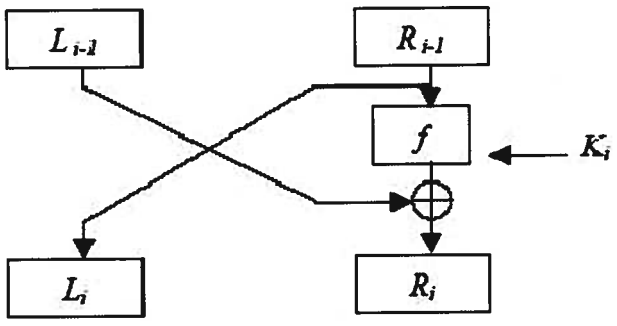


Figura 5- Estrutura de Feistel para Cifrar
Fonte: (Foundation, 1998)

[29] *S-boxes* / Caixas S
[30] *E-boxes* / Caixas E
[31] *P-boxes* / Caixas P

a cada um dos pares concatenados das 16 sub-chaves.

As fases de transformação do bloco de dados são: permutação inicial, tratamento de chaves, fases intermédias e tratamento final.

O DES começa por uma permutação inicial que permite trocar de posição os 64 bits do bloco. Através de uma tabela de permutação os bits são deslocados para novas posições, como por exemplo o bit 58 para a posição um, o bit 50 para a posição dois e assim por diante. O novo bloco de 64 bits é dividido em dois blocos, metade esquerda e

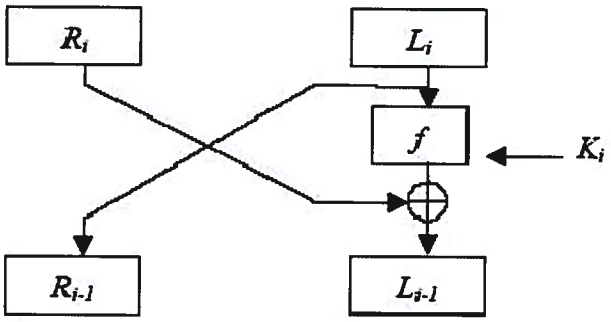


Figura 6 - Estrutura de Feistel para Decifrar
Fonte: (Foundation, 1998)

direita, cada um com 32 bits. A etapa seguinte é um ciclo de 16 voltas de operações idênticas, a "função-F"[32] e operação lógica XOR[33]. A "função-F" combina metade do bloco de dados, o lado direito, junto com a sub-chave da respectiva volta. O resultado devolvido pela "função-F" é operado através de um XOR, com a outra metade do bloco de dados, a parte esquerda e as metades são trocadas antes da próxima volta. No passo seguinte e para terminar o processo de cifra de um bloco, as metades são unidas e é executada uma permuta final, inversa à permuta inicial.

A "função-F" opera com a metade do bloco de dados do lado direito e executa quatro passos. No primeiro passo o bloco de dados de 32 bits é expandido par 48 bits, servindo-se de uma tabela de permutação expansiva. A segunda fase combina o resultado da primeira com a sub-chave da volta correspondente, aplicando uma operação lógica XOR. A terceira e última separa o bloco em oito partes de 6 bits cada, e muda para oito blocos de 4 bits, invocando as oito tabelas de conversão. Termina este método da "função-F" com uma permutação (Stallings, 2005).

O algoritmo era seguro, contudo o incremento do poder computacional, nas mãos dos criptoanalistas, tornou o DES inseguro. Recentemente o DES foi quebrado em

[32] Função *Feistel*-Bagunça metade do bloco com uma chave.
[33] Operação lógica, onde a resposta é verdadeira quando as variáveis, bits, assumem valores diferentes entre si (0 - Falso, 1- Verdadeiro).

menos de um dia para uma chave de 56 bits. Actualmente o DES possui uma versão mais fortalecida composta com três chaves de 56 bits, 156 bits no total. Esta nova versão têm o nome de 3-DES. Em 2002 o DES foi finalmente substituído pelo AES (Stallings, 2005).

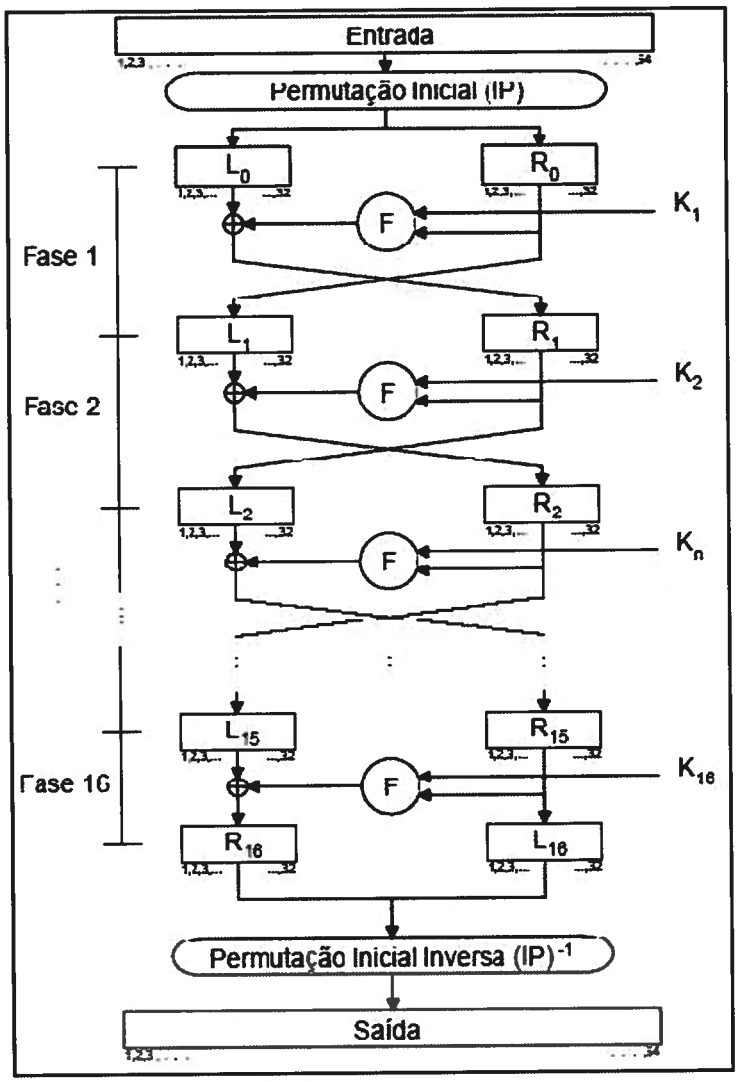


Figura 7- Fluxograma da cifra DES
Fonte: (Foundation, 1998)

Para se analisar a estrutura do AES é importante reflectir sobre a definição de estado, no contexto do algoritmo. O estado é uma matriz de dados, quatro linhas e quatro colunas, que se irá manipular entre os *rounds*[34]. Cada item da matriz tem de comprimento um byte ou oito bits, no total cada quadro tem 128 bits. No AES o número de voltas depende do tamanho da chave, sendo 10, 12 ou 14, para chaves de 128, 192 ou 256 bits. A chave é expandida usando-se o escalonamento de chaves do *Rijndael*.

[34] Ciclos ou Voltas

A cada volta do processo de cifra, o AES realiza quatro etapas: *AddRoundKey*, *SubBytes*, *ShiftRows* e *MixColumns*.

