Criptografia, Segurança de Dados e Privacidade - Até que ponto pode-se confiar na discrição dos computadores?

Verônica Lagrange Moutinho dos Reis

TESE SUBMETIDA AO CORPO DOCENTE DA COORDENAÇÃO DOS PROGRAMAS DE PÓS-GRADUAÇÃO DE ENGENHARIA DA UNIVERSIDADE FEDERAL DO RIO DE JANEIRO COMO PARTE DOS REQUISITOS NECESSÁRIOS PARA A OBTENÇÃO DO GRAU DE MESTRE EM CIÊNCIAS EM ENGENHARIA DE SISTEMAS E COMPUTAÇÃO.

Aprovada por:

Profa. Dina Feigenbaum Cleiman, D. Sc.
(Presidente)

Prof. Paulo Mario Bianchi França, Ph. D.

Prof. Valmir Carneiro Barbosa, Ph. D.

RIO DE JANEIRO, R.J. - BRASIL
AGOSTO DE 1989

REIS, VERÔNICA LAGRANGE MOUTINHO DOS
Criptografia, Segurança de Dados e Privacidade: Até que ponto pode-se confiar na discrição dos computadores? [Rio de Janeiro] 1989
XI, 128 p. 29,7 cm (COPPE/UFRJ, M.Sc., Engenharia de Sistemas, 1989)
Tese - Universidade Federal de Rio de Janeiro, COPPE
1. Criptografia I. COPPE/UFRJ II. Título (série).

Qbsb Nbsjb Bohfmb ,

Xbmfsjb , Dmbvejp ,

Xjshjojb f Lvmjfub.

Agradecimentos:

Sem a menor sombra de dúvida, a existência dessa tese se deve à brilhante orientação recebida da Profa. Dina.

Aos professores Valmir e Bianchi agradeço a honra que me deram em participar da minha banca e às valiosas sugestões recebidas.

À minha família e ao meu namorado pelo apoio, pelo incentivo e pelas horas de lazer que não pude compartilhar com eles.

Às diversas pessoas que de alguma forma contribuíram com sugestões, dicas, referências sugeridas e emprestadas: Aldner Oliveira, Alfredo, Antenor, Carlos Mendes, Eduardo, Gonzalo, Maria Angela, Maria Regina Barbosa, Sergio Guedes, Stelling e Valéria (perdoem-me se esqueci alguém).

Ao NCE e seus funcionários pelo incentivo.

À Claudia Noé pelos desenhos da máquina rotor e ao Claudio e ao Antenor pelos desenhos usados durante a apresentação da tese.

Resumo da Tese apresentada à COPPE/UFRJ como parte dos requisitos necessários para obtensão do grau de Mestre em Ciências (M. Sc.)

CRIPTOGRAFIA, SEGURANÇA DE DADOS E PRIVACIDADE: ATÉ QUE PONTO PODE-SE CONFIAR NA DISCRIÇÃO DOS COMPUTADORES?

Verônica Lagrange Moutinho dos Reis

Agosto de 1989

Orientador: Profa. Dina Feigenbaum Cleiman

Programa: Engenharia de Sistemas e Computação

Esta dissertação descreve um estudo do estado-da-arte da criptologia. São analisadas sua história, aplicações, os algoritmos da criptografia clássica e computacional e as técnicas criptoanalíticas. A seguir propomos uma nova estratégia para ataque dos algoritmos atualmente considerados seguros, através de busca heurística. Procuramos, também, situar a criptografia dentro da segurança de dados dos sistemas de computadores e abordamos sob um enfoque social o problema de sigilo e completeza de informações consideradas importantes.

Abstract of Thesis presented to COPPE/UFRJ as partial fulfillment of the requirements for the degree of Master of Science (M. Sc.)

CRYPTOGRAPHY, SECURITY OF DATA AND PRIVACY: UP TO WHICH POINT CAN WE TRUST COMPUTER'S DISCRETION?

Verônica Lagrange Moutinho dos Reis

August, 1989

Thesis Supervisor: Prof. Dina Feigenbaum Cleiman
Department: Systems Engineering and Computer Science

In this thesis we develop a study about the state-of-the-art of cryptology. We analyze its history, applications, the algoritms of the classic and computational cryptography and the techniques available for cryptanalysis. We then propose a new strategy to attack the algoriths that are now considered safe. This new strategy is based on heuristic search. We also try to place cryptography inside the field of computer systems' data security. Lastly, we talk about the problem of secrecy of important data and privacy from a social point of view.

ÍNDICE:

ÍNDICE DE FIGURAS

CAPÍTULO I - INTRODUÇÃO

Esta dissertação se propõe a fazer um estudo sobre o que vem a ser uma ciência ao mesmo tempo muito antiga e bastante atual: a CRIPTOGRAFIA. A criptografia, do grego kryptós (esconder) e grápho (escrita), estuda formas de se camuflar uma mensagem a ser transmitida por meios não totalmente imunes a "xeretas" de maneira que estes não consigam entender o que é comunicado.

É uma ciência muito antiga, pois tem-se notícia de textos criptografados desde 1900 A.C. [1] e a primeira descrição de um sistema de criptografia militar conhecida data do quinto século A.C. [1]. Verificou-se que a criptografia se desenvolve muito em tempos de guerra, por motivos óbvios. No entanto, com o advento do computador e mais modernamente, com os novos usos que se tem dado aos sistemas de computadores, "criptografar" e "decriptografar" se tornaram atividades diárias em muitos CPD's, agências bancárias, bolsas de valores etc.

Para melhor situarmos a criptografia computacional contemporânea, analisaremos outras técnicas utilizadas na proteção de sistemas de computadores (dados e programas). Analisaremos também os problemas sociais, políticos e econômicos que provocaram o grande desenvolvimento da criptografia nas últimas décadas.

A motivação para a escolha do tema se deu a partir de um estudo de códigos (semiologia) ocorrido durante o curso de Informática e Sociedade I e de um trabalho sobre

criptografia desenvolvido no curso de Sistemas Operacionais I. Tópicos como a quantidade de informação de cada símbolo de determinado código (por exemplo, cada letra do nosso alfabeto), redundância e entropia (ruído) que são diretamente aplicáveis à criptografia, onde se pretende esconder a informação (o significado de cada símbolo) através da inserção de ruído e da retirada da redundância.

No próximo capítulo estudaremos as origens da criptografia e como esta se desenvolveu ao longo do tempo.

No terceiro capítulo veremos quais as aplicações atuais e qual a sua razão de ser. Em outras palavras, qual a sua importância e quais as técnicas utilizadas para proteção de dados (em arquivos e quando são transmitidos) e para verificação de identidade (para saber se a "pessoa" com quem eu estou "falando" realmente é quem ela diz ser).

A seguir, descreveremos com detalhe as principais técnicas e algoritmos utilizados pela criptografia clássica e pela criptografia computacional.

No quinto capítulo propomos novas estratégias para ataque criptoanalítico, isto é, uma metodologia de ataque a algoritmos considerados inquebráveis até então. Esta metodologia foi desenvolvida a partir do estudo feito anteriormente das técnicas atuais tanto de criptografia quanto de criptoanálise e de ferramentas recentemente desenvolvidas na área de inteligência artificial, em especial, a busca heurística.

No sexto capítulo tentaremos localizar a criptografia num contexto maior em que ela se insere que é a chamada Segurança de Dados. A segurança de dados se preocupa com a

integridade das informações armazenadas, transmitidas e utilizadas como um todo. Ela deverá garantir que a informação seja verdadeira, coerente com os outros dados disponíveis e pode ser acessada por quem de direito.

O sétimo capítulo enfoca o problema do ponto-de-vista social e legal: a necessidade de legislação para proteção da informação, seja ela na forma de programas de computadores (software) ou na forma de dados sobre pessoas físicas ou jurídicas. No primeiro caso, têm sido criadas leis à base do direito autoral. No segundo, a questão passa pelo direito à vida privada.

Nosso estudo nos levou à conclusão da importância estratégica da criptografia, no sentido tanto de domínio da informação e da técnica quanto de autonomia econômica, política e social.

Consta ainda dessa dissertação um glossário dos termos mais comuns em criptografia, criptoanálise e segurança de dados e um anexo sobre distribuição de frequência das letras na língua portuguesa.

Como principais contribuições dessa dissertação destacamos a proposiçãao de uma nova metodologia para desenvolvimento da criptoanáalise através de técnicas de inteligência artificial e a avaliação dos impactos causados na sociedade em conseqüência da utilização da teleinformática no tratamento de dados sensíveis.

CAPÍTULO II - PEQUENA HISTÓRIA DA CRIPTOGRAFIA

Neste capítulo retornaremos muitos anos na história da humanidade na tentativa de buscar as origens da criptografia e analisar como ela vem se desenvolvendo. Nesses dois aspectos, a principal fonte de referência é, sem dúvida, The Codebreakers, de DAVID KAHN [1], que narra com detalhes toda a história da criptografia e criptanálise dando ênfase aos acontecimentos criptológicos das duas Grandes Guerras.

II.1 - Criptografia ao Longo do Tempo

Kahn nos conta que o primeiro vestígio de criptografia conhecido data de 1900 A.C., que a primeira descrição de um sistema de criptografia militar de que se tem notícia foi feita pelos gregos no quinto século antes de Cristo e que a criptanálise surgiu com os árabes nos anos 600. Existem listas dos diversos tipos de criptografia conhecidos na época incluindo sistemas de transposição e substituição. Essa criptanálise árabe, no entanto, se confundia com estudo léxico e vice-versa pois encontram-se também textos tratando da freqüência das palavras do Corão tanto para estudo da cronologia dos diversos capítulos (quais os mais antigos e quais os mais recentes) quanto para estudos criptanalíticos.

Não há nenhum documento que comprove terem sido esses conhecimentos criptológicos repassados à Europa Medieval. Lá, a criptologia (ciência que engloba a criptografia e a

criptoanálise) só começou a se desenvolver em alguns principados italianos a partir de 1400, tendo realmente se expandido com o aparecimento da diplomacia moderna. Pela primeira vez, estados mantinham relações permanentes com outros. Os embaixadores mandavam para casa, regularmente, relatórios. Intrigas, suspeitas e ciúmes entre as cidades-estado italianas muitas vezes tornavam necessário o uso de criptografia.

Daí pra frente a criptologia deslanchou, tendo inclusive surgido a profissão de "criptólogo do rei". Na Renascença, surgiu a substituição polialfabética, que foi considerada inquebrável durante muito tempo. Somente em 1863 um criptólogo amador, o oficial prussiano Kasiski publicou um pequeno livro que resolvia o problema insolúvel há mais de 300 anos: como chegar a uma solução geral de textos criptografados por substituição polialfabética com chave repetida. Apesar dessa solução ter aberto as portas para a criptologia moderna, não despertou muito interesse na época, tendo o próprio Kasiski se desinteressado da criptologia e se tornado um antropologista amador.

A próxima revolução da criptografia viria com o surgimento do telégrafo. Podemos dizer que o telégrafo democratizou o uso da criptografia, antes só usada em comunicações militares ou diplomáticas de altíssima importância (e por amantes). O telégrafo deveria transmitir desde simples sinais até as mensagens acima citadas. Seriam então necessários mais de um sistema de criptografia, cada um deles adequado a um nível de comunicação. Assim, o

telégrafo estimulou a invenção de novos sistemas de criptografia e conseqüentemente, de criptanálise. Segundo Kahn, foi o telégrafo que deu à criptografia moderna sua estrutura e seu conteúdo.

O rádio, inventado em 1895, seria o fator que revolucionaria a criptanálise. Se o telégrafo havia tornado as comunicações militares mais eficientes mesmo aumentando a possibilidade de interceptação, o rádio aumentou muitas vezes tanto a eficiência da comunicação quanto a da interceptação. A partir desse momento não era mais necessário ter acesso físico à linha de comunicação para interceptá-la: bastava sintonizar na mesma freqüência do inimigo. Dessa forma, o rádio introduziu dois novos fatores para o desenvolvimento da criptanálise: quantidade e continuidade.

Assim, telégrafo e rádio se complementaram no sentido de que enquanto o primeiro criou a criptografia moderna, o segundo provocou o desenvolvimento da moderna criptanálise. Um desenvolveu a criptologia internamente e o outro externamente. O rádio completou o trabalho que o telégrafo havia começado trazendo a criptologia para o mundo real: pela primeira vez, durante a Primeira Guerra Mundial a criptologia foi largamente utilizada.

A partir de então, a criptologia passou a ser uma arma, e necessária. Durante a Primeira Guerra os EUA empregaram 400 pessoas em criptografia/criptanálise e na Segunda Guerra esse número cresceu para 16 mil.

Outros marcos importantes são Friedman que associou a criptologia à estatística em seu livro chamado **The Index of**

Coincidence. Antes dele, todas as contagens de freqüência, características lingüísticas, enfim, técnicas desenvolvidas eram peculiares e particulares da criptologia. A partir daí, a criptologia passou a contar com todo o poderoso ferramental estatístico. Foi ele também que cunhou o termo CRIPTOANÁLISE, em 1920. Até então havia o problema da ambiguidade do verbo DECIFRAR, usado tanto para "traduções" autorizadas quanto não autorizadas [1]; SHANNON [2], com sua Teoria Matemática da Informação, que sem dúvida tem grande influência na criptologia contemporânea , através principalmente da demonstração matemática de sistemas criptográficos inquebráveis e do conceito de redundância (de acordo com [3], os sistemas criptográficos da IBM, inclusive o DES, têm suas raizes no artigo em que Shannon associa criptografia à Teoria da Informação [4]); e a indústria criptológica, surgida durante a Primeira Guerra e que fornece criptografadores a diversas instituições governamentais e privadas no mundo inteiro até os dias de hoje. Seu empresário mais bem sucedido foi Hagelin, que ganhou milhões em royalties com suas invenções durante a guerra e quando esta acabou se estabeleceu na Suíça, vendendo equipamentos a vários países. Uma descrição da máquina criada por ele se encontra no capítulo IV.

Muitas descobertas da área foram feitas por amadores como Kasiski e também várias personalidades famosas em outras áreas eram criptólogos amadores. Entre eles destacamos: Charles Babbage, Alan Turing, Thomas Jefferson, Edgar Allan Poe e Julio Verne.

Uma aplicação importante e muito interessante de criptoanálise que vale a pena mencionar aqui é a decifração de escritas antigas como, por exemplo, os hieróglifos egípcios.

Kahn menciona o computador pela primeira vez ao falar da N.S.A. (National Security Agency) e seu papel durante a guerra fria. Diz ele em seu livro que a NSA tem provavelmente mais computadores que qualquer outra instalação no mundo. Ela tem não só máquinas comerciais de uso geral como computadores construídos por ela para tarefas específicas.

Com o aparecimento do computador em cena, surgiram novas aplicações para a criptografia como transferência automática de fundos, processamento distribuído e sua necessidade de verificação de identidade, proteção de dados em arquivos etc.

II.2 - Criptografia Computacional

MEYER [5] considera "marcos" da criptografia computacional a adoção do DES(Data Encryption Standard) em Janeiro de 1977 pela NBS(National Bureau of Standards) como padrão federal e em Dezembro de 1980 pela ANSI(American National Standards Institute) para uso comercial nos EUA e a proposição de sistemas de chave pública.

O DES, desenvolvido pela IBM, será descrito com mais detalhes no capítulo IV. Desde a sua padronização, ele vem sendo extensivamente utilizado, também por ser muito rápido. Vale a pena mencionar que é um algoritmo cercado de muita polêmica, principalmente na época em que foi proposto. Foram levantados além de problemas técnicos, problemas políticos. Alegava-se que sua adoção evitaria muito trabalho à N.S.A., pois esta não precisaria perder tempo criptoanalisando outros algoritmos que pudessem surgir, principalmente na indústria que estava nascendo [6]. No entanto, o DES parece razoavelmente seguro, apesar da observação citada por MEYER [6] sobre vulnerabilidade cruzada: "Quando um número muito grande de pessoas se utiliza de um mesmo algoritmo criptográfico, esse fato aumenta grandemente o possível retorno econômico que se poderia ter na quebra daquele sistema criptográfico." De toda a literatura consultada, 16 livros, 1 tese e 62 artigos, num total de 79 referências, 31 delas mencionam o DES. Das que não mencionam, 11 referências foram escritas antes de 1977 e 37 tratam de questões específicas como a

descrição de um algoritmo ou tratam de outros assuntos como inteligência artificial ou privacidade.

A proposição de sistemas de chave pública foi feita por W. Diffie e M. Hellman em 1975 [7]. A idéia é fazer com que existam duas chaves: uma de encriptação e outra de decriptação. Uma delas é pública e conhecida de todos e a outra é secreta e de conhecimento apenas de seu proprietário. É óbvio que as coisas devem ser feitas de maneira que não se consiga descobrir a chave secreta a partir da pública. O algoritmo mais famoso e utilizado é o RSA, de Rivest, Shamir e Adleman, proposto em 1978. Praticamente toda a segurança bancária se baseia nele. O algoritmo, explicado mais detalhadamente no capítulo IV, foi desenvolvido a partir do fato de ser muito difícil, computacionalmente falando, fatorar grandes números que são produtos de 2 números primos também muito grandes. Atualmente muitos pesquisadores se ocupam desse problema de fatoração, utilizando inclusive o fato de as máquinas serem cada vez mais velozes. Em 1982, Wunderlich e Simmons usando um CRAY-1 conseguiram fatorar números de 60 dígitos em apenas 1 hora [7]. Dessa data em diante era necessário usar primos de 100 dígitos por motivo de segurança ... só que em Outubro de 1988 a imprensa noticiou que dois pesquisadores americanos, Manasse e Lenstra, conseguiram fatorar um produto de dois primos com 100 dígitos! A mesma notícia afirma que a partir de agora os números seguros são os de mais de 200 dígitos ... [8]. Apesar disso, a criptografia de chave pública vem se desenvolvendo já existindo, inclusive, chips que implementem algoritmos de

chave pública [9].

Voltando à NSA, trata-se de uma cidade inteira dedicada à criptologia. Essa agência se manteve tão secreta durante tanto tempo que muitos americanos, mesmo os políticos, nunca ouviram falar dela. Alguns até brincam dizendo que NSA significa "No Such Agency" ou "Never Say Anything". Até o final dos anos 60 a NSA procurava se manter pelo menos 5 anos à frente do estado-da-arte em criptologia. Com o surgimento da criptologia computacional, no entanto, essa tarefa ficou bem mais difícil e a NSA começou uma espécie de intercâmbio tecnológico, primeiro com companhias como a IBM (no desenvolvimento do Lucifer, primeira versão do que viria a ser o DES) e a seguir com a comunidade científica. A partir de então, uma série de conflitos entre a comunidade (que queria divulgar suas descobertas) e a NSA (que queria controlar todo o conhecimento da área, divulgando apenas o que não comprometesse a "segurança nacional") acabou levando à aprovação, pelo governo, de uma portaria (em fevereiro de 1981) que permite à NSA censurar toda a literatura produzida na área. Entre os fatos que levaram a isso destacam-se: a homologação do DES com uma chave de 56 bits em vez da de 112 bits originalmente proposta (a comunidade científica argumentava que essa redução foi forçada pela NSA para que ela, e apenas ela, pudesse quebrá-lo); a proposição de sistemas de chave pública, totalmente criada e desenvolvida em universidades e a criação da revista Cryptologia, em 1977. Isso sem falar dos diversos congressos e seminários que começavam a ocorrer nos Estados Unidos, contando,

inclusive, com conferencistas estrangeiros [10].

II.3 - Criptografia no Brasil

Por aqui a criptografia computacional vem conquistando seus espaços apesar de ainda estar longe de ser uma técnica amplamente difundida e dominada. Segundo reportagem da Revista Info [11], até 1984 apenas os órgãos governamentais que desempenham atividades estratégicas consideradas de segurança nacional usavam criptografia, e mesmo assim, até 1975 (quando surgiu o Cepesc (Centro de Pesquisas para Segurança das Comunicações)) não existia nada que tivesse sido desenvolvido por aqui. Somente quando surgiram as primeiras redes de teleprocessamento privadas, devido à automação bancária, a criptografia se tornou conhecida dos profissionais de informática. A partir de então cresceu a demanda por cursos de pós-graduação com especialização em criptografia. Ao mesmo tempo crimes eletrônicos começaram a ocorrer no pais, levando o tema a ser discutido em congressos [11].

O primeiro crime eletrônico de que se tem notícia no Brasil ocorreu no início dos anos 80, quando quatro pessoas interceptaram as transmissões entre as agências do Banco Meridional de Guarulhos(SP) e de Porto Alegre(RS) e fizeram transferências eletrônicas de fundos que, na época, renderam Cr$ 2,5 milhões jamais recuperados. O roubo foi descoberto pelos próprios computadores algumas semanas mais tarde quando os "hackers" brasileiros tentaram nova investida

[11].

Em 1985, a SEI(Secretaria Especial de Informática) criou uma comissão especial para tratar da questão de segurança de dados. Em março de 1987, o Comitê Brasileiro de Informática instalou a Comissão de Estudos de Técnicas Criptográficas, presidida pelo engenheiro Almir Paz de Lima, um dos maiores especialistas do assunto no Brasil. Essa comissão está empenhada em normatizar o uso da criptografia seguindo os passos da ISO(International Organization for Standardization) [11].

Também a ABNT(Associação Brasileira de Normas Técnicas) vem se preocupando com o problema, visando principalmente a criação de uma rede eletrônica de transferência de fundos. Ainda segundo a Info, a tendência é que a ABNT adote um sistema híbrido mesclando o DES com sistemas de chave pública.

O Cepesc, vinculado ao SNI(Serviço Nacional de Informações), é o único laboratório nacional gerador de tecnologia. A Prólogo S.A., subsidiária da Imbel(Indústria de Material Bélico do Exército) fabrica equipamentos criptográficos com tecnologia do Cepesc [11]. A Revista Dados & Idéias de março de 1988 nos informa, numa pequena nota, que a Microlab começa a produzir criptografadores que serão fornecidos ao Centro Tecnológico do Exército. A Microlab também utiliza tecnologia do Cepesc.

A história do surgimento da Prólogo é bem interessante: no início da década de 80, o Itamaraty verificou que havia quebra de sigilo nas transmissões criptografadas com

equipamentos Gretag (Suíço) pois nas concorrências de preço mínimo da Bolsa de Commodities de Londres o café brasileiro nunca obtinha bom preço. Desconfiava-se que os concorrentes conheciam de antemão a proposta brasileira. Ciente do perigo de usar aparelhos estrangeiros para criptografar mensagens estratégicas, o governo decidiu desenvolver e fabricar seus próprios equipamentos. Um ano depois a Prólogo recebia autorização do Cepesc para fabricar o primeiro equipamento: o CD-1. Atualmente a linha de equipamentos da Prólogo para voz e dados é de seis produtos, todos eles competitivos a nível de preço e desempenho com os equipamentos mais simples fabricados no mercado mundial [11].

Na figura II.1 apresentamos um dos criptogramas modernos mais difíceis de serem criptoanalisados...

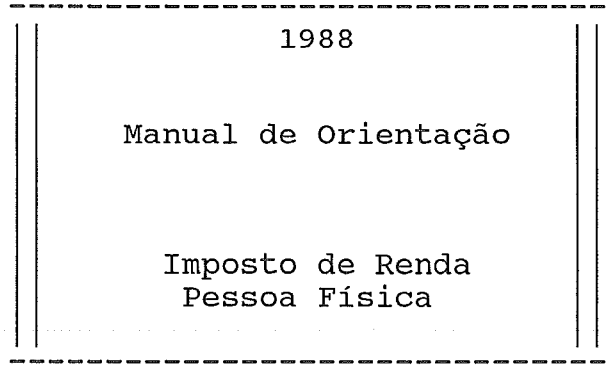
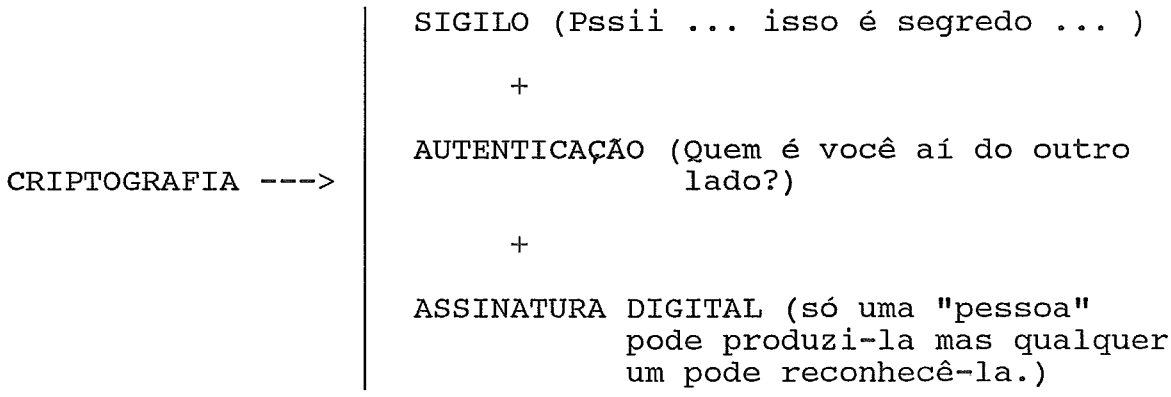


FIGURA II.1: O Criptograma mais indecifrável do mundo ...
(Adaptado de [L7])

CAPÍTULO III - IMPORTÂNCIA E ATUALIDADE DA CRIPTOGRAFIA

Desde que foi inventada, na antigüidade, até o advento do computador, a criptografia foi usada principalmente para proteger o **sigilo** da comunicação. Em outras palavras, sua função era evitar que terceiros se inteirassem do conteúdo da mensagem. Assim, as milhares de técnicas desenvolvidas (das quais algumas serão vistas no capítulo IV) tinham basicamente aplicações militares e diplomáticas (e sentimentais). Uma outra aplicação se encontra na área do entretenimento: diversos jornais e revistas trazem em suas colunas de diversões pequenos problemas criptográficos.

