Trabalho de Introdução à Criptografia: Criptossistema de Rabin

Alunos: Antônio Isaac (2019006361) e João Antonio Pedrosa (2019006752)

Introdução

O algoritmo do *Criptossistema de Rabin* foi oficialmente publicado em 1979 por Michael O. Rabin (conhecido também pela publicação do Teste de Primalidade de Miller-Rabin) e usa criptografia com chave assimétrica para a comunicação entre dois agentes e para criptografar a mensagem compartilhada entre eles.

Esse criptossistema tem segurança relacionada com a dificuldade de fatorização de um inteiro n grande, de forma parecida com o algoritmo *RSA*.

Porém, diferentemente do algoritmo RSA, o Criptossitema de Rabin possui uma prova matemática formal que diz que ele é computacionalmente seguro contra um *Chosen-Plaintext Attack*, um modelo de ataque que assume que o atacante pode obter as cifras correspondentes à qualquer mensagem.

A desvantagem desse criptossistema, contudo, se encontra no fato de que ele produz o mesmo *output* para 4 *inputs* distintos, ou seja, ao descriptografar uma cifra que fora criptografada com o criptossistema de Rabin, você receberá 4 mensagens distintas, cada uma com igual possibilidade de ser a mensagem intencionada.

Portanto, há uma complexidade extra nesse criptossistema necessária no processo de descriptografia para identificar qual dos 4 possíveis *inputs* foi o que gerou a cifra recebida. Obviamente há uma facilidade maior no caso em que a mensagem representa um texto, pois, após receber a cifra, muito provavelmente é uma questão trivial descobrir qual a mensagem, dentre as 4 possíveis, é a intencionada. A questão fica mais complicada quando esse criptossistema é utilizado para enviar apenas números.

O Método

Contexto

Imagine duas pessoas, Bob e Alice, e que Bob deseja mandar uma mensagem para Alice, de maneira privada. Para tal, Bob gostaria de definir uma chave pública e uma chave privada, de modo que a chave pública possa ser usada por todos enquanto a chave privada só possa ser usada por Alice, que está recebendo a mensagem.

Encontrando as Chaves

Bob começaria encontrando uma chave pública e uma privada pelo seguinte processo:

1. Escolher 2 primos grandes p e q tais que $p \equiv q \equiv 3 \mod 4$
2. Calcular n tal que $n = pq$

Os números calculados serão as chaves:

- Chave Pública: Um inteiro $n = pq$
- Chave Privada: Par ordenado (p, q).

Encriptação

Lembrando que Bob tem acesso apenas à chave pública n, para encriptar uma mensagem M, ele seguiria dois passos:

1. Transformar M em um número qualquer tal que $m < n$
2. Encontrar c tal que $c \equiv m^2 \mod n$

O c encontrado é a mensagem criptografada.

Decriptação

Para decriptar a mensagem, Alice tem acesso à c e (p, q) e irá encontrar o valor original m. A decriptação segue três passos:

1. Calcular dois valores m_p e m_q tais que:

$$m_p \equiv c^{\frac{p+1}{4}} \mod p$$
$$m_q \equiv c^{\frac{q+1}{4}} \mod q$$

   Vale ressaltar que $\frac{p+1}{4}$ e $\frac{q+1}{4}$ são inteiros pois $p \equiv q \equiv 3 \mod 4$.

2. Encontrar valores de x_p e x_q que satisfaçam a seguinte equação diofantina:

$$p \cdot x_p + q \cdot x_q = 1$$

3. Encontrar quatro possiveis mensagens originais resolvendo as seguintes equações:

$$m_1 \equiv x_p \cdot p \cdot m_p + x_q \cdot q \cdot m_q \mod n$$
$$m_2 \equiv n - m_1$$
$$m_3 \equiv x_p \cdot p \cdot m_p - x_q \cdot q \cdot m_q \mod n$$
$$m_4 \equiv n - m_3$$

Informação Adicional

Vale ressaltar que apenas uma das 4 possíveis mensagens encontradas é a mensagem original e é impossível para Alice determinar qual é a correta sem alguma informação adicional. No caso do valor original representar uma mensagem de texto, é fácil adivinhar qual dos resultados é o verdadeiro, afinal, muito provavelmente apenas um deles fará sentido em linguagem natural. Entretanto, nem sempre a adivinhação é possível, e.g. quando o valor original é a representação de um valor númerico. Nesses casos, Bob deve enviar, juntamente com o texto criptografado, algo que permita à Alice definir qual é a mensagem correta.

Técnicas Computacionais

Para a escolha de chaves e a encriptação, basta utilizar um algoritmo de geração de primos pseudo aleatórios e realizar o cálculo de um quadrado e um módulo. Para a decriptação, precisamos fazer uso de algumas técnicas computacionais conhecidas.

