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

Rio de Janeiro, RJ - Brasil

Criptografia
O que é, porque utilizar e quais suas fragilidades?

GRIS-2011-A-003

Thiago Gonçalves Escobar

A versão mais recente deste documento pode ser obtida na página oficial do GRIS: http://www.gris.dcc.ufrj.br.

GRIS - Grupo de Resposta a Incidentes de Segurança
Av. Brigadeiro Trompowski, s/n°
CCMN – Bloco F1 - Decania
Cidade Universitária - Rio de Janeiro/RJ
CEP: 21949-900
Telefone: +55 (21) 2598-9491

Última atualização em: 21 de fevereiro de 2011

Sumário

1 Introdução

Criptografia (do grego *kryptós*, "escondido", e *gráphein*, "escrita") é o estudo dos princípios e técnicas pelas quais a informação pode ser transformada da sua forma original para outra ilegível, de forma que possa ser conhecida apenas por seu destinatário, o que a torna difícil de ser lida por alguém não autorizado. Assim sendo, só o receptor da mensagem pode ler a informação com facilidade. É um ramo da Matemática, parte da criptologia.

Este artigo visa esclarecer o leitor em relação a criptografia assimétrica, através de noções básicas e exemplos práticos de uso da mesma, utilizando um dos métodos mais utilizados hoje em dia, o RSA.

Pesquisar sobre criptografia é importante, pois cada vez mais estamos ligados ao mundo através da Internet, e com isso, cada vez mais, informações sensíveis estão sendo enviadas via web, como contas de banco e informações pessoais, como documentos, que podem ser utilizados por terceiros com a finalidade de realizar um desfalque. Logo, entender como se proteger é extremamente importante.

Após a apresentação inicial da criptografia, e sua comparação com a esteganografia, darei alguns exemplos práticos de situações o uso da mesma é cada vez mais importante. Então, mostrarei como funciona, e no que é baseado um dos métodos criptográficos mais utilizados atualmente (RSA), além do que gera sua segurança. Também serão abordadas algumas políticas para escolhas dos componentes do RSA.

2 Criptografia: O que é?

A criptografia pode ser entendida como uma função que modifica a informação e a torna ilegível para quem não tenha a chave de decodificação. Tornando, assim, a informação só disponível para o destinatário, que é o detentor da chave de decodificação. Ela veio para resolver problemas de envio de informação sensível em meios não confiáveis.

Muitos confundem a criptografia com a esteganografia, visto que o objetivo de ambas é o mesmo(ocultar o conteúdo da mensagem a todos que não sejam destinatários legítimos da mesma). A grande diferença entre ambas é que a criptografia é facilmente identificável, já a esteganografia, em geral, é feita de um modo que não seja facilmente identificável. Ou seja, podemos dizer que a criptografia é um método explícito, enquanto a esteganografia, é um método implícito.

2.1 A criptografia RSA

Um dos métodos de criptografia mais utilizados atualmente é chamado de RSA (nome que surgiu através da união do nome dos seus criadores – R.L.Rivest, A. Shamir e L.Adleman). O mesmo foi inventado em 1978, quando os três trabalhavam no Massachussets Instute of Technology(M.I.T). Ele é um método de criptografia que utiliza chaves públicas e por ser um dos métodos mais utilizados atualmente, será abordado seu funcionamento.

Para se criptografar utilizando o método RSA precisamos de dois números primos, p e q, além da mensagem que será criptografada, transcrita para uma codificação que um computador possa entender (como a tabela ascii ou uma tabela que veremos no exemplo prático), que chamaremos de m. Definimos n como o produto entre p e q, ou seja, $n = pq$, e como p e q são primos, n só possui dois fatores primos. Além disso, definimos[1] $\Phi(n)$ (lê-se fí de n) como $(p-1)(q-1)$.

Definimos agora e como um número que o $mdc(e, \Phi(n)) = 1$(que é correspondente a dizermos que e é um número inversível[2] módulo $\Phi(n)$). Em geral, e é escolhido como um número primo, para facilitar a sua escolha. Este par (n, e) também é conhecido como *chave pública* e é a *chave de codificação* do sistema RSA que estamos utilizando. Note que embora o usuário conheça e, que é inversível módulo $\Phi(n)$, o mesmo não conhece $\Phi(n)$, pois caso contrário, seria fácil quebrar o RSA.

Seja m uma mensagem que será codificada, chamaremos de $C(m)$ o resultado do processo de codificação. Para calcular $C(m)$ tudo que devemos fazer é calcular :

$$\boldsymbol{C(m)} = \text{resto da divisão de } m^e \text{ por } n$$

Além disso, devemos ter o cuidado de pegar $m < n$, pois caso contrário, mais de uma decodificação seria possível.