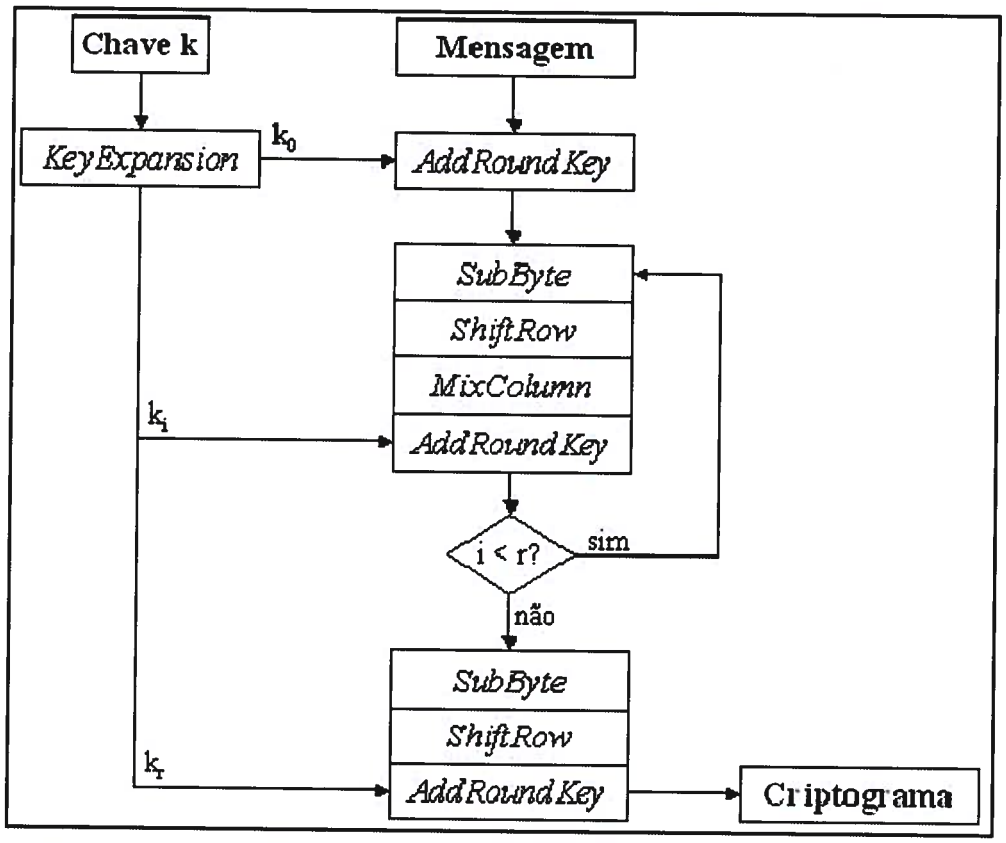


Figura 8- Fluxograma do sistema criptográfico AES
Fonte: (Stallings, 2005)

Na volta final, o passo *MixColumns* é eliminado. *AddRoundKey* é uma operação de XOR entre o estado e a chave da volta, com o mesmo número de bytes. O XOR é executado byte a byte entre o estado e a chave. O processo *SubBytes* utiliza uma tabela de 16 por 16 bytes, que contêm valores hexadecimais. Cada byte do estado é substituído por outro da caixa de substituição[35]. Conforme a Figura 8 a troca é feita da seguinte maneira: os quatro primeiros e os últimos quatro bits do byte a substituir representam, em hexadecimal, respectivamente a linha e a coluna onde se encontra o novo byte. Por exemplo, e segundo a Tabela 11, o valor hexadecimal "c2" deverá ser substituído pelo valor "25". A Figura 9 mostra o quadro usado no AES para efectuar este processo. A operação inversa chama-se *InvSubBytes* e aplica uma tabela inversa. A transformação *ShiftRows*, consiste em deslocar à esquerda as linhas do estado, trocando assim as posições dos bytes. O número de posições a serem deslocadas depende do número da

[35] S-Box

linha. A primeira linha não sofre qualquer tipo de rotação, na segunda linha os bytes são deslocados uma posição, na terceira e última linha as rotações são da ordem de duas e três posições. *InvShiftRows* é o nome da operação inversa e para aplicá-la basta executar as mesmas rotações em sentido contrário. Na etapa *MixColumns*, os quatro bytes de cada coluna são tratados separadamente. Isto significa que o resultado de cada determinada coluna não altera o resultado das outras. No entanto cada byte influencia os outros bytes da mesma coluna. Este processo fornece difusão à cifra e podemos representar essa transformação por uma multiplicação de matrizes em GF(2^8). O novo estado E' será o resultado da multiplicação de uma matriz fixa F pela matriz E e representa-se por E' = E □ F. O símbolo □ caracteriza a multiplicação matricial (Daemen & Rijmen, 2002; Stallings, 2005)

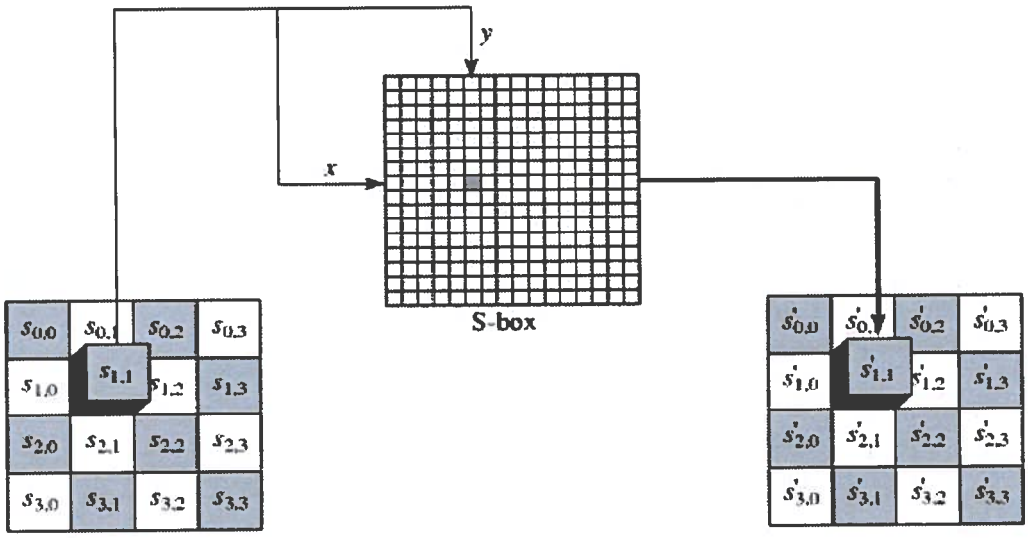


Figura 9 - Substituição de bytes para a cifra AES
Fonte: (Stallings, 2005)

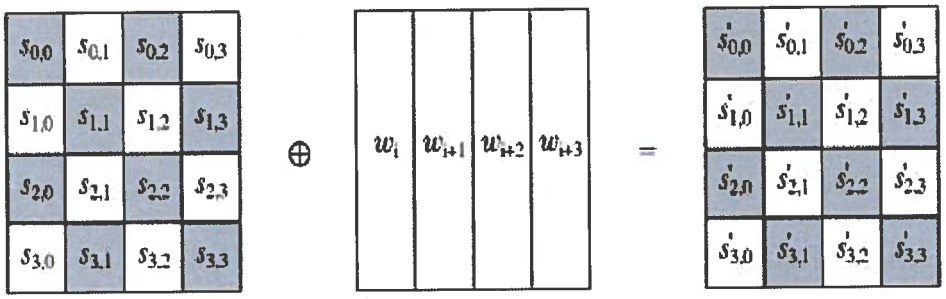


Figura 10 - Transformação *AddRoundKey* da cifra AES
Fonte: (Stallings, 2005)

x \ y	0	1	2	3	4	5	6	7	8	9	A	B	C	D	E	F
0	63	7C	77	7B	F2	6B	6F	C5	30	01	67	2B	FE	D7	AB	76
1	CA	82	C9	7D	FA	59	47	F0	AD	D4	A2	AF	9C	A4	72	C0
2	B7	FD	93	26	36	3F	F7	CC	34	A5	E5	F1	71	D8	31	15
3	04	C7	23	C3	18	96	05	9A	07	12	80	E2	EB	27	B2	75
4	09	83	2C	1A	1B	6E	5A	A0	52	3B	D6	B3	29	E3	2F	84
5	53	D1	00	ED	20	FC	B1	5B	6A	CB	BE	39	4A	4C	58	CF
6	D0	EF	AA	FB	43	4D	33	85	45	F9	02	7F	50	3C	9F	A8
7	51	A3	40	8F	92	9D	38	F5	BC	B6	DA	21	10	FF	F3	D2
8	CD	0C	13	EC	5F	97	44	17	C4	A7	7E	3D	64	5D	19	73
9	60	81	4F	DC	22	2A	90	88	46	EE	B8	14	DE	5E	0B	DB
A	E0	32	3A	0A	49	06	24	5C	C2	D3	AC	62	91	95	E4	79
B	E7	C8	37	6D	8D	D5	4E	A9	6C	56	F4	EA	65	7A	AE	08
C	BA	78	25	2E	1C	A6	B4	C6	E8	DD	74	1F	4B	BD	8B	8A
D	70	3E	B5	66	48	03	F6	0E	61	35	57	B9	86	C1	1D	9E
E	E1	F8	98	11	69	D9	8E	94	9B	1E	87	E9	CE	55	28	DF
F	8C	A1	89	0D	BF	E6	42	68	41	99	2D	0F	B0	54	BB	16

