CONICT 12º Congresso de Inovação, Ciência e Tecnologia
Pró-reitoria de Pesquisa e Pós-Graduação
INSTITUTO FEDERAL São Paulo Câmpus Cubatão

12º Congresso de Inovação, Ciência e Tecnologia do IFSP - 2021

NOÇÕES DE CRIPTOGRAFIA RSA

JANAINA MANZALI DA SILVA[1], BRUNO CHIODEROLI GREGIO[2]

[1]Cursando Técnico em Administração Integrado ao Ensino Médio, Instituto Federal de Educação, Ciência e Tecnologia de São Paulo - IFSP, Câmpus Birigui, janaina.m@aluno.ifsp.edu.br.
[2]Mestre em Matemática, professor, Instituto Federal de Educação, Ciência e Tecnologia de São Paulo - IFSP, Câmpus Birigui, bruno.gregio@unesp.br.

Área de conhecimento (Tabela CNPq): 1.01.01.03-9 Teoria dos Números.

RESUMO: Para proteger a imensa quantidade de informação compartilhada a todo momento no mundo atual, requer-se um método eficiente para criptografar e descriptografar mensagens. Nesse cenário, surge a criptografia RSA (Rivest-Shamir-Adleman), uma das mais utilizadas atualmente. Sua fundamentação está numa importante área da Matemática, a Teoria dos Números, mais precisamente nos conceitos de divisibilidade e números primos, estando a segurança do método atrelada à ineficiência de algoritmos de fatoração diante da utilização de chaves numéricas com 500 algarismos ou mais. Uma pequena amostra de tal dificuldade é apresentada neste resumo, bem como os principais cálculos a serem efetuados num processo de encriptação e decriptação de uma mensagem, como forma de evidenciar o modelo matemático que o embasa.

PALAVRAS-CHAVE: criptografia RSA; divisibilidade; números primos; aritmética.

RSA ENCRYPTION NOTIONS

ABSTRACT: To protect the immense amount of information shared at all times in today's world, an efficient method of encrypting and decrypting messages is required. In this scenario, RSA (Rivest-Shamir-Adleman) encryption appears, one of the most used today. Its foundation is in an important area of Mathematics, the Theory of Numbers, more precisely in the concepts of divisibility and prime numbers, with the security of the method being linked to the inefficiency of factoring algorithms when faced with the use of numerical keys with 500 digits or more. A small sample of such difficulty is presented in this summary, as well as the main calculations to be performed in a message encryption and decryption process, as a way of highlighting the mathematical model that supports it.

KEYWORDS: RSA cryptography; divisibility; prime numbers; arithmetic.

INTRODUÇÃO

Um método de criptografia tem como principal objetivo garantir segurança e privacidade. A palavra "criptografia" deriva do grego kryptós, "escondido" e gráphein, "escrita", representando uma escrita

escondida, ou seja, tornando algo legível em ilegível. De forma resumida, a criptografia é definida como um conjunto de regras e técnicas utilizadas para cifrar (codificar) algo escrito, resultando disto um código incompreensível para quem não está autorizado a ter acesso ao seu conteúdo.

Segundo Singh (2005 apud BONFIM, 2017), "a natureza humana tem uma necessidade à privacidade, mesmo que seja para guardar simples segredos. É possível observar na história que a linguagem de códigos sempre foi muito usada como recurso militar, político, em questões comerciais, em guerras e até mesmo questões sentimentais". Com isso, percebe-se que a criptografia surge e se aperfeiçoa para atender as demandas da humanidade ao longo da história, de modo que, atualmente, até mesmo aplicativos de mensagem, como o Whatsapp, utilizam criptografia para garantir segurança e privacidade aos seus usuários.

Nesse cenário, destacam-se dois tipos de criptografia: simétrica e assimétrica. Na criptografia simétrica (convencional) temos uma única chave, responsável por cifrar e decifrar, de modo que o seu compartilhamento é necessário, ocasionando assim um problema de segurança. Tal problema é revertido com a criptografia assimétrica, que utiliza duas chaves, uma pública para codificar e uma privada para decodificar, como pode ser visto na figura a seguir.