O uso da criptografia para **autenticação do remetente** (como saber se a mensagem foi realmente mandada pela pessoa que deveria e não por um "inimigo" se fazendo passar pelo "meu amigo"?) surgiu um pouco antes da "era do computador": na guerra, quando era preciso saber se os aviões que se aproximavam eram amigos ou não. No entanto, esta aplicação realmente se intensificou e passou a figurar lado a lado com o sigilo quando os computadores passaram a ser usados para comunicação de dados e processamento distribuído. **É preciso garantir que o terminal que está tendo acesso remoto ao sistema é o que tem autorização para isso**, que os dados recebidos pelo satélite realmente vieram da máquina cujo "nome" aparece na mensagem. É preciso também garantir que a mensagem recebida não foi alterada, ou seja, há necessidade de se **autenticar a mensagem**, o que também é conseguido através da criptografia.



SIGILO (Pssii ... isso é segredo ...)

+

AUTENTICAÇÃO (Quem é você aí do outro lado?)

+

ASSINATURA DIGITAL (só uma "pessoa" pode produzi-la mas qualquer um pode reconhecê-la.)

FIGURA III.1 - Principais Aplicações da Criptografia

Outra aplicação surgida mais recentemente é a da **assinatura digital**, necessária a partir do momento em que bits e bytes passaram a substituir papéis em transações comerciais e bancárias como, por exemplo, transferência eletrônica de fundos. Assinatura digital deve ter as mesmas características de uma assinatura num papel: só uma pessoa pode produzi-la mas qualquer um a reconhece, sendo inclusive viável provar sua autenticidade ou não perante um tribunal, caso ocorra litígio entre as partes. A figura III.1 apresenta as principais aplicações da criptografia.

A criptografia pode ser usada não apenas na proteção dos dados, mas também na proteção dos programas (software), por exemplo, evitando que cópias não autorizadas possam ser executadas pelo sistema.

É interessante notar que os mecanismos (técnicas e algoritmos) que garantem sigilo (que também poderíamos chamar de **"autenticação do destinatário"**), autenticação da mensagem e do rementente e assinatura digital serão vistos no capítulo IV.

Como não poderia deixar de ser, também na era do computador, a diversão é importante e tem seu mercado assegurado. Assim sendo, encontram-se vários jogos criptográficos como Crypto-Mania da IBM [12], Code Quest [13], MasterChip [14], Cryptarithms [15] e o Cryptease [16].

Será apresentada a seguir uma série de aplicações da criptografia, vamos listar várias delas (as mais representativas), explicando por que esse uso se faz necessário e como ele é feito.

Controle de satélites - comandos tipo correção de órbita são enviados a partir de estações terrestres. Nada impede que uma estação não autorizada também envie comandos... para evitar que o satélite "obedeça" esses comandos equivocados, a estação responsável por ele envia seus comandos criptografados.

Transmissão de rádio customizada - são canais especiais contratados por prédios comerciais em que apenas música é transmitida (sem anúncios). Esses sinais são enviados criptografados para que apenas os clientes possam utilizá-los. Segundo nota na revista Dados e Idéias [17] algumas emissoras já o fazem utilizando, inclusive, uma faixa especial de freqüência recentemente aprovada pelo governo para esse fim.

Passwords - segundo [3], "passwords" são a aplicação mais utilizada de criptografia nos meios computacionais. Ocorre que em todo sistema de computadores multiusuário, cada pessoa autorizada a usá-lo deve ter em seu poder uma senha, que é digitada sempre que essa pessoa quer entrar em

sessão. Para proteger essas senhas, elas são guardadas criptografadas pelo computador, que criptografará a senha fornecida e a comparará com a que ele tem guardada.

Proteção de software - criptografia é um dos diversos mecanismos utilizados na proteção do software como, por exemplo, não permitir que cópias ilegais do produto possam ser executadas.

Tranferência eletrônica de fundos - uma das maiores, senão a maior, aplicação bancária de teleprocessamento. Retiradas e depósitos são feitos em qualquer agência ou caixa automático para crédito/débito imediato na agência do cliente. Todos os grandes bancos do país já oferecem esse serviço. As ordens de depósito/retirada trafegam pelos canais da Embratel (principalmente o TRANSDATA), se bem que alguns já começam a implementar suas redes privadas através de satélites [18],[19] e [20]. Não é preciso lembrar que aqui justamente ocorreu o único crime eletrônico de que se tem notícia no Brasil (vide capítulo II) e da importância de se garantir o sigilo e a autenticidade das operações realizadas.

Teleconferências - serviço oferecido pela Embratel e pelas demais companhias de telecomunicações nos outros países. Trata-se de se realizar reuniões, conferências, cursos etc. à distância. Cada grupo de pessoas se reune numa sala numa cidade e através de circuito de televisão tudo o que acontece numa sala pode ser visto e ouvido das outras. Como nem sempre as reunioes tratam de assuntos de domínio público, pode ser necessário garantir o seu sigilo, e uma boa forma de fazê-lo é através da criptografia [21].

Sigilo/autenticação de informações transmitidas - a primeira das aplicações da criptografia englobando, inclusive, a maioria das até aqui citadas. Trata-se de garantir o sigilo das informações transmitidas através de canais inseguros, sujeitos à "escuta", como linhas telefônicas, rádio, telégrafo, sinais de fumaça etc.

Sigilo/autenticação de informações armazenadas - quando as informações armazenadas são de fácil acesso por parte de uma comunidade e é necessário restringi-las por algum motivo, pode-se usar a criptografia. Um exemplo seria um banco de dados de pacientes num hospital. Alguns dados podem ser consultados por pesquisadores para fins estatísticos, desde que se preservem os dados pessoais do doente, como o nome. Essa informação só deverá estar disponível ao médico que cuida dele. Outro exemplo seria o cadastro de funcionários da empresa; não é qualquer pessoa que pode consultar o salário dos diretores... ou mesmo qualquer dado desse arquivo.

Diversão - como já vimos anteriormente, existem vários jogos eletrônicos baseados na criptologia.

Aplicações futuras - quando cada casa tiver uma "tomada de informação" ao lado da tomada do telefone, como prevê MASUDA [22], algumas aplicações da "tomada" exigirão o equivalente a uma "assinatura digital", como compras feitas em casa, pagamento de contas, relatórios enviados à empresa para a qual se trabalha etc. Como veremos no próximo capítulo, assinaturas digitais podem ser implementadas através de criptografia. Assim, quando o futuro chegar,

será possível renovar o estoque de oxigênio da casa de campo na Lua sem sair de casa ...

CAPÍTULO IV - ALGORITMOS E TÉCNICAS UTILIZADOS

Criptografar consiste em camuflar a mensagem a ser transmitida por um meio que não é totalmente confiável de maneira que apenas o destinatário consiga entendê-la. Para tanto, é preciso que apenas o destinatário e o remetente tenham conhecimento de um dado necessário para a correta conversão da mensagem. Esse dado é a **chave**. Assume-se que qualquer outra pessoa conhece o algoritmo utilizado para encriptar/decriptar e que apenas a chave é secreta. Geralmente, um criptossistema funciona como o descrito na figura IV.1.

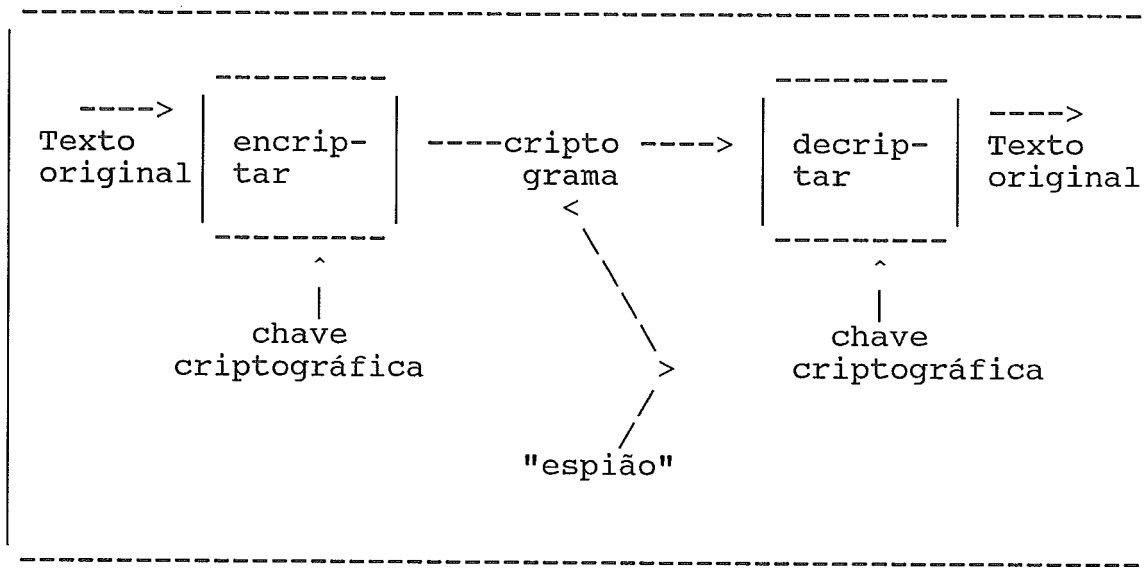


FIGURA IV.1:Criptossistema

Veremos, a seguir, algumas das técnicas mais utilizadas e usuais de criptografia, tanto a tradicional quanto a computacional. No que se refere à criptografia tradicional, analisaremos com detalhes os métodos de substituição

monoalfabética, substituição polialfabética, substituição poligráfica, transposição, códigos e máquina Rotor. Já os métodos de criptografia computacional serão subdivididos em duas categorias: algoritmos de chave secreta e algoritmos de chave pública, a saber:

- Chave Secreta:
  - Registros de Deslocamento;
  - DES;
  - Algumas técnicas para distribuição de chaves.
- Chave Pública:
  - RSA;
  - Mochila;
  - Como esses algoritmos podem ser usados para autenticar e assinar mensagens.

Também faremos um resumo dos princípios em que se baseia a criptografia computacional e de como se pode classificar os algoritmos em relação à transformação sofrida pela mensagem. Mais adiante daremos uma olhada no método de trabalho do "espião", o criptoanalista. Será feita uma análise geral dos recursos disponíveis à criptoanálise e de como ela é conduzida.

IV.1 - Métodos da Criptografia Clássica

IV.1.1 - O Método de Substituição Monoalfabética

Este tipo de substituição consiste na permutação do alfabeto utilizado. Por exemplo, QERTPOIUASDFGLJHZXCVBMN é uma permutação do nosso alfabeto. Usando-se esta permutação, o texto claro BOM DIA se transformaria no criptograma ELF TAQ. Um caso particular de substituição simples é a Cifra de Cesar (acredita-se que Júlio Cesar foi o primeiro a utilizá-la), que consiste num deslocamento de 3 posições no alfabeto. Por exemplo a letra A é substituída por D, B por E, C por F etc. Generalizando, pode-se deslocar o alfabeto de N posições, neste caso N é a chave, e pode variar de 0 a 22. Para o criptanalista, o método de tentativa e erro não é de maneira nenhuma inviável uma vez que só existem 23 possibilidades... Em se tratando da substituição monoalfabética de maneira geral, a criptanálise é trivial, pois apesar de o número de chaves ser o número de permutações de alfabeto utilizado (no caso do exemplo acima $23! > 2 \times 10^{22}$), a distribuição de freqüência da mensagem original é mantida. Em outras palavras, na linguagem corrente, nem todas as letras aparecem com a mesma freqüência. Por exemplo, na língua portuguesa as letras que mais aparecem são E, A, O, S, I, R, N e T (essas letras formam 69,7% do texto) e a que menos aparece é o Z. Quando criptografamos um texto por substituição monoalfabética, mantem-se a mesma freqüência e isso pode (e deve) ser usado

na criptoanálise: se J é a letra que mais aparece no criptograma, provavelmente J substitui E, e assim por diante.

O anexo B mostra várias estatísticas das letras do alfabeto na língua portuguesa como, por exemplo, distribuição de freqüência de cada letra.

IV.1.2 - O Método de Substituição Polialfabética

Surgida durante a Renascença, a substituição polialfabética foi considerada inquebrável durante muito tempo [1]. Nesse caso, cada letra do texto original é substituída por um alfabeto diferente. Geralmente a chave consiste de N alfabetos o que fará com que a enésima-primeira letra do texto original seja substituída pelo mesmo alfabeto utilizado na primeira. A substituição polialfabética mais polular é a de Vigenère, que consiste numa tabela com o alfabeto deslocado de 0 até N-1 posições (sendo N o número de letras do alfabeto utilizado) e de uma palavra-chave que indicará quais as linhas da tabela serão aplicadas ao texto claro. Sendo a chave: ONTEM , o texto claro: PARABENS ANIVERSARIO PROXIMO DIA DOZE e a tabela de Vigenère da figura IV.3, o criptograma ficaria: ENMENSBN EAXJZVFOFDS CGCRNZD QDE QDMZ.

chave	O N T E M O N T E M O N T E M O N T E M O N T...
texto claro	P A R A B E N S A N I V E R S A R I O P R O X
criptograma	E N M E N S B N E A X J Z V F O F D S C G C R

FIGURA IV.2: Exemplo de Substituição Polialfabética

Chave v	Texto Claro: A B C D E F G H I J L M N O P Q R S T U V X Z
A	A B C D E F G H I J L M N O P Q R S T U V X Z
B	B C D E F G H I J L M N O P Q R S T U V X Z A
C	C D E F G H I J L M N O P Q R S T U V X Z A B
D	D E F G H I J L M N O P Q R S T U V X Z A B C
> E	E F G H I J L M N O P Q R S T U V X Z A B C D
F	F G H I J L M N O P Q R S T U V X Z A B C D E
G	G H I J L M N O P Q R S T U V X Z A B C D E F
H	H I J L M N O P Q R S T U V X Z A B C D E F G
I	I J L M N O P Q R S T U V X Z A B C D E F G H
J	J L M N O P Q R S T U V X Z A B C D E F G H I
L	L M N O P Q R S T U V X Z A B C D E F G H I J
> M	M N O P Q R S T U V X Z A B C D E F G H I J L
> N	N O P Q R S T U V X Z A B C D E F G H I J L M
> O	O P Q R S T U V X Z A B C D E F G H I J L M N
P	P Q R S T U V X Z A B C D E F G H I J L M N O
Q	Q R S T U V X Z A B C D E F G H I J L M N O P
R	R S T U V X Z A B C D E F G H I J L M N O P Q
S	S T U V X Z A B C D E F G H I J L M N O P Q R
> T	T U V X Z A B C D E F G H I J L M N O P Q R S
U	U V X Z A B C D E F G H I J L M N O P Q R S T
V	V X Z A B C D E F G H I J L M N O P Q R S T U
X	X Z A B C D E F G H I J L M N O P Q R S T U V
Z	Z A B C D E F G H I J L M N O P Q R S T U V X

FIGURA IV.3: Tabela de Vigenère

O número de chaves numa cifra de Vigenère de período N é 23^N, ou seja, muito menos do que o total de substituições

polialfabéticas de mesmo período. No entanto, se N tende a infinito ou se é do mesmo tamanho da mensagem, temos uma "chave sem repetição", também conhecida como cifra de Vernam que é comprovadamente inquebrável desde que nunca se repita o mesmo trecho de chave para duas mensagens [4]. O grande problema dessa cifra é a dificuldade de atualização da chave, que varia tanto quanto a própria mensagem, e só vale a pena para aplicações extremamente delicadas como a "hot line" entre Moscou e Washington [23].

IV.1.3 - O Método de Substituição Poligráfica

Nesse caso, em vez da substituição ser 1 por 1 (um caráter de cada vez) é N por N (N caracteres de cada vez). No caso mais simples, N=2 (substituição digráfica). O digrafo mais famoso é o do cientista inglês Playfair, em que o alfabeto é misturado num quadrado 5X5 (uma letra é omitida, geralmente o J):

D	B	M	W	I
C	O	X	G	E
Q	Y	R	F	S
Z	A	K	T	P
L	U	H	M	V

FIGURA IV.4: Exemplo de um Digrafo do Playfair [24]

Sendo as regras para criptografar:

1) Se P1 e P2 são 2 vértices diametralmente opostos de um retângulo, então C1 e C2 são as outras bordas com C1 na mesma linha que P1. Por exemplo, RE é criptografado como SX.

2) Se P1 e P2 estão na mesma linha, C1 e C2 são as letras à direita de P1 e P2 (A primeira coluna é considerada à direita da última). Exemplo: GE é criptografado como EC.

3) Se P1 e P2 estão na mesma coluna, C1 e C2 estão logo abaixo de P1 e P2. (A coluna do topo é considerada como logo abaixo da última.) Exemplo: IS é criptografado como EP.

4) Não existe letra dobrada. Caso alguma ocorra, uma letra nula (X por exemplo) é inserida para eliminar o problema.

Para decriptografar é só inverter as regras acima.

Nesta categoria também se encontram os sistemas algébricos, onde cada "polígrafo" do texto original é multiplicado por uma matriz (a chave). Para decifrar é só multiplicar o trecho do criptograma pelo inverso da matriz usada para cifrar. Esses sistemas são conhecidos como Cifra

de Hill [25].

Assim, sendo d=3 o número de letras a ser transformado de cada vez, $M = m_1m_2m_3$ a mensagem original e $\underline{K}$ a matriz 3X3 usada como chave, o criptograma será $C = c_1c_2c_3$ onde:

$$c_i = \Sigma_{i=1,3}\ k_{ij} \cdot m_i \pmod N$$

Sendo N o número de letras do alfabeto utilizado.

Por exemplo: vamos criptografar o digrama AB, onde as letras são substituídas por números (A = 0, B = 1 etc.) usando como chave a matriz $K = \begin{vmatrix} 3 & 2 \\ 3 & 5 \end{vmatrix}$

$$c_1 = (0 \text{ X } 3 + 1 \text{ X } 3) \bmod 23 = 3$$

$$c_2 = (0 \text{ X } 2 + 1 \text{ X } 5) \bmod 23 = 5$$

Logo o criptograma correspondente a AB será DF.

Para decifrar é só repetir a operação usando a matriz inversa de K, que nesse caso é $K^{-1} = \begin{vmatrix} 21 & 10 \\ 15 & 8 \end{vmatrix}$, pois:

$$K \cdot K^{-1} \bmod 23 = \begin{vmatrix} 1 & 0 \\ 0 & 1 \end{vmatrix}, \text{ que é a identidade.}$$

IV.1.4 - Transposição

No ciframento por transposição, os caracteres do texto original são embaralhados de acordo com algum padrão pre-estabelecido. Este padrão será a chave. Por exemplo, vamos criptografar a palavra INCONSTITUCIONAL num padrão zig-zag de profundidade 2:

```
I   N   T   O
 N O S I U I N L
  C   T   C   A
```

Tomando-se uma linha de cada vez o criptograma fica: INTONOSIUINLCTCA. Poderíamos também ter escrito a palavra num quadrado e a seguir tomado uma coluna de cada vez:

```
  1 2 3 4 5
 -----------
| I N C O N |
| S T I T U |
| C I O N A |
| L         |
 -----------
```

Nesse caso, se a chave for as colunas 3-5-2-1-4, o criptograma será: CIONUANTIISCLOTN.

IV.1.5 - Códigos

Um livro de códigos é usado como se fosse um dicionário, que é consultado quando se deseja criptografar ou decriptografar algum texto. Veja o exemplo da figura IV.5.

Texto claro:Bom dia. Tudo bem?

Criptograma: 20 93

Trecho do livro de códigos usado:

....	....
Boa noite. - 45	20 - Bom dia.
Boa tarde. - 60	45 - Boa noite.
Bom dia. - 20	60 - Boa tarde.
....	....
Tudo bem? - 93	93 - Tudo bem?
....	....

FIGURA IV.5: Exemplo de codificação.

Uma das vantagens da utilização de códigos é a compressão de dados, que além de diminuir a quantidade de "bits" transmitidos (custo de transmissão), diminui a redundância o que dificulta o trabalho do criptoanalista. O maior problema, no entanto, é a distribuição da chave, nesse caso o livro de códigos.

IV.1.6 - Máquinas Rotor

Rotores são máquinas eletro-mecânicas que implementam nada mais nada menos que substituições polialfabéticas.

Um rotor é um disco com todas as letras do alfabeto utilizado em sua volta. O disco é feito de material isolante, com um contato elétrico, correspondente a cada letra, de cada lado do disco. Um condutor interno ligará um contato correspondente a uma letra do lado esquerdo a outro contato correspondente a outra letra do lado direito. Ou seja, um rotor é uma substituição monoalfabética. Na figura IV.6 temos um esquema de um rotor.

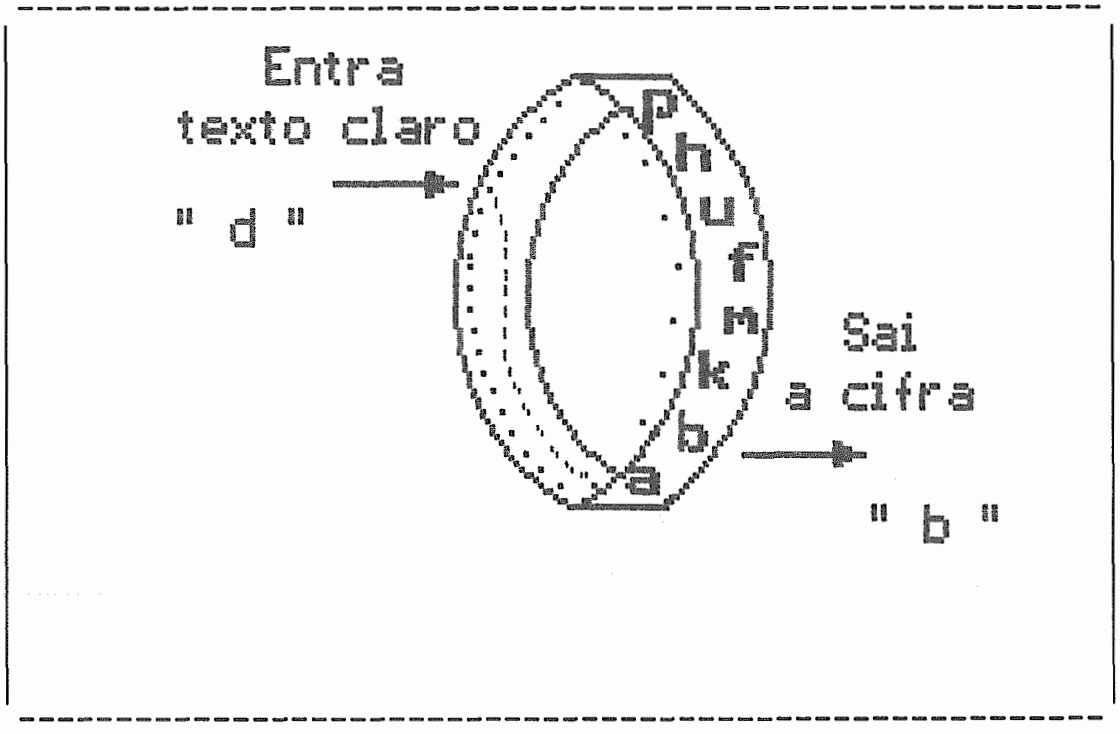


FIGURA IV.6: Um sinal passando num rotor [3]

Já a máquina rotor será uma seqüência de vários rotores, cada um com uma substituição monoalfabética própria. Como o nome já diz, um rotor roda. Assim, após a substituição de cada letra, os diversos rotores podem ser arrumados de outra maneira, o que acarretará numa substituição através de outro alfabeto e assim por diante. Temos, então, uma substituição polialfabética (ver item IV.1.2). A figura IV.7 nos mostra uma seqüência de rotores.

