DATA PROTECTION

FERNANDO BOAVIDA
MÁRIO BERNARDES

INTRODUÇÃO À CRIPTOGRAFIA

FCA

Edição
FCA – Editora de Informática, Lda.
Av. Praia da Vitória, 14 A – 1000-247 Lisboa
Tel: +351 213 511 448
fca@fca.pt
www.fca.pt

Distribuição
Lidel – Edições Técnicas, Lda.
Rua D. Estefânia, 183, R/C Dto. – 1049-057 Lisboa
Tel: +351 213 511 448
lidel@lidel.pt
www.lidel.pt

Livraria
Av. Praia da Vitória, 14 A – 1000-247 Lisboa
Tel: +351 213 511 448 * Fax: +351 213 522 684
livraria@lidel.pt

ISBN edição impressa: 978-972-722-902-4
1.ª edição impressa: maio 2019

Paginação: Carlos Mendes
Impressão e acabamento: Tipografia Lousanense, Lda. – Lousã
Depósito Legal n.º 456229/19
Capa: José M. Ferrão – *Look-Ahead*

Marcas registadas de FCA – Editora de Informática, Lda. –

Todos os nossos livros passam por um rigoroso controlo de qualidade, no entanto aconselhamos a consulta periódica do nosso site (www.fca.pt) para fazer o *download* de eventuais correções.

Não nos responsabilizamos por desatualizações das hiperligações presentes nesta obra, que foram verificadas à data de publicação da mesma.

Os nomes comerciais referenciados neste livro têm patente registada.

ÍNDICE

Os Autores VIII

Lista de Siglas e Acrónimos IX

1 Introdução 1
- **1.1** Motivação e âmbito 2
- **1.2** Objetivos e abordagem 4
- **1.3** Organização do presente texto 6

2 Fundamentos 9
- **2.1** O que é a criptografia? 10
- **2.2** Perspetiva histórica 12
- **2.3** Princípios e conceitos 15
  - **2.3.1** Algoritmo criptográfico seguro 15
  - **2.3.2** Princípio de Kerckhoffs 16
  - **2.3.3** Criptografia e matemática 17
  - **2.3.4** Números aleatórios *versus* pseudoaleatórios 19
  - **2.3.5** Força das chaves criptográficas 21
  - **2.3.6** Autenticação 22
  - **2.3.7** Criptografia simétrica *versus* criptografia assimétrica 23
- **2.4** Tipos de ataques 25
  - **2.4.1** Ataques genéricos ou específicos 25
  - **2.4.2** Ataques passivos ou ativos 26
  - **2.4.3** Ataques com base no texto 27
  - **2.4.4** Ataques de colisão 28
  - **2.4.5** Outros tipos de ataques 29
- **2.5** Árvores de ataque 30
- **2.6** Desempenho e complexidade 31

Conclusão 32

Exercícios 33

Bibliografia 37

Outros recursos 37

3 Visão Geral da Criptografia Simétrica 39
- **3.1** Princípios gerais 40
- **3.2** Técnicas clássicas de encriptação 41
  - **3.2.1** Técnicas de substituição 42
  - **3.2.2** Técnicas de transposição 47
  - **3.2.3** Máquinas de rotores 48
- **3.3** Segurança computacional 51
- **3.4** Modos de encriptação 52

3.4.1 Encriptação em sequência 52
3.4.2 Encriptação em blocos 53
Conclusão 58
Exercícios 59
Bibliografia 62
Outros recursos 62

4 Encriptação de Blocos 63
4.1 Introdução 64
4.2 Redes de substituição-permutação 65
4.2.1 Paradigma de confusão-difusão 65
4.2.2 Substituição-permutação 66
4.2.3 Esquema de chaves 69
4.2.4 Efeito de avalanche 69
4.2.5 Força de uma rede de substituição-permutação 70
4.3 Redes de Feistel 70
4.4 *Data Encryption Standard* 72
4.4.1 Visão geral 72
4.4.2 Função de estágio 75
4.4.3 Esquema de chaves 78
4.4.4 Fragilidades do DES 80
4.4.5 Alternativas para o reforço do DES 82
4.5 *Advanced Encryption Standard* 84
4.5.1 Estrutura 85
4.5.2 Função de estágio do AES 89
4.5.3 Segurança do AES 97
4.6 Outros algoritmos de encriptação de blocos 98
Conclusão 99
Exercícios 101
Bibliografia 104
Outros recursos 104

5 Funções de *Hash* Criptográficas 105
5.1 Introdução 106
5.2 Áreas e exemplos de aplicação 108
5.3 Requisitos de segurança 116
5.4 Exemplos de funções de *hash* 118
5.4.1 Dois exemplos simples 118
5.4.2 Funções de *hash* baseadas em CBC 120
5.4.3 MD-5 120
5.4.4 Estrutura geral de uma função de *hash* 121
5.5 *Secure Hash Algorithm* 122
5.5.1 Estrutura do SHA-512 123
5.5.2 Inicialização do *buffer* de estado 124
5.5.3 Função de compressão 'F' 124

5.5.4 Função de estágio 127
Conclusão 128
Exercícios 129
Bibliografia 130
Outros recursos 130

6 Códigos de Autenticação de Mensagens 131
6.1 Aspetos gerais 132
6.2 MAC baseados em funções de *hash* 138
6.2.1 *Hash-based Message Authentication Code* 138
6.3 MAC baseados em encriptação de blocos 140
6.3.1 *Data Authentication Algorithm* 141
6.3.2 *Cipher-based* MAC 142
6.4 Encriptação autenticada 143
6.4.1 *Counter with Cipher Block Chaining Message Authentication Code* 144
6.5 Geração de números pseudoaleatórios 146
Conclusão 148
Exercícios 148
Bibliografia 151
Outros recursos 151

7 Criptografia Assimétrica 153
7.1 Aspetos gerais 154
7.1.1 Conceito de criptografia assimétrica 154
7.1.2 Formatos e extensões de ficheiros 158
7.2 Requisitos 160
7.3 Algoritmo RSA 160
7.3.1 Apresentação geral do algoritmo 161
7.3.2 Ataques ao algoritmo RSA 164
7.3.3 Exemplos práticos de utilização do RSA 165
7.4 Outros algoritmos de encriptação assimétrica 169
7.4.1 Diffie-Hellman 169
7.4.2 Elgamal 171
7.4.3 *Digital Signature Algorithm* 173
7.4.4 Criptografia de curvas elípticas 173
7.5 Geração de números pseudoaleatórios com base no RSA 175
Conclusão 176
Exercícios 178
Bibliografia 179
Outros recursos 180

8 Infraestruturas de Chaves Públicas 181
8.1 Introdução 182
8.1.1 Motivações 182
8.1.2 Soluções para distribuição de chaves públicas 184

8.1.3 Exemplos de utilização de PKI 186
8.2 Componentes 189
8.3 Arquiteturas 190
8.4 Aspetos organizacionais 193
8.4.1 Certificado raiz 194
8.4.2 Certificados multinível 194
8.4.3 Separação de CA e RA 194
8.4.4 Identificação de utilizadores 195
8.4.5 Expiração de certificados 195
8.4.6 Revogação de certificados 196
8.4.7 Ciclo de vida das chaves criptográficas 196
8.5 Estruturas de dados 197
8.5.1 Certificados X.509 197
8.5.2 Listas de certificados revogados 201
8.6 Exemplos práticos 202
8.6.1 Obtenção de certificados multidomínio 202
8.6.2 Verificação de certificados 207
Conclusão 207
Exercícios 208
Bibliografia 209

