¿QUÉ ES UN HASH INFORMÁTICO?

AUTOR: JOSÉ MALDODANO.

MARZO: 2020

San Marcos

Introducción

La palabra hash es usada en el mundo de la informática para describir a una cadena de texto codificada. Una cadena formada por número y letras de longitud fija y en un orden único e irrepetible que representan a una serie de datos. Esta cadena de texto es creada gracias a una función criptográfica única conocida como función hash.

Ahora bien, ¿Qué importancia tienen los hashes en la actualidad? ¿Cuál es la historia? ¿Qué funciones hash existen en la actualidad? Las respuestas a dichas interrogantes son las que a continuación se pretende abordar en la lectura propuesta.

Contenido

¿Qué es un Hash?

Es una cadena de texto codificada. Una cadena formada por número y letras de longitud fija y en un orden único e irrepetible que representan a una serie de datos, creada por una función criptográfica única.

Origen de los hashes

El uso de los hashes se remonta a los inicios de la electrónica y la informática. En el año de 1.953, el investigador de IBM, Hans Peter Luhn propuso una forma de buscar y validar rápidamente información y documentos que se transformaría en lo que hoy llamamos función resumen o hash.

Luego de este trabajo inicial, se produjo una intensa investigación que dio origen al trabajo Indexing for rapid random-access memory, presentado por Arnold I. Dumey en 1956. Finalmente, el investigador W.W. Peterson trabajo en una aproximación práctica en 1957 y en 1961, Peterson creó la primera función hash conocida, la Cyclic Redundancy Check (Comprobación de Redundancia Cíclica) o CRC.

El objetivo de esta función era la de comprobar los datos transmitidos en redes y en sistema de almacenamiento digital. En ese entonces, su uso estaba bastante limitado a espacios académicos y militares. Pero su utilidad rápidamente permitió la expansión de la misma a la industria civil. Así en la actualidad CRC es la función más usada en el mundo y podemos verla implementada en casi todo aparato electrónico que existe en la actualidad.

Con el avance tecnológico de los años 70 y 80, especialmente en informática y computación, los investigadores comenzaron a investigar más sobre este tipo de funciones. En 1973, se presentó el trabajo The Art of Computer Programming, de D. E. Knuth y seguidamente G. D. Knott presentó Hashing Functions en 1975. Ambos libros sentaban las bases para el diseño de funciones hash en sistemas informáticos.

Finalmente 1977, J. Lawrence Carter y Mark N. Wegman propusieron The Universal hash functions o las definiciones que Ralph Merkle pública en 1979.

Todo este trabajo teórico se transforma en 1989 en la primera función hash diseñado específicamente para computadores, la función MD2, creada por Ralph Merkle. Fue acá donde comenzó la verdadera revolución de las funciones hash hasta nuestros días.

¿Cómo funciona esta tecnología?

Una función hash trabaja en base a tomar una serie de datos que son organizados en una serie de bloque de datos. Estos bloques de datos son luego sometidos a una serie de procesos matemáticos y lógicos. La finalidad de estos procesos es transformar todos los bloques de datos en una cadena alfanumérica única, irrepetible y de longitud fija. En dicha cadena se plasma toda la información de los bloques de datos sometidos al proceso de hashing.

Dicho de una forma más sencilla, esta función lo que hace es resumir una gran cantidad de datos en una cadena mucho más pequeña, irrepetible y con una longitud fija. De allí a que estas funciones sean también conocidas como funciones de resumen o funciones de verificación.

Este funcionamiento presenta una enorme ventaja. Y es que, el proceso de hashing solo puede hacerse en un sentido. Es decir, es imposible obtener los datos originales de los bloques teniendo tan solo el resultado en nuestras manos. Gracias a esto es posible crear un documento digital, someterlo a un proceso de hashing y usar dicho resultado como una prueba de autenticidad y no modificación. Esto pues cualquier modificación que hagamos en el documento dará como resultado un hash distinto y podremos darnos cuenta de la manipulación del mismo.

¿Qué tan seguras son estas funciones?

A continuación, se presenta información relacionada con la seguirdad en las funciones de Hash.

Una de las principales razones que llevaron a desarrollar estas funciones fue para conseguir un medio que garantizara la seguridad de la información digital. Esto significa que dichas funciones debían ser también muy seguras para evitar que un hacker pudiera romperlas y así alterar la información que estas pudieran proporcionar.

Esto se logró usando una aproximación muy cercana a la criptografía, tanto así, que se usan prácticamente los mismos principios para desarrollar funciones hash. Si bien, la seguridad absoluta no existe, pero las funciones hash actuales son muy seguras y confiables.

Por supuesto, existen funciones más seguras que otras. Especialmente cuando se habla de funciones certificadas o bien estudiadas como el caso de SHA-256 o SHA-512. Pero en todos los casos, la seguridad de las funciones hash actuales garantiza que nadie podrás saltarse su seguridad.

Un ejemplo de esto es SHA-256, una función muy utilizada en la actualidad y que es catalogada como muy segura. Para que un hacker logre romper SHA-256 necesitaría un enorme poder computacional. Tanto que ni todas las supercomputadoras del mundo le bastarían para lograrlo en un corto periodo de tiempo. Si en lugar de SHA-256, nombramos funciones como SHA-3 (Keccak), Blake3 o Scrypt este nivel de seguridad aumenta exponencialmente.

Características de las funciones hash

Una de las principales características de las funciones hash es que estas no son reversibles. Esto significa que un hash no puede convertirse en la serie de datos que dieron origen al mismo. Esta imposibilidad de obtener el mensaje original a partir del resumen obtenido se puede explicar con la aritmética modular (mod). Y es que estos deben ser computacionalmente fáciles y rápidos de obtener.

Por otro lado, los hashes ofrecen la capacidad de ser altamente resistentes a las colisiones. Es decir, dos series de datos distintos no pueden dar origen a un mismo resultado. Esto garantiza que las funciones siempre den como resultado cadenas de textos únicas e irrepetibles bajo cualquier condición.

Otra de las características elementales de los hashes es que el resultado de sus funciones tiene una longitud fija. Las funciones hash dependiendo de su construcción ofrecerán resultado de 16, 32, 48, 64, 128, 256 o más caracteres fijos.

Todo esto tiene un importante impacto en la usabilidad. Especialmente cuando se trata de verificación, autenticación y manejo de integridad de datos. Lo mejor de todo, es que

la mayoría de estas funciones son de dominio público y sus implementaciones son software libre.

Conclusiones y recomendaciones

En conclusión, las funciones Hash logran tomar una serie de datos que al mismo tiempo son organizados en una serie de bloques de datos. Estos bloques de datos son luego sometidos a una serie de procesos matemáticos y lógicos. La finalidad de estos procesos es transformar todos los bloques en una cadena alfanumérica única, irrepetible y de longitud fija. En dicha cadena se plasma toda la información de los bloques de datos sometidos al proceso de hashing. Todo ello, se realiza con el fin de conseguir un medio que garantice la seguridad de la información digital, y pese que hay unas más seguras que otras, en general las funciones Hash son bastante seguras, debido a que logran evitar que un hacker pueda romperlas y así alterar la información que estas proporcionen.
Uno de los beneficios que se obtienen de las funciones Hash es que son accsesibles a todo usuario y resumen la información a cedenas muy pequeñas lo que agiliza su ubicación.

Referencias bibliográficas

Maldodano J (2020). Qués es un Hash Informático. Recuperado de https://es.cointelegraph.com/explained/whats-a-hash

San Marcos

www.usanmarcos.ac.cr

San José, Costa Rica