UNIVERSIDAD MAYOR DE SAN ANDRÉS

FACULTAD DE CIENCIAS PURAS Y NATURALES

CARRERA DE INFORMÁTICA

PERFIL DE TESIS DE GRADO

COMPARACIÓN Y REFINAMIENTO DE ALGORITMOS DE FACTOORIZACIÓN DE NÚMEROS ENTEROS

Tesis de Grado para obtener el Título de Licenciatura en Informática

Mención Ciencias de la Computación

POR: ROLANDO TROCHE VENEGAS

TUTOR METODOLÓGICO: MsC. ROSA FLORES

ASESOR: M.SC. JORGE TERAN POMIER

LA PAZ – BOLIVIA
Junio, 2021

ÍNDICE

1. INTRODUCCIÓN

Factorizar números enteros es importante, hoy más que nunca la factorización de números enteros juega un papel crucial en la vida de todas las personas. Los métodos de encriptación actuales se basan, en gran medida, en la complejidad y el tiempo que toma factorizar números grandes.

También se debe notar la definición de número grande, si se le pregunta a un niño que es un número grande este puede decirnos que es el 100 o hasta el 1000 y si se le preguntara a una persona adulta esta podría decirnos 1 000 000 o 100 000 000, pero esto no es nada si se piensa en los problemas que hoy en día lidian los matemáticos.

Factorizar un número se refiere a encontrar todos los factores primos por los que está compuesto dicho número. El teorema fundamental de la aritmética nos dice que para todo número entero positivo mayor a 1 es un numero primo o bien un único producto de números primos (Euclides, 300 AC).

$$n = \prod_{i=1}^{k} p_i^{\alpha_i}$$

Este es el objetivo que buscamos, encontrar esos factores para números, ahora sí, grandes.

Los algoritmos de encriptación actuales, como RSA, usan números de, por ejemplo, 250 dígitos (829 bits) que fue el último en ser factorizado febrero de 2020. Estos son los números grandes que nos interesan. Los números RSA por ejemplo son números con exactamente dos factores primos, estos números primos, de al menos, la mitad de cantidad de dígitos que el resultado.

Hoy en día tenemos diferentes métodos de factorización como los algoritmos simples de factorización, fracciones continuas, curva elíptica, algoritmos de cribado y hasta nuevos algoritmos basados en programación cuántica.

En este trabajo se estudiará los diferentes métodos de factorización de números y su implementación con diferentes enfoques y refinamientos de implementación que el autor propondrá para ciertos métodos. Con esto se podrá revisar los alcances de los algoritmos de encriptación actuales y a donde se puede llegar en base a estos y realizar una comparación de los mismo, mostrando que algoritmo es el adecuado dada ciertas condiciones.

2. ANTECEDENTES

En la parte de los antecedentes, se debe escribir lo referente a los trabajos previos relacionados con el problema que será abordado en la siguiente sección, es decir; estado del arte, investigaciones realizadasanteriormente que guarda

alguna vinculación con el problema elegido, reportes de investigación en los que se basará el presente trabajo,etc.

3. PLANTEAMIENTO DEL PROBLEMA

Con el avance de las computadoras y el mayor poder de cómputo, toda la seguridad basada en factorización y factores primos corre riesgo de quedar deprecada algún día.

Por esto es necesario tener un compendio de gran parte de los métodos y algoritmos de factorización existentes y sus limitaciones, con el poder de cómputo actual. Por otro lado, hoy en día con la ayuda de los nuevos y mejores procesadores muchos algoritmos pueden ser paralelizados lo que reduciría su tiempo de ejecución.

3.1. PROBLEMA CENTRAL

¿Cuál es el comportamiento de los métodos de factorización de números enteros de acuerdo a las características de los números y los diferentes enfoques en su implementación?

