

Criptografia

Mestrado Integrado em Engenharia de Telecomunicações e
Informática

J. Bacelar Almeida (jba@di.uminho.pt)

Departamento de Informática
Universidade do Minho

Parte I

Conceitos Preliminares

O que é a Criptografia?

- Historicamente, a **Criptografia** consiste no estudo de técnicas que procuram tornar possível a comunicação secreta entre duas partes sobre um canal aberto.
- Por extensão, a criptografia hoje procura dar resposta a numerosas propriedades do processo de comunicativo para além da *confidencialidade* (segredo), como sejam:
 - *integridade*;
 - *autenticidade*;
 - *anonimato*;
 - ...
- Designamos essas propriedades genericamente por **propriedades de segurança**.

Cripto-análise

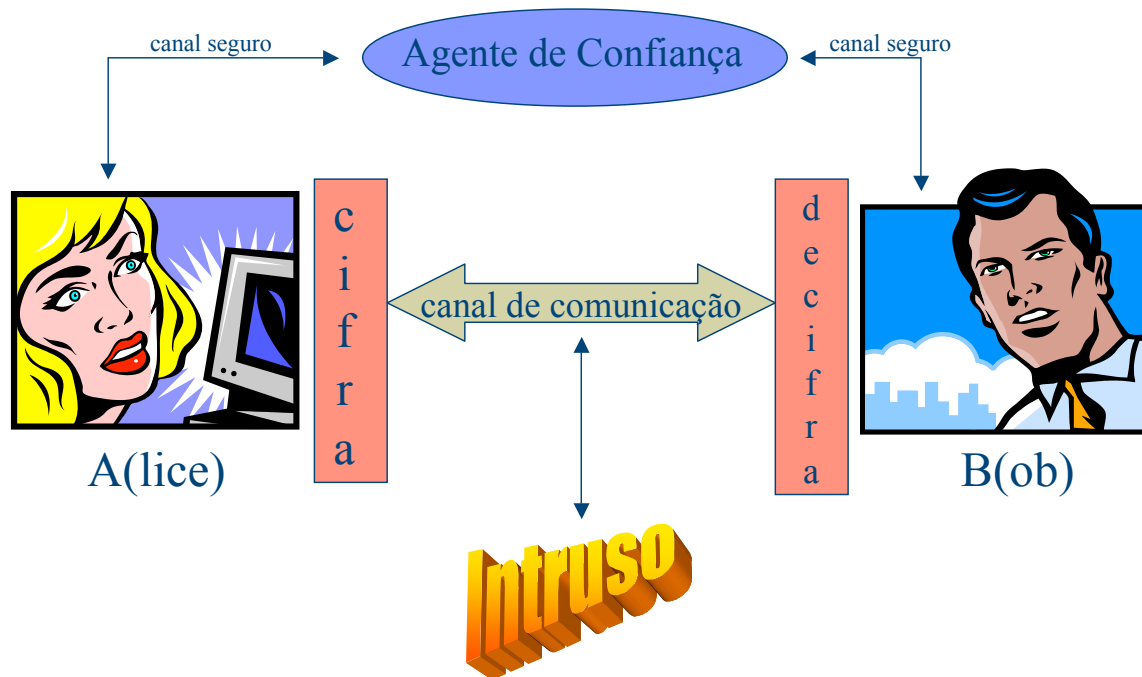
- Em oposição, a **Cripto-análise** tenta gorar os objectivos da Criptografia.
- Quando se tem sucesso ao comprometer (**atacar**) os objectivos da técnica criptográfica, dizemos que essa técnica foi **quebrada**.
- Conjuntamente, a Criptografia e a Cripto-análise formam uma área a que podemos chamar **Criptologia**.
- Como área científica, a *Criptologia* dispõe de profundos pontos de contacto com a Matemática e as Ciências da Computação.

- A Criptografia existe desde a antiguidade, normalmente associada a actividades militares e diplomáticas.
- A segurança dependia, em grande parte, do secretismo que rodeava as técnicas utilizadas (o que, historicamente, se revelou “catastrófico”).
- Esta tendência fez-se notar ainda no Século XX, durante as 1^a e 2^a Guerras Mundiais e prolongou-se durante as primeiras décadas da Guerra Fria.
- Só no princípio dos anos 70 surgiu como área de investigação académica de reconhecimento generalizado.
- Hoje é reconhecida a importância de eliminar o obscurantismo como factor na segurança dos sistemas criptográficos.

