公开可验证的零知识水印检测*

何永忠[1,2+], 武传坤[1], 冯登国[1]

[1](信息安全国家重点实验室(中国科学院 软件研究所),北京 100080)

[2](中国科学院 研究生院,北京 100049)

Publicly Verifiable Zero-Knowledge Watermark Detection

HE Yong-Zhong[1,2+], WU Chuan-Kun[1], FENG Deng-Guo[1]

[1](State Key Laboratory of Information Security (Institute of Software, The Chinese Academy of Sciences), Beijing 100080, China)

[2](Graduate School, The Chinese Academy of Sciences, Beijing 100049, China)

+ Corresponding author: Phn: +86-10-62528254 ext 803, E-mail: yzhe@is.iscas.ac.cn

Received 2004-01-05; Accepted 2004-06-10

He YZ, Wu CK, Feng DG. Publicly verifiable zero-knowledge watermark detection. *Journal of Software*, 2005,16(9):1606−1616. DOI: 10.1360/jos161606

Abstract: As the detection key in symmetric watermarking scheme can be used to forge or remove watermarks from digital works, it is required that the detection key be secret in watermark detection procedures. Based on zero-knowledge and proof of knowledge concepts and protocols in Cryptology, zero-knowledge watermark detection protocols can make the verifier believe the presence of a watermark in a disputed digital work while not compromising the detection key. The security requirements of a publicly verifiable zero-knowledge watermark detection scheme are outlined in this paper. Then a publicly verifiable commitment scheme and a zero-knowledge proof of knowledge protocol which proves knowing the discrete logarithm of a committed value are presented. Finally, using the above scheme and protocol as building blocks, a publicly verifiable zero-knowledge watermark detection protocol is proposed and its security considerations are addressed.

Key words: digital watermark; zero-knowledge; publicly verifiable

摘 要: 对称水印方案的水印检测密钥可以被用来伪造和移去水印,因此要求它在检测过程中也是保密的.零知识的水印检测方案利用密码学中零知识和知识证明的思想和算法,实现在水印检测时使得验证者确信水印存在性的同时又不泄漏水印检测密钥.提出了公开可验证的零知识水印检测的安全需求,给出一个公开可验证的承诺方案和一个证明知道被承诺值的离散对数的零知识知识证明协议.在此基础上提出了一个公开可验证的零知识水印方案,并讨论了它的安全性.

关键词: 数字水印;零知识;公开可验证

* Supported by the National Natural Science Foundation of China under Grant Nos.60025205, 90304007 (国家自然科学基金); the National High-Tech Research and Development Plan of China under Grant No.2004AA147070 (国家高技术研究发展计划(863))

作者简介: 何永忠(1969－),男,重庆人,博士生,主要研究领域为密码学,系统安全;武传坤(1964－),男,博士,研究员,博士生导师,主要研究领域为密码学;冯登国(1965－),男,博士,研究员,博士生导师,主要研究领域为密码学.

© 中国科学院软件研究所 http://www.jos.org.cn

中图法分类号: TP309 文献标识码: A

大部分数字水印方案水印的嵌入密钥与检测密钥相同,这类方案一般被称为对称水印方案.当发生版权争端时,版权所有者需要把水印检测密钥给仲裁者,而该密钥的泄漏可以被利用来伪造或者移去水印.一种解决方法是采用类似于密码学中公开密钥体制的水印方案,即水印检测密钥和嵌入密钥不同.这样就可以公开水印检测密钥用于水印的检测而不会泄漏嵌入密钥.但是这类方案不能抵抗敏感性攻击[1].另一种解决方法是采用零知识证明的思想,证明者能够证明自己的水印确实存在图像中,但是同时又不泄漏检测水印的密钥.Hirotsugu[2]的方案是最早提出的零知识水印检测方案.该方案将水印嵌入在图像的最低有效位平面,采用图同构零知识的知识证明协议证明该水印对应的图与该图像生成的图同构,由于这种水印很容易被移去,因此不适合作为版权证明方案.Gopalakrishnan[3]的方案基于 Cox 等人的扩频水印[4],证明者向验证者证明用 RSA 加密的水印存在于加密的图像中.由于该方案会泄漏相关值,所以不是零知识的.Craver[5]的方案通过对图像像素的位置进行置换的方法来证明水印的存在,由于置换后的像素值不变,攻击者可以据此猜测像素的本来位置,因此也不是零知识的.对上述各方案的详细分析可参见文献[6].

Adelsbach 等人[7,8]提出的基于扩频水印的检测协议可以严格证明是零知识的.文献[9]中的协议与 Adelsbach 等人的协议类似.但这两个协议都没有考虑下面的两个问题.一个问题是,协议中的承诺方案在参数设置时,需要可信第三方的参与或者证明者和验证者执行交互协议产生.这样得到的协议要么必须依赖可信第三方,要么只有证明者参与才能进行验证,因此不具有公开可验证性.另一个问题是,不能验证嵌入的水印是否合法.文献[10]提出的模糊攻击方法可以使得攻击者从任何一幅图像中找到伪水印并声称对该图像拥有版权,因此我们需要限制水印的性质,规定什么样的水印是可接受的、合法的,从而阻止模糊攻击.文献[5,10]采用的水印合法的条件是,合法的水印由计算不可逆的单向函数生成,版权证明者需要证明他知道生成该水印的单向函数的原象.

在密码学中,公开可验证性的目的是降低或者消除协议中某一方对协议其他参与方的可信度的依赖,减少对协议的验证者身份的限制,从而提高整个方案的安全性,并且当协议参与各方出现争端时,方便争端的仲裁.

本文把公开可验证的思想引入到水印的检测协议中来,在第 1 节中说明公开可验证的涵义并给出水印检测协议的安全目标.第 2 节提出公开可验证的承诺方案和一个证明知道被承诺值的离散对数的零知识知识证明协议.第 3 节利用前面提出的公开可验证的承诺方案,以及证明知道被承诺值的离散对数的零知识知识证明协议,给出一个基于 Adelsbach 协议的公开可验证的水印方案.第 4 节分析该方案的安全性.

