自嵌入水印算法的安全性分析

和红杰[1],陈 帆[2]
(1. 西南交通大学信号与信息处理四川省重点实验室; 2. 西南交通大学信息科学与技术学院, 四川成都 610031)

摘 要: 指出自嵌入水印算法中由于偏移值密钥空间较小使算法存在安全隐患,为进一步说明偏移值密钥在自嵌入水印算法的重要性,提出一种同步伪造攻击算法并对最近提出的自嵌入水印算法成功实施了攻击,仿真结果表明偏移值已知会严重破坏自嵌入水印算法的可信性. 最后给出一种随机生成偏移值的方法,实现简单且具有大的密钥空间,有效提高了自嵌入水印算法抵抗同步伪造攻击的能力.

关键词: 脆弱水印; 自嵌入; 同步伪造攻击; 偏移值;
中图分类号: TP391 **文献标识码**: A **文章编号**: 0372-2112 (2007) 03-0557-06

On the Security of the Self-Embedding Watermarking Scheme

HE Hong-jie[1], CHEN Fan[2]
(1. *Sichuan Key Lab of Signal and Information Processing, Southwest Jiaotong University, Chengdu* 610031, *China*;
2. *School of Information Science&Technology, Southwest Jiaotong University, Chengdu* 610031, *China*)

Abstract: A security holes resulting from a small key space of the offset in the exiting self-embedding watermarking schemes have been pointed out. To further illustrate their insecurity, a novel forgery technique, called a 'synchronous counterfeiting (SC) attack', is proposed and implemented on the latest self-embedding watermarking scheme proposed in the literature successfully. We also describe a possible solution to this problem of a small key space that makes our attack computationally intractable.

Key words: fragile watermarking; self-embedding; synchronous counterfeiting attack; the offset

1 引言

计算机技术的飞速发展与信息媒体的数字化,以及各种功能强大的图像编辑软件的出现,使得数字图像传播、编辑、合成和复制等变得方便快捷的同时,也使攻击者可以毫不费力且不留痕迹地篡改其内容. 中国有句古话:"耳听为虚,眼见为实","亲眼所见"是人们区分真伪的依据,但是通过眼睛人们很难辨别数字图像的真伪. 因此,数字图像的完整性、真实性、可信性等问题引起了研究者的广泛关注,基于数字水印的认证技术是解决该问题的有效方法之一[1].

根据容忍图像被修改的程度,基于数字水印的图像认证技术可分为脆弱水印[2~5]和半脆弱水印[6]两类,它们各自有不同的应用场合. 自嵌入水印[4,5]作为脆弱水印算法的一种具有其独特的优点,它不仅能够检测并定位图像被篡改位置,而且可以近似恢复被篡改的图像内容. 1999 年 Fridrich[4]提出了一种基于 DCT 的分块自嵌入脆弱水印算法,该算法将图像分为互不相交的 8 ×8 的图像块,对其高 7 位的 DCT 系数按一定的码长量化编码后,采用固定"偏移值"嵌入另一图像块的最低位 (LSB: Least Significant Bit),从而在定位图像篡改块的同时,可以利用水印信息近似恢复被篡改图像块的内容; 文献[5]通过分析 DCT 系数的特性,重新设定了 DCT 系数的码长分配,使篡改恢复质量的得到进一步的提高. 同时,文献[5]指出"偏移值"应由用户密钥控制以增加系统的安全性.

在自嵌入水印算法中,基于图像块内容生成的水印不是嵌入在该图像块自身的最低位,而是根据一定的偏移值嵌入在另一图像块的最低位. 这样做不仅有利于被篡改内容的恢复[5],而且引入了图像块之间的相关性,提高了算法抵抗 VQ 攻击(vector quantization attack)[7]的能力. 然而,偏移值已知仍然会使算法存在安全隐患,因为自嵌入水印算法中的图像内容和水印是分开的且每个水印块只包含一个图像块的内容信息,根据"偏移值"容易找到图像块对应的水印. 换句话说,攻击者不需要知道水印生成算法和加密私钥,根据含水印图像就能得

收稿日期:2005-12-12;修回日期:2006-06-08
基金项目:四川省青年基金项目(No. 03ZQ026-033)

到众多图像块内容及相应的水印块,这些信息为在不含水印的图像中伪造水印提供了可能.