3.2. PROBLEMAS SECUNDARIOS

- Métodos de encriptación insuficientes para la factorización avanzada
- Diferencias entre enfoques clásicos y modernos donde se distinguen los cambios entre ellas
- Complejidad de implementación de técnicas avanzadas de factorización
- Tiempos de ejecución bastante elevados

4. DEFINICIÓN DE OBJETIVOS

4.1. OBJETIVO GENERAL

Determinar los métodos que se comportan mejor en números con diferentes características.

4.2. OBJETIVOS ESPECÍFICOS

- Definir las diferencias entre los enfoques clásicos y modernos.
- Calcular los límites de algoritmos o métodos de encriptación basados en factorización.
- Simplificar el entendimiento de las técnicas de factorización.
- Paralelizar métodos de factorización

5. HIPÓTESIS

Los métodos de cribado son los mejores en general para cualquier tipo de número, mientras otros como el método de Fermat responde mejor cuando la distancia entre los factores primos es pequeña.

5.1. OPERALIZACIÓN DE VARIABLES

Variable independiente.

- Números enteros

Variables dependientes

- Factores primos
- Tiempo de ejecución
- Espacio de memoria

6. JUSTIFICACIÓN

6.1. JUSTIFICACIÓN ECONÓMICA

Con el constante avance de la tecnología y la ciencia en diferentes campos hace que cada día se necesite de mejores equipos, maquinaria, software, hardware entre otros, lo cual implica costos gigantescos, por lo cual hacer una redefinición en las herramientas teóricas resulta ser un gran ahorro y reducción de gastos para las diferentes investigaciones realizadas por universidades, instituciones del estado, comunidades científicas.

Además de una reducción de tiempos, el cual es representado a su vez en una reducción de costos, mejorando así la accesibilidad a nuevos campos de investigación.

6.2. JUSTIFICACIÓN SOCIAL

La sociedad en su conjunto se beneficia indirectamente ya que la presente investigación está enfocada al área teórica pero enteramente ligada a futuros campos de investigación

y aplicación para mejorar el estilo de vida, hambre de conocimientos y experimentación por parte de la población en general.

6.3. JUSTIFICACIÓN CIENTIFICA

La presente investigación brinda al campo científico tecnológico un complejo análisis de diferentes algoritmos de factorización, lo cual conlleva a tener nuevos y/o actualizaciones de los mismos generando avances en diferentes campos no solo del área sino también de otros campos afines, generando propuestas o aplicaciones de la presente investigación.

Por lo tanto, al emprender la investigación de los algoritmos de factorización de números enteros grandes y su evaluación, profundiza el conocimiento, aportando así a futuras investigaciones, ya que se está trabajando en un área en desarrollo. Así también como bases prácticas y teóricas para la aplicación de dichos algoritmos en áreas como el análisis complejo de números primos, representación del conocimiento, teoría de números, criptografía, combinatoria, entre otros

7. ALCANCES Y LÍMITES

7.1. ALCANCES

- Implementación de los métodos
- Recolección de datos
- Demostración de los métodos
- Paralelización de algoritmos
- Refinamiento de algoritmos
- Comparativa de datos

7.2. LÍMITES

- Encontrar factores primos para números RSA grandes
- Romper seguridad actual basada en primos
- Encontrar nuevos números primos
- Algoritmos no estándares

8. APORTES

8.1. PRÁCTICO

La investigación aportará con los resultados de la implementación de algoritmos de factorización. Los resultados podrían ser analizados para su aplicación en campos como criptografía, criptomonedas y aplicaciones que requieran la factorización de números grandes.

8.2. TEÓRICO

La investigación aportará con las posibles modificaciones a algoritmos de factorización, estas modificaciones requerirán el respectivo análisis de complejidad. Con los resultados se aportará con modificaciones de algoritmos aplicados al campo de teoría de números.

9. METODOLOGÍA

Para realizar la presente investigación se basará en el tipo descriptivo, empleando el método lógico partiendo de casos particulares de factorización hasta llegar a generalización de las mismas, dando a su vez solución a varias técnicas de factorización.