9 Aplicações 211
9.1 Introdução 212
9.2 Aplicações de sistema 213
9.2.1 Autenticação forte 214
9.2.2 Encriptação de dispositivos 215
9.2.3 Assinaturas digitais 216
9.2.4 Encriptação de conteúdos 218
9.2.5 Acesso físico 218
9.2.6 Federações de identidade 219
9.3 Aplicações de rede 220
9.3.1 IP *Security* 220
9.3.2 SSL/TLS 223
9.3.3 *Secure Shell* 225
9.3.4 Correio eletrónico 227
9.3.5 *Network link encryption* 227
9.3.6 Redes virtuais privadas 228
9.3.7 Redes sem fios 230
9.3.8 Redes celulares 233
9.3.9 *Domain Name System Security Extensions* 233
9.4 Computação em nuvem 235
9.4.1 Conceito 235
9.4.2 Modelos 236
9.4.3 Arquitetura de referência 238

9.4.4 Proteção de dados em ambiente de computação em nuvem 239
9.4.5 Segurança como um serviço de computação em nuvem 240
9.5 *Blockchain* 241
9.5.1 Conceito 242
9.5.2 Aspetos estruturais 245
9.5.3 Construção da *blockchain* 249
9.5.4 Tipos de *blockchain* 252
9.5.5 Modelos de consenso 253
9.5.6 Contratos inteligentes 254
9.5.7 Aplicações 256
9.5.8 Limitações, problemas e mitos 260
9.5.9 Plataformas de *blockchain* 261
Conclusão 262
Exercícios 263
Bibliografia 265
Outros recursos 266

10 Gestão de Chaves, Políticas e Procedimentos 269
10.1 Introdução 270
10.2 Gestão de chaves 271
10.2.1 Hierarquia de chaves 271
10.2.2 Distribuição de chaves 272
10.2.3 Distribuição com base em encriptação simétrica 273
10.2.4 Distribuição com base em encriptação assimétrica 275
10.2.5 Distribuição de chaves públicas 278
10.2.6 Kerberos 283
10.3 Armazenamento de segredos 288
10.4 Práticas de gestão de chaves e certificados 291
10.5 Depósito de chaves em terceiros 292
Conclusão 292
Exercícios 293
Bibliografia 294
Outros recursos 294

11 Conclusão 295
11.1 Normalização 296
11.2 Critérios para seleção de produtos 297
11.2.1 Certificação de produtos 298
11.2.2 Publicação *versus* não publicação 298
11.2.3 Código aberto *versus* código proprietário 299
11.3 Especialistas em criptografia 300
11.4 Computação e criptografia quânticas 300
11.5 Considerações finais 301

Índice Remissivo 303

OS AUTORES

Fernando Boavida
Professor Catedrático da Faculdade de Ciências e Tecnologia da Universidade de Coimbra, docente do Departamento de Engenharia Informática. Possui uma extensa experiência de ensino, investigação e engenharia nas áreas de informática, redes e protocolos de comunicação, planeamento e projeto de redes, redes móveis e redes de sensores. Membro da Ordem dos Engenheiros e do *Institute of Electrical and Electronics Engineers*. Coautor dos livros *Engenharia de Redes Informáticas, Administração de Redes Informáticas, TCP/IP – Teoria e Prática* e *Redes de Sensores sem Fios*, publicados pela FCA.

Mário Bernardes
Ex-Diretor do Serviço de Gestão de Sistemas e Infraestruturas de Informação e Comunicação da Universidade de Coimbra, onde desenvolveu atividade de planeamento, projeto e administração de sistemas e redes informáticos durante largos anos. Mestre (pré-Bolonha) em Arquitetura de Sistemas e Tecnologias de Informação, pela Faculdade de Ciências e Tecnologia da Universidade de Coimbra. Participa regularmente como docente convidado em cursos de mestrado e pós-graduação na área das tecnologias da informação e comunicação. Como consultor, tem ainda desempenhado a atividade de avaliador de projetos. Coautor dos livros *Administração de Redes Informáticas* e *TCP/IP* – Teoria e Prática, publicados pela FCA.

1 INTRODUÇÃO

A criptografia é, hoje em dia, indispensável para o funcionamento das redes, dos sistemas informáticos e das aplicações, por forma a garantir a integridade dos dados, a autenticação dos utilizadores, o controlo de acesso a recursos e a confidencialidade da informação. O estudo dos mecanismos que lhe estão subjacentes, dos algoritmos que a suportam, das políticas e procedimentos que garantem a sua utilização eficaz e, ainda, dos serviços que presta são, assim, fundamentais para a compreensão global dos sistemas criptográficos. O presente livro foi escrito tendo em vista esse estudo, que se revela fundamental para engenheiros, especialistas e técnicos que desenvolvam a sua atividade na área das tecnologias da informação e comunicação.

Neste capítulo introdutório começa-se por detalhar a motivação e âmbito do presente livro. De seguida, explicam-se os seus objetivos, bem como a abordagem adotada para os atingir, e identifica-se o público-alvo. Por fim, descreve-se a estrutura de capítulos utilizada, fazendo-se uma breve introdução a cada um deles.

1.1 MOTIVAÇÃO E ÂMBITO

O que é a criptografia e como evoluiu ao longo dos tempos? Quais os seus princípios fundamentais? Que tipo de ataques existem? Qual a diferença entre números aleatórios e números pseudoaleatórios? O que é uma chave criptográfica e qual a sua força? O que é a criptografia simétrica? E a criptografia assimétrica? Que modos de encriptação existem? O que são redes de substituição-permutação? O que são as redes de Feistel? Quais os principais algoritmos para encriptação de blocos? O que é uma função de *hash* criptográfica e quais os principais algoritmos usados para a sua implementação? O que são, para que servem e como funcionam os códigos de autenticação de mensagens? Que algoritmos de encriptação assimétrica existem e como funcionam? O que é e para que serve uma infraestrutura de chaves públicas? Quais as principais aplicações da criptografia? O que são e como funcionam as *blockchains*? Que soluções existem para a gestão de chaves criptográficas? Qual a importância das políticas e procedimentos na gestão de sistemas criptográficos? Que critérios devem ser utilizados na seleção de produtos criptográficos? A resposta a estas e a várias outras questões relacionadas é a principal motivação para o presente texto.

A criptografia é uma disciplina que conta com muitos séculos de história. Conhecer a história da criptografia, o seu contexto, os seus princípios fundamentais, bem como os seus sucessos e fracassos é fundamental para a robustez dos atuais algoritmos e sistemas criptográficos. Um bom algoritmo criptográfico deve ter em atenção os potenciais tipos de ataques e representar um compromisso equilibrado e eficaz entre a complexidade que

lhe está subjacente e o desempenho dos sistemas que terão de executar as operações de encriptação e desencriptação.

A criptografia simétrica, na qual uma mesma chave secreta é utilizada para encriptar e desencriptar texto em claro, assenta num conjunto de técnicas e modos de encriptação que importa conhecer. Neste contexto, é fundamental perceber as vantagens e desvantagens de utilizar uma mesma chave para cifragem e decifragem, bem como os aspetos positivos e negativos da encriptação em blocos *versus* encriptação em sequência.