为说明自嵌入算法中存在的安全性隐患,本文从分析其偏移值密钥空间入手,讨论了根据含水印图像得到偏移值密钥的可能性,然后提出一种新的伪造攻击算法并通过实验仿真验证该攻击对自嵌入水印算法的破坏性.由于该攻击是通过同步修改图像块内容(高7位)和相应图像块的最低位(水印)来实现的,因此我们称之为"同步伪造攻击".最后,本文提出一种具有较大密钥空间的偏移值生成方法来提高自嵌入水印算法抵抗同步伪造攻击的鲁棒性.

2 得到偏移值密钥的可能性

在自嵌入水印算法中,水印嵌入偏移值 p_i[4,5]指定了图像块水印的嵌入位置,即第 i 个图像块的水印嵌入位置 $f(i)$ 与该图像块偏移值的关系为:

$$f(i) = i + p_i, \quad i = 1, 2, \cdots, N \tag{1}$$

其中,N 表示图像块数.根据自嵌入水印算法易知,$f(i)$ 须满足两个必要条件:

(1) 对 $\forall i \in [1, N]$,有 $f(i) \in [1, N]$,即每一个图像块水印都必须放在该图像中某个图像块的最低位;

(2) 对 $\forall i, j \in [1, N], i \neq j$,有 $f(i) \neq f(j)$,即两个图像块的水印不能放在同一个图像块的最低位.

2.1 蛮力攻击(Brute Force Attack)

为减小密钥的传送量,在自嵌入水印算法中对图像块水印的加密采用公钥密码体制.因此,攻击者根据得到的含水印图像和公钥,在偏移值密钥空间不是足够大的情况下,通过蛮力攻击(Brute Force Attack)有可能估计出偏移值密钥.下面分别分析文献[4]和[5]中偏移值的密钥空间.

(1) 文献[4]中的偏移值密钥空间

文献[4]固定选取近似 $N/3$ 为偏移值,偏移值的方向随机选择(向前或向后),为满足 $f(i)$ 的两个必要条件,含水印图像中所有图像块偏移值的方向应是相同的,否则会出现不同图像块的水印嵌入位置相同,即 $\exists i, j (\in [1, N])$ 使的 $f(i) = f(j)$.因此,如果将文献[4]中的偏移值看作密钥的话,该密钥仅有两个取值,显然该算法的偏移值密钥很容易得到.

(2) 文献[5]中偏移值的密钥空间

文献[5]利用线性函数的系数作为控制水印嵌入偏移值的密钥.线性函数如下:

$$f(i) - 1 = k_0 + k_1 * i \pmod N \tag{2}$$

其中,k_0 和 k_1 为偏移值的密钥.根据数论知识可知[8],在同余运算中所有整数可分为 N 个剩余类,即 k_0 和 k_1 的取值仅需考虑模数 N 的一个完全剩余系$\{1, 2, \cdots, N\}$.另外,由于 $f(i)$ 是第 i 个图像块的水印嵌入位置,即对 $\forall i, j \in [1, N], i \neq j$,有 $f(i) \neq f(j)$.也就是说,$f(i)$ 通过模数 N 的完全剩余系[8],这就要求 k_1 满足 $\gcd(k_1, N) = 1$,因此文献[5]中偏移值的密钥空间为:

$$K = \phi(N) * N \tag{3}$$

其中 $\phi(N)$ 为欧拉 ϕ 函数[8],其值等于与 N 互素且小于 N 的正整数个数.由公式(3)可知,对不同大小的含水印图像,偏移值密钥空间有较大的差别.这是因为:(1)密钥空间与图像的大小(图像块的大小一定)成正比;(2)欧拉 ϕ 函数的性质.

