UNIVERSIDADE FEDERAL DE SERGIPE

PRÓ-REITORIA DE PÓS- GRADUAÇÃO E PESQUISA

PROGRAMA DE PÓS-GRADUAÇÃO EM MATEMÁTICA

MESTRADO PROFISSIONAL EM MATEMÁTICA

REDE NACIONAL - PROFMAT

Vinícius Matos de Oliveira

Criptografia e a Matemática

São Cristóvão - SE
2021

Vinícius Matos de Oliveira

Criptografia e a Matemática

Dissertação apresentada ao Programa de Pós - Graduação em Matemática da Universidade Federal de Sergipe, como parte dos requisitos para obtenção do título de Mestre em Matemática.

Orientador: Prof. Dr. Naldisson dos Santos

São Cristóvão - SE
2021

FICHA CATALOGRÁFICA ELABORADA PELA BIBLIOTECA CENTRAL
UNIVERSIDADE FEDERAL DE SERGIPE

Oliveira, Vinícius Matos de
O48c Criptografia e Matemática / Vinícius Matos de Oliveira ; orientador Naldisson dos Santos. – São Cristóvão, 2020.
54 f.

Dissertação (mestrado em Matemática) – Universidade Federal de Sergipe, 2020.

1. Matemática. 2. Números primos. 3. Congruências e restos. 4. Criptografia. I. Santos, Naldisson dos orient. II. Título.

CDU 51

UNIVERSIDADE FEDERAL DE SERGIPE
PRÓ-REITORIA DE PÓS-GRADUAÇÃO E PESQUISA
PROGRAMA DE PÓS-GRADUAÇÃO EM MATEMÁTICA

Dissertação submetida à aprovação pelo Programa de Pós-Graduação em Matemática da Universidade Federal de Sergipe, como parte dos requisitos para obtenção do grau de Mestre em Matemática.

Criptografia e a Matemática

por

Vinicius Matos de Oliveira

Aprovada pela banca examinadora:

Naldisson dos Santos
Prof. Naldisson dos Santos - UFS
Orientador

Fábio dos Santos
Prof. Fabio dos Santos - UFS
Primeiro Examinador

Prof. Gleidson Gomes da Silva - UFPE
Segundo Examinador

São Cristóvão, 22 de Fevereiro de 2021

Agradecimentos

Agradeço primeiramente à minha família por todo apoio dado neste período de luta e, sem dúvida, esta conquista eu divido com minha mãe, Gorete, minhas irmãs, Valéria e Vanessa e meus sobrinhos, Luís Fernando e Isabela. Grato por toda palavra de incentivo nos momentos de desânimo e por toda força a cada viagem de 440 quilômetros semanais para poder realizar este sonho.

Aos professores do departamento de matemática da UFS, agradeço a todos que foram meus professores durante o curso: Prof. Dr. Almir, Prof. Dr. Anderson, Prof. Dr. Alysson, Prof. Dr. Evilson, Prof. Dr. Fábio e, especialmente, ao meu orientador Prof. Dr. Naldisson, por toda dedicação, ensinamento e compreensão durante esta jornada. Muito obrigado a todos.

Aos meus colegas de Profmat: Carlos Educardo, Tony e Lucas. Obrigado pelo auxílio, companheirismo e troca de experiências.

Não posso deixar de agradecer a todos meus amigos e familiares que contribuíram direta ou indiretamente por esta conquista.

À CAPES, pois o presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Código de Financiamento 001.

Resumo

A presente dissertação visa alinhar o conhecimento matemático ao sistema de criptografia RSA. Inicialmente, é feita uma linha histórica sobre a evolução da criptografia, mostrando a necessidade de governantes, reis, comandantes passarem informações confidenciais ao longo da história. A seguir, é estudado o conhecimento matemático essencial para entender o sistema criptográfico mais utilizado atualmente: o sistema RSA. Entre este conhecimento matemático, os números primos se destacam na aritmética básica, e o sistema de congruência na aritmética modular. Em seguida, é demonstrado como funciona o sistema RSA de criptografia, assim como é garantida enorme dificuldade em quebrar a mensagem criptografada a partir dos números primos escolhidos para sua composição.

Palavras-chave: Números primos; Congruências; Criptografia; Sistema RSA.

Abstract

The present dissertation aims to align the mathematical knowledge to the RSA encryption system theme. Initially, a historical line is made on the evolution of cryptography, showing the need for rulers, kings, commanders to pass confidential information throughout history. Next, the essential mathematical knowledge is studied to understand the most used cryptographic system today: the RSA system. Among this mathematical knowledge, prime numbers stand out in basic arithmetic, and the congruence system in the modular arithmetic. Soon after, it is demonstrated how works the system Encryption RSA, as well as is guaranteed enormous difficulty in breaking the message encrypted from the prime numbers chosen for its composition

Keywords: Prime numbers; Congruences; Cryptography; RSA system.

Sumário

Introdução

A necessidade de se comunicar é inerente ao ser humano. Hoje se vive em um mundo onde o volume de troca de mensagens é imensurável e a quantidade de dados criados cresce exponencialmente. Guardadas as devidas proporções, em toda a história da humanidade, houve essa necessidade de troca de informações e, quanto mais sigilosa essa troca de informações, mais privilegiado seria aquele que a conhecia.

O que diferencia as trocas de informações atuais daquelas de outrora é claramente o avanço da tecnologia para codificar. Paralelamente, evoluíram as pessoas e as tecnologias que decodificavam essas informações sigilosas. O método de codificar sempre precisava de novos detalhes ou uma mudança radical diante do surgimento de métodos eficazes para quebrar o código apresentado.

A matemática entra nesse contexto porque é o centro do sistema de codificação mais utilizado atualmente: o sistema RSA. Este sistema garante o sigilo do imenso fluxo de informações, desde as transações bancárias, passando pelas mensagens nos aplicativos e por todas as formas de se garantir, ou ao menos se tentar, que não vazem informações pessoais da população.

O sistema RSA utiliza conceitos matemáticos difundidos nos bancos escolares, como o uso de números primos e as congruências, baseando sua logística na dificuldade de se decompor números cujos fatores são números primos grandes. Durante o trabalho será estudado detalhadamente todos os conteúdos matemáticos essenciais para entender como funciona tal sistema.

É importante frisar que o sistema RSA é recente e surge junto ao uso dos computadores nas transações comerciais, políticas e pessoais. Antes disso, a humanidade passou por

diversos modelos de criptografia, sempre com visão de que tal modelo nunca seria superado e suas mensagens não seriam decodificadas. Puro engano. Diversos governantes perderam batalhas ou a própria vida por ter informações interceptadas e decodificadas pelo oponente. Os decodificadores passaram a ter fundamental importância nos rumos de uma guerra.

Diante desse cenário de necessidade de ter um fluxo de informações que fosse ao mesmo tempo secreto e rápido, a evolução nas últimas décadas é notória. Com o advento dos computadores e sua rápida possibilidade de se trocar dados, foi necessário também se pensar como fazer com que todo esse fluxo fosse seguro para quem utilizava.

Na última década, com a popularização das redes sociais, puderam-se constatar diversas situações em que os usuários tiveram seus dados expostos devido a alguma fragilidade no sistema de codificação das informações. Toda essa situação coloca em xeque a real proteção que tais redes sociais disponibilizam para seus usuários.

Uma situação de exposição de dados não pode ocorrer com o sistema bancário, pois basta imaginar o caos que seria se as contas bancárias da população fossem expostas com suas respectivas senhas, ou então os dados do cartão de crédito levados ao público, Hodiernamente, grande parte das transações bancárias são realizadas por meio de aplicativos no próprio celular. O usuário precisa ter a confiança de que tais informações serão realmente sigilosas ou o aplicativo não será utilizado pelos consumidores.

Diante de tanta necessidade e comodidade do uso da internet para as transações bancários, para a troca de mensagens pessoais, para os projetos realizados no trabalho, o que sustenta toda confiabilidade neste uso é o sistema de criptografia. Sem este, seria possível interceptar e decodificar facilmente mensagens e dados trocados entre usuários.

Quando se fala em troca de mensagens eficientes e sigilosas, recai-se sobre o sistema RSA, criado por Ronald Rivest, Adi Shamir e Leonard Adleman - daí o nome RSA, inicias dos sobrenomes de seus criadores. O ponto central da eficiência desta sistema é a facilidade em se encontrar números primos grandes e a dificuldade em fatorar números formados pelo produto destes números primos.

Rivest, Shamir e Adleman resolveram um problema do sistema criptográfico anterior ao sistema RSA, idealizado pelos americanos Whitfield Driffie, Martin Hellman e Ralph Merkle. Esse sistema denominado DHM já introduziu definitivamente a Matemática, com a Teoria dos Números, utilizando as congruências. O sistema serviu de base para o RSA, com a diferença que no sistema DHM havia a necessidade de troca de chaves secretas entre os indivíduos. Este

sistema era simétrico, pois precisava trocar essa informação inicial para poder ser utilizado, diferentemente do sistema RSA que é assimétrico, ou seja, há uma chave pública para cifragem e uma chave privada para decifrar mensagens.

Driffie já havia percebido esse problema e a dificuldade em se compartilhar chaves secretas com o grande fluxo de informações mundiais que já ocorria. Não era nada prático. Mas o sistema proposta com o uso de congruência era de uma simplicidade e confiabilidade incrível. Faltava apenas resolver o problema das chaves simétricas. Driffie não conseguiu, porém deixou o legado para que outros pudessem fazer.

É este contexto que se encontra a presente dissertação, com foco no sistema RSA e toda a utilização matemática envolvida. Dessa forma, além desta introdução e das considerações finais, esta dissertação tratou das seguintes temáticas em seus capítulos:

No primeiro capítulo, houve um passeio histórico a respeito da criptografia, percebendo-se a importância histórica deste sistema para os governantes de todo o mundo. Ainda, são apresentados os principais modelos de criptografia já criadas pelos homens, apontando a evolução deste modo de propagar uma informação sigilosa. Na sequência, é dado espaço para diferencias as criptografias simétricas e assimétricas, informações essenciais para entender mais à frente porque o sistema RSA funciona.

O segundo capítulo foi dedicado ao estudo dos conteúdos matemáticos essências para compreender o sistema RSA. Inicialmente, viu-se como funciona a divisão nos inteiros e o algoritmo de Euclides. A seguir, o foco serão os números primos, base central para o sistema RSA, com a apresentação do Teorema Fundamental da Aritmética, a infinidade dos números primos e a apresentação de alguns métodos para encontrar primos, além de uma passagem rápida nos números de Fermat e de Mersenne. Foi estudada também a Aritmética Modular, com a congruência e suas propriedades para, a seguir, se chegar ao Pequeno Teorema de Fermat. Finalmente, neste capítulo será explanado a respeito do Teorema de Euller.

No terceiro capítulo, foi trabalhado especificamente o sistema RSA de criptografia. Para compreendê-lo, foram discutidas as chaves públicas e privadas presentes no sistema, bem como o processo de codificação e decodificação. Um exemplo prático do sistema RSA é apresentado na sequência e foi mostrado porque o sistema é seguro.

Dessa forma, a dissertação buscou mostrar a necessidade da humanidade em transferir informações sigilosas, culminando com o sistema de criptografia atual, baseado basicamente na Matemática. Ou seja, o segredo das informações bancárias, das mensagens nas redes sociais,

a produção nas empresas, necessitam de informações matemáticas. Situação não imaginada ou não difundida a somente um século. A Matemática veio garantir o fluxo sigiloso de informações e mostra que se faz essencial na maioria dos aspectos cotidiano do homem.

CAPÍTULO 1

A História da Criptografia

Durante a história da humanidade, reis, rainhas, presidentes e generais valeram-se da comunicação eficiente para governar e fazer com que toda a região governada recebesse as orientações adequadas. Em momento de guerras, essa comunicação poderia custar a vida da população e de seus governantes. Por isso, diversos líderes tentavam passar essas mensagens de forma secreta, uma vez que a interceptação da mensagem e sua leitura pelo inimigo poderiam mudar o rumo da guerra.

> Esta busca pelo segredo levou as nações a criarem departamentos para a elaboração de códigos, responsáveis por garantirem a segurança das comunicações inventando e utilizando os melhores códigos possíveis. Ao mesmo tempo, os decifradores de códigos inimigos tentam quebrar esses códigos, para roubar seus segredos. Os decifradores de códigos são os alquimistas linguísticos, uma tribo mística que tenta invocar palavras que tenham significado a partir de uma mistura de símbolos sem sentido. A história dos códigos e de suas chaves é a história de uma batalha secular entre os criadores de códigos e os decifradores, uma corrida armamentista intelectual que teve um forte impacto no curso da história humana. (SINGH, 2007, p. 11)

Dessa forma, a comunicação eficiente e sigilosa torna-se essencial para um governo próspero, cujo receio de invasão inimiga ou vazamento de informações privilegiadas são pro-

blemas constantes na história de quase todas as civilizações. Assim, a comunicação sigilosa é ponto crucial para o desenvolvimento daquela sociedade, tornando a informação uma valiosa arma: valiosa para aqueles que a mantém em segredo e, mais ainda, para os grupos inimigos que conseguem interceptá-las e entender a mensagem a ser passada.

