MD5原理

13331233 孙中阳

MD5简介

MD5系 `Message Digest Algorithm` 的简称，译为消息摘要算法第五版，是一种计算机安全领域广泛使用的散列函数，能为消息提供完整性保护，主流编程语言普遍已有MD5实现。

MD5特点

1、压缩性：任意长度的数据，算出的MD5值长度都是固定的
2、容易计算：从原数据计算出MD5值很容易
3、抗修改性：对原数据进行任何改动，哪怕只修改1个字节，所得到的MD5值都有很大区别
4、强抗碰撞：已知原数据和其MD5值，想找到一个具有相同MD5值的数据（即伪造数据）是非常困难的

实现

简单的说，MD5首先将需要加密的信息分为512位（64 Byte）大小的数据块，对每一块进一步划分为16个小块进行线性运算，得到一个128位散列值，然后将其和之前得到的散列值加以合并即生成结果

一、填充

由于数据大小不总是512位的整数倍，故需要对其填充：如果输入信息的长度(bit)对512求余的结果不等于448，就需要填充使得对512求余的结果等于448。填充时对需要填充的部分先填一个1，剩下的填0。填充完后，信息的长度就为N * 512+448(bit)
然后，对于剩下的64位（512 - 448），用数据的大小的值进行填充，最终得到(N+1) * 512位的数据

二、循环运算

首先将四个幻数填充进128个散列值，然后循环进行四轮运算即可
四个线性运算：

```
F(X,Y,Z)=(X&Y)|((~X)&Z)
G(X,Y,Z)=(X&Z)|(Y&(~Z))
H(X,Y,Z)=X^Y^Z
I(X,Y,Z)=Y^(X|(~Z))
```

其中XYZ为当前散列值中的三个

对应的四种操作：

```
FF(a,b,c,d,Mj,s,ti)表示a=b+((a+F(b,c,d)+Mj+ti)<<<s)
GG(a,b,c,d,Mj,s,ti)表示a=b+((a+G(b,c,d)+Mj+ti)<<<s)
HH(a,b,c,d,Mj,s,ti)表示a=b+((a+H(b,c,d)+Mj+ti)<<<s)
II(a,b,c,d,Mj,s,ti)表示a=b+((a+I(b,c,d)+Mj+ti)<<<s)
```

将线性运算和小数据块结合起来

四轮运算：

```
第一轮
a=FF(a,b,c,d,M0,7,0xd76aa478)
b=FF(d,a,b,c,M1,12,0xe8c7b756)
c=FF(c,d,a,b,M2,17,0x242070db)
d=FF(b,c,d,a,M3,22,0xc1bdceee)
a=FF(a,b,c,d,M4,7,0xf57c0faf)
b=FF(d,a,b,c,M5,12,0x4787c62a)
c=FF(c,d,a,b,M6,17,0xa8304613)
d=FF(b,c,d,a,M7,22,0xfd469501)
a=FF(a,b,c,d,M8,7,0x698098d8)
b=FF(d,a,b,c,M9,12,0x8b44f7af)
c=FF(c,d,a,b,M10,17,0xffff5bb1)
d=FF(b,c,d,a,M11,22,0x895cd7be)
a=FF(a,b,c,d,M12,7,0x6b901122)
b=FF(d,a,b,c,M13,12,0xfd987193)
c=FF(c,d,a,b,M14,17,0xa679438e)
d=FF(b,c,d,a,M15,22,0x49b40821)
```

第二轮

```
a=GG(a,b,c,d,M1,5,0xf61e2562)
b=GG(d,a,b,c,M6,9,0xc040b340)
c=GG(c,d,a,b,M11,14,0x265e5a51)
d=GG(b,c,d,a,M0,20,0xe9b6c7aa)
a=GG(a,b,c,d,M5,5,0xd62f105d)
b=GG(d,a,b,c,M10,9,0x02441453)
c=GG(c,d,a,b,M15,14,0xd8a1e681)
d=GG(b,c,d,a,M4,20,0xe7d3fbc8)
a=GG(a,b,c,d,M9,5,0x21e1cde6)
b=GG(d,a,b,c,M14,9,0xc33707d6)
c=GG(c,d,a,b,M3,14,0xf4d50d87)
d=GG(b,c,d,a,M8,20,0x455a14ed)
a=GG(a,b,c,d,M13,5,0xa9e3e905)
b=GG(d,a,b,c,M2,9,0xfcefa3f8)
c=GG(c,d,a,b,M7,14,0x676f02d9)
d=GG(b,c,d,a,M12,20,0x8d2a4c8a)
```

