Criptografia

A criptografia permite implementar uma série de pontos de segurança indispensáveis para uma comunicação segura em uma rede computadores publica, como a Internet.

De um modo geral, os seguintes requisitos são necessários à implementação de aplicações seguras na Internet:

Confidencialidade: é a garantia de que a comunicação entre o transmissor e receptor seja implementada de forma sigilosa, sem o conhecimento de terceiros. Por exemplo, quando utilizamos o Internet Banking, a comunicação entre o cliente e o servidor do banco deve ser feita de forma confidencial.

Integridade: é a garantia de que a informação trocada entre o transmissor e receptor não será alterada. Por exemplo, quando enviamos um email, nada impede que alguém intercepte a mensagem e altere seu conteúdo.

Autenticidade: é a garantia de que a pessoa que realizou a operação é de fato aquela que diz ser. Por exemplo, quando recebemos um email de alguém, não temos qualquer garantia de que a mensagem foi enviada pela pessoa que consta no cabeçalho do email, pois essa informação é facilmente forjada

Não repúdio: é a possibilidade de provar a terceiros a autoria de uma operação após a mesma ter ocorrido. Por exemplo, quando enviamos um email, a pessoa que recebe a mensagem não tem como provar quem foi a pessoa que o enviou.

A criptografia moderna está baseada no conceito de chaves criptográficas, que consistem em uma seqüência de bits.

Métodos de criptografia

Os métodos de criptografia têm sido divididos em duas categorias: as cifras de substituição e as de cifras de transposição.

Cifras de Substituição

Cada letra ou grupo de letras é substituído por outra letra ou grupo de letras. Por exemplo:

Cada uma das 26 letras do alfabeto tem seu correspondente em outra letra.

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

QWERTYUIOPASDFGHJKLZXCVBNM

Esse sistema é conhecido como substituição mono alfabética, sendo a chave a string de 26 letras correspondente ao alfabeto.

Substituindo as letras da palavra "atacar" pela correspondente resultaria em "qzqeqk".

Todavia em um texto pequeno, a cifra poderia ser descoberta facilmente. Pois o intruso começaria contando as letras mais freqüentes do texto cifrado e depois disso ele atribuiria a letra "a" a letra mais comum. Em seguida ele poderia verificar os trigamas e encontrar um no formato gXi, o que poderia sugerir que X poderia ser "u".

Embora a criptografia moderna utilize as mesmas idéias básicas da substituição tradicional, sua ênfase atual é diferente, ela tem como objetivo tornar o algoritmo complexo e emaranhado para que o intruso não seja capaz de obter qualquer sentido da mensagem.

Cifras de Transposição

Muda a ordem das letras. Por exemplo:

Para cifrar o texto "transferir um mil"

A cifra se baseia em uma chave que é uma palavra ou frase. No exemplo a palavra é "disco".

A chave servirá de apoio para enumerar as colunas.
Palavra chave = "d i .s c o"
(2 3 5 1 4)
A mensagem é escrita abaixo da chave, de 5 em 5 letras (que é a mesma quantidade de letras da chave).

Palavra chave = "d i .s c o"
(2 3 5 1 4) (2 3 5 1 4) (2 3 5 1 4)
t .r .a. n. s f .e r .i .r u m m i l

O texto é lido na vertical, conforme ordenado pela palavra chave. Resultando em "niitfuremsrlarm"

Chave

A chave consiste em uma string que pode ser alterada sempre que necessário. Desse modo o algoritmo de criptografia pode ser conhecido. Quando o algoritmo se torna público, vários especialistas tentam decodificar o sistema. Se após alguns anos nenhum deles conseguirem a proeza, significa que o algoritmo é bom.

Tamanho da chave

O fator de trabalho para decodificar o método através de uma pesquisa no espaço da chave é exponencial em relação ao tamanho da chave. Por exemplo, uma chave com um tamanho de dois bytes significa que existem 65536 possibilidades, e um tamanho de chave de quatro bytes significa 4294967296 de possibilidades, portanto quanto maior for a chave, maior será o fator de trabalho com que o intruso terá de lidar. Mas é importante lembrar que o algoritmo usado terá influencia direta no tamanho da chave.

Tipos de Chave

O tipo de chave usada depende do tipo da criptografia usada. Existem dois tipos de criptografia:

Criptografia simétrica, que usa uma chave privada.

Criptografia de chave pública, que usa um par de chaves, conhecida com chaves publica e privada.