Os códigos utilizados, então, tornam-se ferramentas para o progresso ou derrota de uma civilização. No momento em que a informação das estratégias de guerra, das rotas do comércio, dos detalhes do exército torna-se decisivo no andamento da sociedade, a necessidade de passar essa informação a seus aliados de forma codificada representa a vida ou morte de parte da população e até(que palavra é essa) da própria comunidade.

> À medida que a informação se torna uma mercadoria cada vez mais valiosa e a revolução nas comunicações muda a sociedade, o processo de codificação de mensagens vai desempenhar um processo cada vez maior na vida diária. Hoje em dia nossas chamadas telefônicas saltam entre satélites e nossos e-mails passam por vários computadores. Ambas as formas de comunicação podem ser interceptadas facilmente, ameaçando nossa privacidade. De modo semelhante, à medida que mais negócios são realizados através da Internet, devem ser instalados mecanismos de proteção para a segurança das empresas e de seus clientes. (SINGH, 2007, p. 12)

Atualmente, o sistema de criptografia possibilita uma garantia que os dados pessoais, bancários, sigilosos, de modo que se possam fazer desde operações bancárias com altos valores até conversas íntimas, sem que se sinta insegurança, na maioria das vezes, quando se realizam tais ações. Claro que, da mesma forma que o sistema de criptografia se torna mais eficiente, também evolui o sistema de quebra dessa criptografia, de modo que se podem verificar casos de vazamento de informações por parte de algumas empresas, a exemplo dos recentes casos do Facebook, onde parte dos usuários teve 50 milhões de contas expostas com dados pessoais compartilhados, com informações dadas pelo próprio criador da rede, Mark Zuckberg.

A codificação, de acordo com Singh (2007, p. 13), é o único meio de proteger a privacidade e garantir o sucesso do mercado digital, de forma que a criptografia fornecerá os fechos e as chaves da Era da Informação. Hodiernamente, as informações circulam com impressionante velocidade e a espera pode se tornar um prejuízo ou desvantagem em relação a concorrentes comerciais.

É importante destacar que a necessidade de vantagem na antecipação de informações percorre toda a história. Sempre foi necessário saber algo antes de seus inimigos, seus concorrentes, para poder se preparar diante na nova situação apresentada. Então, quando se fala que os tempos atuais exigem uma maior velocidade no recebimento de informações, na verdade, deve-se relativizar esse entendimento, uma vez que a velocidade de informações à possibilitada de acordo com as tecnologias de cada época. Dessa forma, sobressai-se aquele que encontra formas de acelerar essa troca de informações e, obviamente, aquele que consegue interceptar e decodificar as informações das mensagens secretas.

Na verdade, as formas de cifrar as mensagens atualmente são bastante semelhantes às formas tradicionais de cifragem. Para exemplificar, tivemos a Enigma, máquina de codificação usada nas primeiras décadas do século 20, principalmente na Segunda Guerra Mundial. Singh aponta que há apenas três diferenças significativas entre a cifragem mecânica que foi a base de cifras como a Enigma e a cifragem atual:

> A primeira diferença é que uma máquina de cifras é mecânica, é limitada pelo que se pode construir na prática, enquanto o computador pode simular uma máquina de cifragem hipotética de imensa complexidade. A segunda diferença é simplesmente uma questão de velocidade. A eletrônica pode operar muito mais rapidamente do que os misturadores mecânicos: um computador programado para imitar a cifra Enigma pode cifrar em um instante uma longa mensagem. Além disso, um computador programado para efetuar uma forma de cifragem muito mais complexa pode realizar a tarefa dentro de um tempo razoável. A terceira diferença, e talvez a mais significativa, é que um computador mistura números no lugar de letras do alfabeto. Os computadores lidam apenas com números binários - sequências de um e zero conhecidas como dígitos binários, ou, abreviadamente, bits (de binary digits, em inglês). Esta conversão pode ser realizada de acordo com vários protocolos, tais como o American Standard Code for Information Interechange (Código Padrão Amercicano para Troca de Informações), conhecido pela sigla ASCII, que se pronuncia 'ass-key'. (SINGH, 2007, p. 269)

O que não mudou no decorrer de toda a história foi a necessidade de codificar e decodificar mensagens. A humanidade foi construída por disputas, sejam de grupos, de comunidades ou de países, e a necessidade de sigilo das táticas utilizadas sempre foi importante para a manutenção do poder. No momento atual, a situação não é diferente: tanto há a necessidade de

proteção das informações no âmbito de uma nação, cuja exposição desencadearia intrigas e guerras, como a necessidade de proteger dados e conversas pessoais do simples cidadão.

1.1 Evolução da Criptografia

Inicialmente, os modelos de criptografia eram mais simples, não por falta de necessidade, mas por serem suficientes para o momento histórico. A evolução do sistema criptográfico é impulsionado pelas tentativas de se decodificar estes códigos para que se possam conhecer os segredos ali criptografados. Assim, a história dos códigos secretos é marcada pela criação de um novo método seguido pela sua tentativa de resolução pelos oponentes - situação que poderia ocorrer rapidamente ou demorar séculos.

Alvarenga (2017, p. 48) considera o modelo de criptografia presente na Antiguidade "marcada por métodos extremamente simples, mas que eram suficientes para a necessidade da época". Neste rol de civilizações da Antiguidade, estão incluídos a Grécia, Egito e Roma, civilizações com desenvolvimento tecnológico e intelectual considerável, corroborando a ideia de que a criação de um sistema de criptografia mais complexo não seria problema, o que não foi feito porque as civilizações oponentes não dispunham de preparação adequada para decodificar a criptografia ingênua utilizada.

Esses métodos mais simples de criptografia presentes na Antiguidade foram esquecidos e/ou aperfeiçoados aos pouco e novos métodos mais complexos foram criados. Na Idade Média, principalmente na civilização árabe-islâmica, devido à necessidade de gerir um grande território, houve um avanço tanto nos métodos de criptografia quanto nos estudos para decodificar os sistemas criptografados. Surgem, então, estudos mais aprofundados sobre criptoanálise, cuja tentativa era entender os textos criptografadas, geralmente analisando a frequência com que as letras apareciam nas mensagens. Cada alfabeto possui letras que são usadas com mais frequência e, ao codificar, essas letras necessariamente transformam-se em um mesmo código. Assim, os criptoanalistas analisavam a frequência dos códigos utilizados, fazendo a desencriptação até o texto formar um texto coerente.

Quando as potências europeias começaram a se lançar ao mar em busca de especiarias e, consequentemente, conquistando novos mundo, começou também a disputa mais acirrada por estes elementos. As nações europeias passam, então, por períodos de aliança e, sequencialmente, por rompimento dessas alianças. A necessidade de comunicação, ligado à desconfiança

a outras nações, levam ao estudo mais sério da criptografia, fazendo evoluir os sistemas monoalfabéticos para os polialfabéticos, como argumenta Alvarenga (2017, p. 51). Este autor ainda acrescenta:

> A partir do século, XVI, com o início do Renascimento, a contínua animosidade entre os países europeus e as contínuas alianças e rompimento de alianças levara a que os governantes procurassem meios de ler a correspondência diplomática dos inimigos, atuais ou futuros. Com este objetivo, criaram-se vários gabinetes, chamados Câmaras Negras, cujo objetivo era tanto decifrar os códigos secretos, quanto criar códigos invioláveis. (ALVARENGA, 2017, p. 50)

Nos séculos seguintes, a criptografia foi tendo um enorme avanço, justamente pela disputa entre as nações. Essa necessidade de proteger suas conquistas, territórios e, ao mesmo tempo, descobrir as ideias e intenções dos oponentes, possibilitou esse progresso na criptografia: cada vez mais se construíam códigos mais complexos e, do outro lado, a arte de decifrar tais códigos ia se aperfeiçoando.

Com a criação do telégrafo, em 1838, pelo norte-americano Samuel F.B. Morse, as informações começaram a chegar ao destino com rapidez. O código Morse utiliza justamente um processo de substituição de letras por pontos e traços. Logo depois disso, surgiu a necessidade de codificar a mensagem telegrafada. Deve-se perceber que para codificar uma mensagem há a necessidade de o receptor ter a informação sobre como decodificar, e esta informação também poderia ser interceptada pelo oponente. Mais a frente será discutida essa problemática que percorreu toda a história da criptografia.

Com o aumento no fluxo de mensagem com a criação do telégrafo, tornou-se mais urgente encontrar novas formas de codificação para tornar essa ação mais rápida. Neste ponto da história, surgem as máquinas de codificação assistida, como aponta Alvarenga (2017, p. 52). Ele aduz que a demanda por rapidez tanto na codificação quanto na decodificação das mensagens, levou a o emprego de processos automáticos, através destas máquinas de codificação.

No século, há de se destacar o aperfeiçoamento das máquinas de codificação, principalmente na Alemanha, na Inglaterra e nos Estados Unidos. Até meados do século XIX, não havia quase nenhuma informação a respeito da criptografia estadunidense, que ganha destaque após a primeira guerra mundial. A seguir, os principais sistemas de criptografia destas potências econômicas e bélicas, com as máquinas criadas tal finalidade.

Talvez a máquina mais famosa criada para codificar informações seja a "Máquina Enigma", de origem na Alemanha nazista, surgida em 1918, configurada pelo alemão Arthur Scherbius. Depois da primeira guerra mundial. O exército alemão sentiu necessidade de cifrar suas informações para tornar seguras as ordens estabelecidas. Alvarenga (2017, p. 61) fala que, a partir de 1925 o exército alemão concluiu que esta máquina servia aos propósitos bélicos alemãs, e em 1928 começou a usar uma versão dela, a Enigma G97 que tinha como garantidor de segurança a troca periódica mensal de suas chaves.. A máquina contava com três rotores, cuja função era misturar o sistema para modificar o código a depender do interesse.

Quanto à criptografia inglesa, destaca-se a criação da máquina "Typex", que tinha como vantagem sobre a "Enigma"maior quantidade de rotores. A cronologia da invenção inglesa acontece da seguinte forma:

> Em agosto de 1934 surgiu o primeiro protótipo de uma máquina codificadora, a Typex Mark I. Por volta de 1937 já havia cerca de 30 máquinas em operação. A Typex era baseada na máquina comercial Enigma que incorporava um número de características a mais, para aumentar a segurança. Por exemplo, tinha cinco rotores (contra três ou quatro na Enigma), sendo que os dois primeiros ficavam estacionários, embora pudessem ser acionados a mão. Os rotores continham múltiplos entalhes que poderiam eventualmente acionar os rotores vizinhos. (ALVARENGA, 2017, p. 62)

Quando se fala sobre a criptografia estadunidense, pouco se tem informação sobre o sistema apresentado antes do século XX, como mencionado anteriormente. Após a primeira guerra mundial, os Estados Unidos aprimoraram a segurança de seus códigos, culminando com a criação da máquina de cifras Sigaba, mais utilizada pela Marinha e pelo Exército. Ela possui o mesmo processo da Enigma, com um sistema de rotores para a codificação das mensagens. Porém, a máquina estadunidense possui quinze rotores, com clara vantagem sobre os três rotores da máquina alemã. Ainda, a Sigaba possuía um dispositivo que fazia movimentar os rotores de forma aleatória, dificultando sua decodificação pelos oponentes.

Diante da história da evolução da criptografia, percebe-se que os sistemas foram se tornando mais complexos a partir da necessidade das civilizações e das situações em que o sistema era decodificado pelos oponentes. A informação secreta sempre foi essencial para manter uma civilização em vantagem em relação às demais. Esta vantagem tornou-se determinante no momento em que as informações circulam mais rapidamente. As nações aperfeiçoaram os sistemas

para ter seu domínio predominante e, posteriormente, as empresas passaram a necessitar destes sistemas de codificação. Atualmente, diversas empresas, como bancos e redes sociais, investem para tornar seus negócios cada vez mais seguros e poder alavancar um maior número de clientes. A forma como essas empresas realizam a codificação e decodificação das informações é estudado em capítulo posterior.

1.2 Modelos de Criptografia

Como falado anteriormente, os sistema de codificação foram se aperfeiçoando ao longo dos tempos de forma que a complexidade foi aumentando na proporção em que os oponentes instituíam formas eficazes para decodificar seus códigos. É importante conhecer alguns principais modelos de criptografia utilizados pelas civilizações, que vão desde a utilização de substituição de letras - mono ou polialfabético -, por transposição e por esteganografia.

A encriptação por substituição de letras serviu por muitos séculos aos interesses das civilizações, apesar de seu sistema ser bastante simples, principalmente o monoalfabético que, como o próprio nome indica, utiliza unicamente um alfabeto para fazer a substituição de letras. Na Roma Antiga, Júlio César usava um simples sistema onde cada letra do alfabeto era trocada por outra letra do mesmo alfabeto. Esse procedimento era bastante utilizado em Roma, quando César precisava se comunicar com os generais.