Figura 11 - Tabela de substituição para cifra AES (S-Box)
Fonte: (Stallings, 2005)

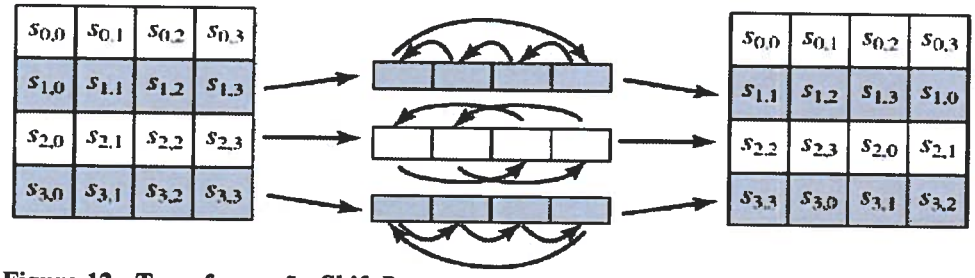


Figura 12 - Transformação *Shift Row*
Fonte: (Stallings, 2005)

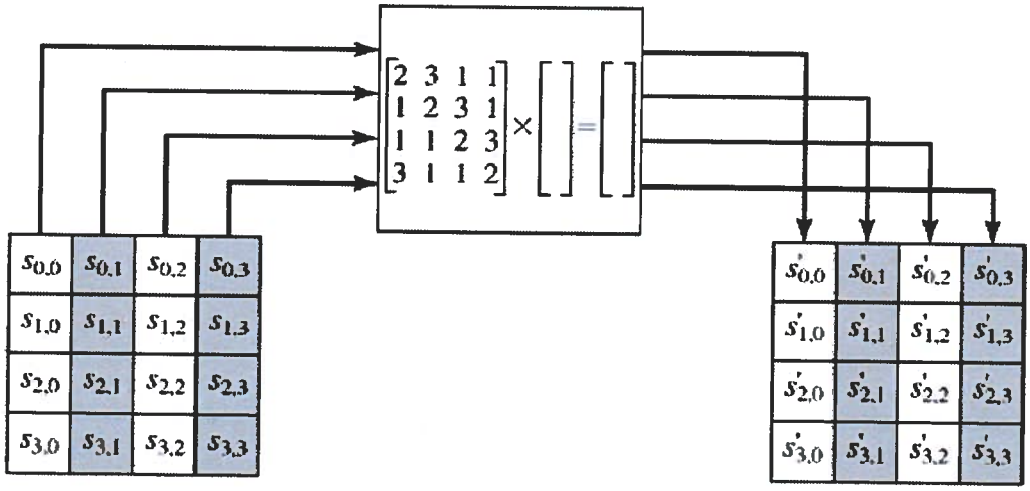


Figura 13 - Transformação *Mix Columns*
Fonte: (Stallings, 2005)

A expansão da chave inclui as seguintes funções: *Rotword*, *SubWord* e *Rcon[i/Nk]*. A função *RotWord* usa uma palavra[36] $[a_0,a_1,a_2,a_3]$ como input[37], executa uma permutação cíclica e devolve a palavra $[a_1,a_2,a_3,a_0]$. A função que recebe uma palavra de entrada, aplica uma tabela de substituição[38] a cada um dos quatro bytes de forma a devolver uma outra palavra de saída, designa-se por *SubWord* (Daemen & Rijmen, 2002; Stallings, 2005).

3.5.4 Segurança dos algoritmos simétricos

A segurança de um algoritmo de chave única ou simétrica depende de dois componentes, a sua robustez e o comprimento da chave. O algoritmo utiliza a chave para alterar o texto original e convertê-lo em texto cifrado. Para recuperar a mensagem em formato ilegível é necessário usar a mesma chave e executar a operação. Na criptografia simétrica, a chave que é utilizada para cifrar os dados também é usada para decifrá-los. O tamanho de uma chave é um factor muito importante na segurança do algoritmo, uma chave de dois dígitos possui 100 possibilidades de combinações, uma chave de seis dígitos têm 1 milhão de possibilidades. Um número maior de possibilidades de combinações, requer ao criptoanalista mais tempo para decifrar o texto. Partindo do princípio que o algoritmo é forte, significa que só há uma maneira de quebrar o algoritmo criptográfico. Calcular a complexidade de um ataque à chave pela via da Força Bruta não é complicado. A rapidez para executar a quebra, pela Força Bruta, de um algoritmo depende de três factores: o comprimento da chave, o número de chaves e a velocidade de processamento. A força bruta aplicada ao DES significa testar todas as possibilidades da chave, 2^{56} ou cerca de $7{,}2 \times 10^{16}$ chaves. Existem outros ataques conhecidos que podem quebrar os 16 ciclos do algoritmo DES, contudo são estudos teóricos e não são viáveis na prática. Esses ataques são conhecidos por criptoanálise diferencial[39], criptoanálise linear[40] e *Davies' attack*[41] (Stallings, 2005).

3.6 Criptografia Assimétrica

[36] Conjunto de quatro bytes forma uma palavra.
[37] Dados de entrada.
[38] S-Box
[39] Os seus criadores são os israelitas Eli Biham e Adi Shamir
[40] Descoberta por Mitsuru Matsui
[41] Criada por Donald Davies e melhorada por Biham e Biryukov

O RSA é um algoritmo de criptografia de dados, que deve o seu nome a três professores do MIT, actuais fundadores da empresa *RSA Data Security Inc*, Ronald Rivest, Adi Shamir e Leonard Adleman. Este algoritmo até à data, é o mais bem sucedido na implementação de chaves assimétricas, fundamenta-se na teoria clássica dos números e a sua segurança reside na dificuldade, em parte, de factorizar números grandes. O RSA funciona com uma chave pública, que é conhecida por várias entidades e uma chave privada que deve ser mantida escondida. A mensagem é cifrada pela chave pública e a decifra pela chave privada.

O método permite a comunicação segura sem transporte de chaves e permite assinar documentos. Este método tem resistido a todos os ataques por um período de tempo suficiente, pelo que pode ser usado com confiança (Rivest, Shamir, & Adleman, 1978).

3.7 Criptografia Simétrica vs. Assimétrica

A criptografia de chave simétrica, quando comparada com o sistema de chaves assimétricas, é a indicada para garantir a confidencialidade de grandes volumes de dados, pois o seu processamento é mais rápido. Torna-se pouco segura quando pretendemos um canal de comunicação seguro para compartilhar a chave secreta. É necessário que a chave secreta seja previamente combinada por um canal de comunicação seguro, para não comprometer a confidencialidade da chave.

O sistema assimétrico, apesar de ter um processamento mais lento que o método simétrico, resolve o problema de distribuição de chaves, pois dispensa a necessidade de um canal de comunicação seguro para compartilhar as chaves.

Consequentemente, para o mesmo grau de segurança, as chaves simétricas são consideravelmente menores do que as chaves assimétricas. Actualmente o comprimento das chaves ronda os 2048 bits para as chaves assimétricas e 128 bits para chaves simétricas.

Para aproveitar as vantagens de cada um destes métodos, o ideal é o uso combinado de ambos. A criptografia simétrica é usada para cifrar informação e o sistema assimétrico é utilizado para compartilhar a chave secreta ou de sessão e assim dispensa a necessidade de um canal de comunicação seguro para distribuição da chave.

Pode-se então traçar um comparativo entre a criptografia simétrica e assimétrica, conforme mostra a Tabela 11 (Menezes, van Oorschot, & Vanstone, 1996).