1 水印验证协议及其安全性

水印协议一般由密钥生成、水印生成、水印嵌入、水印检测这 4 种算法构成.密钥生成算法生成一对水印的嵌入和检测密钥,对于对称水印方案来说,这两个密钥一般是相同的.水印生成算法生成满足一定性质的水印数据.水印嵌入算法在水印嵌入密钥的控制下,将水印嵌入隐藏到数字图像中.水印检测算法在水印检测密钥的控制下,检测水印是否存在.扩频水印方案是对称水印方案,并且水印嵌入/检测密钥即水印本身.

零知识的水印检测协议是一种零知识的知识证明协议.证明者向验证者证明图像中存在水印,但是不泄漏水印检测密钥.从零知识的知识证明相关概念中[11,12],我们提出公开可验证的零知识水印检测的安全目标.与以前的零知识水印检测相比,除了要求证明者能够证明图像中确实存在水印以外,我们还要求任何验证者(包括不可信的验证者)能够验证相关参数的合法性和水印的合法性,防止证明者或者可信第三方的欺骗,因此称为公开可验证的零知识水印检测协议.

定义 **1**(公开可验证的零知识水印检测协议). 一个零知识水印检测协议是公开可验证的,如果具有下列性质:版权所有者完成相关参数生成和水印嵌入后,任何人都可以作为验证者验证参数的选取是否合法、图像中是否存在证明者的水印,以及该水印是否合法;在参数生成和水印嵌入过程中可以不需要可信第三方和验证者的参与;如果转换为非交互式协议,在验证过程中可以不需要证明者的参与.公开可验证的零知识水印检测协议

© 中国科学院软件研究所 http://www.jos.org.cn

的安全性包括:

- 完全性.一个诚实的证明者总能让验证者相信他确实在图像中嵌入了水印,并且该水印以及参数的选择都是合法的.
- 合理性.一个不诚实的证明者,他能欺骗验证者相信他在图像中嵌入了合法水印的可能性可以忽略.
- 零知识性.验证协议执行后,验证者获得的关于水印嵌入密钥的信息与他验证前知道的一样.
- 有效性.存在一个多项式时间的知识提取器.

需要说明的是,当嵌入水印的图像被修改时,验证水印是否存在仍然需要证明者的参与.否则,如同公钥水印一样,会遭受敏感性攻击.

2 相关密码协议

本节我们简要介绍这几个基本的密码协议,包括公开可验证的承诺方案、证明知道被承诺值的离散对数零知识证明协议以及证明被承诺值是 1 或者–1 的零知识协议.我们提出的公开可验证水印验证协议将利用这几个基本的密码协议,其中公开可验证的承诺方案和证明知道被承诺值的离散对数零知识的知识证明协议是在本文中首次提出的.

2.1 公开可验证的承诺方案

Pedersen 在文献[13]中提出的承诺方案与文献[14]中提出的承诺方案是一种类型.这类承诺方案的一个问题是,在生成公开参数的时候要么需要一个可信第三方,要么需要承诺的验证者的参与.我们在 Pedersen 方案的基础上,提出一个任何验证者都可以公开验证的承诺方案.

协议 **1**(**Pedersen** 承诺方案[13]). 协议的公开参数为$\langle p, g, f\rangle$.

在群 Z_p^* 中的两个生成元 g, f ,其中 p 是素数.没有人知道 $\log_g f$.对一个数 $s\in Z_p$ 承诺.随机选择一个 $t\in_R Z_p^*$,计算: $com(s)=g^s f^t \bmod p$.计算结果 $com(s)$即为 s 的承诺值.

下面我们称 s 为被承诺值,$com(s)$简称承诺.Pedersen 证明了该承诺方案的安全性.一方面,承诺者找到对应一个承诺的另一个被承诺值在计算上不可行(称为计算上绑定的);另一方面,验证者不能以高于随机猜测成功的概率计算出正确的被承诺值(称为无条件隐藏的).在该方案中,要求任何人都不知道 $\log_g f$,否则,比如证明者知道 $\log_g f$,那么他就可以在承诺揭示时候更换被承诺值.为了避免这种情况,Pedersen 推荐了两种方法来生成 g 和 f,一种是由可信第三方来选取,另一种是承诺者和验证者运行一个抛硬币协议来选取.这两种方法都存在缺陷,前者需要可信的第三方的参与,后者要求证明者的参与,因此不能达到公开可验证的要求.

实际上,我们可以要求选择的生成元都是已知原象的单项哈希函数值来解决这个问题.为了说明该方法的可行性,先介绍两个数论中的相关引理.

引理 **1**[15]. 如果 p 是奇素数, $p-1$ 有 $p_1, p_2, ..., p_r$ 个不同的素因子. $g\in Z_p^*$ 是 Z_p^* 的生成元当且仅当 $g^{(p-1)/p_i} \bmod p\neq 1$,其中 i=1,…,r.

引理 **2**[16]. 如果 p 是奇素数, Z_p^* 的生成元的个数为 $\phi(p-1)$ 个. ϕ 是欧拉函数.

假设 $H(x)$是安全哈希函数,输出的二进制位长等于$|p|$(记号$|p|$表示 p 的二进制表示的长度,下同).如果$|H(x)|<|p|$,可以通过多个哈希拼接的方式构造一个哈希函数,使得$|H(x)|+1=|p|$.随机生成 u,v,使 $g=H(u), h=H(v)$,且 g,h 是生成元.文献[17]提出一种在给定区间中均匀生成素数 p 同时可以给出 $p-1$ 的因子分解的多项式时间算法,这样,根据上面的引理,可以容易地判定一个随机数是否是生成元,并且一个随机选择的数是生成元的可能性是可观的(noticeable)[18],所以,寻找这样的(u,g)和(v,h)在计算上是可行的.承诺者向验证者出示(u,g)和(v,h),验证它们是否确实是哈希函数两对不同的原象和值,并且 g,h 是生成元.容易证明,在 random oracle 模型下[19],求 $\log_g f$ 是困难的.

在我们提出的水印验证方案中,因为要计算承诺的平方根,因此我们要求 g 和 f 的阶为 q.这时,为了保证公开可验证性,参数的验证方法应相应改变.

© 中国科学院软件研究所 http://www.jos.org.cn

协议 **2**(公开可验证承诺方案 **PVC**(**publicly verifiable commitment**)). 协议的公开参数为$\langle p,q,u,g,f\rangle$.