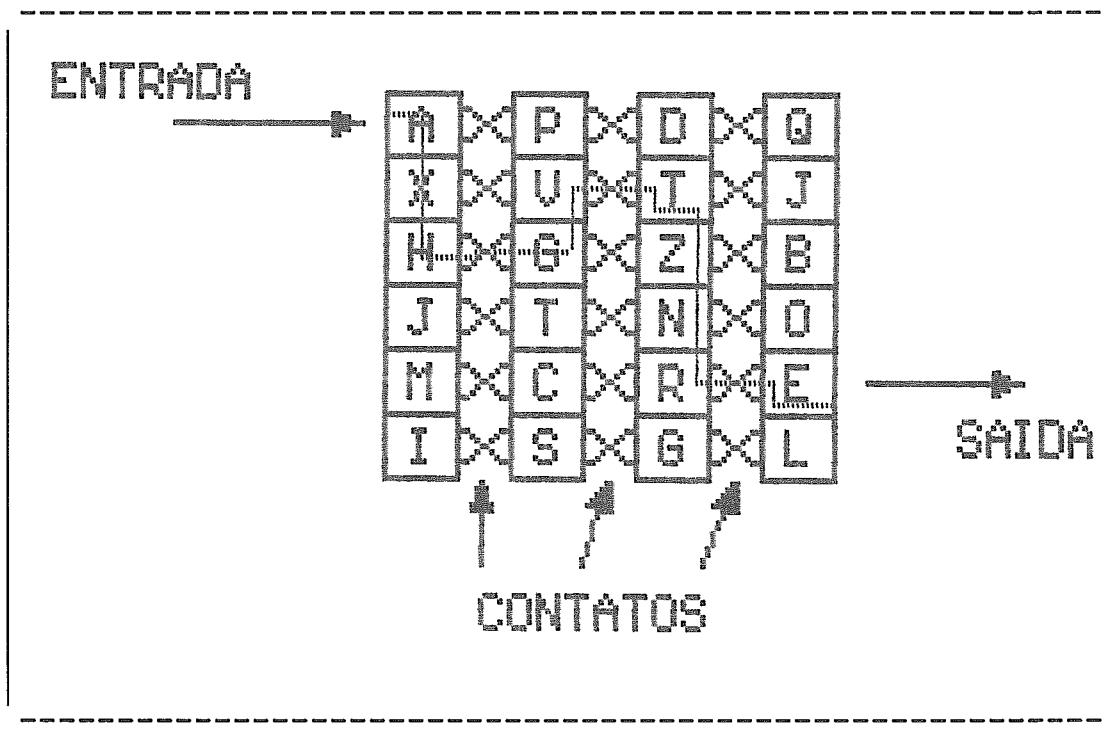


FIGURA IV.7: Sequência de rotores [3]

Sendo T uma letra do texto original a ser cifrado e S1, S2 e S3 as substituições monoalfabéticas de três rotores, o caráter C do texto cifrado correspondente será:

$C = S3 (S2 (S1 (T)))$

O movimento dos rotores devem apresentar as seguintes características [3]:

1) O período deve ser longo.(Quanto maior o período, mais nos aproximaremos da "chave sem repetição")

2) Cada mudança de estado deve ser grande, ou seja, todos os rotores deveriam mudar em relação uns aos outros, de maneira que a diferença de um caráter para o seguinte não seja de apenas uma transformação.

São exemplos de máquina rotor a Enigma, usada pelos alemães durante a Segunda Guerra e a Máquina de Hagelin.

A Enigma é um conjunto de 4 rotores acrescida de um "rotor refletor", um rotor com contatos somente de um lado, ligados aos pares. A substituição promovida pelo rotor refletor R é uma involução, ou seja, R^2 é a identidade, o que significa que a Enigma aplica a cada passo uma substituição que é uma involução. Assim cifrar e decifrar são a mesma operação [26]. No caso da Enigma, sendo T o caráter do texto claro, o caráter correspondente C do texto cifrado será:

$$C = S1^{-1} (S2^{-1} (S3^{-1} (S4^{-1} (R(S4(S3(S2(S1(T)))))))))$$

Constituem a chave da Enigma as quatro permutações S1, S2, S3 e S4, o rotor refletor R e as posições iniciais dos rotores.

Os rotores da Enigma variavam como um odômetro. Isso garante o maior período possível, mas a segunda propriedade não é satisfeita de maneira nenhuma, uma vez que todos os

rotores permanecerão na mesma posição enquanto o último avança uma posição de cada vez até dar a volta completa, quando então o penúltimo rotor avançará uma posição e assim sucessivamente. Ou seja, cada mudança é a menor possível.

Já a Máquina de Hagelin é implementada de maneira ligeiramente diferente. Ela é composta de seis discos, cada um deles com um número diferente de dentes. Depois que um caráter é convertido, cada disco se desloca de uma posição. Por exemplo, no modelo M-209, muito usado na Segunda Guerra Mundial e um dos poucos sistemas cuja descrição completa é de domínio público [3], cada disco possui respectivamente: 17, 19, 21, 23, 25 e 26 posições. Como esses números são primos entre si, a chave só se repetirá depois de 26 X 25 X 23 X 21 X 19 X 17 caracteres (aproximadamente 101 milhões).

Cada disco tem associado a cada uma de suas posições um pino, que pode estar setado ou não correspondendo a 0 ou 1. A série de 6 zeros ou uns corresponderá ao próximo caráter da chave, que será subtraída módulo 26 (que é o número de caracteres do alfabeto usado) ao próximo caráter do texto a ser cifrado, de maneira muito parecida com a substituição de Vigenère (aqui o caráter é subtraído ao invés de ser somado).

IV.2 - Métodos de Criptografia Computacional

A seguir apresentaremos algoritmos desenvolvidos para uso em computador. Não que os anteriores não possam ser implementados em computadores, mas os que se seguem foram feitos baseando-se em peculiaridades dos nossos amigos eletrônicos como rapidez na manipulação de grandes quantidades de dados e a intratabilidade de alguns problemas (até o presente momento, pelo menos).

Os algoritmos da criptografia computacional podem ser divididos em simétricos e assimétricos. Algoritmos simétricos são aqueles em que a mesma chave é usada para encriptar e decriptar e algoritmos assimétricos são aqueles que requerem duas chaves, uma para cada operação. Algoritmos simétricos são também chamados "algoritmos de chave secreta" e os assimétricos são conhecidos como "algoritmos de chave pública". A seguir analisaremos alguns algoritmos de cada categoria.

IV.2.1 - Métodos de Chave Secreta

Os métodos descritos a seguir são, como todos os vistos anteriormente, simétricos. Isso significa que a mesma chave que é usada para cifrar é usada para decifrar ...

IV.2.1.1 - Registros de Deslocamento

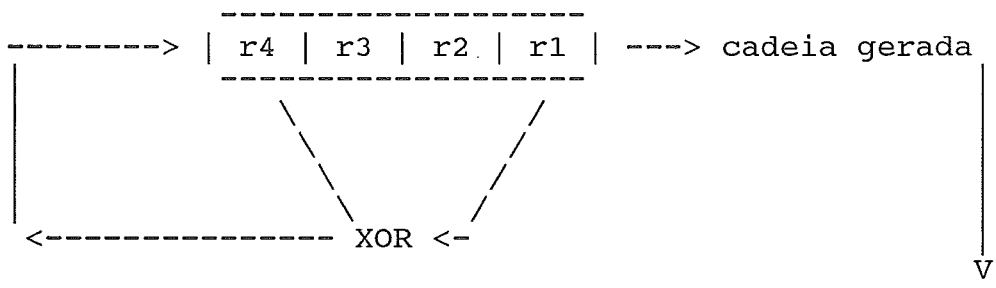
Esse tipo de sistema foi desenvolvido com o intuito de se gerar longas seqüências de aparência aleatória a partir de uma chave pequena. Isso pode ser conseguido com o uso de registros de deslocamento [3], tradução para Linear Feedback Shift Register (LFSR).

A figura IV.8 mostra um LFSR de 4 bits. O primeiro e o quarto bits são adicionados módulo 2 (operação de ou-exclusivo) para produzir a próxima entrada do registro cujo conteúdo é deslocado para a direita. O bit que "sai" do registro é usado para criptografar o próximo bit do texto claro também através de um ou-exclusivo. Isto fará com que as operações de cifrar e decifrar sejam inversas, usando a mesma chave.

A "chave", nesse caso, é a configuração inicial do registro e a operação escolhida. O tamanho da cadeia pseudo-aleatória gerada dependerá dessa escolha inicial. No exemplo da figura IV.8, temos a seqüência máxima possível (Maximal Length Shift Register Sequence - MLSRS). De acordo com DIFFIE e HELLMAN [3], para qualquer inteiro m existe uma

MLSRS gerada a partir de um registro de m bits que não se repetirá em $2^m - 1$ bits.

Apesar de esse sistema procurar imitar a cifra de Vernam, comprovadamente segura, ele é comprovadamente inseguro e pode ser quebrado em alguns segundos num minicomputador [3]. Se a operação for uma só (como a do exemplo dado), tudo se resumirá a resolver um sistema linear de m equações e m incógnitas (dados os m primeiros bits do texto).



Sendo a configuração inicial 0001 teremos:

0	0	0	1
1	0	0	0
1	1	0	0
1	1	1	0
1	1	1	1
0	1	1	1
1	0	1	1
1	1	0	1
0	1	1	0
0	0	1	1
1	0	0	1
0	1	0	0
0	0	1	0

FIGURA IV.8: Esquema de um LFSR e a seqüência gerada [25]

No entanto, alterações podem ser feitas no sentido de aumentar a segurança dos LFSR, como aplicar uma lógica não

linear à saída do LFSR antes de utilizá-lo na geração do criptograma, como esquematizado na figura IV.9. Alguns exemplos de funções não lineares podem ser vistos em [27].

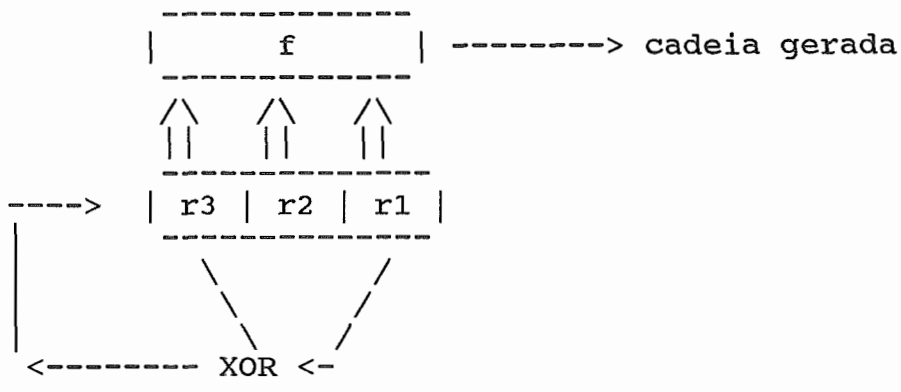


FIGURA IV.9: LFSR mais lógica não linear (f). [3]

IV.2.1.2 - O Padrão "Data Encryption Standard (DES)"

O sistema criptográfico adotado em 1977 pelo NBS (National Bureau of Standards) americano [25], conhecido como DES (Data Encryption Standard), baseia-se no LUCIFER, desenvolvido na IBM por Horst Feistel [26].

O DES cifra blocos de 64 bits de texto claro usando uma chave de 56 bits. O processo de cifrar e decifrar é o mesmo e consiste numa série de transposições e substituições (chamam-se "algoritmos produto" os algoritmos que misturam transposição e substituição).

O algoritmo funciona como mostrado na figura IV.10. O bloco T, da mensagem original, passa primeiro por uma permutação inicial IP. Em seguida são executadas 16 interações da função F para, finalmente, ser novamente permutado dessa vez pela inversa de IP, IP^{-1}, resultando no criptograma C.

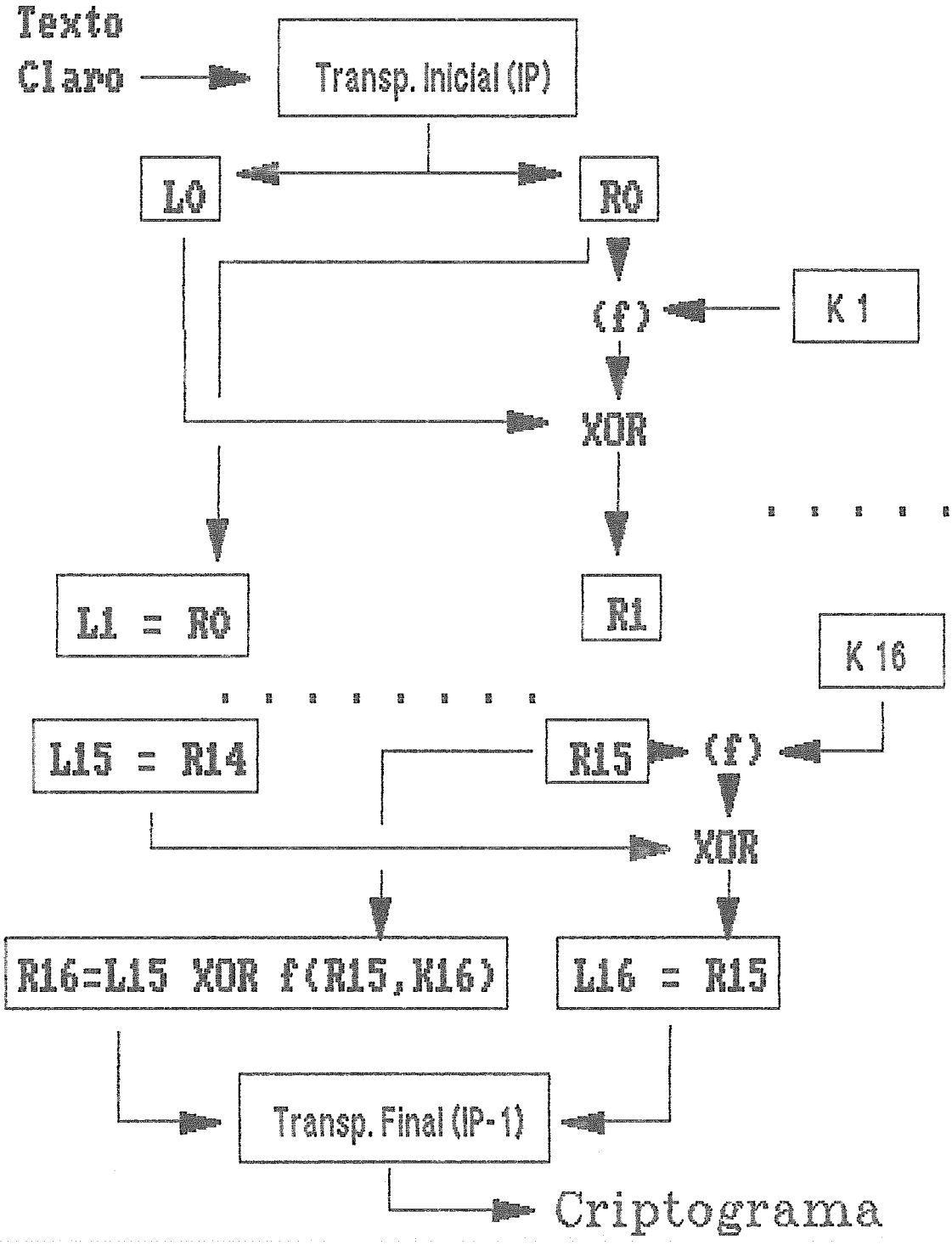


FIGURA IV.10: O DES [25]

Na figura IV.11 temos as permutações IP e IP^{-1}. De acordo com a tabela para IP, essa transformará $T=t_1t_2...t_{64}$

em $T_0=t_{58}t_{50}...t_7$. Tanto essas tabelas quanto as que serão vistas mais tarde são fixas.

Entre as duas transposições, são executadas 16 interações da função F que mistura substituição e transposição. Sendo T_i o resultado da i-ésima interação e R_i e L_i as metades direita e esquerda de T_i respectivamente ($T_i=L_iR_i$) onde

$$L_i = t_1 \ldots t_{32}$$

$$R_i = t_{33} \ldots t_{64}$$

Então:

$$L_i = R_{i-1}$$

$$R_i = L_{i-1} \text{ XOR } F(R_{i-1}, K_i)$$

sendo K_i uma chave de 48 bits descrita mais adiante. Um esquema da função F é descrito na figura IV.12. Primeiro R_{i-1} é expandido para 48 bits ($E(R_{i-1})$) usando a tabela E mostrada na figura IV.13. Esta tabela é usada da mesma maneira que as permutações inicial e final, só que alguns bits são repetidos.

É feito então um OU-exclusivo de $E(R_{i-1})$ com K_i e o resultado quebrado em blocos de 6 bits $B_1,...,B_8$. Cada bloco B_j é então usado como entrada para uma substituição (S_j) que retorna um bloco de 4 bits: $S_j(B_j)$. Esses blocos são concatenados e passam pela transposição P, mostrada na figura IV.14.

Permutação inicial IP							
58	50	42	34	26	18	10	2
60	52	44	36	28	20	12	4
62	54	46	38	30	22	14	6
64	56	48	40	32	24	16	8
57	49	41	33	25	17	9	1
59	51	43	35	27	19	11	3
61	53	45	37	29	21	13	5
63	55	47	39	31	23	15	7

Permutação final IP^{-1}							
40	8	48	16	56	24	64	32
39	7	47	15	55	23	63	31
38	6	46	14	54	22	62	30
37	5	45	13	53	21	61	29
36	4	44	12	52	20	60	28
35	3	43	11	51	19	59	27
34	2	42	10	50	18	58	26
33	1	41	9	49	17	57	25

FIGURA IV.11: Permutações IP e IP^{-1} [25]

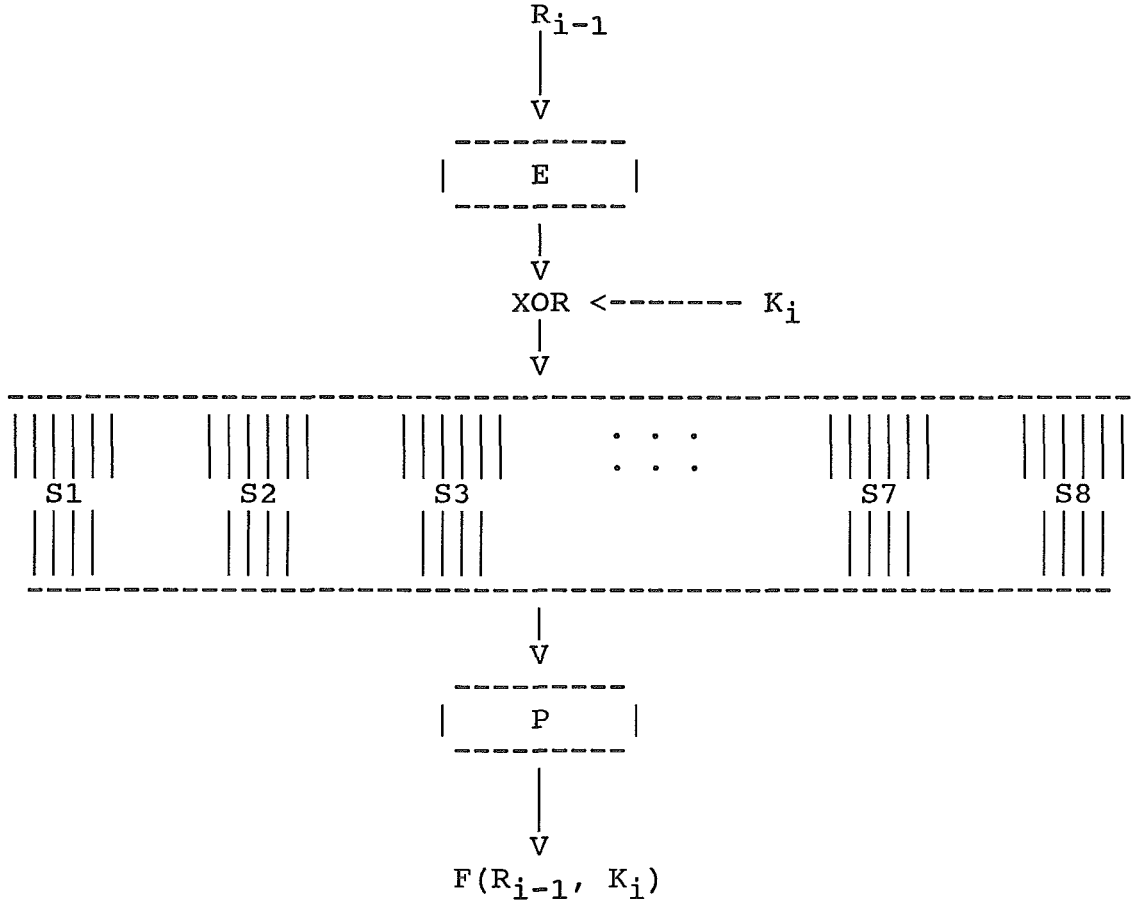


FIGURA IV.12: Cálculo da função $F(R_{i-1},K_i)$ [25]

32	1	2	3	4	5
4	5	6	7	8	9
8	9	10	11	12	13
12	13	14	15	16	17
16	17	18	19	20	21
20	21	22	23	24	25
24	25	26	27	28	29
28	29	30	31	32	1

FIGURA IV.13: Tabela E de seleção de bits. [25]

16	7	20	21
29	12	28	17
1	15	23	26
5	18	31	10
2	8	24	14
32	27	3	9
19	13	30	6
22	11	4	25

FIGURA IV.14: Transposição P. [25]

Cada substituição S_j troca um bloco de 6 bits por um de 4 de acordo com a tabela da figura IV.15 da seguinte maneira: o inteiro correspondente a b_1b_6 seleciona uma linha e o correspondente a $b_2b_3b_4b_5$ seleciona uma coluna. O valor retornado será a representação em 4 bits do valor constante dessa linha e coluna. Por exemplo, se B_1 = 010011, então S_1 retornará o valor na linha 1 e coluna 9, ou seja, 6, que é representado como 0110.

		Coluna														
	Linha	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
	0	14	4	13	1	2	15	11	8	3	10	6	12	5	9	0	7
S1	1	0	15	7	4	14	2	13	1	10	6	12	11	9	5	3	8
	2	4	1	14	8	13	6	2	11	15	12	9	7	3	10	5	0
	3	15	12	8	2	4	9	1	7	5	11	3	14	10	0	6	13
	0	15	1	8	14	6	11	3	4	9	7	2	13	12	0	5	10
S2	1	3	13	4	7	15	2	8	14	12	0	1	10	6	9	11	5
	2	0	14	7	11	10	4	13	1	5	8	12	6	9	3	2	15
	3	13	8	10	1	3	15	4	2	11	6	7	12	0	5	14	9
	0	10	0	9	14	6	3	15	5	1	13	12	7	11	4	2	8
S3	1	13	7	0	9	3	4	6	10	2	8	5	14	12	11	15	1
	2	13	6	4	9	8	15	3	0	11	1	2	12	5	10	14	7
	3	1	10	13	0	6	9	8	7	4	15	14	3	11	5	2	12
	0	7	13	14	3	0	6	9	10	1	2	8	5	11	12	4	15
S4	1	13	8	11	5	6	15	0	3	4	7	2	12	1	10	14	9
	2	10	6	9	0	12	11	7	13	15	1	3	14	5	2	8	4
	3	3	15	0	6	10	1	13	8	9	4	5	11	12	7	2	14
	0	2	12	4	1	7	10	11	6	8	5	3	15	13	0	14	9
S5	1	14	11	2	12	4	7	13	1	5	0	15	10	3	9	8	6
	2	4	2	1	11	10	13	7	8	15	9	12	5	6	3	0	14
	3	11	8	12	7	1	14	2	13	6	15	0	9	10	4	5	3
	0	12	1	10	15	9	2	6	8	0	13	3	4	14	7	5	11
S6	1	10	15	4	2	7	12	9	5	6	1	13	14	0	11	3	8
	2	9	14	15	5	2	8	12	3	7	0	4	10	1	13	11	6
	3	4	3	2	12	9	5	15	10	11	14	1	7	6	0	8	13
	0	4	11	2	14	15	0	8	13	3	12	9	7	5	10	6	1
S7	1	13	0	11	7	4	9	1	10	14	3	5	12	2	15	8	6
	2	1	4	11	13	12	3	7	14	10	15	6	8	0	5	9	2
	3	6	11	13	8	1	4	10	7	9	5	0	15	14	2	3	12
	0	13	2	8	4	6	15	11	1	10	9	3	14	5	0	12	7
S8	1	1	15	13	8	10	3	7	4	12	5	6	11	0	14	9	2
	2	7	11	4	1	9	12	14	2	0	6	10	13	15	3	5	8
	3	2	1	14	7	4	10	8	13	15	12	9	0	3	5	6	11

FIGURA IV.15: Substituições S [25]

As chaves K_j são derivadas da chave inicial K, de acordo com o esquema da figura IV.16. As permutações P1 e P2 das chaves são mostradas na figura IV.17. A permutação

P1 descarta os 8 bits de paridade e o resultado PC-1(K) é dividido em duas metades C e D de 28 bits cada. Essas metades são deslocadas para a esquerda para produzir as diversas chaves K_i.

