基于位承诺的数字证书敏感信息保护

蔡国明[1] 王亚弟[1] 汪 淼[2] 徐开勇[1]
(解放军信息工程大学电子技术学院 郑州 450004)[1] (解放军信息工程大学理学院 郑州 450001)[2]

摘 要 目前数字证书缺少敏感信息保护机制,容易泄露用户的隐私信息。针对数字证书的敏感信息泄露问题,采用基于单向函数的位承诺协议实现数字证书中敏感信息的隐藏和选择性披露,详细讨论了基于位承诺的证书申请、证书颁发及证书查询验证过程,并分析了位承诺协议的安全性和效率。

关键词 位承诺,数字证书,敏感信息保护,选择性披露
中图法分类号 TP309.2 **文献标识码** A

Sensitive Content Protection in Digital Certificate Based on Bit Commitment

CAI Guo-ming[1] WANG Ya-di[1] WANG Miao[2] XU Kai-yong[1]
(Institute of Electronic Technology, Information Engineering University, Zhengzhou 450004, China)[1]
(Institute of Science, Information Engineering University, Zhengzhou 450001, China)[2]

Abstract The absence of sensitive content protection in digital certificate makes the holder's private information to be easily disclosed. The protection and selective disclosure of sensitive content in digital certificate can be achieved by using bit commitment protocol based on one-way hash function. This paper described the process of certificated application, certificated distribution and certificate access, and finally analyzed the security and efficiency.

Keywords Bit commitment, Digital certificate, Sensitive content protection, Selective disclosure

1 引言

数字证书是各类实体(持卡人/个人、商户/企业、网关/银行等)在网上进行信息交流及商务活动的身份证明,等同于现实世界中的身份证。数字证书的格式一般采用 X.509 国际标准,证书的内容通常包括版本号、序列号、主体名称、有效期和权威机构的签名等。目前最广泛使用的数字证书标准是 X.509v3,它允许对证书的内容进行扩展[1],支持在一个数字证书内将额外的特征信息,如主体单位、年龄、职务级别和社会安全号码等集成到证书中,并同证书的拥有者的公钥绑定在一起。

证书的这一扩展机制提高了证书的灵活性,但也给证书的使用带来了安全问题,最突出的问题是用户隐私泄露[2]。目前用户披露证书的方式有两种:一种是由用户自身将自己的证书提交给对方,另一种是对方可以通过目录服务器查询得到。无论采用哪种方式,对方一旦得到用户的证书后,就可以察看证书的所有内容。X.509v3 证书本身并没有机制来保护证书拥有者的隐私,存储于证书内部的重要特征属性和其他内容一样是公开的。这就意味着,如果一个证书提交给某人,那么证书拥有者的其他信息也跟着暴露了。这样恶意攻击者就可以通过收集用户证书来达到收集证书拥有者隐私的目的。

目前对于证书隐私的保护研究甚少。加密证书[3]是一种选择性的揭示 X.509v3 证书扩展项的方法,能够保护证书隐私不被泄露。它使用 CA 公钥对证书中的扩展项进行加密,以隐藏证书拥有者的敏感信息。但是加密证书易遭受字典攻击,而且计算代价大,实用性不强。隐秘证书也可以实现对关键信息的可选择性揭示,但是它采用的是专用的公钥加密系统,不具备通用性,而且有专利障碍,使用成本较高。

位承诺[4,5,7]允许一个人事先向别人提交一个数值而不用立即揭露该数值所表示的意义,他可以在某个时间以后才揭示它的意义。文献[6]提出了一种利用位承诺和隐秘属性的证书敏感信息保护方案,但是它采用将隐私信息之后直接串接一段随机位串作为预映射值,可能会泄露用户的隐私信息。本文提出的数字证书敏感信息保护技术对敏感信息进行了保护,不会泄露用户的敏感信息,而且计算代价在可承受范围之内。

本文提出了利用位承诺思想的数字证书敏感信息保护技术,既可以阻止恶意用户察看证书主体的隐私信息,还可以对合法用户揭示隐秘信息并确保信息的可信性。分析结果表明,该技术安全性高、计算代价低。

2 位承诺

位承诺是许多密码学协议的基本组成元素,发送者 S 向接收者 R 发送关于位 b 的承诺 c,使得接收者 R 无法从承诺 c 中得到任何有关位 b 的信息,同时协议结束时,R 可以得到