Esse mesmo modelo foi utilizado pelos hebreus, com um detalhe que possibilitava uma pequena vantagem na forma utilizada por César: "as 22 letras do alfabeto eram escritas onze a onze, e cada letra seria substituída pela correspondente superior ou inferior"(ALVARENGA, 2017, p. 8). Esse modelo era chamado de "atbash", sendo relativamente fácil e ser decodificado devido à simetria na troca das letras e na utilização do alfabeto de cima para baixo para codificar e o inverso para decodificar.

O Código Mecânico, também criado utilizando o sistema monoalfabético de substituição de letras, valia-se de um posicionar as letras sequencialmente em quatro grades, de forma que a letra seria substituída pelos traços da grade ao redor da letra, adicionado a pontos que serviam para diferenciar grupos de letras. Outro modelo monoalfabético, criado por italiano Leon Battista Alberti, utilizou um disco de cifras, com diâmetros diferentes e em que estão inseridos os alfabetos. Com o movimento dos discos em torno de um eixo, criava-se com facilidade nossas cifras a partir do giro desses discos.

O aperfeiçoamento dos métodos de codificação ocorre, como se discutiu anteriormente, com a evolução dos métodos de decodificação, de forma que o surgimento do sistema polialfabético foi introduzido justamente porque o sistema monoalfabético deixou de exercerem com eficiente seu objetivo. Como se deve imaginar, o sistema polialfabético utiliza vários alfabetos embaralhados, dificultando o deciframento da mensagem.

Além dos métodos de substituição de letras, as civilizações utilizaram situações em que as letras do alfabeto eram deslocadas de sua posição original ou ainda sistemas em que cada letra tinha a função de uma palavra ou frase. Este último era a chamada "encriptação por esteganofragia", que, de acordo com ALVARENGA o método consistia em:

> Uma forma bastante astuciosa de esteganografia foi utilizada pelo abade beneditino Johannes Trithemius. Seu sistema era conhecido como as Ave Marias. Ele era composto por 14 alfabetos, nos quais, a cada letra, corresponde uma palavra ou grupo de palavras. O texto cifrado final, quando lido normalmente, apresenta ser um texto coerente, como se fosse uma oração ou glorificação religiosa, em latim. (2017, p. 34)

O deslocamento das letras de sua posição original também foi bastante utilizado, sendo chamado de método de "transposição". Esse sistema tinha uma vantagem em relação ao sistema de substituição de letras, uma vez que as palavras de um idioma possuem agrupamentos mais comuns de letras, ou até mesmo, obrigatórios, facilitando a descoberta por quais letras foram substituídas as originais. Veja o exemplo da letra "Q"no idioma português: obrigatoriamente deve haver a letra "U"na sequência. Esse fato poderia facilitar a decodificação da mensagem. Com transposição das letras, essa junção não vai acontecer, tornando mais difícil de decodificar a mensagem.

1.3 Criptografia simétrica e assimétrica

Todas as possibilidades de codificar uma mensagem apresentadas até então se baseiam em algum método onde o emissor deve compartilhar com o receptor a mensagem e também o algoritmo para decodificar a mensagem. Ou seja, antes de alguém compartilhar um segredo com outro, precisaria ter compartilhado outro segredo. Imaginando que um professor X da Universidade Federal de Sergipe - UFS -, no polo de São Cristóvão, deseja transmitir uma mensagem secreta para um professor Y da UFS, no polo de Itabaiana, antes ele irá precisar

estabelecer os critérios para que a mensagem possa ser decifrada, ou seja, precisa indicar qual o algoritmo para decodificação da mensagem. Esse algoritmo é a chave da cifragem. Neste contexto, a forma de decodificar é apenas o algoritmo inverso da codificação. Deve-se perceber que este tipo de cifragem corre risco se a chave privada se tornar pública.

Este tipo de sistema onde se usa a mesma chave para cifrar e decifrar a mensagem é chamada de simétrica. Há também o sistema assimétrico, mais utilizado atualmente, que se baseia numa chave pública e em uma privada. Neste caso, todos conhecem a forma para codificar uma mensagem para alguém, a chave pública, mas somente a pessoa tem a chave privada para decodificar.

> [No sistema simétrico], ambos, o emissor e o receptor, possuem, efetivamente, um conhecimento equivalente e ambos usam a mesma chave para cifrar e decifrar - seu relacionamento é simétrico. Por outro lado, num sistema de chave assimétrico, como o próprio nome sugere, a chave de cifragem e a chave de decifragem não são idênticas. Em uma cifra assimétrica, se Alice sabe a chave de cifragem, ela pode cifrar mas não pode decifrar a mensagem. Para decifrá-la, ela deve ter acesso à chave de decifragem. Esta distinção entre cifragem e decifragem é o que torna a cifra assimétrica tão especial. (SINGH, 2007, p. 294)

Na criptografia assimétrica, existem duas chaves, a chave pública e a chave privada. Para cifrar uma mensagem para alguém, precisa-se ter a chave pública, de conhecimento geral, e utilizar esta chave para cifrar a mensagem; o receptor decifra a mensagem secretamente. Voltando ao caso dos professores da UFS, para o professor X enviar uma mensagem cifrada ao professor Y, ele utiliza a chave de ciframento pública PY e envia a mensagem para o professor Y. Este tem a chave de deciframento secreta SY e decodifica a mensagem. A chave não está presente em nenhum nível na mensagem, sendo os algoritmos de ciframento e deciframento iguais. Assim, garante-se o sigilo da mensagem entre o emissor e o receptor, porque somente ambos têm conhecimento da chave secreta.

Alvarenga (2017, p. 118) compara o sistema a partir de três características: velocidade, segurança e limitações. O autor foca que cada sistema - simétrico ou assimétrico - possui vantagens e desvantagens, podendo o emissor e o receptor valer-se de um ou outro ou até de ambos os sistemas. Deve-se lembrar de que a problemática do sistema sistemático é o fato de o emissor também precisar compartilhar a chave privada para que o destinatário possa decifrar a mensagem. Para aumentar a segurança, o receptor poderia compartilhar a chave privada por

uma mensagem cifrada pelo sistema assimétrico, garantido, assim, o sigilo. E por que não enviar toda a mensagem pelo sistema assimétrico? Claro que é uma das possibilidades, porém os algoritmos da chave simétrica possuem menor complexidade, influenciando diretamente em sua velocidade e custo computacional.

Hodiernamente, o sistema RSA, anagrama formado pelos criadores - Rivest, Shamir e Adleman -, é o sistema mais utilizado. Baseia-se no uso do produto de grandes números primos para a construção da chave, sendo que a decodificação somente seria possível conhecendo tais fatores primos. A dificuldade de decifrar tal código sem a chave privada reside no fator de dificilmente fatorar um número produto de fatores primos grandes unicamente na "força bruta".

> Em teoria, quando o RSA foi inventado em 1977, ela ofereceu um antídoto para o cenário do Grande Irmão, porque os indivíduos seriam capazes de criar suas próprias chaves públicas e particulares, enviando e recebendo mensagens perfeitamente seguras. Contudo, na prática, havia um grande problema porque o processo real da cifragem pela RSA exigia uma boa dose de poder de computação em comparação com as formas simétricas de cifragem, tais como o DES. Consequentemente, na década de 1980, apenas o governo, os militares e as grandes empresas possuíam computadores suficientemente poderosos para rodar a RSA. Não é surpresa que a RSA Data Security Inc., a empresa fundada para comercializar a RSA, tenha desenvolvido seus produtos de cifragem tendo em mente apenas esses mercados. (SINGH, 1999, p. 324)

Foram Ronald Rivest, Adi Shamir e Leonard Adleman do Laboratório de Ciência de Informação do Massachusetts Institute of Technolgy (MIT), que deram em 1978 o passo decisivo para a implementação do sistema criptográfico com sistemas assimétricos, idealizado por Driffie. (HEFEZ, 2016, p. 274)

A esta altura, faz-se importante mencionar o código ASCII, American Standford Code for Information Interchange, de tradução Código Padrão Americano para o Intercâmbio de Informação. Este código transformou as informações para o sistema adotado pelos computadores - o sistema binário. Não configura um sistema de cifragem, mas uma tradução para o código binário os símbolos mais utilizados. O RSA é um sistema que cifra a mensagem anteriormente codificada em ASCII e somente o seu devido destinatário pode decifrá-la.

> Em traços gerais, são gerados dois pares de números - as chaves - de tal forma

> que uma mensagem encriptada com o primeiro par possa ser apenas desencriptada com o segundo par; entretanto, o segundo número não pode ser derivado do primeiro. Esta propriedade assegura que o primeiro número possa ser divulgado a alguém que pretenda enviar uma mensagem encriptada ao detentor do segundo número, já que apenas essa pessoa pode desencriptar. (ALVARENGA, 2017, p. 174)

O sistema RSA é um dos mais seguros atualmente, sendo utilizado nas principais trocas de mensagens de teor privado e/ou sigiloso. Diante de uma sociedade onde a quantidade de conteúdo é aumentada exponencialmente e grande parte deste conteúdo é sigiloso, como mensagens pessoais, senha de contas bancárias e cartões de crédito, torna-se essencial proteger esse sistema de troca. O sistema RSA parece ser o auge entre os sistemas que buscam codificar informações, originado na criação dos primeiros sistemas simples de substituição de letras, surgidos na Antiguidade.

Para entender como funciona mais detalhadamente o sistema de criptografia RSA, temática abordada no quarto capítulo, faz-se necessário conhecer alguns conceitos matemáticos, especialmente da Aritmética, para poder visualizar de forma mais clara como ocorrem a troca de mensagens. Este sistema tem por concepção principal a relativa facilidade que se encontram números primos grandes e da enorme dificuldade de fatorar um número que seja produto de dois destes números primos. Assim, o próximo capítulo se debruça sobre tais conhecimentos matemáticos necessários para entender mais a fundo o sistema RSA.

CAPÍTULO 2

Números Primos e a Aritmética Modular

2.1 Divisão nos Inteiros e Algoritmo de Euclides

A divisão nem sempre é possível entre números inteiros, podendo expressar tal possibilidade através da relação de divisibilidade. E mesmo quando não existir essa relação de divisibilidade entre dois números inteiros, ainda tem-se a possibilidade de uma divisão com resto, chamado de divisão euclidiana.

Dados dois números inteiros a e b, diz-se que a divide b, na forma $a \mid b$, quando existir um $c \in \mathbb{Z}$ tal que $b = ca$. Assim, diz-se que a é um divisor de b, b é um múltiplo de a, ou ainda, b é divisível por a. Caso não exista nenhum número inteiro c tal que $b = ca$, diz-se que b não é divisível por a. Diz-se que $4 \mid 12$, pois $12 = 4 \times 3$, temos também que 12 é múltiplo de 4 e 4 é divisor de 12. E 4 não divide 14, pois não existe $c \in \mathbb{N}$ tal que $14 = c \times 4$ é satisfeita.

Sejam $a, b, c \in \mathbb{Z}$. Tem-se que:

(i) $1 \mid a$, $a \mid a$ e $a \mid 0$. Isso decorre das igualdades $a = a \times 1$, $a = 1 \times a$ e $0 = 0 \times a$. Como consequência dessas proposições, todo número inteiro a é divisível por ± 1 e por $\pm a$.

(ii) $0 \mid a \Leftrightarrow a = 0$. Suponha que $0 \mid a$; logo existe $c \in \mathbb{Z}$ tal que $a = c \times 0$.

(iii) se $a \mid b$ e $b \mid c$, então $a \mid c$. Para $a \mid b$ e $b \mid c$, implica que existem $f, g \in \mathbb{Z}$, tais que

$b = fa$ e $c = gb$. Substituindo o valor de b da primeira equação na outra, obtém-se:

$$c = gb = g(fa) = (gf)a,$$

o que mostra que $a \mid c$.

Dados dois números inteiros a e b, ambos diferentes de zero e pertencentes aos números inteiros, pode-se associar cada um desses números ao conjunto de divisores de a e b, ou $D(a)$ e $D(b)$. Fazendo a interseção desses conjuntos nunca resultará em conjunto vazio, uma vez que sempre estará presente, ao menos, o número 1. Considerando o maior número na interseção entre os divisores de dois números, tem-se o *Máximo Divisor Comum*.

Assim, dado $d \geq 0$, d é *máximo divisor comum* de a e b se atende a essas duas propriedades: i) d é um divisor comum de a e b; e ii) d é divisível por todo divisor de a e b. Para a notação de máximo divisor comum, será utilizado $mdc(a, b) = d$, de forma que não importa a ordem de a e b, então $mdc(a, b) = mdc(b, a)$.

É importante destacar que é sempre possível efetuar a divisão de a por b com resto, mesmo quando um número inteiro b não divide o número inteiro a. Em seu livro *Elementos*, Euclides já expressava tal possibilidade, sendo tal algoritmo nomeado como uma divisão euclidiana. Essa divisão serve de base para o *Algoritmo de Euclides*, que pode ser usado de forma eficiente para o cálculo do máximo divisor comum entre dois números.

