# Hash相关算法加密与解密

## Hash算法

一般翻译做“散列”，也有直接音译为"哈希"的，就是把任意长度的输入（又叫做预映射， pre-image），通过散列算法，变换成固定长度的输出，该输出就是散列值。这种转换是一种压缩映射，也就是，散列值的空间通常远小于输入的空间，不同的输入可能会散列成相同的输出，而不可能从散列值来唯一的确定输入值。

简单的说就是一种将任意长度的消息压缩到某一固定长度的消息摘要的函数。

HASH主要用于信息安全领域中加密算法，他把一些不同长度的信息转化成杂乱的128位的编码里,叫做HASH值. 也可以说，hash就是找到一种数据内容和数据存放地址之间的映射关系

了解了hash基本定义，就不能不提到一些著名的hash算法，MD5 和 SHA1 可以说是目前应用最广泛的Hash算法，而它们都是以 MD4 为基础设计的。那么他们都是什么意思呢？这里简单说一下：

### **MD4**

MD4(RFC 1320)是 MIT 的 Ronald L. Rivest 在 1990 年设计的，MD 是 Message Digest 的缩写。它适用在32位字长的处理器上用高速软件实现--它是基于 32 位操作数的位操作来实现的。

### **MD5**

MD5(RFC 1321)是 Rivest 于1991年对MD4的改进版本。它对输入仍以512位分组，其输出是4个32位字的级联，与 MD4 相同。MD5比MD4来得复杂，并且速度较之要慢一点，但更安全，在抗分析和抗差分方面表现更好

### **SHA1**

SHA1是由NIST NSA设计为同DSA一起使用的，它对长度小于264的输入，产生长度为160bit的散列值，因此抗穷举(brute-force)性更好。SHA-1 设计时基于和MD4相同原理,并且模仿了该算法。

那么这些Hash算法到底有什么用呢？

Hash算法在信息安全方面的应用主要体现在以下的3个方面：

#### 文件校验

我们比较熟悉的校验算法有奇偶校验和CRC校验，这2种校验并没有抗数据篡改的能力，它们一定程度上能检测并纠正数据传输中的信道误码，但却不能防止对数据的恶意破坏。

MD5 Hash算法的"数字指纹"特性，使它成为目前应用最广泛的一种文件完整性校验和(Checksum)算法，不少Unix系统有提供计算md5 checksum的命令。

#### 数字签名

Hash 算法也是现代密码体系中的一个重要组成部分。由于非对称算法的运算速度较慢，所以在数字签名协议中，单向散列函数扮演了一个重要的角色。 对 Hash 值，又称"数字摘要"进行数字签名，在统计上可以认为与对文件本身进行数字签名是等效的。而且这样的协议还有其他的优点。

#### 鉴权协议

如下的鉴权协议又被称作"挑战--认证模式：在传输信道是可被侦听，但不可被篡改的情况下，这是一种简单而安全的方法。

### **Emule**

emule里面他具体起到什么作用呢？什么是文件的hash值呢？



emule又称电骡，类似迅雷，基于P2P （Peer-to-peer缩写，指的是点对点的意思的软件）， 它采用了“多源文件传输协议”(MFTP，the Multisource FileTransfer Protocol)。在协议中，定义了一系列传输、压缩和打包还有积分的标准，emule 对于每个文件都有md5-hash的算法设置，这使得该文件独一无二，并且在整个网络上都可以追踪得到。

MD5-Hash-文件的数字文摘通过Hash函数计算得到。不管文件长度如何，它的Hash函数计算结果是一个固定长度的数字。与加密算法不同，这一个Hash算法是一个不可逆的单向函数。采用安全性高的Hash算法，如MD5、SHA时，两个不同的文件几乎不可能得到相同的Hash结果。因此，一旦文件被修改，就可检测出来。

当我们的文件放到emule里面进行共享发布的时候，emule会根据hash算法自动生成这个文件的hash值，他就是这个文件唯一的身份标志，它包含了这个文件的基本信息,然后把它提交到所连接的服务器。当有他人想对这个文件提出下载请求的时候，这个hash值可以让他人知道他正在下载的文件是不是就是他所想要的。尤其是在文件的其他属性被更改之后（如名称等）这个值就更显得重要。而且服务器还提供了,这个文件当前所在的用户的地址,端口等信息,这样emule就知道到哪里去下载了.

一般来讲我们要搜索一个文件，emule在得到了这个信息后，会向被添加的服务器发出请求，要求得到有相同hash值的文件。而服务器则返回持有这个文件的用户信息。这样我们的客户端就可以直接的和拥有那个文件的用户沟通，看看是不是可以从他那里下载所需的文件。

对于emule中文件的hash值是固定的，也是唯一的，它就相当于这个文件的信息摘要，无论这个文件在谁的机器上，他的hash值都是不变的，无论过了多长时间，这个值始终如一，当我们在进行文件的下载上传过程中，emule都是通过这个值来确定文件。

#### 什么是userhash

道理同上，当我们在第一次使用emule的时候，emule会自动生成一个值，这个值也是唯一的，它是我们在emule世界里面的标志，只要你不卸载，不删除config，你的userhash值也就永远不变，积分制度就是通过这个值在起作用，emule里面的积分保存，身份识别，都是使用这个值，而和你的 id和你的用户名无关，你随便怎么改这些东西，你的userhash值都是不变的，这也充分保证了公平性。其实他也是一个信息摘要，只不过保存的不是文件信息，而是我们每个人的信息。

#### 什么是hash文件

我们经常在emule日至里面看到，emule正在hash文件，这里就是利用了hash算法的文件校验性这个功能了，文章前面已经说了一些这些功能，其实这部分是一个非常复杂的过程，目前在ftp,bt等软件里面都是用的这个基本原理，emule里面是采用文件分块传输，这样传输的每一块都要进行对比校验，如果错误则要进行重新下载，这期间这些相关信息写入met 文件，直到整个任务完成，这个时候part文件进行重新命名，然后使用move命令，把它传送到incoming文件里面，然后met文件自动删除，所以我们有的时候会遇到hash文件失败，就是指的是met里面的信息出了错误不能够和part文件匹配，另外有的时候开机也要疯狂hash，有两种情况一种是你在第一次使用，这个时候要hash提取所有文件信息，还有一种情况就是上一次你非法关机，那么这个时候就是要进行排错校验了。

## MD5加密工具

### HashCalc

HashCalc是一款绿色小软件，它能够检验任意文件、文字串、HEX字串的MD5码,超强的文件校验码计算小工具！

速度快，支持算法多，直接拖动即可计算，十分方便！ 当然还可对字串直接计算！

|  |  |
| --- | --- |
| Hash algorithms | Hash Size (in bits) |
| MD2 | 128 |
| MD4 | 128 |
| MD5 | 128 |
| SHA-1 | 160 |
| SHA-2(256) | 256 |
| SHA-2(384) | 384 |
| SHA-2(512) | 512 |
| RIPEMD-160 | 160 |
| PANAMA | 256 |
| TIGER | 192 |
| Checksum algorithms | Checksum Size (in bits) |
| ADLER32 | 32 |
| CRC32 | 32 |
| Other hash algorithms | Hash Size (in bits) |
| eDonkey/eMule (MD4-based hash used in many P2P applications) | 128 |