第三轮

```
a=HH(a,b,c,d,M5,4,0xfffa3942)
b=HH(d,a,b,c,M8,11,0x8771f681)
c=HH(c,d,a,b,M11,16,0x6d9d6122)
d=HH(b,c,d,a,M14,23,0xfde5380c)
a=HH(a,b,c,d,M1,4,0xa4beea44)
b=HH(d,a,b,c,M4,11,0x4bdecfa9)
c=HH(c,d,a,b,M7,16,0xf6bb4b60)
d=HH(b,c,d,a,M10,23,0xbebfbc70)
a=HH(a,b,c,d,M13,4,0x289b7ec6)
b=HH(d,a,b,c,M0,11,0xeaa127fa)
c=HH(c,d,a,b,M3,16,0xd4ef3085)
d=HH(b,c,d,a,M6,23,0x04881d05)
a=HH(a,b,c,d,M9,4,0xd9d4d039)
b=HH(d,a,b,c,M12,11,0xe6db99e5)
c=HH(c,d,a,b,M15,16,0x1fa27cf8)
d=HH(b,c,d,a,M2,23,0xc4ac5665)
```

第四轮

```
a=II(a,b,c,d,M0,6,0xf4292244)
b=II(d,a,b,c,M7,10,0x432aff97)
c=II(c,d,a,b,M14,15,0xab9423a7)
d=II(b,c,d,a,M5,21,0xfc93a039)
a=II(a,b,c,d,M12,6,0x655b59c3)
b=II(d,a,b,c,M3,10,0x8f0ccc92)
c=II(c,d,a,b,M10,15,0xffeff47d)
d=II(b,c,d,a,M1,21,0x85845dd1)
a=II(a,b,c,d,M8,6,0x6fa87e4f)
b=II(d,a,b,c,M15,10,0xfe2ce6e0)
c=II(c,d,a,b,M6,15,0xa3014314)
d=II(b,c,d,a,M13,21,0x4e0811a1)
a=II(a,b,c,d,M4,6,0xf7537e82)
b=II(d,a,b,c,M11,10,0xbd3af235)
c=II(c,d,a,b,M2,15,0x2ad7d2bb)
d=II(b,c,d,a,M9,21,0xeb86d391)
```

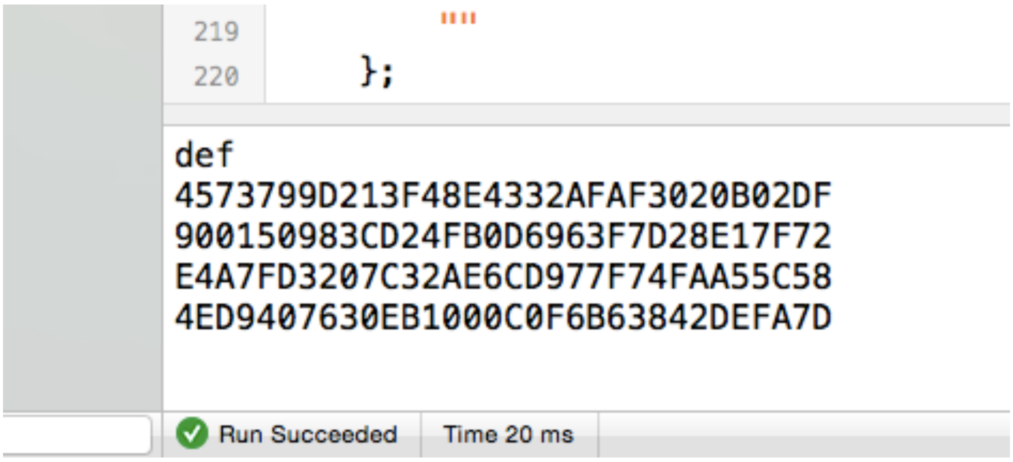
每轮循环后，将A，B，C，D分别加上a，b，c，d，然后进入下一循环。在此过程中一部分信息丢失，使得最终的MD5值不可逆
（1 + 6 = 7，但不能反推出7一定是1与6的和）