(1) 参数生成:承诺者首先选取大素数 q,并且 $p=2q+1$ 也是素数.H 为安全哈希函数,对任意的 $x\in Z_p^*$, $|H(x)|+1=|p|$,并且 $H(x)\in Z_p^*$.选择一个随机数, $u\in_R Z_p^*$,令 $g=H(u)^{\frac{p-1}{q}}=H(u)^2 \bmod p$. $f=H(g)^2 \bmod p$,且 $g\neq 1, f\neq 1$.公开 p,q,u,g,f.

(2) 承诺生成:如果对 $s\in Z_q$ 承诺,随机选择一个 $t\in_R Z_q^*$,计算 $com(s)=g^s f^t \bmod p$.

(3) 承诺参数验证: $q=(p-1)/2$, p,q 都是大素数. $g=H(u)^2 \bmod p$, $f=H(g)^2 \bmod p$.

(4) 承诺揭示和验证:承诺者公开 s,t.验证者验证 $com(s)=g^s f^t \bmod p$.

定理 **1**. 公开可验证承诺方案 PVC 是无条件隐藏和计算上绑定的,并且 $\log_g h$ 是未知的.

证明:首先,易知 g 和 f 的阶都是 q.设 PVC 方案中由 g 生成的 q 阶子群为 G.设 a 为 Z_q^* 的生成元.构造从子群 G 到 Z_q^* 的同构映射 $\varphi: g^i\to a^i$, $i\in Z_{q-1}$.这样,PVC 的安全性等价于 Pedersen 承诺方案的安全性.根据 random oracle 模型,可以把 g 和 f 看作是随机数,由离散对数的难解性假设可知,求 $\log_g h$ 是困难的. □

2.2 证明知道被承诺值的离散对数的零知识知识证明协议

有关零知识协议和知识证明协议的定义参见文献[11,12].

协议 **3**(证明知道被承诺值的离散对数的零知识协议 **ZK-COM-DL**). 协议的公开参数为$\langle p,q,u,g,f,h\rangle$.

假设验证者 V 获得承诺 y,证明者 P 向验证者 V 证明他知道 x,t 满足 $y=g^{h^x} f^t \bmod p$, $|p|=k$ .h 是 Z_q^* 的生成元,其他公开参数与承诺方案 PVC 相同.协议如下:

(1) P: 随机选取 $r\in_R Z_q^*, s=t\cdot h^{r-x}$,计算 $c=g^{h^r} f^s \bmod p$,并把 c 发送给验证者.

(2) V: 随机选择 $b\in_R \{0,1\}$,发送 b 给证明者.

(3) P:

如果 b=0,发送 $z=r,s$ 给验证者;

如果 b=1,发送 $z=r-x,s$ 给验证者.

(4) V:

如果 b=0,验证 $c=g^{h^z} f^s \bmod p$,相等就接受,不成立就拒绝;

如果 b=1,验证 $c=y^{h^z} \bmod p$,相等就接受,不成立就拒绝.

定理 **2**. 协议 ZK-COM-DL 是证明知道被承诺数的离散对数的零知识知识证明协议,并且验证者错误接受的概率小于 $\frac{1}{2}+\frac{1}{Q(k)}$,其中 Q 为任意的多项式.

证明:证明一个协议是零知识知识证明协议包括 4 个方面:完全性,合理性,并存在一个多项式时间的知识提取器,以及构造一个与诚实证明者输出不可区分的模拟器.

(1) 完全性的含义是,如果证明者确实知道被承诺数的离散对数,那么协议正常运行完后,验证者接受.容易看出,当证明者知道 y 中的被承诺值的离散对数并且知道 t 时,总能通过第 4 步的验证,所以验证者总是接受.

(2) 合理性的含义是,如果证明者不知道 y 中的被承诺值的离散对数,则验证者拒绝.我们称这样的证明者为欺骗者.先假定欺骗者依照协议执行第 1 步.如果欺骗者猜测在第 3 步 b=0,那么他在第 1 步随机选择 r,s,验证者在第 4 步接受.如果 b=1,欺骗者必须求 z 和 s,使得:

$$g^{h^r} f^s = y^{h^z} \bmod p .$$

如果欺骗者不知道被承诺值的离散对数,也不知道被承诺值,那么,上面的等式可以变换为

$$g^{h^{r-z}} f^{s\cdot h^{-z}} = y \bmod p .$$

这个问题的难度不亚于对该承诺方案的攻击,所以成功的概率小于任何 $\frac{1}{Q(k)}$.

如果欺骗者知道被承诺值 a,即 $a=h^x \bmod q$,这时,他就需要求解 z,r,使 $a=h^{r-z} \bmod q$.由离散对数假设可知,

© 中国科学院软件研究所 http://www.jos.org.cn

成功的概率也小于任何 $\frac{1}{Q(k)}$.

如果欺骗者不知道 t,则他成功的概率小于求解 $\log_f g$.

另外,当欺骗者在第 1 步不依照协议执行时,可以进行类似的分析,结论也相同.这样,欺骗者成功的概率小于 $\frac{1}{2}+\frac{1}{Q(k)}$.

(3) 知识提取器 E 可以这样构造.第 1 轮和证明者运行协议,得到 $c,0,z_1=r,s$ 这 4 个消息.然后,输入消息 $c,1$ 给证明者问答函数,询问下一个消息,得到 $z_2=r-x,s$.这样可以计算 $x,x=z_1-z_2$.

(4) 模拟器的构造也比较简单.随机地选取 $b' \in_R \{0,1\}$, $r,s \in_R Z_q^*$,如果 $b'=0$,计算 $c=g^{h^r} f^s \bmod p$,否则,计算 $c=y^{h^r} \bmod p$,并把 c 发送给验证者.当 $b=b'$ 时,模拟器把 r 发送给验证者;否则失败.易知,失败的概率不大于 1/2.其他情况下,与任何验证者的交互,模拟器的输出与诚实的证明者的输出计算不可分. □

协议 ZK-COM-DL 每运行一轮,验证者以 1/2 的概率相信证明者确实知道被承诺值的离散对数,为了降低错误接受概率需要运行多轮.如果对效率要求较高,我们可以构造一个错误接受的概率可忽略的 4 轮的知识证明协议,但是不能证明该协议是零知识的.

2.3 证明被承诺值是1或者–1

在我们设计水印验证协议过程中,使用了承诺方案,但是没有揭开承诺步骤,因此需要向水印的验证者证明承诺的正确形式.下面简述我们使用的方案.该方案与文献[5]中的证明知道一组水印中某一个水印的离散对数的零知识协议相同.