设 N 的素分解为:$N = p_1^{e_1} p_2^{e_2} \cdots p_k^{e_k}$ 则:

$$\phi(N) = N\left(1 - \frac{1}{p_1}\right)\left(1 - \frac{1}{p_2}\right)\cdots\left(1 - \frac{1}{p_k}\right) \tag{4}$$

如对 280×208 的灰度图像,$N = 5 * 7 * 2 * 13 = 910$,根据式(4)计算可得 $\phi(N) \approx 0.316N$,因此该图像的偏移值密钥空间约为 $0.316 * 910^2 \approx 2.6 * 10^5$;按同样的方法可以计算出大小为 512×512 图像的偏移值密钥空间为 $2^{23} \approx 8.38 * 10^6$.就目前的计算机速度而言,$10^6$ 基数的密钥空间是不安全的,通过蛮力攻击可以估计出密钥的值.

2.2 认证设备攻击(Verification Device Attack)[2]

所谓认证设备攻击是指攻击者将含水印图像有目的篡改后提交认证设备,根据认证设备输出的篡改检测结果获取算法的秘密信息,一般采用限制访问认证设备的次数来避免认证设备攻击[2].现有自嵌入水印算法中,偏移值密钥为图像块的线性函数,因此攻击者通过几次访问认证设备就可以准确地计算出偏移值密钥.方法如下:

攻击者将含水印图像第 n_1 个图像块的最低位修改后提交认证设备,篡改检测结果指出该图像被篡改并定位被篡改的图像块,记为 i_1.根据篡改检测算法推知,图像块 i_1 的水印被嵌入在第 n_1 个图像块的最低位.即通过一次访问认证设备可以得到一对图像块及其水印的嵌入位置(i_1, n_1).

文献[4]中所有图像块的偏移值相同,因此访问一次认证设备就能计算出偏移值:

$$p_i = n_1 - i_1, \quad i = 1, 2, \cdots, N \tag{5}$$

文献[5]中控制密钥的函数是二元一次方程,因此得到两个图像块对应的水印嵌入位置(即两次访问认证设备),就能准确计算出偏移值密钥.按上述方法,篡改 n_2 图像块的最低位得到另一对图像块及其水印的嵌入位置(i_2, n_2),由式(2)得到下面的二元一次方程组:

$$\begin{cases} n_1 - 1 = (k_0 + k_1 * i_1) \bmod N \\ n_2 - 1 = (k_0 + k_1 * i_2) \bmod N \end{cases} \tag{6}$$

解该二元一次方程组,可以准确地计算出偏移值

密钥 k_0 和 k_1. 这再次说明线性性质是密码设计的大敌[9].

由上述分析可见,现有分块自嵌入水印算法中采用线性函数控制水印嵌入偏移值,其密钥空间较小,攻击者利用得到的含水印图像,通过认证设备攻击或蛮力攻击估计出偏移值密钥的取值是完全可能的.

3　同步伪造攻击及仿真结果

为进一步说明自嵌入水印算法中偏移值已知给算法带来的安全隐患,我们提出了一种同步伪造攻击算法,并通过实验仿真来评估其对自嵌入水印算法安全性的影响.

3.1　同步伪造攻击

假定攻击者得到一幅或多幅利用相同密钥生成的含水印图像并估计出偏移值密钥,攻击者在修改图像块内容的同时,根据水印嵌入偏移值同步修改相应的水印,通过这种方法可以篡改含水印图像或在不含水印的图像中伪造水印,而认证水印算法不能检测此类篡改. 我们称这种伪造含水印图像的方法为"同步伪造攻击".

在描述"同步伪造攻击"算法之前,首先介绍同步伪造攻击的基本操作——"同步替换"$Replace(y_i, y_j)$.

为方便描述,分别用 $\tilde{y}_t$ 和 $\ddot{y}_r$ 表示含水印图像块 y_r 的高 7 位和最低位,用 w_r 表示含水印图像块 y_r 相应的水印. 设 y_i 和 y_j 是两个不同的含水印图像块,"同步替换"操作 $Replace(y_i, y_j)$ 描述如下:

①由水印嵌入偏移值找出 y_i 和 y_j 对应的水印 w_i 和 w_j,即 $w_i = \ddot{y}_{i+p_i}$ 和 $w_j = \ddot{y}_{j+p_j}$;

②篡改操作:用图像块 y_j 的高 7 位替换图像块 y_i 高 7 位,即用 $\tilde{y}_k$ 替换 $\tilde{y}_i$,实现对图像块 y_i 的篡改;