Teorema 2.1.1. *Divisão Euclidiana. Sejam a e b dois números inteiros com $b \neq 0$. Existem dois únicos números q e r tais que*

$$a = bq + r, \ \ com \ \ 0 \leq r <\mid b \mid .$$

A unicidade de q e r é comprovada da seguinte forma. Suponha que

$$a = bq + r = bq' + r', \ \text{ onde } \ q, q', r, r' \in \mathbb{Z}, \ \text{ com } \ 0 \leq r <\mid b \mid \ \text{ e } \ 0 \leq r' < b.$$

Assim, tem-se que

$$-\mid b \mid < -r \leq r' - r \leq r' <\mid b \mid .$$

Por outro lado, $b(q - q') = (r' - r)$, o que implica que $\mid b \mid\mid q - q' \mid = \mid r' - r \mid < \mid b \mid$, sendo possível somente se $q = q'$ e, consequentemente, $r = r'$.

No teorema acima, os números q e r são chamados, respectivamente, de quociente e resto da divisão de a por b. Esse quociente e resto sempre existem e são únicos.

Nos livros escolares de matemática, o método utilizado para o cálculo do mdc entre os números a e b consiste em listar todos os divisores de a e todos os divisores de b. A seguir, destaca-se o maior número na interseção entre os dois conjuntos. Não há nada de errado com esse método, porém, torna-se impraticável com números grandes.

O objetivo do algoritmo euclidiano é calcular o máximo divisor comum entre dois números inteiros. Este algoritmo é descrito por Euclides nas proposições 1 e 2 do Livro 7 dos *Elementos*. Dados a e b inteiros positivos e que $a \geq b$, deseja-se calcular o máximo divisor comum entre a e b. O *algoritmo euclidiano* consiste em dividir a por b, achando o resto r_1. Se $r_1 \neq 0$, divide-se b por r_1, obtendo r_2. Se $r_2 \neq 0$, divide-se r_1 por r_2, obtendo o resto r_3. E assim sucessivamente até que o resto seja 0. O último resto diferente de zero é o máximo divisor comum entre a e b. Veja o exemplo no quadro abaixo com o cálculo do máximo divisor comum entre 240 e 162.

-	1	2	13
240	162	78	6
78	6	0	-

De acordo com a regra exposta anteriormente, o máximo divisor comum entre 240 e 162 é 6, ou seja, $mdc(240, 162) = 6$, que é o último resto não nulo da sequência de divisões. Essa sequência sempre chega a um resto zero, pois, caso contrário, ter-se-ia uma sequência de números naturais $b > r_1 > r_2 > r_3 \ldots$ que não possui menor elemento, o que não é possível pelo Princípio da Boa Ordenação.

2.2 Números Primos

Os números que somente possuem como divisores o número 1 e ele próprio são considerados *números primos*. Dessa forma, os números $2, 3, 5, 7, 11, 13, 17 \ldots$ são números primos. Aqueles números que possuem mais de dois divisores são chamados de números compostos, como os números $4, 6, 9, 15 \ldots$. Os números primos sempre despertaram grande interesse aos estudiosos e diversos questionamentos a respeito ainda são um mistério.

Muitos problemas ainda estão em aberto quando o assunto é número primo. Não se tem

resposta sobre a possibilidade de existirem infinitos números primos gêmeos, ou seja, pares de números ímpares primos, cuja diferença é dois entre eles. Também não se sabe se a sequência de Fibonacci (sequência iniciada em 0 e 1 em que cada termo equivale à soma dos dois números subsequentes) contém infinitos números primos. Ou ainda, sobre a possibilidade de sempre existir um número primo entre n^2 e $(n+1)^2$ para qualquer $n \in \mathbb{N}$. Há ainda o seguinte questionamento: para $n = 0, 1, 2, 3, 4, \ldots, 40$, tem-se que $n^2 - n + 41$ é primo. Para $n = 41$, esta afirmação não procede; porém o questionamento é sobre a existência de infinitos números primos dessa forma.

O que se percebe é que os números primos ainda são cercados de problemas em aberto e outros tantos recentemente esclarecidos. Esse mistério sempre despertou a curiosidade dos estudiosos e, atualmente, é a base para o Sistema de Criptografia RSA, temática estudada no próximo capítulo. Antes disso, serão estudados mais alguns pontos sobre os números primos e outros conteúdos matemáticos essenciais para entender como funciona o Sistema RSA.

2.2.1 Teorema Fundamental da Aritmética

Como visto, um número natural maior do que 1 que só possui como divisores 1 e ele próprio é um número primo. Esses números são suficientes para gerar todos os outros. Isso consiste no *Teorema Fundamental da Aritmética*, demonstrado abaixo por Hefez (2016, p. 123)

Teorema 2.2.1 (Teorema Fundamental da Aritmética)**.** *Todo número natural maior do que* 1 *ou é primo ou se escreve de modo único como um produto de números primos.*

Demonstração. Usaremos a segunda forma do Princípio de Indução. Se $n = 2$, o resultado é obviamente verificado. Suponha-se o resultado válido para todo número natural menor do que n e vamos provar que vale para n. Se o número n é primo, nada temos a demonstrar. Suponhamos, então, que n seja composto. Logo, existem números naturais n_1 e n_2 tais que $n = n_1 n_2$, com $1 < n_1 < n$ e $1 < n_2 < n$. Pela hipótese de indução, temos que existem números primos $p_1, \ldots, p_r$ e $q_1, \ldots, q_s$ tais que $n_1 = p_1 \ldots p_r$ e $n_2 = q_1 \ldots q_s$. Portanto, $n = p_1 \ldots p_r q_1 \ldots q_s$.

Vamos, agora, provar a unicidade da escrita. Suponha que tenhamos $n = p_1 \ldots p_r = q_1 \ldots q_s$, onde os p_i e os q_j são números primos. Como $p_1 | q_1 \ldots q_s$, temos que $p_1 = q_j$ para algum j, que, após reordenamento de $q_1, \ldots, q_s$, podemos supor que seja q_1. Portanto, $p_2 \ldots p_r = q_2 \ldots q_s$. Como $p_2 \ldots p_r < n$, a hopótese de indução acarreta que $r = s$ e os p_i e q_j são iguais aos pares.

2.2.2 Infinidade dos Primos

Euclides, no Livro IX dos *Elementos*, já demonstrava que os números primos são infinitos. Ele faz a demonstração através da redução ao absurdo. Segue abaixo o teorema e a demonstração:

Teorema 2.2.2. *Existem infinitos números primos.*

Demonstração. Suponha que exista um número finito de números primos $p_1, p_2, \ldots, p_r$. Considere o número natural $n = p_1 p_2 ... p_r + 1$. Pelo Teorema Fundamental da Aritmética, o número n possui um fator primo p que, portanto, deve ser um dos $p_1, p_2, \ldots, p_r$ e, consequentemente, divide o produto $p_1 p_2 ... p_r$. Mas isto implica que p divide 1, o que é absurdo.

Sabe-se, então, que existem infinitos números primos. Com essa informação, outros problemas importantes (sem resposta até o momento) surgem:

(i) há um padrão na distribuição dos números primos?

(ii) quando são considerados números cada vez maiores, a presença dos números primos aumenta ou diminui?

(iii) há alguma regularidade nessa "densidade"?

(iv) há alguma forma de medir essa "densidade"?

Muitas foram as tentativas de se responder a estas e outras perguntas, além da tentativa de se contabilizar os números primos até certa numeração, porém, não se obteve método eficiente para tal finalidade. A seguir, será visto algumas tentativas.

2.2.3 Métodos para Encontrar Números Primos

Um dos métodos mais antigos para se encontrar números primos é o método chamado *Crivo de Eratóstenes*, criado pelo matemático grego Eratóstenes, antes de Cristo. Nicômaco descreve o crivo de Eratóstenes em seu livro *Aritmética* da seguinte forma:

O método para obtê-los [os números primos] é chamado por Eratóstenes uma peneira, porque tomamos os números ímpares misturados de maneira indiscriminada e, por este método, como se fosse pelo uso de um instrumento ou peneira, separamos os primos ou indecomponíveis dos secundários ou compostos. (NICÔMACO, apud COUTINHO, 2014, p. 62).

O crivo determina todos os números primos até um certo número natural n. O crivo funciona da seguinte forma: Listam-se os números ímpares de 3 a n, isso porque 2 é o único número par primo. O primeiro número da lista é 3, destaca-se este e riscam-se os demais números da lista, de 3 em 3. Assim, riscam-se todos os múltiplos de 3. Na sequência, passa-se para o menor elemento dessa lista, com exceção de 3, que não tenha sido riscado, ou seja, o número 5. Destaca-se o número 5 e riscam-se os demais números da lista, de 5 em 5. Dessa forma, eliminam-se todos os múltiplos de 5. Segue-se nesse ritmo até chegar ao número n.

Por exemplo, se $n = 48$, a lista de números é

$$3\ \ 5\ \ 7\ \ 9\ \ 11\ \ 13\ \ 15\ \ 17\ \ 19\ \ 21\ \ 23\ \ 25\ \ 27\ \ 29\ \ 31\ \ 33\ \ 35\ \ 37\ \ 39\ \ 41\ \ 43\ \ 45\ \ 47.$$

Ao final da primeira passagem, riscando de 3 em 3, tem-se

$$3\ \ 5\ \ 7\ \ \underline{9}\ \ 11\ \ 13\ \ \underline{15}\ \ 17\ \ 19\ \ \underline{21}\ \ 23\ \ 25\ \ \underline{27}\ \ 29\ \ 31\ \ \underline{33}\ \ 35\ \ 37\ \ \underline{39}\ \ 41\ \ 43\ \ \underline{45}\ \ 47.$$

Ao final da segunda passagem, riscando de 5 em 5, tem-se

$$3\ \ 5\ \ 7\ \ \underline{9}\ \ 11\ \ 13\ \ \underline{15}\ \ 17\ \ 19\ \ \underline{21}\ \ 23\ \ \underline{25}\ \ \underline{27}\ \ 29\ \ 31\ \ \underline{33}\ \ \underline{35}\ \ 37\ \ \underline{39}\ \ 41\ \ 43\ \ \underline{45}\ \ 47.$$

Ao final da terceira passagem, riscando de 7 em 7, nada é modificado na lista, assim como pela quarta passagem, riscando de 11 em 11. Na verdade, mais nenhum número será eliminado e a lista de números primos menores que 48 é

$$3\ \ 5\ \ 7\ \ 11\ \ 13\ \ 17\ \ 19\ \ 23\ \ 29\ \ 31\ \ 37\ \ 41\ \ 43\ \ 47.$$

Há de se destacar que a partir da terceira passagem não se eliminou mais nenhum número, ou seja, todas as passagens seguintes foram desnecessárias. O que leva a imaginar que não há necessidade de riscar os números até chegar ao número n, ou seja, há algum momento em que se podem parar de riscar os números.

Coutinho (2014, p. 63-64) aponta que é possível parar de riscar bem antes de se chegar a n. Ele aduz que, se m é um inteiro da lista, então $m \leq n$. se m for composto, então terá um fator menor ou igual a $\sqrt{m}$. Porém, $\sqrt{m} \leq \sqrt{n}$. Dessa forma, qualquer número composto da lista tem um fator menor ou igual a $\sqrt{n}$. Conclui-se que não há necessidade de riscar números de q em q, quando $q >$ $\sqrt{n}$. No exemplo anterior poderia parar em de riscar de 5 em 5, uma vez

que $\sqrt{48} < 7$.

Claro que o crivo de Eratóstenes não é viável quando se deseja encontrar números primos muito grandes. Essa nem é a intenção do crivo; este encontra os números primos até n. Dessa forma, não se pode dizer que o crivo não seja eficiente; na verdade ele é, pois vale-se de algoritmos simples para se chegar ao objetivo.

2.2.4 Números de Fermat e de Mersenne

Os números de Fermat são da forma $F_n = 2^{2^n} + 1, n = 0, 1, 2, \ldots$. Fermat conjecturou que os números dessa forma seriam todos primos. Para os valores de n entre 0 e 6, os números de Fermat são $3, 5, 17, 257, 65537, 4294967297$ e 18446744073709551617. De fato, para $n = 0, 1, 2, 3, 4$, os números de Fermat são primos e, por isso, são chamados *primos de Fermat*. Porém, quando $n = 5$, tem-se que $F_n = 2^{2^5} + 1 = 4294967297 = 641 \times 6700417$, sendo composto, desacreditando a afirmação de Fermat. Ainda não se sabe se existem outros primos de Fermat além desses cinco primeiros.