该方案的目标是,证明者 P 要证明承诺 X 是对 1 或者–1 的承诺,但是验证者 V 无法以高于 1/2 的概率猜测被承诺值具体是 1 或者–1.协议中,证明者证明他知道$(gX,X/g)$中的一个承诺的以 f 为基的离散对数,但不会泄漏是哪一个承诺.

协议 **4**(证明 X 中被承诺值是 **1** 或者**–1** 的协议 **ZK-COM1-1**). 协议的公开参数为$\langle p,q,u,g,f\rangle$.X 是承诺.

(1) P: 随机选择 $r_1,r_2 \in_R Z_q^*$,计算 $c_1=gXf^{r_1} \bmod p, c_2=Xf^{r_2}/g \bmod p$,并将结果随机置换后发送给验证者.

(2) V: 随机选取 $b \in_R \{0,1\}$ 发送给证明者.

(3) P:

如果 b=0,则将 $\{r_1,r_2\}$ 发送给验证者;

如果 b=1,则将 c_1,c_2 中某一个的离散对数 r_3 发送给验证者.如果 X 的被承诺值是 1,则 $X=gf^t \bmod p$,那么 $r_3=r_2+t$.如果被承诺值是–1,则 $r_3=r_1+t$.

(4) V: 验证:

如果 b=0,验证 $c_1=gXf^{r_1} \bmod p, c_2=Xf^{r_2}/g \bmod p$;

如果 b=1,验证 c_1,c_2 中的一个与 f^{r_3} 相等.

3 安全的水印检测

本节我们将利用前面提出的公开可验证的承诺方案、证明知道被承诺值的离散对数的零知识知识证明协议以及 Adelsbach 等人[7]提出的零知识水印检测协议,构造一个公开可验证的水印检测协议.我们先简单介绍 Adelsbach 等人提出的零知识水印检测协议,然后给出公开可验证的水印检测协议.

3.1 Adelsbach等人的水印检测协议简述

Adelsbach 等人提出的零知识水印检测协议是以 Cox 等人的扩频水印方案为基础的.扩频水印方案有良好的健壮性,是一种广受重视的水印技术.虽然 Cox 等人描述的是基于数字图像的水印方案,然而对其他,如数据,声音、视频等也适用.为方便叙述起见,Adelsbach 等人的水印检测协议也是基于数字图像的.待嵌入的水印表示为 $WM=(WM_1,\ldots,WM_k)$,WM_i 为随机数实数,满足期望为 0,方差为 1 的正态分布 $N(0,1)$,k 为扩频因子.设图像 O 的

© 中国科学院软件研究所 http://www.jos.org.cn

DCT 变换 k 个最大交流分量的系数记为 $DCT(O,k)=(DCT(O)_1,...,DCT(O)_k)$.把水印 WM 嵌入图像 O 的方法是修改相应的 DCT 变换系数:

$$DCT(\bar{O})_i = DCT(O)_i \cdot (1+\alpha \cdot WM_i),$$

式中 α 为权重因子,在满足嵌入水印的健壮性和不可见性之间折衷选择.通过 DCT 逆变换得到嵌入了水印的图像 $\bar{O}$.

通过计算相关性值 $corr$ 来检测水印.设 δ 为检测阈值,如果 $corr \ge \delta$ 就表明检测到了水印.盲水印检测方案中(不需要原始图像 O)$corr$ 的计算如下:

$$corr = \frac{\langle DCT(\bar{O},k), WM\rangle}{\sqrt{\langle DCT(\bar{O},k), DCT(\bar{O},k)\rangle}},$$

记号$\langle X,Y\rangle$表示向量 X 和 Y 的内积.

令 $A=\langle DCT(\bar{O},k), WM\rangle$, $B=\langle DCT(\bar{O},k), DCT(\bar{O},k)\rangle\delta^2$, $C=A^2-B$.那么,检测条件可变换为 $C\ge 0, A\ge 0$.

该水印检测方案的问题是,一旦执行后,水印 WM 就不是秘密的了,知道水印的人也可以很容易地从 $\bar{O}$ 除去水印 WM.而 Adelsbach 等人的水印检测方案的目标就是在水印检测协议执行后,水印 WM 依然是保密的.在这个方案中,使用了 Pedersen 的承诺方案,以及零知识证明被承诺数大于等于 0 的协议,记为 ZK-GE0($com(x)$),以及零知识证明一个被承诺数 x 是另两个被承诺数 y,z 之积的协议,记为 ZK-MUL($com(x),com(y),com(z)$).参见文献[7,20].需要指出的是,采用我们的承诺方案与采用文献[14]的承诺方案不会影响 ZK-GE0,ZK-MUL 等协议的执行.

协议 5(扩频水印的零知识检测协议 **ZK-Adelsbach**).

水印的证明者为 P,水印的验证者为 V.P 向 V 证明图像 $\bar{O}$ 中存在水印 $com(WM)$.记号 $com(WM)$表示对水印 WM 各个分量的承诺, $com(WM)=(com(WM_1),...,com(WM_k))$, com 为前文中的 Pedersen 承诺.

(1) P 和 V 分别计算 $DCT(\bar{O},k)$.

(2) P 和 V 分别计算 B.同时,P 计算 $com(B)$然后发送给 V,并向 V 揭示承诺.V 验证 $com(B)$中被承诺值确实是 B.

(3) P 和 V 分别计算:

$$com(A)=\prod_{i=1}^{k} com(wm_i)^{DCT(\bar{O})_i}.$$

(4) P 和 V 执行零知识协议 ZK-GE0($com(A)$)证明 $com(A)$中的被承诺数 $A\ge 0$.

(5) P 计算 A^2,并和 V 执行 ZK-MUL($com(A^2)$,$com(A),com(A)$),向 V 证明 $com(A^2)$ 中的被承诺数为 A^2.

(6) P 和 V 分别计算:

$$com(C)=\frac{com(A^2)}{com(B)}.$$

(7) 执行 ZK-GE0($com(C)$),P 向 V 证明 $com(C)$中的被承诺数 $C\ge 0$.

(8) 如果 V 接受上面的所有零知识证明,那么承认图像 $\bar{O}$ 中存在水印 $com(WM)$.