Chaves simétricas

No esquema de chaves simétricas, também conhecidas como chaves secretas, o transmissor e o receptor devem possuir a mesma chave criptográfica. O transmissor utiliza sua chave para criptografar a mensagem, que pode então ser enviada. O receptor utiliza a mesma chave para descriptografar a mensagem recebida. Caso a mensagem seja interceptada, não será possível ler o seu conteúdo, a não ser que se descubra a chave secreta utilizada.
A vantagem deste método é ser bastante eficiente em relação ao tempo de processamento, ou seja, o tempo gasto para codificar e decodificar mensagens tem como principal desvantagem a necessidade de utilização de um meio seguro para que a chave possa ser compartilhada entre pessoas ou entidades que desejem trocar informações criptografadas.

O sistema de criptografia simétrico, como visto, possui uma chave única tanto para a criptrografica do transmissor, quanto a descriptografia do receptor. Estes sistemas podem ser divididos em duas categorias: stream ciphers e block ciphers.

Sistemas de criptografia

O stream ciphers operam bit a bit sobre o fluxo de dados (stream) de entrada, combinando-o com outro fluxo de dados que constitui a chave simétrica, que é normalmente uma seqüência pseudo-

aleatória derivada de uma chave de tamanho fixo. Este algoritmo é também conhecido como cifradores bit a bit 1 ou de caracteres (stream ciphers).

O algoritmo cifradore de bloco (block ciphers) atua sobre um bloco de dados de tamanho fixo. Os algoritmos simétricos modernos são, na maioria dos casos, "embaralhadores" de bits derivados a partir de operações matemáticas básicas e funções lógicas tipo Ou-Exclusivo, deslocamento de bits, substituições e permutações simples. Apesar de, teoricamente, ser possível implementar um algoritmo através de operações mais elaboradas, o enfoque acima é mais eficiente em termos de desempenho e implementação em hardware e software.

O objetivo dos algoritmos simétricos conhecidos é obter os princípios de confusão e difusão. O princípio de confusão obscurece a relação entre o texto claro e o cifrado, e é normalmente obtido através de técnicas de substituição. Já a difusão elimina a redundância do texto claro, espalhando-a sobre o texto cifrado.

O mecanismo mais usado para implementar difusão é através de permutações e também através de substituições.

Os algoritmos cifradores de caracteres foram bastante usados antes da existência de computadores. Atualmente, em sistemas de criptografia computacionais, cifradores de blocos de 64 bits são os mais comumente utilizados. Este tamanho de bloco é considerado grande o suficiente para garantir a segurança do algoritmo, e pequeno o bastante para os computadores atuais operarem. Contudo, devido ao avanço nas técnicas de criptoanálise e a presença de arquiteturas computacionais de 64 bits, cifradores de blocos de 128 bits têm sido propostos.

O algoritmo AES (Advanced Encryption Standard), a ser proposto pelo NIST 2 em substituição ao padrão DES, requer operação sobre blocos de 128 bits.

O algoritmo cifrador de blocos mais conhecido e usado é o DES (Data Encryption Standard). Devido à sua importância para o desenvolvimento da criptografia moderna, o DES é também a referência clássica no assunto.

Algoritmo Digital Encryption Standard - DES

Pesquisadores da IBM, no final da década de 1970, decidiram desenvolver um novo algoritmo de cifragem para computadores e criaram o um esquema chamado Lúcifer, um algoritmo, inventado pelo criptográfico Horst Feistel. Em conjunto com National Security Agency (NSA), agencia esta responsável pelos projetos de dados secretos do governo Norte Americano. O fruto deste trabalho foi o DES – Digital Encryption Standard.

O DES é basicamente um cifrador de substituição que utiliza um bloco de 64 bits. Ele possui uma chave de 56 bits e mais oito bits de paridade, completando os 64 bit. O DES, com uma chave de 56 bit nos possibilita ter **72057594037927936** chaves. O DES executa uma série de transposições, substituições, e operações de recombinação em blocos de dados de 64 bits.

Inicialmente, os 64 bits de entrada sofrem uma transposição e são colocados em uma função usando tabelas estáticas de transposição e substituição (conhecidas como caixas-P e caixas-S).

Exemplo do funcionamento de uma caixa-P:

P significa "permuta", então se forem designados 8 bits de entrada "01234567" será efetuada uma transposição, que irá mudar a ordem dos números. A saida dessa caixa-P será "36071245".

Exemplo do funcionamento de uma caixa-S:

S representa a substituição dos números por outros números. Supondo que o número 0 seja substituido por 2, o 1 por 4, o 2 por 5, o 3 por 0, o 4 por 6, o 5 por 7, o 6 por 1 e o 7 por 3.

