



Universidade Federal do Rio de Janeiro
Instituto de Matemática
Departamento de Ciência da Computação
Grupo de Resposta a Incidentes de Segurança

Rio de Janeiro, RJ – Brasil

Fundamentos da Criptologia

Parte I – Introdução e Histórias

GRIS-2005-A-003

Breno Guimarães de Oliveira

A versão mais recente deste documento pode ser obtida na página oficial do GRIS

GRIS – Grupo de Resposta a Incidentes de Segurança
CCMN Bloco I 1º andar
Sala: I1021
Av. Brigadeiro Trompowski, s/nº
Cidade Universitária - Rio de Janeiro/RJ
CEP: 21949-900
Telefone: +55 (21) 2598-3309

Este documento é Copyright© 2005 GRIS. Ele pode ser livremente copiado desde que sejam respeitadas as seguintes condições:

É permitido fazer e distribuir cópias inalteradas deste documento, completo ou em partes, contanto que esta nota de copyright e distribuição seja mantida em todas as cópias, e que a distribuição não tenha fins comerciais. Se este documento for distribuído apenas em partes, instruções de como obtê-lo por completo devem ser incluídas. É vedada a distribuição de versões modificadas deste documento, bem como a comercialização de cópias, sem a permissão expressa do GRIS.

Embora todos os cuidados tenham sido tomados na preparação deste documento, o GRIS não garante a correção absoluta das informações nele contidas, nem se responsabiliza por eventuais consequências que possam advir do seu uso.

Sumário:

1. Introdução e Nomenclatura.....	3
2. Um Pouco de História – A Evolução da Criptologia.....	3
3. Curiosidades.....	8
3.1. A Vida Secreta de Alice & Bob.....	8
3.2. As Cifras de Beale.....	9
4. Bibliografia e Crédito das Imagens.....	10

1. Introdução e Nomenclatura

Não, não é um texto sobre criptas e cemitérios, embora muita gente tenha morrido por causa dela. A criptologia é a área de conhecimento em que estão reunidos os estudos da criptografia e da criptoanálise. Antes de mais nada é preciso ter certeza de que estamos falando das mesmas coisas, então seguem algumas definições:

Criptografia é o conjunto de princípios e técnicas empregadas para cifrar a escrita, de modo que apenas os que têm acesso às convenções combinadas possam lê-la. A mensagem em linguagem cifrada é chamada de *criptograma*. Por motivos óbvios, não se trata de um embaralhamento aleatório em que os passos não são conhecidos. Pelo contrário, a codificação é sempre feita de modo a permitir uma posterior decodificação, senão todo o material ocultado estaria perdido e o trabalho teria sido em vão. Na prática, a mensagem é modificada por uma *função* de codificação com o auxílio de um valor especial ou *chave*, e transformada assim no criptograma. Para que o receptor possa ler a mensagem, ele modifica o criptograma com uma *chave* e uma função de decodificação, obtendo novamente o texto original.

Criptoanálise, por sua vez, é o conjunto de técnicas e métodos para decifrar uma escrita de sistema desconhecido. Note que o termo ‘decifrar’ está sendo usado aqui com o significado de descobrir a mensagem original de um criptograma sem possuir a chave de decodificação, isto é, sem ser um usuário legítimo. Uma tentativa de criptoanálise é comumente chamada de *ataque*.

Nessa série de textos veremos diversos métodos de criptografia existentes, desde os mais elementares até os que ainda são utilizados nos dias de hoje. Será apresentada ainda a criptoanálise dos métodos descritos, demonstrando pelo menos um possível ataque para cada um deles, suas vantagens e desvantagens.

Algum conhecimento da Teoria dos Números faz-se necessário para o entendimento dos métodos abordados e sua justificativa.

2. Um Pouco de História – A Evolução da Criptologia

Durante muitos séculos, a criptografia foi tratada como uma arte. Seu nome vem das palavras gregas *kruptós* (oculto, secreto, obscuro, ininteligível) e *graphía* (com o sentido de ‘escrita’) e, enquanto gregos deliciavam-se com o embaralhamento de palavras, os romanos já utilizavam tais conhecimentos para guerras e segredos de estado. Naturalmente, isso lhes deu grande vantagem, pois, ainda que um mensageiro fosse capturado, a mensagem permaneceria protegida. Mas a “arte de cifrar mensagens” tem origens muito mais antigas.