到稿日期:2008-07-01 返修日期:2008-09-26 本文受 863 国家重点基金项目(2008AA01Z404)资助。
蔡国明 博士研究生,研究方向为密码学与信息安全,E-mail:cgmcai817@sohu.com;王亚弟 教授,博士生导师,研究方向为密码学与信息安全;汪 淼 讲师,研究方向为信息安全;徐开勇 研究员,研究方向为信息安全。

位 b 并验证开始得到的 c 的确是 b 的承诺。位承诺可以当作将信息映射到集合{0,1}上的操作，其特点有点类似于信封或者签名，主要有以下两个特性：

① 一旦发送者声明消息是由他发出的，他就无法修改消息的内容，即发送者 S 可以改变初始承诺位的概率小于 $\mu(x)$；

② 除非发送者解开承诺，否则没有人能够读到消息的内容，即接收者 R 可以猜到初始承诺位的概率等于 $1/2+\mu(x)$。其中 $\mu(x)$为可忽略函数。第一个特性要求发送者使用的映射函数不可替代，第二个特性则保证接收者无法利用接收到的承诺位判断发送者发布的信息。

3 基于位承诺的数字证书敏感信息保护

3.1 问题描述

数字证书敏感信息保护的核心思想是证书敏感信息的隐藏和敏感信息的选择性揭示。证书敏感信息隐藏指主体的特征信息在证书中必须得到保护；敏感信息的选择性揭示又包含两方面内容：用户的选择性揭示和内容的选择性揭示。用户的选择性揭示指证书主体只向那些经过授权的个体出示证书，内容的选择性揭示指证书主体可以出示证书所有内容也可以是部分内容。

利用位承诺可以较好地解决数字证书敏感信息保护问题。位承诺允许一个人事先向别人承诺一个预测，但直到某个时间以后才揭示他的预测。因此，当证书中的敏感信息被用这些信息的承诺取代时，证书中的敏感信息得到了隐藏。

位承诺一般包括两个协议，一个是承诺协议，另一个是披露协议。在承诺协议中，Alice 通过一定的方式向 Bob 承诺她拥有一个位 b(某个信息)。在披露协议中，Alice 告诉 b 给 Bob，Bob 检查位 b 是否与 Alice 原先承诺的一致。

一个位承诺协议必须满足的条件是：

①机密性。即从证书的隐蔽字段推断出原始的特征信息，在计算上不可能。

②不可否认性。证书主体承诺的原始特征信息必须与披露的原始特征信息相吻合。

为了实现证书敏感信息的隐藏和敏感信息的选择性揭示，我们引入位承诺协议，同时对位承诺协议进行了安全性增强。增强的位承诺协议具备以下几个特性。

①机密性。即从证书的隐蔽字段推断出原始的特征信息，在计算上不可能；另外，隐蔽字段对应的原始特征信息在传输过程中不能被泄露。

②不可否认性。证书主体承诺的原始特征信息必须与披露的原始特征信息相吻合。

③访问控制。只有授权的用户才能访问证书的原始特征信息。原始特征信息可以部分披露，也可以完全披露。访问控制策略由用户自行定义。

④性能。产生、签署以及验证证书上的隐蔽信息所付出的额外的计算是能够承受的。

位承诺协议可以采用对称密码、单向函数和伪随机序列发生器来实现[7]。因为数字证书应用场景一般采用的是非对称密码系统，所以本文采用单向函数来实现位承诺协议。

3.2 证书申请过程(承诺协议)

Alice 提交证书申请过程如图 1 所示。

①Alice 产生一个长度为 128 比特的随机位串 r；

②Alice 用该随机位串 r 和欲添加进数字证书的隐蔽特征信息 v_i 产生各字段的预映射值 p_i：

$p_1=v_1\oplus r$；

$p_2=v_2\oplus r$；

……

$p_n=v_n\oplus r$。

如果直接将随机位串 r 串接到特征信息 v_i 的后面，可能会泄露 v_i 的值。例如攻击者从用户 Alice 的 Email 地址预映射值 Alice@sohu.com$*r$ 中仍然可能猜测出用户的真实邮件地址为 Alice@sohu.com。将随机位串和原始值异或后能够保护原始值不被泄露。