Tabela 11-Comparação de Algoritmos Simétrico e Assimétricos
Fonte: (Menezes, van Oorschot, & Vanstone, 1996)

Características	Cifra de Chave Simétrica	Cifra de Chave Assimétrica
Velocidade	Alta	Lenta
Confiabilidade	Boa	Muito Boa
Nível de Segurança	Alto	Alto
Quantidade de chaves	Uma	Duas
Gerência e distribuição de Chaves	Complexa	Simples
Assinatura Digital	Não	Sim

3.8 Conclusão

Neste capítulo estudou-se alguns dos principais sistemas criptográficos que utilizam as técnicas de substituição, transposição, simetrica e assimétrica. O estudo destes métodos, suas origens, funcionamento, assim como a sua segurança, foram aqui analisados.

As técnicas de substituição e transposição tiveram origem na época antiga e foi com base nos seus conceitos que foi criado o sistema simétrico. As cifras simétricas aqui identificadas como as mais importantes são o DES e o AES. Para analisar o sistema criptográfico assimétrico foi escolhido como modelo o método RSA.

Alem de focar o principal problema do sistema simétrico, que consiste na criação e distribuição da sua chave privada, também se elaborou uma análise comparativa entre estes dois métodos, simétrico e assimétrico, que servirá de fundamento para as conclusões finais.

O sistema criptográfico RSA além de oferecer confiabilidade, pode ser utilizado, também, para autenticação.

No próximo capítulo explicam-se os conceitos matemáticos que dão suporte e complexidade à criação das chaves pública e privada do sistema criptográfico RSA.

IV. Conceitos Matemáticos

4.1 Introdução

Este capítulo tem por objectivo principal rever e sistematizar noções básicas de matemática, para melhor entender o funcionamento do algoritmo assimétrico RSA.

Apesar de ter sido um passatempo para Fermat, Euler e Gauss, hoje a teoria dos números é o pilar do sistema criptográfico de chaves assimétricas.

4.2 Divisibilidade

Dados dois número inteiros a e b, com $a \neq 0$ diz-se que a divide b (denotamos a | b) se existe c inteiro tal que $b = a \times c$.

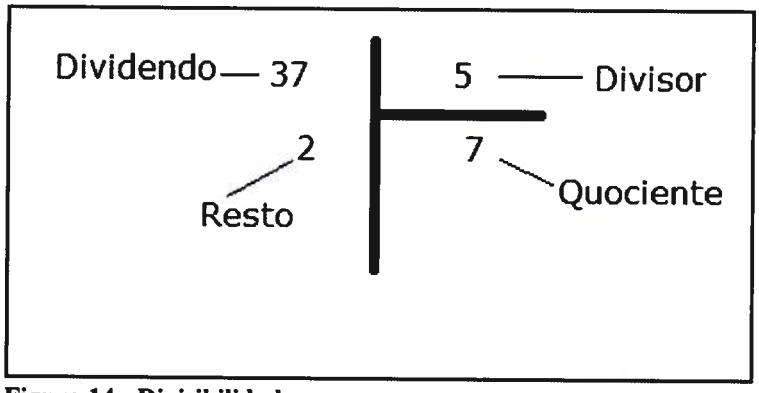


Figura 14 - Divisibilidade
Fonte: (Costa & Anjos, 1967)

Numa divisão o dividendo é igual ao produto do divisor pelo quociente mais o resto, para o resto menor que o divisor (Dividendo = Divisor x Quociente + Resto). Esta regra representa a propriedade fundamental da divisão (Mollin, 2007; Silva, Silva, & Paulo, 1970).

4.3 Números Primos

Um número primo é um número natural que tem dois e apenas dois divisores naturais distintos: o número um e ele mesmo.

Um número se não é primo é chamado número composto. Número composto é todo o número natural que tem mais de dois divisores. O número cinco tem como divisores o um e o cinco, logo é um número primo. Os divisores de 15 são um, três, cinco e 15, neste caso esta-se em presença de um número composto.

Dois ou mais números são primos entre si quando o máximo divisor comum desses números é o um. Quando mdc(a; b) = 1, diz-se que a e b são primos entre si ou coprimos. Por convenção, os números zero, um e menos um, não são considerados primos nem compostos (Mollin, 2007).

2	3	5	7	11	13	17	19	23	29	31	37	41	43	47	53	59	61	67
71	73	79	83	89	97	101	103	107	109	113	127	131	137	139	149	151	157	163
167	173	179	181	191	193	197	199	211	223	227	229	233	239	241	251	257	263	269
271	277	281	283	293	307	311	313	317	331	337	347	349	353	359	367	373	379	383
389	397	401	409	419	421	431	433	439	443	449	457	461	463	467	479	487	491	499
503	509	521	523	541	547	557	563	569	571	577	587	593	599	601	607	613	617	619
631	641	643	647	653	659	661	673	677	683	691	701	709	719	727	733	739	743	751
757	761	769	773	787	797	809	811	821	823	827	829	839	853	857	859	863	877	881
883	887	907	911	919	929	937	941	947	953	967	971	977	983	991	997			

Figura 15 - Tabela de números primos inferiores a 1000
Fonte: (Mollin, 2007)

4.3.1 Identificação de um número primo

Há várias formas, mas um dos processos mais simples, ainda que trabalhoso, é o da Decomposição em Factores Primos. Vai-se testando a divisibilidade do número por cada um dos números primos, iniciando em 2, até que a divisão tenha resto zero ou que o quociente seja menor ou igual ao número primo que se está testando como divisor.

Veja-se a decomposição do número 17 na Tabela 12.

Tabela 12 - Decomposição em factores primos do número 17
Fonte: Própria

Decomposição do número 17			
Dividendo	Divisor	Quociente	Resto
17	2	8	1
17	3	5	2
17	5	3	2

Neste ponto já se pode ter a certeza de que o número 17 é primo, pois nenhum dos divisores primos testados produziu resto zero e o quociente da divisão pelo número primo cinco é igual a três que é menor que o divisor cinco. E o número 91 será primo? Veja-se a demonstração na Tabela 13.

Tabela 13 - Decomposição em factores primos do número 91
Fonte: Própria

Decomposição do número 91			
Dividendo	Divisor	Quociente	Resto
91	2	45	1
91	3	30	1
91	5	18	1
91	7	13	0

Como no último teste a divisão foi exacta, restando zero, concluí-se que o número 91 não é um número primo, de facto ele possui quatro divisores distintos: um, sete, 13 e 91 (Mollin, 2007).

4.3.2 Algoritmo para identificar números primos

O algoritmo aqui descrito permite saber se dois números inteiros são primos entre si (coprimos).

```
#include <stdio.h>
Int main(void)
{
        int a, b,  p, limite, primo=2, mdc=1;
        scanf("%d", &a);
        scanf("%d", &b);
        while ( a != 1 && b != 1)
        {
                If (a % primo == 0)
                {
                        a = a / primo;
                        if ( b % primo == 0)
                        {
                                b = b / primo;
                                mdc = mdc * primo;
                        }
                }
                else
                        if (b % primo ==0)
                                b= b / primo;
                        else
                                if (primo == 2 )
                                        primo = 3
                                else
                                        do
                                        {
```

```
                    Primo = primo + 2;
                    p = 1;
                    limite = primo / 2;
                    do
                    {
                        p = p + 2;
                    }
                    while( primo % p != 0 && p < limite);
                }
                While (primo % 0 p == 0);
    }
    printf("%d\n", mdc);
    return 0;
```

4.4 Potências

A potência é uma operação matemática, escrita como a^n, envolvendo dois números, a base a e o expoente n. Quando n é um número natural maior que um, a potência a^n indica a multiplicação da base a por ela mesma tantas vezes quanto indicar o expoente n. E representa-se matematicamente da seguinte maneira: $a^n = a_1 \times ... \times a_n$.

Da mesma forma que a multiplicação de n por a pode ser vista como uma soma de n parcelas iguais a a, ou seja: $a \times n = a_1 + ... + a_n$.

O expoente geralmente é indicado à direita da base, aparecendo sobrescrito ou separado da base por um circunflexo. Pode-se ler a^n como a elevado à n-ésima potência, ou simplesmente a elevado a n. Alguns expoentes possuem nomes específicos, por exemplo, a^2 costuma ser lido como a elevado ao quadrado e a^3 como a elevado ao cubo.

