Miramontes Sarabia Luis Enrique

Septiembre 12, 2019

Secure Hash Algorithm (SHA-1)

SHA-1 es un algoritmo usado para computar una representación condensada de un mensaje o archivo. Cuando un mensaje de cualquier longitud menor a 2^64-bits se tiene como entrada, SHA-1 producirá una salida (o message digest) con una longitud de 160-bits. Posteriormente se puede usar esta salida como entrada para algún algoritmo de firma digital que verifique la firma de un mensaje. Este procedimiento es normalmente más eficiente que firmar el mensaje completo, ya que en la mayoría de los casos el message digest será de un tamaño menor al mensaje completo.

A SHA-1 se le llama seguro porque no es computacionalmente factible encontrar un mensaje que corresponda a un message digest determinado, o de encontrar dos mensajes diferentes que sean capaces de producir la misma salida. Cualquier cambio a un mensaje producirá muy posiblemente un cambio en la salida y la verificación del Hash será inválida.

SHA-1 utilizará bloques de 512-bits para el cálculo del hash, cada bloque pude ser representando como una secuencia de 16 palabras que a su vez están formadas por 8 dígitos hexadecimales que representan strings de 32-bits. Con este tamaño es posible representar cualquier entero entre cero y 2^32 -1. Para la transformación también se usarán las operaciones lógicas AND, OR, XOR y NOT. Además de una operación X + Y definida como la suma de dos enteros x, y aplicando un módulo 2^32. También se hará uso de un shift circular a la izquierda y derecha.

El algoritmo necesita realizar un procedimiento de padding a los bloques de 512-bits, necesitamos que el número de bits de entrada totales del mensaje sean múltiplos de 512, por lo que se usará un padding que consiste en un ‘1’ seguido de m ‘0’. A su vez seguidos por un entero de 64-bits que representa el número de bits en el mensaje original sin padding.

Secure Hash Algorithm hace uso de una secuencia de 80 funciones que operan en tres parabras de 32-bits y producen una salida de 32-bits como respuesta. También se ocupan constantes K(t) con valores fijos dependiendo el valor de t.

Existen dos métodos para calcular el Hash usando SHA-1, el primero hace uso de dos buffers de 32-bits y una secuencia de ochenta parabras de 32-bits. Las palabras del primer buffer son llamadas A, B, C D, E. Las palabras del segundo buffer son H0, H1, H2, H3, H4, cuyos valores son fijos. Las secuencia de ochenta palabras serán llamadas W(0), W(1),... W(79).

Para generar el message digest los bloques obtenidos en el proceso de padding, llamados M(1), M(2),…, M(n) se procesarán en orden. El procesamiento de cada M(i) involucra 80 pasos. Se debe dividir cada bloque en 16 palabras W(0), W(1),… W(15) donde W(0) es la palabra menos significativa. Con estas palabras se van calculando los valores de W(t) haciendo uso de funciones y variables definidas anteriormente.

El segundo método sirve cuando se tiene almacenamiento limitado, su funcionamiento se basa en una implementación de lista circular que contendrá las 16 palabras de 32-bits de un bloque. Las operaciones usadas en este caso serán haciendo uso de una máscara y operaciones lógicas.

Para concluir el algoritmo SHA-1 parece ser más sólido, si lo pensamos y relacionamos con lo visto anteriormente en clase se podría decir que cada paso que tiene el algoritmo (los ochenta necesarios para obtener el message digest) lo hace más fuerte y resistente a ataques, se podría ver como si fuera un análogo del número de rondas en los ciphers como AES y DES.

Actualmente SHA-1 ha sido roto de manera práctica, investigando en internet encontré un sitio llamado shattered.io donde investigadores del CWI y Google lograron generar dos archivos PDF que tienen el mismo message digest y el cambio que tienen es un fondo de color.