③Alice 通过调用一个强的安全的单向散列函数 H，产生各预映射 p_i 对应的散列值：

$a_1=H(p_1)$；

$a_2=H(p_2)$；

……

$a_n=H(p_n)$。

a_i 即对应于 v_i 的隐蔽特征信息被放到原来 v_i 所在的位置，并作为证书申请的一部分，一起提交给证书签发中心(Certificate Authority，CA)。

最后 Alice 提交的证书申请内容包括：主体名称、主体标识符、有效期、随机数 r 和各种敏感信息的预映射值 p_1，p_2，…，p_n 及其对应的散列值 a_1，a_2，…，a_n。

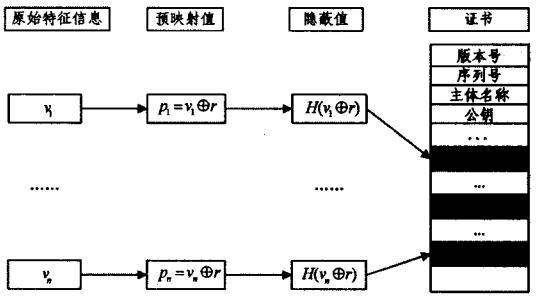

图 1 证书申请过程

3.3 证书颁发过程

CA 收到用户的证书申请后，首先验证证书申请中的散列值与用户所承诺的属性值是否一致，然后再签署证书。

①CA 计算各预映射值 p_i 的散列值 a_i；

②CA 比较计算出的 a_i 与收到的证书申请中的 a_i 是否一致；

③如果一致，则 CA 签署该证书。此时证书中敏感字段的值 v_1，v_2，…，v_n 就不是明文，而是以一个个散列值 a_1，a_2，…，a_n 代替。

3.4 查询验证过程(披露协议)

假设用户 Bob 想查询用户 Alice 的证书中的隐蔽字段的值，此时他要做的工作包括两部分：一是获取隐蔽字段的值，二是验证隐蔽字段的值是真实可信的。查询验证过程如图 2 所示。

查询验证协议如下：

①Bob 向目录服务器发起证书查询请求，此时一般要提交待查证书的序列号。

(下转第 100 页)

[J]. Communications of the ACM, 2006, 49(2): 76-80
[2] 王玲,钱华林. 计算机取证技术及其发展趋势[J]. 软件学报, 2003, 14 (9): 1635-1644
[3] [美]Steel C. Windows 取证:企业计算机调查指南[M]. 吴渝,唐红,陈龙,译. 北京:科学出版社, 2007
[4] Kornblum J. Identifying Almost Identical Files Using Context Triggered Piecewise Hashing[J]. Digital Investigation, 2006, 3 (s1): 91-97
[5] Chen Long, Wang Guoyin. An Efficient Piecewise Hashing Method for Computer Forensics[C]// International Workshop on Knowledge Discovery and Data Mining. Adelaide, Australia, January 2008: 635-638
[6] Roussev V, Chen Yixin, Bourg T, et al. md5bloom: Forensic File System Hashing Revisited[J]. Digital Investigation, 2006, 3 (s1): 82-90
[7] 陈龙,王国胤. 一种细粒度数据完整性检验方法[J]. 软件学报, 2009, 20(4): 902-909
[8] 靳蕃,陈志. 组合编码原理及应用[M]. 上海:上海科学技术出版社, 1995
[9] Bose R. Information Theory Coding and Cryptography[M]. China Machine Press, 2003

(上接第 94 页)

②目录服务器首先根据证书序列号查询证书是否存在。如果证书存在,则查询该证书的策略,确定是否允许 Bob 访问其隐蔽字段;如果证书不存在,则返回查询失败(每个证书的策略由证书主体自行定义,目录服务器只进行策略检查)。

③如果策略允许 Bob 访问证书的隐蔽字段,则用 Bob 的公钥加密随机数 r,并将证书、预映射值 $p_1, p_2, \cdots, p_n$($p_i = v_i \oplus r$)和加密后的结果 $Ep_B(r)$传给 B。

④Bob 首先计算预映射值对应的散列值 $a_1, a_2, \cdots, a_n$($a_i = H(p_i)$),再将计算出的散列值与证书中存储的散列值进行对比。如果两者一致,则说明传递过来的预映射值是可信的。然后 Bob 用私钥脱密 $Ep_B(r)$得到随机数 r,最后将预映射值 $v_i \oplus r$ 和随机数 r 进行异或得到该隐秘字段所对应的原文 v_i。