História da Criptografia Moderna

- **1948-1949** – *Claude Shannon* desenvolve a *Teoria da Informação* e enquadra o estudo da Criptografia nessa teoria.
- **1970-1977** – Desenvolvimento e standardização do Data Encryption Standard (DES).
- **1976** – Artigo científico de (*Diffie & Hellmann*) definindo os princípios da criptografia de chave pública.
- **1978** – Rivest, Shamir e Adleman descobrem a primeira cifra assimétrica: o RSA.
- **1985** – Descoberta da cifra assimétrica El Gamal.
- ...
- **1995** – Standardização do Digital Signature Algorithm.
- ...
- **2001** – Escolha do substituto do DES: Advanced Encryption Standard (AES).
- ...

Modelo de Comunicação

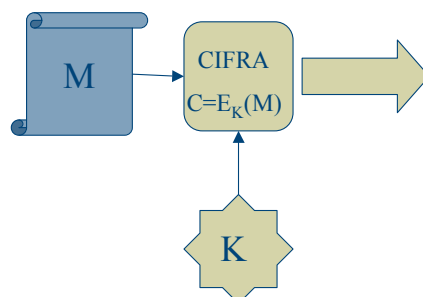


Princípio de Kerckhoff

Para avaliar a segurança de uma técnica criptográfica devemos assumir que esta é do conhecimento de eventuais inimigos.

Corolário: a segurança da cifra é assegurada por um parâmetro explícito — a **chave**.

$$C = \text{enc}_K(M)$$



- A natureza hostil do *ambiente* onde a técnica criptográfica será executada é **personificada** pelo **adversário** (também designado por *intruso*, *spy*, *Eve*, ...).
- Se o objectivo da técnica for comprometido, diz-se que esta foi objecto de um **ataque**.
- Distinguem-se dois tipos de ataques, dependendo das faculdades atribuídas ao adversário:
 - **Passivo**: o adversário dispõe unicamente da capacidade de escutar o canal de comunicação (i.e. de observar todo o tráfego que circula do canal).
 - **Activo**: atribui-se adicionalmente capacidade para manipular a informação que circula no canal de comunicação (alterar/bloquear/injectar mensagens).

Definição de Segurança

Uma técnica criptográfica diz-se segura se nenhum atacante puder ter sucesso a atacá-la.

- Na prática, interessa distinguir os adversários de acordo com a sua **capacidade computacional**. Dependendo das capacidades que lhe são atribuídas, obtém-se assim diferentes **noções de segurança**:
 - **Segurança Absoluta** – quando a segurança é estabelecida perante um adversário sem limitações computacionais.
 - **Segurança Computacional** – quando se considera que o adversário dispõe de limitações do *poder computacional* “realistas” (tempo de processamento, capacidade de memória, etc.).
- ?? como se poderá “demonstrar” a segurança ??

- **texto limpo:** mensagem a transmitir.
- **cifra:** operação que transforma o texto limpo numa mensagem “com significado obscurecido” — o **criptograma**.
- **chave:** parâmetro de segurança da operação de cifra
- **sistema criptográfico:** especificação das operações de “inicialização”; “cifra” e “decifragem”.
- **ataque:** comprometimento dos objectivos da técnica criptográfica (e.g. obtenção do texto limpo sem conhecimento da chave; descobrir a chave utilizada; etc.).
- **intruso/adversário/inimigo/spy:** entidade que personifica quem pretende comprometer os objectivos da técnica criptográfica.

Parte II

Algumas Cifras Clássicas

Cifra de César

- Cifra conhecida desde a antiguidade clássica.
- Diz-se que terá sido utilizada por *Júlio César* na campanha da Gália.
- Operação de cifra consiste em realizar um “deslocamento” das letras do alfabeto.

Texto limpo:	A	B	C	D	E	F	...	T	U	V	X	W	Y	Z
Criptografia ($K = 6$):	G	H	I	J	K	L	...	Z	A	B	C	D	E	F

- Exemplo: cifrar a mensagem `CartagoEstaNoPapo` com chave $K = 6$ resulta em `IGWZGMUKYZGTUVGVU`.
- Número total de chaves possíveis é 26 (uma delas **fraca**).

