MD5 Algorithm

Harley Kozushko

Opening

- MD5 algorithm can be used as a digital signature mechanism.
- This presentation will explore the technical aspects of the MD5 algorithm.

Description of the MD5 Algorithm

- Takes as input a message of arbitrary length and produces as output a 128 bit “fingerprint” or “message digest” of the input.
- It is conjectured that it is computationally infeasible to produce two messages having the same message digest.
- Intended where a large file must be “compressed” in a secure manner before being encrypted with a private key under a public-key cryptosystem such as PGP.

MD5 Algorithm

- Suppose a b-bit message as input, and that we need to find its message digest.

MD5 Algorithm

- Step 1 – append padded bits:
  - The message is padded so that its length is congruent to 448, modulo 512.
    - Means extended to just 64 bits shy of being of 512 bits long.
  - A single “1” bit is appended to the message, and then “0” bits are appended so that the length in bits equals 448 modulo 512.

MD5 Algorithm

- Step 2 – append length:
  - A 64 bit representation of b is appended to the result of the previous step.
  - The resulting message has a length that is an exact multiple of 512 bits.

MD5 Algorithm

- Step 3 – Initialize MD Buffer
- A four-word buffer (A,B,C,D) is used to compute the message digest.
  - Here each of A,B,C,D, is a 32 bit register.

MD5 Algorithm

- Step 3 cont.
- These registers are initialized to the following values in hexadecimal:

word A: 01 23 45 67
word B: 89 ab cd ef
word C: fe dc ba 98
word D: 76 54 32 10

MD5 Algorithm

- Step 4 – Process message in 16-word blocks.
  - Four auxiliary functions that take as input three 32-bit words and produce as output one 32-bit word.

F(X,Y,Z) = XY v not(X) Z
G(X,Y,Z) = XZ v Y not(Z)
H(X,Y,Z) = X xor Y xor Z
I(X,Y,Z) = Y xor (X v not(Z))

MD5 Algorithm

- Step 4 – Process message in 16-word blocks cont.
  - if the bits of X, Y, and Z are independent and unbiased, the each bit of F(X,Y,Z), G(X,Y,Z), H(X,Y,Z), and I(X,Y,Z) will be independent and unbiased.

MD5 Algorithm

- Step 5 – output
  - The message digest produced as output is A, B, C, D.
  - That is, output begins with the low-order byte of A, and end with the high-order byte of D.

Summary

- The MD5 algorithm is simple to implement, and provides a “fingerprint” or message digest of a message of arbitrary length.
- The difficulty of coming up with two messages with the same message digest is on the order of 2^64 operations.