③同步调整水印:用 w_j 替换 w_i,即用 $\ddot{y}_{j+\vec{p}}$ 替换 $\ddot{y}_{i+\vec{p}}$;

由上述描述可知,"同步替换"操作 $Replace()$ 的关键是第①步,即找图像块对应的水印. 而这个操作仅与偏移值有关,与图像、图像大小及图像块的位置均无关,因此在 $Replace(y_i, y_j)$ 操作中,y_i 和 y_j 既可以是同一幅含水印图像的图像块,也可以是采用相同密钥生成的不同含水印图像中的图像块,且含水印图像的大小可以相同也可以不相同.

下面以 $Replace()$ 为基本操作,描述"同步伪造攻击"算法 $X' = \mathrm{forge}(Y_1, \cdots, Y_n, \cdots, Y_m, X)$:

①参数说明:Y_n 为含水印图像($n = 1, 2, \cdots M$, $M > 0$ 为含水印图像的个数);X 为不含水印的图像;X 为伪造的、与 X 视觉相似的含水印图像;

②算法描述:

Begin

　For 对 X 中的任意图像块 x_i

　　从所有含水印图像块集合 $\{y_{nr}, n = 1, 2, \cdots, M, r = 1, 2, \cdots, N_n\}$ 中找出 x_i 近似等价的含水印图像块,记为 y_i,即;

　　$y_i \in \{y_{nr}, n = 1, 2, \cdots, M, r = 1, 2, \cdots, N_n\}$;

　　Replace(x_i, y_i);

End

　End.

从上述攻击算法可以看出,该攻击和向量量化攻击[7]有以下不同:

(1) 替换图像块的方法不同. 向量量化攻击是替换整个图像块,上述攻击是将图像块分为图像内容和水印两部分,通过同步改变图像内容和水印进行伪造攻击;

(2) 对含水印图像的要求不同. 向量量化攻击要求含水印图像的大小是相同的,而上述攻击中含水印图像的大小既可以相同,也可以不同;

(3) 选择近似图像块的范围不同. 向量量化攻击只能在"等价图像块"[7]类中选择,而上述攻击可在所有含水印图像块中选择.

(4) 可以伪造任意大小的含水印图像. 该攻击伪造的含水印图像即可以和含水印图像的大小相同,也可以不同. 例如可以通过剪切、修改等方法伪造不同大小的图像.

因此,与向量量化攻击相比,同步伪造攻击可选含水印图像且图像块的范围更大,攻击力度更强,利用较少的水印图像就可以伪造出质量较好的伪造含水印图像. 尤其需要指出的是:该攻击可以任意修改图像的大小,通过剪切和拼贴操作伪造的含水印图像与合法的含水印图像有相同的视觉效果.

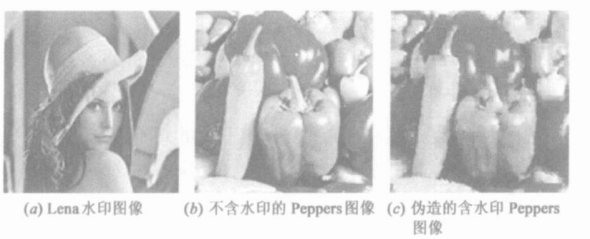
3.2　仿真结果

以灰度图像为例,利用文献[5]算法和相同密钥生成不同大小的含水印图像,在不同条件下对其实施"同步伪造攻击"的部分仿真结果如下.

图 1 为 Peppers 图像及其伪造的含水印图像. (a) 含水印的 Lena 图像,(b) 为不含水印的 Peppers 图像,这两幅图像的内容互不相关. 利用含水印的 Lena 图像,通过同步伪造攻击得到的伪造含水印 Peppers 图像如图(c)所示,该图像与原始 Peppers 图像的峰值信噪比为 22.8888dB. 虽然利用一幅内容与之无关的含水印图像得到的伪造图像质量不是很高,但它基本表达了要伪造图像的意义并能通过认证,充分显示了该攻击的伪造能力及可能带来的安全隐患.