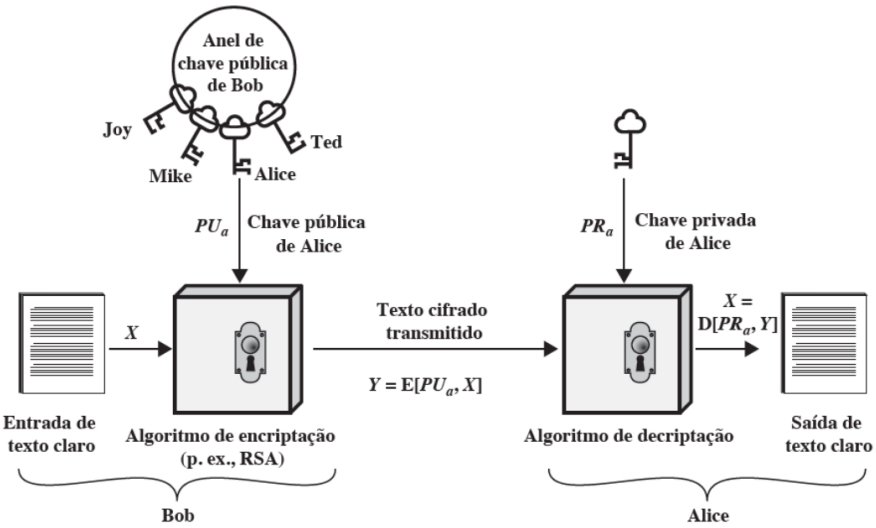

Figura 1: Esquema de criptografia assimétrica com chaves pública e privada.
Fonte: (STALLINGS, 2015, p.202).

No exemplo, Bob envia uma mensagem criptografada para Alice, utilizando a chave pública da mesma para encriptar a mensagem, que por sua vez, só poderá ser decriptada pela própria Alice, única detentora da chave privada. A unicidade da propriedade da chave privada é fundamental para a segurança do método.

No âmbito da criptografia assimétrica surge o método RSA, criado em 1978 e estruturado a partir de conceitos aritméticos tais como primalidade, divisibilidade, congruências lineares, entre outros, e sua segurança está relacionada, de maneira geral, à ineficiência dos métodos de fatoração atuais. Por reunir segurança e eficiência, trata-se da criptografia mais utilizada atualmente, sobretudo em senhas e transações bancárias.

MATERIAIS E MÉTODOS

Como a pesquisa é de cunho teórico, para alcançar os objetivos, fez-se uma revisão de literatura, com pesquisas em livros, dissertações, teses e artigos que versam sobre o tema, além de alguns sites.

Juntamente a isto, aconteceram encontros semanais entre aluna/pesquisadora e orientador, a fim de discutir, explanar e investigar o material estudado.

RESULTADOS E DISCUSSÃO

Nesta seção apresentaremos como funciona, de fato, o algoritmo do RSA, evidenciando os cálculos envolvidos desde o processo de escolha das chaves, até a codificação e decodificação de uma mensagem. De modo geral, mostraremos os conceitos matemáticos necessários à compreensão e execução do método, bem como o que de fato garante a segurança do mesmo.

Para implementá-lo, precisamos, essencialmente, de dois números primos p e q, distintos entre si, de modo que o produto entre eles resulte um número n, isto é, $n = pq$. Sendo assim, associamos n à chave pública (usada para cifrar a mensagem) e os próprios p e q à chave privada (usada para decifrar).

Apresentando o RSA dessa maneira, parece muito fácil quebrar seu código, pois se conhecemos n, basta fatorá-lo para obter p e q, e assim obter a chave privada. De fato, a ideia é exatamente esta, porém a fatoração não é algo tão simples assim quando pensamos em números extremamente grandes e, neste sentido, vale ressaltar que a ideia de números grandes é um tanto genérica, pois pela infinitude dos números inteiros, existirá sempre um inteiro tão grande quanto se queira. Observe que fatorar números desta natureza já deixa de ser algo tão trivial.