Por convenção, como é sabido, estabeleceu-se que:

Tabela 14 - Resultados de potências definidas por convenção
Fonte: (Silva, Silva, & Paulo, 1970)

Expoente zero	$a^0 = 1$
Expoente um	$a^1 = a$
Expoentes inteiros negativos	$a^{-1} = (1/a)$
Indeterminações	0^0
	0^n, quando $n < 0$
	∞^0
	1^∞
Expoente fraccionário	$\sqrt[m]{a^n} = a^{m/n}$
Expoente decimal	$a^{1,5} = a^{15/10} = \sqrt[10]{a^{15}}$

4.5 Operações com potências

Podem-se realizar as seguintes operações: Adição, Subtracção, multiplicação e divisão. Na adição e subtracção o processo para resolver estas potências é idêntico, converte-se o valor de cada uma delas e somam-se ou subtraem-se os valores obtidos. Podem-se aplicar duas regras para a multiplicação de potências.

O produto de potências da mesma base é uma potência com a mesma base e cujo expoente é a soma dos expoentes dos factores:

$a^m \times a^n = a^{m+n}$.

O produto de potências com o mesmo expoente é uma potência com o mesmo expoente e cuja base é o produto das bases dos factores:

$a^m \times b^m = (a \times b)^m$.

Assim com a multiplicação a divisão com potências também se pode aplicar duas regras para operar com este tipo de potência.

O quociente de potências com a mesma base é uma potência com a mesma base e cujo expoente é a diferença entre os expoentes do dividendo e do divisor:

$a^m : a^n = a^{m-n}$.

O quociente de duas potências com o mesmo expoente é uma potência com o mesmo expoente e cuja base é o quociente das bases do dividendo e do divisor:

$a^m : b^m = (a : b)^m$.

A potência de uma potência é outra potência com base da primeira e expoente igual ao produto dos expoentes:

$(a^m)^n = a^{m \times n}$.

Mesmo com todas as simplificações descritas anteriormente, a sua resolução deve ser feita iterativamente[42], para potências com expoente muito grande.

Estes cálculos podem ainda ser simplificados e no caso de módulos e expoentes grandes, essa simplificação ganha maior importância. Uma das formas para calcular eficientemente o a^n mod m passa por escrever o expoente como uma soma de potências de dois. Por exemplo calcular 7^{327} mod 853: as potências de dois são 2^0, 2^1, 2^2, 2^3, 2^4, 2^5, 2^6, 2^7, 2^8, que correspondem aos valores decimais 1, 2, 4, 8, 16, 32, 64, 128 e 256. O valor do expoente 327 é igual a 1 + 2 + 4 + 64 + 256 o que implica a seguinte equação

[42] Gera resultados parciais que são usados nas operações seguintes até atingir o resultado final.

$(7^1 \times 7^2 \times 7^4 \times 7^{64} \times 7^{256})$ mod 853 ou $(7^1$ mod 853 $\times 7^2$ mod 853 $\times 7^4$ mod 853 $\times 7^{64}$ mod 853 $\times 7^{256}$ mod 853) mod 853. O resultado final desta operação é 286 (Costa & Anjos, 1967; Mollin, 2002).

4.6 Mínimo Múltiplo Comum (mmc)

Se a e b são números naturais não nulos, o mínimo múltiplo comum de a e b é o menor múltiplo de a e b. A sua representação simbólica é mmc(a, b).

Quando um número é divisível por outro, isto é, a divisão entre eles tem resto igual a zero, diz-se que os números são múltiplos. O zero é múltiplo de todos os números naturais. Exemplos: Múltiplos de 4 = {0, 4, 8, 12, 16, 20, ...} e Múltiplos de 6 = {0, 6, 12, 18, 24, ...}. O menor múltiplo comum de quatro e seis, diferente de zero, é 12. Diz-se então que o número 12 é o mínimo múltiplo comum de quatro e seis.

Para calcular o mmc de dois ou mais números por um processo mais rápido e prático, utiliza-se a decomposição de um número em factores primos; $4 = 2 \times 2$, $6 = 2 \times 3$, mmc(4,6) = $2 \times 2 \times 3$, mmc(4,6) = 12. O mínimo múltiplo comum de dois ou mais números decompostos em factores primos, é igual ao produto de factores comuns e não comuns de maior expoente (Mollin, 2007; Stallings, 2005).

4.7 Máximo Divisor Comum (mdc)

O maior divisor comum (abreviadamente mdc) entre dois ou mais números inteiros, em que pelo menos um deles não é zero, é o maior número inteiro que divide todos estes números.

Nem sempre existe um máximo divisor comum e nem sempre é único.

É claro que se m = n = 0, então o máximo denominador comum, simbolicamente representado por mdc(m,n), não faz sentido.

O máximo divisor comum de {12,15}, ou seja, mdc(12,15) é três. O número três é o maior divisor natural que divide, com resto zero, o número 12 e 15. Alguns exemplos: mdc(6, 12) = 6, mdc(12, 20) = 4, mdc(20, 24) = 4, mdc(12, 20, 24) = 4, mdc(6, 12, 15) = 3. Quando o máximo divisor comum de dois ou mais números é igual a 1, então os números dizem-se primos entre si. Os números {a, b, c, ...} cujo mdc(a, b, c, ...) = 1 são chamados primos entre si. Os números 35 e 24 são primos entre si, pois mdc(35,24) = 1. Os números 35 e 21 não são primos entre si, pois mdc(35,21) = 7.

Há duas formas de determinar o máximo divisor comum de dois números ou mais números. Um dos processos para calcular o máximo divisor comum entre dois ou mais números é utilizar a decomposição desses números em números primos. O mdc de dois ou mais números, quando factorizados, é o produto dos factores comuns a eles, cada um elevado ao menor expoente. Achemos o mdc de *30* e *12*, note que $30 = 2 \times 3 \times 5$, $12 = 2 \times 2 \times 3 \Rightarrow 12 = 3 \times 2^2$ então mdc(30,12) = 2×3. Os factores comuns aos números 30 e 12 são os números dois e três. No caso do dois tinha-se expoente um e dois, mas considera-se o menor, ficando só o dois e não o dois ao quadrado. O outro processo permite calcular o mdc, efectuando várias divisões até chegar a uma divisão exacta. O divisor da divisão exacta, é o mdc dos números. Calcular, através deste processo, o mdc de 48 e 30. A

Tabela 15 descreve as sucessivas operações para calcular o mdc de 48 e 30.

Tabela 15 - Calculo do MDC de 48 e 30
Fonte: (Burnett & Paine, 2002)

Calcular o máximo divisor comum de 48 e 30			
Dividendo	Divisor	Quociente	Resto
48	30	1	18
30	18	1	12
18	12	1	6
12	6	2	0

A operação começa com a divisão do maior número pelo menor número. O divisor da operação anterior é o novo dividendo e o novo divisor é o anterior resto e assim sucessivamente até ao final da operação. Se a última divisão é exacta então o mdc dos números é o último divisor, caso contrário não existe mdc entre os números. Neste exemplo, o divisor da divisão exacta é *6*, então o mdc(48,30) = 6 (Mollin, 2007; Stallings, 2005)..

4.7.1 Algoritmo para calcular o Máximo Divisor Comum (mdc)

O algoritmo ou programa que permite calcular o máximo divisor comum entre dois números, é calculado através de divisões sucessivas entre dois valores (a , b). O ciclo só termina quando b for igual a zero.

Os valores de a são os valores de b e o valor inicial de a (195). Os valores de b correspondem aos valor inicial de b (150) e os restos da divisão entre a e b. Os valores a

fornecer ao algoritmo podem ser pela ordem crescente ou decrescente, conforme exemplificam a Tabela 16 e a Tabela 17.

Tabela 16 - Calculo do mdc de 195 e 150
Fonte: (Burnett & Paine, 2002)

Valores de **a**	Valores de **b**
195	150
150	45
45	15
15	0

Tabela 17 - Calculo do mdc de 150 e 195
Fonte: (Burnett & Paine, 2002)

Valores de **a**	Valores de **b**
150	195
195	150
150	45
45	15
15	0

```
#include <stdio.h>
int main(void)
{
        int a, b, r;
        scanf("%d", &a;
        scanf("%d", &b;
        while (b != 0)
        {
                r = a % b;
                a = b;
                b = r;
        }
printf("%d\n", a)
return 0;
}
```

4.8 Algoritmo de Euclides

O algoritmo de Euclides é um método simples e eficiente para calcular o máximo divisor comum entre dois números inteiros diferentes de zero. É um dos algoritmos mais antigos ainda em uso, conhecido desde 300 a.C (Mollin, 2007).