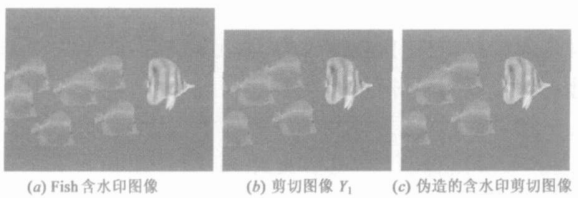
图 2 为剪切图像及其伪造含水印图像. (a) 为 Fish

含水印图像,其大小为 304 ×400;以图像块为单位得到的剪切图像 Y_1 = Fish(41:304,81:400) 如图(b)所示;(c)为利用 Fish 含水印图像得到的 Y_1 的伪造含水印图像,与 Y_1 的峰值信噪比为 51.1645dB. 这是因为剪切图像 Y_1 中任意图像块内容(高 7 位)均可以在(a)中找到完全相同的图像块,因此伪造的含水印图像(c)与(b)的高 7 位完全相同,即伪造含水印图像与真实的含水印图像的质量相同.



(a) Lena 水印图像 (b) 不含水印的 Peppers 图像 (c) 伪造的含水印 Peppers 图像

图 1 Peppers 图像及其伪造的含水印图像



(a) Fish 含水印图像 (b) 剪切图像 Y_1 (c) 伪造的含水印剪切图像

图 2 剪切图像及其伪造含水印造图像

(a)Girl1 水印图像 (b)Girl2 水印图像 (c) Y_1 的伪造含水印图像 (d) Y_2 的伪造含水印图像

图 3 拼贴图像的伪造含水印图像

图 3 为拼贴图像的伪造含水印图像. (a)和(b)分别为 Girl1 和 Girl2 的含水印图像;利用 Photoshop 编辑软件将 Gilr1 的头部替换为 Grl2 的头部,得到的拼贴图像记为 Y_1,利用含水印的 Grl1 和 Grl2 图像得到 Y_1 的伪造含水印图像如图(c)所示,与 Y_1 的峰值信噪比为 34.8220dB. 从图中可以看出,伪造含水印图像仅是篡改边界存在方块效应,伪造质量较好. 将伪造含水印图像(c)的 25:224 列和含水印 Girl1 的 41:240 列(以图像块为单位)拼贴得到的一幅大小为 256 ×400 的拼贴图像 Y_2,利用(a)和(c)通过同步伪造攻击得到 Y_2 的伪造含水印图像如图(d)所示,与 Y_2 的峰值信噪比为 51.1623 dB. 在此条件下也能够得到高质量的伪造含水印图像.

由上述仿真结果可以看出,如果说质量不高的伪造含水印图像有可能通过视角效果判定其为不真实的,那么以含水印图像块为单位拼贴/剪切得到的伪造含水印图像,无论从视觉质量还是认证结果都无法得到伪造或篡改证据. 因此该攻击严重影响自嵌入水印算法的可信性. 也就是说,通过该算法认证的高质量图像也未必真实,从而使算法存在严重的安全隐患.

4 抵抗同步伪造攻击的策略

根据上述分析,为使自嵌入水印算法能够抵抗同步伪造攻击,偏移值密钥空间应该足够大,且必须满足水印嵌入位置 $f(i)$ 的两个必要条件. 下面我们利用混沌系统对初值的极端敏感性和良好的随机性等特点,提出一种基于混沌映射随机生成图像块水印嵌入偏移值 p_i 的方法,并分析该方法的可行性.

偏移值生成方法:

①根据给定的混沌初值 k,利用混沌映射生成长度为 N 的实值混沌序列 $S=(s_1,s_2,\cdots,s_N)$:

$$S=H(k,N) \tag{7}$$

其中,H 表示混沌映射,本文采用文献[3]中的混沌映射:

$$s_{n+1}=(1+0.3(s_{n-1}-1.08)+379s_n^2+1001*z_n^2)\bmod 3 \tag{8}$$

其中,z_n 是任意的混沌序列(本文取 logistic 混沌映射). 该混沌系统初值在(-1.5,1.5)之间时系统具有混沌吸引子,两个初值 s_1 和 s_2 分别记为 $k_{(1)}$ 和 $k_{(2)}$. logistic 混沌映射初值 z_1 在(0,1)之间,该初值记为 $k_{(3)}$,则偏移值密钥 $k=\{k_{(1)},k_{(2)},k_{(3)}\}$.