Exemplo de ataque

- Considere-se o criptograma `FXLNTQCLOPNPDLC`.
- Baixo número de chaves permite uma busca exaustiva sobre todas as possíveis chaves:

Criptograma:	F	X	L	N	T	Q	C	L	O	P	N	P	D	L	C
$K^{-1} = +1$:	G	Y	M	O	U	R	D	M	P	Q	O	Q	E	M	D
...	...														
$K^{-1} = +15$:	U	M	A	C	I	F	R	A	D	E	C	E	S	A	R
...	...														

- A chave utilizada deve então ter sido $K = 26 - 15 = 11$.
- Análise de frequências permite ataques muito mais eficientes... (porquê?)
- E.g.: alta frequência da letra `L` sugere chave $K = L - A = 11$.

Ataque por *Força Bruta*

- O ataque apresentado designa-se por **ataque por força bruta**.
- Esse ataque caracteriza-se por o adversário percorrer todo o *espaço de chaves* na expectativa de encontrar o texto limpo original.
- Pressupõe que:
 - existe suficiente redundância no texto original;
 - ou que se conhece um par “texto limpo/criptograma”.
- No entanto, estes condicionalismos são habitualmente cumpridos pelas aplicações correntes de cifras.
- É assim normalmente tido como *um ataque que é sempre passível de ser aplicado a uma cifra*.
- ... mas cuja *viabilidade* se encontra condicionada pelo tempo que demora percorrer todo o espaço de chaves!!!
- Pode, portanto, ser ultrapassado adoptando **tamanhos razoáveis** para as chaves.

...sobre números grandes...

- O que é que se considera como **tamanho razoável**?
- Obs.: questão só faz sentido se considerarmos adversários limitados computacionalmente.
- Note que a dimensão do *espaço de chaves* é exponencial em relação ao tamanho da chave. Um incremento de um bit no tamanho da chave duplica o espaço de chaves disponível.
- Exemplos de tempos de execução:

Tam. Chave	Tempo ($1\mu\text{sec}/\text{test}$)	Tempo ($1\mu\text{sec}/10^6\text{test}$)
32 bit	35.8 min.	2.15 msec.
40 bit	6.4 dias	550 msec
56 bit	1140 anos	10 horas
64 bit	500000 anos	107 dias
128 bit	$5 * 10^{24}$ anos	$5 * 10^{18}$ anos

- Actualmente, considera-se que 2^{80} fornece um nível de segurança aceitável.

Cifra por substituição mono-alfabética

- Considere-se uma cifra *por substituição mono-alfabética*.
- Generaliza a cifra de César permitindo “deslocamentos” diferentes para as diferentes letras do alfabeto.
- Exemplo...

A	B	C	W	Y	Z
R	X	K	B	I	F

- O número de possíveis chaves é então de $26! \approx 17.5 * 10^{24}$.

Devemos confiar na segurança desta cifra?

Ataque a uma cifra por substituição

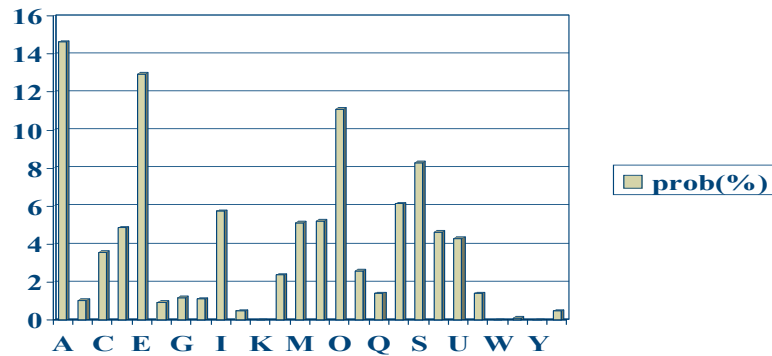
- Considere que numa cifra por substituição mono-alfabética se interceptou o seguinte criptograma:

FPGFBNBVPKFBMSBEMDMGUCDKDGUGDMUSFMMDBEFLEFEQDCPPGIDEXDCBKPMDFKPMFPQBUGPSUGHKEGPF
QBMFPXPKSESEBSURBHKBBHMEQBFUFSDSBGHKPPFCECPHDKQPDHDKQPFDBADVEDFDCDDCEZPKLDZEDGMPPM
NDKPMGDGVPEMPDNUPFQDVMPCPZGEFUQBMCPUGMEOPFSEBHPFBMBFBDUNPCDPXSEQDSDBCBBUQBFBFCPMU
KNEKDGUGDMHPKMBFDNPFMCPBKENPGBIMSUKDSBGJUPGPFQKPQEVPSBFSEOEDIUOBMPGKPMQDUKDFQPMXP
SPFQKESBMPMPMQKDZEDGUGDHPKNUFQDQPKKEVPOBJUPPJUPVDEMUSPCPKKPMHBFCEOAPMJUPFDBMDIED
PPOPMBOADKDGHDKDBHKDQBSBGEFJUEPQDSDB

- Sabe-se que a mensagem transmitida é um texto em Português.

