§5.2 认证函数

- **Hash函数**

若一个合法文件有数兆字节长，为了进行签名认证，自然按160比特分划成一块一块，用相同的密钥独立地签每一个块。然而，这样太慢。

解决办法：引入可公开的密码Hash函数。它将取任意长度的消息做自变量，结果产生规定长度的消息摘要。(如使用数字签名标准DSS，消息摘要为160比特)，然后签名消息摘要。

对数字签名来说，Hash函数h这样使用：

消息：	x	任意长
消息摘要：	Z=h(x)	160bits
签名：	$y=sig_k(Z)$	320bits

§5.2 认证函数

验证签名(x, y)，其中$y = sig_k(h(x))$，使用公开的Hash函数h，重构作$Z' = h(x)$。然后检验$Ver_k(Z, y)$，看是否$Z = Z'$。

用以认证的Hash函数，是否会减弱认证方案的安全性？签名的对象由完整消息变成消息摘要，这就有可能出现伪造。

伪造方式一：攻击者以一个有效签名(x, y)开始，$y = sig_k(h(x))$。首先他计算$Z = h(x)$，并企图找到一个$x' \neq x$满足$h(x') = h(x)$。若他做到这一点，则(x', y)也将成为一个对消息x'的有效签名。为防止这一点，函数h必须具有无碰撞特性。

定义1(弱无碰撞) Hash函数h称为是弱无碰撞的，是指对给定消息$x \in X$，在计算上几乎找不到异于x的$x' \in X$使$h(x) = h(x')$。

§5.2 认证函数

伪造方式二：攻击者首先找到两个消息$x \neq x'$，满足$h(x)=h(x')$，然后攻击者把x 给Bob且使他对x的摘要h(x)签名，从而得到y，那么(x', y)是一个有效的伪造。

定义2(强无碰撞) Hash函数h被称为是强无碰撞的，是指在计算上几乎不可能找到相异的x和x'使得$h(x) = h(x')$。

注：强无碰撞自然含弱无碰撞。

伪造方式三：在某种签名方案中可伪造一个随机消息摘要Z的签名。若h的逆函数h^{-1}是易求的，可算出$h^{-1}(Z)=x$满足$Z=h(x)$，则(x, y)(其中$y=sig_k(h(x))$)为合法签名。

定义3(单向的) Hash函数h被称为是单向的，是指计算h的逆函数h^{-1}在计算上不可行。

§5.2 认证函数

Hash函数应具有的性质：

（a） Hash函数可以作用于一个任意长度的数据分组，产生一个固定长度的输出。

（b）任意给定消息M，计算h = H(M)容易。

（c）任意给定h，找到M满足H(M) = h很难，计算上的不可行性，即单向性。

（d）任意给定的数据块M，找到不等于M的M’，使H(M)=H(M’)在计算是不可行性。即弱无碰撞。

（e）找到任意数据对(x, y)，满足H(x) = H(y)是计算不可行的。即强无碰撞。

§5.2 认证函数

Hash函数的构造形式：

（1）利用数学难题，如因子分解、离散对数问题等

（2）利用某些私钥密码体制，如DES等

（3）通过直接构造复杂的非线性关系，如MD4, MD5, SHA-1等

§5.2 认证函数

对Hash函数的攻击有两类

（1）强力攻击（或穷举攻击）。典型防范：生日攻击

任找23人，从中总能选出两人具有相同生日的概率至少为1/2。根据此特点，一个40bit的消息摘要将是不安全的。因为在100万个随机散列值中将找到一个碰撞的概率为1/2。通常建议，消息摘要的尺寸为128bits。

（2）特殊攻击。

§5.2 认证函数

常用的Hash算法

（1）MD5算法：

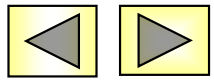
由麻省理工学院(MIT)Rivest提出， MD5不基于任何假设和密码体制，采用直接构造法。

输入：任意长度的消息；

输出：128位消息摘要；

处理：以512位输入数据块为单位；

安全性：采用穷举攻击寻找一个消息具有给定Hash值的计算困难性为2^{128}，若采用生日攻击，寻找有相同Hash值的两个消息需要试验2^{64}个消息。单轮MD5已有攻击结果，但对MD5全部四轮无有效攻击法。



§5.2 认证函数

（2）安全散列算法(SHA)：

由美国国家标准技术研究所(NIST) 提出，作为联邦信息处理标准于1993年发表（FIPS PUB 180），1995年修订，称为SHA-1，SHA-1基于MD4设计。设计目的是用于数字签名标准算法DSS。

输入：最大长度为2^{64}位的消息；

输出：160位消息摘要；

处理：输入以512位数据块为单位处理；

安全性：消息摘要比MD5长32位，采用穷举攻击计算困难性为2^{160}，若采用生日攻击计算困难性为2^{80}，安全性更高。但硬件实现速度较MD5慢25%。