10. MARCO TEÓRICO

Factorización por divisiones sucesivas. Si se quiere encontrar los factores primos de n lo que se hace es tener una lista de todos los números primos menores a n, luego se prueba dividir n entre cada primo p_i, si $p_i|n$ entonces se vuelve a hacer el proceso desde ese primo, pero ahora con n/p_i. Si se llega a $\sqrt{n}$ y no se ha encontrado ningún primo que divida a n, entonces se declara a n como primo.

Método de Fermat. Para factorizar un número impar n, Fermat trato de expresar n como una diferencia de dos cuadrados, $x^2 - y^2$ con el par x, y diferente de $\frac{n+1}{2}, \frac{n-1}{2}$. Este par entrega $x + y = n$ y $x - y = 1$. Cualquier otra representación de n como $x^2 - y^2$ entrega una factorización no trivial $n = (x - y)(x + y)$

Algoritmo

Entrada: Un entero compuesto impar *n*

```
x = √n
t = 2x + 1
r = x² − n
while(r no sea una raíz cuadrada) {
      r = r + t
      t = t + 2
}
x = (t-1) / 2
y = √r
```

Salida: Los factores $x - y$ y $x + y$ de n

Algoritmo de Factorización de una línea de Hart. Hart en 2012 invento una variación del Método de Factorización de Fermat, que es mucha más corto, simple de programar. El da un argumento heurístico de que factoriza n en $O(n^{\frac{1}{3+\varepsilon}})$ pasos.

El algoritmo de Hart comienza verificando si n es una raíz. Si n no es una raíz, entonces hace divisiones sucesivas, pero se detiene cuando p alcanza $n^{\frac{1}{3}}$. En caso que n no ha sido factorizado todavía, realiza los siguientes pasos.

Para $i = 1,2,3,$ ... prueba cualquier $|\sqrt{n_i}|^2 \ mod \ n$ si es raiz. Si este número es igual a t^2 entonces, es un factor de n

Algoritmo después de verificar si es raíz y las divisiones sucesivas

Entrada: Un entero positivo n y un límite L

for(i = 1 hasta L){
    s = $|\sqrt{n}|$
    m = s^2 mod n
    if(m es raíz) {break}
}
t = $\sqrt{m}$
Salida: mcd(s-t, n) es un factor de N

Este algoritmo es especialmente rápido para enteros de la forma $(c^a + d)(c^b + e)$ donde $c, |d|, |e|$ y $|a - b|$ son enteros positivos pequeños

Variación de Fermat de Lehman

En 1985, Lawrence propuso una manera de factorizar n cuando se cree que $n = pq$ con $p \leq q$, donde la proporción p/q es aproximadamente a/b y a y b son coprimos pequeños. Cuando $a = b = 1$, este algoriitmoo es lo mismo que el de Fermat. Asumiendo que $mcd(ab, n) = 1$.

Suponemos primero que ambos a y b sooon impares. Escriba $x = |\sqrt{abn}|$. Se prueban los enteroos $(x - 1)^2 - abn$, $i = 0, 1, 2,$ ... si son una raíz como en Fermat. Se supone que j es el primer valor de i para el cual este número es una raíz, entonces $(x - j)^2 - abn = y^2$

Entonces:

$$abn = (x + j)^2 - y^2 = (x + j + y)(x + j - y)$$

Se remueve los factores de ab de los dos factores del trinomio para obtener los factores de n. Esto es, $mcd(x + j + y, n)$ y $mcd(x + j - y, n)$ serán los factores de n.

