

基于多变量二次多项式的动态图软件水印方案

张茂胜^{1,2}, 孙小雁³

(1 玉林师范学院 数学与信息科学学院, 广西 玉林 537000; 2 武汉大学 国家多媒体软件工程技术研究中心, 湖北 武汉 430072; 3 玉林师范学院 计算机科学与工程学院, 广西 玉林 537000)

摘 要: 针对现有动态图软件水印算法在安全性上存在的不足, 提出一种基于多变量二次多项式的软件水印方案. 首先建立从基本域到扩域的中心映射, 然后利用两个仿射变换将中心映射隐藏, 将版权信息从基本域映射到扩域得到水印信息, 再将水印信息转换为整数, 利用传统动态图的方法嵌入程序当中. 通过实例验证了水印证明版权的有效性.

关键词: 多变量; 仿射; 软件水印; 版权

中图分类号: TP391

文献标识码: A

文章编号: 1000-7180(2012)12-0068-04

Dynamic Software Watermarking Scheme Based on Multivariate Quadratic Polynomials

ZHANG Mao-sheng^{1,2}, SUN Xiao-yan³

(1 College of Math and Information, Yulin Normal University, Yulin 537000, China;

2 National Engineering Research Center for Multimedia Software, Wuhan University, Wuhan 430072, China;

3 College of Computer Science, Yulin Normal University, Yulin 537000, China)

Abstract: An innovative watermark algorithm against security problems of the existing watermark scheme is proposed, in which the multivariate quadratic polynomials and traditional dynamic graph are combined. First, a central map from fundamental field to extension field is designed. And then the central map is hidden by using two affine transformations and the copyright information in fundamental field are mapped to software watermark in extension field. The watermark is embedded after it is transformed into an integer. An example shows the validity of the system.

Key words: multivariate; affine; watermarking; copyright

1 引言

软件水印作为数字水印的一个分支, 是软件的保护技术之一. 常见的软件保护技术如序列号、授权文件、软件狗等, 这些保护技术容易遭受破解攻击, 难以实现盗版跟踪和提供法律上统一的证明版权的依据^[1]. 软件水印是一种与以上方法完全不同的软件保护技术, 通过在软件产品中嵌入版权信息和身份认证信息, 在发生版权纠纷时将这些信息提取出

来以证明软件的所有权^[2]. Collberg 提出了一种动态图水印技术^[3], 把软件水印隐藏在程序动态建立的图拓扑结构中, 其原理是用图的拓扑结构表示一个大整数 N , 并拆分成两个大素数 P 、 Q 的乘积. 由大整数分解困难性, 只有合法的版权所有人才能将 N 分解成 P 和 Q , 由此证明版权^[4]. 然而, 随着计算机硬件计算能力的不断提高, 算法中使用的大整数应该在 1024 位以上, 使得指针数量大幅增加, 这一方面增加了内存开销, 另一方面降低了水印的隐蔽

收稿日期: 2012-04-26; **修回日期:** 2012-05-18

基金项目: 广西壮族自治区教育厅立项项目(201106LX513); 广西自然科学基金资助项目(桂科青 0832101); 广西壮族自治区教育厅立项项目(201106LX516); 玉林师范学院青年科研项目(2010YJQN18)

性^[5-6]. 量子计算机的出现也使得大整数分解不再是不可攻克的难题^[7].

本文利用有限域上多变量多项式分解的困难性问题设计软件水印. 在构造基本域和扩展域后, 在基本域中生成版权信息, 通过仿射变换将版权信息进行置换运算, 然后将版权信息从基本域映射到扩展域, 再用另一个仿射变换经转换运算后得到水印信息, 通过两个仿射变换将中心映射隐藏. 最后将水印信息用动态图的方式嵌入到程序当中. 在提取水印信息后, 由于只有版权所有者才能将版权信息从基本域映射到扩展域, 从而证明版权.

2 多变量多项式中的困难性问题

有限域 F_q 上 n 个变量 m 个方程的二次多项式方程组形式为如式(1)所示.

$$y_i(x_1, \dots, x_n) = \sum_{1 \leq j < k \leq n} c_{ijk} x_j x_k + \sum_{1 \leq i \leq n} b_{ij} x_j + a_i \quad (1)$$

其中 $c_{ijk}, b_{ij}, a_i \in F_q, i=1, \dots, m$.