Primeira Parte - Exponenciação Rápida

Para a primeira parte da decriptação, basta a realização de uma exponenciação modular, que pode ser resolvida por meio de um algoritmo de exponenciação rápida em tempo logarítmico no tamanho do expoente.

Segunda Parte - Algoritmo Extendido de Euclides

Para a segunda parte da decriptação, precisamos resolver uma equação diofantina. Para isso, faremos uso do algoritmo extendido Euclides. O algoritmo extendido de Euclides encontra, dados dois números a e b, o mínimo divisor comum entre esses dois números e x e y tais que $p \cdot x + q \cdot y = \gcd(p, q)$. Como p e q são primos, $\gcd(p, q) = 1$ e portanto o Algoritmo Extendido de Euclides retorna a resposta da equação diofantina.

Terceira Parte - Teorema Chinês do Resto

Para a parte final da decriptação, queremos calcular m_1, m_2, m_3 e m_4 tendo em mãos p, m_p, x_p, q, m_q e x_q. Basta realizarmos as operações, porém, como esses números podem ser muito grandes, iremos utilizar o Teorema Chinês do Resto para realizar os cálculos de maneira eficiente. O Teorema Chinês do Resto é muito utilizado na computação de cálculos com inteiros muito grandes e diz que, sabendo os restos da divisão euclidiana de um inteiro n por um conjunto de inteiros é possível determinar o resto da divisão de n pelo produto desses inteiros, desde que todos esses inteiros sejam primos entre si.

Exemplo de Uso - Digital Signature Algorithm (DSA)

O Criptossistema de Rabin pode ser utilizado para criar e verificar assinaturas digitais, pois criar uma assinatura digital requer uma chave privada (p, q) e verificar requer a chave pública n.

Assinando

Uma mensagem $m < n$ pode ser assinada com a chave privada (p, q) com o seguinte passo-a-passo:

1. Gere um valor aleatório u;
2. Use uma função *hash* criptográfica H (por exemplo, *SHA256*) para computar $c = H(m|u)$, onde a barra denota uma operação de concatenção entre m e u. c deverá ser um inteiro menor que n;
3. Trate c como valor criptografado pelo Criptossistema de Rabin e tente realizar o processo de descriptografia com a chave privada (p, q). Isso irá produzir os 4 resultados usuais: r_1, r_2, r_3 e r_4;
4. Apesar do que pareceria ser o normal, encriptar qualquer r_i não leva, necessariamente, ao nosso valor c inicial, isso só ocorre no caso em que c é um resíduo quadrático módulo p e q; dizemos que um inteiro k é um resíduo quadrático módulo h se ele for congruente à um quadrado perfeito módulo h; ou seja, se existe um inteiro x tal que $x^2 \equiv k \pmod h$. Para determinar se esse é o caso em que estamos, basta criptografar r_1. Se o resultado dessa operação não é igual à c, repetiremos esse algoritmo a partir do passo **1**. A expectativa do número de vezes que teremos que repetir esse algoritmo até acharmos um c apropriado é 4;
5. Tendo encontrado um r_1 cuja cifra é igual a c, a assinatura será (r_1, u);

Verificando uma assinatura

Uma assinatura (r, u) para uma mensagem m pode ser verificada utilizando a chave pública n da seguinte forma:

1. Compute $c = H(m|u)$;
2. Criptografe r utilizando a chave pública n;
3. A assinatura é válida se, e somente se, o resultado da criptografia de r seja igual a c;

Vantagens e Desvantagens

Vantagens

Eficiência

- Para encriptação, um quadrado módulo n deve ser calculado. Isso é mais eficiente que o RSA, por exemplo, que requer o cálculo de pelo menos um cubo.
- Para a decriptação é aplicado o Teorema Chinês do Resto e duas Exponenciações Modulares. Nesse caso, o algoritmo é tão eficaz quanto o RSA.

Segurança

- É provado que qualquer algoritmo que consiga decriptar uma mensagem encriptada com Rabin pode ser utilizado para fatorar n. Sendo assim, quebrar o Criptossistema de Rabin é pelo menos tão difícil quanto o problema da fatoração de inteiros. Vale lembrar que essa prova não existe para diversos outros sistemas de criptografia que são utilizados comumente, e.g. o RSA.

Desvantagens

Eficiência

- A decodificação produz, juntamente com o resultado correto, três resultados falsos. Esta é a grande desvantagem do Rabin, e o principal impedimento para o seu uso prático.
- Esquemas de desambiguação de mensagens incluiriam custos computacionais adicionais, o que prejudicaria a eficiência do algoritmo.

Segurança

- O Criptossistema de Rabin não possui indistinguibilidade contra um *Chosen Plaintext Attack*, i.e., dado duas mensagens e o *Cipher Text* de uma delas, existe uma chance significativa do adversário conseguir acertar à qual das duas mensagens o *Cipher Text* é relacionado, pois o sistema de criptografia é determinístico.