Cuando una de los dos a o b es par y el otro es impar, los cálculos son un poco más complicados porque se debe lidiar con mitades. Lehman evito este problema multiplicando a, b y los otros números en el algoritmo por 2

Método de Pollard Rho. Es un algoritmo de factorización de enteros. Este fue inventado por John Pollard el año 1975. Este no utiliza mucho espacio de memoria, y el tiempo esperado de ejecución es proporcional a la raíz del factor primo más pequeño del número compuesto a ser factorizado. Está basada en la combinación de 2 ideas, que también son útiles para muchos otros métodos de factorización. La primera idea es la bien conocida Paradoja del cumpleaños: un grupo de al menos 23 personas seleccionadas aleatoriamente contiene 2 personas con el mismo cumpleaños en más del 50% de los casos. Más generalmente: si los números son elegidos de manera aleatoria en un conjunto de p números, la probabilidad de elegir el mismo número dos veces excede el 50% después de 1.177 $\sqrt{p}$ números elegidos. El primer duplicado se espera que aparezca después de que $c * \sqrt{p}$ hayan sido seleccionados, para algún pequeño constante c. La segunda idea es la siguiente: si p es algún divisor desconocido de n y las variables x, y son 2 enteros que se piensa son idénticas modulo p, en otras palabras $x \equiv y \, mod \, p$, entonces este puede ser verificado calculando $mcd(|x - y|, n)$; más importante, este cálculo puede revelar una factorización de n, a menos que x, y también sean idénticos modulo n.

Estas ideas pueden ser combinadas en un algoritmo de factorización de la siguiente manera.

Generar una secuencia en $\{0, 1, \dots, n - 1\}$ seleccionando x_0 y definiendo a x_{i+1} como el resto no negativo más pequeño de $x_i^2 + 1 \, mod \, n$, ya quee p divide a n los restos no negativos más pequeños $x_i \, mod \, p$ y $x_j \, mod \, p$ son iguales si y solo si x_i y x_j son identicos modulo p, ya que x_i mod p se comporta más o menos como un entero aleatorio en $\{0, 1, \dots, p - 1\}$ podemos esperar factorizar n calculando $mcd(|x_i - x_j|, n)$ para $i \neq j$ después de que al menos $c * \sqrt{p}$ elementos de la secuencia han sido calculados

11. ÍNDICE TENTATIVO

RESUMEN
CAPÍTULO I INTRODUCCIÓN
1.1. ANTECEDENTES
1.2. PLANTEAMIENTO DEL PROBLEMA
1.3. OBJETIVOS
1.3.1. OBJETIVO GENERAL
1.3.2. OBJETIVOS ESPECÍFICOS
1.4. JUSTIFICACIÓN
1.5. ALCANCES Y LÍMITES
1.6. METODOLOGÍA
CAPÍTULO II MARCO TEÓRICO
2.1. FACTORIZACION DE NUMEROS ENTEROS
2.2. METODOS DE FACTORIZACION DE NUMEROS ENTEROS
2.3. ANÁLISIS DE COMPLEJIDAD DE ALGORITMOS DE FACTORIZACIÓN DE NÚMEROS ENTEROS
CAPÍTULO III MARCO APLICATIVO
3.1. IMPLEMENTACIÓN DEL ALGORITMO 1
3.2. IMPLEMENTACIÓN DEL ALGORITMO 2
3.3. IMPLEMENTACIÓN DEL ALGORITMO 3
3.4. IMPLEMENTACIÓN DEL ALGORITMO 4

3.5. IMPLEMENTACIÓN DEL ALGORITMO 5

CAPITULO IV RESULTADOS Y ANÁLISIS
CAPÍTULO V CONCLUSIONES Y RECOMENDACIONES
BIBLIOGRAFÍA
ANEXOS

12. CRONOGRAMA DE AVANCE

CRONOGRAMA DE AVANCE TESIS DE GRADO																				
ACTIVIDADES	DEL 1 DE JULIO AL 30 DE NOVIEMBRE																			
	JULIO				AGOSTO				SEPTIEMBRE				OCTUBRE				NOVIEMBRE			
	1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4
Redacción del Capitulo II - Marco Teórico																				
Desarrollo del Capitulo III - Marco Aplicativo																				
Selección de algoritmos																				
Análisis de algoritmos seleccionados																				
Análisis de modificaciones en algoritmos seleccionados																				
Implementación de algoritmos seleccionados																				
Análisis de resultados																				
Redacción del Capitulo III - Marco Aplicativo																				
Redacción del Capitulo IV - Estado de la Hipótesis																				
Redacción del Capitulo V - Conclusiones y Recomendaciones																				