②采用稳定排序法生成 S 的索引有序序列 A;

对混沌序列 $S=(s_1,s_2,\cdots,s_N)$ 采用稳定排序法生成有序序列 $S_A=(s_{a_1},s_{a_2},\cdots,s_{a_N})$,即 S_A 中的元素满足 $s_{a_1}\le s_{a_2}\le\cdots\le s_{a_N}$,则索引有序序列

$$A=(a_1,\cdots,a_i,\cdots,a_N);$$

③定义第 i 个图像块的水印嵌入偏移值

$$p_i=a_i-i,\quad i=1,2,\cdots,N$$

根据公式(1)可知,基于密钥 k 得到的水印嵌入位置等于混沌序列的索引有序序列,即 $f(i)=a_i$. 下面证明 $f(i)$ 满足两个必要条件:

(1)基于密钥生成的混沌序列 $S=(s_1,s_2,\cdots,s_N)$,元素 s_i 的下标 i 为元素的地址,故长度为 N 的混沌序列的地址范围为 $1,\cdots,N$,即 $a_i\in[1,N]$. 因此按上述算法生成的水印嵌入位置 $f(i)$ 满足:

$$\forall i\in[1,N],f(i)\in[1,N].$$

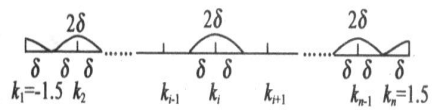
(2)对于序列 $S=(s_1,s_2,\cdots,s_N)$,排序前后序列的元素个数相同且任何一个元素 s_i 在有序序列中有且仅有一个位置,因此 $\forall i,j\in[1,N],i\neq j,s_i$ 和 s_j 在有序序列中的位置 $a_i\neq a_j$,即 $f(i)\neq f(j)$.

基于混沌映射生成偏移值,密钥空间仅依赖于混沌系统对初值的敏感性,与图像的大小(图像块 N)无关. 下面通过实验来考察其密钥空间.

本文采用混沌映射生成水印嵌入位置的密钥 $k=$

$\{k_{(1)}, k_{(2)}, k_{(3)}\}$,其三个子密钥 $k_{(1)}$, $k_{(2)}$ 和 $k_{(3)}$ 分别为不同区间上的实数. 由实数的性质可知任意区间内的实数有无穷多个,由于计算机的位长效应,当 k 变化小于 δ,即 $k \in O(k_i, \delta)$ (如图 4) 时,有 $H(k, N) = H(k_i, N)$. 图 3 中任意 k_i 的 δ 邻域 $O(k_i, \delta)$, $i = 2, 3, \cdots, n-1$ (k_1、k_n 为半邻域) 为产生混沌序列相同的点的集合,该集合可看作密钥的一个取值,因此 *key* 的密钥空间为实数区间 (k_1, k_n) 内互不重叠的 δ 邻域的个数,即:

$$I_k = \frac{k_n - k_1}{2\delta} + 1 \tag{9}$$



为求邻域半径 δ,定义密钥 k 和 $k + \Delta k$ 生成的水印嵌入位置向量的平均距离为:

$$M_{\Delta k}^{k} = \frac{1}{N} \sum_{i=1}^{n} |f_k(i) - f_{k+\Delta k}(i)| \tag{10}$$

其中,Δk 为密钥变化量. 由混沌对初值的敏感性可知,当密钥变化量 Δk 大于 δ 时,水印嵌入位置向量 f_k 和 $f_{k+\Delta k}$ 不同且为随机的;当密钥变化量 Δk 不大于 δ 时,f_k 和 $f_{k+\Delta k}$ 相同,此时 $M_{\Delta k}^{k} = 0$.