No que diz respeito à encriptação de blocos, os algoritmos existentes utilizam uma de duas técnicas básicas: redes de substituição-permutação ou redes de Feistel. A primeira é utilizada pelo *Data Encryption Standard* (DES) e suas variantes, e a segunda está na base da norma mais atual no que respeita a criptografia simétrica, o *Advanced Encryption Standard* (AES).

A garantia da integridade de documentos recorre, em geral, à utilização de funções de *hash* criptográficas que, dado um conjunto de dados de entrada, produzem um código de saída. Dados de entrada diferentes conduzirão, com elevada probabilidade, a códigos de *hash* diferentes, não sendo possível realizar a operação inversa, isto é, determinar os dados de entrada a partir do valor da função de *hash*. Para além da garantia da integridade de documentos, as funções de *hash* podem também ser usadas para autenticação de mensagens. Os códigos de autenticação de mensagens também podem ser gerados com recurso a algoritmos de cifragem de blocos.

A generalização da utilização de algoritmos criptográficos, essencial para a Internet e para as atuais aplicações, só foi atingida com o desenvolvimento de algoritmos de criptografia assimétrica, nos quais se utilizam duas chaves complementares: uma chave privada, que nunca é divulgada, e uma chave pública, que pode ser livremente acedida. Este tipo de criptografia facilita os mecanismos de gestão e distribuição de chaves, permitindo que utilizadores e/ou entidades que não se conhecem nem têm confiança mútua troquem informação de forma segura.

A criptografia assimétrica apoia-se, na prática, em infraestruturas de chaves públicas (*Public Key Infrastructures* – PKI), que são essenciais para a garantia da autenticidade e para a distribuição de chaves de forma generalizada. Estas infraestruturas assentam num conjunto de componentes e organizam-se de acordo com determinados tipos de arquiteturas. As PKI assentam numa série de estruturas de dados essenciais para a sua operação e gestão.

As soluções criptográficas existentes foram desenvolvidas para o suporte de aplicações, que existem em elevado número e que se organizam em dois tipos essenciais: aplicações de sistema e aplicações de rede. As primeiras têm em vista a garantia da confidencialidade, integridade, autenticação e controlo de acesso a sistemas informáticos, serviços ou documentos digitais. As segundas têm por finalidade a proteção das redes, das comunicações

e dos serviços que elas suportam. De entre o enorme conjunto de aplicações que recorrem à criptografia, as *blockchains* têm vindo a ser alvo de atenção crescente, sendo também abordadas no presente texto.

Tal como qualquer outro componente de uma arquitetura de segurança, a criptografia não resolve por si só os problemas de segurança de sistemas e redes. A eficácia dos mecanismos e algoritmos de criptografia exige a adoção de políticas e procedimentos de segurança adequados, que devem ser escrupulosamente cumpridos. Naturalmente, um dos aspetos críticos da criptografia é o da gestão de chaves, sejam elas públicas ou privadas, já que a força de qualquer algoritmo criptográfico reside nas chaves utilizadas. É, assim, fundamental dispor de soluções de gestão de chaves, políticas e procedimentos que não ponham em causa o "edifício" da criptografia.

Para além dos aspetos anteriormente referidos, é fundamental conhecer a normalização existente na área, pois essa normalização está, em regra, na base de produtos criptográficos robustos. Naturalmente, existem outros critérios para a seleção de produtos, que não devem ser descurados, como sejam a certificação e a validação por parte da comunidade criptográfica. Por fim, há que estar atento aos desenvolvimentos e ameaças que surgem constantemente. O âmbito do presente livro abrange todos os aspetos atrás referidos.

1.2 OBJETIVOS E ABORDAGEM

São objetivos do presente livro a apresentação, análise e estudo dos aspetos essenciais da criptografia, que vão desde os seus fundamentos até aos modernos sistemas criptográficos, incluindo os principais algoritmos, políticas e procedimentos. Trata-se, portanto, de um texto abrangente, que fornece uma visão simultaneamente global e detalhada da criptografia atual, sem a qual não seria possível o funcionamento dos sistemas informáticos.

Pretende-se, desta forma, não só introduzir o leitor à criptografia e seus mecanismos, mas também aprofundar todas as questões essenciais para a compreensão do modo como esta funciona.

Neste sentido, a abordagem adotada assenta em dois pilares: uma componente teórica e uma componente prática.

Primeiramente, nos Capítulos 2 a 10, os conceitos são explicados com a profundidade necessária à sua compreensão, de forma a que o leitor perceba os problemas e as respetivas soluções. Estes capítulos podem ser lidos/estudados sem quaisquer preocupações em termos de abordagem ou estudo práticos.

Ao longo do livro, para exemplificar os conceitos apresentados, optou-se por utilizar o OpenSSL (www.openssl.org) que, para além de uma implementação dos protocolos *Transport Layer Security* (TLS) e *Secure Sockets Layer* (SSL), inclui uma biblioteca

criptográfica de uso geral e uma aplicação utilizável em modo de linha de comando (`openssl`) que permite explorar as funções criptográficas implementadas. O OpenSSL é incluído na generalidade das distribuições de sistemas Unix (Linux, macOS, BSD, Solaris), estando também disponíveis binários pré-compilados para sistemas Microsoft Windows. O OpenSSL tem um certificado de conformidade com a norma de segurança FIPS 140-2 emitido pelo programa *Cryptographic Module Validation Program* (CMVP) do *National Institute of Standards and Technology* (NIST).

A aplicação de linha de comando pode ser utilizada para diversos fins, tais como a criação e gestão de chaves públicas e privadas, criação de certificados digitais, encriptação e desencriptação, produção de *digests* de mensagens, teste de servidores SSL/TLS, entre outros. Uma vez instalado o OpenSSL, pode verificar-se a versão instalada:

```
$ openssl version
OpenSSL 1.0.2k-fips 26 Jan 2017
```

Para analisar as funcionalidades oferecidas podem usar-se os comandos OpenSSL:

```
list-standard-commands
list-cipher-commands
list-message-digest-commands
list-cipher-algorithms
list-message-digest-algorithms
list-public-key-algorithms
```

Para além das *man pages* e das páginas do OpenSSL (www.openssl.org), existe um grande volume de informação disponível, impressa ou em formato digital, da qual se destaca o "OpenSSL Command-Line HOWTO" (https://www.madboa.com/geek/openssl/).

Complementarmente, no final de cada um dos referidos capítulos é apresentado um conjunto de questões e exercícios, de natureza teórica, teórico-prática ou prática, com os objetivos de, por um lado, consolidar os conceitos apresentados ao longo do capítulo e, por outro, potenciar o aprofundamento e exploração autónomos das matérias abordadas. Sempre que, nos exercícios, for utilizado código ou ficheiros auxiliares, eles poderão ser encontrados na página do livro em www.fca.pt.

Tendo em atenção os aspetos anteriormente referidos, podem ser identificados vários públicos-alvo. Podem beneficiar deste livro docentes e estudantes de licenciatura e mestrado em disciplinas na área dos sistemas e redes de computadores, redes de telecomunicações, gestão de sistemas e redes e, naturalmente, segurança. O livro é também adequado aos profissionais com responsabilidades na instalação e administração de sistemas e redes informáticos em empresas de pequena, média e grande dimensão, em operadores de telecomunicações, em fornecedores de serviços IP e na administração pública.

