

Reflexión Actividad 5.2

Cristóbal Alberto Escamilla Sada - A00827074

El propósito de esta actividad fue aplicar los conceptos de la tabla de Hash en la bitácora de accesos fallidos que hemos utilizado a lo largo del semestre. Esta bitácora cuenta con registros que contienen el mes, día, hora, dirección IP y la razón de falla. El objetivo fue crear una aplicación mediante la cual el usuario puede ingresar una dirección y recibir la cantidad de accesos a esa dirección así como un resumen de cada uno de los accesos. Para esto utilizamos el `unordered_map` de C++ e hicimos un struct llamado "registro", conteniendo los atributos que resumen dicho registro. El tipo de dato "registro" tiene como atributos el mes, día, hora y la razón de falla. El `unordered_map` organiza la información mediante una tabla de Hash. En este caso la llave consiste en una dirección IP de tipo string y el contenido consiste en un vector de tipo registro. El tamaño del vector es lo que va a determinar cuántos accesos tuvo x dirección IP. Esta implementación resultó en un código sumamente corto debido a la eficiencia de las clases de C++ que elegimos. La función para leer el archivo e insertar los registros al `unordered_map` tiene una complejidad lineal, sólo toma un par de segundos la lectura de más de 90,000 datos. Lo bello del `unordered_map` es que la búsqueda es de orden constante. Nuestra función de búsqueda encuentra el vector de registros de una IP ingresada y despliega todos los resúmenes en el vector. Por lo tanto, es de complejidad lineal. Estas son las únicas dos funciones en nuestra aplicación. Es una aplicación simple pero al mismo tiempo sumamente efectiva.

La naturaleza del problema a resolver es de seguridad cibernética; estamos lidiando con una red infectada. En este caso el uso de las tablas de Hash es sumamente ventajoso ya que la información está almacenada en forma de conjunto, en donde cada dirección IP tiene sus propiedades únicas (resumen). La desventaja es que los elementos no están conectados entre sí (como un grafo) por lo que no se puede simular una red de conexiones. Sin embargo, las tablas de Hash tienen un uso importante en la industria de la seguridad cibernética.

El Hashing no es muy ventajoso en el análisis de redes conteniendo múltiples conexiones. Sin embargo, es ventajoso en el análisis de intercambio de mensajes. Mediante el hashing, uno puede analizar si el mensaje que está recibiendo fue modificado durante el tránsito. También un dato interesante es que muchos de los usuarios/contraseñas en las páginas web se almacenan en tablas de hash. De esta manera la página se asegura que la clave que ingresó el usuario es la correcta. En fin, El hashing es una excelente herramienta para el análisis de conjuntos en una red de internet (Hashing, 2018).

A través de esta actividad aprendí la belleza del hashing y su implementación en el lenguaje de C++. Usando el `unordered_map` se pudieron agrupar los 90,000 datos adecuadamente y con un tiempo sumamente eficiente. Tristemente esta actividad marca el cierre de la unidad de formación. Sin embargo quedo sumamente satisfecho y motivado para seguir aprendiendo sobre la implementación de las estructuras de datos en los lenguajes de programación. Estos ADTs tienen una infinidad de aplicaciones en la vida laboral y en un par de años sé que las estaré utilizando en mi trabajo.

Referencias:

Hashing - Cybersecurity Glossary. (2018, September 28). Retrieved November 28, 2020, from Cybersecurity Glossary website:

<https://cybersecurityglossary.com/hashing/>