定义 1 随机给定一个 F_q 上 n 个变量的二次多项式方程组 $Y=(y_1, \dots, y_m), m \leq n$. 对于已知元素 $Y \in F_q^m$, 求解一个元素 $x \in F_q^n$, 使得 $Y=(y_1(x_1, \dots, x_n), \dots, y_m(x_1, \dots, x_n))$, 称为 MQ 问题.

定义 2 设 P, F 是两个给定的有限域 $GF(q)$ 上 n 个变量 m 个多项式的随机多变量二次多项式组, 求解有限域 $GF(q)$ 上的两个仿射变换 U, T 使得 $F=T \cdot P \cdot U(X)$, 称为 IP 问题. 其中 $U \in GF(q)^n \times GF(q)^n, T \in GF(q)^m \times GF(q)^m$.

MQ 问题和 IP 问题分别是 NP 完全问题和 NP 困难问题^[8-9], 同时没有证据表明量子计算机对 MQ 问题存在高效解法. 因此, 将 MQ 问题代替安全性受到威胁的大整数分解问题用于生成软件水印, 可以提高水印的安全性能.

3 基于多变量二次多项式的软件水印算法

首先, 版权所有者确定基本域 F_q , 设置正整数 $n > m$; 产生仿射变换 U, T 和中心映射 P , 仿射变换 U, T 的表达形式如式(2)所示.

$$U = \begin{bmatrix} u_{11} & u_{12} & \cdots & u_{1n} \\ \vdots & \vdots & \ddots & \vdots \\ u_{n1} & u_{n2} & \cdots & u_{nn} \end{bmatrix} \quad (2)$$

$$T = \begin{bmatrix} t_{11} & t_{12} & \cdots & t_{1m} \\ \vdots & \vdots & \ddots & \vdots \\ t_{m1} & t_{m2} & \cdots & t_{mm} \end{bmatrix}$$

对任意消息 $X=(x_1, x_2, \dots, x_n)$, 计算 $Y=T \cdot$

$P \cdot U(X)$ 可得关于 x 的二次方程组如式(3)所示.

$$\begin{cases} y_1 = \sum_{1 \leq j < k \leq n} c_{1jk} x_j x_k + \sum_{j=1}^n b_{1j} x_j + a_1 \\ \vdots \\ y_i = \sum_{1 \leq j < k \leq n} c_{ijk} x_j x_k + \sum_{j=1}^n b_{ij} x_j + a_i \\ \vdots \\ y_m = \sum_{1 \leq j < k \leq n} c_{mjk} x_j x_k + \sum_{j=1}^n b_{mj} x_j + a_m \end{cases} \quad (3)$$

将该多项式方程组保存, 或制作成数字证书提交第三方权威验证机构保管, 在发生版权纠纷时提取以便证明版权. 计算二次方程组的过程如图 1 所示.

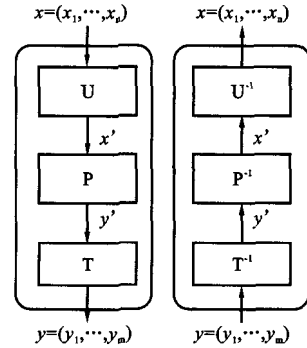


图1 结合MQ和IP问题产生二次方程组框架

对版权信息 $S=(s_1, s_2, \dots, s_m)$, 计算水印向量 $W=(w_1, w_2, \dots, w_n)$ 如式(4)所示.

$$W=U^{-1} \cdot P^{-1} \cdot T^{-1}(S) \quad (4)$$

将水印向量 W 转换成拓扑图嵌入程序源代码当中. 转换可采取两种方式: 链表方式和动态图方式.

(1) 链表法嵌入水印

利用单链表将水印向量嵌入程序当中, 链表中每个结点包含两个域: 数据域和指针域. 数据域存放水印值, 指针域指向下一个结点, 头结点指向第一个水印. 如图 2 所示.