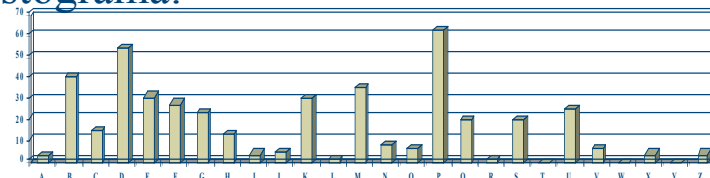
Como proceder à cripto-análise desta cifra?

- Explora-se o facto de, em Português, existirem diferentes probabilidades de ocorrência de letras nos textos.



- Por outro lado, também são distintas as probabilidades com que estas se agrupam (e.g. “as”, “os”, “es”, “que”, “nao”, ...)
- Assim, realizando a análise de frequências no criptograma obtemos:

Histograma:



- Podemos prosseguir por “palpites”: as letras *P*, *D* e *B* deverão corresponder ao *A*, *E* e *O*. Por outro lado, a existência de várias ocorrências dos pares *PM*, *PF*, *MP*, *JUP*, ... sugerem-nos a seguinte *decifragem parcial*:

ME-MO-O-E-MOAS-O-SAS-U-A-A-U-ASU-ESSAO-M--M--A-EE--A--A-O-ESA--ESEM-OU-E-U--
---EM-OSE-E-----O-U-O--O-OS--OMUM-A-O---EEM---E-A--EA-A--EMAO-A--AMA-AA---E-
-A--A-SEES-A-ESA-A-E-SEA-UEM-A-ASE-E---MU-OS-EU-S--EM--O-EMOSOMOAU-E-AE----
A-AO-OUU-OMO-ESU----A-U-AS-E-SOMA-EMS-EO---E-O-S-U-A-O-QUE-EM--E---E-OM----
A-U-OSE--ES-AU-AM-ESE--EM----OSESES--A--A-U-A-E--UM-A-E----E-OQUEEQUE-A-SU-
E-E--ES-OM----ESQUEMAOSA--AEF-ESO--A-A--A-AO--A-O-O--MOU-E-A-AO

- ...que não parece fazer muito sentido!!! :-)

- Fazendo *backtracking* e tentando outra alternativa, obtinhamos:

```
NE-NO-O-ERNOAS-O-SAS-U-ARA-U-ASU-ESSAO-N--N-TA-EE--A--A-ORES-RESENTOU-E-U--R--
ENTOSE-ER----O-U-O-RO-OS-TONUN-A-O--REEN---E-ARTEA-ARTENAO-A--ANA-AA---ER-A--A-
SEES-ARES-A-E-SEA-UENTA-ASE-E---NUTOS-EU-S--EN--O-ENOSONOAU-E-AE---TA-AO-
OOUTONO-ESUR--RA-U-AS-ERSONA-ENS-EOR--E-O-S-URA-O-QUE-ENTRET--E-ON----A-U-OSE-
RESTAURANTESE--ENTR--OSESSESTRA--A-U-A-ER-UNTATERR--E-OQUEEQUE-A-SU-E-ERRES-ON---
-ESQUENAOSA--AEE-ESO--ARA--ARAO-RATO-O--NQU-ETA-AO
```

- Que finalmente nos conduz a:

```
NEMNOGOVERNOASCOISASMUDARAMUMASUCESSAOINFINITADEEMBAIXADORESAPRESENTOUMECUMPRIMEN
TOSEXERCICIOCUJOPROPOSITONUNCACOMPREENDIDEPARTEAPARTENAOHAVIANADAADIZERFAZIAMSEES
GARESAMAVEISEAGUENTAVASEDEZMINUTOSDEUMSILENCIOPENOSONOAUGEDAEXCITACAODOOUTONODESU
RGIRAMUMASPERSONAGENSDEORIGEMOBSCURACOMQUEMENTRETIVECONCILIABULOSEMRESTAURANTESEX
CENTRICOSSESSESTRAZIAMUMAPERGUNTATERRIVELOQUEEQUEVAISUCEDERRESPONDILHESQUENAOSABIA
EELESOLHARAMPARAOPRATOCOMINQUIETACAO
```

