

基于相同图像的抗攻击信息隐藏算法研究

黄 力

(柳州职业技术学院, 广西 柳州 545006)

摘 要: 针对使用相同图像进行信息隐藏时, 可能被攻击泄漏密文, 提出了一种新的信息隐藏算法. 首先对使用同一图像进行两次不同的信息隐藏的情况进行分析, 建立了一套基于概率统计理论的密钥分析算法, 从而可能提取被隐藏的信息. 在此基础上, 设计了一个基于随机序列多存储位置的信息隐藏算法, 该算法在信息隐藏的位置上进行改进, 造成攻击者无法建立有效的概率统计模型, 从而达到提高算法强度的目的. 文中详细阐述了多位置的信息嵌入过程, 并定量地分析了算法抗攻击的性能.

关键词: 信息隐藏; 同一图像; 攻击; 隐藏位置; 算法

中图分类号: TP391

文献标识码: A

文章编号: 1000-7180(2010)04-0186-04

The Anti-attack Algorithm Research of the Information Hiding with the Same Image

HUANG Li

(Liuzhou Vocational & Technical College, Liuzhou 545006, China)

Abstract: The secrete information may be revealed when use information hiding technology with the same image. Aimed to solve this problem, has proposed a new information hiding algorithm. At first, has analyzed hiding different information two times in the identical image. According to it, has established a set of guessing key algorithm based on the probability statistical theory, thus possible to withdraw the information which hidden. Based on this, has designed one hiding algorithm based on the random multi-storage location information, this algorithm makes the improvement in information hiding position, causes the aggressor to be unable to establish the effective probability statistical model, thus achieves that enhancing the algorithm intensity. In the paper elaborated in detail the information embedded in multiposition, and has analyzed the algorithm anti-attack performance in quota.

Key words: information hiding; same image; attack; hiding position; algorithm

1 引言

近年来, 出于信息安全的考虑, 针对图像信息隐藏的算法的研究日益火热, 但与此同时, 相关的信息隐藏攻击算法研究也在同步发展^[1-3]. 如何研究具有较高安全性能的信息隐藏算法成为研究者不断努力的目标. 文中通过研究发现, 在进行信息隐藏时所选取的图片来源经常被人忽视. 然而, 如果所选取的原始图片不当, 比如, 使用相同的图片进行两次或两次以上不同信息隐藏的过程. 这样的信息隐藏过程

将蕴涵很大的风险, 攻击者通过收集发送发基于同一幅图像做的不同信息隐藏操作, 有可能恢复出发送方在发送时所使用的密钥和嵌入的隐藏信息. 在信息隐藏的实际应用过程中, 由于原始图像的获取途径的限制, 经常会选取一些图片库作为信息隐藏的载体. 类似这样的图片库其图片容量总是非常有限的, 如果频繁使用信息隐藏操作, 则不可避免会出现在同一幅图像上嵌入不同秘密信息的情况. 因此, 设计信息隐藏算法时, 需要考虑针对使用同一幅图像进行信息隐藏时, 如何避免被攻击者获取被嵌入

收稿日期: 2009-02-16; 修回日期: 2009-05-26

基金项目: 广西教育厅科研基金(200911LX490)

的隐藏信息.

另外,从信息隐藏的安全性分析,目前常用的评价一个信息隐藏算法的安全性主要分两个层面:(1)攻击能否在大量的图像信息中,找到嵌入有隐秘信息的图片,这是信息隐藏的发现过程^[4-5].(2)如果确定了某幅图片中嵌入了秘密信息,能够通过一定的算法计算,将其中所隐藏的秘密信息提取出来^[6-7].这是指信息隐藏内容的提取过程.目前,一个安全强度非常高的信息隐藏算法,不仅能够做到难以被别人检测到是否嵌入有秘密信息,而且还能够公开信息隐藏算法和告知某幅图片嵌有秘密信息的情况下,仍然能够保证其嵌入信息的安全性^[8].文中设计的信息隐藏算法也将力求达到这一目标.

2 信息泄漏原因分析

设计抗同一原始图像攻击的信息隐藏算法之前,首先需要对针对该情况下的可能攻击情况进行分析,找出问题的关键所在,然后才能设计出具有高安全性的信息隐藏算法.

2.1 信息隐藏过程中的条件假设

根据 LSB 信息隐藏算法的设计思想,下面将分析用户嵌入信息的过程:如果用户需要隐藏 L 比特信息,通常的做法是选取一个密钥 K 通过 HASH 算法,生成 L 个信息嵌入位置,然后将待隐藏的 L 比特信息按照 LSB 算法,在每个像素的最低位将信息嵌入.

