Cifrado de Hill. Primera Parte.

Objetivo.

El estudiante realizará el cifrado de Hill en forma serial para posteriormente desarrollar la versión paralela del mismo.

Antecedentes.

En 1929, Lester S. Hill desarrolla un sistema criptográfico de sustitución polialfabético, es decir, un mismo signo, en este caso una misma letra, puede ser representado en un mismo mensaje con más de un carácter.

Este sistema está basado en las operaciones modulo, la definición del alfabeto a utilizar y la multiplicación de matrices.

Operaciones Módulo. Estas operaciones se realizan con números enteros, estableciendo una distancia dada entre el valor de un modulo y el siguiente. Para ello se definen dos números, *a y b*, los cuales corresponden al número entero *a* y el módulo a aplicar, *b*. El resultado se denomina r.

La operación se denota como *a mod b = r*. De manera práctica, *a* será dividido entre *b* y se obtendrá un resultado *r* que equivale al residuo de la división. Si se toma "módulo 27", se consideran los números enteros 0, 1, 2,… , 26 y el resto se identifica con estos de forma cíclica.

Así, el 27 es igual a 0, el 28 a 1, el 29 a 2, etcétera, y lo mismo con los números negativos, de forma que – 1 es igual 26, – 2 es igual 25, etcétera. Además, se reducen las operaciones aritméticas (suma, resta, multiplicación y división) al conjunto de los números enteros módulo 27 de forma natural, es decir, al operar dos números enteros (módulo 27) el resultado se considera también módulo 27. Por ejemplo, si se realiza la multiplicación de los números 6 y 13, módulo 27, el resultado dará 24 (módulo 27), puesto que 6 x 13 = 78 y 78 = (2 x 27) + 24. O el inverso de 2, es decir, el número a tal que 2^{-1} a es igual a 1 (módulo 27), es 14, puesto que 2 x 14 = 28, que es igual a 1, módulo 27.

Alfabeto. El alfabeto está compuesto por los símbolos que conforman los mensajes a enviar. De inicio, se pueden contemplar sólo las letras del alfabeto, sin incluir las letras especiales para cada idioma. En ese caso, se tendrá la siguiente relación:

A	**B**	**C**	**D**	**E**	**F**	**G**	**H**	**I**
0	*1*	*2*	*3*	*4*	*5*	*6*	*7*	*8*
J	**K**	**L**	**M**	**N**	**O**	**P**	**Q**	**R**
9	*10*	*11*	*12*	*13*	*14*	*15*	*16*	*17*
S	**T**	**U**	**V**	**W**	**X**	**Y**	**Z**	
18	*19*	*20*	*21*	*22*	*23*	*24*	*25*	

Donde se tienen 26 símbolos, iniciando la enumeración desde el valor de cero.

El alfabeto se puede extender a manejar también los dígitos del 0 al 9, símbolos de puntuación, espacio y algún carácter especial.

En el cifrado, ambos extremos deben estar de acuerdo con el mismo sistema de alfabeto.

Algoritmo del cifrado de Hill. En el cifrado de Hill se utiliza una matriz cuadrada de números A como clave, la cual determina la transformación lineal $Y = A \cdot X$, donde Y, X son vectores columna y A y X se multiplican con la multiplicación de matrices.

Si se toma una matriz de dimensión n=3, se tendrá una matriz cuadrada de 3 x 3 y los vectores deberán coincidir con 3 renglones y una columna.

Por ejemplo, tomando el alfabeto anteriormente descrito, se quiere cifrar la frase:

COMPUTO DE ALTO DESEMPEÑO

Para empezar a codificarla, se debe separar la frase en grupos de tres letras. Dado que el espacio no está definido en el alfabeto, este no será tomado en cuenta, por lo que las agrupaciones quedan:

(COM) (PUT) (ODE) (ALT) (ODE) (SEM) (PEÑ) (O)

Revisando la agrupación anterior, se observan dos grupos que no están permitidos:

(PEÑ): La letra Ñ no está definida en el alfabeto, por lo que la cambiaremos por N.

(O): Las agrupaciones deben ser de tres letras, sin embargo, la frase no alcanza a cubrir con esta característica en el último grupo. Se soluciona agregando uno de los símbolos del alfabeto varias veces hasta completar el tamaño del grupo. En este caso, se tomará la letra X como el relleno del último grupo. Esto debe ser convención entre ambos extremos.