O algoritmo de Euclides consiste em efectuar divisões sucessivas entre dois números até obter resto zero. O máximo divisor comum entre os dois números iniciais é o último resto diferente de zero. Este método não requer qualquer factorização. Como exemplo tome-se os números 348 e 156. Com auxílio da

Tabela 18 podemos visualizar as susessivas operações desta técnica.

Como o resto da primeira operação não é zero, substituí-se na segunda linha o dividendo e o divisor, pelo divisor e resto da primeira linha. Na seguinte operação, como o resto não é zero, substituí-se o dividendo e o divisor na terceira linha, pelo divisor e resto da segunda linha. Na terceira e última linha a divisão dá resto zero, portanto, o máximo divisor comum dos inteiros 348 e 156 é 12. Tambem se pode concluir que: mdc(348,156) = = mdc(156,36) = = mdc(36,12) = = 3.

Tabela 18 - Calculo do mdc de 348 e 156 pelo método de Euclides
Fonte: (Burnett & Paine, 2002)

	A	B	C	D
Linhas	Dividendo	Divisor	Quociente	Resto
1	348	156	2	36
2	156	36	4	12
3	36	12	3	0

Junto podemos visualizar um pseudocódigo e um programa para executar o algoritmo de Euclides automaticamente.

Pseudocódigo

```
AlgoritmoDeEuclides(a: inteiro; b: inteiro):inteiro
Variáveis
  divisor: inteiro
  dividendo: inteiro
  c: inteiro
Início
  dividendo = a
  divisor = b
  Fazer enquanto resto(dividendo/divisor) ≠ 0
    c = resto(dividendo/divisor)
    dividendo = divisor
```

```
    divisor = c
  Fim-Fazer
    AlgoritmoDeEuclides = divisor
fim-função
/*
* Programa para calcular o máximo divisor comum de dois números
* Recebe números inteiros não negativos a e b, ambos não nulos e devolve o mdc(a,b).
*/
long euclides_mdc(long a, long b)
{
  long r;
  do
  {
     r = a % b;
      a = b;
      b = r;
  }
  while ( r != 0);
  return a;
}
```

O algoritmo também pode ser declarado utilizando recursividade[43]:

```
AlgoritmoDeEuclides(a: inteiro; b: inteiro): inteiro
Início
  Se b = 0 então
    AlgoritmoDeEuclides = a
  Senão
    AlgoritmoDeEuclides = AlgoritmoDeEuclides(b,resto(a,b))
  Fim-se
Fim-função
/*
* Recebe números inteiros não-negativos a e b, ambos não nulos e devolve o mdc(a,b).
*/
Long Euclides_mdc(long a, long b)
{
  If (b=0) return a;
```

[43] Propriedade de um programa ou função que se pode invocar a si próprio.
http://www.priberam.pt/recursividade

```
  return Euclides_mdc(b, a % b);
}
```

4.9 Inverso Multiplicativo.

Dois números são inversos multiplicativos um do outro, quando a sua multiplicação der o resultado um. Inverso multiplicativo também muitas vezes designado por recíproco. O inverso multiplicativo de a é simbolicamente representado por 1/a. Esta propriedade aplica-se a todos os números e são quase sempre números distintos, excepto poucos casos. O zero é o único número que não tem inverso multiplicativo e tem algumas propriedades peculiares. A fracção 1/0 representa uma operação indefinida. Também o um é um número, que é o seu próprio inverso multiplicativo, por razões óbvias.

O inverso multiplicativo de um número em módulo é o múltiplo do número em questão, cujo resultado dividido pelo módulo, dá resto um. Para exemplificar vai-se usar um modelo simples de fracções que, multiplicadas, resultam na unidade. O inverso multiplicativo de três meios é dois terços e vice-versa, porque a sua multiplicação dá o resultado de seis sextos que representa a unidade. Simbolicamente a sua representação é: $2/3 \times 3/2 = 6/6 = 1$. O algoritmo estendido de Euclides é o melhor método para calcular o inverso multiplicativo de um número em módulo (Mollin, 2007).

4.10 Pequeno Teorema de Fermat.

Se p é número primo e a não é múltiplo de p, então o pequeno teorema de Fermat diz que $a^{p-1} = 1 \pmod p$. Este Teorema não tem nada a ver com o último teorema de Fermat. Utilizando o teorema de Fermat podemos calcular expressões complexas com relativa facilidade. Para calcular a potência $2^{5432675} \pmod{13}$, basta calcular a divisão de 5432675 / 12 = 11. Assim aplicando as regras das potências e o Teorema de Fermat: $2^{5432675} \pmod{13} = (2^{12})^{452722} \pmod{13} \times 2^{11} \pmod{13} = (1)^{452722} \pmod{13} \times 2^{11} \pmod{13} = 1 \times 2^{11} \pmod{13} = 2048 \pmod{13} = 7$ (Mollin, 2007).

4.11 Função de Euler

A função de Euler é o conjunto números inteiros positivos que são menores que o valor de n e que são primos com n. A sua configuração simbólica é φ (n). Se p é

primo, então $\varphi(n) = p-1$. Se $n = p \times q$ com p, q primos, então $\varphi(n) = (p-1) \times (q-1)$. Estes cálculos são muito importantes para a geração da chaves no sistema criptográfico RSA. Com o número 42 e a partir da sua factorização podemos exemplificar: $\varphi(42) = \varphi(2 \times 3 \times 7) = \varphi(2) \times \varphi(3) \times \varphi(7) = (2-1) \times (3-1) \times (7-1) = 1 \times 2 \times 6 = \varphi(42) = 12$. O pequeno teorema de *Fermat* estabelece que, para a e n primos entre si, mdc(a, n) = 1 então $a^{\varphi(n)} \equiv 1 \pmod{n}$.

Fazendo n = 42, temos: $a^{\varphi(42)} \equiv 1 \pmod{42} = a^{12} \equiv 1 \pmod{42}$. Finalmente, multiplicando por a, o que retira a necessidade de a e n serem primos entre si: $a^{13} \equiv a \pmod{42}$.

Permite calcular o inverso multiplicativo de um número em módulo, mas nem sempre é possível o seu uso. Se mdc(a, n) = 1 então $a^{\varphi(n)} \equiv 1 \pmod{n}$ o seu inverso é $a^{\varphi(n)-1} \bmod n = x$, x é o inverso de n. Para melhor compreender este raciocínio vamos exemplificar calculando o inverso de cinco, modulo sete. Sete é um número primo logo $\varphi(7) = (7-1) = 6$, para calcular o inverso de 5, modulo 7 basta executar a seguinte operação $5^{6-1} \bmod 7 \Rightarrow 5^5 \bmod 7 = 3$. O três é o inverso de 5, modulo7 (Mollin, 2007; Stallings, 2005).

4.12 Conclusão

Neste capítulo analisou-se e reflectiu-se sobre os conceitos matemáticos que suportam um dos métodos criptográficos mais popular actualmente, o RSA. Podemos concluir que estes teoremas matemáticos não foram criados para este efeito, no entanto é com base neles que o sistema criptográfico RSA obtem a sua segurança. A segurança do método RSA baseia-se na complexidade dos fundamentos matemáticos inseridos na teoria dos números. No capítulo seguinte mostrar-se-á como funciona o algoritmo RSA e falar-se-à sobre os cuidados a ter para implementar este método com segurança.

V. Algoritmo RSA

5.1 Introdução

O documento publicado por Diffie e Hellman, em 1976 originou outro estudo que foi publicado, em 1978, por R. Rivest, A. Shamir e L. Adleman (Rivest, Shamir, & Adleman, 1978). Neste documento eles descrevem um novo sistema de criptografia, chamado RSA.

O algoritmo RSA, faz parte dos algoritmos criptográficos de chaves assimétricas, onde a chave de cifrar é diferente da chave para decifrar, sendo a sua robustez baseada na dificuldade, custo computacional, de se factorizar números inteiros. Estas duas chaves diferentes, mas matematicamente relacionadas, são denominadas por chave pública e privada. Este capítulo visa mostrar como funciona o algoritmo e alertar para os cuidados a ter para impedir que se torne vulnerável às técnicas da criptoanálise. Baseada no problema da factorização de inteiros, pode-se resumir o RSA em três partes: Geração de chaves, Codificação e Descodificação.