13. BIBLIOGRAFÍA

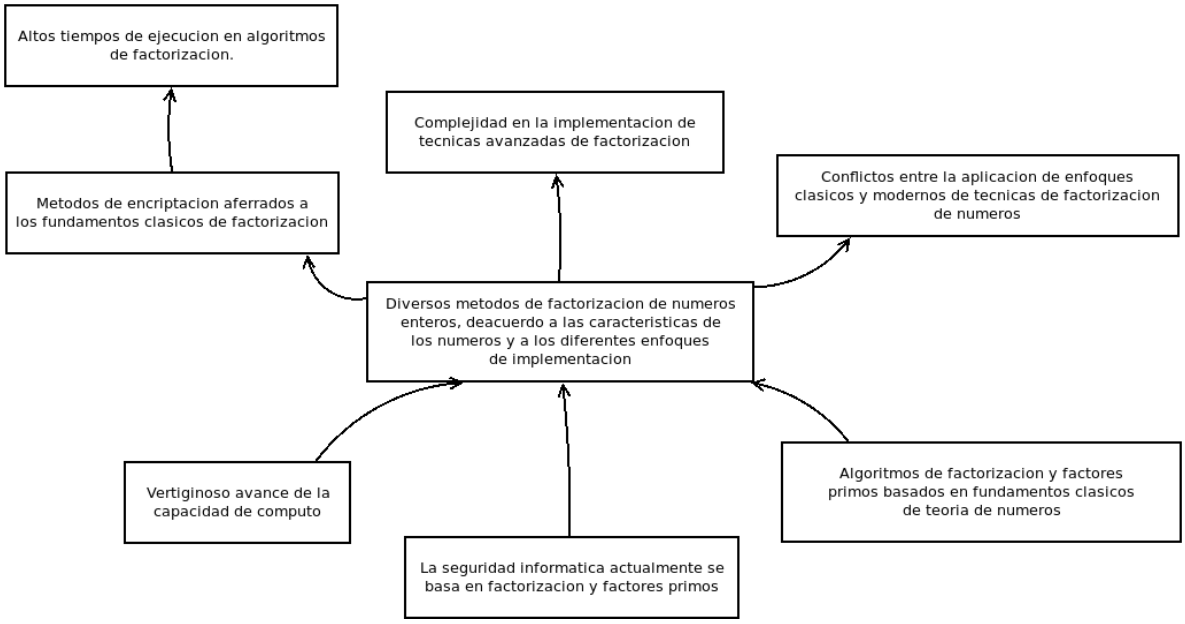
Abiodun E. Adeyemi, 2019, On Odd Perfect, MultiPerfect and Harmonic Number [en linea], Department of Mathematics University of Ibadan,<https://arxiv.org/pdf/1906.05798.pdf> [13 de junio de 2019]

KAM HUNG YAU, 2019, REPRESENTATION OF AN INTEGER AS THE SUM OF A PRIME IN ARITHMETIC PROGRESSION AND A SQUARE-FREE INTEGER WITH PARITY ON THE NUMBER OF PRIME FACTORS [en linea],<https://arxiv.org/pdf/1904.06783.pdf>[13 de junio de 2019]

Figura 2. 1 Redes Sociales más Utilizadas
Fuente: Camacho, 2017

ANEXOS

• ANEXO A – ÁRBOL DE PROBLEMAS



• DECLARACIÓN JURADA

• CARTA DOCENTE ASESOR PROYECTO DE GRADO

• CARTA REVISIÓN PERFIL DE PROYECTO DE GRADO

• CARTA ACEPTACIÓN COMO DOCENTE ASESOR

• CARTA CONFORMIDAD DEL PERFIL DEL ASESOR

• CARTA APROBACIÓN DE TEMA Y PERFIL DEL TUTOR

• CARTA AVAL DE LA INSTITUCIÓN