Cifra de Vigenère (substituição poli-alfabética)

- Inventada por Blaise Vigenère (finais sec. XVI). Conhecida como *“le chiffre indéchiffrable”*.
- Intercala múltiplas cifras de César.
- Quebrada no final do sec. XIX por *Charles Babbage* e *Friedrich Kasiski*.
- Descrição da cifra:
 - Chave é uma “frase” em que cada letra determina uma substituição
 - Cada substituição é um simples deslocamento determinado pelo carácter respectivo da chave ((e.g. A = 0, B = 1, ...); c.f. cifra de César).
 - Tamanho da chave determina número de substituições utilizadas.

Exemplo de utilização

- Pretende-se cifrar a mensagem $M = \text{CifraIndecifavel}$ com a chave $K = \text{BACO}$.
- Tabela de Vigenère:

	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
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
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
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

- Operação de cifra:

Chave:	B	A	C	O	B	A	C	O	B	A	C	O	B	A	C	O	B
Texto limpo	C	I	F	R	A	I	N	D	E	C	I	F	R	A	V	E	L
Criptograma	D	I	H	F	B	I	P	R	F	C	K	J	S	A	X	S	M

- Observações:
 - A mesma letra não é sempre cifrada da mesma forma;
 - mas se o texto limpo for muito maior do que a chave, os padrões no texto limpo vão-se repercutir no criptograma.
 - As técnicas de cripto-análise desenvolvidas dispõe já de um nível de sofisticação considerável...

Cifras por Transposição

- O mecanismo de cifra pode afectar as posições dos caracteres, em vez de manipular esses mesmos caracteres.

- E.g. considere-se a permutação:

1	2	3
2	1	3

- Operação de cifra:

- Para cifrar: AINDAOUTRACIFRA

- organiza-se a mensagem numa matriz

2	1	3
A	I	N
D	A	O
U	T	R
A	C	I
F	R	A

- lê-se a matriz seguindo a ordem da chave...
 - ...resultando em: IANADOTURCAIRFA
- ? Como Quebrar esta cifra ?

Combinação de Cifras

- Tendo visto diferentes *cifras simples*...
- ...fará sentido construir uma *cifra complicada* combinando várias dessas cifras mais simples?
- Em particular, será que é legítimo dizer que a *segurança* dessa nova cifra é acrescida?
- **Depende!!!**
 - pode acontecer de não trazer valor acrescentado nenhum (e.g. combinando duas cifras por substituição)
 - mas há padrões de combinação que podem ser vantajosos (e.g. intercalando substituições com permutações)

Cifra *One-Time-Pad*

- Generaliza a cifra de Vigenère com:
 - comprimento da chave é o mesmo da mensagem a cifrar;
 - a chave é completamente aleatória.
- Cifra demonstrada *incondicionalmente segura* por *Claude Shannon* (1949).
- É normalmente descrita operando sobre um alfabeto binário:
 - Operações de cifra/decifragem são simplesmente o *xor* com a chave.

$$C_i = T_i \oplus K_i \quad M_i = C_i \oplus K_i$$

- Chaves só podem ser utilizadas numa única operação de cifra.
- Segurança da cifra resulta do facto de o conhecimento do criptograma não resultar na **diminuição de incerteza** relativa ao conhecimento do texto limpo.
- Os problemas inerentes à geração e distribuição da chave tornam a cifra **inviável**.

Em Resumo...

- Na cripto-análise, explora-se *toda a informação* disponível, como sejam:
 - a natureza na mensagem transmitida;
 - informação parcial dessa mensagem;
 - histórico sobre a utilização da cifra (e.g. existência de mensagens cifradas com a mesma cifra/chave);
 - possíveis *vícios de utilização* da cifra (e.g. deficiências na escolha das chaves, etc.).
- Mesmo se existem técnicas **incondicionalmente seguras**, elas impõe normalmente requisitos de tal maneira fortes à sua utilização que se tornam impraticáveis.
- Por isso, a generalidade das técnicas criptográficas utilizadas actualmente baseiam-se numa noção de segurança onde se limitam as capacidades computacionais do adversário: **segurança computacional**.