3.2 公开可验证的水印方案

在文献[1,10]中提出一种模糊攻击法(也称为可逆攻击),即攻击者可以通过对图像的分析,构造一个伪水印满足相关值范围的检测.为了防止这类攻击,可以要求合法的水印是以原始图像为种子在单向函数作用下生成.对于零知识的水印验证协议,在水印的验证过程中不能泄漏水印和原始图像的信息,因此,我们的方案公开原始图像 DCT 的承诺,然后计算该承诺的哈希函数值,再选取一个随机数 ω,将原始图像承诺的哈希值乘以 $h^{\omega} \bmod q$ 得到水印 WM.由于随机数 ω 和 h^{ω} 是秘密的(只公开 h^{ω} 的承诺),所以水印也是秘密的.水印的生成函数为

$$WM = h^{\omega}\cdot H(g^{DCT(O)_1} f^{r_1} \bmod p, g^{DCT(O)_i} f^{r_i} \bmod p,..., g^{DCT(O)_k} f^{r_k} \bmod p) \bmod q.$$

我们这里分析一下该生成函数的单向性,即已知水印 WM 求解 $\omega, DCT(O)_i, r_i$ ($i=1,2,...,k$)的困难性.由于 H

© 中国科学院软件研究所 http://www.jos.org.cn

是安全哈希函数,所以当 ω 确定时,攻击者成功计算 $DCT(O)_i, r_i$ 的概率小于哈希函数求逆的概率,因此在计算上不可行.如果所有 k 对 $DCT(O)_i, r_i$ 确定了,那么由于离散对数的难解性,求 $h^{\omega} \bmod q$ 的离散对数 ω 是不可行的.还应该考虑的一种攻击是,当已知一组解 $\omega, DCT(O)_i, r_i$,求另一组解 $DCT(\hat{O})_i, \hat{r}_i$ 以及 $\hat{\omega}$ 是否困难.如果 $\hat{\omega} \neq \omega$,求 $DCT(\hat{O})_i, \hat{r}_i$ 的成功概率小于哈希函数求逆;否则,如果 $\hat{\omega} = \omega$,并且攻击者希望避免求哈希的逆,则需要计算 $g^{DCT(\hat{O})_1} f^{\hat{r}_1}$ 使得:

$$g^{DCT(O)_1} f^{r_1} = g^{DCT(\hat{O})_1} f^{\hat{r}_1} \bmod p .$$

根据承诺的计算绑定性,该计算不可行.因此,该水印生成函数是单向的.

在本方案中,水印的每个分量对应于 WM 二进制表示的一位.另外,我们计算并公开对水印 WM 的各个分量 WM_i 的承诺 $com(WM_i)$,以及对 h^{ω} 的承诺 $m = g^{h^{\omega}} f^{t} \bmod p$.为了验证水印的合法性,证明者需要证明他知道 m 的被承诺值的离散对数,并且 m 的离散对数与原始图像承诺的哈希值之积就是以 $com(WM_i)$中被承诺值为分量的水印 WM.不失一般性,我们定义水印的各个分量在$\{-1,1\}$上均匀取值.该方法可以容易地扩展到定义域为其他整数范围或者实数区间.我们要求水印分量在$\{-1,1\}$上取值而不是在$\{0,1\}$上取值,目的是为了不改变嵌入水印后图像 DCT 系数的平均值.

文献[21]提出对通过单向函数生成合法水印的方案的一种在计算上可行的穷举攻击方法,即攻击者随机猜测随机种子 ω,检测生成的水印能否通过水印检测,反复尝试,直到找到一个能够通过水印检测的随机种子.该文献提出避免这种攻击的方法是,同时要求证明者证明未嵌入水印的原始图像 O 与同一水印的相关值很低.我们的方案中也需要这一步骤,由于原始图像和水印都要求是保密的,所以我们可以采用类似于 ZK-Adelsbach 的零知识协议.我们把这个零知识协议记为 ZK-Negative-Adelsbach,它要求在公开水印承诺以及原始图像 DCT 的承诺后,证明者证明该水印与原始图像的相关值很低.该协议与 ZK-Adelsbach 的协议不同之处有:验证者只知道 $com(DCT(O)_i)$,不知道 $DCT(O)_i$,所以在计算 $com(A)$的时候需要执行 ZK-MUL($com(A_i)$,$com(WM_i)$, $com(DCT(O)_i)$)协议,然后验证 $com(A)$等于所有 $com(A_i)$之积.计算 $com(B)$的方法类似.另外,不需要在第(4)步证明 $A \geq 0$.在第(7)步变更为证明 $com(C)$中 $C \leq 0$.该协议的具体步骤这里不再详细列出.

需要说明的是,证明者还必须证明该原始图像 O 与图像 $\bar{O}$ 是相似的,否则攻击者可以容易找到另外一个图像通过 ZK-Negative-Adelsbach 协议验证.我们可以通过零知识证明 k 对 $com(DCT(O)_i)$, $com(DCT(\bar{O})_i)$ 中被承诺值之差的平方和足够小(小于阈值 λ 乘以 k 个 $DCT(\bar{O})_i$ 的平方和)来表明两个图像是相似的.该协议可以容易地使用文献[20]中各个协议来构建,我们把它记为 ZK-Sim($com(DCT(O,k))$, $com(DCT(\bar{O},k))$, λ).

另外,为了增强 ZK-Adelsbach 的安全性,对于水印和图像的相关值的判定,我们采用一种更为严格的判定规则:

$$-\delta \leq \frac{\langle DCT(\bar{O},k), WM\rangle - \alpha \cdot \sum_{i=1}^{k} DCT(\bar{O})_i}{\sqrt{\langle DCT(\bar{O},k), DCT(\bar{O},k)\rangle}} \leq \delta .$$

令 $A = \langle DCT(\bar{O},k), WM\rangle - \alpha \cdot \sum_{i=1}^{k} DCT(\bar{O})_i$, $B = \langle DCT(\bar{O},k), DCT(\bar{O},k)\rangle \delta^2$, 那么判定规则变为 $C = B - A^2 \geq 0$.这样,ZK-Adelsbach 协议需要作相应的改变,得到的新协议称为 ZK-Tight-Adelsbach 协议,具体的步骤这里也不再列出.

下面具体描述我们提出的公开可验证的水印协议.在协议中还用到了离散对数的零知识证明协议,记为 ZK-DL.这是一个常见的零知识协议,这里不作介绍.

协议 **6**(公开可验证的零知识水印检测方案 **PZW**).