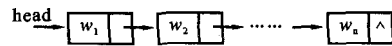


图2 链表法表示水印示意图

(2) 动态图法嵌入水印

传统的动态图法将一个整数转换成拓扑图嵌入程序, 如基数 k 编码、排列图编码、PPCT 编码等. 本文的水印信息是一个向量, 因此需要先将该向量转换成整数后再使用动态图法. 以二元基本域为例, 水印向量按照公式(5)转换为整数 N , 然后再将 N 转换为传统动态图.

$$N = \sum_{i=1}^n w_i \times 2^{n-i} \quad (5)$$

完整的水印生成过程如图 3 所示。

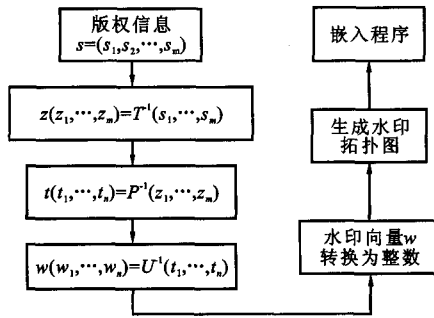


图 3 多变量水印嵌入流程图

对向量长度为 m 的同一个版权信息,能生成 $(n-m)^2$ 个不同的水印,因此水印资源非常丰富。也可以通过在版权信息中添加 $(n-m)^2$ 个校验位,使版权信息与水印之间建立一一对应的关系。

水印的提取过程以及版权的证明与传统拓扑图结构类似,在 Collberg 和 Thornborson^[3] 提出的动态图水印技术中,检测出大整数 N 后,只有合法用户才能分解成两个大素数的乘积从而证明版权^[10]。本文提出的方案中,首先提取拓扑图,转换为整数后再生成水印向量,由于只有版权所有者才能生成同样的向量,因此可以证明合法版权。

4 多变量软件水印实例

正建立映射 $F: GF(25) \rightarrow GF(23)$, $F = T \cdot P \cdot U$, 其中 T, P, U 分别如式 (6) 所示。

$$\begin{aligned}
 U: \begin{bmatrix} t_0 \\ t_1 \\ t_2 \\ t_3 \\ t_4 \end{bmatrix} &= M_1 X + c_1 = \begin{bmatrix} 1 & 0 & 0 & 1 & 1 \\ 1 & 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 & 0 \\ 1 & 1 & 1 & 1 & 1 \\ 0 & 1 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} x_0 \\ x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} + \begin{bmatrix} 1 \\ 1 \\ 0 \\ 1 \\ 0 \end{bmatrix} \\
 P: \begin{cases} z_2 = t_2 + t_0 t_1 \\ z_3 = t_3 + t_1 t_2 \\ z_4 = t_4 + t_2 t_3 \end{cases} \\
 T: \begin{bmatrix} y_0 \\ y_1 \\ y_2 \end{bmatrix} &= M_2 z + c_2 = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{bmatrix} \begin{bmatrix} z_2 \\ z_3 \\ z_4 \end{bmatrix} + \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \quad (6)
 \end{aligned}$$

通过计算可得到多变量二次多项式如式 (7) 所示。

$$\begin{aligned}
 y_0 &= x_0 + x_1 + x_2 + x_3 + x_0 x_1 + x_0 x_2 + x_1 x_3 + x_1 x_4 + x_2 x_4 + x_3 x_4; \\
 y_1 &= x_2 + x_4 + x_0 x_3 + x_1 x_2 + x_1 x_3 + x_1 x_4 + x_2 x_3 + x_2 x_4 + x_3 x_4; \\
 y_2 &= x_0 + x_2 + x_0 x_2 + x_0 x_3 + x_0 x_4 + x_1 x_2 + x_1 x_3 + x_1 x_4 + x_2 x_3 + x_3 x_4; \quad (7)
 \end{aligned}$$

将式 (7) 作为数字证书,由第三方权威机构保管,其它参数由版权所有者保管。

假设版权信息为 $S=(1,1,0)$,按如下步骤计算水印信息:

(1) 通过 T 的逆变换可计算得 $z = M_2^{-1}(Y - c_2) = (1,1,1)$;