1.3 ORGANIZAÇÃO DO PRESENTE TEXTO

Além do Capítulo 1, no qual são apresentadas as motivações, âmbito, objetivos e público-alvo, o livro é composto por 10 capítulos que abordam, sucessivamente, os fundamentos da criptografia, os princípios subjacentes à criptografia simétrica, a encriptação de blocos de dados, as funções de *hash* criptográficas, os códigos de autenticação de mensagens, a criptografia assimétrica, as infraestruturas de chaves públicas, as aplicações da criptografia, a gestão de chaves, políticas e procedimentos e, por fim, a normalização, os critérios para a seleção de produtos e as perspetivas de evolução dos sistemas criptográficos.

O Capítulo 2 – intitulado "Fundamentos" – introduz e explica uma série de aspetos essenciais da criptografia, nomeadamente o que é e como evoluiu ao longo dos séculos, o que é um algoritmo criptográfico seguro, o que estabelece o princípio de Kerckhoffs, qual a relação entre criptografia e matemática, quais as diferenças entre aleatoriedade e pseudoaleatoriedade, o conceito de força de uma chave criptográfica, em que consiste a autenticação, o que são chaves públicas e chaves privadas, os tipos de ataques, e os conceitos de desempenho e de complexidade dos algoritmos criptográficos.

O Capítulo 3 fornece uma visão geral da criptografia simétrica. Começa por apresentar os princípios gerais deste tipo de criptografia passando, seguidamente, para a apresentação de uma série de técnicas clássicas de encriptação. Discute-se depois o conceito de segurança computacional. Na parte final do capítulo apresentam-se e analisam-se os dois modos básicos de encriptação simétrica: o modo de encriptação em blocos e o modo de encriptação em sequência.

Dada a importância da encriptação simétrica de blocos de dados, o Capítulo 4 é dedicado a este tipo de encriptação. Inicialmente, são explicadas as abordagens de encriptação baseadas em redes de substituição-permutação e em redes de Feistel. Em seguida, explicam-se detalhadamente as normas DES e AES para encriptação de blocos. Por fim, identificam-se outros algoritmos de encriptação de blocos.

O Capítulo 5 é dedicado às funções de *hash* criptográficas. Começa-se por explicar o conceito e identificar os requisitos que este tipo de funções devem satisfazer. Seguidamente, são apresentados e discutidos alguns exemplos de funções de *hash*, incluindo funções baseadas em cifragem de cadeias de blocos e o algoritmo MD-5. O capítulo termina com uma secção dedicada ao *Secure Hash Algorithm* (SHA) e suas variantes, por se tratar da norma mais atual para funções de *hash* criptográficas.

O Capítulo 6 aborda as soluções para geração de códigos de autenticação de mensagens. Após uma breve apresentação do conceito, objetivos e principais abordagens, apresentam-se soluções para a implementação de códigos de autenticação de mensagens baseados em funções de *hash* e em cifragem de cadeias de blocos. O capítulo inclui, ainda, a

apresentação de uma solução para encriptação autenticada, bem como algumas soluções para a geração de números pseudoaleatórios.

O Capítulo 7 é dedicado à criptografia assimétrica, que está na base da generalização da utilização de criptografia na sociedade atual. Começa-se por apresentar os princípios que lhe estão subjacentes, a terminologia, utilizações típicas e os requisitos. Subsequentemente, apresenta-se o algoritmo que atualmente é mais utilizado para criptografia assimétrica, nomeadamente, o algoritmo Rivest, Shamir, and Adleman (RSA). Também são brevemente apresentados outros algoritmos de encriptação assimétrica. Na parte final do capítulo aborda-se a geração de números pseudoaleatórios com base no algoritmo RSA.

A criptografia assimétrica requer, na prática, a existência de sistemas de apoio, denominados infraestruturas de chaves públicas. Estas infraestruturas são abordadas no Capítulo 8, que detalha os seus componentes, arquitetura e estruturas de dados. O capítulo abrange ainda aspetos organizacionais e práticos deste tipo de infraestruturas.

No Capítulo 9 é apresentado o vasto leque de aplicações da criptografia, segundo duas perspetivas essenciais: aplicações de sistema e aplicações de rede. As primeiras incluem a autenticação forte, a encriptação de dispositivos, as assinaturas digitais, a encriptação de conteúdos, o acesso físico a recursos e instalações, e as federações de identidade. As segundas incluem toda uma série de protocolos e aplicações para garantir a integridade e confidencialidade das comunicações na Internet. Dada a sua crescente importância, o capítulo aborda, ainda, a criptografia em ambientes de computação em nuvem (também designado em inglês por *cloud computing*) e a tecnologia *blockchain*, esta última com particular detalhe.

A gestão de chaves criptográficas, essencial para a robustez de qualquer solução de criptografia, é abordada no Capítulo 10, sendo dada particular atenção à solução Kerberos. Para além disso, o capítulo aborda aspetos importantes das políticas, procedimentos e práticas de gestão de material criptográfico, já que são, também, aspetos essenciais de qualquer sistema.

Por fim, o Capítulo 11 é dedicado à apresentação e discussão de uma série de aspetos de grande importância que, frequentemente, são descurados por parecerem secundários. O primeiro é o da importância da normalização das soluções criptográficas, como forma de garantir a sua qualidade e robustez a ataques. O segundo é o dos critérios a utilizar na seleção de produtos, já que a vasta maioria dos responsáveis e especialistas de segurança terá a seu cargo não o desenvolvimento, mas sim a seleção, aquisição e integração de sistemas de criptografia. O capítulo aborda ainda as oportunidades e ameaças que se abrem com a computação e criptografia quânticas, cujo desenvolvimento se antevê para os próximos 10 a 15 anos.

Tal como referido na secção 1.2, no final de cada capítulo é apresentado um conjunto de questões e exercícios propostos. No caso de os exercícios requererem ferramentas informáticas, serão privilegiadas as ferramentas de acesso livre.

2 FUNDAMENTOS

Nos dias de hoje, a criptografia está em todo o lado, sendo absolutamente essencial para o funcionamento da sociedade tal como a conhecemos. Sem a criptografia não poderíamos utilizar cartões bancários, não poderíamos realizar operações de comércio eletrónico, não poderíamos utilizar telefones móveis, a Internet não poderia funcionar, não poderíamos confiar nos sistemas operativos e em todo o *software* que é executado nos computadores, não poderíamos aceder a qualquer tipo de conta, etc. É claro que nem sempre foi assim e, por isso, importa não só perceber a evolução histórica da criptografia mas, também, conhecer os seus fundamentos. São esses os objetivos do presente capítulo.

O capítulo começa por explicar o que é a criptografia e fornecer uma perspetiva histórica desta disciplina. Seguidamente, são apresentados alguns dos seus princípios e conceitos mais importantes. Discutem-se, ainda, os principais tipos de ataques, bem como aspetos de desempenho e complexidade dos algoritmos criptográficos. Todos os aspetos abordados são fundamentais, sendo transversais às matérias abordadas no livro.

2.1 O QUE É A CRIPTOGRAFIA?

A palavra criptografia, de origem grega, é composta por dois elementos básicos. O primeiro – cripto – significa oculto ou secreto e o segundo – grafia – significa escrita. A criptografia é, portanto, a escrita de textos, ou dados, de forma oculta ou secreta, isto é, de forma a que só quem de direito possa entender a informação cifrada[1].