De fato, as cifras existem desde que o ser humano aprendeu a riscar uma pedra na parede da caverna. Isso porque qualquer escrita desconhecida pode ser considerada uma cifra. Nesse caso, não há uma chave de codificação particular, e cada objeto, idéia, etc., possui sua própria representação. A Pedra de Roseta é um exemplo claro em que a criptoanálise foi essencial para o estudo da cultura de povos antigos. Descoberta por tropas napoleônicas na



Trecho da Pedra de Roseta

cidade de Roseta, no Egito, em julho de 1799, e atualmente no Museu Britânico de Londres, a pedra de 196 a.C. contém o mesmo texto em três idiomas: grego, demótico (um tipo de escrita egípcia) e em hieróglifos. A compreensão do verdadeiro significado dos hieróglifos havia sido a muito perdida, e somente com essa “tabela de conversão” secular que os historiadores puderam finalmente decifrá-los.

Os métodos criptográficos desenvolvidos na antiguidade eram baseados essencialmente em técnicas de substituição e transposição simples, já que o uso de contas matemáticas complexas era pouco prático. A substituição troca letras e/ou conjuntos de letras por outras letras ou símbolos, seguindo regras específicas. Já os métodos de transposição consistem na reorganização das letras da mensagem, numa ordem conhecida apenas por remetente e destinatário.

Uma das primeiras descrições conhecidas de uma cifra de substituição está no Kama-sutra, baseada em manuscritos de aproximadamente 400 anos antes de Cristo. A 45ª das 64 artes que o Kama-sutra recomenda que as mulheres estudem é a *mlecchita-vikalpa*, a “arte da escrita secreta”, concebida para ajuda-las a salvaguardar detalhes de suas relações. Uma das técnicas envolve combinar aleatoriamente cada letra do alfabeto com outra qualquer, e substituir as letras da mensagem por seus respectivos pares.

As cifras de substituição também eram o método criptográfico preferido de Júlio César. Em seu livro intitulado “Guerras Gaulesas”, ele conta como enviava mensagens substituindo as letras romanas por letras gregas, tornando a mensagem inteligível aos inimigos. Valerius Probus chegou a escrever um tratado inteiro apenas sobre as diferentes cifras utilizadas por César, mas tal documento não sobreviveu aos dias de hoje. No entanto, o documento “Vida dos Césares LVI” de Caius Suetonius descreve o método mais tarde conhecido como a *Cifra de César*, que veremos na próxima seção.

As cifras de transposição, por sua vez, também foram largamente utilizadas no ramo militar, em especial pelos espartanos. Em torno do século V a.C., esse povo guerreiro cifrava e ocultava suas mensagens usando um bastonete denominado escútala e uma cinta, que era enrolada no mesmo. A



Simulação da codificação Espartana

mensagem original era gravada na cinta seguindo o comprimento da escútala, enquanto letras sem sentido eram colocadas nos espaços restantes da cinta, que era então desenrolada e enviada ao destinatário. Este precisava possuir uma escútala com o mesmo diâmetro, do contrário a mensagem não faria sentido para ele.

Outra cifra de transposição no mínimo curiosa foi encontrada nas paredes de vilas romanas em Pompéia (Itália) e Cirencester (Grã-Bretanha), e ficou conhecida como o “Quadrado Latino” ou “Quadrado Sator”. Trata-se da frase “*sator arepo tenet opera rotas*, que pode ser traduzida como “O plantador Arepo guia o arado com as mãos”, escrita num quadrado formando o palíndromo geométrico à direita. O quadrado pode ser lido em qualquer direção, formando sempre a frase citada. No entanto, o texto pode ter suas letras transpostas, revelando então a mensagem “*pater noster*” repetida duas vezes (com a letra ‘n’ sendo comum a ambas, formando uma cruz como nas palavras cruzadas de hoje em dia), e as letras ‘a’ e ‘o’, que restaram, também repetidas duas vezes. “*Pater noster*” é “pai nosso” em latim, e, junto com a cruz, formam uma clara alusão ao cristianismo, cujos fieis eram perseguidos e devorados por leões em eventos públicos realizados pelos romanos na época. As duas letras ‘a’ e duas letras ‘b’ restantes provavelmente representavam alfa e ômega – início e fim – também de significado cristão. E essa é apenas uma das