参数和承诺生成:设安全参数(素数的二进制长度)为 l,扩频因子为 k,水印检测阈值 δ,以及某个图像与原始图像相似判定的阈值 λ,权重因子 α.为方便描述,假设 l=k.

公开参数的生成:证明者首先随机选取大素数 q,同时给出 $q-1$ 的因子分解,并且使 $p = 2q+1$ 也是素数.H 为安全哈希函数,对任意的 $x \in Z_p^*$, $|H(x)| = |l|$, $|p| = l+1$, 并且 $H(x) \in Z_p^*$.选择一个随机数, $u \in_R Z_p^*$, 令 $g = H(u)^2 \bmod p \neq 1$, $f = H(g)^2 \bmod p \neq 1$.随机选择 $h \in_R Z_q^*$ 且为 Z_q^* 的生成元.

© 中国科学院软件研究所 http://www.jos.org.cn

对原始图像 DCT 的承诺:对所有的 $i,1\le i\le k$, $com(DCT(O)_i)=g^{DCT(O)_i}f^{r_i} \bmod p$,并计算:

$$\beta=H(com(DCT(O)_1),com(DCT(O)_2),...,com(DCT(O)_k)) \bmod q.$$

秘密参数:随机选取 $\omega\in Z_q^*$,计算水印 $WM=h^{\omega}\cdot\beta \bmod q$,设 WM 的二进制形式为 $b_0,b_1,...,b_{k-1}$,高位在前.第 i 位 b_i 如果为 1,WM_i=1,如果为 0,WM_i=−1 $(1\le i\le k)$.

水印承诺生成:计算 $m=com(h^{\omega})=g^{h^{\omega}}f^{t} \bmod p$.对所有的 $i,1\le i\le k$,计算 $com(WM_i)=g^{WM_i}f^{t_i} \bmod p$.

水印嵌入:同 Cox 等人的方案. $DCT(O')_i=DCT(O)_i\cdot(1+\alpha\cdot WM_i)$.

公开可验证的水印检测协议:验证者获知 $\overline{O}$,m, $com(DCT(O)_i)$, $i=1,2,...,k$,水印检测阈值 δ ,图像相似判定的阈值 λ ,权重因子 α 以及其他公开参数.

(1) 公开参数的验证:验证 $q-1$ 的因子分解是否正确(包括各个不同的因子是否是素数), $q=(p-1)/2$,p,q 都是大素数. $g=H(u)^2 \bmod p\ne 1$, $f=H(g)^2 \bmod p\ne 1$,并且验证 h 是 Z_p^* 生成元(根据引理 1).

(2) 证明者和验证者执行零知识证明协议 ZK-COM-DL,证明知道 m 的被承诺值的离散对数 ω .即 $m=g^{h^{\omega}}f^{t} \bmod p$.

(3) 证明者将水印分量的承诺 $com(WM_i)=g^{WM_i}f^{t_i} \bmod p$, $i=1,2,...,k$,发送给验证者,然后证明者和验证者执行协议 ZK-COM1-1 证明所有 k 个承诺 $com(WM_i)$ 中的被承诺数是 1 或者−1.

(4) 验证者计算 β ,证明者和验证者执行离散对数零知识证明协议 ZK-DL,证明知道 T,使得:

$$f^T=m^{\beta}\prod_{i=1}^{k}(g\cdot com(WM_i))^{\frac{q+1}{2}\cdot 2^{k-i}} \bmod p .$$

(5) 证明者和验证者执行 ZK-Negative-Adelsbach 协议,证明原始图像中不存在水印 WM.

(6) 证明者和验证者执行 ZK-Sim $(com(DCT(O,k)),com(DCT(\overline{O},k)),\lambda)$.证明原始图像 O 和图像 $\overline{O}$ 相似.

(7) 证明者和验证者执行 ZK-Tight-Adelsbach 协议,证明图像 $\overline{O}$ 中存在水印 WM.

(8) 对以上步骤(2)~步骤(7),根据错误概率的需要反复执行多次,如果都通过,验证者接受,即认可 $\overline{O}$ 中存在证明者的水印(如果每个协议一次执行的错误概率小于 1/2,则 n 次执行验证者错误接受的概率小于 $1/2^n$).否则拒绝.

4 PZW 的安全性分析

在公开可验证的水印检测协议 PZW 中,第(1)步是参数合法性检测.第(2)步~第(4)步是为了防止攻击者采用模糊攻击从图像中提取伪水印欺骗验证者的,我们称其为水印的合法性验证.对应地,第(5)步~第(7)步称为水印存在性验证.在下面的分析中,我们称一个水印存在于图像中,如果水印与图像的相关值较大,并且与该图像对应的原始图像的相关值很小.

4.1 协议的完全性

当证明者确实是待检图像的水印嵌入者时,即水印存在图像中.他按要求选择参数,知道 m 的被承诺数的离散对数 ω ,因此,根据各子协议的完全性易知,在第(1)~(3)步以及第(5)~(7)步,验证者总是接受.在第(4)步,因为 g 和 f 的阶都是 q,且 $2\cdot\frac{q+1}{2} \bmod q=1$,所以,

$$g^{1/2} \bmod p=g^{1/2 \bmod q} \bmod p=g^{(q+1)/2} \bmod p .$$

对于 f 也有同样的性质,所以,

$$\begin{aligned}\prod_{i=1}^{k}(g\cdot com(WM_i))^{\frac{q+1}{2}2^{k-i}} \bmod p&=\prod_{i=1}^{k}(g\cdot g^{WM_i}f^{t_i})^{\frac{q+1}{2}2^{k-i}} \bmod p\\&=\prod_{i=1}^{k}g^{(1+WM_i)\frac{q+1}{2}2^{k-i} \bmod q}f^{t_i\frac{q+1}{2}2^{k-i} \bmod q} \bmod p\\&=\prod_{i=1}^{k}g^{(1+WM_i)\frac{1}{2}2^{k-i} \bmod q}f^{t_i\frac{q+1}{2}2^{k-i} \bmod q} \bmod p\end{aligned}$$

© 中国科学院软件研究所 http://www.jos.org.cn

$$= g^{\sum_{i=i}^{k}(1+WM_i)\frac{1}{2}2^{k-i} \bmod q} f^{\sum_{i=1}^{k} t_i \frac{q+1}{2} 2^{k-i} \bmod q} \bmod p$$
$$= g^{WM} f^{T+t\cdot\beta} \bmod p$$
$$= m^{\beta} \cdot f^{T} \bmod p,$$

其中 $T = -t\cdot\beta + \sum_{i=1}^{k} t_i \frac{q+1}{2} 2^{k-i} \bmod q$.所以,诚实的证明者知道 T,在第(4)步验证者总是接受.

