Algoritmo MD5

Paulo Eduardo Bolher

Introdução

- O que é um Algoritmo de Hash?

    - Função que mapeia dados de tamanho variável para uma string de tamanho fixo.

    - Amplamente utilizado em segurança da informação, integridade de dados, e autenticação.

Histórico do MD5

- Criador: Ronald Rivest
- Ano de Criação: 1991
- Sucessor do MD4: Desenvolvido para corrigir fraquezas do algoritmo MD4
- Publicação: RFC 1321

Características do MD5

- Comprimento Fixo: Produz uma hash de 128 bits (32 caracteres hexadecimais)
- Velocidade: Rápido e eficiente
- Aplicações Comuns: Verificação de integridade de arquivos, armazenamento de senhas (não recomendado atualmente)

Estrutura Interna do MD5 - Rounds e Operações:

Quatro Rounds: 16 operações em cada round

- Round 1 (Operação F): Processa dados usando uma combinação de operações lógicas e aritméticas. Cada um dos 16 passos do round aplica a função F e uma adição de módulo 2^32.
- Round 2 (Operação G): Similar ao primeiro round, mas utiliza a função G para processar os dados. Este round também inclui 16 passos.
- Round 3 (Operação H): Utiliza a função H e processa os dados com 16 passos, novamente usando combinações de operações lógicas e aritméticas.
- Round 4 (Operação I): O round final usa a função I, com 16 passos de operações para completar o processamento dos dados.

Estrutura Interna do MD5 - Funções de Compressão:

- F (Round 1): (B AND C) OR ((NOT B) AND D)
    - Esta função faz a combinação bit a bit dos valores de entrada utilizando AND, OR e NOT.
- G (Round 2): (B AND D) OR (C AND (NOT D))
    - Utiliza uma combinação diferente das operações AND e NOT para adicionar complexidade.
- H (Round 3): B XOR C XOR D
    - A função H é uma simples combinação de XOR (ou exclusivo) entre os três valores de entrada.
- I (Round 4): C XOR (B OR (NOT D))
    - Utiliza uma combinação de XOR, OR e NOT para a mistura final dos valores.

Estrutura Interna do MD5 - Operações Usadas:

- AND, OR, XOR, NOT: Operações lógicas básicas que são usadas em diferentes combinações das funções F, G, H e I.
- Adições módulo 2^32: Durante o processamento, são realizadas adições onde os resultados são considerados módulo 2^32 (ou seja, só os últimos 32 bits são mantidos). Isso ajuda a manter a complexidade e a mistura dos dados.

Estrutura Interna do MD5 - Processo Detalhado:

- Inicialização: Quatro variáveis de 32 bits são inicializadas com valores específicos:
  - A = 0x67452301
  - B = 0xEFCDAB89
  - C = 0x98BADCFE
  - D = 0x10325476
- Processamento de Blocos: A mensagem é dividida em blocos de 512 bits. Cada bloco é processado separadamente pelos quatro rounds.
- Atualização das Variáveis: Após o processamento de cada bloco, as variáveis A, B, C e D são atualizadas com os resultados das operações.
- Concatenação Final: Os valores finais das variáveis A, B, C e D são concatenados para formar a hash final de 128 bits (32 caracteres hexadecimais).

Estrutura Interna do MD5 - Complemento Visual:

- Funções F, G, H, I é aplicada aos valores B, C, e D.
- M[k] é uma porção do bloco de dados atual.
- T[i] é uma constante de tabela específica.
- <<< s indica uma rotação para a esquerda dos bits.

$$A = B + ((A + F(B,C,D) + M[k] + T[i]) <<< s)$$

$$A = B + ((A + G(B,C,D) + M[k] + T[i]) <<< s)$$

$$A = B + ((A + H(B,C,D) + M[k] + T[i]) <<< s)$$

$$A = B + ((A + I(B,C,D) + M[k] + T[i]) <<< s)$$

MD5 VS MD4

Aspecto	MD4	MD5
Rounds	3 rounds de 16 passos cada	4 rounds de 16 passos cada
Funções de Compressão	3 funções (F, G, H)	4 funções (F, G, H, I)
Constantes	Não possui constantes específicas	Tabela de constantes específicas (T[i])
Deslocamentos de Bits	Menos variados e previsíveis	Mais variados e complexos
Resistência a Colisões	Vulnerável a ataques de colisão	Melhor resistência a colisões

Porque não deve ser utilizado

Referências

[1]https://www.ietf.org/rfc/rfc1321.txt

[2]https://www.avast.com/pt-br/c-md5-hashing-algorithm

[3]https://staff.emu.edu.tr/alexanderchefranov/Documents/CMPE412/MD5%20Message%20Digest%20Algorithm%20260220218.pdf