Então, temos agora a mensagem codificada. Mas como fazemos para decodificá-la? Tudo que precisamos é de n, que já possuímos, e de um número que chamaremos de d, que é o inverso modular de e em $\Phi(n)$. Isso é equivalente a dizer que $de = 1mod(\Phi(n))$. Como usaremos d como potência devemos escolhê-lo como número positivo. Chamamos o par ***(n, d)*** de *chave de decodificação* ou também de *chave de privada*. A chave de decodificação no RSA deve ser privada pois queremos que qualquer pessoa seja capaz de codificar informação utilizando-o, porém somente pessoas autorizadas (os detentores da chave privada) sejam capazes de decodificá-la.

Seja a uma mensagem codificada, chamaremos de $D(a)$ o resultado do processo de decodificação. Para calcular $D(a)$ tudo que devemos fazer é calcular :

$$\boldsymbol{D(a)} = \text{resto da divisão de } a^d \text{ por } n$$

Com isso, temos novamente a mensagem original, ou seja, $D(C(n)) = n$. Com isso, podemos observar que as funções de codificação e de decodificação são inversas. Isso será importante para o método de assinaturas que é utilizado pelo RSA(e também por todos os métodos de criptografia de chave pública).

[1] Na verdade, A função $\Phi(n)$ é algo muito mais complexo do que foi abordado. Para mais informações a respeito, veja o Livro *Números Inteiros e Criptografia RSA- S.C.Coutinho*

[2] Na verdade, isso é só uma aplicação da Teoria dos Grupos, que diz que um número que seja primo com a base é sempre inversível nessa base. Ou seja, existe algum fator tal que esse número multiplicado por este fator inteiro dividido por n deixa resto 1

2.1.1 Exemplo prático de criptografia RSA

Vamos ver agora como o RSA funciona na prática.

Primeiramente, vamos criar nosso sistema RSA. Para isso, nós devemos escolher dois números primos para servir como base para nossa criptografia.

Para facilitar os cálculos do exemplo, iremos escolher dois números muito pequenos (o que nunca deve ser feito na prática, pois a segurança do RSA está diretamente ligada ao tamanho das chaves criptográficas – fato que será abordado posteriormente).

Escolhemos nesse exemplo $p = 13$ e $q = 19$. Temos então $n = pq = 247$. Podemos, então, calcular $\Phi(n) = (p-1)(q-1) = 1218 = 216$.

Escolhemos então $e = 5$, que é o menor número primo que não divide 216. Encontramos agora d tal que $de = 1\ mod(\Phi(n))$. Temos então $d = 173$. Com isso, temos a chave pública $(n, e) = (247, 5)$ e a chave privada $(n, d) = (247, 173)$. Após isso, podemos destruir toda informação que não sejam as chaves pública e privada. Devemos fazer isso para aumentar a segurança do sistema de criptografia.

Finalmente, iremos criptografar uma mensagem. A mensagem escolhida é **"GRIS"**. Devemos então pré-codificá-la através de uma tabela de substituição de letras por números. Utilizaremos a tabela a seguir :

A	B	C	D	E	F	G	H	I	J	K	L	M
10	11	12	13	14	15	16	17	18	19	20	21	22

N	O	P	Q	R	S	T	U	W	V	X	Y	Z
23	24	25	26	27	28	29	30	31	32	33	34	35

Uma coisa que devemos nos preocupar é não utilizar uma tabela que seja ambígua, além de sempre utilizar notação completa. Por exemplo, se nossa tabela começasse por 1 e fosse até 26, quando pegássemos a mensagem "AA" seria pré-codificada como "11" e não teríamos como saber se a mesma seria correspondente a "AA" ou "K", ou seja, teríamos uma ambigüidade. Porém, se quando pré-codificássemos a mesma utilizássemos 01 a 26, a mensagem "AA" seria pré-codificada como "0101" e não haveria ambigüidade quando fôssemos decodificar a mensagem.

Obtemos então $\mathbf{G = 16, R = 27, I = 18, S = 28}$, que é nossa mensagem que será enviada. Dividimos a mensagem em blocos, que serão enviados individualmente. Devemos sempre tentar dividir a mensagem de forma que o bloco atual seja o maior possível. Mas não devemos nunca utilizar um bloco que seja maior do que n, que nesse caso é 247. Obtemos então a mensagem dividida nos blocos: **[162] [71] [82] [8]**.

Lembramos então que para codificar uma mensagem, fazemos $C(b)$ = resto da divisão de b^e por n. Ou seja, pegamos cada bloco b, elevamos a $e = 5$ e então achamos o resto da divisão entre esse número e $n = 247$. Então, codificamos a mensagem (sempre bloco por bloco) e obtemos os seguintes blocos: **[41] [67] [62] [164]**.