为考察整个密钥空间的平均特性,从密钥空间 I_k 中随机选取一组密钥 $k^{(j)} = \{k_{(1)}^{(j)}, k_{(2)}^{(j)}, k_{(3)}^{(j)}\}$ $(j = 1, 2, \cdots, n)$,定义密钥变化量的平均距离 $\overline{M}_{\Delta k}$:

$$\overline{M}_{\Delta k} = \frac{1}{n} \sum_{j=1}^{n} M_{\Delta k}^{k^{(j)}} \tag{11}$$

$k_{(1)}$ 密钥空间:为求 $k_{(1)}$ 密钥的邻域半径 δ_1,取 $N = 1024$,在密钥空间中随机选取三组密钥,密钥个数分别为 100、300 和 1000. 固定密钥 $k_2^{(j)}$ 和 $k_3^{(j)}$,仅改变密钥 $k_{(1)}^{(j)}$,变化量 Δk 分别取 10^{-i}, $i = 1, 2 \cdots, 20$,根据公式 (11) 分别计算出这三组密钥的 $\overline{M}_{\Delta k}$,绘制出当 $N = 1024$ 时 $k_{(1)}$ 的 $\Delta k - \overline{M}_{\Delta k}$ 三条曲线,如图 5 上部所示,三条曲线分别用“ * ”、“ — ”和“ + ”显示. 图中横坐标为密钥 $k_{(1)}$ 变化量 Δk 的负对数,纵坐标为密钥变化量的平均距离 $\overline{M}_{\Delta k}$. 由图 5 可以看出这三条曲线基本重合,说明本文采用的混沌系统具有良好的随机性.

为验证密钥空间与图像块数目 N 的关系,取 $N = 100$,重复上述实验绘制出当 $N = 100$ 时 $k_{(1)}$ 的 $\Delta k - \overline{M}_{\Delta k}$ 三条曲线,如图 5 (下部) 所示. 无论 N 等于 100 还是等于 1024,由图 5 的 $\Delta k - \overline{M}_{\Delta k}$ 曲线可以看出,当 $\Delta k \geq 10^{-12}$ 时,$\overline{M}_{\Delta k}$ 不等 0,可见算法的密钥空间与图像 N 的个数无关. 密钥 $k_{(1)}$ 的邻域半径 $\delta_1 < 10^{-12}$,由式 (9) 计算可得 $k_{(1)}$ 的密钥空间 $I_1 = \frac{1.5 + 1.5}{2\delta_1} > \frac{3}{2} * 10^{12}$.

按相同的方法对密钥 $k_{(2)}$ 和 $k_{(3)}$ 测试,测试结果分别如图 6 和图 7 所示,按相同的计算方法可得 $k_{(2)}$ 和 $k_{(3)}$ 密钥空间分别为 $I_2 > \frac{3}{2} * 10^{15}$ 和 $I_3 > 10^{16}$. 因此该算法的密钥空间 $I = I_2 * I_3 > \frac{9}{4} * 10^{43}$,且不受图像大小的影响. 可见基于混沌随机生成偏移值的密钥空间比文献 [5] 的密钥空间大得多,有效降低了攻击者利用蛮力攻击估计出偏移值密钥的可能性. 因此,基于混沌随机生成偏移值能有效提高自嵌入水印算法的安全性.