Ao longo da história a criptografia foi encarada de diferentes formas. Inicialmente, era entendida como uma arte – a arte de esconder informação, através da utilização de um segredo, ou seja, a arte de escrever e resolver códigos. Em parte, essa forma de ver a criptografia ainda é partilhada por muitos. Mais tarde, essencialmente até ao século XIX, era encarada com uma ciência, dada a crescente base científica que a sustentava. Atualmente, é encarada como uma disciplina da matemática e das ciências da computação, já que a criptografia moderna recorre de forma intensiva a estes dois ramos da ciência e não seria possível sem qualquer um deles.

A visão atual sobre a criptografia começou a surgir nas décadas de 1970 e 1980, à medida que se foi desenvolvendo uma base teórica sólida para esta disciplina. No entanto, só na década de 1990 é que a criptografia deixou o âmbito meramente militar e governamental, passando a ter uma aplicação verdadeiramente global. Para isso contribuíram de forma decisiva, por um lado, o aparecimento da criptografia assimétrica – que abordaremos no Capítulo 7 – e, por outro, a generalização do uso de sistemas computacionais.

[1] Ao longo do livro utilizaremos indistintamente os verbos cifrar e encriptar – o primeiro deles de origem italiana e, portanto, latina, e o segundo de origem grega – apesar de se poderem encontrar acesas discussões sobre as diferenças entre estes dois termos.

A Figura 2.1 apresenta um cenário muito simples, no qual dois utilizadores – Alice e Bruno – pretendem trocar uma mensagem, 'm', sem que terceiros possam aceder ao seu conteúdo. O que terão de fazer para evitar que, por exemplo, a utilizadora Eva tenha acesso ao conteúdo da mensagem?

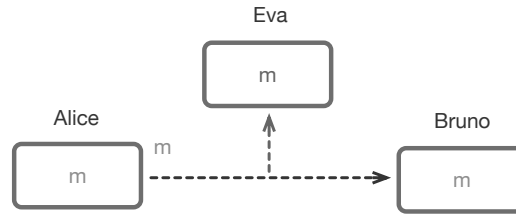

Figura 2.1 • Cenário de envio de mensagem não encriptada

Uma solução simples para garantia da confidencialidade das comunicações entre a Alice e o Bruno será a encriptação de todas as mensagens, utilizando uma chave secreta, conhecida apenas pelos dois, e um algoritmo de encriptação.

Seja:

m – o texto da mensagem, não encriptado
k – uma chave secreta conhecida apenas pelos interlocutores
E – o algoritmo de encriptação utilizado
D – o correspondente algoritmo de desencriptação
c – o texto cifrado (ou encriptado)

Então pode escrever-se:

$c = E (k, m)$
$m = D (k, c)$
$D (k, E (k, m)) = m$

Ao enviar-se a mensagem encriptada, 'c', qualquer outro utilizador sem acesso à chave secreta será incapaz de decifrar, ou seja, entender, o seu conteúdo, pelo que a confidencialidade será preservada. Este cenário de confidencialidade é ilustrado na Figura 2.2.

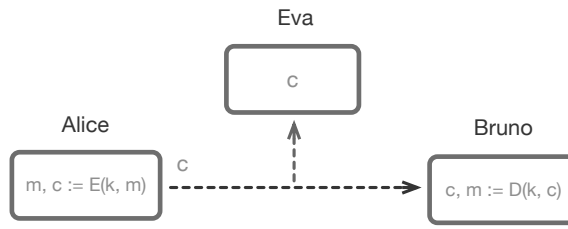

Figura 2.2 • Cenário de envio de mensagem encriptada

Naturalmente, ficam por analisar neste cenário simplificado aspetos de extrema importância, que serão alvo de atenção ao longo do presente livro. Por exemplo, que algoritmo de encriptação utilizar para garantir determinado nível de proteção? Como será gerada e partilhada a chave secreta? A que ataques está sujeito este cenário?

Qualquer sistema de segurança é tão forte como o seu componente mais fraco, pelo que é crucial que as soluções criptográficas sejam as mais robustas possíveis. De facto, se a criptografia for quebrada, o risco de um ataque passar despercebido é muito elevado, pelo que esta se reveste de um carácter crítico. Mais ainda, uma má solução criptográfica cria uma falsa sensação de segurança, o que pode tornar ainda mais difícil detetar e impedir ataques.

Para além disso, há que ter em conta que a criptografia é apenas uma parte de qualquer sistema de segurança, sendo que aspetos como as políticas e os procedimentos a utilizar – a abordar no Capítulo 10 – são, no mínimo, tão críticos como a robustez das soluções criptográficas em si.

2.2 PERSPETIVA HISTÓRICA

O desejo e a necessidade de encriptação de textos existem praticamente desde o aparecimento da comunicação escrita, ela própria uma forma de codificar informação, podendo ser encontrados exemplos no Egito, Mesopotâmia, Grécia e Roma.

Uma forma muito simples e muito conhecida de encriptação, que recorre a técnicas de substituição de caracteres, foi utilizada no império romano, sendo atribuída a Júlio César (apesar de existirem evidências de utilização anterior), que a utilizava para envio de mensagens de natureza militar aos seus generais. A cifragem de César, como é conhecida, consiste na substituição de cada letra pela letra que se encontra três posições à frente no alfabeto. Assim, a letra "A" será substituída pela letra "D", a letra "B" pela letra "E" e assim sucessivamente. Neste esquema, é curioso notar que a chave secreta (o número de posições a deslocar no alfabeto) está embutida no próprio algoritmo (a substituição de uma letra por outra que se obtém por deslocação de três posições no alfabeto), o que não é uma boa prática, pois chave e algoritmo devem estar sempre completamente dissociados. Curiosamente, esta técnica simples foi utilizada com êxito durante alguns séculos.

Uma variante um pouco mais forte desta forma de encriptação é a cifragem por deslocamento (*shift cipher*), na qual o algoritmo se mantém, mas a chave – ou seja, o número de posições a deslocar – é mantida secreta não tendo, portanto, o valor fixo de três unidades. Considerando um alfabeto de 26 caracteres, a chave pode tomar 26 valores distintos, de 0 a 25, ou seja, sendo 'i' a posição (isto é, o índice) do carácter no texto em claro ou no texto cifrado e 'mod' a operação módulo,

$$0 \leq k \leq 25$$

3

VISÃO GERAL DA CRIPTOGRAFIA SIMÉTRICA

A utilização de uma mesma chave secreta para cifrar e decifrar informação está na génese da criptografia. Esta abordagem da criptografia, denominada simétrica precisamente pelo facto de as funções criptográficas direta e inversa usarem uma só chave, é utilizada por alguns dos mais seguros algoritmos criptográficos atualmente existentes e é o objeto principal do presente capítulo. Assim, começa-se por apresentar os princípios gerais da criptografia simétrica. Segue-se uma resenha das principais técnicas clássicas de encriptação e suas fragilidades, apresentando-se, de seguida, o conceito de segurança computacional, essencial para os modernos sistemas criptográficos. O núcleo do capítulo é dedicado à apresentação de uma série de algoritmos, que se organizam em dois modos básicos de encriptação: o modo sequencial e o modo de blocos.

