Programación de estructuras de datos y algoritmos fundamentales (Gpo 602)

Act 5.2 - Actividad Integral sobre el uso de códigos hash (Evidencia Competencia)

investigación y reflexión de la importancia y eficiencia de uso de las tablas hash para tales fines

Fecha

29/11/2025

Profesor

Daniel Perez R

Alumno

José Luis Gutiérrez Quintero A01739337

Investigación sobre las tablas hash y SHA-256

Las tablas hash son una estructura de datos muy usada cuando se necesita acceder rápido a elementos sin tener que buscarlos uno por uno. Básicamente, guardan datos en un arreglo, y para saber en qué lugar van se usa una función hash que convierte una “clave” (por ejemplo, una dirección IP) en un número. Ese número indica en qué posición del arreglo se guarda la información.

Si dos claves distintas generan el mismo número (esto se llama colisión), se tiene que manejar usando estrategias como direccionamiento abierto o listas enlazadas. En esta actividad usamos direccionamiento abierto, así que, si una posición ya está ocupada, se busca la siguiente libre.

En nuestro caso, la clave es la red de una IP (los dos primeros números), por ejemplo "145.25" y lo que guardamos como valor es: cuántas veces aparece esa red (número de accesos), cuántas IPs únicas hay (conexiones) y una lista ordenada de esas IPs. Esto sirve para ver qué redes están siendo más usadas o tienen más actividad.

Sobre SHA-256, es un tipo especial de función hash. Pero a diferencia de una función hash normal que guarda información en una tabla, SHA-256 se usa para generar una “huella digital” de un archivo. Esto sirve para saber si un archivo fue alterado, porque, aunque solo se cambie una letra, el resultado del SHA-256 será totalmente diferente.

SHA-256 no se puede revertir (o sea, no se puede obtener el archivo original desde el hash), y es muy seguro, por eso se usa en criptomonedas, contraseñas y firmas digitales. En este proyecto puede servir para asegurar que el archivo bitacora.txt no ha sido manipulado antes de ejecutarse el programa.

Reflexión personal

Durante esta actividad entendí mejor cómo funcionan las tablas hash y por qué son tan útiles para organizar información. Me pareció muy práctico que podamos usar los primeros dos números de una dirección IP (lo que llaman la red) como llave para agrupar datos relacionados. Esto nos permite tener un resumen rápido por red, como saber cuántos accesos hubo, cuántas direcciones únicas se conectaron y cuáles fueron exactamente.

Una cosa que también me quedó clara es que la eficiencia de una tabla hash depende mucho de la función hash que se utilice. Si la función está bien diseñada, los datos se distribuyen de forma uniforme y se evitan colisiones. Eso hace que insertar, buscar o eliminar datos sea muy rápido, en promedio con una complejidad O(1). Pero si hay muchas colisiones o la tabla está mal distribuida, entonces esas operaciones pueden volverse más lentas, hasta O(n) en el peor caso.

Sobre el uso de SHA-256 en este contexto, creo que sí es adecuado y suficiente para el objetivo de asegurar la integridad del archivo bitacora.txt. SHA-256 no es una función pensada para buscar rápido en una tabla, pero sí para detectar si un archivo ha sido modificado. Una de sus principales ventajas es que si cambias un solo carácter del archivo, el hash resultante cambia completamente. Eso permite verificar si el archivo es exactamente el mismo que se esperaba.

Otras ventajas de SHA-256 son que es muy seguro, no es fácil de invertir (es decir, no puedes saber qué texto generó ese hash), y es resistente a colisiones, lo cual significa que no se generan dos hashes iguales

para entradas distintas (al menos no de forma práctica). Por eso es usado también para guardar contraseñas, verificar archivos en descargas, y en blockchain como Bitcoin.

Sin embargo, una de las desventajas de SHA-256 es que no es muy rápido, y su complejidad computacional es O(n), ya que necesita recorrer todo el archivo para calcular el hash y hacer varias operaciones internas pesadas (mezclas, rotaciones, funciones lógicas, etc.). Por eso no sería buena idea usarlo como función hash dentro de una tabla, donde buscamos rapidez y no tanto seguridad. Para eso es mejor usar funciones simples como sumar valores ASCII o multiplicar por potencias.

En conclusión, usar tablas hash para este tipo de actividad me pareció bastante útil porque organizan bien los datos y permiten encontrar patrones por red. Y en cuanto a SHA-256, sí es muy útil para asegurar que la información no ha sido alterada, aunque no es ideal para tareas donde se necesita rendimiento, como una tabla hash directa. Aprendí que, dependiendo del problema, hay que elegir la herramienta correcta.

Referencias

- GeeksforGeeks. (2021). *Hashing Data Structure*. https://www.geeksforgeeks.org/hashing-data-structure/
- Cloudflare. (s. f.). *What is SHA-256?*. https://www.cloudflare.com/learning/ssl/what-is-sha-256/Cloudflare. (s. f.). *What is SHA- 256?*.https://www.cloudflare.com/learning/ssl/what-is-sha-256/
- Wikipedia contributors. (2023). *Hash function*. Wikipedia. https://en.wikipedia.org/wiki/Hash_function