Outros números especiais são os *números de Mersenne*, escritos da seguinte forma $M_p = 2^p - 1$, onde p é um número primo. Hefez (2016, p. 145) explica o seguinte sobre os números de Mersenne:

No intervalo $2 \leq p \leq 5000$ os números de Mersenne que são primos, chamados de *primos de Mersenne*, correspondem aos seguintes valores de p:

$$2, 3, 5, 7, 13, 19, 31, 61, 89, 107, 127, 521, 607, 1279, 2203, 2281, 3217, 4253 \text{ e } 4423.$$

Até o ano de 2016, o maior número primo conhecido era o número de Mersenne $M_{57885161}$, descoberto em janeiro de 2013, e que possui no sistema decimal 17425170 de dígitos.

Para se encontrar números primos grandes, as fórmulas exponenciais são as mais propícias. Através dos números de Fermat e, principalmente, dos números de Mersenne os matemáticos vêm encontrando números primos cada vez maiores; inclusive o maior número primo até 2016, como falado anteriormente, é um número de Mersenne. No número de Mersenne, se n é composto, então $M(n)$ também será composto.

Demonstração. Se $n = rq$, então

$$2^n - 1 = (2^r)^q - 1 = (2^r - 1)(2^{r(q-1)}) + 2^{r(q-2)} + \cdots + 2^r + 1).$$

Portanto, $M(r)$ é fator de $M(n) = M(rq)$. Claro que $M(q)$ também é um fator de $M(n)$. Então, somente se encontram números primos através dos números de Mersenne com $M(n)$, com n primo. O que não garante que todos os números de Mersenne com $M(n)$, com n primo sejam também primos. O que garante é um caminho possível para se encontrar cada vez maiores números primos.

2.3 Aritmética modular

2.3.1 Congruência e Propriedade

Diz-se que dois números a e b são congruentes módulo m quando o resto da divisão de a e b por m resulta no mesmo número, sendo representado por $a \equiv b \; mod \; m$. Como exemplo, tem-se que $31 \equiv 43 \; mod \; 4$, pois o resto da divisão de 31 por 4 é o mesmo resto da divisão de 43 por 4. Quando essa afirmação não for verdadeira, diz que não são congruentes, ou seja, $a \not\equiv b \; mod \; m$.

Hefez (2016, p. 166) aduz que não há necessidade de efetuar a divisão euclidiana para saber se dois números são congruentes módulo m. Basta aplicar a seguinte proposição:

Proposição 2.3.1. *Suponha que $a, b, m \in \mathbb{Z}$, com $m > 1$. Tem-se que $a \equiv b \mod m$ se, e somente se, $m|b - a$.*

Demonstração. Sejam $a = mq + r$, com $0 \leq r < m$ e $b = mq' + r'$, com $0 \leq r' < m$, as divisões euclidianas de a e b por m, respectivamente. Logo, $b-a = m(q'-q)+(r'-r)$. Portanto, $a \equiv b \mod m$ se, e somente se, $r = r'$ o que, em vista da igualdade acima, é equivalente a dizer que $m|b - a$, já que $|r - r'| < m$.

A congruência módulo um número inteiro m é uma relação de equivalência. Decorre então as seguintes situações:

(i) $a \equiv a \; mod \; m$;

(ii) se $a \equiv b \; mod \; m$, então $b \equiv a \; mod \; m$;

(iii) se $a \equiv b \; mod \; m$ e $b \equiv c \; mod \; m$, então $a \equiv c \; mod \; m$;

(iv) se $a \equiv b \; mod \; m$ e $c \equiv d \; mod \; m$, então $ac \equiv bd \; mod \; m$;

(v) se $a \equiv b \; mod \; m$, então $a^n \equiv b^n \; mod \; m$.

Para demonstrar esta última proposição, basta fazer por indução em n. Para a proposição iv, Hefez (2016, pág. 168) indica que basta notar que $bd - ac = d(b - a) + a(d - c)$, de onde se conclui que $m|bd - ac$.

A congruência é uma relação importante de equivalência na adição e na multiplicação dos números inteiros, como serão mostradas nas proposições a seguir:

Proposição 2.3.2. *Sejam* $a, b, c, d, m \in \mathbb{Z}$*, com* $m > 1$.

(i) Se $a \equiv b \, mod \, m$ *e* $c \equiv d \, mod \, m$*, então* $a + c \equiv b + d \, mod \, m$.

(ii) Se $a \equiv b \, mod \, m$ *e* $c \equiv d \, mod \, m$*, então* $ac \equiv bd \, mod \, m$.

Demonstração. Suponhamos que $a \equiv b \, mod \, m$ e $c \equiv d \, mod \, m$. Logo, tem-se que $m|b - a$ e $m|d - c$. i) Observa-se que $m|(b - a) + (d - c)$ e, portanto, $m|(b + d) - (a + c)$, o que prova este resultado; ii) Deve-se notar que $bd - ac = d(b - a) + a(d - c)$ e concluir que $m|bd - ac$.

A proposição seguinte mostra que o cancelamento sempre é válido com relação à adição na congruência modular.

Proposição 2.3.3. *Sejam* $a, b, c, m \in \mathbb{Z}$*, com* $m > 1$. *Tem-se que* $a + c \equiv b + c \, mod \, m \Leftrightarrow a \equiv b \, mod \, m$.

Demonstração. Se $a \equiv b \, mod \, m$, segue-se, das proposições anteriores que $a+c \equiv b+c \, mod \, m$, pois $c \equiv c \, mod \, m$. Reciprocamente, se $a + c \equiv b + c \, mod \, m$, então, $m|b + c - (a + c) \equiv b + c \, mod \, m$, o que implica que $m|b - a$ e, assim, $a \equiv b \, mod \, m$.

Como visto anteriormente, todo número inteiro é congruente módulo m ao seu resto. Assim, á congruente a um dos números $0, 1, 2, 3, \ldots, m - 1$, sendo que dois desses números diferentes não são congruentes módulo m. Com essa informação, é possível afirmar que, para encontrar o resto de uma divisão de um inteiro por m, somente é preciso encontrar um número natural de 0 a $m - 1$ que seja congruente a a módulo m.

Hefez (2016, p. 167) aponta que um *sistema completo de resíduos* módulo m refere-se a todo conjunto de números inteiros cujos restos pela divisão por m são os números $0, 1, 2, 3, \ldots, m-1$, sem repetições e numa ordem qualquer. Assim, um sistema completo de resíduos módulo m possui m elementos.

As proposições apresentadas até então servirão de apoio para a apresentação de um teorema importante: *Pequeno Teorema de Fermat.*

Teorema 2.3.1 (Pequeno Teorema de Fermat)**.** *Dado um número primo p e a um natural não divisível por p, tem-se que p divide o número $a^{p-1} - 1$.*

Demonstração. Considerando os $p-1$ múltiplos de a, tem-se: $a, 2a, 3a, \ldots, (p-1)a$. Sabe-se que nenhum desses números é congruente módulo p com outros deles e nem congruentes a zero módulo p. Tomando $ra \equiv as \, mod \, p$, com $1 \leq r \leq s \leq (p-1)$, como $p-a$, temos $r \equiv s \, mod \, p$, o que é uma contradição. Assim, tais números formam um sistema completo de resíduos. Logo, eles são congruentes a $1, 2, 3, \ldots, (p-1)$ em alguma ordem. Dessa forma,

$$a \times 1 \times (2a) \times (3a) \ldots (p-1)a \equiv 1 \times 2 \times 3 \times \ldots \times (p-1) \, mod \, p,$$

onde

$$a^{p-1}(1 \times 2 \times 3 \times \ldots \times (p-1)) \equiv 1 \times 2 \times 3 \times \ldots \times (p-1) \, mod \, p.$$

Observa-se que o $mdc(p, 1 \times 2 \times 3 \times \ldots \times (p-1)) = 1$, segue-se que $a^{p-1} \equiv 1 \, mod \, p$.

De acordo com Hefez (2016, p. 137), o Pequeno Teorema de Fermat, "fornece-nos um teste de não primalidade. De fato, dado $m \in \mathbb{N}$, com $m > 1$, se existir algum $a \in \mathbb{N}$, com $mdc(a, m) = 1$, tal que m não divide $a^{m-1} - 1$, então m não é primo".

2.4 Teorema de Euller

Será apresentado neste momento o *Teorema de Euller*. Este teorema e suas implicações são de fundamental importância para o estudo dos sistema de criptografia apresentados no próximo capítulo. Para chegar ao Teorema de Euller, deve-se apresentar primeiramente que um *sistema reduzido de resíduos* módulo m é um conjunto de números inteiros $r_1, r_2, \ldots, r_s$ tais que:

(i) $mdc(r_i, m) = 1$, para todo $i = 1, 2, 3, \ldots, s$;

(ii) $r_i \not\equiv r_j \, mod \, m$, se $i \neq j$;

(iii) Para cada $n \in \mathbb{Z}$ tal que $mdc(n, m) = 1$, existe i tal que $n \equiv r_i \, mod \, m$.

Para obter um sistema reduzido de resíduos $r_1, r_2, \ldots, r_s$ módulo m, pode-se utilizar um sistema completo de resíduos $a_1, a_2, \ldots, a_s$ módulo m, eliminando-se os elementos a_i que não são primos com m. Será designado por $\varphi(m)$ o "número de elementos de um sistema reduzido de

resíduos módulo $m > 1$, que corresponde à quantidade de números naturais entre 0 e $m-1$ que são primos com m."(HEFEZ, 2016, p. 195). Colocando $\varphi(1) = 1$, isso definira um importante função dada por:

$$\varphi : \mathbb{N} \to \mathbb{N} \ \text{(denominada função fi de Euler)}.$$

Pela definição, tem-se que:

$$\varphi(m) \leq m - 1, \ \text{para todo m} \geq 2.$$

Deve-se chamar atenção para o seguinte fato: se $m \geq 2$, então $\varphi(m) = m - 1$ se e somente se, m é um número primo. Realmente, m é primo se, e somente se, $1, 2, \ldots, m-1$ formam um sistema reduzido de resíduos módulo m, ou seja, $\varphi(m) = m - 1$.

Para calcular $\varphi(m)$, segue-se a seguinte proposição:

Proposição 2.4.1. *Sejam $r, s \in \mathbb{N}$ tais que $mdc(r, s) = 1$. Então,*

$$\varphi(r.s) = \varphi(r)\varphi(s).$$

Demonstração. Para $r = 1$ e $s = 1$ o resultado é trivial. Supondo então que $r > 1$ e $s > 1$, observa-se a construção de uma tabela formada pelos números naturais de 1 e $(r \times s)$:

1	2	$\ldots$	k	$\ldots$	s
$s+1$	$s+2$	$\ldots$	$s+k$	$\ldots$	$2s$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
$(r-1)s+1$	$(r-1)s+2$	$\ldots$	$(r-1)r+k$	$\ldots$	$r \times s$

Como se tem $mdc(t, r \times s) = 1$ se, e somente se, $mdc(t, r) = mdc(t, s) = 1$, para calcular $\varphi(r.s)$, deve-se destacar aqueles números da tabela que são primos com r e com s. Hefez (2016, p. 199) aponta que, se o primeiro número de uma coluna não for primo com s, então os demais elementos destas colunas também não serão. Conclui-se que os números primos com s estão nas colunas restantes em número de $\varphi(s)$, com elementos que são primos com s. Deve-se, agora, saber quais elementos das colunas restantes são primos com r. Como $mdc(r, s) = 1$, então, a sequência $k, s+k, \ldots, (r-1)s+k$ forma um sistema completo de resíduos módulo m e, assim, $\varphi(r)$ desses números são primos com r. Dessa forma, o número de elementos primos ao mesmo tempo com $4r$ e s é $\varphi(r).\varphi(s)$.

Por exemplo, se se deseja descobrir quais os números menores ou iguais a 65 que são primos com este, ou seja, quanto é $\varphi(65)$, deve-se fazer o seguinte:

$$\varphi(5).\varphi(13) = 4.12 = 48.$$

Com as informações acima, chega-se ao Teorema de Euler, descrito abaixo com a demonstração trazida por Hefez (2016, p. 197)

Teorema 2.4.1 (Teorema de Euler)**.** *Sejam $m, a \in \mathbb{Z}$ com $m > 1$ e $mdc(a, m) = 1$. Então, $a^{\varphi(m)} \equiv 1 \mod m$.*

Demonstração. Seja $r_1, r_2, \ldots, r_{\varphi(m)}$ um sistema reduzido de resíduos módulo m. Logo, $ar_1, ar_2, \ldots, ar_{\varphi(m)}$ formam um sistema reduzido de resíduos módulo m e, portanto,

$$ar_1, ar_2, \ldots, ar_{\varphi(m)} \equiv r_1, r_2, \ldots, r_{\varphi(m)} \; mod \, m.$$

Com isso,

$$a^{\varphi(m)} r_1 \times r_2 \times \ldots \times r_{\varphi(m)} = ar_1 \times ar_2 \times \ldots \times ar_{\varphi(m)} \equiv r_1 \times r_2 \times \ldots \times r_{\varphi(m)} \; mod \, m.$$

Como $(r_1 \times r_2 \times \ldots \times r_{\varphi(m)}, m) = 1$, segue-se que

$$a^{\varphi(m)} \equiv 1 \; mod \, m.$$

O Pequeno Teorema de Fermat é um caso particular do Teorema de Euler, uma vez que enquanto o Teorema de Fermat trabalha com congruências envolvendo módulo primo, o Teorema de Euler lida com módulos em números compostos. Basta notar que, se p é primo, então, $\varphi(p) = p - 1$.