Para a entrada "01234567" a substituição resultaria em "24506713".

Os estágios são parametrizados por diferentes funções da chave. A função consiste em 4 etapas, que são executadas em sequência.
Primeiro é feita uma transposição de 64 bits dos dados.
Os 16 estágios restantes são parametrizados por diferentes funções da chave.
O penúltimo estágio troca os 32 bits da esquerda pelos 32 bits da direita (os 32 bits representam uma divisão dos 64 bits).
O último estágio é o inverso da primeira transposição
Em cada uma das 16 iterações, é utilizada uma chave específica. Antes de se iniciar o algoritmo, uma transposição de 56 bits é aplicada a chave. Antes de cada iteração, a chave é particionada em duas unidades de 28 bits, sendo que cada uma delas é roteada para a esquerda por um determinado número de bits. Em cada rodada, um subconjunto de 48 bits dos 56 bits é extraído e permutado. O algoritmo então executa a transposição final e gera 64 bits.

Outros algoritmos de chave simétrica

3DES - Triple DES - O 3DES é uma simples variação do DES, utiliza o ciframento DES três vezes sucessivamente, podendo empregar um versão com duas ou com três chaves diferentes. É seguro, porém muito lento para ser um algoritmo padrão. O tamanho da chave pode ser de 112 ou 168

RC2 - Projetado por Ron Rivest (o R da empresa RSA Data Security Inc.) e utilizado no protocolo S/MIME, voltado para criptografia de e-mail corporativo. Possui chave de tamanho variável, entre 8 a 1024 bits. Rivest também é o autor do RC4, RC5 e RC6.

IDEA - O International Data Encryption Algorithm foi criado em 1991 por James Massey e Xuejia Lai e possui patente da suíça ASCOM Systec. O algoritmo é estruturado seguindo as mesmas linhas gerais do DES. Mas na maioria dos microprocessadores, uma implementação por software do IDEA é mais rápida do que uma implementação por software do DES. O IDEA é utilizado principalmente no mercado financeiro e no PGP, o programa para criptografia de e-mail pessoal mais disseminado no mundo. Possui um chave de 128 bits.

Blowfish - Algoritmo desenvolvido por Bruce Schneier, que oferece a escolhe entre maior segurança ou desempenho através de chaves de tamanho variável, entre 32 a 448 bits.

Chave assimétrica

O principal problema do uso de chave simétrica é o transporte da chave,pois tanto o receptor quanto o transmissor devem ter a mesma chave. Desta forma temos o problema da transmissão da chave, uma vez que alguém intercepte a transferência de chave, terá total conhecimento da mensagem. Com uso de chave assimétricas, este problema ser resolvido.

No esquema de chaves assimétricas, também conhecidas como chaves públicas, o transmissor possui um par de chaves e o receptor outro. O par de chaves é formado por uma chave privada e uma pública. A chave privada é de conhecimento apenas do dono do par, enquanto a chave pública é amplamente divulgada, ou seja, transmitida e pode ser consultada por qualquer usuário.

Quando o transmissor deseja enviar uma mensagem confidencial, ele utiliza a chave pública do receptor que está disponível na rede. O receptor, ao receber a mensagem, utiliza sua chave privada para abrir a mensagem. Caso alguém intercepte a mensagem, não será possível a sua leitura, pois apenas a chave privada do receptor pode ler uma mensagem criptografada com sua chave pública.

Nesse método as chaves de criptografia e decriptografia são diferentes. Quando uma chave criptografa um dado, a outra pode decriptografá o dado recebido.

O usuário tem duas chaves, uma chave pública que é usada por todo mundo que queira enviar mensagens a ele, e a chave privada que o usuário utiliza para decriptografar as mensagens recebidas. Deste modo o usuário divulga sua chave publica para todos que ele queira que recebam e leiam a sua mensagem, e mantém a sua chave privada em sigilo. Um outro modo seria ao contrário, criptografar com chave privada e decriptografar com chave pública. Nesse caso, não existe uma questão de segurança, mas de identificação, certificando a origem do dado, uma vez que somente uma chave privada poderia escrever esta mensagem.

RSA

O RSA é um algoritmo assimétrico que possui este nome devido a seus inventores: Ron Rivest, Adi Shamir e Len Adleman, que o criaram em 1977 no MIT. É, atualmente, o algoritmo de chave pública mais utilizado, além de ser uma das mais poderosas formas de criptografia de chave pública conhecidas até o momento. O RSA utiliza números primos.