S	A	T	O	R
A	R	E	P	O
T	E	N	E	T
O	P	E	R	A
R	O	T	A	S

diversas combinações possíveis com o Quadrado Latino, que incluem frases como “Petro et reo patet rosa sarona” (“A rosa sarônica está aberta a (São) Pedro e ao réu”) ou mesmo “satan, ter oro te, reparato opes” (“Satanás, eu lhe peço três vezes, recupere minha riqueza”). Outro resultado

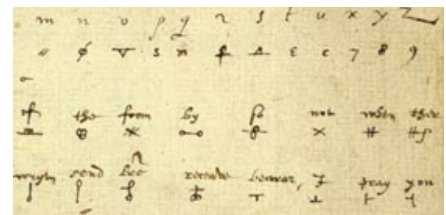


S	A	T	O	R
A	R	E	P	O
T	E	N	E	T
O	P	E	R	A
R	O	T	A	S

interessante é mostrado na figura da direita e acontece quando, a partir das letras “o” de “rotas” e “sator”, seguirmos pelo tabuleiro com o mesmo movimento do cavalo em um jogo de xadrez, obtemos duas vezes a frase “oro te, pater” (“oro a ti, pai”), e as letras que sobraram no tabuleiro sem terem sido percorridas formam “sanas” (“cure”). Acredita-se que o Quadrado Latino era colocado ao lado das casas que ofereciam refúgio aos cristãos. O Quadrado à esquerda encontra-se exibido no Museu de Manchester, Inglaterra.

Mas essa não foi a única vez em que religião e criptografia estiveram juntas. Durante a Idade Média o mundo árabe era muito mais avançado que o ocidente nessa arte, e a criptologia ocidental só se manteve através de monges europeus que, entre uma oração e outra, analisavam textos bíblicos e outros documentos religiosos. Seu interesse se dava principalmente porque os textos originais do Antigo Testamento, escritos em hebraico, continham algumas palavras codificadas, embora mais como truques literários do que para guardar segredos. Tais criptogramas eram gerados pela cifra de substituição conhecida como “*atbash*”, que consistia na inversão direta das letras do alfabeto. Se fôssemos codificar nosso alfabeto usando o *atbash*, deveríamos trocar A com Z, B com Y, C com X, D com W e assim por diante. O Livro de Jeremias, por exemplo, fala sobre o reino de “Sheshach”, que em hebraico é o equivalente *atbash* para “Babel”.

No século XVI, Mary, Rainha da Escócia e herdeira legítima do trono da Inglaterra, foi mantida prisioneira em Chartley Hall, sob a vigilância constante de Sir Amias Paulet, leal a Elizabeth I. Seu único contato com o exterior era através de Gilbert Curle, que enviava mensagens criptografadas dentro de barris de cerveja inglesa para aliados que conspiravam por sua libertação e pelo assassinato da Rainha. Sua cifra, como mostra a figura ao lado, substituiu cada letra não por uma outra equivalente, mas por um símbolo inventado. Além disso, não eram apenas as letras que eram codificadas: palavras inteiras como “of” e “you” também possuíam seus símbolos próprios. Havia ainda a sofisticação do uso de quatro caracteres nulos – sem significado algum – inseridos por toda a mensagem para confundir os criptoanalistas. Mas não foi o suficiente, e espiões da Rainha Elizabeth I infiltraram-se na conspiração e acabaram por quebrar o frágil código de Mary, que foi condenada no dia 1 de fevereiro de 1587 e guilhotinada sete dias depois, junto com seus comparsas, enquanto sinos tocavam por todo o país em comemoração.

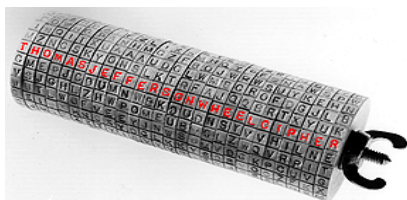


Cifras de substituição como a da Rainha Mary, que trocam letras e expressões por símbolos, também foram muito utilizadas no século XVIII pela maçonaria, para manter seus registros privados. A Cifra Pigpen (“pigpen” é uma caneta usada para marcar porcos) adotada por eles troca cada letra do alfabeto por um símbolo, utilizando a tabela da direita. Para escrever o criptograma, basta substituir as letra da mensagem original pelo símbolo formado pelas linhas ao redor da letra, e o ponto,

d	c	b	a	p	o
f	e	x	q	m	n
g	h	i	j	k	

x	w	z
u	v	y
t	s	





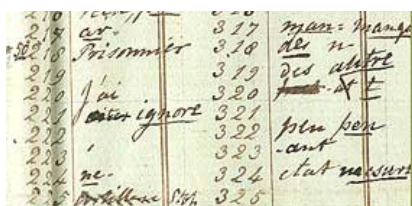
No final do século XVIII, centenas de anos após a consagração do método espartano, Thomas Jefferson desenvolveu uma técnica similar – que, no entanto, é uma cifra de substituição, e não de transposição – conhecida como “A Roda de Jefferson”. Ela consiste em vinte e seis pedaços cilíndricos de madeira do mesmo tamanho, atravessados concêntricamente por um eixo de ferro. As letras do