Por lo anterior, la agrupación finalmente queda como:

(COM) (PUT) (ODE) (ALT) (ODE) (SEM) (PEN) (OXX)

Ahora, cada uno de los grupos de letras será convertido al valor numérico correspondiente, formando el vector que corresponde:

(2 14 12) (15 20 19) (14 3 4) (0 11 19) (14 3 4) (18 4 12) (15 4 13) (14 23 23)

Para su cifrado, se requiere una matriz A no singular, esto es, una matriz que tenga A^{-1}. Se propone la siguiente:

$$A = \begin{bmatrix} 1 & 2 & 3 \\ 0 & 4 & 5 \\ 1 & 0 & 6 \end{bmatrix}$$

Teniendo la matriz, se procede a realizar la operación de multiplicar A por el vector columna X. Cada vector es uno de los que se obtuvieron al separar el mensaje y que después se han transpuesto para poder realizar la operación.

Para (COM) = (2 14 12):

$$\begin{bmatrix} 1 & 2 & 3 \\ 0 & 4 & 5 \\ 1 & 0 & 6 \end{bmatrix} \begin{bmatrix} 2 \\ 14 \\ 12 \end{bmatrix} = \begin{bmatrix} 66 \\ 116 \\ 74 \end{bmatrix}$$

Al resultado obtenido en X se le aplica el modulo de acuerdo al tamaño del alfabeto, esto es, mod 26:

$$\begin{bmatrix} 66 \\ 116 \\ 74 \end{bmatrix} mod 26 = \begin{bmatrix} 14 \\ 12 \\ 22 \end{bmatrix}$$

Reinterpretando en base al alfabeto los valores obtenidos:

(14 12 22) = (O M W)

Con lo que se ha cifrado el primer grupo de letras del texto en claro. Se procede de la misma manera para los grupos restantes, quedando el texto cifrado como:

OMWITZGGMBJKGGMKYMKDPZZW

Descripción del proyecto.

Para la realización del proyecto se realizará un programa serie que realice el cifrado de Hill en base al alfabeto indicado en este documento. De parámetros de entrada recibirá las dimensiones de la matriz, *n*, los valores de la matriz A y el mensaje en texto claro a cifrar.

Teniendo el programa serie, se convertirá a un programa paralelo, indicando las secciones del programa serie que se han paralelizado, si la descomposición es funcional o de dominio y las directivas de MPI a utilizar.

Puntos a evaluar:

Programa serie: 30% de la calificación.

1. Lectura de los valores *n* y la matriz A.
2. Lectura desde un archivo de texto del mensaje en claro.
3. Separación del mensaje en grupos de *n* letras.

4. Conversión de las letras a valores numéricos de acuerdo al alfabeto.
5. Multiplicación de la matriz por los vectores y conversión a valores módulo.
6. Reconversión de los valores a módulo a valores del alfabeto y salida a archivo de texto.

Programa Paralelizado: 40% de la calificación.

1. Justificación del tipo de descomposición utilizado.
2. Tipo de operaciones utilizadas en la paralelización.
3. Justificación del tipo de operación utilizada.
4. Escalabilidad y portabilidad del programa.

Análisis de los programas: 20% de la calificación.

1. Tiempo de ejecución del programa serial con diferentes tamaños del mensaje.
2. Tiempo de ejecución del programa paralelo con diferentes tamaños del mensaje y diferentes número de procesos.
3. Cálculo del SpeedUp obtenido con los programas para diferentes tamaños de mensaje y diferentes procesos.
4. Análisis de los resultados.
5. Graficas a página completa, indicando que es lo que se mide en cada eje.

Presentación: 10% de la calificación.

1. Contener los apartados de:
    a. Índices. De contenido, figuras, tablas y gráficas.
    b. Objetivo.
    c. Antecedentes.
    d. Desarrollo, códigos fuente, diseño y análisis.
    e. Conclusiones.
    f. Bibliografía. La bibliografía deberá estar citada dentro del documento utilizando el formato APA. Las referencias a páginas electrónicas deberá ser en el mismo formato. Los sitios deberán ser de universidades, centros de investigación o artículos en sitios reconocidos.