4.2 协议的合理性

首先,根据参数和水印合法性验证协议的合理性,当参数或者水印不合法时,不能通过第(1)步~第(4)步的验证.下面我们说明,如果由 m 的离散对数和原始图像单向生成的水印 $WM = (WM_1, WM_2, ..., WM_k)$ 不存在于图像 $\bar{O}$ 中,证明者不能在每一个验证步骤都使验证者接受.协议中第(2)~(4)步保证了所有 $com(WM_i)$ 中被承诺数构成的序列就是 WM 对应的水印分量序列 $WM = (WM_1, WM_2, ..., WM_k)$.根据协议 ZK-Tight-Adelsbach, ZK-Negative-Adelsbach 以及 ZK-Range 的合理性,由于 $com(WM_i)$ 对应的水印不在图像 $\bar{O}$ 中,那么它不能同时通过第(5)~(7)步的验证.需要指出的是,证明者使用 $WM_i \pm q$ 代替 WM_i 计算不影响验证结果.

4.3 零知识性和有效性

零知识理论中的一个基本事实是,由一组顺序执行的零知识协议组成协议也是零知识的.由于水印检测的各步都是零知识的,所以整个验证协议是零知识的.另外,由于水印检测协议第(2)步是一个知识证明协议,因此,存在一个知识提取器在多项式时间可输出秘密 ω,从而求得水印 WM,所以水印检测协议满足有效性.

4.4 模糊攻击和穷举攻击

首先说明,由于我们的水印是从原始图像的 DCT 以及随机种子单向生成的,因此可以抵抗模糊攻击.如果一个证明者(此时为攻击者)使用模糊攻击,从图像中找到一个伪水印 WM',从而通过水印存在性验证,却不能同时通过第(2)~(4)步的水印合法性验证.因为,当 WM' 及承诺 $com(WM_i')$ 已经确定,所以证明者必须找到伪原始图像 $\hat{O}$,以及 $\hat{\omega}$ 和 $\hat{m}$,使得:

$$\hat{m}^{\beta} = g^{WM'} f^{t\cdot\beta} \bmod p,$$
$$\hat{m} = g^{h^{\hat{\omega}}} f^{t} \bmod p,$$

即攻击者要找到 k 组 $DCT(\hat{O})_i, \hat{r}_i$ 以及 $\hat{\omega}$,使得:

$$h^{\omega} \cdot H(g^{DCT(\hat{O})_1} f^{\hat{r}_1} \bmod p, g^{DCT(\hat{O})_i} f^{\hat{r}_i} \bmod p, ..., g^{DCT(\hat{O})_k} f^{\hat{r}_k} \bmod p) \bmod q = WM'.$$

由于水印生成函数的单向性,攻击者成功的概率可以忽略.

其次,我们说明本方案可以防止穷举攻击.显然,当随机选择 $DCT(\hat{O})_i, \hat{r}_i$ 以及 $\hat{\omega}$ 以后,水印 WM' 也确定了,根据文献[21]的分析,攻击者不能使得它通过第(5)~(7)步的水印存在性检测.我们注意到存在另外一种随机攻击步骤,即攻击者先随机选择 $com(DCT(\hat{O})_i)$ 和 $\hat{\omega}$,找到合适的水印 WM' 后,计算

$$DCT(O')_i = DCT(\bar{O})_i / (1 + \alpha \cdot WM_i'),$$

从而可以通过存在性检测.但是,根据承诺方案的计算绑定性质,攻击者不能伪造对应于同样承诺 $com(DCT(\hat{O})_i)$ 的不同被承诺值 $DCT(O')_i$.所以,本方案可以防止穷举攻击.

从以上的分析可以看到,由于采用了零知识的证明协议,并且所有参数的选择都不依赖于可信第三方,所以任何人都可以验证参数和水印是否合法,水印是否存在.由于所有的交互式零知识协议都可以转换为非交互式协议,所以当转变为非交互式协议以后,验证过程不需要证明者的参与.需要说明的是,当验证协议转变为非交互协议以后,对图像进行任何的改动都会改变 C 的值,因此不会通过该协议的验证,所以,该协议可以抵抗敏感性攻击.

© 中国科学院软件研究所 http://www.jos.org.cn

4.5 其他安全相关的问题

4.5.1 扩频因子较小的情况

在实际使用的、目前推荐的安全素数的位数 l 至少在 1 024 位,当水印的扩频因子较小时,比如 k=1000,相应的验证协议也要更改.具体的方法是,把协议的第(4)步变更为:证明者计算 $L = g^{WM \bmod 2^{l-k}} \bmod p$,将 L 发送给验证者.证明者用零知识协议 ZK-DL 证明知道 L 基于 g 的离散对数,以及知道 T 使得:$f^T = \frac{L}{m}\prod_{i=1}^{k}(g \cdot com(WM_i))^{\frac{q+1}{2}\cdot 2^{k-i}} \bmod p$.

4.5.2 更大的水印分量取值范围

前面要求水印的取值在{−1,+1}上,实际上,可以容易扩展到更大的区间.比如,可以要求 WM 中每4位对应一个水印分量,这样,每个分量有 8 种可能的取值.具体的实现这里不再详述.

4.5.3 多文档合谋攻击

所谓多文档合谋攻击,就是同一个文档 D 分别嵌入 t 个不同的水印,生成 t 个文档 $D_1, D_2, ..., D_t$,由这 t 个文档恢复文档 D,从而恢复水印的攻击.Cox 指出[4],如果水印分量在均匀分布的区间取值,那么一般只需要 5 个文档就可以攻击成功.一种简单的避免这种攻击的方法是,要求对同一文档不能嵌入不同水印.这种要求对版权所有者鉴别来说是可行的.