5.2 Geração das Chaves Pública e Privada no RSA

Para geração da chave pública e chave privada, escolhem-se aleatoriamente dois números primos grandes p e q. O tamanho dos números depende da robustez que se quer implementar ao algoritmo. Um número grande implica maior segurança mas menor desempenho.

Para obter o módulo, calcula-se o inteiro n, a partir do produto de p e q ($n = p \times q$). Os números p e q são mantidos privados e o n é do domínio público. Antes de começar a cifrar há que converter o texto m em números usando para isso os códigos ASCII. Determinar a função de Euler para os número primos p e q, através da seguinte equação: $\varphi(n)= (p-1) \times (q-1)$. Escolhe-se um valor e, tal que este seja com o valor $\varphi(n)$, primos entre si, ou seja: $1 < e < \varphi(n)$ e $mdc(e, \varphi(n)) = 1$. Onde mdc significa maior divisor comum de ambos os números. Para calcular a variável d, deve-se usar o algoritmo de Euclides estendido, de tal forma que $d \times e \equiv 1 \pmod{\varphi(n)}$.

Os valores obtidos e e d, serão usados na chave pública e privada. A chave pública é formada pelo par de números n e e e a chave privada pelo par n e d (Mollin, 2002; Rivest, Shamir, & Adleman, 1978; Stallings, 2005).

5.3 Cifra e Decifra

Para cifrar usa-se a chave publica PU={e, n} através da fórmula, $C = M^e \bmod n$ e para decifrar tem-se a chave privada PR={d, n} através da fórmula, $M = C^d \bmod n$. A mensagem a tratar deve ser, em comprimento, inferior ao valor de n. A próxima tabela mostra como gerar chaves, cifrar e decifrar com o algoritmo RSA.

Tabela 19 - Exemplo de calculo de chaves no algoritmo RSA
Fonte: (Stallings, 2005)

Algoritmo RSA exemplo	
Descritivo da operação	Cálculos/fórmulas/variáveis
Escolha de número primos	p = 17 e q = 11
Calcular módulo	$n = p \times q \Rightarrow 11 \times 17 = 187$
Função de Euler	$\varphi(n) = (p-1) \times (q-1) \Rightarrow 16 \times 10 = 160$
Seleccionar e,	$mdc(e, \varphi(n)) = 1 \Rightarrow mdc(e, 160) = 1$; escolhemos o 7
Determinar o d	$d \times e \equiv 1 \pmod{\varphi(n)}$ e $d < 160$
Valor de d	$d=23; 23 \times 7 = 161; 10 \times 16 + 1$
Chave Pública	$PU=\{7, 187\}; C = M^7 \bmod 187$
Chave Privada	$PR=\{23, 187\}; M = C^{23} \bmod 187$
Legenda	M é a mensagem original C é a mensagem cifrada

Após o cálculo da chave pública e privada vai-se exemplificar como cifrar e decifrar uma mensagem com a letra "M". O seu valor em número é 88. O valor de C e D resultam das seguintes operações: $C = 88^7 \bmod 187 \Rightarrow C = 11$, $D = 11^{23} \bmod 187 \Rightarrow D = 88$ (Mollin, 2007; Stallings, 2005).

5.4 Ataques ao RSA

Desde a sua publicação, o sistema criptográfico RSA tem sido analisado por vários investigadores, que estudam as suas vulnerabilidades. Este algoritmo apoia-se em conceitos matemáticos do século XVIII e por essa circunstância um dos possíveis ataques, além da força bruta, baseia-se em técnicas matemáticas. A força bruta concentra-se em testar todas as chaves privadas possíveis. A defesa contra os ataques de força bruta reside, exactamente, em gerar chaves grandes e num intervalo tão grande,

que seja computacionalmente muito caro definir qual a chave a ser usada. Torna-se também uma enorme dificuldade em obter a chave privada a partir da chave pública, que é do domínio público. A outra abordagem, apoia-se em técnicas matemáticas, que se podem classificar em três tipos de ataques. A factorização de n, em dois factores primos p e q, a determinação directa de φ(n) sem primeiro determinar p e q e obter o valor de d directamente, sem possuir o valor de φ(n). Para qualquer um dos casos tratam-se de métodos com grandes dificuldades, pois actualmente não existem algoritmos de factorização eficiente para números primos grandes. Na factorização de um valor n de 200 dígitos, seriam necessárias um total de 1,2 × 1023 operações, levando $3{,}8 \times 109^{44}$ anos a processar. O valor de φ(n) deve ser razoavelmente grande a fim de não se tornar um ponto de fraqueza a ser explorado por eventuais atacantes (Mollin, 2007; Stallings, 2005).

5.5 Cuidados necessários em novas aplicações

A maior vulnerabilidade do RSA encontra-se na falta de atenção quanto aos cuidados que seguem, a serem observados ao se desenvolver uma nova aplicação que utilize o algoritmo RSA, envolvendo implementação, escolha das chaves e utilização do algoritmo

Os primos p e q devem ser grandes para evitar ataques de força bruta. A ordem de grandeza deles deve ser analisada conforme o contexto da aplicação (necessidade de segurança, capacidade computacional relacionada à aplicação). Hoje em dia, o tamanho mínimo para p e q considerado seguro é de 2048 bits (Mollin, 2002; Rivest, Shamir, & Adleman, 1978; Stallings, 2005).

5.6 Assinatura Digital

Alguns algoritmos criptográficos, de chave pública, permitem gerar o que se chama de assinatura digital. O primeiro método descoberto foi o esquema de assinatura RSA, que permanece até hoje uma das técnicas mais práticas e versáteis disponíveis (Menezes, van Oorschot, & Vanstone, 1996).

O algoritmo criptográfico RSA além da operação normal de cifra com chave pública e decifra com a chave privada, permite a cifra com a chave privada e a decifra

[44] Usando o algoritmo de *Richard Schroeppel*.

com a chave pública. Obviamente que este método não assegura o sigilo da informação, uma vez que qualquer um pode decifrar a mensagem, dado que a chave pública é do conhecimento público (Rivest, Shamir, & Adleman, 1978).

Para obter confidencialidade com a assinatura digital, basta combinar os dois métodos. O emissor assina a mensagem, utilizando a sua chave privada e volta a cifrar a mensagem com a chave pública do receptor. O receptor ao receber a mensagem, deve primeiro decifrá-la com a sua chave privada, o que garante a sua privacidade. Em seguida é novamente decifrada, ou seja, verifica a sua assinatura utilizando a chave pública do emissor, garantindo assim a sua autenticidade. Na

Tabela 20 podem-se analisar as operações possíveis de executar com o sistema criptográfico RSA (Rivest, Shamir, & Adleman, 1978; Preneel & Rijmen, June 1997).

Tabela 20 - Operações do sistema criptográfico RSA
Fonte: (Rivest, Shamir, & Adleman, 1978)

Descritivo	Emissor (a)	Receptor (b)	Objectivo criptográfico
Cifra RSA	Cb(m)	Db(m)	Confidencialidade
Assinatura Digital	Da(m)	Ca(m)	Autenticidade
Assinatura Digital + Cifra RSA	Cb(Da(m))	Ca(Db(m))	Confidencialidade + Autenticidade
Legenda	C - Chave pública D - Chave Privada m - mensagem de a para b		

5.7 Conclusão

O sistema criptografico de chave pública RSA é um padrão mundial, tanto para a confidencialidade como para a assinatura digital. Analisou-se o seu funcionamento compreendendo a importância da matemática para a segurança deste método.

A maior vulnerabilidade do método RSA é a falta de conhecimento e atenção quanto aos cuidados a ter ao se implementar uma aplicação que utilize o algoritmo RSA.

Tanto a geração da chave como a implementação e utilização do algoritmo devem seguir todos os cuidados determinados.