为了后续分析的方便,此处给出如下符号定义: $c_{k,j}$ 为原始图像中坐标为 (k,j) 的像素点的位图值, $C = \{c_{k,j}\}$ 为采用位图值表示的原始图像.原始图像的尺寸为 $M \times N$. S 为包含秘密信息的图像. k_1, k_2 分别代表两次不同信息隐藏过程中所选取的密钥. m_1, m_2 分别代表两次不同信息隐藏过程中所隐藏的信息. λ_1, λ_2 分别代表两次不同信息隐藏过程中信息的嵌入率. $X = \{x_i\}_{i=1}^n, Y = \{y_i\}_{i=1}^n$ 分别代表两次不同信息隐藏过程中单个位图值($\{x_{k,j}\}, \{y_{k,j}\}, x_{k,j} \in [0, 255], y_{k,j} \in [0, 255], 1 \leq k \leq M, 1 \leq j \leq N, n = M \times N$)的向量表示形式. $\xi_{1i} = x_i - F(N_0(x_i)), x_i \in X$ 为图像中的噪声信息.

2.2 信息泄漏原理分析

假设用户使用同一原始图像,做了两次不同的信息隐藏操作,得到了两个不同的含有秘密信息的图像.那么攻击者通过对这两幅图像的对比,可以定

位出所有位图值不相同的位置.计为: $A = \{i \mid x_i \neq y_i\}, 1 \leq i \leq n$.

通过分析可知, x_i, y_i 当中只有一个的值等于 c_i .且此时两幅图像的噪声有如下分布特征:

ξ_{1i} 服从 $N(0, \sigma^2)$ 且 (ξ_{2i} 服从 $N(0, \sigma^2)$ 或 ξ_{2i} 服从 $N(-1, \sigma^2)$);

(ξ_{1i} 服从 $N(1, \sigma^2)$ 或 ξ_{1i} 服从 $N(-1, \sigma^2)$) 且 ξ_{2i} 服从 $N(0, \sigma^2)$.

再定义比特值内的距离为 d_{1i}, d_{2i}

$$d_{1i} = |x_i - F(N_0(x_i))|;$$

$$d_{2i} = |y_i - F(N_0(y_i))|.$$

通过比较 d_{1i} 和 d_{2i} 的大小对 x_i 与 c_i 是否相等进行判断,判断准则如下:

① 若 $d_{1i} \leq d_{2i}$, 表明 x_i 未经过修改,有 $t_{1,i} = 0, t_{2,i} = 1$;

② 若 $d_{1i} > d_{2i}$, 表明 x_i 已经过修改,有 $t_{1,i} = 1, t_{2,i} = 0$;

③ 其余情况 $i \notin A$, 有 $t_{1,i} = 0, t_{2,i} = 0$.

分析可得两幅图像中所有数值不等的位置,这些位置可被认为是隐藏有秘密信息的位置.接着将问题转化成概率中的判别问题.

有 $H_0: x_i \neq c_i, H_1: x_i = c_i$, 将假的当成真的概率为 $P_{w1} = p\{H_0 \mid H_1\}$, 将真的当成假的概率为 $P_{z1} = p\{H_1 \mid H_0\}$, 假设 $s_i^{(\bar{k})}, x_i^{(\bar{k})}, 1 \leq i \leq n, \bar{k} \in K$ 表示用户所使用的隐藏密钥范围,则有:

$$s_i^{(\bar{k})} = \begin{cases} 1, & \text{密钥 } \bar{k} \text{ 产生位置 } i \\ 0, & \text{否则} \end{cases}$$

$$x_i^{(\bar{k})} = \begin{cases} 1, & t_{1,i} = 1 \text{ 且 } s_i^{(\bar{k})} = 1 \\ 0, & \text{否则} \end{cases}$$

再定义一个随机空间 $\zeta_m^{(\bar{k})}, 1 \leq m \leq n$ 且其满足 $\zeta_m^{(\bar{k})} = \sum_{i=1}^m x_i^{(\bar{k})}$. 式中 $\zeta_m^{(\bar{k})}$ 为在截获的图像中出现信息不等的地方,假设由 $\bar{k}$ 能够确定嵌入的秘密信息存放处,那么 $\zeta_m^{(\bar{k})}$ 和 $\bar{k}$ 正好从两个不同途径确定秘密信息隐藏位置,由此,可在所有的图像对比空间中分析 $\bar{k}$ 值.

再定义: $\tilde{H}_0: \bar{k}$ 代表正确 key 值, $\tilde{H}_1: \bar{k}$ 代表错误 key 值. 定义判别门限 T , 遍历所有的 $\bar{k} \in K$, 当 $\zeta_m^{(\bar{k})} \geq T$, 表面 $\tilde{H}_0$ 是需要的结果; 否则 $\zeta_m^{(\bar{k})} < T$, 结果输出为 $\tilde{H}_1$. 对概率分布公式 $P_{w2} = \{\tilde{H}_0 \mid \tilde{H}_1\}$ 、 $P_{z2} = \{\tilde{H}_1 \mid \tilde{H}_0\}$ 计算可知, 只要通过 $P_{w2} = 2^{-|K|}$ 次的运算, 就可能分析出信息隐藏时嵌入秘密信息的 KEY 值.