Na decifração, os K_i são aplicados na ordem inversa, isto é, de K_{16} até K_1 [23].

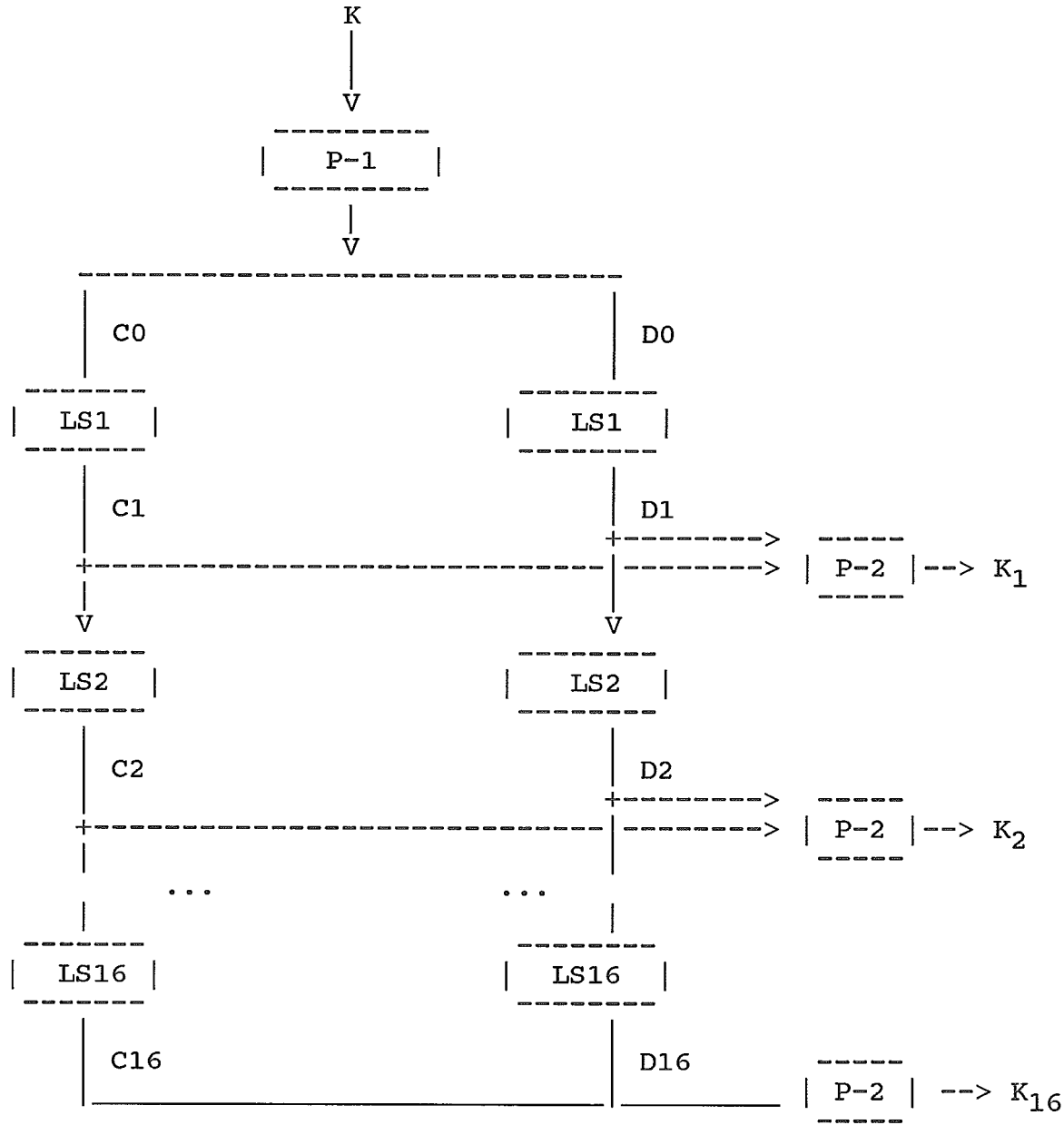


FIGURA IV.16: Cálculo das chaves. [25]

Transposição P-1

57	49	41	33	25	17	9
1	58	50	42	34	26	18
10	2	59	51	43	35	27
19	11	3	60	52	44	36
63	55	47	39	31	23	15
7	62	54	46	38	30	22
14	6	61	53	45	37	29
21	13	5	28	20	12	4

Transposição P-2

14	17	11	24	1	5
3	28	15	6	21	10
23	19	12	4	26	8
16	7	27	20	13	2
41	52	31	37	47	55
30	40	51	45	33	48
44	49	39	56	34	53
46	42	50	36	29	32

FIGURA IV.17: Tabelas das transposições P-1 e P-2 [L11]

O DES tem sido implementado tanto em hardware quanto em software.

Um dos pontos fortes do DES está no fato dele implementar transformações não lineares [3]. Essas transformações são conseguidas através das caixas S, provocando o efeito que Shannon chamou de confusão [23]. Confusão significa tornar a relação entre as diversas variáveis a mais complexa possível, de forma a maximizar o tempo requerido para a criptanálise [23].

As polêmicas sobre o DES pouco antes dele ser adotado, descritas no capítulo II pairavam principalmente sobre o tamanho da chave e as caixas S. Os críticos alegavam que a NSA aprovou chaves que pudessem ser quebradas por ela e por mais ninguém (daí a diminuição da chave do Lucifer de 128 bits [23] para a chave do DES aprovado, de 56 bits), além de não permitir a liberação do projeto das caixas S [23].

IV.2.1.3 - Gerenciamento de Chaves

Como se observou até aqui, a chave é o ponto crítico de todos os métodos analisados. Portanto, deve-se ter cuidados especiais na sua geração, guarda e distribuição. Os cuidados requeridos na geração das chaves são intrínsecos de cada algoritmo. Já os cuidados necessários à guarda da chave dependem principalmente de outros fatores da segurança do sistema (relacionados ao sistema operacional) como, por exemplo, um arquivo que possa ser acessado somente pelo usuário autorizado. O que mostraremos em seguida serão algumas maneiras de se distribuir as chaves entre os usuários de forma segura. Pois, como afirma SILVA FILHO [28], "não basta um bom algoritmo criptográfico para que a segurança seja efetiva mas que, na verdade, a segurança depende diretamente do sistema de gerenciamento de chaves criptográficas utilizado".

IV.2.1.3.1 - O Método do Quebra-Cabeças

MERKLE [29] propõe uma forma para as pessoas interessadas em manterem comunicação privativa informarem uma à outra qual a chave a ser usada através do canal normal de comunicação e sem que terceiros também tenham acesso a esta chave, ou, pelo menos, que estes últimos tenham algum trabalho para descobri-la.

Vamos supor, então, que X e Y desejem se comunicar e que Z quer descobrir quais as informações trocadas entre X e Y. O algoritmo criptográfico utilizado é de conhecimento dos três, mas a chave utilizada é de conhecimento apenas de X e Y e, preferencialmente, bem difícil de Z descobrir.

O autor é ainda mais severo em suas restrições: ele assume que o espião Z tem perfeito conhecimento de tudo que é enviado pelo canal e afirma que, mesmo assim é possível que X e Y mantenham uma conversação secreta.

X e Y podem apenas se comunicar através do canal em questão. Sendo assim, nenhuma combinação prévia foi feita. O caso, então, é fazer com que X e Y cheguem a um acordo a respeito da chave que pretendem usar e que o trabalho que Z deverá ter para descobrir esta chave seja muito maior que o trabalho de X e Y para chegarem ao acordo.

O primeiro conceito a ser definido é o de quebra-cabeça. Um criptograma, em princípio, não pode ser resolvido. Já um quebra-cabeça é passível de ser resolvido, desde que se dedique a ele o tempo e esforço necessários.

O que se faz então? Escolhe-se um bom algoritmo de

criptografia e uma chave não muito difícil, mas que só possa ser descoberta testando-se todas as combinações possíveis. Se a chave é, normalmente de 128 bits, usa-se uma de 30 bits; se a busca em 2^{128} chaves é inviável, o mesmo não ocorre com 2^{30} chaves. Neste caso, o trabalho pode ser tedioso, mas certamente pode ser realizado num tempo razoável. Uma outra coisa, se o criptograma não tiver nenhuma redundância, não é possível saber se a chave descoberta realmente é a correta. Assim, uma parte da mensagem deverá ser conhecida de modo que, se esta surgir ao se tentar uma chave nova, ficaremos sabendo que esta é a chave correta.

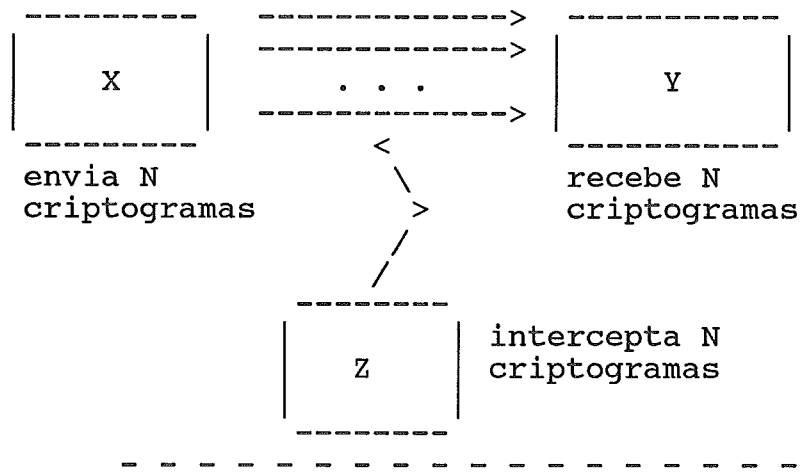
Vamos ver, então, como é que X e Y estabelecerão a sua chave secreta. Caberá a X a geração de várias chaves, uma das quais será escolhida por Y. X gerará N chaves (aqui chamadas "chaves aleatórias") e as enviará a Y. Só que essas chaves serão criptografadas com outras chaves, também geradas aleatóriamente por X, e aqui chamadas de "chaves do quebra-cabeças". A cada chave aleatória é associada uma identificação, que vai no mesmo criptograma. Os N criptogramas são enviados. Y escolherá um deles e tentará resolvê-lo. Quando conseguir, enviará a X a identificação da chave aleatória que deverá ser usada por ambos. A partir desse ponto, eles podem se comunicar através da chave escolhida. X teve o trabalho de gerar N criptogramas, Y teve o trabalho de resolver 1 criptograma.

E Z, como fica nessa história? Ele tem conhecimento dos N criptogramas enviados e da identificação que Y mandou de volta. Resta a ele tentar resolver cada um dos

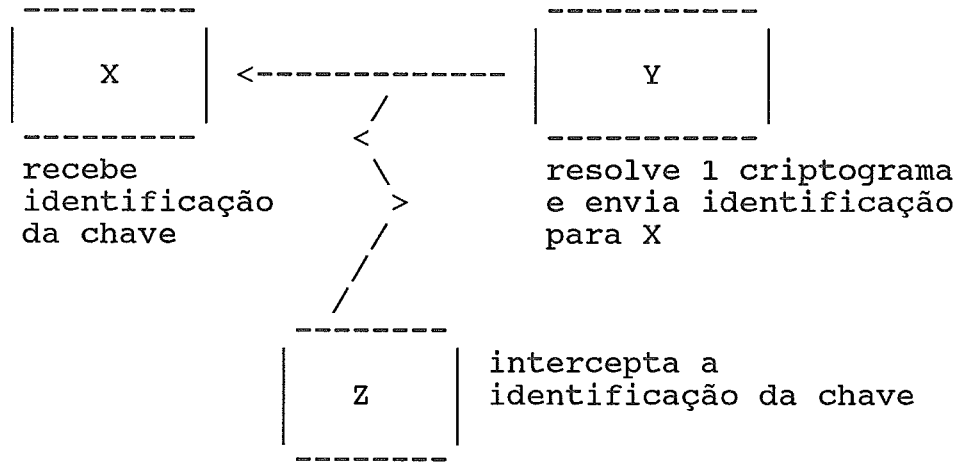
criptogramas até encontrar um que contenha a mesma identificação que a recebida por X, e só assim Z também terá conhecimento da chave. Para isso ele deverá resolver, em média, N/2 quebra-cabeças, o que, sem dúvida, é muito mais trabalhoso do que o trabalho que tiveram X e Y juntos ...

A figura IV.18 ilustra o método do quebra-cabeças.

Primeiro Passo:



Segundo Passo:



Daí em diante...

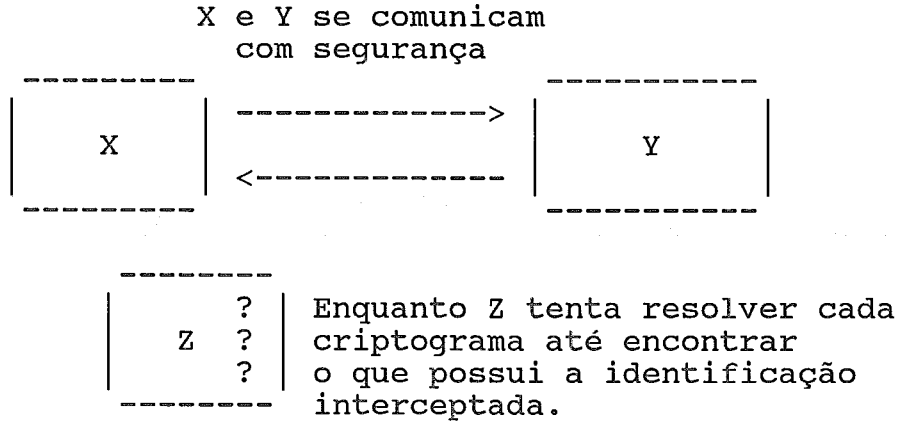


FIGURA IV.18: O Método do Quebra-Cabeças.

IV.2.1.3.2 - Sistema de Chaves Hierarquizadas

Este sistema consiste de 2 tipos de chaves com status diferente. A princípio, o algoritmo pode ser o mesmo para as duas categorias de chave. Uma classe será chamada de **chave mestre** e a outra de **chave de sessão.** A chave efetivamente usada na comunicação é a de sessão, sendo a chave mestre utilizada apenas para informar qual a chave de sessão em vigor. Ou seja, a chave de sessão muda a cada sessão enquanto a chave mestre se mantém constante.

A vantagem é que só será necessário encontrar um meio seguro para transmitir a chave mestre, que não muda com muita freqüência.

Nessa categoria se encaixam também os sistemas híbridos, em que o algoritmo da chave mestre é assimétrico (esses algoritmos serão vistos a seguir). Quer dizer, a chave mestre é pública enquanto as chaves de sessão continuam como as vistas até agora.

IV.2.1.3.3 - Sistema de Chaves Ocultas

Este sistema, descrito em [28], se aplica a algoritmos criptográficos simétricos e transitivos. Se A e B desejarem se comunicar com segurança, podem executar o seguinte procedimento:

1. A e B criam, aleatoriamente, chaves mestres ocultas M_a e M_b, respectivamente. Essas chaves são de conhecimento apenas de quem as criou.

2. Quem iniciar a comunicação, por exemplo A, gera uma chave de distribuição S_{ab}, e a transmite a B usando sua chave mestra: $CM_a(S_{ab})$.

3. B não poderá saber qual é a chave, mas pode criptografar o criptograma recebido e mandá-lo de volta a A: $CM_b\ (CM_a\ (S_{ab}))$.

4. A decifra a mensagem recebida e envia o resultado a B: $DM_a\ (CM_b\ (\ CM_a\ (S_{ab}))) \text{ ---> } CM_b\ (S_{ab})$

5. Agora B pode decifrar o criptograma e ele terá a chave de distribuição criada por A sem que esta tenha sido compreendida por qualquer outra pessoa que estivesse ouvindo a conversa.

IV.2.2 - Métodos de Chave Pública

Sem dúvida uma solução brilhante para o problema de distribuição de chaves num sistema criptográfico são os algoritmos que utilizam uma função assimétrica em que a chave usada para encriptar é distinta da usada para decriptar. Além disso uma "armadilha" é criada de maneira que uma chave não possa ser deduzida a partir da outra, ou pelo menos que a chave tornada pública não dê nenhuma dica sobre a chave secreta. Alguns desses algoritmos ainda permitem a assinatura digital, ou seja, que um usuário tenha certeza (e possa prová-lo perante um juiz) de que a mensagem recebida foi enviada pelo usuário X. Em outras palavras, a mensagem pode ser "assinada" pelo rementente.

Os sistemas de chave pública foram inicialmente propostos por Diffie e Hellman em 1976 [30].

Esses criptossistemas se baseiam em problemas conhecidos como NP-completos, que são problemas para os quais não se conhece nenhum algoritmo que "rode" em tempo polinomial e resolva o problema geral mas, uma vez dada uma provável solução é fácil verificar se ela é verdadeira ou não.

O RSA, primeiro sistema de chave pública, foi proposto em 1977 e o criptossistema baseado no Problema da Mochila em 1978. Ambos serão analisados a seguir.

IV.2.2.1 - Algoritmo de Rivest,Shamir e Adleman (RSA)

Este algoritmo [31] se baseia no fato de não existir até agora um algoritmo computacional eficiente que fatore um número que seja produto de dois números primos muito grandes (este problema pertence ao conjunto dos problemas NP-completos).

A mensagem é dividida em blocos de forma que cada bloco possa ser representado como um inteiro entre 0 e n-1.

A chave pública do RSA será o par ordenado de números inteiros (e,n) e a chave secreta será (d,n) de tal maneira que, sendo a mensagem representada como um inteiro entre 0 e n-1,esta seja criptografada assim:

$C \equiv E(M) \equiv M^e \pmod{n}$ e decriptografada assim:

$M \equiv E(M) \equiv C^d \pmod{n}$, onde M é a mensagem e C o criptograma.

Para chegar as chaves (e,n) e (d,n), a primeira coisa é calcular n, que será o produto de dois números primos muito grandes p e q:

$$n = p * q$$

O sucesso da chave reside nesses dois números. Para que um espião a descubra, ele deverá primeiro descobrir quem são p e q. Uma vez que esses números são muito grandes, será extremamente trabalhoso fatorar n (que será público). De acordo com DENNING [25], o melhor algoritmo conhecido é o de Schroeppel que, para fatorar n leva

exp(sqrt(ln(n)ln(ln(n)))) ciclos de máquina. Por exemplo, usando-se um computador que efetue uma multiplicação em 10^{-6} segundos, os tempos para fatorar n são dados na tabela da figura IV.19.

n	Tempo de fatoração via algoritmo de Schroepel
50	3,9 horas
75	104 dias
100	74 anos
200	$3,8X10^{7}$ séculos
300	$4,9X10^{13}$ séculos
500	$4,2X10^{23}$ séculos

FIGURA IV.19: Tempos necessários para fatorar n [32].

Voltando ao cálculo das chaves do RSA, a seguir calcula-se d, que deverá ser um inteiro, bem grande e não ter nenhum fator em comum com (p-1)*(q-1):

m.d.c.((p-1)*(q-1),d) = 1

Por último, e é calculado a partir de p,q e d e será o inverso multiplicativo de d módulo (p-1)*(q-1) de tal forma que:

$e * d \equiv 1 \pmod{(p-1)*(q-1)}$

Demonstra-se que as funções D e E são inversas, ou seja, que:

$$D(E(M)) \equiv (E(M))^d \equiv (M^e)^d \equiv M^{(e*d)} \pmod{n}$$

$$E(D(M)) \equiv (D(M))^e \equiv (M^d)^e \equiv M^{(e*d)} \pmod{n}$$

O cálculo de $M^e \pmod{n}$ requer, no máximo, 2*log(e) multiplicações e 2*log(e) divisões. Na figura IV.20 mostramos o algoritmo usado:

---

1. Seja $e_k e_{k-1} \ldots e_1 e_0$ a representação binária de e.
2. C := 1
3. FOR i := k TO 0 STEP -1 DO
   - 3.1. C := $C^2 \pmod{n}$
   - 3.2. IF e_i = 1 THEN C := C*M (mod n)
4. HALT. C é o criptograma de M.

FIGURA IV.20: RSA usado para criptografar M

Segundo os autores, este algoritmo não é o melhor mas "metade do melhor". Entretanto, uma vez que os procedimentos de criptografar e decriptografar são idênticos, a implementação é simples, e se tornará muito mais eficiente se for feita em hardware. MULLER-SCHLOER [33] descreve a implementação em hardware de um criptossistema

híbrido usando o DES e o RSA.

O RSA é, sem dúvida, o mais famoso de todos os algoritmos de chave pública e também bastante utilizado. Como já tivemos a oportunidade de observar no capítulo II, usa-se o RSA em aplicações bancárias, o que pode nos dar uma idéia do impacto causado pela fatoração de um produto de dois primos de 100 dígitos, conseguida por Lenstra e Manasse...

IV.2.2.2 - Algoritmo Baseado no Problema da Mochila

O método da mochila, exaustivamente analisado por OLIVEIRA [35] em sua tese de mestrado, foi desenvolvido por Merkle e Hellman (por isso é algumas vezes chamado de MH, as iniciais dos autores) e se baseia no problema da mochila, também conhecido como problema da soma dos subconjuntos.

Este problema tem a seguinte formulação básica:

" Dada uma mochila de tamanho S e um conjunto de n cilindros, todos de mesmo diâmetro da mochila e alturas a_1, a_2, ... a_n achar um subconjunto de cilindros que preencha completamente a mochila. Ou seja, achar um subconjunto cuja soma seja igual a S."

Se não for feita nenhuma restrição quanto aos valores a_i, o problema poderá não ter nenhuma solução ou várias. Além disso, o único algoritmo conhecido para se encontrar uma solução é o de busca exaustiva, ou seja, o problema da mochila pertence à classe dos NP-completos.

Dada uma mochila de tamanho $\underline{n}$, ou seja, o vetor mochila

contém n elementos, serão necessárias 2^n tentativas para determinar a solução pelo método de exaustão. Portanto, para valores grandes de n, torna-se quase impossível determinar a solução por esse processo.

No entanto, existem alguns casos particulares de mochila para os quais a solução do problema mochila associado pode ser facilmente determinada. É o caso de mochilas supercrescentes.

Mochila supercrescente é aquela em que cada elemento é maior que a soma de seus antecessores:

$$a_i > \Sigma_{J=1,I-1}\ a_j,\ I>=2$$

O fato de a seqüência ser supercrescente implica que:

1) a representação de S, quando possível, é única (Sendo $S = \Sigma_{I=1,n}\ a_i \cdot x_i$, onde a_i são os elementos da mochila supercrescente e x_i um vetor de zeros e uns, o que significa que S pode ser representado por no máximo uma configuração de a_is);

2) esta representação pode ser encontrada com, no máximo, n passos computacionais.

Uma mochila supercrescente fácil pode ser transformada numa mochila difícil. Para isso escolhe-se um par de inteiros w e M, primos entre si e $M > \Sigma_{I=1,n}\ a_i$.

Calcula-se, então, um vetor $b = (b_1, \ldots, b_n)$ da seguinte forma:

$$b_i = w.a_i \pmod M,\ 1 <= i <= n,\ \text{ou seja},\ 0 <= b_i < M$$

$a = (a_1, a_2, \ldots, a_n)$ é o vetor fácil;

$b = (b_1, b_2, \ldots, b_n)$ é o vetor "difícil".

Espera-se ser difícil resolver o problema usando-se $\underline{b}$, desde que w e M sejam desconhecidos.