Conclusão

Com este trabalho reflectiu-se sobre a criptografia nos tempos actuais, assim como a sua importância nos tempos antigos. Descreveram-se os objectivos da criptografia e explicou-se o funcionamento de algumas cifras para haver uma familiarização com o verdadeiro papel da criptografia. Falou-se na constante disputa entre criptógrafos e criptoanalistas. Abordou-se o problema da criptografia simétrica em gerir, distribuir e guardar a sua chave privada. Analisou-se as duas cifras simétricas, DES e AES, consideradas as mais importantes para este estudo. Outro dos objectivos foi o estudo dos fundamentos matemáticos, que contribuíram para o aparecimento do sistema criptográfico RSA. Elaborou-se um estudo comparativo entre o método simétrico e assimétrico, que forneceu dados importantes para a conclusão final deste trabalho. Por último, utilizou-se o algoritmo RSA como modelo de criptografia assimétrica, mostrando os pontos em que se baseia a sua segurança e exemplificando o processo de cifrar, decifrar e assinatura digital.

Analisando a informação contida neste trabalho pode-se concluir que a compreensão da evolução histórica da criptografia ajuda não só a entender a sua importância para a humanidade, como também o seu desenvolvimento ao longo dos tempos. O seu desenvolvimento é explicado não só pelas evoluções tecnológicas mas também devido ao trabalho dos criptoanalistas cujo objectivo é quebrar os algoritmos criptográficos. Desta forma, é necessário que os criptógrafos estejam sempre um passo à frente dos criptoanalistas. A constante disputa entre criptógrafos e criptoanalistas culminou numa nova ciência, a Criptologia.

Apresentaram-se os fundamentos básicos sobre o funcionamento dos algoritmos simétricos DES e AES. Contudo, recomenda-se o uso da cifra AES, que é o actual padrão de criptografia do governo americano e que foi escolhido a partir de um concurso público lançado em 1977 pelo NIST. Havia a necessidade de escolher um novo algoritmo mais seguro e eficiente para substituir o DES, que apresentou fragilidades.

Como já foi dito anteriormente, a segurança do sistema criptográfico simétrico depende da sua chave privada, pois o adversário pode identificar todos os detalhes do algoritmo, mas não pode conhecer a chave privada que é usada no processo de codificação e descodificação. Este é o principal desafio do método simétrico, o processo de criação, distribuição e guarda da sua chave privada.

Relativamente à chave, pode-se concluir que quanto maior for, mais seguro é o algoritmo, contudo, o seu processamento torna-se mais lento.

O conceito criptográfico de chave pública foi introduzido por Diffie e Helman para solucionar os dois problemas associados à criptografia convencional: distribuição da chave e autenticação (assinatura digital). Pode-se concluir que a criptografia de chave pública não foi criada para substituir a criptografia simétrica, mas para complementá-la, tornando-a mais segura.

Foram explicados os conceitos elementares do algoritmo RSA, que se trata de um sistema criptográfico que usa conceitos matemáticos simples, que o faz parecer fácil de ser quebrado, mas até à data o algoritmo resistiu a todas as provas a que foi submetido e por isso continuam as expectativas quanto ao seu futuro.

O artigo de Don Boneh, "Twenty years of attacks on the RSA cryptosystem", faz um estudo de vinte anos de ataques sofridos pelo RSA e termina concluindo que, passados duas décadas ainda não existem técnicas de ataque ao RSA que não sejam evitáveis através da aplicação de cuidados durante a sua implementação.

Após recolha e análise dos dados, como é visível na Tabela 12, concluiu-se que as futuras aplicações criptográficas devem utilizar uma abordagem híbrida, ou seja, combinar as vantagens da segurança simétrica e as vantagens da criptografia assimétrica.

Ficou claro que o uso da criptografia no mundo actual é imprescindível e um método seguro para proteger os dados que circulam em canais inerentemente inseguros como a internet. O RSA garante a transmissão de informações sigilosas através de redes inseguras e também pode garantir autenticação do utilizador, o que o torna extremamente útil em transacções bancárias.

Este trabalho foi muito importante para mim, pois a aprendizagem efectuada aquando do desenvolvimento deste tema contribuiu para um enriquecimento pessoal e profissional.

No decurso deste trabalho surgiram algumas dificuldades, primeiro na escolha e identificação dos objectivos, posteriormente no manuseamento das bases das pesquisas, e por último no acesso e selecção à informação disponível sobre o tema. Optou-se por utilizar informação contida em livros e materiais digitais e apesar das dificuldades, confirmo que todos os objectivos propostos foram atingidos.

Bibliografia

Al-Kadi, I. A. (1992). *The origins of cryptology: The Arab Contributions.*

Boneh, D. (1999). Twenty years of attacks on the RSA cryptosystem. *In Notices of the American Mathematical Society (AMS)* , Vol. 46, No. 2, pp. 203-- 213.

Brassard, G. (1988). *Modern Cryptography - A Tutorial. Lecture Notes in Computer Science.* Springer-Verlag.

Burnett, S., & Paine, S. (2002). *RSA Security's Official Guide to Cryptography.* EUA: Macgraw-Hill.

Cohen, F. (1995). *A Short History of Cryptography.*

Costa, A. A., & Anjos, A. O. (1967). *Compêndio de Matemática.* Porto Editora, Lda.

Daemen, J., & Rijmen, V. (2002). *The Design of Rijndael - AES the Advance Encryption Standard.* Springer.

Dent, A. W., & Mitchell, C. J. (2005). *User's Guide to Cryptography and Standards.* Artech House.

Feistel, H. (1973). *Cryptography and Computer Privacy.* Scientific American.

Foundation, E. F. (1998). *Cracking DES: Secrets of Encryption Research, Wiretap Politics & Chips Design.* O'Reilly Media.

Friedman, W. F. (1920). *The Index of Coincidence and Its Applications in Cryptoanalysis.* Aegeon Park Press.

Hellman, W. D. (1976). *New Directions in Criptography, IEEE transactions on information Theory.*

Kahn, D. (1967). *The codebreakers: The story of Secret Writing.* New American Library.

Koblitz, N. (1994). *A Course in Number Theory and Cryptography* . Springer-Verlag (Second Edition).

Menezes, A., van Oorschot, P., & Vanstone, S. (1996). *Handbook of Applied Cryptography.* CRC Press.

Mollin, R. A. (2007). *An Introduction to Cryptography-(Discrete Mathematics and its Applications), Second Edition.* Taylor & Francis Group.

Mollin, R. A. (2002). *RSA and Public-Key Criptography (Discrete Mathematics and Its Applications).* Chapman and HAll/CRC.

Preneel, B., & Rijmen, V. (June 1997). *State of the Art in Applied Cryptography.* Leuven Belgium: Springer.

Rivest, R., Shamir, A., & Adleman, L. (1978). *A Method for Obtaining Digital Signatures and Public-Key Cryptosystems*. Obtido em 15 de Fevereiro de 2014, de URL: http://cr.yp.to/bib/entries.html#1978/rivest

Schneier, B. (1996). *Applied cryptography: Protocols, Algorithms and source code in C, 2 edition.* John Wiley & Sons.

Shannon, C. E. (1948). *A Mathematical Theory of Communication.* Bell System Tech. J.

Shannon, C. E. (1949). *Communication Theory of Secret Systems*. Obtido em 10 de Novembro de 2013, de http://netlab.cs.ucla.edu/wiki/files/shannon1949.pdf

Silva, J. S., Silva, J. D., & Paulo, S. (1970). *Compêndio de Álgebra, Segunda Edição.* Braga: Livraria Cruz, Braga.

Singh, S. (2002). *The Code Book, How to Make it, Break it, Hack it, Crack it.* Delacorte Press.

Singh, S. (1999). *The Code Book: The Science of Secrety from Ancient Egypt to Quantum Cryptography.* Anchor Book.

Singh, S. (2006). *The Code Book: The Science of Secrety from the Ancient Egypt to Quantum Cryptography.* Freshman Seminar.

Stalling, W. (2007). *Network Security Essentials: Applications and standards.* Prentice Hall.

Stallings, W. (2005). *Cryptography and Network Security: Principles and Pratice.* Prentice Hall.

Stinson, D. (1995). *Cryptography: Theory and Pratice, Third Edition.* Obtido em 27 de Julho de 2013, de http://www.icst.pku.edu.cn/course/Cryptography/CryptographyTheoryandpractice(3ed).pdf