- As técnicas de cripto-análise são tanto mais efectivas quanto se consigam afastar significativamente de um ataque por força bruta.
- O **tamanho de chave recomendado** para uma determinada cifra deve ser definido com base no **nível de segurança** pretendido (e.g. 2^{80}) e de qual a percentagem do tamanho de chave **é consumida** pelas técnicas de cripto-análise conhecidas.

Parte III

Segurança da Informação

Propriedades de Segurança

A criptografia é hoje utilizada para fornecer garantias referentes a um vasto leque de *propriedades de segurança*:

- **Confidencialidade:** garantir que o conteúdo da mensagem só é do conhecimento dos intervenientes legítimos.
- **Integridade:** garantir que o receptor não “aceita” mensagens que tenham sido manipuladas.
- **Autenticidade:** assegurar a “origem” da mensagem.
- **Não repúdio:** demonstrar a “origem” da mensagem.
- **Anonimato:** não fornecer qualquer informação sobre a origem da mensagem.
- **Indentificação:** assegurar a “identidade” do interveniente na comunicação.
- ...

- Estamos normalmente interessado numa combinação de propriedades (e.g. num *canal seguro* entre duas partes pretende-se garantir a confidencialidade, autenticidade e integridade).
- Por outro lado, algumas das propriedades referidas não resultam directamente de uma técnica criptográfica específica, mas antes de uma conjugação de técnicas.
- Esta combinação de técnicas resultam tipicamente no que se designa por **protocolos criptográficos** — aí especificam-se as trocas de mensagens (e as técnicas criptográficas utilizadas) para se atingirem os fins pretendidos.
- A segurança de protocolos criptográficos (i.e. se eles realmente cumprem os requisitos para que foram desenvolvidos) não depende unicamente da segurança das técnicas que os suportam.

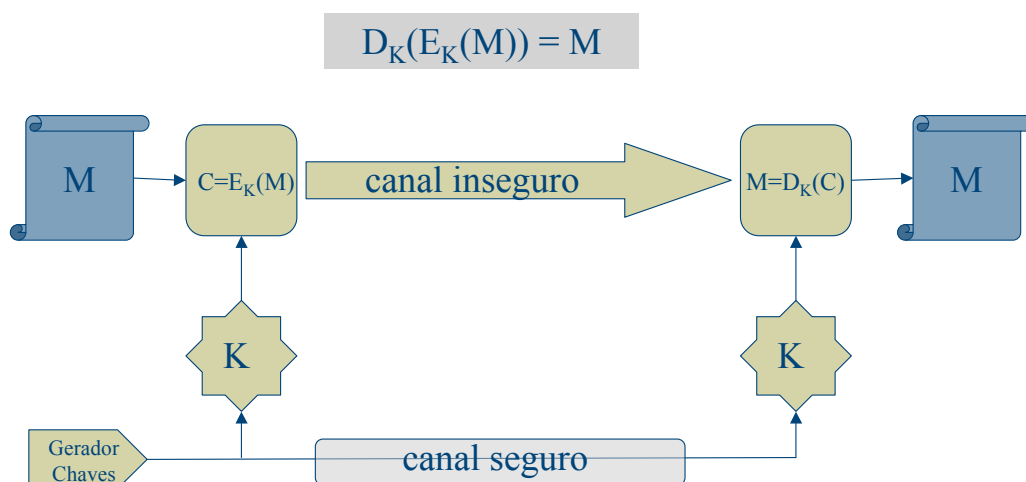
Criptografia e Segurança

A segurança das técnicas criptográficas constituem um ingrediente fundamental e necessário na segurança de sistemas informáticos, mas não suficiente.

- Podemos distinguir (pelo menos) os seguintes níveis no estabelecimento da segurança de um sistema informático:
 - Técnica criptográfica;
 - Protocolo;
 - Implementação;
 - Utilização.
- Uma brecha de segurança em qualquer um destes níveis compromete a segurança de todo o sistema.

- Cifras.
- Assinaturas Digitais.
- Funções de *hash*
- *Message Authentication Codes*.
- ...

Cifras Simétricas



- A mesma chave é utilizada na operação de *cifra/decifragem*.
- Pressupõe por isso que, numa fase prévia à comunicação, se procedeu ao **acordo de chaves**.
- ...operação que “tipicamente” envolve a utilização de **canais seguros**.
- Exemplos: RC4; DES; IDEA; AES.