Rule 7 In deciphering regard it to be had to the place, order, combination, frequency, & number of Letters.

Arnold Conradus, que continha uma série de regras e princípios para criar e quebrar criptogramas. O livro fornecia ainda uma série de exercícios e instruções para tratar cifras em inglês, alemão, holandês, latim, francês e italiano. Com base nesse conhecimento, Scovell elaborou um método geral para proteger as mensagens britânicas dos espiões franceses de Napoleão. Para cada quartel-general inglês era enviada uma cópia do mesmo dicionário. Mensagens codificadas por seu sistema substituiriam então cada palavra por um código indicando a posição relativa da palavra naquele dicionário. O criptograma “74C9”, por exemplo, levaria o receptor à página 74, coluna C, nona palavra de cima para baixo. Um código simples, mas de muito sucesso na época.

Já as tropas napoleônicas codificavam suas mensagens usando métodos simples de substituição conhecidos como *petits chiffres*, baseados em 50 números e desenvolvidos justamente para serem escritos e decifrados rapidamente no campo de batalha. Continham essencialmente pequenas instruções ou ordens de comando para as tropas. Cartas maiores e de conteúdo mais sensível eram enviadas com um código mais robusto, baseado na combinação de 150 números. Essa cifra ficou conhecida como o Código do Exército de Portugal (que estava sob domínio francês na época). Mesmo tratando-se de uma cifra mais complexa que a anterior, George Scovell conseguiu quebrá-la em apenas dois dias.

Depois disso, novas tabelas criptográficas foram enviadas de Paris para todos os líderes militares franceses, dessa vez baseadas num código diplomático do século XVIII e utilizando 1400 números para substituição. O novo guia possuía ainda indicações sobre como enganar o inimigo, por exemplo, adicionando código sem sentido no final das cartas, já que os criptoanalistas costumam tentar decifrar o final da carta primeiro, procurando por frases padrão que encerram correspondências formais. Esse método criptográfico ficou conhecido como a Grande Cifra de Paris, e durante um ano George Scovell analisou documentos interceptados, fazendo progressos graduais que permitiam a coleta de algumas informações, como movimento de tropas e identificação de pessoas e lugares. Em dezembro de 1812, uma carta de José Bonaparte para Napoleão foi interceptada, e o código já havia sido comprometido o



210	100	210	100
218	108	317	108
219	109	318	109
220	110	319	110
221	111	320	111
222	112	321	112
223	113	322	113
224	114	323	114
225	115	324	115
226	116	325	116

Análise de Scovell da Grande Cifra de Paris

suficiente para que os ingleses pudessem decifrar a maior parte dos avisos sobre planos e operações. Isso permitiu que as tropas inglesas se preparassem para a batalha final pelo controle da Espanha, em 21 de junho de 1813, com considerável vantagem. Naquela mesma noite os ingleses invadiram a central de comando de José Bonaparte e descobriram sua cópia da tabela da Grande Cifra de Paris, quebrando finalmente o código por completo.

Mas não é só de segredos de Estado que vive a criptografia. Pensadores, religiosos e poetas foram estimulados a codificar seus trabalhos ao longo da história. Leonardo da Vinci (1452-1519), por exemplo, escrevia seus projetos com letras invertidas, de modo que não fizessem sentido ao leitor ocasional, mas que pudessem ser compreendidas facilmente quando exibidas através de um espelho. No Brasil, durante a ditadura militar (1964-1985), jornalistas trocavam o contexto de notícias para burlar a censura, como relatar um evento como se fosse uma receita de culinária, ou previsão do tempo.