Teorema 2.4.2 (Pequeno Teorema de Fermat)**.** *Sejam $a \in \mathbb{Z}$ e p um número primo tais que $mdc(a, p) = 1$. Tem-se que*

$$a^{p-1} \equiv 1 \; mod \, p.$$

Utilizando os teoremas e suas consequências, qual o resto da divisão de 5^{103} por 33? Note que

$$\varphi(33) = \varphi(3 \times 11) = \varphi(3) \times \varphi(11) = 2 \times 10 = 20.$$

Pelo Teorema de Euler, tem-se que $5^{20} \equiv 1 \; mod \; 33$, logo,

$$5^{103} = 5^{20 \times 5 + 3} \equiv 26 \; mod \; 33.$$

Portanto, 26 é o resto da divisão de 5^{103} por 33.

CAPÍTULO 3

Criptografia RSA

O método mais conhecido de criptografia de chave pública é o RSA. Ao contrário da Diffie-Hellman, ela teve bastante sucesso porque era simples e muito segura. Apesar dos problemas informáticos associados à codificação e decodificação de mensagens, a RSA "foi criada em um momento conveniente quando os computadores estavam recebendo um investimento tecnológico"(KAHN, 1996, p. 251).

> Este código foi inventado em 1978 por R. L. Rivest, A. Shamir e L. Adleman, que na época trabalhavam no Massachussets Institute of Technology (M.I.T.). As letras RSA correspondem às iniciais dos inventores do código. Há vários outros códigos de chave pública, mas a RSA é, atualmente, o mais usado em aplicações comerciais. Este é o método utilizado, por exemplo, no Netscape, o mais popular dos softwares de navegação da Internet (COUTINHO, 2014, p. 3).

De fato, desde seu início, a criptografia RSA é encontrada em várias implementações de computadores, tais como segurança de bancos de dados bancários, tráfego de informações confidenciais em redes abertas e protocolos de comunicação na Internet, principalmente no comércio eletrônico.

> O sucesso da Era da Informação depende da capacidade de proteger essas informações enquanto elas fluem ao redor do mundo e isto depende do poder da

> criptografia. A cifragem pode ser vista como a fonte das chaves e trancas da Era da Informação. Durante dois mil anos ela foi importante apenas para o governo e os militares, mas hoje ela também tem um papel a desempenhar na facilidade dos negócios e, no futuro, pessoas comuns dependerão da criptografia para proteger sua privacidade. (SINGH, 1999 p. 318)

O princípio baseia-se na relativa facilidade em encontrar números primos grandes e ao mesmo tempo na enorme dificuldade prática em fatorar o produto de dois desses números, além do uso de propriedade relativamente elementares da Teoria dos Números, como a variante do Teorema de Euler. (HEFEZ, 2016, p. 274)

De uma maneira muito introdutória e conversadora, a codificação RSA funciona da seguinte maneira: dois números primos p_1 e p_2, de preferência grandes, são selecionados, e, a seguir, é feito o cálculo $n = p_1 \times p_2$. Então um número menor que n e que seja primo é escolhido, chamado aqui de e. Deve-se divulgar estes números n e e, uma vez que eles são necessários para a cifragem da mensagem, sendo consideradas chaves públicas. A mensagem a ser enviada é convertida em um número M e cifrada para produzir o texto C, de acordo com a fórmula $C = M \mod N$. Para decifrar, basta encontrar o número f, que é o módulo inverso a $\varphi(n)$.

> Para poder implementar o RSA precisamos de dois parâmetros básicos: dois números primos que vamos chamar de p e q. Para codificar uma mensagem usando RSA é suficiente conhecer o produto dos dois primos, que vamos chamar de n. Para decodificar uma mensagem precisamos conhecer os primos p e q. A chave de codificação do RSA é portanto constituída essencialmente pelo número $n = pq$. Cada usuário do método tem sua própria chave de codificação. Esta chave é tornada pública: todos ficam sabendo que, para mandar uma mensagem para o banco Acme, deve ser usada a chave n. Por isso, n também é conhecida como "chave pública". Já a chave de decodificação é constituída pelos primos p e q. Cada usuário tem que manter sua chave de decodificação secreta ou a segurança do método estará comprometida. (SINGH, 2007, p. 4)

Naturalmente, para o leitor ilegítimo, a explicação acima é bastante complicada, e os objetos exibidos nela, bem como os resíduos de fissão, podem parecer matemáticas avançadas e tediosas. Os matemáticos tinham que realizar este cálculo de uma maneira simples e demonstrar a segurança desta criptografia.

3.1 Tipos de Chaves

A RSA consiste em uma criptografia de chave pública, portanto, ela deve ter duas chaves: pública e privada. Esta chave pública pode ser revelada a qualquer pessoa sem comprometer a segurança do código. Com ele, os remetentes podem criptografar suas mensagens para serem lidas pelo destinatário com sua chave privada.

Para tornar as chaves públicas e privadas, você deve selecionar o número primo. Normalmente apenas dois números primos ($u = 2$) são escolhidos para as chaves, e quando se faz esta escolha chamamos isso de criptografia RSA dupla. "Quando mais de dois primos ($u > 2$) são selecionados, chamamos de codificação *multiprime RSA*"(KAHN, 1996). Ao importar o número de primos utilizados, deve-se ter em mente que a escolha dos primos é feita pelo destinatário da mensagem, e de preferência em segredo.

Uma vez selecionados os primos, a primeira chave que pode ser feita é a chave pública. Para fins de criptografia RSA, a chave pública consiste em dois números: n, chamado de módulo RSA; e, chamado de display público. Por definição, n é o produto de todos os primos selecionados, ou seja, $n = p_1 \times p_2 \times \ldots \times p_u$. Para o número e, com $e < n$ e $mdc(e, \varphi(n)p_1 \times p_2 \times \ldots \times p_u) = 1$. Em forma simplificada, deve-se:

$$n = p_1 \times p_2 \times \ldots \times p_u,$$

$$(n, e) = \{e; mdc(e, \varphi(n)) = 1\}.$$

O par (n, e) é a chave pública da codificação RSA em questão.

A decodificação requer uma chave privada. Para os primos que compõem a chave pública, a chave privada RSA consiste em dois números: n, novamente o módulo RSA, e f, chamado de display privado RSA. O valor de n é igual à chave pública, onde $n = p_1 \times p_2$. Para o cálculo de f deve-se encontrar o inteiro positivo inverso do número e em $\varphi(n)$. O par (n, d) é a chave privada de criptografia RSA em questão.

3.2 Processo de Codificação

Uma vez recebidas as chaves públicas e privadas, você pode criptografar suas mensagens. O remetente recebe do destinatário a chave pública, que não deve ser enviada em segredo, sem comprometer a segurança do código. O processo de codificação ou codificação é feito em

duas etapas: Criptografia e criptografia preditiva.

> Para acelerar a cifragem e a decifragem, Zimmermann empregou um truque hábil que usa a cifragem assimétrica RSA associada com a velha cifragem simétrica. A cifragem simétrica tradicional pode ser tão segura quanto a cifragem assimétrica e é muito mais rápida de ser feita, mas sofre com a necessidade de exigir a distribuição de uma chave, que terá que ser transportada em segurança do remetente ao destinatário. é aí que a RSA vem em nossa ajuda, porque ela pode ser usada para cifrar a chave simétrica. (SINGH, 1999, 324).

Na pré-codificação, a mensagem a ser enviada deve ser processada de tal forma que possa ser modificada utilizando aplicações em fórmulas matemáticas, o que significa converter a mensagem em uma sequência de números que formam um número completo (COUTINHO, 2014, p. 181). Há várias maneiras de fazer isso, que variam de acordo com as necessidades.

Por exemplo, o próprio documento de padronização da RSA menciona o OS21P (Octec-String-to-lnteger-Primitive), que converte dados de uma forma octal para um número inteiro positivo. Isto porque este documento se refere à criptografia usada em computadores que trabalham com seus dados em forma de bits e bytes e, portanto, em agrupamento de 8 bits (forma octal). No campo dos compactadores, há outra forma de converter caracteres não numéricos em inteiros positivos, então, tem-se uma tabela ASCII (American Standard Code for lnformation lnterchange), que corresponde a cada caractere legível por computador, um inteiro de O a 255. (CAHN, 1996)

Mas não há obrigação de utilizar uma certa conversão de dados. é importante manter o método entre o remetente e o receptor. Neste trabalho, será utilizada uma conversão muito simples, pedagógica, de contagem de cartas sugerida por Cutião (2007), de acordo com a tabela abaixo:

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

Nesta transformação, os espaços entre as palavras são representados pelo número 99. Após converter os dados em inteiros correspondentes, o número longo resultante deve ser dividido em blocos de outros inteiros, com números sempre abaixo de n (módulo RSA) e o maior número possível.

Com uma mensagem pré-codificada, pode-se realizar a codificação. Uma vez codificada a mensagem, é exibida uma sequência de blocos de números que devem ser processados um após o outro. E a mensagem codificada com a chave pública (n, e) representada por $(C_{n,e})$ é uma concatenação de blocos C_n, ou seja:

$$C_{n,e} = C_{n,e_1} C_{n,e_2} \dots C_{n,e_z}.$$

3.3 Processo de Decodificação

O processo de decodificação é semelhante à codificação reversa e é feito em uma única etapa. Com a chave privada (n, f), o destinatário pode decodificar as mensagens usando exposição e álgebra modular. A primeira etapa de decodificação é reorganizar a mensagem criptografada recebida em blocos inteiros. Da mesma forma, uma mensagem numérica deve ser dividida em blocos de números inteiros menores do que n e em blocos maiores. Isto é possível, ou seja, a sequência de blocos deve ser obtida novamente. Alterou a ordem dos blocos de números e a propriedade da chave privada (n, d). Você obtém os blocos originais com o resto da divisão a^f por n, onde a, refere-se a cada bloco codificado.

A seguir, será mostrado como a operação anterior nos devolve aos blocos de origem das mensagens. É interessante notar que os processos de codificação e decodificação são matematicamente opostos, mas é impossível derivar um do outro sem conhecer o primo do qual as chaves são feitas. Após decodificar os blocos, eles devem ser concatenados para obter uma codificação preliminar ao longo de todo o comprimento. Ou até mesmo: $P' = P'_1 P'_2 \dots P'_z$.

Após a concatenação, deve-se realizar um processo de pré-codificação reversa usando a conversão reversa e receber a mensagem original novamente.

Para facilitar a compreensão e mostrar que os processos de codificação e decodificação são matematicamente opostos, serão dados dois exemplos. Um primeiro mais simples, para efeito de exercício da teoria, deve-se codificar/decodificar a mensagem *PRIMO*.

Escolhida a mensagem *PRIMO*, deve-se pré-codificar de acordo com uma tabela. Nessa pré-codificação, será utilizada a seguinte tabela de conversão, com cada letra correspondendo a

um número de dois algarismos e o espaço entre as palavras pelo número 99.

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

Na tabela, foram consideradas apenas letras maiúsculas, e mais o espaço cujo código foi convencionado como 99. Então, a mensagem pré-codificada fica assim:

$$2527182224.$$

Utilizando os parâmetros do sistema RSA, escolhem-se dois números primos distintos, denotados aqui por p e q. Seja $n = pq$. Para este exemplo, tem-se que $p = 7$ e $q = 17$, então $n = 119$. Agora, deve-se quebrar em blocos o número produzido anteriormente, de forma que os blocos sejam menores que n. Não existe uma única forma de quebrar esses números em blocos, mas se devem evitar algumas situações, como iniciar um bloco por 0 (zero). A mensagem codificada, então, ficará com os seguintes blocos de números:

$$25 - 27 - 18 - 22 - 24.$$

Após a etapa de pré-codificação, passa-se a codificação em si. Para codificar a mensagem, necessita-se de n, que é o produto dos primos q e r, e de um inteiro positivo e que seja inversível módulo $\varphi(n)$. Ou seja, $mdc(e, \varphi(n)) = 1$. Como visto no capítulo anterior, deve-se calcular $\varphi(n)$, sabendo que

$$\varphi(n) = \varphi(p) \times \varphi(q) = (p-1)(q-1).$$

Cada bloco b é codificado assim: $C(b) = b^e \; mod \; n$, onde e representa o elemento que compõe a chave pública (e, n). Neste caso, foi considerado $e = 3$ como expoente de cada código de letra, uma vez que 3 é um dos coprimos com $\varphi(n) = \varphi(119)$. Agora devem ser calculados os termos b do bloco numérico.

