基于二元 Golay隐写码的快速隐写算法

李玉辉, 刘九芬, 张 卫明

(信息工程大学 信息工程学院, 河南 郑州 450002)

摘要: 文章首先研究了基于二元 Golay码的隐写编码的性能, 然后研究了基于二元 Golay隐写码的快速实现问题, 提出了一种快速隐写算法。该算法在不提高载体数据修改率的前提下, 有效降低了编码的计算复杂度, 提高了隐写算法的实现效率。

关键词: 信息隐藏; 隐写码; Golay码; 快速算法
中图分类号: TN918; O15704　　文献标识码: A　文章编号: 1671－0673(2007)03－0269－03

Fast Steganographic Algorithm Based on the Golay Codes

LI Yu-hui, LIU Jiu-fen, ZHANG Wei-ming

(Institute of Information Engineering, Information Engineering University, Zhengzhou 450002, China)

Abstract: This paper researches two problems, which are the capability of the steganographic code based on the binary Golay codes and the fast realization of this code. Then the paper presents a fast algorithm. In condition of keeping the modification rate, this algorithm could reduce computational complexity of encoding. Hence, it improves the efficiency of the steganographic algorithm.

Key words: information hiding; steganographic code; golay code; fast algorithm

在当前 Internet高速增长、电子商务和电子政务等社会活动信息化不断发展的背景下, 信息隐藏已经成为信息安全领域研究的热点。隐写术是信息隐藏技术的一个重要分支, 主要研究如何将秘密信息隐藏于公开的数字媒体之中实现隐蔽通讯。秘密信息的伪装性和与载体的不可分离性使隐写术除了具有传统加密通信的优点外, 还大大减少了受攻击的可能性。因此隐写术是一种利用公共信道传递涉密信息的可靠方法, 可以保证国家政治、军事、经济信息在公共网络中的通信安全。

提高隐藏效率, 即对载体作尽量少的修改而隐藏尽可能多的消息, 是关系隐写术安全性和有效性的一个重要问题。应用编码技术是解决这一问题的有效途径。Ron Crandall[1]提出应用矩阵编码可以实现高效率的隐写方案, 而且矩阵编码被用于基于 JPEG图像的 F5算法[2]。基于二值图像的 CPT算法[3,4], 以及隐写软件 Steganos Security Suite 5.06[5]也都采用了编码技术来提高隐写效率。

在特定环境下, 通讯双方传输的秘密信息常常具有时效性, 需要双方在最短的时间内安全地完成通讯。所以如何实现快速隐写, 降低信息嵌入的计算复杂度是隐写算法设计关心的另一重要问题。Dagmar Schonfeld[6]对基于 BCH码的隐写编码利用生成多项式给出了一种快速实现方法, 该方法通过适当提高载体的修改率来降低编码的计算复杂度。

本文首先研究了基于二元 Golay码的隐写编码的性能, 结果表明二元 Golay隐写码是目前已知的性能优良的隐写编码方法之一, 利用它可以达到低的修改率并实现较大的信息嵌入量, 可以满足多数隐写算法的需求。然后研究了基于二元 Golay隐写码的快速实现问题, 提出了一种快速隐写算法。该算法在保持原有修改率的前提下, 有效降低了编码的计算复杂度, 因而可同时满足好的隐蔽性和网络实时通讯的高效性, 具有很好的应用前景。

收稿日期: 2007－03－09; 修回日期: 2007－04－04
基金项目: 国家自然科学基金资助项目(60473022)河南省自然科学基金资助项目(0511011300)
作者简介: 李玉辉(1981－), 男, 河南信阳人, 信息工程学院硕士研究生, 主要研究方向为信息隐藏技术。