Os avanços tecnológicos do século XX foram sem dúvida essenciais para a criptologia, e surgiram as primeiras máquinas criptográficas. No entanto, a manipulação e computação dos criptogramas por essas máquinas ainda eram pouco eficientes, e técnicas alternativas continuavam sendo utilizadas. Um dos casos de maior sucesso foi o uso de idiomas de tribos indígenas estadunidenses, conhecidas apenas por seus nativos, para enviar mensagens sigilosas. Durante a Primeira Grande Guerra, o *Choctaw* foi muito utilizado, e na Segunda Grande Guerra, o *Navajo*. De fato, as máquinas criptográficas estadunidenses levavam 30 minutos para codificar, enviar e decodificar uma mensagem em inglês de três linhas, enquanto que os índios Navajos recrutados levavam 20 segundos para a mesma tarefa.

Com o advento da computação, no entanto, técnicas mais avançadas puderam ser desenvolvidas e a criptologia passou a ser tratada como uma ciência de fato, com grande base matemática, e ainda estimulada pela indústria bélica. Isso possibilitou a criação de máquinas criptográficas mais complexas e ao mesmo tempo relativamente pequenas e portáteis, podendo ser rapidamente instaladas em navios, submarinos e bases militares. A captura da máquina criptográfica alemã “Enigma” durante a Segunda Grande Guerra, e a quebra de seu código pelo grupo britânico e polonês liderado por Alan Turing foram, sem dúvida, essenciais para a vitória dos Aliados, que puderam ler todas as mensagens inimigas interceptadas.



"Enigma"



Após a Segunda Grande Guerra veio a Guerra Fria (1945-1991), entre EUA e URSS, e a espionagem era a maior arma de ambos os lados. Durante esse período, enormes estímulos militares foram concedidos para o avanço da criptografia, na tentativa de evitar que um lado descobrisse informações importantes sobre o outro, e vice-versa. Ao mesmo tempo, a criptoanálise era um assunto levado muito a sério. Em 1943, antes mesmo da “formalização” da Guerra Fria, os EUA já haviam iniciado um projeto secreto mais tarde chamado “Venona”, com o único objetivo de interceptar e explorar comunicações diplomáticas soviéticas codificadas. Em 1968, os soviéticos conseguiram acesso à máquina criptográfica KW-7 e seus manuais, e, durante 17 anos, puderam ouvir mensagens e relatórios que a CIA e as forças armadas estadunidenses emitiam.

Com o término da Guerra, foi a vez da indústria estimular o campo da criptologia. Empresas de grande porte, com filiais espalhadas pelo mundo, sofrem a cada ano enormes prejuízos com a chamada “espionagem industrial”. Bancos e empresas que realizam comércio eletrônico também são alvos constantes de criminosos, que fazem uso da tecnologia para roubar informações como números de cartão de crédito de clientes ou mesmo para forjar transações comerciais em benefício próprio. Assim, sempre que informações importantes e confidenciais são enviadas de um lugar a outro, criptografia forte é – ou deveria ser – utilizada. A criptologia é, portanto, uma ciência importantíssima para a Segurança Nacional, bem como para qualquer empresa que não deseja ver seus segredos expostos à concorrência, ou prejudicar seus clientes.

3. Curiosidades

3.1 A Vida Secreta de Alice & Bob

Os criptólogos ocidentais, ao escreverem em inglês sobre seus métodos e técnicas, costumavam usar frases como “A está se comunicando com alguém que afirma ser B. Para ter certeza, A envia para B...” Com o passar do tempo, para tornar as definições mais amigáveis, as letras foram transformadas em nomes, e ‘A’ se tornou Alice, enquanto ‘B’ virou Bob. Esses personagens criptológicos foram usados para explicar tantos protocolos e definições que atualmente Alice e Bob são duas celebridades no mundo das cifras, com relatos biográficos e tudo mais.

Ao longo dos anos, Alice e Bob já tentaram fraudar companhias de seguro, jogaram pôquer com apostas altíssimas por e-mail, e conversam ativamente em telefones grampeados. Ninguém sabe ao certo, mas Bob parece ser um corretor subversivo da bolsa de valores, e Alice, por tentar comprar tantas ações com ele, uma especuladora. Mas Alice não quer que seu marido saiba sobre as transações entre ela e Bob, e muitos suspeitam que eles tenham um caso. Seus tópicos principais de conversa são fraude no imposto de renda e derrubadas de Governo, o que os deixou com inimigos poderosos, como os Fiscais da Receita e a Polícia Secreta.