Merkle e Hellman usam essas duas mochilas no seu sistema criptográfico. A mochila difícil é tornada pública e é usada como chave de criptografar e a mochila fácil, w e M são secretos e constituem a chave de decifração. O valor S, associado ao texto $x = (x_1, x_2, \ldots, x_n)$ será o criptograma enviado:

Texto: $x = (x_1 \ldots x_n)$ $x_i = 0$ ou 1

Criptograma: $S = \Sigma_{I=1,n}\ x_i.b_i$, onde $b = (b_1 \ldots b_n)$ é a chave pública.

Para decifrar S:

Criptograma transformado: $S' = w^S \pmod M$

$S' = (a_1x_1, \ldots, a_nx_n)$, onde $a = (a_1 \ldots a_n)$ é a mochila fácil e $x = (x_1 \ldots x_N)$ é o texto original.

Os autores propuseram iterações sucessivas do algoritmo acima como forma de fortalecer o método, mas DESMEDT e outros [35] mostram justamente como essas iterações podem ajudar a quebrar o sistema.

Como vantagens do método apresentado, OLIVEIRA [34] destaca o fato de este ser um método de chave pública, o que facilita a distribuição das chaves. Além disso, é um sistema de implementação fácil e que requer pouco esforço

computacional.

As desvantagens são a expansão dos dados (a relação entre o número de bits do criptograma e o do texto original é alta) e o tamanho da chave (normalmente maior que 10000 bits = 100 elementos X 100 bits por elemento da chave pública). Um outro problema é o de segurança. OLIVEIRA se dedicou ao estudo de várias formas de se quebrar o sistema da mochila, chegando à conclusão de que este apresenta várias falhas apesar de não ser ainda possível provar definitivamente que o sistema de chave pública tipo mochila não deva mais ser usado como sistema criptográfico.

IV.2.2.3 - Autenticação e Assinaturas

Além de garantir a privacidade das informações transmitidas, a criptografia também se propõe a garantir a autenticidade do remetente e do conteúdo da mensagem, isto é, o destinatário tem a garantia de que a mensagem não foi alterada por terceiros e de que ela foi realmente enviada por quem de direito. Isso é chamado autenticação do usuário e autenticação da mensagem.

Em algumas aplicações, no entanto, pode ser necessário provar a origem da mensagem. Nesse caso, será preciso "assinar" o "documento" transmitido. A essência da assinatura é que apenas uma pessoa pode produzi-la mas várias podem reconhecê-la.

Os algoritmos de chave secreta em si, não trazem a facilidade da assinatura, pois o destinatário, que também conhece a chave pode forjar uma mensagem cifrada e afirmar tê-la recebido...[3]

Vale a pena lembrar que o conceito de assinatura digital surgiu com os sistemas de chave pública, tendo sido inicialmente chamada de "autenticação one-way" por Diffie e Hellman (os pais dos sistemas de chave pública) [30].

A autenticação de mensagens já é parcialmente garantida pelo algoritmo criptográfico utilizado. Para que se evite que mensagens antigas sejam enviadas novamente (ataque da meia-noite), um "timestamp" (data e hora) pode ser criptografado junto com o texto. Dessa maneira o destinatário poderá verificar a "idade" da comunicação. Pode ocorrer também da mensagem ser alterada, nesse caso é necessário usar técnicas de ciframento com propagação de erro. Se alguns bits forem incluidos, alterados ou removidos do criptograma, o texto claro resultante ficará completamente deturpado. O destinatário então acusará o ocorrido para que as devidas providências sejam tomadas. Essas técnicas serão estudadas no item IV.3 (Taxonomia da Criptografia Computacional).

A autenticação de usuários está diretamente relacionada com o problema de gerenciamento de chaves, já tendo sido explorado no item IV.2.1.3. Outra forma mais específica de autenticação de usuário se faz na hora em que a conexão entre usuários é realizada, uma espécie de saudação. A saudação é um procedimento que assegura que a comunicação

entre dois nós genuínos foi estabelecida e que as mensagens que serão trocadas não são antigas, pré-gravadas [26]. Um exemplo de protocolo de saudação entre um terminal e o computador central é o seguinte [26]:

1 - O terminal envia sua identidade T ao sistema;

2 - O sistema gera uma chave aleatória de sessão CS e a envia cifrada pela chave mestre: CT(CS);

3 - O terminal recebe e decifra a mensagem obtendo CS. O terminal gera outro valor aleatório r, que independe de CT e CS, e envia ao sistema a mensagem CS(r);

4 - O sistema decifra CS(r) e envia a mensagem CS(r+1) ao terminal;

5 - O terminal decifra CS(r+1) e verifica que de fato o resultado é um a mais do que o aleatório r gerado.

O valor r e r+1 são incluídos para prevenir o "ataque da meia-noite".

Outra forma de autenticação de usuários em sistemas de computadores multi-usuário são as **senhas**. Segundo DIFFIE e HELLMAN [3], as senhas são de longe a aplicação de criptografia **mais difundida** em segurança de computadores. De maneira geral, as senhas são armazenadas nos computadores em arquivos criptografados por uma função "one-way", ou seja, uma função não inversível. Ainda assim, há a possibilidade de que o inimigo observe a transmissão da senha (ou do criptograma correspondente à senha) e a utilize mais tarde... LAMPORT, em [36], descreve mecanismos que garantem que a mesma senha nunca é usada duas vezes.

No caso de sistemas de chave pública, apenas o destinatário é capaz de traduzir corretamente a mensagem,

sendo assim garantida a autenticidade do destinatário. E para que ele possa ter garantias da origem da mensagem, esta deverá vir "assinada" da seguinte maneira:

Sendo um algoritmo de chave pública aplicado sobre uma sequência de bits, a lógiga não é alterada se a aplicamos a um texto claro ou não. Então, a sequência:

$C = (\text{Chave Pública}_B\ (\text{Chave Secreta}_A\ (\ M\ )\ )$

pode ser calculada e enviada de A para B. Essa mensagem segue secreta e assinada. Secreta porque só B pode traduzi-la corretamente:

$C_1 = (\text{Chave Secreta}_B\ (\text{Chave Pública}_B(\text{Chave Secreta}_A(M)))$

$C_1 = (\text{Chave Secreta}_A\ (M))$

C_1 é a mensagem assinada. B aplicará sobre ela a chave pública de A e terá a mensagem original:

$\text{Chave Pública}_A\ (\text{Chave Secreta}_A\ (M)) = M$

e guardará C_1 e M. Caso A mais tarde negue ter enviado M, B tem como prová-lo perante um juiz. Basta repetir a operação acima.

Caso não haja necessidade de segredo, mas apenas da assinatura, A enviará:

$C = (\text{Chave Secreta}_A\ (M))$

e qualquer usuário do sistema será capaz de traduzir C.

O único problema desse protocolo é quando a chave secreta de A é comprometida. Ou quando A, querendo negar uma

mensagem por ele enviada, afirma ter sido sua chave comprometida. A solução, nesse caso é a "contratação" de um "tabelião" que terá conhecimento de todas as chaves públicas em vigor e autenticará a assinatura por ocasião do recebimento desta [37]. O tabelião, também chamado de "servidor de autenticação" pode ser usado em sistemas que utilizem tanto chave pública quanto chave secreta. Numa rede, podem inclusive existir mais de um [38].

Encontra-se na literatura vários protocolos de assinatura digital baseados no acima descrito como em [39], [40], [41] e [42].

Dos sistemas de chave pública estudados, apenas o RSA se presta a assinar mensagens exatamente como descrito acima. O MH não serve porque não há como aplicar a chave pública ao criptograma para traduzi-lo (este é um problema NP-completo!).

IV.3 - Taxonomia da Criptografia Computacional

Estudaremos, a seguir, princípios em que se baseia a criptografia computacional e como se pode classificar os algoritmos de acordo com a transformação sofrida pela mensagem: ciframento bit a bit ou bloco a bloco; técnicas de sincronismo e propagação de erro.

IV.3.1 - Ciframento Bit a Bit ou Bloco a Bloco

Ciframento bloco a bloco é aquele em que se divide o texto claro em blocos, geralmente de vários caracteres ou bits e se criptografa cada bloco com a mesma chave **c**.

Sendo M a mensagem e M_1M_2 ... os diversos blocos em que M foi dividida, um ciframento bloco a bloco funciona assim:

$$E_c(M) = E_c(M_1)E_c(M_2) \ldots$$

Exemplos de algoritmos que usam ciframento bloco a bloco podem ser vistos na figura IV.21. A substituição monoalfabética é considerada nesta categoria apesar de só cifrar um caracter de cada vez porque a mesma chave é usada para cada caracter [25].

ALGORITMO	TAMANHO DO BLOCO
Substituição Monoalfabética	1 caracter
Transposição de Período d	d caracteres
Playfair	2 caracteres
Hill com matriz dXd	d caracteres
DES	64 bits
Mochila de tamanho n	n bits
RSA com chave (e,n) [$C = M^e \bmod n$]	$\log_e n$ bits

FIGURA IV.21 : Exemplos de ciframento bloco a bloco. [25]

No ciframento bit a bit a mensagem é dividida em bits ou caracteres e cada elemento M_i da mensagem é criptografado com o i-ésimo componente c_i da chave $c=c_1c_2$ ... , ou seja:

$$E_c(M) = E_{c1}(M_1)E_{c2}(M_2) \ldots$$

O ciframento bit a bit é considerado periódico se a chave se repete depois de d caracteres, sendo d um valor fixo. Caso contrário, o ciframento é aperiódico.

Na figura IV.22 temos exemplos de ciframento bit a bit.

ALGORITMO	PERÍODO
Vigenère de período d	d
Máquina Rotor com t rotores	26^t
Máquina de Hagelin com t rodas cada uma com p_i pinos	$p_1 X p_2 \cdots p_t$
Vernam	
Registros de deslocamento com reg. de n bits	2^n

FIGURA IV.22 : Exemplos de ciframento bit a bit. [25]

IV.3.2 - Ciframento Síncrono ou Encadeado

Cifras diversas podem também ser divididas em **síncronas** e **encadeadas.** O ciframento é **síncrono** se o próximo bit ou bloco do texto não depende dos que o precederam. Tudo o que será levado em conta é o fato de que ele é o i-ésimo caracter/bit/bloco, e como tal será cifrado. Nesse tipo de ciframento não há propagação de erro. Já nos cifamentos **encadeados** os caracteres anteriores do texto cifrado fazem parte da chave, de maneira que se houver algum erro em cifragens anteriores este se propagará.

Existem dois motivos básicos para se usar ciframento encadeado: redundância e autenticação da mensagem. O encadeamento camuflará mais ainda prováveis repetições de padrão do texto claro no texto cifrado, evitando que a criptoanálise se torne viável através da análise de frequência dos blocos do texto cifrado.

Se, num ataque ativo, alguns bits do criptograma forem

alterados, o destinatário receberá uma mensagem completamente deturpada, sabendo que algo de errado ocorreu.

Por outro lado, esta propriedade (propagação de erro) não é muito desejável no caso de ciframento de arquivos, pois a perda acidental de um bloco poderá tornar inacessíveis os blocos subsequentes. Além disso, deve-se evitá-la se o acesso ao arquivo não for sequencial...

A seguir daremos um exemplo de encadeamento por blocos com realimentação de texto original e cifrado [26]:

Dada a chave de ciframento (pública ou secreta) C e um valor inicial z (chamado "semente"), o i-ésimo bloco do texto cifrado y_i será:

$$y_i = C(x_i + u_i) \ ,$$

onde x_i é o i-ésimo bloco do texto original; + é a operação de ou-exclusivo bit a bit e u_i é dado por:

$$u_1 = z$$

$$u_i = x_{i-1} + y_{i-1} \quad (i = 2,3, \ldots).$$

Assim, o i-ésimo bloco y_i de texto cifrado depende da chave C, do bloco do texto claro x_i e dos blocos anteriores do texto original e cifrado x_{i-1} e y_{i-1}, respectivamente.

Para decifrar o criptograma usando a chave de decifrar D, observa-se que:

$$x_i = D(y_i) + u_i$$

O algoritmo acima também se aplica se o ciframento for bit a bit. Basta tomar y_i como sendo o i-ésimo bit do

criptograma, a chave $c = c_1c_2...$ e x_i o i-ésimo bit do texto claro.

IV.4 - Algumas Técnicas de Criptanálise

O "espião" pode se interpor entre o remetente e o destinatário de maneira ativa ou passiva. Ataque ativo é aquele em que o "inimigo" altera a mensagem, inclui novas mensagens ou as remove. Já no ataque passivo as mensagens são apenas interceptadas, tanto para conhecimento do seu conteúdo quanto para análise de trânsito (o inimigo quer saber a quantidade de informações transmitidas e a frequência).

Algumas formas de se prevenir o ataque ativo foram vistas no item IV.2.2.3 (Autenticação e Assinaturas).

Procuraremos mostrar aqui as "armas" de que dispõe o inimigo para chegar ao conteúdo de mensagens criptografadas por ele interceptadas. Além de partirmos do pressuposto que o espião conhece o algoritmo utilizado, o conhecimento inicial dele poderá ser:

a) apenas o texto cifrado (ataque unicamente por texto cifrado);

b) o texto cifrado e o respectivo texto claro (ataque com texto claro conhecido);

c) o texto cifrado de um texto claro escolhido por ele (ataque com texto claro escolhido).

Quanto mais dados tiver o criptanalista, mais forte

será o ataque. Portanto, criptossistemas que resistam a ataques mais fortes, por exemplo ataque com texto claro escolhido, serão mais seguros.

Um exemplo de (b) seriam os ataques em que o espião, tendo alguma idéia do assunto, pode assumir a existência de algumas palavras no texto. Se o mote é fornecimento de canhões, provavelmente a palavra "canhões" aparece no texto...

Um exemplo famoso de (c) (ataque com texto claro escolhido) ocorreu durante a Segunda Guerra. Os japoneses usavam uma mistura de cifras e códigos. Os americanos já tinham quebrado a cifra, e interceptaram uma mensagem em que os japoneses acertavam os detalhes para atacar M. Os americanos achavam que M era o código para Midway, mas não tinham certeza. Eles então mandaram uma mensagem via cabo marítmo para Midway pedindo que eles mandassem de volta uma mensagem não cifrada "comunicando" que haveria falta d'água em Midway devido a um problema inventado. A mensagem foi mandada e alguns dias mais tarde foi interceptada uma comunicação japonesa avisando Tokio da interceptação de uma mensagem americana que falava da previsão de falta d'água em M para os próximos dias ! [1]

Sem dúvida, cada tipo de criptossistema terá um tipo de criptanálise mais eficiente. Por exemplo, o estudo de frequência em substituição é eficiente na determinação da chave mas não o será em transposição, uma vez que para qualquer chave usada a frequência das letras do criptograma será a mesma. Mesmo o método de tentativa e erro pode ser eficiente em alguns casos. Suponha, por exemplo, que o

campo salário de um registro foi criptografado usando-se um sistema de chave pública. Uma vez que esse valor estará um intervalo de digamos, 100000 a 5000000, é perfeitamente viável ao criptanalista aplicar a chave de cifrar a todos os valores nesse intervalo, comparando o resultado com o original até descobrir o valor correto [25].

Mesmo casos particulares da cifra de Vernam podem ser criptanalisados. Se a chave também for um texto em linguagem corrente (nesse caso a cifra é conhecida como "running key"), o estudo de freqüência pode levar à solução. Esse tipo de ataque foi proposto por Friedman. Veja o exemplo da figura IV.23.

Nesse exemplo, 12 das 19 letras do criptograma foram gerados por pares de letras muito comuns (de alta freqüência). Dos 7 pares que sobraram, em 6 deles ou a letra da chave ou a do texto claro pertencem à categoria de alta freqüência e no sétimo par as duas letras são de média freqüência.

Friedman recomenda que se comece assumindo que todos os pares são de alta freqüência, diminuindo assim as possibilidades de cada letra do texto claro e da chave. Os chutes iniciais são relacionados a distribuições de digramas e trigramas (e palavras prováveis) para determinar os pares.

```
Texto Claro:  THETREASUREISBURIED....
Chave:        THESECONDCIPHERISAN....
Criptograma:  MOILVGOFXTMXZFLZAEQ....
```

Considerando as primeiras 3 letras do criptograma MOI, existem 26 possibilidades de emparceiramento. Examinando os possíveis pares, só 3 caem na classe de alta freqüência:

```
                        |       |               |
Texto claro:   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
Chave:         M L K J I H G F E D C B A Z Y X W V U T S R Q P O N
Texto cifrado: M M M M M M M M M M M M M M M M M M M M M M M M M M
```

Os pares de alta freqüência para as três letras são:

```
M --> E-I, I-E, T-T
O --> A-O, O-A, H-H
I --> A-I, I-A, E-E, R-R
```

Agora temos 3 X 3 X 4 = 36 possibilidades para os pares acima. A maioria é pouco provável como EAA e IOI. O trigrama THE, que ocorre simultameamente no texto e na chave tem grandes chances. O criptoanalista então o escolhe e prossegue tentando resolver os próximos pares.

FIGURA IV.23 : Ataque a cifra do tipo "running key" [25]

Shannon, na sua Teoria Matemática da Criptografia [4], procurou generalizar as técnicas de criptoanálise. Muitos conceitos desenvolvidos por ele são muito úteis até hoje como **redundância** ("a criptanálise só é possível devido à redundância"), **unicity distance** (o tamanho mínimo que um criptograma deve ter para que sua solução seja única para quem não conhece a chave) e **fator trabalho** (a quantidade de trabalho que o criptoanalista terá para resolver o criptograma).

Em princípio, tendo-se um criptograma maior que a "unicity distance", o método de tentativa e erro acabará

resolvendo o problema. No entanto, a menos de casos muito especiais, não é viável. Shannon [4], afirma que a criptoanálise realmente envolve grandes parcelas de tentativa e erro, só que de forma metódica. O que se faz é procurar dividir o universo de chaves em subconjuntos e testar determinadas características desses subconjuntos de maneira a determinar a qual deles pertence a chave correta. A seguir o subconjunto selecionado é novamente subdividido até que se chegue à chave ou a um conjunto muito pequeno que pode ser testado um a um. Na divisão dos conjuntos e nos testes são usados os mais diversos recursos como características matemáticas da cifra usada, estatísticas tanto do criptograma quanto da linguagem do texto claro, prováveis palavras do texto claro etc.

O trabalho de se desenvolver uma cifra é muito complexo pois é necessário pensar em todas as possibilidades de ataque existentes e nas que possam passar pela cabeça do inimigo!

CAPÍTULO V - METODOLOGIA PARA A CRIPTOANÁLISE DOS SISTEMAS ATUALMENTE CONSIDERADOS INQUEBRÁVEIS

Neste capítulo propomos uma metodologia para o desenvolvimento da criptoanálise baseada em conceitos de inteligência artificial.

A primeira parte faz um apanhado geral das técnicas estudadas nos capítulos anteriores. Na segunda parte apresentamos uma nova metodologia para criptoanálise dos sistemas atualmente considerados inquebráveis, ressaltando como foi possível desenvolvê-la a partir de técnicas já existentes e justificando a escolha de algoritmos baseados em inteligência artificial para a sua implementação.

V.1 - Revisão do estado-da-arte da criptologia

Como pudemos observar, as técnicas criptográficas evoluem em função dos avanços da criptoanálise, da tecnologia existente e das novas aplicações que surgem a cada dia. Nos tempos antigos, as mensagens tinham de ser escondidas mesmo depois de recebidas pois o conhecimento dessas pelo adversário poderia comprometer novas comunicações que se utilizassem do mesmo algoritmo/chave. Nessa época, as embaixadas precisavam se preocupar em alterar as palavras da mensagem recebida, mantendo o sentido do texto, antes de divulgá-lo. Ocorria que os primeiros algoritmos eram completamente vulneráveis a ataques do tipo

B (**texto claro conhecido e respectivo criptograma**) e ataques do tipo C (**texto claro escolhido**). (Vide capítulo IV, item 4).

A criptografia moderna já constrói algoritmos que resistem aos ataques tipo B e C e mais: temos algoritmos que são de domínio público (por exemplo, DES - vide capítulo IV, item 2.1.2) e outros em que até parte da chave é de domínio público (RSA, Mochila - vide capítulo IV, itens 2.2.1 e 2.2.2) e mesmo assim a segurança é alcançada.

Dos algoritmos de domínio público atualmente utilizados, os mais consagrados são, sem dúvida, o DES e o RSA. Ambos vêm resistindo aos mais variados ataques criptoanalíticos [43].

Do ponto-de-vista prático, existe uma certa tendência de se utilizar sistemas híbridos, misturando-se a velocidade de algoritmos de chave secreta como o DES com a segurança do RSA no que diz respeito ao gerenciamento das chaves criptográficas. Mas, segundo READ [44], a longo prazo o DES será substituído por sistemas de chave pública apesar de que, na área comercial o DES ainda deverá ser muito utilizado na próxima década, devido aos investimentos já feitos. Alguns autores [44], [6] acreditam que muitas organizações preferirão investir em algoritmos próprios, apesar do custo adicional que isso representa e do risco de se gerar sistemas que não apresentem muita segurança. BRICKELL e ODLYZKO [43] relacionam alguns produtos existentes no mercado que utilizam algoritmos próprios que foram quebrados, alguns inclusive por ataque unicamente ao texto cifrado (tipo A).

Com relação aos sistemas de chave pública, que fizeram dez anos em 1985 [45], nos seus três primeiros anos de existência, surgiram três algoritmos (RSA, Mochila e o de McEliece). O esquema de McEliece é muito complexo para ser usado na prática e o da Mochila é considerado inseguro por alguns autores. Sobra o RSA, que atualmente reina sozinho no mundo da criptografia de chave pública [45].

Segundo DIFFIE [45], a principal motivação atualmente da pura pesquisa criptográfica é a busca de formas de se substituir papéis, canetas e apertos de mão por computadores e telecomunicações. É a autenticação de mensagens e usuários e assinatura digital levadas às últimas conseqüências. Estão sendo pesquisadas atualmente, além da "zero knowledge proofs" (essa técnica será mais bem explicada no capítulo VI), formas de se dividir dados secretos entre n pessoas de maneira que sejam necessárias pelo menos k delas para que os dados possam ser recuperados ($k < n$) [45].

V.2 - Uma Proposta para Criptoanálise dos Sistemas Atualmente Considerados Inquebráveis

Em 1982 teve início um projeto japonês cuja finalidade era criar máquinas capazes de trabalhar de forma "inteligente". O projeto de Quinta Geração visava a elaboração de máquinas que aceitariam comandos orais, que fariam tradução de texto, que seriam capazes de "ver" e "sentir" e tomariam decisões. Para isso, 9 empresas (entre elas Fujitsu, Nec, Hitachi, Sharp, Mitsubishi e Toshiba) se juntaram num projeto de 850 milhões de dólares que seria desenvolvido em 10 anos sob a direção do Icot (Institute for New Generation Computer Technology) [46].

Apesar de já se saber que esse projeto não atingiu plenamente seus objetivos, a inteligência artificial abriu novos campos de pesquisa e trouxe de volta a interdisciplinaridade entre as ciências. De uma maneira geral pode-se dividir a pesquisa em inteligência artificial da seguinte forma [46]:

1 - Capacidade de resolver problemas:

A - "Raciocínio" automático:

- **Busca heurística;**
- Programação lógica;
- Prova de teoremas.

B - Sistemas especialistas:

- Inferência;
- Representação;
- Aquisição e aprendizado.

2 - Capacidade sensorial e cognitiva:

A - Interfaces de linguagem natural;

B - Visão por computador;

C - Reconhecimento da fala.

3 - Robótica:

A - Modelos geométricos e "racionínio";

B - Realimentação sensitiva e integração;

C - Movimento e planejamento de tarefas.

Vamos nos deter na busca heurística, que vai nos interessar mais tarde.

Segundo SLAGLE [47] os principais objetivos da parte da inteligência artificial conhecida como busca heurística são: compreender a inteligência humana e usar a "inteligência da máquina" para adquirir conhecimento e resolver problemas intelectualmente difíceis.

As principais diferenças entre a programação convencional e a programação de inteligência artificial, segundo LUCENA [48] são:

- a primeira é algorítmica, ela define todos os passos explicitamente para a solução de um determinado problema enquanto a segunda é baseada no conceito de busca heurística. Ou seja, projeta-se o espaço de solução dos problemas, de forma explícita ou implícita, e aplica-se sobre ele um método de solução do problema que , em geral, é implícito e não explícito como no caso da programação convencional;

- na programação convencional, a informação e o controle sobre a área do programa que está tratando do problema aparecem de forma integrada, enquanto na

programação de inteligência artificial elas estão dissociadas;

- a diferença mais sutil e fundamental é que na programação convencional sempre se espera a resposta correta, o que não ocorre na inteligência artificial.