5 结 论

本文基于 Pedersen 的承诺方案提出了公开可验证的承诺方案和证明知道被承诺值的离散对数的零知识知识证明协议.在此基础上,改进了 Adelsbach 的零知识水印验证方案,实现了可防止模糊攻击、公开可验证的零知识水印验证协议.对需要有第三方参与的版权所有者证明方案[8],利用我们提出的方案可以大大降低对第三方可信度的依赖.需要指出的是,为了降低验证者错误接收的概率,本文中的零知识协议需要执行多轮,因此效率较低.我们进一步的工作将研究如何设计具有高效率的水印检测协议.一种可行方法是设计常数轮的、错误概率可忽略的高效率零知识协议,另一种可能的方法是在保证安全的前提下,采用非零知识的知识证明协议.

References:

[1] Kalker T, Linnartz JP, Dijk MV. Watermark estimation through detector analysis. In: Proc. of the IEEE Int'l Conf. on Image Processing (ICIP'98). Los Alamitos: IEEE Computer Society Press, 1998. 425−429.

[2] Kinoshita H. An image digital signature system with zkip for the graph isomorphism problem. In: Proc. of the IEEE Conf. on Image Processing (ICIP'96), Vol 3. Los Alamitos: IEEE Computer Society Press, 1996. 247−250.

[3] Gopalakrishnan K, Memon N, Vora P. Protocols for watermark verification: Multimedia and security. IEEE Multimedia, 2001,8(4): 66−70.

[4] Cox IJ, Kilian J, Leighton T, Shamoon T. Secure spread spectrum watermarking for multimedia. IEEE Trans. on Image Processing, 1997,6(12):1673−1687.

[5] Craver S. Zero knowledge watermark detection. In: Information Hiding: The 3rd Int'l Workshop. LNCS 1768, Berlin: Springer-Verlag, 2000. 101−116.

[6] Adelsbach A, Katzenbeisser S, Sadeghi AR. Watermark detection with zero-knowledge disclosure. ACM Multimedia Systems Journal, 2003,9(3):266−278.

[7] Adelsbach A, Katzenbeisser S, Sadeghi AR. Cryptography meets watermarking: Detecting watermarks with minimal or zero knowledge disclosure. In: Proc. of the European Signal Processing Conf. (EUSIPCO 2002). 2002. 446−449.

[8] Adelsbach A, Sadeghi R. Zero knowledge watermark detection and proof of ownership. In: Information Hiding: The 4th Int'l Workshop. LNCS 2137, Berlin: Springer Verlag, 2001. 273−287.

[9] Zou XX, Dai Q, Huang C, Li JT. Zero-Knowledge watermark verification protocols. Journal of Software, 2003,14(9):1645−1651 (in Chinese with English abstract). http://www.jos.org.cn/1000-9825/14/1645.pdf

[10] Craver S, Memon N, Yeo BL, Yeung MM. Resolving rightful ownerships with invisible watermarking techniques: Limitations, attacks, and implications. IEEE Journal on Selected Area in Communications, 1998,16(4):573−586.

© 中国科学院软件研究所 http://www.jos.org.cn

[11] Bellare M, Goldreich O. On defining proofs of knowledge. In: Advances in Cryptology, Crypto'92. LNCS 740, Berlin: Springer-Verlag, 1993. 390–420.

[12] Goldreich O, Oren J. Definitions and properties of zero-knowledge proof systems. Journal of Cryptology, 1994,7(1):1–32.

[13] Pedersen TP. Non-Interactive and information-theoretic secure verifiable secret sharing. In: Advances in Cryptology, CRYPTO'91. Berlin: Springer-Verlag, 1991. 129–140.

[14] Fujisaki E, Okamoto T. A practical and provably secure scheme for publicly verifiable secret sharing and its applications. In: Advances in Cryptology—EUROCRYPT'98. LNCS 1403, Berlin: Springer-Verlag, 1996. 32–46.

[15] Pan CD, Pan CB. Elememtary Number Theory. Beijing: Peking University Press, 1992 (in Chinese).

[16] Song YY. Number Theory for Computing. Berlin: Springer-Verlag, 2000.

[17] Bach E. Analytic Methods in the Analysis and Design of Number-Theoretic Algorithms. Cambridge: MIT Press, 1985.

[18] Goldreich O. Foundations of Cryptography. Cambridge: Cambridge University Press, 2001.

[19] Bellare M, Rogaway P. Random oracles are practical: A paradigm for designing efficient protocols. In: Proc. of the1st ACM Conf. on Computer and Communications Security. New York: ACM Press, 1993. 62–73.

[20] Boudot F. Efficient proofs that a committed number lies in an interval. In: Advances in Cryptology, Eurocrypt 2000. LNCS 1807, Berlin: Springer-Verlag, 2000. 431–444.

[21] Ramkumar M, Akansu A. Image watermarks and counterfeit attacks: Some problems and solutions. In: Proc. of the Symp. on Content Security and Data Hiding in Digital Media. Newark: New Jersey Institute of Technology, 1999. 102–112.

附中文参考文献:

[9] 邹萧湘,戴琼,黄晁,李锦涛.零知识水印验证协议.软件学报,2003,14(9):1645–1651. http://www.jos.org.cn/1000-9825/14/1645.pdf

[15] 潘承洞,潘承彪.初等数论.北京:北京大学出版社,1992.

****

中国计算机学会设立创新奖

中国计算机学会创新奖是一项完全由社会力量设立的在计算机领域的科学技术创新奖。中国计算机学会负责该奖项评选和颁奖，每年评选一次。2005 年是首次。

中国计算机学会设立此奖的目的是为推动中国计算机及相关领域的科技创新和进步，促进科研成果的转化，促进 IT 产业的发展，推动科技界学术共同体评价体系的建立，发现和激励创新型科技人才。

中国计算机学会创新奖的评奖活动严格按照公开、公正的原则，根据《中国计算机学会创新奖评奖条例》执行。创新奖评委机构由 CCF 设立，评委会分为初评委员会和终评委员会，成员均为业内专家。为保证评选的公正性不受干扰，评奖委员会成员名单在最终评奖结果宣布以前不对外公布。主办单位对入围项目将予以公布，接受社会对入围项目的署名投诉。此外，对查实的不符合条例的行为将予以处罚。

申报中国计算机学会创新奖不需要缴纳任何费用，拥有科技成果的团体或个人均可申报参评此奖项，此外，获奖者还可获得证书和奖金。

2005 年度中国计算机学会创新奖的申报工作已经开始。

相关信息请登陆中国计算机学会网站 http://www.ccf.org.cn 了解。

© 中国科学院软件研究所 http://www.jos.org.cn