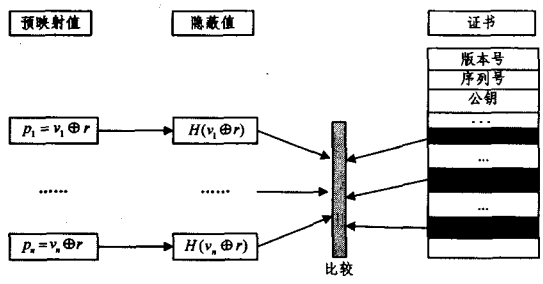

图 2 查询验证过程

4 安全性与效率分析

本文提出的位承诺协议的安全性取决于单向函数和公钥密码算法的强度。我们假定单向函数和使用的公钥密码算法是足够安全的,下面我们对第 3 节提出的位承诺协议的安全性和效率进行分析。

4.1 机密性

对证书中原始特征信息存储的机密性保证是基于单向散列函数和预映射值构造的强度。预映射值是将一个随机位串与特征属性值异或后产生的(如图 1 所示),那么攻击者在确定一个预映射值对应的原始信息时将需比较 2^{128} 次,这足以对抗穷尽攻击。

对证书中原始特征信息传输的机密性保证是基于公钥密码算法。采用公开密钥长度大于 2304 位能够保护随机数 r 足够安全[7]。由于 r 在查询过程中被加密,攻击者即使得到预映射值 $v_i \oplus r$ 也无法还原原始特征信息 v_i。

4.2 不可否认性

Alice 对隐蔽字段承诺的结果是对预映射值 $p_i = v_i \oplus r$ 进行单向函数变换得到的,只要单向函数足够安全,Alice 想伪造预映射值 p' 使得 $H(p_i') = H(p_i)$,在计算上不可行。Bob 通过计算预映射值所对应的散列值并与证书中的散列值进行比较,来验证隐蔽字段的值是否真实可信。

4.3 访问控制

X. 509V3 格式的数字证书提供了丰富的扩展字段,证书主体可以灵活定义自己的访问控制策略,来实现用户的选择性揭示和内容的选择性揭示。由于每个证书主体身份、职务级别的不同,其访问控制策略可能千差万别,因此由证书主体自行定义策略,而不是由签发中心集中定义,可以提高策略定义的灵活性。由目录服务器代替用户进行策略检查,可以实现对证书的统一访问控制。

4.4 性能

以证书内仅存储一个隐秘字段为例,忽略异或运算的计算量,证书申请过程比普通的证书申请过程仅多出一次哈希运算;忽略策略查询所花费的时间,证书查询验证过程比普通的证书查询验证过程多出一次加密运算和一次哈希运算。

结束语 本文针对证书隐私泄露问题,提出了基于位承诺的数字证书敏感信息保护方案。该方案能够阻止非法用户恶意收集他人证书,实现证书敏感信息的隐藏和敏感信息的选择性揭示。针对合法用户恶意收集他人证书及其隐私的问题,将是我们下一步研究的重点。

参 考 文 献

[1] Park J S, Sandhu R. Smart Certificates: Extending X. 509 for Secure Attribute Services on the Web[C]// 22th National Information Systems Security Conference. Crystal City, Virginia, October 1999
[2] Renfro S G. VeriSign CZAG: Privacy Leak in X. 509 Certificates [C]// 11th USENIX Security Symposium. San Francisco, California, August 2002
[3] Persiano P, Visconti I. User Privacy Issues Regarding Certificates and the TLS Protocol[C]// 7th ACM Conference of Computer and Communications Security. Athens, Greece, November 2000
[4] Naor M. Bit Commitment Using Pseudorandomness[C]// Advances in Crypology-Crypto 89, Lecture Notes in Computer Science. Vol. 435, New York: Spring-Verlage, 1990: 128-137
[5] Fischlin M, Fischlin R. Efficient Non - malleable Commitment Schemes. CRYPTO, 2000: 413-431
[6] 王永钊. 信任协商过程中证书上敏感信息的保护[D]. 天津:天津大学, 2006
[7] Schneier B. 应用密码学[M]. 吴世忠,等. 北京:机械工业出版社, 2000