Universidade da Beira Interior
Faculdade de Engenharia
Departamento de Informática

Segurança Informática
Guia para Aula Laboratorial 4
Licenciatura em Engenharia Informática
Licenciatura em Informática Web
Licenciatura em Tecnologias e Sistemas da Informação

Computer Security
Guide for Laboratory Class 4
Degree in Computer Science and Engineering
Degree in Web Informatics
Degree in Information Technologies and Systems

Sumário
Discussão dos conceitos da confusão e difusão de Shannon através de uma experiência simples em que se calcula a entropia de mensagens em texto-limpo e cifradas. Teste de algumas propriedades dos modos de cifras por blocos ECB e CBC através de experiências simples.

Summary
Discussion of the diffusion and confusion concepts given by Shannon via a simple experiment in which the entropy of several plaintexts and cryptograms are calculated. Testing of several properties of the ECB and CBC block cipher modes via simple experiments.

Pré-requisitos:
Algumas das tarefas propostas a seguir requerem o uso de *software* para efetuar cálculos, o acesso a um sistema com compilador de programas escritos em linguagem de programação C e que disponibilize a ferramenta OpenSSL. Sugere-se, assim, o uso de uma distribuição comum de Linux, onde todas estas condições estarão provavelmente preenchidas.

1 Cifragem e Entropia
Encryption and Entropy

Na teoria da informação, a Entropia pode ser entendida como uma medida do **grau de incerteza** associado a determinada variável. Quanto maior for o valor da entropia, maior esse grau de incerteza. Formalmente, a entropia denota-se por $H(X)$ e define-se através da fórmula

$$H(X) = -\sum_{i=1}^{n} p(x_i)\log_b(p(x_i)),$$

em que X simboliza um evento aleatório com n possíveis ocorrências $\{x_1, x_2, ..., x_n\}$, $p(x_i)$ denota a probabilidade de ocorrer x_i e b é a base utilizada (pode usar sempre a base 10 adiante). A Entropia máxima atinge-se quando todas as possíveis ocorrências têm a mesma probabilidade de acontecer (i.e., $p(x_i) = 1/n$ para qualquer x_i).

Q1.: Sabendo isto, consegue deduzir a fórmula que dá o valor máximo da Entropia?
Então não haveria de conseguir? Dá o seguinte:
☐ $u = \sum_{i=1}^{n} p(x_i)\log_b(p(x_i))$
☑ $u = \log_b(n)$
☐ $u = 1/n \sum_{i=1}^{n} \log_b(p(x_i))$
☐ $u = 1/n$

Tarefa 1 *Task 1*

Elabore um programa, numa linguagem à sua escolha, que calcule a entropia dos caracteres (bytes) de um ficheiro.[1] O programa deve devolver os valores da entropia do ficheiro e da entropia máxima.

Deve implementar o programa de forma a que, depois de compilado, possa ser invocado da seguinte forma:

```
$ ./a.out nome_do_ficheiro
```

Q2.: Quantos bytes diferentes há? Por outras palavras, qual é o valor de n para este caso?
☐ 2 ☐ 4 ☐ 127 ☐ 128 ☐ 255 ☑ 256

Q3.: Com base na resposta dada antes, qual o valor máximo para a entropia calculada para os bytes de um ficheiro?
☑ 10 ☑ 5.55 ☑ 2.41 ☑ 0.00390625

[1] Caso ache que não vai conseguir implementar o programa de forma eficaz durante a aula, use o código fornecido com este guia laboratorial (não deixe de tentar implementar em casa).

Tarefa 2 *Task 2*

Crie um ficheiro de texto chamado `texto.txt` e encha-o com o conteúdo do artigo em http://mashable.com/2012/06/08/linkedin-stolen-passwords-list/. Calcule a entropia deste ficheiro e coloque o resultado, bem como o tamanho do ficheiro, na linha correspondente da tabela incluída em baixo.

Nome	Tamanho (em KB)	Entropia
`texto.txt`	1,4	3.202877
`comprimido.zip`	0,938	5.210323
`cifrado.aes`	1,4	5.419597
`comprimido-e-cifrado.zip.aes`	0,944	5.384356
`cifrado-e-comprimido.aes.zip`		

Entropia MÁX: 5.545177

Tarefa 3 *Task 3*

Comprima o ficheiro `texto.txt`, e dê o nome de `comprimido.zip` ao ficheiro obtido. Coloque o tamanho e a entropia deste ficheiro na linha correspondente da tabela acima mencionada. Considere usar sempre a linha de comandos para executar todas as tarefas:

```
$ zip comprimido.zip texto.txt
```

Tarefa 4 *Task 4*

Cifre o ficheiro `texto.txt` com a cifra *Advanced Encryption Standard* (AES) e dê o nome de `cifrado.aes` ao ficheiro resultante. Considere utilizar o comando seguinte na execução desta tarefa:

```
$ openssl enc -aes128 -K
0123456789abcdef0123456789abcdef -in
texto.txt -out cifrado.aes -iv 0
```

Q4.: Qual é o modo de cifra definido pela opção `-aes128`?