1 二元 Golay隐写码

隐写编码通过分组式嵌入来降低修改率，可以应用于各种隐写术，本文以灰度图像 LSB（Least Significant Bit）隐写术为例进行描述。利用 Golay 码可以在 23个像素的 LSB位嵌入 11比特秘密信息，而至多只修改其中 3个像素点。通常用平均修改率（R）和嵌入容量（C）来衡量一个隐写码性能的好坏。对于 Golay隐写码，每个分组中平均修改的像素点个数为

$$E=\frac{1}{2^{11}}\cdot 0+\frac{C_{23}^{1}}{2^{11}}\cdot 1+\frac{C_{23}^{2}}{2^{11}}\cdot 2+\frac{C_{23}^{3}}{2^{11}}\cdot 3=2.853,$$

平均修改率为 $R=\frac{E}{11}=0.259$，

载体可嵌容量为 $C=\frac{11}{23}=0.478$。

表 1给出几种常见隐写码的性能指标。可以看出，Golay隐写码是一类性能比较好的隐写码。

表 1 Golay隐写码与其它码的性能比较

算法名称	平均修改率（R）	嵌入容量（C）
LSB替换/匹配	0.5	100%
LSB Matching Revisited[7]	0.375	100%
F5矩阵编码	0.29	43%
(23, 11, 3) Golay隐写码	0.259	47.8%

2 二元 Golay隐写码的快速隐写算法

本算法是基于 Golay码位置校验集合译码思想[8]而设计的。首先介绍相关知识。

2.1 相关知识

系统 Golay码校验矩阵 $H=(P|I_{11})$，生成矩阵为 $G=(I_{12}|P^{T})$，其中 I_n是一个 11×11单位矩阵，I_{12}是一个的 12×12单位矩阵，P是如下的一个 11×12矩阵

$$P=\begin{pmatrix}
1&0&1&0&0&0&1&1&1&0&1&1\\
1&1&0&1&0&0&0&1&1&1&0&1\\
0&1&1&0&1&0&0&0&1&1&1&1\\
1&0&1&1&0&1&0&0&0&1&1&1\\
1&1&0&1&1&0&1&0&0&0&1&1\\
1&1&1&0&1&1&0&1&0&0&0&1\\
0&1&1&1&0&1&1&0&1&0&0&1\\
0&0&1&1&1&0&1&1&0&1&0&1\\
0&0&0&1&1&1&0&1&1&0&1&1\\
1&0&0&0&1&1&1&0&1&1&0&1\\
0&1&0&0&0&1&1&1&0&1&1&1
\end{pmatrix}$$

下面用 $A\cap B$表示 A与 B的交集，$\bar{A}$表示 A关于全集 $I=\{1,2,\cdots,11\}$的补集。应用位置校验集合表示 P可得到集合

$A=\{A_1\ A_2\ A_3\ A_4\ A_5\ A_6\ A_7\ A_8\ A_9\ A_{10}\ A_{11}\}$，其中位置校验集合为

$A_1=\{1,3,7,8,9,11\}$，$A_2=\{1,2,4,8,9,10\}$，
$A_3=\{2,3,5,9,10,11\}$，$A_4=\{1,3,4,6,10,11\}$，
$A_5=\{1,2,4,5,7,11\}$，$A_6=\{1,2,3,5,6,8\}$，
$A_7=\{2,3,4,5,7,9\}$，$A_8=\{3,4,5,7,8,10\}$，
$A_9=\{4,5,6,8,9,11\}$，$A_{10}=\{1,5,6,7,9,10\}$，
$A_{11}=\{2,6,7,8,10,11\}$。

2.2 快速隐写算法

该算法是以 23个像素为分组单元，每分组单元嵌入 11比特的秘密消息。下面以一组消息嵌入为例，描述算法的具体流程如下：

第 1步 读取图像载体数据，用伪随机生成器控制选取原始数据。记原始数据的一维数组 $a=(a_1\ a_2\ \cdots\ a_{22}\ a_{23})$，记原始数据 a的最低有效位为 LSB（a），欲嵌入的秘密消息为二进制流 $m=(m_{10}\ m_9\ \cdots\ m_0)$，计算 $c=(c_1,c_2,\cdots,c_{23})=LSB(a)\oplus(0\ 0\ \cdots\ 0\ m_{10}\ m_9\ \cdots\ m_0)$。

