



Mensagens Codificadas: Sequencias Equidistantes de Símbolos

O	I	N	I	C	I	O	D	O	S	A	N	O	S	9	0	T	R	E	S	P	E	S	Q	U	I	S	A	D	O	R	E	S	L	A	N	C	A	R	A	M	U	M	A	R	T	I	G	O
C	A	D	E	M	I	C	O	Q	U	E	P	R	O	M	E	T	I	A	R	E	V	O	L	U	C	I	O	N	A	R	T	O	D	O	C	A	M	P	O	C	I	E	N	T	I	F	I	C
U	M	A	R	T	I	G	O	Q	U	E	P	R	O	V	A	R	I	A	A	E	X	I	S	T	E	N	C	I	A	D	E	D	E	U	S	A	E	V	I	D	E	N	C	I	A	E	S	T
R	I	A	O	C	U	L	T	A	P	O	R	M	E	I	O	D	E	U	M	C	O	D	I	G	O	S	E	C	R	E	T	O	C	O	N	T	I	D	O	N	A	B	I	B	L	I	A	O
A	S	O	N	A	O	G	A	N	H	O	U	D	E	I	M	E	D	I	A	T	O	A	P	R	E	O	C	U	P	A	C	A	O	D	A	M	I	D	I	A	A	P	O	S	O	L	I	V
O	C	O	D	I	G	O	D	A	B	I	B	L	I	A	M	I	C	H	A	E	L	D	R	O	S	N	I	N	S	E	R	P	U	B	L	I	C	A	D	O	E	M	9	7	C	E	N	T
N	A	S	D	E	P	E	S	S	O	A	S	S	E	I	M	P	R	E	S	S	I	O	N	A	R	A	M	C	O	M	A	R	E	A	L	P	O	S	S	I	B	I	L	I	D	A	D	E
E	A	C	H	A	R	U	M	A	P	R	O	V	A	D	E	Q	U	E	A	S	E	S	C	R	I	T	U	R	A	S	J	U	D	A	I	C	A	S	S	A	O	D	E	F	A	T	O	D
V	I	N	A	S	A	S	L	E	F	O	I	O	M	E	T	O	D	O	P	A	R	A	D	E	S	V	E	N	D	A	R	O	S	C	O	D	I	G	O	S	S	E	C	R	E	T	O	S
R	A	T	A	S	E	D	E	A	J	E	I	T	A	R	T	O	D	O	T	E	X	T	O	E	M	L	I	N	H	A	S	E	Q	U	I	D	I	S	T	A	N	T	E	S	S	E	M	
S	P	A	C	O	S	E	N	T	R	E	A	S	P	A	L	A	V	R	A	S	I	G	N	O	R	A	N	D	O	A	S	V	I	R	G	U	L	A	S	E	O	S	P	O	N	T	O	S

Suponha que uma organização criminosa utiliza a internet para trocar mensagens entre os seus elementos. No entanto, para evitar que essas mensagens sejam percebidas pelos elementos das forças de segurança, essa organização utiliza um método básico de codificação, designado por “Sequencias Equidistantes de Símbolos” (SES).

Nesta codificação, além da cadeia de símbolos, é também transmitido um valor inteiro positivo, que especifica o espaçamento entre os caracteres que realmente interessam. Por exemplo, na mensagem acima reproduzida e para o espaçamento “48”, seria possível extrair a palavra assinalada a verde (“Dados”). Se o valor de espaçamento fosse “47”, poderiam ser extraídas as palavras assinaladas a laranja.

Neste trabalho, terá acesso a uma base de dados de mensagens da referida organização criminosa (“msgs.bin”), guardada num ficheiro binário com inteiros de 8 bits correspondentes ao código ASCII dos caracteres. Cada mensagem é terminada pelo código numérico “10” (*new line*). O ID das mensagens corresponderá à ordem pela qual elas aparecem nesta base de dados.

Adicionalmente, existirá um ficheiro de texto (“keywords.txt”) onde são especificadas as palavras-chave a procurar na base de dados de mensagens.

O objectivo deste trabalho é o de planejar e implementar um sistema que leia os ficheiros de mensagens e de palavras-chave, e teste todos os espaçamentos possíveis, com vista a encontrar as referidas palavras-chave em algumas das mensagens existentes.

O output do sistema consistirá no ID das mensagens que potencialmente contenham alguma(s) das palavras chave a procurar, ordenadas pela quantidade de palavras-chave encontradas.

Exemplo Output:

Msg 7, “Covilhã” (48), “Ataque” (51) // Indica que a mensagem “7” refere a “Covilhã” se lida com espaçamento 48 e “Ataque”, com espaçamento 51

Elementos para Entrega:

Aquando da entrega do trabalho, devem ser incluídos os seguintes elementos:

1. Código-fonte de todas as funções ou bibliotecas não “ANSI-C” que tenham sido implementadas para a resolução do trabalho.
2. Relatório a descrever de forma esquemática e textual, as opções mais significativas, relativamente à implementação do trabalho.