Geração (e manipulação) de chaves

- Um factor determinante para a segurança das técnicas criptográficas é a *qualidade das chaves* utilizadas.
- A sua principal característica é a *aleatoriedade* e *imprevisibilidade*.
- O tamanho (número de *bits*) depende da técnica concreta. Para cifras simétricas são comuns chaves de 40 a 256 bits.
- Na perspectiva da programação, é apropriado olhar para as chaves como *tipos opacos*.
- ...devendo ser prestada “toda a atenção” à sua manipulação (armazenamento; âmbito de vida na execução do programa; etc.)
- Os requisitos impostos impedem a utilização de *PINs* ou de *palavras/frases* *passé* directamente como chave de uma cifra moderna.
- ...mas existem métodos para derivar chaves a partir dessas *passwords* (*key derivation functions*).

Distribuição/Acordo de Chaves

- A pré-distribuição das chaves constitui a maior dificuldade na utilização das cifras simétricas, já que o estabelecimento de *canais seguros* é oneroso.
- Note que pode ser vista alguma *circularidade* neste domínio: *a criptografia pode ser utilizada para construir canais seguros, mas ela própria depende da existência de canais seguros*.
- Os *Protocolos de Distribuição de Chaves* fazem uso de uma “rede mínima de confiança” para distribuir as chaves entre os intervenientes.
- A criptografia assimétrica abriu novas perspectivas: o *Acordo de Chaves* — em vez de se gerar e distribuir a chave, define-se uma forma de cada interveniente “derivar” uma chave apropriada (sem que se ninguém mais possa ser capaz de derivar essa chave...).

- Faz uso de chaves de *cifra/decifragem* “distintas” — o *conhecimento de uma não revela informação sobre a outra*.
- Só a chave para decifrar necessita ser *secreta* (chave privada).
- Permite torneiar problema da distribuição de chaves — a chave para cifrar pode ser comunicada sem requisitos de confidencialidade.
- ...mas permanecem algumas dificuldades (autenticidade da origem...).
- Exemplos: RSA; El-Gamal.

Funções de Hash

- As funções de *Hash* criptográficas (ou *message digest*; *fingerprint*; etc.) permitem produzir um “resumo” de tamanho fixo a partir de uma mensagem de comprimento arbitrário.
- ...de tal forma que “não é viável” encontrar outra mensagem que disponha do mesmo resumo (*pre-image resistant*).
- Tratam-se, por isso, de **funções one-way** (não invertíveis).
- Exemplos: MD5; SHA-1.
- Trata-se de um exemplo de uma técnica fundamental em criptografia que, por si só, não dá resposta directa a nenhuma propriedade de segurança — *o seu poder resulta da combinação com outras técnicas*.

Message Authentication Codes (MACs)

- Um *MAC* pode ser entendido como uma função de hash cujo resultado depende, para além da mensagem, de um segredo (chave secreta).
- Garante assim a *integridade* de uma mensagem:
 - Quem envia a mensagem gera o MAC respectivo que envia junto;
 - O Receptor por sua vez gera o MAC da mensagem recebida e compara-o com o recebido.
 - Se alguém alterar a mensagem não poderá recalcular o MAC (não dispõe da chave).
 - Assim o receptor não aceitará a mensagem manipulada (porque não verifica o MAC).
- Exemplos: HMAC-MD5; HMAC-SHA1.

Assinaturas Digitais

- As **assinaturas digitais** permitem associar uma mensagem a uma determinada “entidade”.
- Cumprem assim um papel análogo ao das *assinaturas correntes* que associam documentos a pessoas.
- Ao nível das propriedades de segurança, estamos interessados em garantir:
 - **autenticidade**: o destinatário deverá confiar na identidade do signatário.
 - **integridade**: que o documento objecto da assinatura não é manipulado.
 - **não repudiável**: o signatário não poderá negar, posteriormente, que realmente assinou o documento.
- Exemplos: RSA; DSA.
- Pode ser entendido como o “contributo mais significativo” da criptografia assimétrica.

- Autentico **canivete Suíço** para quem trabalha em criptografia.
- Originalmente concebido como uma biblioteca que implementa o protocolo *SSL*.
- ...mas disponibiliza uma *shell* que dá acesso à funcionalidade implementada (cifras, funções de hash, assinaturas, ...)
- Disponível para a generalidade das plataformas (Unix, MacOS, Windows).
- Apontadores:
 - <http://www.openssl.org>
 - <http://www.modssl.org>