3.1 PRINCÍPIOS GERAIS

Tal como vimos na secção 2.3.7, existem duas formas fundamentais de criptografia: a criptografia simétrica e a criptografia assimétrica. Sendo esta última mais recente e estando na base da utilização generalizada da criptografia, poder-se-ia pensar que iria tornar obsoleta a primeira, mas a realidade é bem distinta. De facto, a criptografia simétrica é – e, provavelmente, continuará a ser por muito tempo – a forma mais utilizada de criptografia, pelo simples facto de que é computacionalmente muito menos pesada do que a criptografia assimétrica. Por outro lado, a criptografia assimétrica tem vantagens consideráveis, principalmente no que toca à facilidade com que permite a distribuição de chaves. Com efeito, ambas as formas de criptografia são complementares, pelo que os aspetos positivos de cada uma delas devem ser simultaneamente explorados.

A característica fundamental da criptografia simétrica é o facto de a mesma chave ser utilizada para as operações de cifragem e decifragem, devendo, portanto, essa chave ser mantida secreta. Um cenário típico de utilização desta forma de criptografia é o que se representa na Figura 3.1, no qual a utilizadora Alice pretende enviar ao utilizador Bruno uma mensagem cujo conteúdo deve permanecer secreto. Para tal, a Alice cifra a mensagem usando um algoritmo criptográfico apropriado e uma chave que apenas ela e o utilizador Bruno conhecem. Seguidamente, envia a mensagem cifrada. Por sua vez, o Bruno recebe a mensagem cifrada e, utilizando a chave secreta que comparte com a Alice, poderá aplicar o algoritmo inverso para decifrar a mensagem recebida, obtendo, desta forma, a mensagem original em claro.

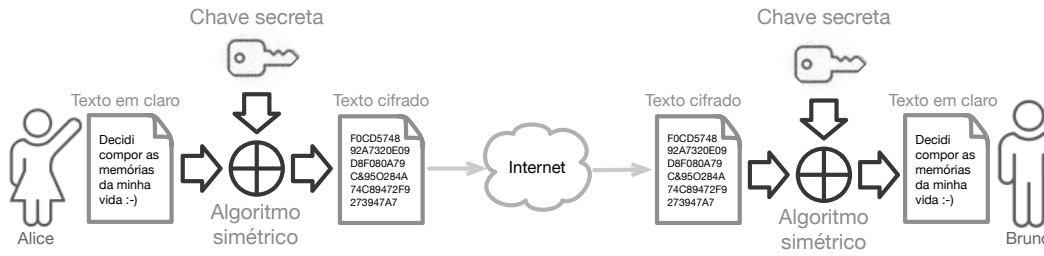

Figura 3.1 • Cenário de utilização de criptografia simétrica

Note-se que, de acordo com o princípio de Kerckhoffs (secção 2.3.2) o algoritmo de encriptação pode e deve ser bem conhecido, sendo que tal facto não pode acarretar qualquer fragilidade adicional para o sistema. De facto, a força da criptografia simétrica depende de dois aspetos:

- O algoritmo deve ser suficientemente robusto para não poder ser quebrado;
- A chave criptográfica tem de ser mantida em segredo.

É por este motivo que a criptografia simétrica também é conhecida como criptografia de chave secreta.

O principal desafio que se coloca à criptografia simétrica decorre, precisamente, da necessidade de distribuir chaves de forma segura. Com efeito, no cenário da Figura 3.1, a Alice e o Bruno tiveram de partilhar de alguma forma a chave secreta entre eles. Para tal, poderão ter-se encontrado ou poderão ter utilizado um canal de comunicação complementar. Por exemplo, poderão ter telefonado um ao outro ou enviado uma carta. É claro que essas alternativas, sendo possíveis para um número limitado de utilizadores (embora pouco práticas), não o são na generalidade dos casos, nos quais o número de pares comunicantes é muito elevado e os intervenientes na comunicação poderão não dispor de canais alternativos ou nem se conhecer.

No presente capítulo, assumiremos que o problema da distribuição segura da chave se encontra resolvido, o que nos permitirá concentrarmo-nos nos algoritmos de encriptação. A distribuição e gestão de chaves criptográficas será abordada parcialmente nos Capítulos 7 e 8, e também no Capítulo 10.

3.2 TÉCNICAS CLÁSSICAS DE ENCRIPTAÇÃO

Até ao aparecimento e vulgarização dos computadores, desenvolveu-se e utilizou-se amplamente uma série de técnicas de encriptação simétrica, aqui designadas por técnicas

clássicas, que abordaremos brevemente nesta secção. Estas técnicas organizam-se em três categorias principais:

- Técnicas de substituição;
- Técnicas de transposição;
- Máquinas de rotores.

Hoje em dia, os tipos de cifragem referidos nesta secção têm um valor meramente histórico, já que são extremamente vulneráveis face ao poder computacional presentemente disponível nos mais vulgares equipamentos informáticos.

3.2.1 Técnicas de substituição

No caso das técnicas de substituição, explora-se a substituição de um ou mais caracteres por outro ou outros, em igual número, sendo essa substituição determinada pelo algoritmo e/ou pela chave criptográfica utilizados. Enquadram-se nesta categoria a cifragem de César, a cifragem por deslocamento e a cifragem monoalfabética (já abordadas na secção 2.2), a cifragem polialfabética, a cifragem de Vigenère e a técnica *one-time pad* (OTP), entre outras.

3.2.1.1 Primeiras técnicas

As primeiras técnicas de criptografia simétrica conhecidas exploram o conceito de substituição, sendo bastante simples, mas, surpreendentemente, foram utilizadas durante largos períodos de tempo na antiguidade. Nestas enquadram-se a cifragem de César, a cifragem por deslocamento e a cifragem monoalfabética, mas também várias outras que não abordaremos pelo facto de o seu interesse ser, atualmente, mais histórico do que técnico e não ser esse o âmbito principal do presente livro.

Uma forma inicial de encriptação por substituição – por muitos reconhecida como a primeira e que foi popularizada através do famoso livro de Dan Brown, "O Código Da Vinci" – não abordada na secção 2.2, é a cifragem de Atbash, que data do século VI a.C. Originalmente foi utilizada para cifrar as 22 letras do alfabeto hebreu, mas pode ser aplicada a qualquer alfabeto. O algoritmo consiste na simples substituição da primeira letra do alfabeto pela última, da segunda letra pela penúltima, da terceira letra pela antepenúltima, e assim sucessivamente. Uma forma muito simples de construir uma tabela que auxilie a cifragem e decifragem é escrever na primeira linha da tabela a primeira metade do alfabeto na sua ordem natural e na segunda linha escrever a segunda metade do alfabeto na ordem inversa. Esse método ilustra-se na Tabela 3.1, usando o alfabeto latino em vez do alfabeto hebreu.

4 ENCRIPTAÇÃO DE BLOCOS

Vimos no Capítulo 3 que a encriptação em blocos é a forma mais comum de encriptação, dadas as limitações práticas da encriptação em sequência. Nesse capítulo abordámos as principais técnicas de encriptação em blocos sem detalharmos, no entanto, a função de encriptação usada para cifrar cada bloco individualmente. Este é o objetivo central do presente capítulo.

Neste capítulo começa-se por apresentar duas abordagens essenciais na encriptação de blocos: as redes de substituição-permutação e as redes de Feistel, já que estas estão na base dos atuais algoritmos de encriptação. Segue-se uma apresentação detalhada dos dois principais algoritmos de encriptação, nomeadamente o DES (incluindo aqui o TripleDES) e o AES. Referem-se, subsequentemente, outros algoritmos de encriptação de blocos.