A mensagem seria então enviada e quando recebida o destinatário efetuaria a decodificação. O que seria feito pelo destinatário após o recebimento da mensagem é achar $D(a)$ = resto da divisão de a^d por n, para cada bloco, sendo a a mensagem que será decodificada. Ou seja, ele deveria pegar cada bloco a, elevar a $d = 173$ e então acharia o resto da divisão entre esse número e $n = 247$. Como na codificação, sempre devemos decodificar bloco por bloco. Se efetuarmos a conta, acharemos os blocos: **[162] [71] [82] [8]**. Que então seriam entendidos como **[16] [27] [18] [28]** e então recodificados como **"GRIS"**.

Como já vimos, as funções de codificação e decodificação são inversas no RSA. Então, poderiamos primeiramente decodificar a mensagem com a chave privada e então qualquer outra pessoa pode codificá-la utilizando a nossa chave pública. Isso poderia ser feito para garantir que a mensagem veio realmente de mim. Isso é constantemente realizado pelo sistema de assinaturas que será abordado a seguir.

Já sabemos que a mensagem **"GRIS"** tem como blocos a ser codificados os blocos **[162] [71] [82] [8]**. Vamos agora codificar a mensagem utilizando a chave pública. Para isso, fazemos $D(b)$ = resto da divisão de b^d por n, sendo b cada um dos blocos a ser codificado.

Fazendo as contas, obtemos: **[223] [184] [36] [164]**.

Após isso, a mensagem seria enviada e então deveria passar por decodificação(ou recodificação, como também é conhecido), só que utilizando a chave pública. Ou seja, faremos $C(a)$ = resto da divisão de a^e por n, sendo a cada bloco que será decodificado.

Fazendo as contas, obtemos: **[162]** **[71]** **[82]** **[8]**. Que como já sabemos, será entendido como **[16]** **[27]** **[18]** **[28]** e então recodificados como **"GRIS"**.

Esse sistema então poderia, agora que já foi testado, ser utilizado. É importante realizarmos alguns casos de teste de codificação e decodificação no nosso sistema RSA para podermos ter certeza que durante a criação das chaves pública e privada não houve nenhum erro, que permitiria gerar um código codificado que não pudesse ser decodificado. Ou até mesmo que o inverso não seria possível (relembrando que as funções de decodificação e codificação são inversas).

2.2 Sistema de Assinaturas

Como o RSA utiliza uma chave pública para codificar uma mensagem, como o mesmo garante que a mesma é legítma? Ele utiliza uma método de assinaturas, que veremos a seguir como funciona.

Sejam $\mathbf{C_e}$ e $\mathbf{D_e}$, respectivamente, as funções de codificação e decodificação de uma empresa, $\mathbf{C_b}$ e $\mathbf{D_b}$, respectivamente, as funções de codificação e decodificação de um banco e a um bloco que será enviado utilizando o Sistema de Assinaturas. Para utilizá-lo, primeiramente a empresa deve utilizar sua chave privada sobre a mensagem(então teríamos $\mathbf{D_e(a)}$, e após isso, codificá-la utlizando a chave pública do banco (então teríamos $\mathbf{C_b(D_e(a))}$ e só então enviá-la ao banco.

Após recebida a mensagem, o banco aplica sua chave de decodifcação sobre a mensagem ($\mathbf{D_b(C_b(D_e(a)))} = \mathbf{D_e(a)}$ – já que $\mathbf{C_b}$ e $\mathbf{D_b}$ são funções inversas) – e então a recodifica utilizando a chave pública da empresa $\mathbf{C_e(D_e(a))}$.

Como as funções de codificação e decodificação são inversas e é muito difícil que existam 2 chaves privadas que codifiquem uma informação de maneira equivalente, o fato de haver uma recodificação possível utilizando a chave pública da empresa é o bastante para garantir que a mensagem original veio de fato da empresa para o banco. No caso onde uma chave privada equivalente à chave privada original foi encontrada é um caso onde a segurança do método de assinaturas foi quebrada.

3 Criptografia: Porque utilizar?

Se formos analisar a história da criptografia, vemos que ela era muito utilizada na antiguidade, por reis que se preocupavam, principalmente, com suas táticas de guerra, que não deveriam ser conhecidas por seus adversários, visto que essa falha na segurança poderia até culminar na perda de uma guerra. É por motivos de privacidade ou ocultação de informação que a criptografia é utilizada.

Cada vez mais, são enviadas informações via web. Quem nunca realizou uma compra online? Ou então um cadastro onde informações pessoais, como CPF e RG fossem necessários? Ou até mesmo uma transação bancária via Internet Banking? Quem é responsável por garantir que nenhum terceiro ficou sabendo dessas informações? A criptografia é responsável por isso.