A premissa por trás do RSA é que é fácil multiplicar dois números primos para obter um terceiro número, mas muito difícil recuperar os dois primos a partir daquele terceiro número. Isto é conhecido como fatoração. Por exemplo, os fatores primos de 3.337 são 47 e 71. Gerar a chave pública envolve multiplicar dois primos grandes; qualquer um pode fazer isto. Derivar a chave privada a partir da chave pública envolve fatorar um grande número, o RSA usa números maiores que 100 dígitos. Se o número for grande o suficiente e bem escolhido, então ninguém pode fazer isto em uma quantidade de tempo razoável. Assim, a segurança do RSA baseia-se na dificuldade de fatoração de números grandes. Deste modo, a fatoração representa um limite superior do tempo necessário para quebrar o algoritmo.

Segundo seus pesquisadores, a fatoração de um número de 200 dígitos requer 4 milhões de anos para ser processada; fatorar um número de 500 dígitos exige 10^{25} anos. Mesmo que os computadores se tornem mais velozes, muito tempo irá passar até que seja possível fatorar um número de 500 dígitos, e até lá poderão escolher a fatoração de um número ainda maior.

Diffie-Hellman - Baseado no problema do logaritmo discreto, e o criptosistema de chave pública mais antigo ainda em uso. O conceito de chave pública, foi introduzido pelos autores deste criptosistema em 1976. Contudo, ele não permite nem ciframento nem assinatura digital. O sistema foi projetado para permitir a dois indivíduos entrarem em um acordo ao compartilharem um segredo tal como uma chave, muito embora eles somente troquem mensagens em público.

ElGamal - É um algoritmo de chave pública utilizado para gerenciamento de chaves. Sua matemática difere da utilizada no RSA, mas também é um sistema comutativo. O algoritmo envolve a manipulação matemática de grandes quantidades numéricas. Sua segurança advém de algo denominado problema do logaritmo discreto. Assim, o ElGamal obtém sua segurança da dificuldade de se calcular logaritmos discretos em um corpo finito, o que lembra bastante o problema da fatoração

Curvas Elípticas - Em 1985, Neal Koblitz e V. S. Miller propuseram de forma independente a utilização de curvas elípticas para sistemas criptográficos de chave pública. Eles não chegaram a inventar um novo algoritmo criptográfico com curvas elípticas sobre corpos finitos, mas implementaram algoritmos de chave pública já existentes, como o algoritmo de Diffie e Hellman, usando curvas elípticas. Assim, os sistemas criptográficos de curvas elípticas consistem em modificações de outros sistemas (o ElGamal, por exemplo), que passam a trabalhar no domínio das curvas elípticas, em vez de trabalharem no domínio dos corpos finitos. Eles possuem o potencial de proverem sistemas criptográficos de chave pública mais segura, com chaves de menor tamanho. Muitos algoritmos de chave pública, como o Diffie - Hellman, o ElGamal e o Schnorr podem ser implementados em curvas elípticas sobre corpos finitos. Assim, fica resolvido um dos maiores

problemas dos algoritmos de chave pública: o grande tamanho de suas chaves. Porém, os algoritmos de curvas elípticas atuais, embora possuam o potencial de serem rápidos, são em geral mais lentos do que o RSA.

Sistemas simétricos X sistemas assimétricos

Nos sistemas simétricos, uma chave secreta deve ser estabelecida para cada par de entidades que queira comunicar-se entre si. Tais chaves devem ser transmitidas de maneira segura, o que pode ser impossível, caso nenhuma comunicação segura tenha sido feita a priori.

Os sistemas assimétricos podem minimizar o problema acima, pois cada entidade interessada em receber mensagens secretas deve apenas publicar sua chave pública, em um diretório de comum acesso. Assim, para enviar uma mensagem sigilosa a esta entidade, basta obter sua chave pública via um canal de comunicação qualquer (inseguro). Somente a entidade que possuir a chave secreta correspondente será capaz de decifrar a mensagem em questão.

Por outro lado, os sistemas assimétricos não são usados diretamente para cifrar mensagens pelos seguintes motivos:

São bastante lentos se comparados aos sistemas simétricos, cerca de 1.000 vezes mais lentos. Como os requerimentos de banda crescem proporcionalmente à velocidade das CPUs é improvável que estes sistemas possam ser usados para cifrar todo o fluxo de dados entre aplicações.

Os sistemas assimétricos são mais vulneráveis a ataques de texto conhecido e escolhido.

O acesso a textos cifrados conhecidos é extremamente facilitado pelo fato da chave de ciframento ser pública. Este problema se agrava quando o número de mensagens possíveis é reduzido. Neste caso, basta cifrar todas as mensagens possíveis com a chave pública disponível para identificar o conteúdo da mensagem enviada.

