RSA 数字签名算法在软件加密中的应用

谭毓安 王 佐 曹元大 （北京理工大学计算机科学与工程系 100081）

摘要：为避免软件的非授权使用，需要对软件进行加密。本文提出一种使用数字签名算法将软件绑定到特定计算机的方法。软件开发商根据主机序号等信息生成一个数字签名，作为对用户的授权。软件运行时，通过验证数字签名来确定软件使用的合法性。将数字签名技术应用到软件加密中，能够在不依赖任何特殊硬件的情况下实现软件保护。

关键词：数字签名 软件保护 公开密码体制

计算机软件的研发过程需要巨大投入，而生产和复制则十分容易，所以加密就成了保护软件的一种必要手段。依据加密原理和方式，目前所采用的加密方法主要分为两大类：硬加密和软加密[1]。硬加密主要有钥匙盘、软件狗、CD指纹等。其特点是提供给用户一个"硬件"实体作为授权，增加了额外的成本开销，"硬件"的生产不能被软件开发商很好地控制，而且不能有效地利用 Internet 来分发软件。

软加密方式是通过软件序列号或注册码的形式来授权用户使用被保护的软件。在安装或软件运行时，对计算机硬件进行检测，以获得特殊指纹信息，即代表这个计算机的主机 ID。用户将主机 ID 通过 Internet、E－mail、电话等方式向软件开发商进行注册，软件开发商再向用户提供一个与主机 ID 相对应的序列号(license number)。用户输入序列号后，软件中的程序检测计算机的主机 ID 和该序列号是否对应。如果用户将软件安装或拷贝到其他计算机上，由于主机 ID 不同，就必须向软件开发商申请新的序列号。这种方式不增加任何的硬件成本，不存在兼容性问题，而且授权过程完全由软件开发商控制，因此受到软件开发商和用户双方的支持。微软最新发布的 Windows XP 和 Office 系列软件即采用了这种软加密方式。这种方式也被称为许可证加密方式或序列号加密方式。

1 注册码加密

1.1 注册码的计算和校验

主机 ID 可以根据计算机的 CPU 序列号、网卡序列号、硬盘卷标、BIOS 的日期及版本等信息产生。主机 ID 对每一个计算机应是惟一的。

用户将主机 ID 提供给软件开发商后，得到一个<ID，注册码>二元组，该二元组就作为对该计算机的授权。

如图 1 所示，注册码加密包括了两个模块：注册码计算模块和校验模块。注册码计算模块由软件开发商掌握，它根据用户提供的识别码计算出对应的注册码。而校验模块内嵌于分发给用户的软件程序中，它首先产生一个主机 ID。

在运行软件程序时，校验模块首先确认 ID 的合法性。如果本机的 ID 与二元组中的 ID 不一致，说明该授权不是对本机的，程序终止。接着再检验<ID，注册码>二元组是否匹配，即注册码是否正确。如果不匹配，说明注册码是伪造的，程序终止。

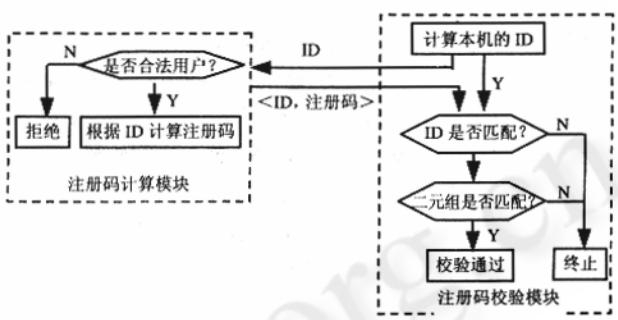

图 1 注册码计算模块和校验模块

<ID，注册码>二元组由软件开发者的计算机上由注册码计算模块产生，在用户的计算机上由软件程序中的校验模块来校验。校验过程和计算过程之间使用的算法和密码之间必须存在一定的联系，才能识别出<ID，注册码>二元组是否正确。