☐ *Electronic Code Book;*
☑ *Cipher Block Chaining;*
☐ *Output Feedback Mode.*
☐ Ahh? Nem nunca ouvi falar em tal coisa!

Q5.: Qual é o tamanho da chave de cifra que usou no comando anterior?

☐ 16 bits ☐ 32 bytes ☐ 64 octetos ☑ 128 bits

De modo análogo ao que fez antes, coloque o tamanho e a entropia deste ficheiro na respetiva linha da tabela. **Pare por aqui, e discuta os resultados na aula.**

Tarefa 5 *Task 5*

Considere executar os dois passos seguintes com calma, para evitar erros:

1. Cifre o ficheiro `comprimido.zip` com o AES e dê o nome de `comprimido-e-cifrado.zip.aes` ao ficheiro resultante;
2. Comprima o ficheiro `cifrado.aes` e dê o nome de `cifrado-e-comprimido.aes.zip` ao ficheiro que resulta deste procedimento. Preencha as linhas correspondentes na tabela que está a popular.

Parta da análise da tabela para responder à questão seguinte. **Q6.: Se tivesse de cifrar e comprimir um ficheiro, o que é que faria primeiro?** Justifique a sua resposta.

☐ QUE CONFUSÃO!
☐ Primeiro cifrar e depois comprimir.
☑ Primeiro comprimir e depois cifrar.
☐ É impossível cifrar e comprimir um ficheiro.

›

---

2 Modos de Cifra: ECB vs. CBC

Cipher Modes: ECB vs. CBC

O conjunto de tarefas seguinte foi desenhada de

forma a motivar o estudo das propriedades dos modos de cifra ECB e CBC. Considere, ainda assim, analisar previamente a aula teórica onde este assunto é tratado.

Tarefa 6 *Task 6*

Crie o ficheiro `texto-limpo.txt` com a sequência `111111111111222222222222` (doze 1s e doze 2s) repetida 3 vezes:

```
$ echo -n "111111111111222222222222" > confidential.txt
$ echo -n "111111111111222222222222" >> confidential.txt
$ echo -n "111111111111222222222222" >> confidential.txt
```

Tarefa 7 *Task 7*

Cifre o ficheiro `texto-limpo.txt` com a cifra AES no modo *Electronic Code Book* (ECB) (i.e. `-aes-128-ecb`) com a chave de cifra `11232233`. O ficheiro resultante deve chamar-se `criptograma.ecb`. No final, tente visualizar o conteúdo este ficheiro, e.g., com `$ hexdump -C criptograma.ecb`.

Q7.: O que significa o acrónimo AES?

A advanced E encryption S standard

Q8.: Qual é o tipo da cifra AES?

☐ Cifra de chave simétrica por blocos contínua.
☐ Cifra de chave simétrica contínua.
☑ Cifra de chave simétrica por blocos.
☐ Cifra de chave assimétrica.
☐ Cifra de chave assimétrica por blocos.
☐ Cifra de chave assimétrica contínua.

Q9.: Consegue detetar padrões no ficheiro cifrado?

☐ Sim.
☐ Não. Nunca fui bom criptanalista... ou lá como isso se chama!
Justifique a sua resposta.

Nas linhas 1 e 4 nota-se um padrão

Q10.: De acordo com a experiência que fez antes, qual é o tamanho do bloco que esta cifra utiliza?

☐ 8 bytes. ☑ 16 bytes. ☐ 32 bytes. ☐ 64 bytes.

Q11.: O modo ECB precisa de vetor de inicialização? ☐ Sim, precisa. ☑ Não, não precisa.

Tarefa 8 *Task 8*

Use o seguinte comando para cifrar o ficheiro `texto-limpo.txt` novamente:

```
$ openssl enc -aes-128-cbc -K 11232233 -in texto-limpo.txt -out criptograma.cbc
```

Q12.: Não me diga que o comando enunciado antes não funcionou?

☑ Digo, digo. ☐ Não, não digo.

Q13.: Como se resolve a situação?

Metendo lá o vetor de inicialização

Nota: quando souber como resolver o problema anterior, emita mesmo o comando, porque vai precisar dele a seguir...

Q14.: O que significa CBC?

C ypher B lock C haining

Q15.: O que é o CBC?

☐ O nome de um algoritmo de cifra.
☑ Um modo de utilização de uma cifra.
☑ O nome de uma estação de televisão norte americana.

Tarefa 9 *Task 9*

Verifique o conteúdo do `criptograma.cbc`. **Q16.: Consegue detetar algum padrão no texto cifrado?**

☑ Que interessante, não vejo qualquer padrão.
☐ Sim, tem os mesmos patrões que o ficheiro original.

Q17.: O que é que aconteceria se cifrasse outra vez o mesmo ficheiro com a mesma chave no modo CBC, mas com diferente vetor de inicialização?

☐ Os dois criptogramas seriam totalmente iguais.
☐ Os dois criptogramas seriam parecidos.
☑ Os dois criptogramas seriam totalmente distintos.