As redes hoje em dia sofrem cada vez mais ataques que podem comprometer a segurança da informação. Se proteger contra isso é extremamente importante. Um forte exemplo disso, são os ataques de *Sniffing*, que é um tipo de ataque que visa ter acesso a informações que estão trafegando por uma rede e que não eram endereçadas ao seu computador. Um dos modos de se proteger de ataques desse tipo é criptografar seus dados antes de enviá-los via web.

Vamos pensar em um exemplo: Um usuário está realizando uma compra online, em seu computador, que é seguro, através de um site desconhecido por ele. Nele, o usuário inclui seu CPF, RG e número de seu cartão, para realizar a compra. Até aí, tudo normal.

Agora podem existir duas situações: 1 - O site utiliza um sistema de criptografia forte. 2- O site não utiliza um sistema de criptografia forte, ou nem utiliza criptografia. Na primeira situação, tudo está certo, e suas informações continuarão seguras. Já na segunda, temos que pensar em todo o trajeto que sua informações percorre quando sai do seu computador até o seu destino final.

Em uma situação normal, a informação percorre diversos roteadores e cada roteador pode ser alvo de *Sniffers* (que é como os atacantes que utilizam Sniffing são conhecidos) que poderão então visualizar dados da compra, como o CPF, o RG e o número do cartão do usuário, por exemplo. E com isso, o atacante terá todas as informações necessárias para realizar qualquer compra se passando pelo usuário. Isso é um sério problema.

E embora não seja o único, esse problema já seria motivo suficiente para utilizarmos a criptografia. Mas existem inúmeros outros, como roubo de usuários e senhas de e-mails, por exemplo. Ou seja, em geral, a criptografia é utilizada para tentar proteger uma informação contra furtos.

4 Criptografia: Quais suas fragilidades?

A segurança de um sistema de criptografia por chaves públicas está inteiramente ligada à dificuldade de se obter a chave privada, que é utilizada para se decodificar uma informação, após o processo de codificação (que utiliza a chave pública, também conhecida como chave primária).

Dizer que a obtenção da chave privada a partir da chave pública para um determinado sistema de criptografia é fácil de ser realizada equivale a dizer que o sistema pode ser quebrado muito facilmente. Logo, devemos nos preocupar com a dificuldade de se obter a chave privada através do conhecimento da chave pública.

Se alguém consegue obter a chave privada através da chave pública, dizemos que a criptografia foi quebrada.

4.1 Em que está baseada a segurança do RSA

A segurança do RSA é baseada no fato de ser difícil obter a chave privada (n, d) a partir da chave pública (n, e). O modo mais direto de se obter (n, d) sabendo (n, e) é fatorar n, obtendo assim p e q.O passo seguinte é calcular $\Phi(n) = (p-1)(q-1)$ para poder calcular d tal que $de = 1mod(\Phi(n))$. E como nenhum algoritmo eficiente para fatorar um número é conhecido atualmente, para números suficientemente grandes, quebrar o RSA não é fácil.

Porém, alguns algoritmos de fatoração conhecidamente funcionam muito bem em situações específicas. É o exemplo do algoritmo de Fermat, que funciona muito bem quando $|p-q|$ é pequeno. Então, devemos sempre tentar conhecer novos algoritmos de fatoração, mesmo que alguns deles só funcionem em situações específicas, pois assim podemos nos proteger e não cairmos numa dessas situações.

Há também alguns outros algoritmos conhecidos para fatorar um número da forma $n = pq$, que não serão abordados neste artigo. Porém, é importante notar que os mesmos funcinam bem quando: $(p-1), (q-1), (p+1) ou (q+1)$ possuem fatores primos pequenos. Logo, devemos também tomar esse cuidado na hora de gerar números primos para o RSA.

Em 1994, o matemático Peter Shor propôs um algoritmo – que veio a ser conhecido c omo *Algoritmo de Shor* – que é capaz de fatorar um número composto em tempo polinomial (em relação ao número de dígitos do mesmo). No entanto, este algoritmo necessita de um computador quântico para ser implementado e até hoje ninguém foi capaz de construir um computador quântico útil e funcional.

5 Conclusões

A conclusão que conseguimos obter através deste artigo é o fato do uso da criptografia ser cada vez mais importante, a fim de protegermos informações sensíveis, que estão cada vez mais presentes em um meio não-seguro. Conclui-se também que o RSA (assim como qualquer outro sistema de criptografia que se baseie no fato de ser difícil realizar fatorações de sua chave pública) teriam fim caso um computador quântico capaz de implementar o Algoritmo de Shor fosse feito.

6 Referências Bibliográficas

Números Inteiros e Criptografia RSA - S.C.Coutinho - Rio de Janeiro - IMPA - 2007

http://en.wikipedia.org/wiki/RSA

http://www.di-mgt.com.au/rsa_alg.html