B1 25: $25^5\ mod\ 119 = 9.765.625\ mod\ 119 = 9$

B2 27: $27^5\ mod\ 119 = 14.348.907\ mod\ 119 = 6$

B3 18: $18^5\ mod\ 119 = 1.889.568\ mod\ 119 = 86$

B4 22: $22^5\ mod\ 119 = 5.153.632\ mod\ 119 = 99$

B5 24: $24^5\ mod\ 119 = 47.962.624\ mod\ 119 = 96.$

Obtém-se, assim, a seguinte sequência de blocos:

$$9\ -\ 6\ -\ 86\ -\ 99\ -\ 96.$$

Na etapa de decodificação, é calculado o parâmetro da chave privada (d, n). Neste caso só cálculos matemáticos envolvem aritmética modular, conceitos de inversibilidade de módulos de números. Para esses valores de p e q, basta aplicar o algoritmo estendido de Euclides a e e $\varphi(n)$. Para o exemplo acima, obtém-se

$$96 = 5 \times 19 + 1, \text{ de onde } 1 = 96 + (-19) \times 7.$$

Logo o inverso de 5 módulo 96 é -19. Porém, se se deseja f positivo para ser usado na potência, de forma que $f = 96 - 19 = 77$, que é o menor inteiro positivo congruente a 96. Dessa forma, para decodificar o primeiro bloco da mensagem codificada $-9-$, calcula-se a forma reduzida de $9^{77}\ mod\ 119$, como segue abaixo:

C1 9: $9^{77}\ mod\ 119$ = $2,99690673E + 73$ = 25

C2 6: $6^{77}\ mod\ 119$ = $8,27268102E + 59$ = 27

C3 86: $86^{77}\ mod\ 119$ = $9,04442050E + 148$ = 18

C4 99: $99^{77}\ mod\ 119 = 4,61221967E + 153 = 22$

C5 96: $96^{77}\ mod\ 119 = 4,31404771E + 52 = 24$

A decodificação, assim, aponta para a mensagem original, de acordo coma tabela apresentada: *PRIMO.*

Para complementar o estudo sobre o sistema RSA, mais um exemplo será dado. Isto exige que dois professores da UFS, indicados aqui por PRO-I (Itabaiana) e por PRO-SC (São Cristóvão) estejam distantes e que precisam compartilhar informações sensíveis de forma segura. Eles irão usar criptografia e esta será a RSA dos dois primos. PRO-I é o remetente do sistema, e PRO-SC é o receptor. O receptor é responsável pela criação das chaves públicas e privadas.

Para usar o método RSA, é preciso inicialmente converter a mensagem em uma sequência de números. Para exemplificar, será feita codificada uma mensagem sem a presença de números, apenas palavras. Dessa forma, a mensagem é constituída pelas palavras e pelos espaços entre elas. Nessa pré-codificação, será utilizada a seguinte tabela de conversão, com cada letra correspondendo a um número de dois algarismos e o espaço entre as palavras pelo número 99.

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

A mensagem a ser codificada por PRO-I é *UNIVERSIDADE FEDERAL DE SERGIPE.* Utilizando os códigos acima, a frase ficará da seguinte forma:

$$30231831142728181310131499151413142710219913149928142716182514.$$

Utilizando os parâmetros do sistema RSA, PRO-SC escolhe dois números primos distintos,

denotados aqui por p e q. Seja $n = pq$. Agora, deve-se quebrar em blocos o número produzido anteriormente, de forma que os blocos sejam menores que n. Para este exemplo, tem-se que $p = 7$ e $q = 17$, então $n = 119$. Não existe uma única forma de quebrar esses números em blocos, mas se devem evitar algumas situações, como iniciar um bloco por 0 (zero). A mensagem codificada, então, ficará com os seguintes blocos de números:

$30 - 23 - 18 - 31 - 14 - 27 - 28 - 18 - 13 - 101 - 31 - 49 - 91 - 51 - 41 - 31 - 42 - 7 - 102 - 19 - 91 - 31 - 49 - 92 - 81 - 42 - 71 - 61 - 82 - 51 - 4.$

Os blocos em que a mensagem foi quebrada não tem nenhuma unidade linguística e, portanto, não pode ser decodificada por contagem de frequência. Diferentemente de outros sistemas criptográficos primitivos que utilizavam a troca de letras e/ou palavras e que fazia a mensagem ter certa lógica devido as frequências de letras no alfabeto e ao agrupamento de determinadas letras, como o Q e o U na língua portuguesa.

Após a etapa de pré-codificação, passa-se a codificação em si. Para codificar a mensagem, necessita-se de n, que é o produto dos primos q e r, e de um inteiro positivo e que seja inversível módulo $\varphi(n)$. Ou seja, $mdc(e, \varphi(n)) = 1$. Como visto no capítulo anterior, deve-se calcular $\varphi(n)$, sabendo que

$$\varphi(n) = \varphi(p) \times \varphi(q) = (p-1)(q-1).$$

O par (n, e) será a *chave de codificação* do sistema RSA. Essa é a chamada *chave pública* e PRO-SC poderá divulgar tal chave. Cada bloco será codificado separadamente, e a mensagem codificada será a sequência de blocos codificados. Seja b cada um dos blocos, este número é um inteiro menor do que n. Para calcular a codificação de cada bloco deve-se conseguir o resto da divisão de b^e por n. Ou seja, a codificação é a forma reduzida de b^e módulo n.

Voltando a mensagem considerada, tem-se que $p = 7$ e $q = 17$, logo $n = 119$ e $\varphi(n) = 96$. Para escolher o número e, neste exemplo vai se escolher 5 que é o menor número primo que não divide 96. Assim, o primeiro bloco - 30 - da mensagem é codificado como o resto da divisão de 30^5 por 119, ou seja, deve-se calcular $30^5 \; mod \; 119$:

$$30^5 = 30^2 \times 30^2 \times 30 \equiv 67 \times 67 \times 30 \equiv 86 \times 30 \equiv -33 \times 30 \equiv -38 \equiv 81 \; mod \; 119.$$

Codificando toda a mensagem, obtém-se a seguinte sequência de blocos:

$81 - 109 - 86 - 12 - 63 - 6 - 112 - 86 - 13 - 33 - 12 - 70 - 7 - 102 - 62 -$

$12 - 77 - 28 - 51 - 66 - 7 - 12 - 70 - 113 - 30 - 77 - 22 - 108 - 80 - 102 - 72.$

Para decodificar um bloco de mensagem codificada, PRO-SC deve ter duas informações: os números n e o inverso de e em $\varphi(n)$, chamado aqui de f. Ou seja, o para (n, f) será a *chave de decodificação*. Essa é a chamada *chave privada*.

Assim, a chave privada é construída. Uma vez que as chaves públicas e privadas tenham sido construídas, PRO-SC não precisará manter seus primos encontrados e eles poderão ser descartados. Na verdade, ele não deveria sequer ter que mantê-los comprometidos pela segurança da criptografia se esses primos forem encontrados por um *hacker* que possa facilmente encontrar os valores n, e, f. Rejeitar os primos é necessário para manter a segurança e não prejudica o uso do código.

Ao fazer as chaves, o receptor PRO-SC revela a chave pública a PRO-I por meios não seguros. Quem intercepta a chave pública é incapaz, de forma prática e rápida, de decifrar a informação, podendo, no melhor dos casos, enviar a informação também para PRO-SC. PRO-I, que recebe a chave pública, é capaz de enviar mensagens.

Para calcular a decodificação, precisa-se verificar o resto da divisão de a^f por n, sendo a cada um dos blocos a serem decodificados. Assim, estar-se calculando a forma reduzida de a^f módulo n.

Sendo n e e conhecidos, fica fácil o cálculo de f. Basta aplicar o algoritmo estendido de Euclides a e e $\varphi(n)$. Para o exemplo acima, obtém-se

$$96 = 5 \times 19 + 1, \text{de onde } 1 = 96 + (-19) \times 7.$$

Logo o inverso de 5 módulo 96 é -19, porém se deseja f positivo para ser usado na potência, de forma que $f = 96 - 19 = 77$, que é o menor inteiro positivo congruente a 96. Dessa forma, para decodificar o primeiro bloco da mensagem codificada - 81 -, calcula-se a forma reduzida de $81^{77}\ mod 119$. Neste caso, o cálculo deverá ser feito através de programas de computadores adequados, uma vez que os valores dificilmente são encontrados sem recursos tecnológicos.

3.4 Segurança

Em teoria é possível decifrar o código na criptografia RSA, e qualquer pessoa com um conhecimento mínimo de matemática pode fazer isso. A chave pública da RSA consiste nos números n e e enquanto a chave privada, além de n, também tem um display privado f. Deve-

se lembrar de que estes três números são compostos de primos escolhidos no início, enquanto o número f depende do número e selecionado. Tanto n como d podem ser identificados pelos primos escolhidos e pela exibição pública.

> O único problema para a segurança da criptografia de chave pública RSA é que, em alguma época no futuro, alguém possa encontrar um modo rápido de fatorar. É concebível que daqui a uma década, ou mesmo amanhã, alguém possa descobrir um método para a fatoração rápida e aí a RSA se tornará inútil. Contudo, por dois mil anos os matemáticos têm tentado e fracassado em encontrar um atalho, e por enquanto, a fatoração continua sendo um cálculo muito trabalhoso. A maioria dos matemáticos acredita que a fatoração é uma tarefa inerentemente difícil e que existe alguma lei matemática que proíbe a existência de qualquer atalho. Vamos presumir que eles estejam certos: deste modo, a RSA estará segura durante o futuro previsível. (SINGH, 2005, p. 303)

Acontece que o número n é o produto de todos os primos utilizados para criar as chaves públicas e privadas, e este número é revelado na chave pública (n, e). Usando o processo de fatorar em números primos, n pode ser identificado e todos os primos encontrados.

Observe que a segurança do sistema de criptografia na RSA depende dos primos geradores selecionados (COUTINHO, 2014, p. 182). Se selecionados corretamente, eles podem tornar o trabalho do criptanalista mais difícil. Por exemplo, para evitar que o número n seja levado em conta de uma maneira possível, podemos escolher geradores de números primos. Há implementações comerciais baseadas na RSA, com chaves públicas, cujo n módulo da RSA tem 2467 números (COUTINHO, 2007). Para considerar um número com tal tamanho, devem-se esperar 1012^{33} operações computacionais. Se o computador atual (3,0 GHz Quad Core) realiza algo como 1012 operações por segundo, levaria até 1012^{21} segundos para o fator n. Este tempo corresponde a, pelo menos, $3,2 \times 10$ anos.

> Desde a implementação do sistema RSA, tem-se procurado informações mais precisas a respeito dos métodos de fatoração, ainda sem sucesso, uma vez que o sistema ainda não foi quebrado. Porém, ainda se sabe se quebrar o código RSA significa necessariamente descobri a fatoração do número n ou se haveria outra forma mais simples de se fazer. De toda forma, até então, não se teve sucesso em quebrar o código utilizando outra técnica que não seja a fatoração de n. (SINGH, 1999, p. 317)

Coutinho (2014, p. 187) aponta que não basta ter números primos grandes, há outro ponto fundamental:

Um ponto importante refere-se à escolha dos primos p e q. É claro que se forem pequenos, o sistema será fácil de quebrar. Mas não basta escolhê-los grandes. De fato, se p e q são grandes, mas $|p - q|$ é pequeno, então é fácil fatorar $n = pq$ usando o algoritmo de Fermat. (COUTINHO, 2014, p. 187)

O módulo RSA n não deve ter mais de 2467 dígitos para garantir a segurança. Para testar a segurança do próprio código, o Laboratório RSA executa constantemente uma tarefa chamada *RSA Factoring Challenge*, recompensando aqueles que alcançam o fator chave público RSA disponível em seu website com uma pequena quantia em dólares. Um dos problemas em 1999 foi a inclusão de uma chave pública RSA de 512 bits (155 dígitos). A chave foi a seguinte (RSA, 2000):

$$1094173864157052742180970732204035761200373294544920599091384213147634998428893478471799725789126733249762575289978183379707653724402714674353159335433389 7.$$

Esta edição foi revista em agosto de 1999 por um grupo de usuários liderado por A. Lenstra e H. te Riehl usando 300 computadores em série, o que levou sete meses de trabalho (RSA, 2000). Mais tarde, em 2005, outro desafio ofereceu um prêmio de fatorar de 641 bits (193 dígitos), que exigiu cinco meses de trabalho utilizando 80 computadores (COUTINHO, 2007).

Há algoritmos que podem fatorar um módulo RSA não utilizado, bem como a necessidade de levar em consideração os primos. Quando os primos p_1 e p_2 são tais que é difícil considerar o módulo, RSA (2000) os chama de *primos fortes*.

Dependendo da finalidade da codificação, a RSA (2000) oferece um tamanho mínimo para um módulo RSA. Para fins comerciais, recomenda-se usar pelo menos 1024 bits (aproximadamente 308 dígitos) e 2048 bits (mais de 615 dígitos) de chaves para dados sensíveis, tais como transações financeiras. Para uso pessoal, uma chave de 768 bits (cerca de 231 dígitos) é suficiente.