Os sistemas simétricos e assimétricos são na verdade complementares. Os de chave pública são normalmente empregados no processo de autenticação e ciframento de chaves de sessão, que por usa vez são usadas por algoritmos simétricos para cifrar o fluxo de dados entre aplicações. Sistemas de criptografia que se utilizam destas duas abordagens são freqüentemente denominados sistemas híbridos.

Criptografia Simétrica.	Criptografia Assimétrica.
Rápida.	Lenta.
Gerência e distribuição das chaves é complexa.	Gerência e distribuição simples.
Não oferece assinatura digital	Oferece assinatura digital.

Tipos de ataques

Infelizmente, até mesmo os sistemas de criptografia mais sofisticados podem, eventualmente, ser comprometidos. Como tudo que é secreto ou pode ser secreto, tem grande atração de pessoas que querem descobrir o segredo, por isso, existe uma grande variedade de ataques aos quais os mesmos estão sujeitos.

Para analisar a pior situação, normalmente pressupõe-se que o oponente tenha pleno acesso ao meio de comunicação e conhecimento do sistema de criptografia a ser atacado. Veremos os tipos básicos de ataques, bem como à terminologia usada para descrevê-los.

Ataques de força bruta

Este ataque consiste em tentar todas as chaves possíveis para decifrar uma determinada mensagem. Neste caso, a segurança do algoritmo de criptografia depende do tamanho da chave implementada.

Criptoanálise

Criptoanálise é a ciência de decifrar um texto sem ter acesso direto à chave de deciframento ou ciframento. O objetivo da criptoanálise é obter, a partir de um texto cifrado, o texto claro original ou até mesmo a chave usada para cifrá-lo. Normalmente é explorada uma fraqueza matemática do algoritmo e modo de operação em uso, ou do protocolo de comunicação.

O modo com o acesso que um oponente tenha sobre o par texto cifrado/texto claro em um ataque, estes são classificados em quatro categorias básicas:

Texto cifrado puro

Este ataque é o mais difícil do ponto de vista da criptoanálise. Nele o oponente tem acesso somente à várias mensagens cifradas por uma determinada chave.

Texto claro conhecido

Neste caso, o oponente tem conhecimento de vários pares de texto cifrado e claro. Este ataque é particularmente facilitado pela natureza repetitiva e bem conhecida dos cabeçalhos e demais estruturas dos protocolos de comunicação.

Texto claro escolhido

Este tipo de ataque é o mais poderoso em se tratando de criptoanálise do algoritmo de criptografia em si. O oponente, além de ter acesso a pares de texto cifrado e claro, é capaz de escolher blocos específicos de texto para serem cifrados, de maneira que estes possam induzir a informações a respeito da chave de ciframento.

A criptoanálise diferencial, introduzida em 1990 por Eli Biham e Adi Shamir, usa o princípio de escolher textos claros com diferenças particulares e analisar os textos cifrados produzidos, para derivar alguns bits da chave de algoritmos baseados no DES.

Criptoanálise linear é outra técnica que utiliza pares de texto cifrado/claro para também derivar probabilisticamente partes da chave de ciframento, de algoritmos também baseados no DES.

Os ataques de texto claro escolhido são mais fáceis de se implementar do que se poderia imaginar em uma primeira instância. Somente para citar um exemplo, este ataque pode ser facilmente aplicado por usuários de sistemas que empregam criptografia ponto a ponto, como no caso de redes virtual (VPN).

Ataque do nó intermediário (Man-in-the-Middle Attack)

O controle de algum nó intermediário no percurso entre duas entidades, por parte de um oponente, é estrategicamente muito interessante. Se este for capaz apenas de grampear o meio de comunicação, isto é, for um oponente passivo, todos os blocos cifrados por um determinado algoritmo vão estar disponíveis para o processo de criptoanálise. Isto aumenta significativamente a probabilidade da criptoanálise ser bem sucedida.Por outro lado, se o oponente tiver a habilidade de interceptar e retransmitir mensagens no meio de comunicação, ou seja, for um oponente ativo, ataques ainda mais sutis entram em questão. O nó intermediário pode, neste caso, personificar cada uma das extremidades (entidade) à outra.

Outro ataque possível é combinar duas mensagens cifradas com a mesma chave, para produzir uma terceira de significado distinto. Esta técnica é também conhecida como ataque de recorte e colagem. Finalmente, ataques baseados na repetição de mensagens são também mais facilmente implementados por oponentes que disponham do posicionamento estratégico descrito acima.