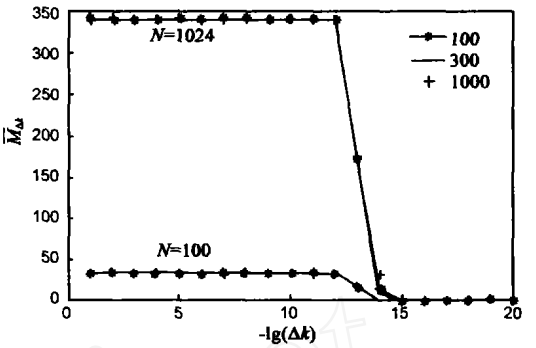


图 5　密钥分量 $k_{(1)}$ 的 Δk-$\overline{M}_{\Delta k}$ 关系曲线

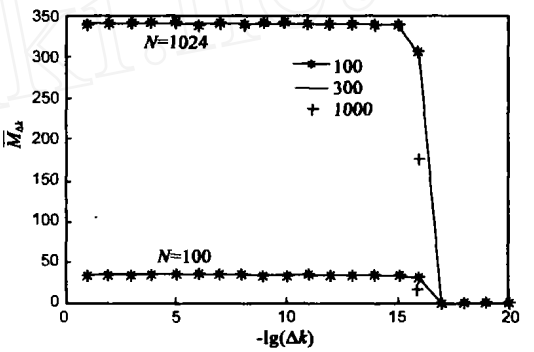


图 6　密钥分量 $k_{(2)}$ 的 Δk-$M_{\Delta k}$ 关系曲线

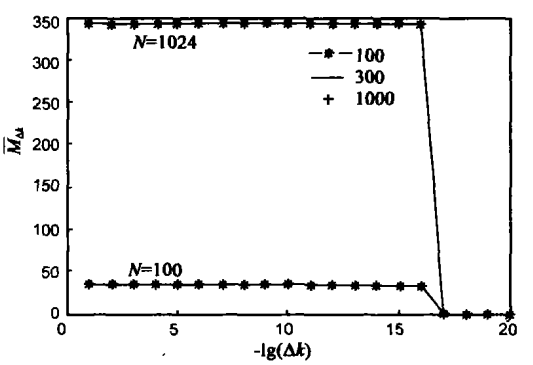


图 7　密钥分量 $k_{(3)}$ 的 Δk-$M_{\Delta k}$ 关系曲线

5　结论

本文首先指出在现有自嵌入水印算法中,水印嵌入偏移值已知会给算法带来严重的安全隐患,并讨论了根据含水印图像得的偏移值密钥的可能性. 为进一步展示偏移值密钥的重要性,本文提出一种已知偏移值条件下实施的同步伪造攻击,并通过实验仿真来验

证该攻击对自嵌入水印算法性能的破坏性.最后,利用混沌映射的对初值的敏感性和伪随机性提出一种具有较大密钥空间的偏移值生成方法.

参考文献:

[1] B B Zhu, M D Swanson, A H Tewfik. When seeing isn't believing[J]. IEEE Signal Processing Magazine, 2004, 21(2): 40 - 49.

[2] F Deguillaume, S oloshynovskiy, T Pun. Secure hybrid robust watermarking resistant against tampering and copy attack[J]. Signal Processing, 2003, 83(10): 2133 - 2170.

[3] 和红杰,张家树,田蕾. 能区分图像或水印篡改的脆弱水印方案[J]. 电子学报, 2005. 33(9): 1557 - 1561.
HE Hong-jie, ZHANG Jia-shu, TIAN Lei. A fragile watermarking scheme with discrimination of tampers on image or watermark[J]. Acta Electronica Sinica, 2005. 33(9): 1557 - 1561. (in Chinese)

[4] J Fridrich, M Goljan. Images with self-Correcting capabilities [C]. IEEE Press, Kobe, Japan: ICIP'99, 1999. 25 - 28.

[5] 张鸿宾,杨成. 图像的自嵌入及篡改的检测和恢复算法[J]. 电子学报 2004. 32(2): 196 - 199.
ZHANG Hong-bin, YANG Cheng, Tamper detection and self-recovery of image using self-embedding[J]. Acta Electronica Sinica, 2004. 32(2): 196 - 199. (in Chinese)

[6] C W Wu. On the design of content-based multimedia authentication systems[J]. IEEE Trans Multimedia, 2002. 4(9): 385 - 393.

[7] M Holliman. N Memon Counterfeiting attacks on oblivious block-wise independent invisible watermarking schemes[J]. IEEE Trans on Image Processing, 2000. 3(9): 432 - 441.

[8] 柯召,孙琦. 数论讲义[M]. 北京:高等教育出版社, 2001.
KE Zhao, SUN Qi. Lectures on the Modern Number Theory [M]. Beijing: By the Higher Education Press, 2001. (in Chinese)

[9] Douglas R Stinson. Cryptography Theory and Practice[M]. By CRC Press, 1995. 3.

作者简介:

和红杰 女,1971 年生于河南宝丰,西南交通大学信号与信息处理省重点实验室博士研究生,主要研究方向为图像处理、数字水印技术、信息隐藏技术等. E-mail: hehojie@sohu.com.

陈 帆 男,1971 年生于河南新蔡,讲师,硕士,主要研究方向:网络技术、信息安全.
E-mail: fchen@home.swjtu.edu.cn