Além do tamanho da chave, é importante usá-la por menos de dois (02) anos, e sugere-se duplicar o número de dígitos de uma nova chave. Ao dobrar o número de dígitos, o trabalho de codificação e decodificação das mensagens pode aumentar quatro vezes, mas o fator de trabalho para a inclusão desta nova chave é oito vezes maior (RSA, 2000).

Em épocas diferentes, durante os últimos dois mil anos, os criptógrafos já acre-

> ditaram que a cifra monoalfabética, a cifra polialfabética e as máquinas de cifragem com a Enigma eram inquebráveis. E, em cada um desses casos, mostrou-se que os criptógrafos estavam errados porque suas afirmações eram baseadas meramente no fato de que a complexidade das cifras tinha superado a engenhosidade e a tecnologia dos criptoanalistas naquele ponto da história. Hoje, podemos ver que os criptoanalistas iriam, inevitavelmente, descobrir um meio de quebrar cada cifra, ou desenvolveriam uma tecnologia para fazer isso para eles. (SINGH, 1999, p. 378)

3.5 Assinaturas Digitais

Assim como na assinatura escrita, a assinatura digital visa certificar a autenticidade e autoria da mensagem enviada. Caso a informação não seja sigilosa, mas sua autoria tem de ser certificada, fica mais simples codificar a assinatura de forma única do que codificar toda a mensagem. Atualmente, há alguns sistemas de assinaturas disponíveis, porém o mais utilizado é o sistema RSA. A assinatura digital já foi imaginada pelo trio que criou o sistema RSA, logo após sua criação.

A assinatura digital garante que a mensagem não foi modificada, uma vez que se não estivesse íntegra a conferência entre os itens calculados pelo destinatário e o obtido na decodificação seria diferente. Além disso, a assinatura digital também garante que o remetente é realmente quem diz ser, porque somente ele possui a chave privada e, assim, somente ele poderia produzir tal assinatura digital. Ainda, o sistema permite que o remetente não possa negar que enviou tal mensagem, já que somente ele é detentor da chave privada.

Resumidamente, uma assinatura digital X de uma mensagem Y é criado a partir da chave privada de algum usuário e pode ser confirmada pela chave pública. Somente determinado usuário poderia produzir sua assinatura digital válida, sendo inviável de alguém o fazer baseando-se em mensagens anteriores.

A função *Message Digest*, ou simplesmente MD - resumo da mensagem, é a função utilizada para assinar a mensagem a ser enviada, onde o processamento do documento produz um punhado de dados, chamado de *hash*. Nitidamente, criptografar o *hash* leva bem menos tempo que codificar toda a mensagem, uma vez que o MD constituem blocos pequenos e com tamanho fixo. Assim, o processo se torna mais eficiente.

A assinatura digital pode ser verificada matematicamente, de forma que o documento

e sua assinatura digital podem ser facilmente confirmados em sua integridade e originalidade. O sistema funciona da seguinte forma: executa-se a função MD, produzindo, assim, um *hash* para aquele documento e, a seguir, pode-se decifrar a assinatura digital com a chave pública do remetente. Ao decifrar a assinatura, deve-se produzir o mesmo *hash*, caso isto não ocorra o documento e/ou a assinatura foram alterados.

3.6 Considerações Finais

Hodiernamente, as comunicações entre as pessoas vêm sendo exercidas de modo progressivo através de redes sociais, bem como as transações bancárias estão migrando predominantemente para o meio virtual. Diante dessa demanda também há uma grande necessidade de garantir sigilo na troca de informações, bem como segurança nas transações bancárias. Caso não haja total segurança, as pessoas e as instituições teriam todos seus dados pessoais e patrimoniais correndo riscos.

Esse problema não surgiu junto à internet. O segredo na troca de mensagens foi decisivo para a disputa entre nações ou mesmo dentro de uma nação, sendo responsável por diversos ataques e contra-ataques que mudaram o rumo de disputas e de guerras. A troca de informações sigilosas foi indispensável para que os líderes pudessem espalhar suas estratégias de combate somente para subordinados e aliados.

Esse sigilo era garantido através de uma codificação própria estabelecida anteriormente, o que garantia que, mesmo interceptada, não seria conhecido o teor da mensagem por pessoas alheias ao processo de codificação. Obviamente, à medida que mais mensagens eram interceptadas, maiores eram as chances de se descobrirem os segredos do processos de codificação. Surgiam, dessa forma, pessoas específicas para trabalhar na tentativa de decodificar as mensagens codificadas.

Esses decodificadores buscavam padrões no processo de decomposição. Os primeiros códigos basicamente trocavam a posição das letras no alfabeto antes de codificar a mensagem. Com o passar o tempo, e com as tentativas certeiras de decodificar as mensagens, os sistemas foram se aperfeiçoando, porém ainda poderiam se perceber alguns padrões presentes em determinadas línguas.

Tomando como exemplo a língua portuguesa, a presença da letra Q é seguida obrigatoriamente da letra U, o que torna sua decodificação mais facilitada. Esse sistema monoalfabético

foi perdendo a eficiência ao longo da história, sendo substituído pelo sistema polialfabético, que, como o próprio nome indica, utiliza mais de um alfabeto no processo de codificação de mensagens.

Além dos métodos de substituição de letras, foram utilizados sistemas em que haviam deslocamento das letras do alfabeto de sua posição original ou ainda sistemas em que cada letra tinha a função de uma palavra ou frase. O deslocamento das letras de sua posição original era chamado de método de transposição. Esse sistema tinha uma vantagem em relação ao sistema de substituição de letras, uma vez que as palavras de um idioma possuem agrupamentos mais comuns de letras, facilitando a descoberta por quais letras foram substituídas as originais.

Todos esses métodos têm em comum o fato de haver uma mesma chave para codificar e para decodificar as mensagens, considerado um sistema simétrico. Ou seja, a pessoa que recebe a mensagem codificada precisa conhecer o processo em que foi cifrada a mensagem para poder entendê-la posteriormente, o que exige uma troca de informação prévia entre o emissor e o receptor. Então, antes de enviar um segredo, os envolvidos já teriam de compartilhar um outro segredo anteriormente: o sistema de codificação.

Essa forma, única até algumas décadas atrás, dificultava a troca de mensagens importantes e também era maior o risco de fracasso, caso o primeiro segredo - o sistema de codificação - fosse logo interceptado. Surgiam, então, ideias de um código onde a chave para decodificar fosse diferente da chave para codificação. A ideia surgia, mas seu desenvolvimento não era algo simples.

Foi a matemática quem impulsionou e revolucionou a forma de trocar mensagens de forma sigilosa. Aliada ao surgimento dos computadores, a forma de manter informações e dados de pessoas e empresas passou a ser algo real, aderindo à forma onde a chave para decodificação não seria a mesma para codificação, chamado de sistema assimétrico.

Foram Ronald Rivest, Adi Shamir e Leonard Adleman do Laboratório de Ciência de Informação do Massachusetts Institute of Technolgy (MIT), que possibilitaram, em 1978, a implementação do sistema criptográfico com sistemas assimétricos. O sistema ficou conhecido como RSA, pelo uso das iniciais dos criadores. Importante destacar que os computadores utilizam a linguagem binária e é através do código padrão americano para intercâmbio de informações, conhecido pela sigla ASCII, que a mensagem é transformada para o sistema adotado nos computadores. Então, o código ASCII não á necessariamente um sistema de cifragem, mas somente uma forma de traduzir a mensagem para a linguagem dos computadores.

O RSA consiste em uma criptografia de chave pública, portanto, ela deve ter duas chaves: pública e privada. Esta chave pública pode ser revelada a qualquer pessoa sem comprometer a segurança do código. Com ele, os remetentes podem criptografar suas mensagens para serem lidas pelo destinatário com sua chave privada. Para as chaves públicas e privadas, devem selecionar números primos. Os números primos sempre despertaram grande interesse aos estudiosos e diversos questionamentos a respeito ainda são um mistério. O Teorema Fundamental da Aritmética aponta que todo número natural maior do que 1 ou é primo ou se escreve de modo único como um produto de números primos. Euclides, no Livro IX dos *Elementos*, já demonstrava que os números primos são infinitos, fazendo a demonstração através da redução ao absurdo. Essa infinidade dos números primos garante ao sistema RSA de criptografia um campo amplo de possibilidades, uma vez que se vale da dificuldade de fatorar um número que é produto de dois números primos bem grandes.

Além dos números primos, o sistema RSA vale-se da congruência entre números. Dois números a e b são congruentes módulo m quando o resto da divisão de a e b por m resultam no mesmo número, sendo representado por $a \equiv b \; mod \; m$. Além disso, o Teorema de Euler é necessário no processo de codificação e decodificação das mensagens criptografadas por RSA. Neste teorema, sejam $m, \; a \in \mathbb{Z}$ com $m > 1$ e $mdc(a, \; m) = 1$. Então,

$$a^{\varphi(m)} \equiv 1 \mod m.$$

Feitas estas considerações matemáticas, e uma vez selecionados os primos para uso no sistema RSA, a primeira chave que pode ser feita é a chave pública. Para fins de criptografia RSA, a chave pública consiste em dois números: n, chamado de módulo RSA; e, chamado de display público. Por definição, n é o produto dos primos selecionados, ou seja, $n = p_1 \times p_2$. Para o número e, com $e < n$ e $mdc(e, \; varphi(n) = p_1 \times p_2) = 1$.

Mais precisamente, o método de criptografia RSA funciona do seguinte modo: I) escolhem-se dois primos p e q, distintos entre si; II) define-se $N = p \times q$; III) deve-se escolher um número e, que faz parte da chave pública, de forma que o máximo divisor comum (mdc) entre ele e $\varphi(N)$ seja 1; IV) de acordo com uma tabela pré-formulada e de domínio público é feita a transformação de todos os caracteres da mensagem em números, obtendo-se a mensagem numérica M, após conversão em dígitos binários ASCII, em um único bloco que será dividido em blocos b, de forma que: $b < N$. Isso garante que, ao utilizar congruência, obtenha-se um único resultado na decodificação; V) de posse da chave pública $(e; N)$, criptografam-se os blocos b de

acordo com a congruência: $C = M^e \ mod \ N$, onde M é a mensagem numérica ao qual a mensagem original foi convertida ao ASCII, e $C(b)$ é a mensagem criptografada; VI) de posse da chave privada $(d; N)$, descriptografa-se, e D é a mensagem descriptografada; VII) cada bloco D deve ser colocado em sequência e de acordo com a mesma tabela usada no item IV os números devem ser convertidos em caracteres.

O sistema RSA é assimétrico porque garante uma função de via única, que somente poderia ser revertida com as informações privilegiadas dos números primos p e q utilizados. Como um dos números da chave pública é $N = p \times q$, a segurança do sistema é baseada na dificuldade de se fatorar o número N. Assim, é essencial que os primos escolhidos sejam muito grandes. Claro que isso não é feito de forma braçal, uma vez que os computadores se encarregam de encontrar cada vez maiores números primos para essa função.

Parece ser o sistema RSA extremamente seguro devido a dificuldade de reversão apresentada anteriormente. Porém, viu-se que em todos os sistemas utilizados durante a história, sempre se tinham a impressão de inviolabilidade destes. Fica a dúvida se em algum momento o sistema RSA poderá ser quebrado. Imagina-se que, se isso um dia vir a acontecer, poderá ser de duas formas: encontrar alguma forma de fatorar os números utilizados que até então não se conheça; ou perceber outro caminho para decodificar sem a necessidade de fatorar o número utilizado. Por enquanto, o sistema RSA é seguro e viável para estabelecer a segurança das transações bancárias, o segredo das grandes instituições e inclusive das mensagens particulares trocadas nos aplicativos das redes sociais.

Referências

[1] COUTINHO S. C., Severino Collier. *Números Inteiros e Criptografia RSA*. Rio de Janeiro: IMPA, 2014.

[2] COUTINHO S. C. *Programa de Iniciação Científica da OBMEP - Criptografia*. Rio de Janeiro. OBMEP, 2008.

[3] HEFEZ, A.. *Aritmética*. Rio de Janeiro: SBM, 2016. (Coleção PROFMAT).

[4] HEFEZ, A..*Exercícios Resolvidos de Aritmética*. Rio de Janeiro SBM, 2016. (Coleção PROFMAT).

[5] KAHN, D. *The codebreakers: The story of Secret Writing*. New York: Scribner, 1996.

[6] RSA Laboraroties. *Frequently Asked Questions About Today's Cryptography*. RSA Labs, 2000. Disponível em http://www.rsa.com/rsalabs/faq/files/rsalabsfaq41.pdf. Acesso em 21 de agosto de 2020.

[7] RSA Laboraroties. *RSA Cryptography Standard. New York*. Disponível em ftp://ftp.rsasecurity.com/pub/pkcs/pkcs-1/pkcs-1v2-1.pdf. Acesso em 21 de agosto de 2020.

[8] SINGH, Simon. *O Livro dos Códigos*: a ciência do sigilo - do antigo Egito à criptografia quântica. Trad. Jorge Calife. 6ª ed. Rio de Janeiro: Record, 2007.