4.1 INTRODUÇÃO

Tal como vimos no Capítulo 3, existem diversas técnicas de encriptação em modo de blocos como, por exemplo, as técnicas ECB, CBC, OFB e CTR. Todas elas requerem a utilização de uma função de encriptação de blocos. Também vimos no Capítulo 3 que as funções de encriptação de blocos operam sobre blocos de tamanho fixo, produzindo um texto cifrado que tem o mesmo tamanho do bloco de texto em claro. A generalidade dos atuais algoritmos de encriptação de blocos utiliza um tamanho de blocos de 128 bits.

Dado que as funções de encriptação utilizam algoritmos de encriptação simétrica, é usada a mesma chave secreta quer para a encriptação quer para a desencriptação. As chaves utilizadas têm, tipicamente, um tamanho de 128 ou 256 bits. A utilização de chaves longas é útil porque dificulta ou inviabiliza ataques de força bruta. Tal como veremos no presente capítulo, a chave secreta pode ser utilizada para gerar várias chaves derivadas, que serão utilizadas em diferentes iterações ou estágios da função de encriptação.

Em geral, as funções de encriptação de blocos são reversíveis – ou seja, é possível construir uma função inversa, utilizada para a desencriptação – embora, tal como vimos no Capítulo 3, algumas técnicas de encriptação de blocos operem também com funções não reversíveis. Um aspeto fundamental de qualquer função de encriptação é o de que o respetivo algoritmo deve ser bem conhecido, de acordo com o princípio de Kerckhoffs, residindo a força deste tipo de encriptação na forma como é construído e na chave secreta.

Idealmente, uma função de encriptação de blocos deve comportar-se como uma permutação aleatória. Note-se que sendo 'm' o número de bits de um bloco, então existirão '$(2^m)!$' permutações possíveis.

NOTA – O operador matemático '!' representa a operação fatorial. Assim, 'p! = p × (p-1) × (p-2) × ... × 2 × 1'.

Dado que nos algoritmos atuais 'm ≥ 128', o número de permutações possíveis é extremamente elevado, inviabilizando ataques de força bruta.

O desafio está, assim, na construção de funções de encriptação que, dependendo de uma chave secreta e executando um número relativamente pequeno de operações, se comportem como permutações aleatórias, ou seja, de uma maneira informal, funções para as quais a alteração de apenas um bit do texto em claro conduza a um texto cifrado para o qual a probabilidade de alteração de cada um dos seus bits seja de 50%. Funções deste tipo serão descritas nas secções 4.4 e 4.5. Entretanto, nas secções 4.2 e 4.3 apresentaremos alguns conceitos que lhes estão subjacentes e que são fundamentais para a sua compreensão.

4.2 REDES DE SUBSTITUIÇÃO-PERMUTAÇÃO

As redes de substituição-permutação (*Substitution-Permutation Networks* – SPN) têm por objetivo gerar permutações pseudoaleatórias a partir de um bloco de dados de entrada e de uma chave de encriptação. Nesta secção serão apresentados os aspetos principais deste tipo de redes.

4.2.1 Paradigma de confusão-difusão

O paradigma de confusão-difusão, desenvolvido pelo matemático, engenheiro eletrotécnico e criptógrafo americano Claude Shanon (n. 1916 – f. 2001), está na base da construção de permutações aparentemente aleatórias, sendo utilizado nas redes de substituição-permutação. Numa primeira fase – a fase de confusão – são executadas diversas permutações dos dados de entrada, para, na fase seguinte – a fase de difusão –, o efeito das permutações ser espalhado por todo o bloco.

A título de exemplo considere-se que a função F deve gerar uma permutação pseudoaleatória a partir de um bloco, 'x', utilizando uma chave 'k'. Tipicamente, a função F será decomposta em várias permutações, 'f_i', que operam sobre blocos mais pequenos, 'x_i', sendo os resultados dessas permutações parciais concatenados para produzir o valor final, '$F_k(x)$'. Tomando como exemplo um bloco de 128 bits, divididos em 16 sub-blocos de 8 bits, teremos:

$x \in \{0,1\}^{128}$

x_i são sub-blocos de 8 bits cada, com i = 1, ..., 16

f_i são permutações aplicadas a x_i, com i = 1, ..., 16
$F_k(x) = f_1(x_1) \,||\, ... \,||\, f_{16}(x_{16})$

'$F_k(x)$' é o resultado da fase de confusão, na qual os 128 bits que compõem o bloco 'x' foram permutados de acordo com as funções 'f_i'.

Na fase de difusão, espalha-se as alterações feitas localmente em cada um dos sub-blocos por todo o bloco. Para tal, os bits que compõem o resultado da fase de confusão, isto é, os bits de '$F_k(x)$' são reorganizados de acordo com um esquema de permutação. Usando como exemplo o caso dos blocos de 128 bits subdivididos em 16 sub-blocos de 1 byte, cada um dos bits do primeiro byte poderá ser permutado com os últimos bits dos bytes com os números de ordem 2, 4, 7, 10, 11, 13, 15 e 16, e assim sucessivamente para cada um dos 16 bytes do bloco.

Ao conjunto das fases de confusão e difusão chama-se iteração ou estágio. Em regra, uma função de encriptação é composta por vários estágios, de forma a melhorar o carácter pseudoaleatório do resultado. Por exemplo, numa função com dois estágios (o que é manifestamente pouco e serve apenas de exemplo) as operações seriam as seguintes:

- **Estágio 1:**
  - Confusão – calcular $F_k(x) = f_1(x_1) \,||\, ... \,||\, f_{16}(x_{16})$;
  - Difusão – reordenar os bits de $F_k(x)$ de modo a obter um novo bloco x';
- **Estágio 2:**
  - Confusão – calcular $F'_k(x') = f'_1(x'_1) \,||\, ... \,||\, f'_{16}(x'_{16})$;
  - Difusão – reordenar os bits de $F'_k(x')$ de modo a obter um novo bloco x".

Deve notar-se que é fundamental que a chave secreta desempenhe de alguma forma um papel na construção das funções $F_k(x)$ e $F'_k(x')$, por forma a que estas só sejam reversíveis para quem conhecer essa chave. Se tal não acontecesse, qualquer atacante poderia inverter as operações de difusão e de confusão, pois essas operações devem ser conhecidas devido ao princípio de Kerckhoffs.

4.2.2 Substituição-permutação

As redes de substituição-permutação (SPN) implementam o paradigma da confusão-difusão apresentado na secção 4.2.1, sendo compostas por múltiplos estágios. Em cada estágio, uma ou mais *substitution box* (S-box) implementam uma função de substituição, o que constitui a fase de confusão referida na mesma secção. É nesta fase que se incorpora a chave secreta, ou uma subchave derivada da chave secreta (ver secção 4.2.3). Tipicamente, essa incorporação é feita executando um OU exclusivo (XOR, $\oplus$) entre a chave e os bits de entrada, ou seja:

$f(x) = S (k \oplus x)$

4 ENCRIPTAÇÃO DE BLOCOS

Vimos no Capítulo 3 que a encriptação em blocos é a forma mais comum de encriptação, dadas as limitações práticas da encriptação em sequência. Nesse capítulo abordámos as principais técnicas de encriptação em blocos sem detalharmos, no entanto, a função de encriptação usada para cifrar cada bloco individualmente. Este é o objetivo central do presente capítulo.