Diante disso, uma nova questão se põe: como escolher os primos p e q suficientemente grandes? Isto é, como saber, de fato, que p e q são primos sem fatorá-los, já que isto é difícil? A resposta para esta questão é que existem alguns testes de primalidade sem a necessidade da fatoração, dentre os quais destacamos os números do tipo $M(n) = 2^n - 1$ e $F(n) = 2^{2^n} + 1$, com $n \in \mathbb{Z}_+$, que ficaram conhecidos, respectivamente, como números de Mersenne e de Fermat. Com relação aos números de Mersenne, para que sejam primos, n deve, necessariamente, ser primo, porém tal condição não é suficiente para garantir a primalidade, podendo assim $M(n)$ ser composto, mesmo com n primo. Quanto aos números de Fermat, temos que $F(n)$ é primo para n igual a 0, 1, 2, 3 e 4, contudo, acredita-se que, por sua fórmula duplamente exponencial, tais números tenham sido pouco estudados, com isso, até hoje não se conhece nenhum primo de Fermat com $n \geq 5$. Ressaltamos também o Crivo de Eratóstenes, que separa todos os primos de 3 até um inteiro n ímpar arbitrário por um sistema de peneira, e o teste de primalidade de Fermat, que usa o Pequeno Teorema de Fermat, no sentido de que, se p é primo, e a é um inteiro não divisível por p, então $a^{p-1} \equiv 1 \bmod(p)$, de modo que, se a congruência não se mantém, então p é composto. Ademais, testes como o dos Números de Carmichael, o Teorema de Korselt, o teste de Lucas, dentre outros, podem ser vistos com mais detalhes em (COUTINHO, 2011).

Voltando à execução do RSA, o algoritmo fica dividido em três etapas: pré-codificação, codificação e decodificação. Na primeira, devemos ajustar os códigos, convertendo todas as letras numa sequência de números. Para simplificar, vamos escolher um texto sem caracteres numéricos, por exemplo "RSA é eficaz". Usaremos 99 para identificar os espaços e, para as letras, a seguinte tabela de conversão:

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	V	W	X	Y	Z
23	24	25	26	27	28	29	30	31	32	33	34	35

Tabela 1: Tabela de conversão para a pré-codificação.

Com isto, o texto é codificado no número: 272810991499141518121035. Em seguida, escolhe-se dois primos distintos p e q, e calcula-se $n = pq$. Para facilitar as contas utilizaremos $p = 13$ e $q = 17$, resultando $n = 221$. Tendo feito isso, quebramos a mensagem codificada em blocos, de modo que cada bloco seja menor que n. Além disso, devemos tomar o cuidado para que não comece com 0, algo que traria problemas no momento de decodificar, e que não corresponda a nenhuma unidade linguística, sendo assim, podemos quebrá-la, por exemplo, em : $2-72-8-109-9-149-91-41-51-81-210-35$.

Para realizar a codificação da mensagem, determinamos $\phi(n) = (p-1)\cdot(q-1)$, conhecido como função fi de Euler, a qual representa a quantidade de números inteiros positivos coprimos e menores que n, e um inteiro positivo e tal que mdc $(e, \phi(n)) = 1$. Assim, $\phi(n) = 12 \cdot 16 = 192$, portanto, podemos tomar $e = 5$, que é o menor primo que não divide 192, com isso a chave de codificação do sistema RSA será representada pelo par $(e, n) = (5, 221)$, que se tornará público. Faremos os cálculos do primeiro bloco, sendo todos os demais realizados de modo análogo. O código referente ao primeiro bloco, ou seja, 2, será o resto da divisão de 2^e por n, isto é, de 2^5 por 221, resultando em 32, que em notação de congruência linear é expressado por $2^5 \equiv 32 \,\mathrm{mod}(221)$.

Para decodificar são necessários n e o inverso de e em $\phi(n)$, que denotaremos d. Se conhecemos $\phi(n)$ e e, pode-se dizer que é relativamente fácil calcular d, pois basta aplicar o algoritmo euclidiano estendido. Neste caso, como $\phi(221) = 192$ e $e = 5$, o algoritmo nos dá $d = 77$. A chave de decodificação será o par (d, n), com o qual calculamos o resto da divisão de 32^d por n, isto é, de 32^{77} por 221.

Assim, temos que:

$$32^5 \equiv 2 \bmod (221)$$
$$\left(32^5\right)^{15} \equiv 2^{15} \equiv 60 \bmod (221)$$
$$32^{75} \cdot 32^2 \equiv 60 \cdot 32^2 \equiv 2 \bmod (221)$$

portanto

$$32^{77} \equiv 2 \bmod (221)$$

ou seja, o resto da divisão de 32^{77} por 221, como já se esperava, resulta em 2, que consiste no código do bloco original.