1.2 使用公开密码体制计算和校验注册码

计算模块和校验模块分别用于产生和验证<ID，注册码>二元组，因此这两个模块之间必须共享一些信息。但是，校验过程不能等同于计算过程，也要确保不能够从校验过程推导出计算过程。一些使用注册码方式加密的软件中，其校验模块暴露了太多的信息，甚至在校验模块包括了二元组的计算过程，由校验模块根据 ID 计算出其注册码，再与用户输入的<ID，注册码>二元组进行匹配。从表面上，这种方式能够起到软件保护的作用，但由于校验模块内嵌于软件之中，破解者通过静态分析和动态跟踪等手段完全可以获得校验模块，因此能够获得注册码的计算方法。然后，破解者就计算任何一台计算机的 ID 所需的注册码，软件保护屏障被完全突破。例如，在研究了 UltraEdit、著名调试工具 Soft-

ICE等的校验方法后,破解者已经开发出这些软件的注册码计算程序(又称注册机)。

使用注册码加密必须利用公开密码体制。软件开发者选定一个加密算法和一个私钥/公钥对。私钥用来为ID产生数字签名,这个数字签名即作为软件的注册码。公钥存在于软件程序中,校验模块得到计算机的ID后,利用公钥来验证数字签名(即注册码)。如果签名正确,则说明注册码是由软件开发者签发的,是一个合法授权的拷贝。

公开密码体制利用公钥和私钥来进行数据的加解密和验证数字签名,从公钥不能推导出私钥。使用暴力穷举法或其他攻击方法不能破解具有一定位数的密钥。使用基于公开密码体制的数字签名算法来计算和校验注册码,在软件程序中包含公钥和签名的验证过程,尽管破解者通过分析跟踪软件程序可以获得注册码的校验过程和公钥,进而获得注册码的计算方法,但他能得到计算过程中所需要的私钥,因此不能计算出注册码。

2 基于RSA算法的软件加密

数字签名算法包括两大类:基于公开密钥的数字签名和基于共享密钥的数字签名。前者包括ElGamal、DSA、RSA等,后者包括HMAC、Asmuth－Bloom等。在软件加密中只能使用基于公开密钥的数字签名。计算机的ID作为签名算法的消息,得到的签名作为注册码。算法的公钥可以公开,即存在校验模块中,私钥则必须秘密保存,只存在于计算模块中。注册码的计算和校验方法也不需要保密。

本文以RSA算法为例说明基于公开密码体制的数字签名算法在软件加密中的应用,也可以使用其他算法,如DSA、ElGamal等。

RSA算法[2]既可用于数据的加密解密,也可用于数字签名。RSA算法产生的数字签名由于其长度固定,适合作为软件的注册码。

2.1 数字签名的计算和校验

计算机的ID的长度不定,通过MD5报文摘要算法,产生固定长度的报文摘要,再作为签名算法的消息输入。

计算注册码的过程实际上就是产生数字签名的过程:

(1) 生成密钥对:公钥为e,模数为n,私钥为d。

(2) 将ID作为消息输入m,使用MD5算法生成ID的报文摘要 $g = MD5(m)$。g的长度为160位。

(3) 用私钥d对g进行加密,得到签名:$s = g^d \bmod n$。s就是与ID对应的注册码。

(4) 将s作为注册码提供给用户。

验证注册码的过程就是对数字签名进行校验的过程:

① 用户输入注册码s。公钥e和模数n以常数的形式存在于校验模块中。

② 计算:$g' = s^e \bmod n$。

③ 计算:$g = MD5(m)$。

④ 比较g和g'。如果相等则证实s是m的正确数字签名,即可确认用户输入的注册码是由软件开发者签发的。

这里使用的是MD5报文摘要算法,也可以根据需要选择SHA－1等其他报文摘要算法。

2.2 注册码的Base－24编码表示法

由于 $s = g^d \bmod n$,所以注册码的长度取决于n的长度。n的位数越多越安全,但注册码也越长。

注册码可能以电话、传真、E－mail等方法传送给用户。为避免传输过程中出错,注册码一般以字符串的形式提供,其中仅包括数字和大写字母。对容易混淆的字符,如数字0和字母O,数字8和字母D,数字2和字母Z,字母U和V等,只取其中的一个。这里使用Base－24编码方案,每个字符位有24种可能的取值,如表1所示:

表1 Base－24编码方案

字符	B	C	D	F	G	H	J	K	M	P	Q	R	T	V	W	X	Y	2	3	4	6	7	8	9
取值	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23

其中B代表0,C代表1,…,9代表23。将注册码s转化为24进制的形式,每个数位取值为0到23之间,再将其替换为表1中的字符。n的位数为l,则注册码字符串的长度k为:

$$k = l \div \log_2 24 = l \div 4.585$$

k为整数,结果向上取整。当n取1024时,k = 224。

2.3 带日期限制和软件序列号的注册码

某些情况下,例如在软件的测试阶段和用于演示培训等,需要一个带时间限制的注册码。可以在<ID,注册码>二元组中再加入一个时间参数,扩展为<ID,截止时间,注册码>的形式,表示该注册码在该时间终止。

在计算这种形式的注册码时,取截止时间的年月日构成一个8位十进制数,年份取4位,月和日分别取2位,如"20040225"。将ID和截止时间合并后作为消息输入m。对不带时间限制的注册码,截止时间设为"99999999"。

此外,还可以为每一个软件拷贝设置惟一的一个序列号,序列号也作为消息输入m的一部分。序列号是一个数字,随软件提供给客户。对序列号进行签名后得到产品码

(product key)。在软件的安装过程中,可以要求用户输入产品码,验证产品码中的序列号和签名是否匹配,签名不匹配时终止安装。使用产品码不能完全达到保护软件的目的,同一个产品码可以被用于多台计算机上。Windows XP 要求用户提供产品码作为申请注册码的依据之一,同一个产品码只能申请到一个注册码。签名必须达到一定的长度。Windows XP 的产品码仅包括 55 位的签名,在一台普通的 P4 电脑上,使用暴力穷举法能够只需一天就能计算出到数十个有效的产品码。

2.4 计算机 ID 的唯一性

识别码对每一个计算机应是惟一的,软件使用识别码来区分各个计算机。当计算机的某些部件被更换或升级时,识别码可能会改变。如果任何一个小的硬件变动度要求用户重新申请注册码,无疑会加重用户和软件开发者双方的负担。因此,识别码的设计应该能容忍一部分变化。

以 Windows XP 为例,它的识别码共 64 位,包括了 10 项硬件信息,如 CPU 序列号、RAM 容量、硬盘卷标 ID、网卡 MAC 地址等。如果仅有 3 项(或更少)的变化,则认为是正常的硬件升级或维修[3]。硬件部件的信息以哈希值的形式构成识别码,例如 CPU 序列号的哈希值在识别码中占 6 位。

软件程序的注册码校验模块<ID,注册码>中得到 ID 后,与本机的识别码相比较。如果完全匹配或者只有很小的硬件变化,则匹配成功。

3 结论

本文介绍了一种使用公开密码体制数字签名算法来生成和校验软件注册码的方法。与依赖于硬件的方案相比,它具有兼容性好、易于实现、适应范围广等特点。应用成熟的数字签名技术来生成注册码,可以充分保证其安全性。此外,还可以结合数字水印、功能隐藏等其他技术来保护注册码校验模块,以避免该模块被绕开或替换。

参考文献

1 Gareth Cronin. A Taxonomy of Methods for Software Piracy Prevention.

2 吴世忠、祝世雄、张文政,应用密码学:协议、算法与 C 源程序[M],机械工业出版社,2000. 24-26。

3 Microsoft Corporation. Windows XP Product Activation, http://www.microsoft.com/windowsxp/pro/evaluation/overviews/activation.asp, 2001 June