Neste capítulo começa-se por apresentar duas abordagens essenciais na encriptação de blocos: as redes de substituição-permutação e as redes de Feistel, já que estas estão na base dos atuais algoritmos de encriptação. Segue-se uma apresentação detalhada dos dois principais algoritmos de encriptação, nomeadamente o DES (incluindo aqui o TripleDES) e o AES. Referem-se, subsequentemente, outros algoritmos de encriptação de blocos.

4.1 INTRODUÇÃO

Tal como vimos no Capítulo 3, existem diversas técnicas de encriptação em modo de blocos como, por exemplo, as técnicas ECB, CBC, OFB e CTR. Todas elas requerem a utilização de uma função de encriptação de blocos. Também vimos no Capítulo 3 que as funções de encriptação de blocos operam sobre blocos de tamanho fixo, produzindo um texto cifrado que tem o mesmo tamanho do bloco de texto em claro. A generalidade dos atuais algoritmos de encriptação de blocos utiliza um tamanho de blocos de 128 bits.

Dado que as funções de encriptação utilizam algoritmos de encriptação simétrica, é usada a mesma chave secreta quer para a encriptação quer para a desencriptação. As chaves utilizadas têm, tipicamente, um tamanho de 128 ou 256 bits. A utilização de chaves longas é útil porque dificulta ou inviabiliza ataques de força bruta. Tal como veremos no presente capítulo, a chave secreta pode ser utilizada para gerar várias chaves derivadas, que serão utilizadas em diferentes iterações ou estágios da função de encriptação.

Em geral, as funções de encriptação de blocos são reversíveis – ou seja, é possível construir uma função inversa, utilizada para a desencriptação – embora, tal como vimos no Capítulo 3, algumas técnicas de encriptação de blocos operem também com funções não reversíveis. Um aspeto fundamental de qualquer função de encriptação é o de que o respetivo algoritmo deve ser bem conhecido, de acordo com o princípio de Kerckhoffs, residindo a força deste tipo de encriptação na forma como é construído e na chave secreta.

Idealmente, uma função de encriptação de blocos deve comportar-se como uma permutação aleatória. Note-se que sendo 'm' o número de bits de um bloco, então existirão '$(2^m)!$' permutações possíveis.

NOTA – O operador matemático '!' representa a operação fatorial. Assim, 'p! = p × (p-1) × (p-2) × ... × 2 × 1'.

Dado que nos algoritmos atuais 'm ≥ 128', o número de permutações possíveis é extremamente elevado, inviabilizando ataques de força bruta.

O desafio está, assim, na construção de funções de encriptação que, dependendo de uma chave secreta e executando um número relativamente pequeno de operações, se comportem como permutações aleatórias, ou seja, de uma maneira informal, funções para as quais a alteração de apenas um bit do texto em claro conduza a um texto cifrado para o qual a probabilidade de alteração de cada um dos seus bits seja de 50%. Funções deste tipo serão descritas nas secções 4.4 e 4.5. Entretanto, nas secções 4.2 e 4.3 apresentaremos alguns conceitos que lhes estão subjacentes e que são fundamentais para a sua compreensão.

4.2 REDES DE SUBSTITUIÇÃO-PERMUTAÇÃO

As redes de substituição-permutação (*Substitution-Permutation Networks* – SPN) têm por objetivo gerar permutações pseudoaleatórias a partir de um bloco de dados de entrada e de uma chave de encriptação. Nesta secção serão apresentados os aspetos principais deste tipo de redes.

4.2.1 Paradigma de confusão-difusão

O paradigma de confusão-difusão, desenvolvido pelo matemático, engenheiro eletrotécnico e criptógrafo americano Claude Shanon (n. 1916 – f. 2001), está na base da construção de permutações aparentemente aleatórias, sendo utilizado nas redes de substituição-permutação. Numa primeira fase – a fase de confusão – são executadas diversas permutações dos dados de entrada, para, na fase seguinte – a fase de difusão –, o efeito das permutações ser espalhado por todo o bloco.

A título de exemplo considere-se que a função F deve gerar uma permutação pseudoaleatória a partir de um bloco, 'x', utilizando uma chave 'k'. Tipicamente, a função F será decomposta em várias permutações, 'f_i', que operam sobre blocos mais pequenos, 'x_i', sendo os resultados dessas permutações parciais concatenados para produzir o valor final, '$F_k(x)$'. Tomando como exemplo um bloco de 128 bits, divididos em 16 sub-blocos de 8 bits, teremos:

$x \in \{0,1\}^{128}$

x_i são sub-blocos de 8 bits cada, com i = 1, ..., 16

f_i são permutações aplicadas a x_i, com i = 1, ..., 16
$F_k(x) = f_1(x_1) \,||\, ... \,||\, f_{16}(x_{16})$

'$F_k(x)$' é o resultado da fase de confusão, na qual os 128 bits que compõem o bloco 'x' foram permutados de acordo com as funções 'f_i'.

Na fase de difusão, espalha-se as alterações feitas localmente em cada um dos sub-blocos por todo o bloco. Para tal, os bits que compõem o resultado da fase de confusão, isto é, os bits de '$F_k(x)$' são reorganizados de acordo com um esquema de permutação. Usando como exemplo o caso dos blocos de 128 bits subdivididos em 16 sub-blocos de 1 byte, cada um dos bits do primeiro byte poderá ser permutado com os últimos bits dos bytes com os números de ordem 2, 4, 7, 10, 11, 13, 15 e 16, e assim sucessivamente para cada um dos 16 bytes do bloco.

Ao conjunto das fases de confusão e difusão chama-se iteração ou estágio. Em regra, uma função de encriptação é composta por vários estágios, de forma a melhorar o carácter pseudoaleatório do resultado. Por exemplo, numa função com dois estágios (o que é manifestamente pouco e serve apenas de exemplo) as operações seriam as seguintes:

- **Estágio 1:**
  - Confusão – calcular $F_k(x) = f_1(x_1) \,||\, ... \,||\, f_{16}(x_{16})$;
  - Difusão – reordenar os bits de $F_k(x)$ de modo a obter um novo bloco x';
- **Estágio 2:**
  - Confusão – calcular $F'_k(x') = f'_1(x'_1) \,||\, ... \,||\, f'_{16}(x'_{16})$;
  - Difusão – reordenar os bits de $F'_k(x')$ de modo a obter um novo bloco x".

Deve notar-se que é fundamental que a chave secreta desempenhe de alguma forma um papel na construção das funções $F_k(x)$ e $F'_k(x')$, por forma a que estas só sejam reversíveis para quem conhecer essa chave. Se tal não acontecesse, qualquer atacante poderia inverter as operações de difusão e de confusão, pois essas operações devem ser conhecidas devido ao princípio de Kerckhoffs.

4.2.2 Substituição-permutação

As redes de substituição-permutação (SPN) implementam o paradigma da confusão-difusão apresentado na secção 4.2.1, sendo compostas por múltiplos estágios. Em cada estágio, uma ou mais *substitution box* (S-box) implementam uma função de substituição, o que constitui a fase de confusão referida na mesma secção. É nesta fase que se incorpora a chave secreta, ou uma subchave derivada da chave secreta (ver secção 4.2.3). Tipicamente, essa incorporação é feita executando um OU exclusivo (XOR, $\oplus$) entre a chave e os bits de entrada, ou seja:

$f(x) = S (k \oplus x)$