O esquema abaixo ilustra os procedimentos descritos anteriormente:

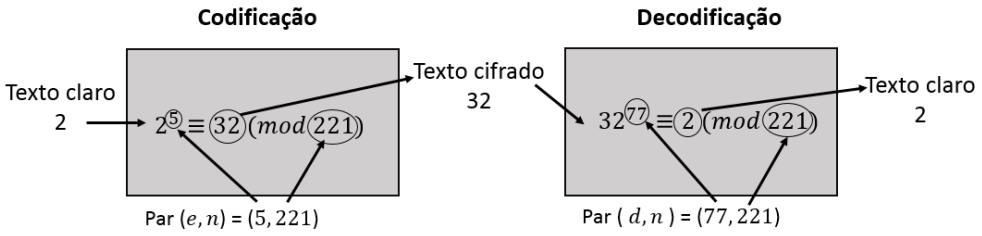

Figura 2: Processo de codificação e decodificação.
Fonte: Elaborado pelo autor.

Inicialmente, de modo simplista, afirmamos que codificamos usando n e decodificamos usando p e q. O que acontece, de fato, é que para calcular d precisamos conhecer $\phi(n)$, que por sua vez é determinado conhecendo-se p e q. Como $\phi(n)$ não é revelado, a segurança do método consiste na dificuldade de

calcular d conhecendo apenas a chave pública (e, n).

Um método alternativo para a decodificação é o de Quisquater-Couvreur, que consiste na utilização do algoritmo chinês do resto. De acordo com Menezes, Oorschot e Vanstone (1997), o tempo de execução com este método é cerca de 4 vezes mais rápido que o empregado na aplicação direta (a que exemplificamos anteriormente).

A criptografia, em si, evolui ao seu máximo na era digital, com cálculos extremamente complexos sendo efetuados em curtos períodos de tempo, gerando códigos complicados e difíceis de serem quebrados. Com este trabalho, conseguimos transmitir uma ideia básica de alguns dos cálculos necessários e esclarecer dúvidas em relação à funcionalidade do método.

Na busca pela segurança de dados, aliada ao aumento do poder de processamento dos computadores, constantemente temos alguém tentando descobrir números primos maiores, sendo assim possível tanto atribuir mais segurança ao método, quanto quebrar os códigos, gerando uma disputa de velocidade para ambos os lados. Atualmente o RSA requer uma chave de 2048 bits, que representa um número com 617 dígitos decimais, para garantir segurança até 2030.

CONCLUSÕES

Como definição, a criptografia admite a criação de protocolos que restringem uma mensagem privada, impossibilitando o acesso do público ou de terceiros. Na atualidade, diante da expansão digital, a integridade, a confidencialidade e a necessidade de autenticação dos dados se fazem necessárias para nos protegermos, dentre tantas outras coisas, dos perigos da Internet. Neste ambiente, qualquer descuido pode tornar públicas informações sigilosas como número e senha de um cartão de crédito, documentos pessoais, fotos pessoais, entre outras coisas. O aumento exponencial de pessoas na internet, comprando itens, trocando informações importantes, trabalhando ou estudando, reflete a importância de se dedicar à proteção citada. O método de criptografia RSA, conforme descrito neste trabalho, tem a sua segurança fundamentada num importante conceito matemático, e como vimos, é capaz de permitir a nossa imersão neste sistema global de redes, que é a internet, de forma saudável e segura.

AGRADECIMENTOS

Agradeço ao IFSP - Campus Birigui por me fornecer o privilégio de realizar minha primeira experiência em contato com um projeto de Iniciação Científica. Também agradeço ao meu professor orientador Bruno Chioderoli Gregio por aceitar minha ideia e me ajudar a construir um projeto acadêmico.

REFERÊNCIAS

BONFIM, D. H. *Criptografia RSA*. Dissertação (Mestrado) — Instituto de Ciências Matemáticas e de Computação, Universidade de São Paulo, 2017.

COUTINHO, S. C. *Números Inteiros e Criptografia RSA*. Rio de Janeiro: IMPA, Coleção Matemática e Aplicações, 2011.

MENEZES, A. J.; OORSCHOT, P. C. V.; VANSTONE, S. A. *Handbook of applied cryptography*. 1. ed. Boca Raton: CRC press, 1997.

STALLINGS, W. *Criptografia e segurança de redes: princípios e práticas*. 6. ed. São Paulo: Pearson Education do Brasil, 2015. 202 p.