E não é só isso. Apesar de tantas transações comerciais, Alice e Bob parecem nunca ter acesso a um telefone decente, e a linha deles está sempre cheia de ruído, de modo que eles dificilmente conseguem ouvir o que um ou outro está dizendo. Mas eles não estão sozinhos: Carol, Dave, Eve, Trent, Victor e Walter são alguns de seus amigos e inimigos, nessa eterna novela mexicana da criptologia.

3.2 As Cifras de Beale

Desafios populares de métodos mais simples tornaram-se relativamente comuns a partir do século XIX. O escritor Edgar Allan Poe costumava resolver desafios de cifras de substituição enviadas a ele por leitores. Publicações periódicas e livros como os de Sherlock Holmes continham desafios intelectuais que entretinham leitores entusiastas da criptoanálise. Na maioria dos casos, tais exercícios não interessavam muito os profissionais da área. Mas houve uma grande exceção.

Em 1885, num panfleto intitulado “As Cifras de Beale”, um autor anônimo relatava uma história contada a ele por Robert Morriss, dono do Hotel Washington em Lynchburg, Estados Unidos. Conta ele que, em 1820 e 1822, um homem chamado Thomas Beale foi hóspede do Hotel e acabou tornando-se amigo de Morriss. Em sua última visita, deixou com Morriss uma caixa lacrada com instruções escritas dizendo que a caixa continha papéis de grande valor e importância, e que não deveria ser aberta por dez anos. As instruções diziam que tais papéis seriam ininteligíveis sem o auxílio de uma “chave”, e que essa “chave” seria entregue a ele por alguém após os dez anos.

Mas 1832 passou e Morriss não teve notícias de Beale ou da pessoa misteriosa com a chave. Beale havia confiado em Morriss, e este provou sua confiança abrindo a caixa somente em 1845. A caixa continha três documentos cifrados numa lista de números, e uma nota escrita por Beale. Ela dizia que ele e uns colegas haviam ido para o oeste em direção a cidade de Santa Fé e que, rumando um pouco para o norte, encontraram por acidente uma enorme quantidade de ouro. A garimpagem lhes rendeu muito ouro e prata, e foi confiada a Beale a tarefa de levar o tesouro de volta ao leste, e coloca-lo em local seguro. Beale fez duas viagens, enterrou o tesouro em Lynchburg e deixou com Morriss os documentos para que alguém distribuisse o tesouro aos parentes, caso alguma tragédia acontecesse aos aventureiros.

Aparentemente alguma tragédia de fato aconteceu, já que Beale e seu grupo desapareceram completamente. Morriss tentou por quase duas décadas decifrar os documentos, mas em 1862 ele percebeu que não tinha muito mais tempo de vida, e contou ao autor anônimo do panfleto toda sua história. No final do panfleto, o autor lamenta ter ouvido sobre o tesouro e a cifra, pois passou 23 anos tentando encontrar o tesouro de Beale, gastando todos os seus recursos e levando sua família à pobreza. Para eliminar a tentação, decidiu tornar a história pública, assim como suas descobertas.

Ele conseguiu decifrar completamente um dos três textos, que descobriu ser uma “cifra de livro”, em que os números do criptograma indicavam a posição das palavras em algum texto externo. No caso do primeiro documento, o texto utilizado foi a declaração da independência dos Estados Unidos. O documento revela que o tesouro era de aproximadamente uma tonelada de ouro, duas de prata e algumas pedras preciosas, obtidas para fins de portabilidade e troca, e foi guardado em potes de ferro.

Desde então houve incontáveis tentativas para decifrar os dois últimos documentos, mas até hoje não se sabe de ninguém que tenha obtido sucesso. A maior dificuldade encontra-se na quantidade muito pequena de texto cifrado, além do fato de que, se os documentos também são “cifras de livro”, o número de possíveis textos base é praticamente ilimitado. E então? Acha que está com sorte?

4. Bibliografia & Crédito das Imagens:

Página Oficial de Simon Singh – <http://www.simonsingh.net/> (30/09/05)

“National Archives of the English Public Record Office” – <http://www.pro.gov.uk/> (30/09/05)

“Codes and Ciphers in The Second World War” - <http://www.codesandciphers.org.uk/> (22/03/03)

“Navajo Code Talkers” - <http://www.americanindians.com/CodeTalkers.htm> (22/09/05)

GECEM – Grupo Escoteiro Caio Martins - <http://www.gecem.org.br/> (03/05/03)

“Alice and Bob After-Dinner Speech” – John Gordon - <http://www.conceptlabs.co.uk/alicebob.html>
(10/09/05)