Samples

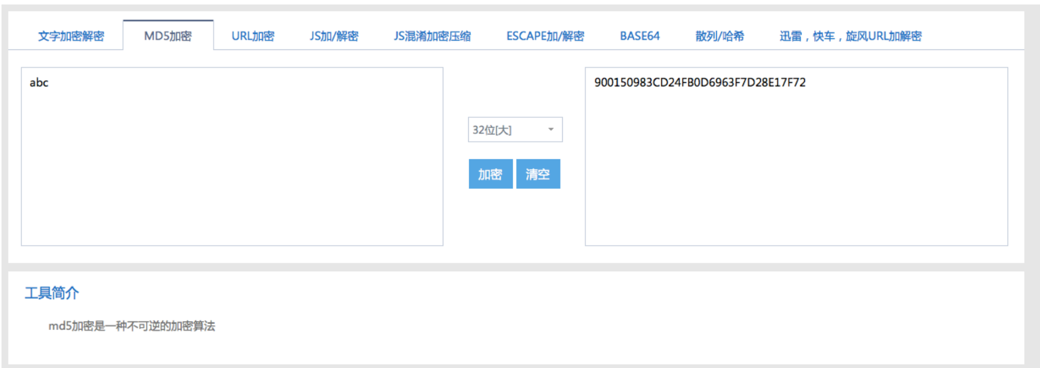
在本次试验中，选择以下四个字符串作为输入数据

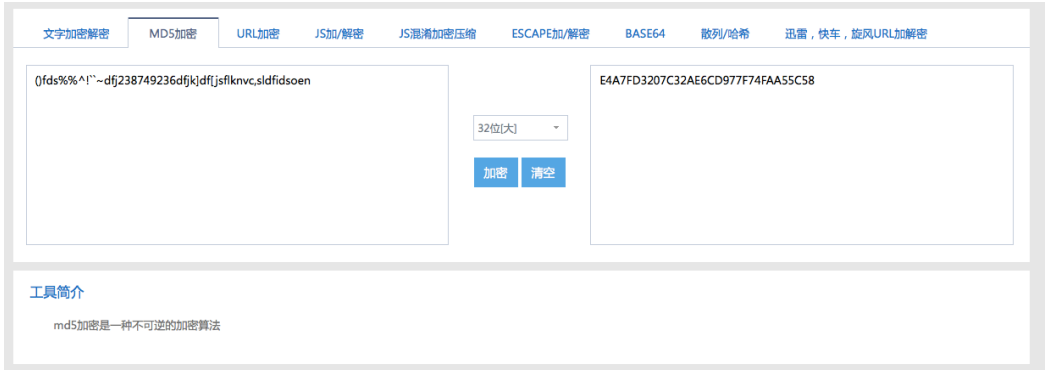
```
"abcabcabcabcabcabcabcabcabcabcabcabcabcabcabcabcabcabcabcabcabca
bcabcabcabcabcabcabcabcabcabcabcabcabcabcabcabcabcabcabcabcabcabc
abcabcabcabcabcabcabcabcabcabcabcabcabcabcabcabcabcabcabcabcabcab
cabcabcabcabcabcabcabcabcabcabcabcabcabcabcabc",
"abc",
"()fds%%^!``~dfj238749236dfjk]df[jsflknvc,sldfidsoen",
"def"
```

最终得到的结果为



与网上工具得到的值进行对比，结果相同




文字加密解密
MD5加密
URL加密
JS加/解密
JS混淆加密压缩
ESCAPE加/解密
BASE64
散列/哈希
迅雷，快车，旋风URL加解密
(}fds%%^!`~dfj238749236dfjk]df[jsflknvc,sldfidsoen
32位[大]
加密
清空
E4A7FD3207C32AE6CD977F74FAA55C58
工具简介
md5加密是一种不可逆的加密算法


文字加密解密
MD5加密
URL加密
JS加/解密
JS混淆加密压缩
ESCAPE加/解密
BASE64
散列/哈希
迅雷，快车，旋风URL加解密
def
32位[大]
加密
清空
4ED9407630EB1000C0F6B63842DEFA7D
工具简介
md5加密是一种不可逆的加密算法