Assim sendo, podemos resumir as diferenças entre o tratamento convencional da informação e as novas estratégias de resolução através das características próprias dessas últimas, que incluem a busca heurística, o que lhes permite modelar estratégias para resolução de problemas específicos dentro de um domínio particular. **Através da inteligência artificial e, particularmente, da busca heurística, pode-se procurar novas soluções para problemas antigos.**

E o que vem a ser heurística? Heurística é intuição [49], é uma estratégia, método ou truque usado para aumentar a eficiência de um sistema que procura resolver problemas complexos [47]. Algumas são específicas para um problema, outras são gerais e se aplicam a um conjunto de problemas. Essa, no entanto, não é uma técnica completamente dominada e ainda requer alguma pesquisa.

Acreditamos, porém, que um vasto campo de pesquisa pode vir a se abrir para a criptoanálise. Através das novas técnicas de inteligência artificial que vêm se tornando cada vez mais viáveis através de máquinas cada vez mais rápidas e de novas tecnologias como os chips óticos [50] e processamento vetorial, novas possibilidades nunca dantes pensadas surgem para a criptoanálise.

Essas novas técnicas poderão permitir a elaboração de

novos algoritmos para a exploração das "fraquezas" dos algoritmos criptográficos atualmente considerados seguros. "Fraquezas" essas que atualmente, através de algoritmos convencionais, levam anos para serem quebradas.

Sugere-se, então, uma pesquisa mais profunda de criptoanálise no campo da inteligência artificial e a elaboração de algoritmos baseados na busca heurística.

Como já observamos, a inteligência artificial, através da busca heurística, procura solucionar problemas formais complexos, onde os significados não são totalmente explícitos nem totalmente independentes do contexto. São exemplos desse tipo de problema jogos incomputáveis como o xadrez, problemas combinatórios complexos como planejamentos e demonstrações de teoremas que exijam procedimentos não mecânicos (intuição).

O estudo das técnicas criptográficas e criptoanalíticas nos levou à conclusão de que a CRIPTOANÁLISE nada mais é do que um exemplo de um problema combinatório complexo, pertencendo, portanto, à classe de problemas que podem ser tratados através de heurística. A figura V.1 ilustra a proposta apresentada.

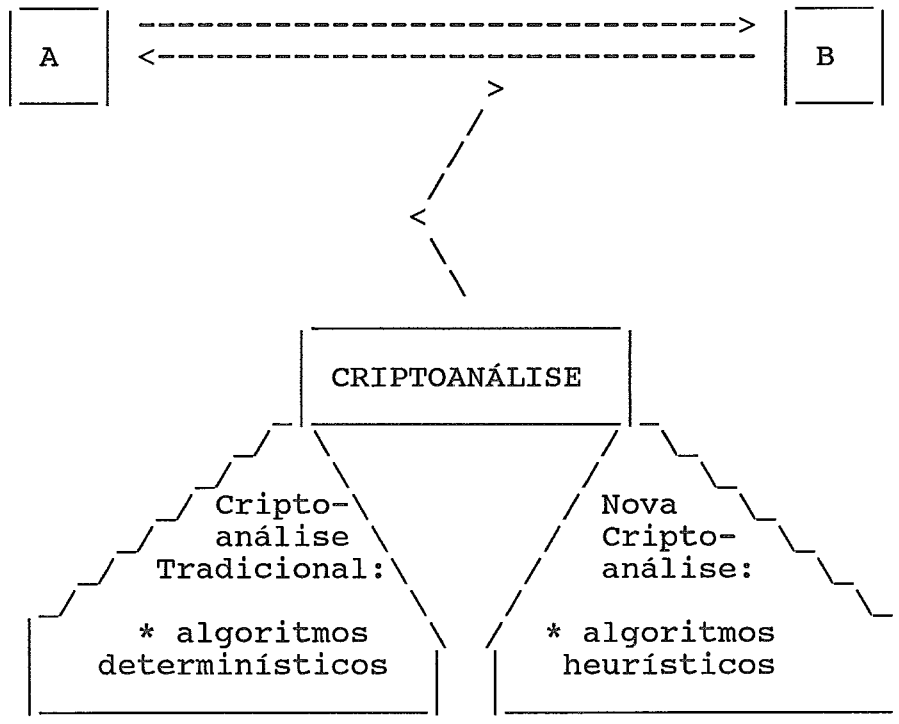


FIGURA V.1: Proposta para uma nova criptoanálise.

A busca heurística nem sempre dá resposta exata mas ela pode, rapidamente, determinar o intervalo em que esta se encontra. A partir daí pode-se usar um velho algoritmo determinístico que terminará o trabalho, ou seja, determirá a solução exata no "conjunto calculado heuristicamente".

Vale lembrar que, na criptoanálise tradicional (antes do computador), a intuição e os macetes eram umas das principais ferramentas de trabalho e que o único método criptográfico até hoje comprovadamente inquebrável é o one-time-pad, a substituição polialfabética sem repetição.

CAPÍTULO VI - CRIPTOGRAFIA E SEGURANÇA DE DADOS

O único método prático conhecido para proteger dados transmitidos através de redes de comunicação que usam linhas terrestres, satélites ou microondas é a criptografia. Em alguns casos, é a forma mais econômica de se proteger arquivos de dados. Procedimentos criptográficos também são usados para autenticação de mensagens, assinatura digital e identificação pessoal para autorização de transferência de fundos e operações com cartão de crédito [5]. Isso sem falar em identificação pessoal para uso de recursos computacionais, o que acontece em todo sistema de computadores multiusuário, as famosas "passwords".

Neste capítulo procuraremos relacionar criptografia a segurança de dados e de sistemas: além de algumas definições, analisaremos os "tipos" de segurança desejáveis e possíveis, "quantidade" de segurança, controles de acesso a recursos computacionais e segurança em redes de computadores.

O conceito de segurança de dados, segundo a IBM (conforme o citado em [51]) é o seguinte:

"A segurança de dados pode ser definida como a proteção de dados contra a revelação acidental ou intencional a pessoas não autorizadas, e contra alterações não autorizadas."

Segundo DENNING [25], existem dois objetivos principais na criptografia usada para segurança de dados: segredo (ou privacidade), que evitaria a divulgação não

autorizada da informação e autenticidade (ou integridade), que evitaria a modificação não autorizada. Segundo DENNING e DENNING [52], outras técnicas de proteção dos dados seriam:

- **controles de acesso:** onde cada acesso direto a dado ou programa para leitura, modificação ou execução só é autorizado se de acordo com a política vigente no sistema, através de mecanismos que implemetem corretamente a política adotada; (vide figura VI.1)

- **controles de fluxo de dados:** que regulam a disseminação de informação entre os diversos usuários; (vide figura VI.2)

- **controles de inferência:** que protegem bancos de dados estatísticos evitando que dados confidenciais possam ser deduzidos a partir das estatísticas fornecidas. (vide figura VI.3)

PARKER [53] cita ainda um outro ponto importante: segurança "orientada ao pessoal", que seria a educação das pessoas envolvidas no trabalho de maneira a evitar alguns "furos" que poderiam comprometer a segurança do sistema. "Furos" esses que podem ocorrer tanto por descuido quanto por desonestidade.

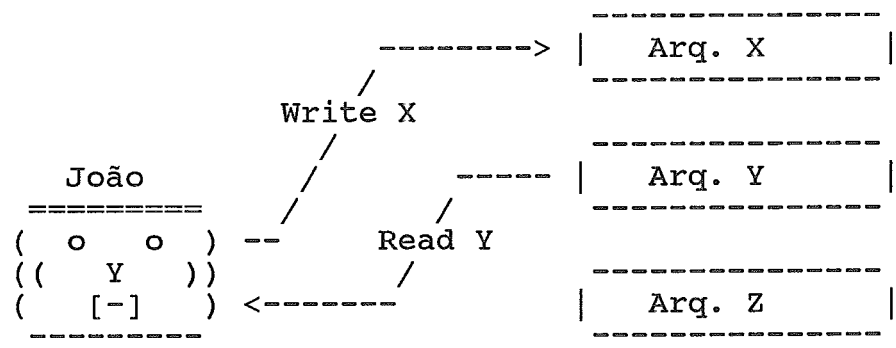


FIGURA VI.1: CONTROLE DE ACESSO. João pode ler o arquivo X, escrever no arquivo Y e não tem nenhum acesso ao arquivo Z. [52]

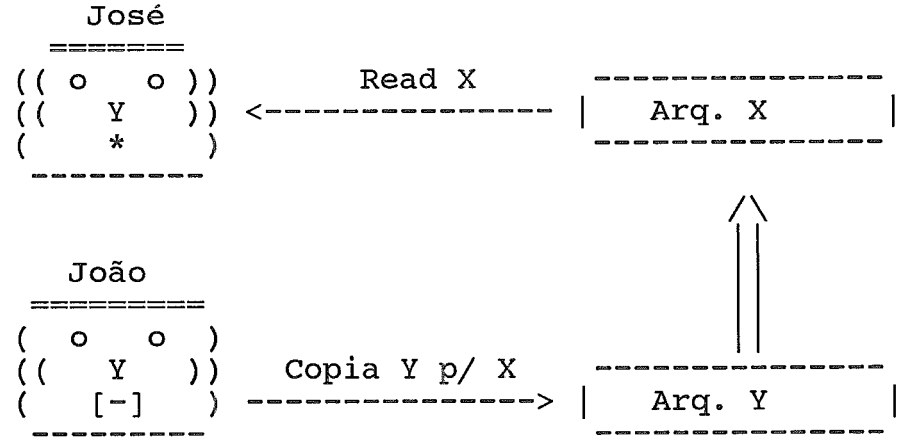


FIGURA VI.2: CONTROLE DE FLUXO DE DADOS. Impossibilitado de ter acesso ao arquivo Y, José pede ao amigo João que faça uma cópia para ele. Controle de fluxo evitaria isso... [52]

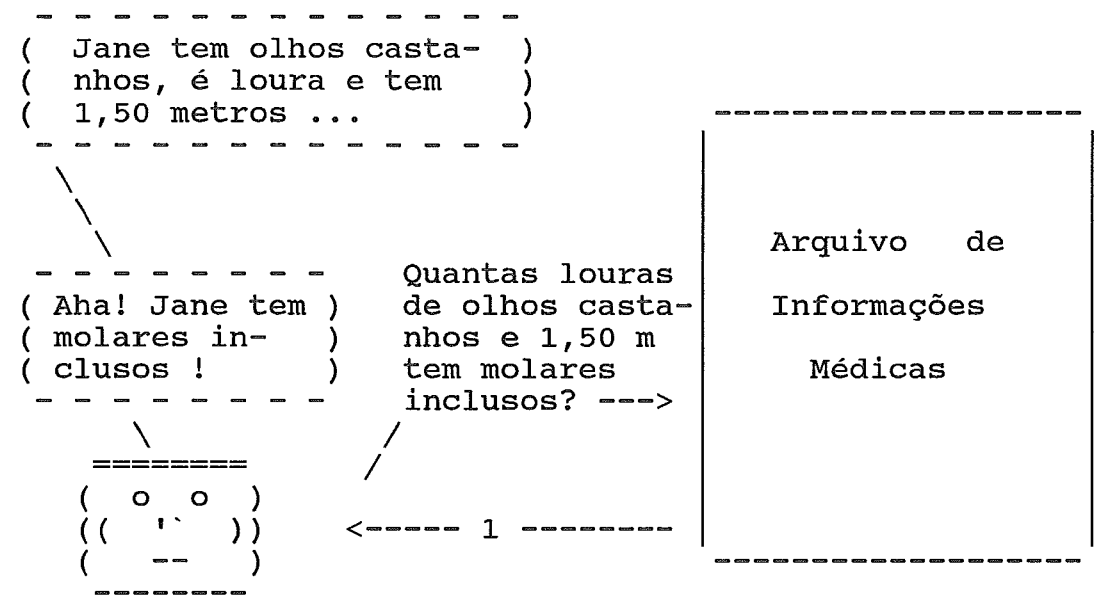


FIGURA VI.3: INFERÊNCIA. Previne perguntas que permitam a dedução de informações confidenciais a partir de dados estatísticos. [52]

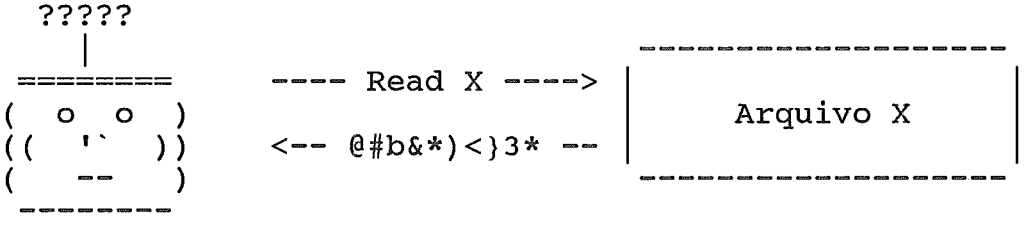


FIGURA VI.4: ENCRIPTAR. João obtém acesso ilícito ao arquivo X, mas não consegue entender seu conteúdo que está criptografado. [52]

A segurança do sistema como um todo e segurança de dados do sistema se confundem uma vez que, em última instância, o sistema nada mais é que um grande conjunto de dados de naturezas diversas (os programas do sistema, os aplicativos, os dados da folha de pagamentos etc.) arrumados de forma ordenada para que todo o conjunto possa gerar

informações úteis de forma segura, íntegra e otimizada.

Assim sendo, tanto a criptografia quanto as outras técnicas citadas acima, são ferramentas da segurança de dados e da segurança do sistema. Por exemplo, a criptografia é usada tanto para proteger dados transmitidos de um terminal a um computador via linha telefônica quanto para verificar a identidade de um usuário que deseja se utilizar dos recursos do sistema (passwords).

Esta segurança pode ser:

- interna;
- externa;
- da interface.

Segurança interna se preocupa com o software e o hardware utilizado e os dados que estão sendo processados. Já a segurança **externa** seriam os backups dos dados guardados em fita num cofre-forte, chaves nos terminais, controle de acesso físico de pessoas à sala do sistema etc. Consideramos problemas de segurança **da interface** entre o meio externo e interno o acesso de usuários aos recursos do sistema, problemas de autenticação etc., estes também tratados através das técnicas acima descritas e de outras próprias, como o reconhecimento de uma característica física do usuário como impressão digital ou a recentemente desenvolvida por Fiat e Shamir chamada "zero knowledge proofs", onde o usuário prova ao computador que ele dispõe de um número sem que este número precise ser revelado [54]. A substituição dos cartões magnéticos pelos "smart cards" também é outro avanço da segurança de interface dos

sistemas. Consta que os já tradicionais cartões magnéticos são facilmente copiáveis, o que não ocorre com os "smart cards" [55]. A "zero knowledge proof", inclusive, é implementada através de "smart cards". Um outro aspecto de segurança **externa** que será visto no capítulo VII é o legal (a existência de uma legislação que proteja os dados confidenciais de cidadãos e empresas).

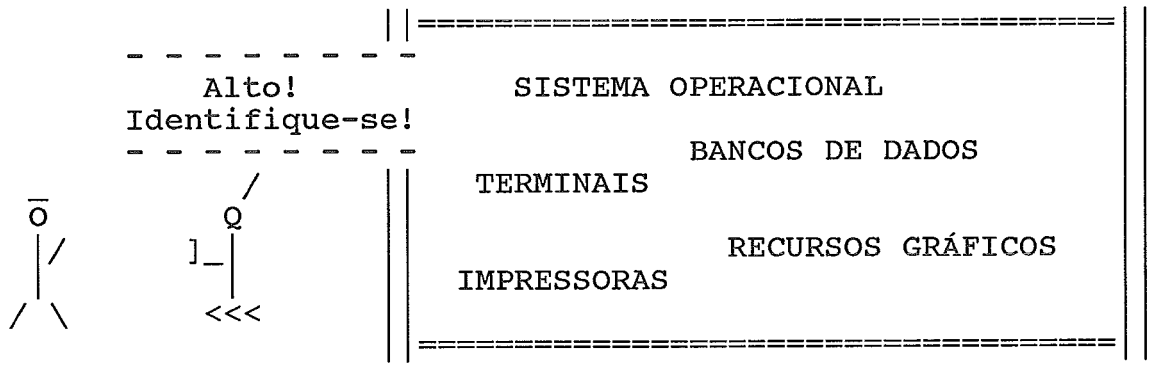


FIGURA VI.5: Segurança de Interface.

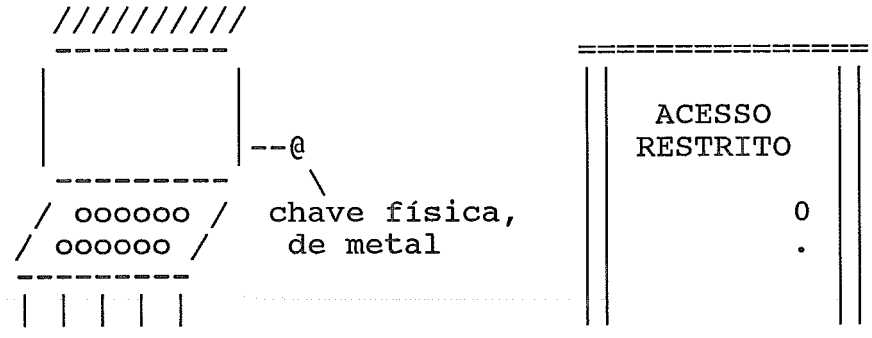


FIGURA VI.6: Segurança Externa.

Nessa dissertação, quando nos referirmos à segurança, fica implícito que se trata da interna. De outra forma

mencionaremos explicitamente que se trata de segurança de interface ou externa.

Os projetistas de sistemas se deparam com o problema de proteger todo o sistema, ou seja, um enorme conjunto de dados. Através da criptografia, esse problema é reduzido ao de se proteger um pequeno grupo de dados (as chaves criptográficas) [5]. Para tal, parte-se do pressuposto de que um algoritmo criptográfico robusto esteja sendo utilizado.

Mesmo o problema reduzido (de se proteger as chaves) deve ser cuidadosamente analisado juntamente com todas as peculiaridades dos dados a serem tratados pois: "A segurança de um sistema não deve depender do segredo de alguma coisa que não possa ser facilmente trocada em caso de comprometimento" [3].

De qualquer maneira, assim como não é possível haver cofres de bancos invioláveis, por maior que seja a proteção, não é possível construir-se sistemas de computadores absolutamente seguros [52]. Quanto mais proteção se deseja, mais se precisa gastar. SHANNON [4], ao aplicar sua Teoria Matemática da Informação ao sistemas criptográficos, faz uma distinção clara entre o que ele chamou de SEGURANÇA INCONDICIONAL e SEGURANÇA COMPUTACIONAL. A primeira é conseguida a um custo elevadíssimo e o único algoritmo criptográfico que comprovadamente a alcança é a substituição polialfabética que utiliza chaves aleatórias sem repetição (one-time pad), como já tivemos a oportunidade de observar. Entretanto, a nível de sistema, tal segurança

não existe. Segurança computacional é o comprometimento custo-segurança desejado, é a relação entre o que se gastará na proteção e o que gastaria um "inimigo" para quebrá-la. SHANNON [4] define como segurança incondicional aquela que resiste a ataques de um inimigo com recursos computacionais infinitos e segurança computacional aquela que resiste ao ataque de alguém que dispõe de quantidade finita de recursos. Assim, a técnica a ser utilizada está diretamente relacionada à aplicação.

Até bem pouco tempo atrás, a segurança do sistema estava bastante calcada na segurança externa. Como o processamento era centralizado, mecanismos físicos restringiam o acesso a um grupo pequeno e bem conhecido de usuários. Isso se aplica também à primeira etapa da automação bancária, que praticamente não usava teleprocessamento. Atualmente, os "hackers" são uma dor de cabeça para os mantenedores de sistemas que permitem o acesso de seus usuários através de linhas telefônicas e os chamados "vírus" se espalham, sendo notícia no mundo inteiro. Para se "escutar" milhares de transações de várias firmas comerciais e governos basta adquirir uma antena e apontá-la ao satélite correto ... [56] Aliás, uma outra aplicação de criptografia é no controle de satélites, os comandos enviados da base ao satélite devem ser criptografados para evitar que terceiros resolvam alterar a sua órbita, por exemplo.

A Universidade Federal do Rio de Janeiro também começa a ingressar nessa nova era. Suas diversas máquinas começam a ser interligadas através de redes como a DECNET (Digital

Equipment Corporation Network) da Digital, BNA(Burroughs Network Architecture) da Unisys e a SNA(Systems Network Architecture) da IBM. Estão sendo feitas também ligações com outras universidades brasileiras como a UNICAMP (Universidade de Campinas), através da RENPAC (Rede Nacional de Pacotes) da Embratel e com universidades estrangeiras, como a UCLA (University of California at Los Angeles), via BITNET(Because It's Time Networking). No caso da universidade e centros de pesquisa, a privacidade não é um fator de suma importância. Entretanto é preciso evitar que pessoas estranhas à comunidade acadêmica utilizem recursos computacionais pagos pelo erário público ou que comprometam a integridade dos vários sistemas interligados.

As diversas redes acima citadas não oferecem a facilidade de enviar mensagens criptografadas. No entanto, elas se preocupam com a integridade da mensagem e com a autenticidade dos diversos nós que a compõem. Além disso, segundo DENNING [25], como a maioria das redes divide a mensagem em pacotes que são enviados "embaralhados" e muitas vezes por caminhos diferentes, é difícil para alguém na escuta ordená-los e entendê-los. É esse também o argumento usado pelos técnicos da Embratel quando perguntados pela privacidade dos dados que trafegam na rede pública via RENPAC [11].

Na BNA, por exemplo, cada máquina tem uma senha, que é verificada quando da conexão desta à rede [57].

Então, na atual conjuntura, cabe aos usuários da rede se preocuparem com o sigilo de seus dados. No mercado

nacional podem ser encontrados algumas software-houses que desenvolvem e comercializam software de proteção para microcomputadores. Segundo reportagem da PCMundo [58], existem nove fornecedores que comercializam 11 produtos que implementam controle de acesso, controle de fluxo e criptografia, além de algumas funções de auditoria. No que diz respeito às máquinas de grande porte, geralmente o próprio fabricante dispõe de algum produto que criptografa arquivos e dados transmitidos. Os preços dos produtos nacionais (software) para micros variam de 4 a 95 OTNs (em Janeiro de 1989) [58], o que, em dólares, representa uma variação de 24,68 a 586,15. Já no mercado americano, o preço dos pacotes de software criptográfico variam de 79,95 dólares (para micros que utilizam o sistema CP-M) até 10 mil dólares, para versões em assembler para grande porte [59]. Outros pacotes, com outras funções como auditoria, podem custar mais de 12 mil dólares [60]. Além do preço, para que possam ser exportados, qualquer aparelho/software criptográfico americano necessita de autorização especial do Departamento de Estado [61]. A SEI (Secretaria Especial de Informática) também tem restrições para a sua importação [6].

Pelo que se pôde observar na literatura consultada, a maioria dos produtos (software e hardware) usam o DES ou pequenas variações dele. Alguns já utilizam o RSA e pudemos encontrar outros que utilizam máquina rotor e registros de deslocamento, além de alguns algoritmos totalmente produzidos pelo fabricante, como o B152 da British Telecom [62]. E no que diz respeito à criptografia das máquinas de

grande porte, muitos sistemas empregam software no computador central e hardware nos terminais remotos [63].

Algoritmos consagrados e de domínio público como o DES, o RSA, máquinas rotor e registros de deslocamento foram descritos no capítulo IV, podendo servir de base para essa discussão.

Deve-se notar que existe também a preocupação de se padronizar protocolos para transmissão de dados criptografados, mesmo que o algoritmo fique a critério de cada um. A ISO (International Organization for Standardization) já tem o seu padrão, e no Brasil um grupo de especialistas vem trabalhando no sentido de definir um padrão nacional [11]. Mais recentemente e a nível internacional, foi aprovado pela ANSI (American National Standards Institute) um padrão de assinatura digital para a área bancária [9].

Como, segundo PARKER [64], a segurança futura dos computadores será toda baseada em criptografia, esse é um mercado promissor, além de se constituir num campo de pesquisa muito interessante. Existe também o problema estratégico de se confiar questões delicadas de segurança a outras empresas ou outros países. Dependendo do caso em questão, é necessário se definir as políticas e se desenvolver os mecanismos internamente passando, inclusive, pela formação de mão-de-obra especializada.