第 2步 将 c视作二元 Golay码的接收码字，计算其伴随式

$$s=Hc^{T}=(s_{10}\ s_9\ \cdots\ s_0)^{T}。$$

第 3步 找出标记集合。s中为 1的位表示对相应的位置校验集合进行了标记，为 0的位表示对相应的位置校验集合未进行标记，如 $s=(1\ 1\ 0\ 0\ 0\ 0\ 0\ 0\ 1\ 0)$，则说明 s只对 A_1，A_2，A_{10}标记，而对其它集合未进行标记。设 ω是 s中非零元素的个数，有 $p=\min\{\omega,11-\omega\}\leqslant 5$则要么有 p个标记的位置校验集合，要么有 p个未标记的位置校验集合 B_1，B_2，…，B_p。

第 4步 对 s进行如下操作，计算码字 c信息位（即前 12位码元）的错误图样 e：

（1）当 $\omega\leqslant 3$时，则 $e=\phi$。

（2）当 $\omega=4$时，则

① $e=\{i\}(\{i\}=B_1B_2B_3B_4)$；

② $e=\{i,12\}(\{i\}=\bar{B}_1\bar{B}_2\bar{B}_3\bar{B}_4)$；

③ $e=\{i,j,k\}(\{i,j,k\}=B_1B_2\bar{B}_3\bar{B}_4)$，这种类型的交集运算共有 6种，即 $\{a,b\}\cup\{c,d\}=\{1,2,3,4\}$。

（3）当 $\omega=5$时，则

① $e=\{i\}(\{i\}=B_1B_2B_3B_4B_5)$；

② $e=\{i,12\}(\{i\}=\bar{B}_1\bar{B}_2\bar{B}_3\bar{B}_4\bar{B}_5)$；

③ $e=\{i,j,12\}(\{i,j\}=B_aB_bB_cB_d\bar{B}_e)$，这种交

集运算的种类共有 10种,即 $\{a,b,c\}\cup\{d,e\}=\{1,2,3,4,5\}$;

④ $e=\{i,j\}$ ($\{i\}=B_aB_b\bar{B}_c\bar{B}_d\bar{B}_e$ 和 $\{j\}=\bar{B}_a\bar{B}_b\bar{B}_cB_dB_e$),这种类型的运算共有 10种。

(4)当 $\omega=6$时,则

① $e=\{i,12\}$ ($\{i\}=B_1B_2B_3B_4B_5$);

② $e=\{i\}$ ($\{i\}=\bar{B}_1\bar{B}_2\bar{B}_3\bar{B}_4\bar{B}_5$ 或 $\{i\}=B_a\bar{B}_b\bar{B}_c\bar{B}_d\bar{B}_e$),这种交集运算的类型共有 5种;

③ $e=\{i,j\}$ ($\{i,j\}=B_aB_bB_c\bar{B}_d\bar{B}_e$),这种交集运算的类型共有 10种。

(5)当 $\omega=7$时,则

① $e=\{i\}$ ($\{i\}=\bar{B}_1\bar{B}_2\bar{B}_3\bar{B}_4$);

② $e=\{i,j\}$ ($\{i,j\}=B_aB_b\bar{B}_c\bar{B}_d$),其中 $\{a,b\}\cup\{c,d\}=\{1,2,3,4\}$,或者 ($\{i,j\}=B_aB_bB_c\bar{B}_d$),其中 $\{a,b,c\}\cup\{d\}=\{1,2,3,4\}$)。

(6)当 $\omega=8$时,则

① $e=\{i\}$ ($\{i\}=\bar{B}_1\bar{B}_2\bar{B}_3$);

② $e=\{i,j,k\}$ ($\{i,j,k\}=B_1B_2\bar{B}_3\cup B_1\bar{B}_2B_3\cup\bar{B}_1B_2B_3$)。