(2) 将 z 的值代入中心映射 P 计算 t ,此时 t 的值不唯一,有如下 4 个可能值:

$(0,0,1,1,0), (0,1,1,0,1), (1,0,1,1,0), (1,1,0,1,1)$

(3) 将 t 的四种可能值代入仿射变换 U 的逆变换中计算水印向量 $w = M_1^{-1}(Y - c_1)$,四种可能值如下:

$(1,1,0,1,1), (1,0,0,1,1), (1,0,0,0,1), (1,1,1,0,1)$

(4) 将水印向量转换为整数可得 $w=27, 19, 17, 29$,选择其中的一个作为水印信息。

(5) 将水印数字转换为拓扑图。

在第 (1) 步中计算 T 后增加 2 位校验位就可以得到唯一的水印向量。

软件水印的验证十分便捷,版权所有者在提取拓扑图后还原为水印向量,通过出示保存的多变量二次多项式或第三方权威机构保管的证书,由版权信息和 T, P, U 计算得到与提取的水印向量一致的向量,即可证明合法版权。

5 结束语

本文提出基于多变量二次多项式的软件水印方案,通过同时利用 MQ 问题和 IP 问题,结合传统动态图技术,将版权信息进行变换后嵌入程序,与基于大整数分解问题的软件水印方案相比安全性更好、运算效率更高。方案中最后将水印向量转换成整数后再使用传统动态图,该过程既降低了运算效率,也影响了安全性。设计一种方案直接将水印向量转换成拓扑图,是下一步的研究内容。

参考文献:

- [1] 芦斌,罗向阳,刘粉林.一种基于混沌的软件水印算法框架及实现[J].软件学报,2007,18(2):351-360
- [2] 王慧娇,沙宗鲁,轩爱成.基于 PPCT 和基数 k 的动态图混合编码方案[J].Computer Engineering and Applications, 2010,46(25):109-111.
- [3] Christian Collberg, Clark Thornborson. Software watermarking: models and dynamic embeddings [C]// Conference Record of the Annual ACM Symposium on

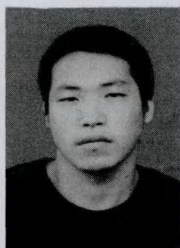
- Principles of Programming Languages. San Antonio, TX, USA:ACM, 1999:311-324.
- [4] Christian Collberg, Andrew Huntwork, Edward Carter, et al. More on graph theoretic software watermarks: Implementation, analysis, and attacks[J]. Information and Software Technology, 2009, 51(1): 56-67.
- [5] Jian Qi Zhu, Ai Min Wang, Yan Heng Liu. Tamper-proofing software watermarking scheme based on constant encoding[C]//Education Technology and Computer Science (ETCS), 2010 Second International Workshop on. Dunedin, New Zealand: Australian Computer Society, Inc., 2010:129-132.
- [6] Zaidi S J H, Wang H. On the analysis of software watermarking[C]//Software Technology and Engineering (ICSTE), 2010 2nd International Conference on. Chengdu, China, IEEE Conference Publications, 2010: 26-30.
- [7] Shor P W. Algorithms for quantum computation: discrete logarithms and factoring [C]//Foundations of Computer Science, 1994 Proceedings, 35th Annual Symposium on. Santa Fe: USA, 1994:124-134.
- [8] Johnson D S, Garey M R. Computers and intractability: a guide to the theory of NP-completeness[M]. Freeman&Co, San Francisco, 1979.
- [9] Jacques Patarin. Hidden Fields Equations (HFE) and Isomorphisms of Polynomials (IP): Two New Families of Asymmetric Algorithms[C]//Advances in Cryptology-EUROCRYPT'96. Berlin, Heidelberg: Springer, 1996:33-48.
- [10] Zhang L, Yang Y, Niu X, et al. A survey on software watermarking[J]. Journal of Software, 2003, 14(2): 268-277.

作者简介:



张茂胜 男,(1981-)博士研究生,讲师.研究方向为信息安全、多媒体.

(上接第 67 页)



万磊 男,(1986-),硕士研究生.研究方向为无线传感器网络.



龚仁彬 男,(1964-),博士,高级工程师,副教授.研究方向为物联网应用.