CAPÍTULO VII - SEGURANÇA DE DADOS: ASPECTOS SOCIAIS E LEGAIS

> "Criminoso de colarinho branco é o homem que aprendeu a roubar com um lápis. Nas últimas décadas, os mais perspicazes descobriram que os ganhos são maiores e os riscos menores se o roubo é feito com um computador."
>
> DENNING & DENNING [52]

A segurança de dados em si, englobando técnicas como a criptografia, é questão tecnológica e operacional. É o meio pelo qual as políticas e regras relativas ao que é confidencial podem ser colocadas em prática, de forma correta e eficiente [64].

Nos capítulos IV e VI vimos os mecanismos para proteção de dados e programas. Seria o análogo ao cofre do banco. Se, apesar do cofre, o banco é assaltado, existe toda uma legislação que determina punições para os assaltantes. No caso dos dados e dos programas, tal legislação também existe. Em alguns países ela é incipiente, devido à pouca idade dos problemas de que trata, em outros já se encontra bastante adiantada.

Num seminário sobre Computadores e Privacidade na Próxima Década, ocorrido em 1979, um dos artigos (NYCUM [65]) questiona se no final dos anos 80 temas como Transferência Eletrônica de Fundos, Terminais Ponto de Venda, Correio Eletrônico e Teleconferências seriam tão comuns quanto uma chamada telefônica, e chega à conclusão de que não, uma vez que mudanças sociais são muito mais lentas que inovações tecnológicas.

Neste capítulo analisaremos os aspectos legais da proteção de programas através de expedientes como patentes e "copyrights". Serão vistos também aspectos da proteção dos dados manipulados (problemas de privacidade, sigilo etc.) tanto Brasil quanto em outros países.

VII.1 - Proteção de Programas

"A impossibilidade de se tratar informação como mercadoria tem-se refletido nos sérios problemas que o sistema de poder enfrenta para reduzi-la às categorias simbólicas pelas quais exerce a dominação. O caso mais flagrante envolve o programa de computador (informação por excelência) que se mantém irredutível, exceto por decisões políticas ... exigindo piruetas mentais para enquadrá-lo seja no código de propriedade intelectual, seja no código da propriedade industrial" [66].

JØRGEN [67] define programa de computador como algo imaterial, que não tem forma ou substância mas que necessita de alguma forma de proteção legal devido aos investimentos de desenvolvimento e teste. O autor discute que tipo de proteção poderia ser esta (direito autoral; patente; direito de propriedade - nesse caso o programa seria apenas alugado; contratos etc.) e até que ponto um programa modificado permanece o mesmo, ou seja, qual o grau de semelhança que dois programas devem ter para que um seja considerado plágio do outro. JØRGEN conclui que a proteção mais adequada seria a "norma do catálogo" que, no entanto, é uma legislação

existente apenas nos países escandinavos.

"Norma do catálogo" é uma lei existente em todos os países nórdicos. Até certo ponto é muito semelhante ao direito autoral, e foi feita visando a proteção de catálogos, compilações de fórmulas, tabelas etc. Ou seja, trabalhos que contêm uma grande quantidade de dados organizados de forma sistemática. Essa lei inclui, além do aspecto literário do direito autoral, o aspecto técnico, de "know-how", também presente nesse tipo de obra [67]. A "norma do catálogo" estaria apta a proteger o programa independentemente da forma em que ele se apresenta, pois protege a coleção de dados que é a sua essência.

A lei de proteção ao software brasileira, aprovada em janeiro de 1988, trata o software como obra literária e regulariza sua comercialização [68]. Entretanto, por aqui, a maior polêmica não girou em torno da categoria em que se deve enquadrá-lo, e sim da questão da similaridade para os casos de importação de software: só podem ser importados aqueles que não tiverem similar no mercado interno. Esse foi o ponto mais discutido pela imprensa especializada [68], [69], [70]. Todo software deve ser registrado na SEI(Secretaria Especial de Informática) assim como todo livro deve ser registrado na Biblioteca Nacional para ter reconhecido o direito do autor. É então de competência da SEI a averiguação da existência de plágio ou similaridade.

A legislação americana, que também trata o software através do direito autoral, evoluiu de inicialmente punir apenas a simples cópia do software até o estágio atual onde

o plágio, quando comprovado, é condenado. Essa evolução se deu através de jurisprudência [69].

Com a aprovação da lei do software, a pirataria, antes atividade institucionalizada no Brasil (podia-se até encontrar anúncio de "pirato-houses" nos jornais) passou a significar crime. Muitos piratas entrevistados pela revista Micro Sistemas estão abandonando o ramo e as Associação Norte-Americana dos Produtores de Software, a Abes (Associação Brasileira das Empresas de Software) e Assespro (Associação Brasileira das Empresas de Serviço de Informática), já começaram uma ação conjunta de combate à pirataria em nossas terras [71].

Do ponto de vista técnico, a pirataria seria evitada através do CONTROLE DE FLUXO, visto no capítulo VI. Contudo, enquanto não existir o controle de fluxo absoluto, os piratas atacarão, restando às "software-houses" investir nessa linha de pesquisa além, é claro, de recorrer às autoridades competentes.

VII.2 - Proteção dos Dados Manipulados

"Ter vida privada é um direito humano. Não é uma questão tecnológica, mas sim social, jurídica e política. Entretanto, o impacto da tecnologia sobre a sociedade está se tornando tão diversificado e profundo que as questões sociais e tecnológicas hoje se combinam", afirma PARKER [64].

O abuso do direito que as pessoas têm à vida privada pode vir a ocorrer com a coleta irrestrita de dados a seu respeito, armazenagem de dados inexatos ou incompletos, revelação não autorizada de dados, conclusões errôneas ou prejudiciais baseadas neles [64].

Muitas pessoas honestas acreditam que não têm nada a temer quanto à violação de sua intimidade, uma vez que suas vidas são um "livro aberto". Assim sendo, a automação total da burocracia só lhes será benéfica, por exemplo, em qualquer estabelecimento comercial essas pessoas teriam crédito levantado imediatamente não tendo que esperar para fazer compras. No entanto, o risco está na armazenagem de dados incorretos ou incompletos. PARKER [64] nos dá um exemplo de diagnóstico médico equivocado ("possíveis palpitações cardíacas" quando se tratava de um problema de estômago) e que essa informação poderia prejudicá-lo no futuro, por exemplo, no caso de ele querer fazer um seguro de vida.

Uma questão polêmica no que diz respeito a ameaças à vida privada é a adoção de um número único que identificaria

cada cidadão. Se por um lado o número único tornaria alguns serviços mais eficientes, por outro ele permite o cruzamento de vários arquivos distintos para os mais variados fins. A esse respeito, existe nos Estados Unidos uma lei do Direito à Vida Privada de 1974 que proíbe o uso do Social Security Number (candidato americano natural ao cargo de número único, uma vez que cada pessoa tem um) entre os órgãos governamentais, exceto quando determinado por lei [64]. No Brasil, o uso do CIC como número único já é uma realidade, pelo menos nos meios financeiros. Ao solicitar a emissão de ações da Telebrás, adquiridas via Plano-de-Expansão, o CIC foi suficiente para me identificar... SWAINE [72] sugere que nomes e números sejam criptografados, ou seja, que se use mnemônicos em cada base governamental em que o cadastramento for necessário. Esses mnemônicos protegeriam a privacidade do cidadão. Esta, no entanto, seria uma solução técnica que exigiria mais esforços do próprio cidadão (que forneceria os mnemônicos) para um problema político. Uma maneira muito mais simples de se resolver isso seria simplesmente não se comparar arquivos distintos, ou seja, respeitar-se o princípio de que cada informação só deve ser usada para o que foi coletada.

Os 5 princípios da "prática justa da informação", segundo RULE e colaboradores [73] são:

- não deve haver sistemas de arquivos de informações pessoais cuja existência seja secreta;

- deve haver uma forma de um indivíduo saber que informação sobre ele está armazenada e como é usada;

- deve haver uma forma de um indivíduo impedir que

informações coletadas para um fim sejam usados para outros sem o seu consentimento;

- ele deve poder corrigir ou incluir registros que contenham informações a seu respeito;

- qualquer organização que crie, mantenha, use ou dissemine informações pessoais deve assegurar a confiabilidade dos dados, que eles sejam usados para o fim esperado e deve tomar as precauções necessárias para evitar o mau uso dos dados sob sua guarda.

Privacidade, no que diz respeito a dados pessoais armazenados, está relacionada ao direito dos indivíduos de controlar e influenciar que informações a seu respeito podem ser coletadas e armazenadas, por quem, para que fins e para quem essas informações podem ser cedidas/reveladas. Também se relaciona ao direito de saber que dados a seu respeito foram compilados, se estão corretos e completos para o fim a que se destinam. Além disso, as pessoas devem poder esperar que essas informações não serão cedidas a terceiros sem autorização explícita e que elas têm o direito de verificar a acurácia de tal informação. Esta é a definição de MEYER [5]. Já BLEKELI [74] nos diz que: **Privacidade** relacionada a sistemas de informção significa o interesse de uma pessoa (física ou jurídica) em ter controle sobre todos os dados a seu respeito. Podemos então resumir assim:

PRIVACIDADE = SIGILO + COMPLETEZA + SER INFORMADO

Sigilo para proteger dados sensíveis ou que não venham ao caso para a situação presente. **Completeza** no sentido da

informação estar correta e completa. **Ser informado** é o direito de cada um de saber o que existe a respeito próprio, aonde e para quê.

Nos EUA, leis que tratam explicitamente da vida privada existem desde 1973 [64].

Nos países nórdicos, o assunto já vem sendo tratado há bastante tempo. Na Noruega, por exemplo, foi aprovada em maio de 1978 uma lei para proteção de "registros de dados pessoais e outras facilidades em que informações pessoais são utilizadas", além de criar um órgão (Data Surveillance Service) para fazer vigorar a lei [75]. O DSS é responsável pela fiscalização de bancos de dados, fichários, por definir quais as informações/aplicações que devem ser mais controladas, transferência interna e externa de informações, fluxo de dados transfronteira etc. Essa legislação se baseia em outras já existentes na Noruega como sigilo, proteção da personalidade (relacionada com espionagem de comunicação privada e vida privada), direito de ser informado (livre acesso a documentos mantidos pela administração, menos quando se trata de certos tipos de informação como saúde do paciente ou o que a polícia sabe sobre determinado criminoso) e direito restrito de registrar informação.

A Constituição Brasileira [76], promulgada em 1988, em seu artigo 5, tem três parágrafos que procuram garantir a privacidade, conforme as definições acima. São esses:

Parágrafo XIV: "é assegurado a todos o acesso à informação e resguardado o sigilo da fonte, quando necessário ao exercício profissional".

Parágrafo XXXIII: "todos têm direito a receber dos órgãos públicos informações de seu interesse particular, ou de interesse coletivo ou geral, que serão prestadas no prazo da lei, sob pena de responsabilidade, ressalvadas aquelas cujo sigilo seja imprescindível à segurança da sociedade e do Estado;".

E o parágrafo LXXII diz: "conceder-se-á "habeas-data":

a) para assegurar o conhecimento de informações relativas à pessoa do impetrante, constantes de registros ou bancos de dados de entidades governamentais ou de caráter público;

b) para a retificação de dados, quando não se prefira fazê-lo por processo sigiloso, judicial ou administrativo;".

KATZAN [51] nos faz ver que os computadores não criaram o problema de segurança de dados, nem o problema do sigilo, mas o uso efetivo dos computadores aumentou o alcance da coleta de dados. Quantidades cada vez maiores de dados podem ser coletados, armazenados e mais tarde recuperados a qualquer momento. Nisto reside o problema. Se por um lado a eficiência dos diversos organismos que lidam com esses dados cresce enormemente, cresce também, e nas mesmas proporções, o controle desses grupos sobre toda a sociedade. Além disso, se quando tudo dá certo o resultado é maravilhoso (pode-se inclusive dizer que o controle até certo ponto torna cada um mais consciencioso de seus atos), em caso de falha os prejuízos são incalculáveis. Esta é uma questão de **Privacidade & Autonomia VERSUS Eficiência.** RULE

[77] conclui que nos resta escolher entre um mundo eficiente e centralmente controlado ou um mundo mais privado e não tão eficiente assim, com pequenos sistemas de informação distribuídos localmente.

Uma pesquisa citada em [64] nos informa que a utilização de computadores na manutenção de registros contribui para que:

(1) os dados tenham se tornado mais atualizados;

(2) os dados fossem usados com mais eficiência;

(3) houvesse um atendimento mais rápido aos pedidos de informação;

(4) houvesse também de um novo intercâmbio de atividades entre organizações;

(5) surgissem sistemas maiores de registros;

(6) os registros se tornassem mais exatos, com menos erros, mas diferentes e mais sérios;

(7) fosse possível maior segurança do que nos sistemas manuais, mas pouco se fez nesse sentido, além de proporcionar segurança física. O potencial de segurança , porém é muito maior nos sistemas computadorizados.

Por outro lado, RULE e colaboradores [73] narram a seguinte experiência:

"Em 1971 um grupo de especialistas em computação, comunicação e supervisão foi reunido e recebeu a seguinte tarefa: Suponha que vocês são consultores da KGB e que vocês precisam projetar um sistema para que se supervisione todos os cidadãos e visitantes na URSS. O sistema não deve ser muito intrometido ou óbvio, ou seja, deve fiscalizar discretamente. Qual foi a decisão do grupo? Construir um

sistema de Tranferência Eletrônica de Fundos. Não apenas esse sistema controlaria toda a contabilidade financeira e proveria estatísticas cruciais para uma economia centralizada, mas seria também o melhor sistema de fiscalização que se pôde imaginar sem ser escancaradamente intrometido."

Isto nos revela que o problema não é tão fácil quanto parece, ou seja, que as leis por si só não são suficientes. É preciso que toda a sociedade esteja atenta e controle seus controladores, pois agências que atuem à margem da lei ou até mesmo empresas podem dispor de uma infinidade de recursos e subterfúgios para controlar o que quiserem. Um exemplo é o processo dos trabalhadores de uma fábrica de Bolonha, Itália, contra a IBM e o sistema implantado na fábrica para "controle total" das atividades e produtividade dos empregados [78]. A IBM alegou problemas de segurança do banco de dados quando o sindicato solicitou que o controle fosse feito por grupos ao invés de individualmente. O juiz absolveu a empresa por falta de provas, tendo o caso ocorrido em 1984.

Existem situações, não há dúvida, em que algum controle é necessário, como por exemplo, por parte da Receita Federal sobre os sonegadores ou da polícia sobre os criminosos conhecidos. No entanto é discutível até que ponto esse controle deve ser exercido e como. O que alguns acham justíssimo que seja controlado, outros podem achar que não. Por exemplo, as "Filhas da Revolução Americana" podem achar de suma importância que a CIA mantenha arquivos das

atividades dos partidos políticos de esquerda, e por outro lado achar excessivo o controle da Receita Federal sobre instituições de caridade [77]. Chegar-se a um consenso social é um problema, a meu ver, infinitamente mais complicado do que a escolha de um bom algoritmo criptográfico!

CAPÍTULO VIII - CONCLUSÕES

Como tivemos a oportunidade de observar, a criptografia é uma técnica muito antiga, mas nem por isso desatualizada ou ultrapassada. Muito pelo contrário, ela é uma ciência que vem crescendo em importância devido às mais recentes aplicações que a comunidade vem dando aos sistemas de computadores e telecomunicações. Em contrapartida, os avanços tecnológicos vêm impulsionando o desenvolvimento da criptografia e da criptoanálise de forma bastante acentuada.

Sabemos que a criptografia não é infalível. No entanto, ela vem sendo aperfeiçoada a cada dia que passa, tornando-se cada vez mais sofisticada.

Vale lembrar mais uma vez da importância das técnicas criptográficas na segurança dos dados tratados tanto por empresas comerciais e financeiras quanto por estados e nações. Assim sendo, quem não quiser ficar para trás, deverá investir na formação de massa crítica para acompanhar e promover o desenvolvimento na área. É preciso formar pessoas para que essas aperfeiçoem técnicas já existentes e desenvolvam novas. Deve-se também incentivar a indústria criptográfica de software e de hardware para que o conhecimento adquirido não morra nos laboratórios e possa vir a beneficiar a sociedade que sustentou o seu desenvolvimento.

Não podemos esquecer que quando se trata de questões delicadas como sigilo e integridade de informações vitais para um bom andamento dos negócios, é preferível não se

confiar nos "adversários", como no caso do Itamaraty e do preço do café brasileiro na Bolsa de "Commodities" de Londres, citado no capítulo II.

Conforme demonstramos, são várias as possibilidades da criptografia. A cada dia surgem novas aplicações para as técnicas criptográficas existentes, assim como as novas tecnologias vão ampliando os horizontes da criptologia como um todo.

Finalmente, é importante lembrar que não são apenas os avanços técnicos que garantem o uso coerente e consciente da informação. Essa utilização só ocorrerá a partir de decisões políticas de uma sociedade senhora de si.

Pode-se, então, confiar nos computadores? Do ponto de vista técnico, a resposta é sem dúvida positiva. São muitos os mecanismos desenvolvidos para garantir a integridade dos dados tratados eletronicamente, apesar de também serem muitas as técnicas criadas para se burlar as primeiras. O caso é que os computadores não existem sozinhos, nem são autosuficientes, o que significa dizer que a resposta à esta pergunta se torna muito mais complexa sob um enfoque social. Nesse contexto, a resposta seria: não totalmente.

CAPÍTULO IX - REFERÊNCIAS BIBLIOGRÁFICAS

[1] KAHN, David. The Codebreakers - the story of secret writing The macmillan Company, New York 1972. 1164 pags.

[2] SHANNON, C. E. A Mathematical Theory of Communication (Parts I and II) The Bell System Technical Journal, 27(3):379-423 e 27(4):623-56, Jul. 1948.

[3] DIFFIE, W. HELLMAN, M. Privacy and Authentication: An Introduction to Cryptography. Proceedings of the IEEE. 67(3):397-427, Mar 1979.

[4] SHANNON, C. Communication Theory of Secrecy Systems. Bell System Technical Journal pags 656-715.

[5] MEYER, C. H. MATYAS, S. M. CRYPTOGRAPHY:a new dimension in computer data security. John Wiley & Sons Inc., 1982. 755 pags.

[6] NICOLETTI, M. C. Criptografia:Situação Atual e Tendências. Anais da Sucesu 87 pags 277-81.

[7] DEVLIN, K. MICRO-MATHS: mathematical problems and theorems to consider and solve on a computer. Camelot Press Ltd, Southampton, Great Britain, 1984. 103 pags.

[8] BROWNE,Malcolm W. Feito Matemático põe em Xeque Segurança de Códigos Bancários. O Globo, 13 de outubro de 1988.

[9] LOWE,Sue J. A Refined Protocol for Digital Signatures Draws New Interest. Data Communications. New York, 14(10):68-74, Sept. 1985.

[10] BRAMFORD, J. The Puzzle Palace. Sidgwick & Jackson, London, 1982, 465 pags.

[11] LUCA, C. Criptografia:A ciência do segredo. INFO. (61):18-22, Fevereiro de 1988.

[12] DUDEK,V. A Look into IBM's Box of Homegrow Software: Crypto-Mania. PCMagazine, 3(21):39, Oct 30, 1984.

[13] DANA,A. Code Quest. Teaching and Computers,2(6):58-59, March 1985.

[14] SOWARD,E.R.;DESHPANDE,V.G. MasterChip: A game of logic. Infoage, 2(2):28-29, Feb 1983.

[15] KUTZ,R.E. Computer Solutions of Cryptarithms Computing Teacher, 12(1):48, Aug-Sept 1984.

[16] SCHWARTZ,S.A. Cryptease:Pleasant Change from the Business of Computing. PCM, 2(12):69-70, June 1985.

[17] "Frequência Ampliada" Nota na Dados e Idéias, 14(128), Fev. 1989.

[18] MAHLMEISTER, A.L. Bancos na era do satélite. Dados e Idéias, 14(128):36-37, Fev. 1989.

[19] Nota na Dados e Idéias, 14(127):10, Jan. 1989.

[20] FONSECA,M. Mudança estética, melhora de serviço. Dados e Idéias, 14(127):22, Jan. 1989.

[21] BRICKELL, E.F.;LEE, P.J.;YACOBI,Y. Secure Audio Teleconference. Advances in Cryptology - Crypto 87 Proceedings Aug. 1987, Berlin pags 16-20

[22] MASUDA, Y. A Sociedade da Informação. Editora Rio, 1982, 210 pags.

[23] LEMPEL, A. Cryptology in Transition. ACM Computing Surveys. New York, 11(4):285-303, Dec. 1979.

[24] SINKOV,A. Elementary cryptanalysis;a mathematical approach. [New York,Random House,c1968] 189 p.

[25] DENNING, D. E. R. Cryptography and Data Security Addison-Wesley Publishing Company, 1983. 400 pags.

[26] LUCCHESI, C. L. Introdução à Criptografia Computacional. Editora da Unicamp, Campinas, 1986. 132 pags.

[27] CHAMBERS, W.G. Clock-controlled shift registers in binary sequence generators. IEE Proceedings. Vol 135 Part E N. 1, January 1988 pags 17-24.

[28] SILVA FILHO, Joel Guilherme. Gerenciamento de chaves em sistemas criptográficos. IN: CONGRESSO NACIONAL DE INFORMÁTICA, 19. Rio

[29] MERKLE, Ralph C. Secure communications over insecure channels. Communications of the ACM, 21(4):294-9, Apr. 1978.

[30] DIFFIE, W. HELLMAN, M. New Directions in Cryptography. IEE Transactions on Information Theory, 22(6):644-54, Nov. 1976.

[31] RIVEST, R. L. et all. A method for obtaining digital signatures and public-key criptosystems. Communications of the ACM, 21(2):120-6, Feb. 1978.

[32] TERADA, R. Criptografia, uma chave de alta segurança. PC Mundo. Rio de Janeiro, CWB, 5(42):20-21, Jan. 1989.

[33] MULLER-SCHLOER,Christian A Microprocessor-based Cryptoprocessor. IEEE Micro, 3(5):5-15, October 1983.

[34] OLIVEIRA, Aldner Peres de. Estado da Arte em Criptografia de Chave Pública baseada no "Problema da Mochila", [Rio de Janeiro] 1988. 300p. Tese - Universidade Federal do Rio de Janeiro, COPPE.

[35] DESMEDT,Y. et all. How iterative transformations can help to crack the Merkle-Hellman cryptographic scheme. Electronics Letters, 18(21):910-1, Oct 1982.

[36] LAMPORT, L. Password Authentication with Insecure Communication. Communications of the ACM. New York, 24(11):770-2, Nov. 1981.

[37] BOOTH, K. Authentication of signatures using public key encryption. Communications of the ACM. New York, 24(11):772-74, Nov. 1981.

[38] NEEDHAM, R. SCHROEDER, M. Using Encryption for Authentication in Large Networks of Computers. Communications of the ACM. New York, 21(8):993-9, Dec. 1978.

[39] DEMILLO,Richard & MERRIT,Michael Protocols for Data Security. Computer, 16(2):39-51, February 1983.

[40] AKL,Selim G. Digital Signatures: A Tutorial Survey. Computer, 16(2):15-24, Feb. 1983.

[41] DENNING,Dorothy E. Protecting Public Keys and Signature Keys. Computer, 16(2):27-35, February 1983.

[42] DAVIES,Donald W. Applying the RSA Digital Signature to Eletronic Mail. Computer, 16(2):55-62, February 1983.

[43] BRICKELL,Ernest F.;ODLYZKO,Andrew M. Cryptanalysis: A Survey of Recent Results. Proceedings of the IEEE. 76(5):578-93, May 1988.

[44] READ,Richard Data Security - a question of control. Communications Systems Worldwide. pags 42-45, February, 1989.

[45] DIFFIE,Whitfield. The First Ten Years of Public-key Cryptography. Proceedings of the IEEE. 76(5):560-77, May 1988.

[46] L'Intelligence artificielle aux Etats-Unis, Athena, 24, Oct. 1986.

[47] SLAGLE,J. Artificial Intelligence: The heuristic Programming Aproach. McGraw-Hill, 1971, 196 págs.

[48] LUCENA, C. P. Inteligência Artificial. Análise & Conjuntura, 2(3):182-97, Belo Horizonte, 1988.

[49] PEARL,J. Heuristics. Addison-Wesley Publishing Company, 1984, 382 págs.

[50] Unicamp fará chip 36 vezes mais rápido do que o normal. Jornal do Brasil, 7/7/1989.