(7)当 $\omega\geqslant 9$时,则 $e=\{12\}$。

第 5步　计算码字 c的错误图样

$E=(E_1,E_2,\cdots,E_{23})=(c\oplus LSB(a_1,a_2,\cdots,a_{12}))G\oplus C$

第 6步　将数组 a与 E中非零元素相对应位置的数据随机加 1或者减 1,得到载密数据 a'。秘密信息的提取比较简单,在接收方,必有 $H(LSB(a'))^T=S$,可以据此提取嵌入的秘密信息。

2.3　算法计算复杂度

算法的复杂度用算法需要进行异或操作的次数进行度量。在本文算法中:第 2步寻求 s共需要 100次异或操作;第 4步中最坏的一种情况是当 $\omega=5$时,①和②各需要 44次异或,③和④各需要 440次异或操作,因此该步骤总共需要 1045次异或操作;第 5步计算 E共需要 196次异或操作。因此该算法总计需要 1341次异或操作。

针对二元 Golay码,"奇偶校验矩阵 $H_{k\times n}$"算法[1]共需要的异或操作次数为 $((k-n)n+k)2^n=((23\times 11)\times 23+11)\times 2^{23}=5830\times 2^{23}$。

"奇偶校验矩阵结合查表 ($H_{k\times n}$ look-up tables)算法[9]共需要的异或操作次数为 $n\cdot 2^{n-k}=23\times 2^{12}$。

基于生成多项式的改进算法 (g(x) improved)[6]共需要的异或操作次数为 $(n-k)+((n-k)(k+1)+k)(n-k+1)=2027$。

本文算法在保持平均修改率的前提下,使计算复杂度显著降低,明显优于另外 3种算法 (见表 2)。

表 2　针对 Golay码不同隐写算法异或操作次数与平均修改率比较

算法名称	异或操作次数	平均修改率
$H_{k\times n}$	5830×2^{23}	0.259
$H_{k\times n}$ look up tables	23×2^{12}	0.259
g(x) improved	2027	0.311
本文算法	1341	0.259

3　结论

本文首先研究了基于二元 Golay码的隐写编码的性能,结果表明二元 Golay隐写码是目前已知的性能优良的隐写编码方法之一。然后研究了基于二元 Golay隐写码的快速实现问题,应用 Golay码位置校验集合译码思想,设计了一种快速隐写算法。该算法具有相对大的嵌入容量,低的平均修改率,且实现速度比以前的算法有显著提高,适合应用于网络实时通讯环境下各种隐写算法的设计。我们下一步将研究基于三元 Golay码的隐写编码的性能和基于三元 Golay隐写码的快速实现问题。

参考文献:

[1] Crandall R. Some notes on steganography[EB/OL]. [1998-05-20]. http://os.inf.tu-dresden.de/~westfeld/Crandall.pdf

[2] Westfeld A. High capacity despite better steganalysis (F5-A steganographic Algorithm)[C]// Proc. Information Hiding, 4th International Wordshop. Pittsburgh PA, USA, LNCS 2137. New York: Springer-Verlag, 2001: 289-302

[3] Chen Y Y, Pan H K, Tseng Y C. A Secure Data Hiding Scheme for Two Colormages[J]. IEEE Transactions on Communications, 2000, 50(8): 1227-1231

[4] Tseng Y C, Pan H K. Data Hiding in 2-Color Images[J]. IEEE Transactions on Computers, 2002, 51(7): 873-890

[5] Steganos Security Suite 5. Help Document[EB/OL]. [2004-06-30]. http://www.steganos.com/

[6] Schonfeld D, Winkler A. Embedding with Syndrome Coding Based on BCH Codes[C]// ACM, 2006: 214-223

[7] Mielikainen J. Member. LSB Matching Revisited[J]. IEEE Processing Letters, 2006, 13(5): 285-287.

[8] Blaum M, Bruck J. Decoding the Code with Venn Diagrams[J]. IEEE Transactions On Information Theory, 1990, 36(4): 906-910

[9] Dijk M, Willems F. Embedding Information in Grayscale Images[C]// Proc. 22nd Symp. inform. theory in the Benelux. The Netherlands, 2001: 147-154.