[51] KATZAN Jr., H. Segurança de Dados em Computação. Livros Técnicos e Científicos Editora S. A., Rio de Janeiro, 1977. 136 pags.

[52] DENNING,D.E.;DENNING,P.J. Data Security. ACM Computing Surveys. 11(3):227-49, Sept. 1979.

[53] PARKER,D. Nobody Knows How Much Computer Crime There Is. Computerweek 7(18):22-23, May 7, 1984.

[54] LENNON,R.E. et al. Cryptography in data processing. Data Processing Guildfold, Surey, UK, Butterworths, 26(7):36-38, Sept. 1984.

[55] JACKSON,K. Tandata Smart Card System. Computer Fraud Secur. Bulletin 10(12):21-23, Oct. 1988.

[56] TANENBAUM, Andrews. Computer Networks. Prentice-Hall, Englewood Cliffs, New Jersey, 1981, 517 pags.

[57] Burroughs Network Architecture (BNA) Architectural Description. Reference Manual Vol. 1, U.S.A. April 1981.

[58] MENDES, L. Os dados para sempre. PC Mundo. Rio de Janeiro, CWB, 5(42):16-19, Jan. 1989.

[59] LEGG,G. Encryption Software Guards Valuable Data. EDN 28(14):258-62, July 7, 1983.

[60] MCLELLAN,V. Electronic Auditor: a 'smarter' safeguard. Digitasl Review 4(15):78, Aug. 3, 1987.

[61] JOHNSTON,R.E. Choosing Encryption:DES or Private Key? Infosystems, 32(7):36, July, 1985.

[62] GREEN-ARMYTAGE,J. BT Clues Up on Cryptographics. Computer Weekly, June 20, 1985.

[63] MELLEN, G. E. Cryptology, Computers and Common Sense. IN: AFIPS CONFERENCE PROCEEDINGS, New York, June 4-8, 1973. National computer conference and exposition. Montvale, NJ, AFIPS PRESS, 1973. v.42 p. 569-79.

[64] PARKER, D. B. Crime por Computador. Agents Editores Ltda., Rio de Janeiro, 1977. 259 pags.

[65] NYCUM, S.C. "Privacy in Electronic Funds Transfer, Point of Sale and Electronic Mail Systems in the Next Decade". pags 39-42. IN: HOFFMAN, L. J. (ed.) Computers and Privacy in the Next Decade. Academic Press, 1980, 215 pags.

[66] LOUREIRO, M. D. E agora? Joseh, Joseph ou José? Anais da Sucesu 87 pags 45-54.

[67] JØRGEN,Bull "Legal Protection of Computer Programs", pags 410-431. IN: Jon Bing and Knut S. Selmer (ed.) A Decade of Computers and Law, Oslo, Universitetsforlaget, 1980.

[68] BENSIMOL, C. Lei Aprovada, mas nem tudo resolvido. Dados e Idéias, 13(118):20-22, Mar. 1988.

[69] Mesa Redonda. Os Retoques Finais. INFO, 6(60):8-12, Jan. 1988.

[70] Os EUA recuam. INFO, 5(55):32-33, Ago. 1987.

[71] MORAES, L. F. A Passagem para a Clandestinidade. MicroSistemas. Rio de Janeiro, ATI, 7(78):8-10, Maio 1988.

[72] SWAINE, M. Taking a Pseudonym Can Prevent 'Dossier Society'. InfoWorld, 5(37):19, Sept. 12, 1983.

[73] RULE, J. et all. The Politics of Privacy. New American Library, New York, 1980. 212 pags.

[74] BLEKELI,Ragnar Dag "Framework for the Analysis of Privacy and Information Systems", pags 21-31, IN: Jon Bing and Knut S. Selmer (ed.) A Decade of Computers and Law, Oslo, Universitetsforlaget, 1980.

[75] SELMER, KNUT S. "Norwegian Privacy Legislation", pags 45-58. IN: Jon Bing and Knut S. Selmer (ed.) A Decade of Computers and Law, Oslo, Universitetsforlaget, 1980.

[76] Brasil. Leis, decretos, etc. Constituição [da] República Federativa do Brasil; promulgada em 5 de outubro de 1988. [São Paulo] Isto é, [1988] 42 p.

[77] RULE,J.;MCADAM,D.;STEAMS,L.;UGLOW,D. "Preserving Individual Autonomy in an Information-Oriented Society", pags 65-87. IN: HOFFMAN, L. J. (ed.) Computers and Privacy in the Next Decade. Academic Press, 1980, 215 pags.

[78] Documento/3 - "Sistemas de informatica e controle dos trabalhadores. O caso IBM e a sentença do juiz de Milão de 5 de dezembro de 1984", pags 119-126 IN: MAGGIOLLINI, P. "As negociações trabalhistas e a introdução de inovações tecnológicas na Europa". Ed. Vozes Ltda/IBASE 190 pags.

[79] FRIEDMAN, W. F. & CALLIMAHOS, L. D. Military Cryptanalytics, Part I, Vol 2., Aegean Park Press.

Bibliografia não Referenciada

DEGIOVANI, R. Os Vírus do Computador. MicroSistemas Rio de Janeiro, ATI, 7(79):6-7, Junho 1988.

EVANS Jr.,A. KANTROWITZ, W. A User Authentication Scheme not Requiring Secrecy in the Computer. Communications of the ACM. New York, 17(8):437-42, Aug. 1974.

FERRAZ, Inhaúma Neves. Classes de Equivalência de Sistemas Criptográficos Algébricos. Tese - IME. Janeiro de 1979.

FRAGA, J. POWELL, D. Introduction a la Securite: Confidentialite et Tolerance aux Instrus. Rapport de Recherche N. 84.013, Fevrier 1984, 58 pags.

HERLESTAM,Tore. Critical remarks on some public-key cryptosystems. BIT, Copenhagem K, Denmark, BIT Data, 18(4):493-96,1978.

HELLMAN, Martin E. An Extension of the Shannon Theory Approach to Cryptography. IEEE Transactions on Information Theory. 23(3):189-94, May 1977.

KONHEIN,Alan G. Cryptography: a primer. New York, Wiley & Sons, 1981. 432 pags.

KULLBACK,Solomon Statistical Methods in Cryptanalysis. Aegean Press, 1976.

MASUDA, Yatosi. Criptografia em Segurança de Arquivos Confidenciais,[Rio de Janeiro] 1973. 133p. Tese - Universidade Federal do Rio de Janeiro, COPPE.

MILLEN,J.K. & CLARK,S.C. & FREEDMAN,S.B. The Interrogator: Protocol Security Analysis. IEEE Transactions on Software Engineering. 13(2):274-288, February 1987.

MOORE,J.H. & SIMMONS,G.J. Cycle Structure of the DES for Keys Having Palindronic (or Antipalindronic) Sequences of Round Keys. IEEE Transactions on Software Engineering. 13(2):262-273, Feb 1987.

MOREIRA, M. E. A complicada proteção do bem intangível. PC Mundo. Rio de Janeiro, CWB, 5(42):19-20, Jan. 1989.

PIERCE, J. R. Symbols, Signals and Noise - the nature and process of communication. Hutchinson of London, 1962. 305 pags.

POPEK,G.J.;KLINE,C.S. Encryption and Secure Computer Networks. ACM Computing Surveys. 11(4):331-56, Dec. 1979.

POWELL, D. DESWARTE, Y. Study of Fault-Tolerant Techniques for Space Craft Data Handling: Candidate Solutions Proposed for OTV and DSO Missions. LAAS Report N. 85.227, October 15, 1985 - 16 pags.

PURDY, G. A High Security Log-in Procedure. Communications of the ACM. New York, 17(8):442-5, Aug. 1974.

RANEA, P. FRAY, J. Rapport Technique Saturne: Aspect Securite. Etude de la Dissemination et de l'Organisation de l'Arquivage. Rapport LAAS N. 87167, Juin 1987, 40 pags.

SALOMAA,A. & YU,S. On a Public-key Cryptosystem based on Iterated Morphisms and Substitutions. Theoretical Computer Science. 48(2-3):283-296, 1986.

SHAMIR,A. ZIPPEL,R.E. On the security of the Merkle-Hellman cryptographic scheme. IEEE Transactions on Information Theory, IT-26(3):339-40, May 1980.

ANEXO A - GLOSSÁRIO

ALGORITMO CRIPTOGRÁFICO - conjunto de normas segundo as quais um texto claro é transformado num texto incompreensível para quem não souber como fazer a transformação inversa. O mesmo que cifra.

ASSINATURA DIGITAL - O equivalente eletrônico a uma assinatura num papel. A assinatura digital deve ter as seguintes características: só uma "pessoa" pode produzi-la mas qualquer um pode reconhecê-la.

ATAQUE ATIVO - Ocorre quando o "inimigo" altera a mensagem que está sendo transmitida ou envia falsas mensagens fazendo-se passar por outrem.

ATAQUE PASSIVO - Corresponde à observação das mensagens transmitidas com o objetivo de ter acesso à informação sem, no entanto, alterá-las.

ATM - Automatic Teller Machine - Caixa Automático.

AUTENTICAÇÃO (DA MENSAGEM) - A garantia de que a mensagem não foi alterada, ou seja, não sofreu nenhum "ataque ativo". É o "selo de garantia" da genuinidade da mensagem.

AUTENTICAÇÃO (DO REMETENTE) - A garantia de que o remetente da mensagem é a pessoa que se identificou como tal e de que essa identidade não foi forjada por terceiros.

CHAVE CRIPTOGRÁFICA - Informação secreta disponível apenas ao remetente e ao destinatário do criptossistema. Pode também ser definida como o conjunto de parâmetros usados unicamente numa aplicação do algoritmo.

CHAVE PÚBLICA - chave usada em sistemas criptográficos assimétricos, em que a chave de decriptografar não pode ser descoberta a partir da chave de criptografar, sendo essa última mantida num diretório público disponível a qualquer usuário do sistema.

CIFRA - é um esquema universal que emprega uma série de transformações visando camuflar um texto. É uma técnica criptográfica (como, por exemplo, transposição e substituição) que atua sobre o texto claro sem se preocupar com sua estrutura linguística.

CIFRAMENTO BIT A BIT - quando a "unidade" de transformação utilizada pelo algoritmo é uma unidade da linguagem (um bit ou uma letra) sendo o i-ésimo componente da chave criptográfica aplicado ao i-ésimo componente do texto original.

CIFRAMENTO BLOCO A BLOCO - quando a "unidade" de transformação usada no algoritmo é um conjunto na linguagem original (56 bytes, uma frase etc.), sendo toda a chave criptográfica aplicada a cada bloco do texto.

CIFRAMENTO ENCADEADO - quando os elementos encriptados anteriormente fazem parte da chave que será usada no próximo elemento do texto.

CIFRAMENTO SÍNCRONO - quando o próximo elemento a ser cifrado é completamente independente dos que o precederam.

CIFRAR - o mesmo que criptografar.

CÓDIGO - é um tipo especial de substituição em que um "livro de códigos" é utilizado e palavras ou frases inteiras são substituídas por um código através do livro, que é usado

como um dicionário de sinônimos. O termo é algumas vezes usado como referência a qualquer tipo de cifra.

COINCIDÊNCIA - a repetição de elementos do texto (letras, dígitos, digrafos etc.) que aconteça numa mensagem ou entre mensagens [79].

CRIPTOANALISAR - aplicar a criptoanálise a (um criptograma).

CRIPTOANÁLISE - é definida como sendo os estudos feitos no sentido de se recuperar a mensagem criptografada sem que se tenha nenhuma informação 'a priori' de como recuperá-la, ou seja, é o estudo da resolução de um criptograma sem a posse da chave e muitas vezes sem o conhecimento do algoritmo utilizado.

CRIPTOGRAFAR - transformar texto claro em criptograma.

CRIPTOGRAFIA - arte de escrever em cifra ou em código. Conjunto de técnicas que permitem criptografar escritas.

CRIPTOGRAFIA COMPUTACIONAL - técnicas criptográficas aplicadas aos sistemas de informação. [51]

CRIPTOGRAMA - o texto criptografado de uma mensagem. Mensagem transformada, codificada.

CRIPTOLOGIA - é a ciência que engloba criptografia e criptoanálise.

CRIPTOSSISTEMA - é composto basicamente de dois elementos: um algoritmo criptográfico ou conjunto de regras fixas para transformação de mensagens e um conjunto de chaves criptográficas variáveis [54].

DECIFRAR - o inverso de criptografar, ou seja, transformar criptograma em texto claro.

DECRIPTAR - o mesmo que decifrar.

DES - Data Encryption Standard, algoritmo criptográfico adotado como padrão pelo National Bureau of Standards (NBS) americano para aplicações que não envolvam segurança nacional.

DIGRAFO - um par de letras [79].

ENCRIPTAR - o mesmo que criptografar.

FATOR TRABALHO - a quantidade de trabalho que o criptoanalista terá para resolver o criptograma.

INDICE DE COINCIDÊNCIA - é a razão entre o número observado de coincidências num dado texto e o número de coincidências esperado numa amostra de texto aleatória do mesmo tamanho [79].

PRIVACIDADE (DE DADOS) - É o direito de cada pessoa física ou jurídica sobre os dados a respeito próprio. Ela inclui: o direito de cada um saber o que existe de informação a seu respeito, aonde está armazenado e por quê; a garantia de que a informação está completa e correta; e a garantia de que esses dados serão usados apenas para o fim ao qual foram cedidos e que não serão divulgados sem autorização prévia.

REDUNDÂNCIA - excesso ou desperdício de sinais ou de signos na transmissão da mensagem, que serve, contudo, para neutralizar os efeitos do ruído no canal de comunicação [Aurélio].

ROTOR - um disco projetado para rodar numa máquina de cifrar de maneira a produzir variações em algum texto ou elemento de chaveamento [79].

RSA - As iniciais de Rivest, Shamir e Adleman.

Algoritmo de chave pública desenvolvido pelos três.

SEGURANÇA (DE DADOS) - Tecnologia que garanta a confiabilidade dos dados do sistema. Pode ser definida como a proteção de dados contra a revelação acidental ou intencional a pessoas não autorizadas, e contra alterações não autorizadas [51].

SUBSTITUIÇÃO - algoritmo criptográfico em que os caracteres do texto claro são substituídos por outros segundo uma chave.

TESTE DE COINCIDÊNCIA - teste estatístico aplicado a duas mensagens criptografadas para determinar se ambas foram criptografadas pela mesma sequência de alfabetos (se houve substituição polialfabética com a mesma chave).

TEXTO CIFRADO - o mesmo que criptograma.

TEXTO CLARO - a mensagem original, em linguagem corrente, que se deseja camuflar.

TEXTO CORRIDO - o mesmo que texto claro.

TRANSPOSIÇÃO - algoritmo criptográfico em que os caracteres da mensagem original são remanejados em sua ordem seguindo um plano pré-estabelecido.

UNICITY DISTANCE - o tamanho mínimo que um criptograma deve ter para que sua solução seja única para quem não conhece a chave utilizada.

ANEXO B - FREQÜÊNCIA DAS LETRAS NA LÍNGUA PORTUGUESA

As informações constantes desse anexo foram retiradas do livro "Military Cryptanalytics" de William Friedman e Lambros Callimahos [79]. São estatísticas relativas a relatórios militares. Caso se deseje usar estatísticas desse tipo em outras aplicações, é aconselhável recalculá-las em função de textos no mesmo formato do que se pretende atacar.

1.1) Frequência absoluta das letras em portugues, em ordem alfabética, baseado num texto de 45106 letras.

A.... 5362	G.... 724	L.... 1245	Q.... 348	V.... 737
B.... 470	H.... 304	M.... 1699	R.... 3292	W.... 24
C.... 2285	I.... 3314	N.... 2912	S.... 3409	X.... 166
D.... 1900	J.... 160	O.... 5001	T.... 2679	Y.... 22
E.... 5441	K.... 17	P.... 1377	U.... 1491	Z.... 207
F.... 520				------
				45106

1.2) Índice de Coincidência = 1.94 (vide glossário)

1.3) Distribuição de freqüência das letras baseado no texto de 45106 letras, reduzidas a 1000, arrumadas de acordo com sua freqüência relativa.

E.... 121	N.... 65	U.... 33	F.... 11	X.... 4
A.... 119	T.... 59	P.... 30	B.... 10	J.... 3
O.... 111	C.... 51	L.... 28	Q.... 8	W.... 1
S.... 76	D.... 42	V.... 16	H.... 7	Y.... -
I.... 73	M.... 38	G.... 16	Z.... 5	K.... -
R.... 73				

1.4) Percentagem de vogais, consoantes mais frequentes,

mais ou menos frequentes e menos frequentes. Percentagem das 8 letras mais frequentes:

Vogais A, E, I, O, U e Y = 45,8%

Consoantes mais frequentes N, R e S = 21,3%

Consoantes mais ou menos frequentes C, D, L, M, P e T = 24,8%

Consoantes menos frequentes B, F, G, H, J, K, Q, V, W, X e Z = 8.1%

8 letras mais frequentes (em ordem decrescente de freqüência) E, A, O, S, I, R, N e T = 69,7%

1.5) Frequência absoluta de iniciais em 7058 palavras num texto claro, ordenadas por freqüência.

P.... 847	M.... 405	I.... 264	B.... 113	Z.... 14
C.... 731	T.... 348	F.... 222	G.... 111	W.... 11
E.... 608	R.... 316	Q.... 222	J.... 92	K.... 7
S.... 601	N.... 299	O.... 187	U.... 77	Y.... 4
A.... 597	V.... 271	L.... 143	H.... 60	X.... 2
D.... 506				----
				7058

2.1) Distribuição de freqüência de digrafos baseado no texto de 45106 letras reduzido a 5000 digrafos.

Primeira Letra \ Segunda Letra	A	B	C	D	E	F	G	I	L	M	N	O	P	Q	R	S	T	U	V
A	11	11	52	60	15	9	14	18	38	36	56	49	23	8	68	72	22	8	16
B	11			1	10			5	2	1		9			9	2	1	2	
C	60		2		30			39	5		1	85			7		8	12	
D	45				61			33		1		61			2	1	1	5	
E	15	5	48	22	11	11	23	27	31	44	97	6	18	6	76	95	20	7	12
F	9				14			13	1			15			2			3	
G	15				14			4	1		1	14			14			15	
H	10				8			3				11						1	
I	42	3	34	31	6	7	9	1	16	22	53	26	5	2	25	39	27	2	10
L	24	1	4	4	24	1	5	21	2	4	2	14	4	2	1	4	7	6	2
M	41	10	3	4	51	1		26	1	2	1	16	15	1	3	5	2	6	2
N	31		29	35	14	7	8	18				25	1			19	114	4	4
O	21	9	32	25	27	10	7	20	20	36	79	5	35	8	71	85	18	12	22
P	26		2		25			2	4		1	60	1	1	28	1	1	3	
Q					1													37	
R	75	2	14	9	86	3	7	46	2	18	8	34	7	3	11	8	18	4	6
S	41	6	22	10	62	6	3	23	3	12	5	23	35	7	4	40	47	18	5
T	65		1	1	69	1		26			1	88			33		1	13	
U	22	5	5	7	26	1	4	18	14	11	17	2	4		9	6	11		1
V	11				37			23				9			1				

2.2) Índice de Coincidência = 5,68

2.3) Abaixo os 91 digrafos que correspondem a 75% do texto, de acordo com a tabela do item 2.1, em ordem de

freqüência.

NT... 114	SE... 62	IA.. 42	IC... 34	PA... 26	ED... 22
EN... 97	DO... 61	MA... 41	TR... 33	TI... 26	AT... 22
ES... 95	DE... 61	SA... 41	DI... 33	PE... 25	UA... 22
TO... 88	AD... 60	SS... 40	OC... 32	IR... 25	OA... 21
RE... 86	PO... 60	CI... 39	EL... 31	OD... 25	LI... 21
CO... 85	CA... 60	IS... 39	ID... 31	NO... 25	OL... 20
OS... 85	AN... 56	AL... 38	NA... 31	LA... 24	ET... 20
ON... 79	IN... 53	VE... 37	CE... 30	LE... 24	OI... 20
ER... 76	AC... 52	QU... 37	NC... 29	AP... 23	NS... 19
RA... 75	ME... 51	OM... 36	PR... 28	EG... 23	SU... 18
AS... 72	AO... 49	AM... 36	IT... 27	VI... 23	RT... 18
OR... 71	EC... 48	2505(*)	OE... 27	SO... 23	EP... 18
TE... 69	ST... 47	ND... 35	EI... 27	SI... 23	UI... 18
AR... 68	RI... 46	OP... 35	UE... 26	OV... 22	----
TA... 65	DA... 45	SP... 35	MI... 26	SC... 22	3755
1224(*)	EM... 44	RO... 34	IO... 26	IM... 22	

2.4) Digrafos mais frequentes, cujo inverso também é frequente, a partir da tabela de 5000 digrafos.

ES... 95	SE... 62	OR... 71	RO... 34	ME... 51	EM... 44
RE... 86	ER... 76	CA... 60	AC... 52	EC... 48	CE... 30
CO... 85	OC... 32	AD... 60	DA... 45	MA... 41	AM... 36
RA... 75	AR... 68	PO... 60	OP... 35	CI... 39	IC... 34
AS... 72	SA... 41	AN... 56	NA... 31	DI... 33	ID... 31

2.5) Digrafos frequentes cujo inverso é raro.

NT... 114	TN... 1	ST... 47	TS... 0	ND... 35	DN... 0

2.6) Letras dobradas.

SS... 40	EE... 11	OO... 5	LL... 2	II... 1	TT... 1
AA... 11	RR... 11	CC... 2	MM... 2	PP... 1	

2.7) Os 20 digrafos que mais aparecem como início de palavra em 6803 palavras.

CO... 464	DE... 259	PR... 169	MA... 130	MI... 105
PO... 386	QU... 220	PA... 143	PE... 122	NO... 104
SE... 333	IN... 188	NA... 133	VE... 115	TR... 103
RE... 276	ES... 173	TE... 132	ME... 111	DI... 102

3.1) Os 59 trigrafos que aparecem mais de 100 vezes num

texto de 45106 letras.

ENT.. 474	EST.. 186	ADE.. 143	SPO.. 130	DES.. 123	EDI.. 107
NTO.. 457	ACA.. 182	STA.. 143	ADA.. 129	ECO.. 121	ASE.. 105
ONT.. 303	RES.. 181	ICA.. 142	TRA.. 129	ODE.. 118	ITO.. 104
NTE.. 284	QUE.. 172	OCO.. 140	NDO.. 127	ECE.. 115	ELE.. 103
CON.. 255	NTA.. 167	ARA.. 136	ENC.. 126	NCI.. 114	ERI.. 103
PON.. 236	POR.. 159	DOS.. 134	OSE.. 126	REC.. 113	PRO.. 102
CAO.. 227	ACO.. 158	OES.. 134	ARE.. 125	PAR.. 112	AME.. 101
ADO.. 211	COM.. 154	IDA.. 133	ESE.. 124	ESS.. 110	OSS.. 101
MEN.. 205	ERE.. 150	TER.. 132	OVE.. 124	DAD.. 109	IME.. 100
TOS.. 191	CIA.. 145	OPO.. 130	SSA.. 124	ORE.. 108	

3.2) Os 19 trigrafos que aparecem 50 vezes ou mais no início de 6803 palavras.

CON.. 224	EST.. 105	NAO.. 86	TRA.. 66	IND.. 52
PON.. 213	PAR.. 93	QUA.. 83	MIL.. 61	RES.. 52
COM.. 136	PRO.. 93	DES.. 71	REF.. 56	REC.. 51
QUE.. 109	POR.. 88	SER.. 70	VEX.. 53	

4) Os 38 tetragrafos que aparecem mais de 50 vezes num texto de 45106 letras.

ONTO.. 233	NCIA.. 95	COES.. 73	ADOS.. 60	RENT.. 52
PONT.. 221	PORT.. 87	IDAD.. 71	IMEN.. 60	TELE.. 52
MENT.. 183	DADE.. 86	CENT.. 70	CONS.. 58	EGRA.. 51
ENTO.. 173	ESTA.. 85	INTE.. 70	NTES.. 58	NFOR.. 51
ENTE.. 147	ENCI.. 83	CONT.. 68	ANDO.. 57	OPON.. 51
ACAO.. 142	SPON.. 83	FORM.. 67	ANTE.. 57	LEGR.. 50
NTOS.. 141	AMEN.. 81	OCON.. 66	ORMA.. 54	
ENTA.. 97	PARA.. 81	ELEG.. 61	VEXA.. 54	

5) Tamanho médio das palavras em portugues = 6,48 letras.