3 抗攻击信息隐藏算法

3.1 算法思想

通过对上述信息隐藏的攻击过程分析可知,虽然该算法的被攻击成功的条件是发送方使用了同一幅图像进行两次不一样的信息隐藏操作,且两次的信息隐藏后的图像都给攻击者所截获.但是仔细分析可以发现在该算法的攻击过程中,很重要的一点是利用了被嵌入的信息只在图像像素值的最低位,以此作为条件,再进行基于概率的统计分析后得出最终攻击结果.

文中设计的高安全强度的信息隐藏算法,在嵌入信息的位置设定上,打破传统的 LSB 算法,只把秘密信息嵌入在最低比特位的约束.而是将待嵌入的信息比特位分散在最低比特位、最高两个比特位.通过将信息散开在不同位置,从而形成一个信息隐藏链,而不是一个信息隐藏点.以此使得攻击方无法通过基于概率统计的原理进行秘密信息的提取攻击.并且由于信息隐藏位置分散开了,无法有效地建立概率统计分析模型.使得攻击方即使知道发送方做信息隐藏时的具体算法,在不知道密钥 K 的情况下,依然是分析出隐藏的信息.

3.2 算法流程

高安全度的隐藏算法关键是如何将隐藏信息的位置进行随机化,使得攻击者无法通过正常的概率统计等分析手段,获取所隐藏的信息.为了得到随机效果非常的处理方法,文中在设计信息隐藏算法时,采用了一个 m 序列,从数学的统计分析可知, m 序列在数据分布的随机性能非常好,具有良好的随机性,且该随机过程同时具有很强的抗分析能力,即使在值得随机算法的处理过程,攻击者也不能根据多次统计的随机结果,推测其使用的随机种子,也无法推测下一次将会得到的结果.因此,该使用该序列来产生随机数能够得到很好的信息分散的效果.算法设计中,将 m 序列作为图像信息中存储位置的选取函数,根据 m 序列的输出结果,确定秘密信息隐藏时的嵌入位置,然后再根据嵌入位置,对信息进行修改完成秘密信息的嵌入过程.整个信息隐藏算法的流程如图 1 所示.

在该算法流程中,首先进行参数的初始化,获取待嵌入的信息值、密钥信息、嵌入的信息长度和原始图像、 m 序列初始状态等参数.然后计算信息嵌入时的步长,并根据初始状态计算 m 序列的数值.再根据 m 序列的状态值执行信息嵌入的不同分支.在

每个分支中根据待隐藏的秘密信息和 m 序列状态值,在原始的图像中做相应的修改.其中 h_i 是待嵌入秘密信息的象素 $\{c_i\}_{i=1}^n$ 数值最高比特, p_i 是待嵌入秘密信息的象素 $\{c_i\}_{i=1}^n$ 数值次高比特.

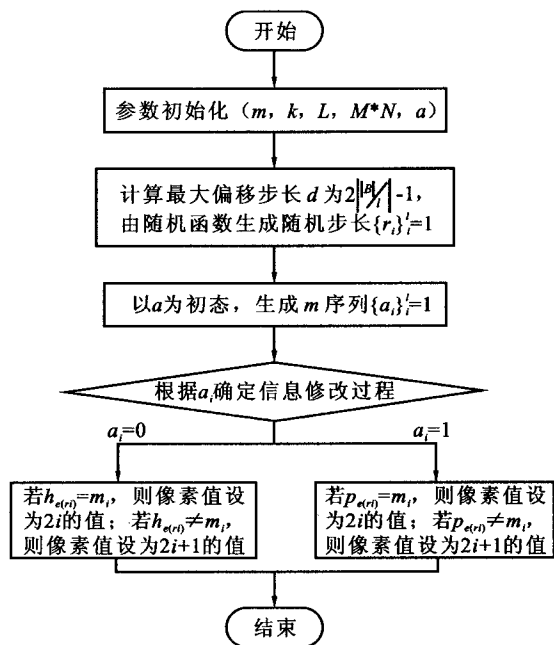


图 1 基于 m 序列的抗攻击信息隐藏算法的流程

4 信息隐藏算法性能分析

为了检验文中设计的信息隐藏算法的安全性,同样还采用前面提出的秘密信息提取算法进行分析.同样假设用户在同一幅图像上做的两次不同的信息隐藏操作结果都被攻击者所截获.攻击者通过对截获的数据进行分析,同样可以获得两幅图像中所有不相同的最低比特位,并定位出这些比特位的详细分布情况.按照前面的分析思想,此时攻击者可以建立基于概率统计模型的秘密信息提取方案,从而破解用户在嵌入秘密信息时所使用的加密密钥.然而,在文中设计的抗统计攻击的信息隐藏算法中,攻击者从最低比特位所提取出来的信息并非是直接嵌入的秘密信息,而是秘密信息嵌入时的相关位置参数.如果要最终正确还原出被嵌入的秘密信息,还需要结合 m 序列值、象素点的最高两个比特的数值一起分析.很遗憾的是,攻击者通过网络信息截获能够拿到两幅嵌有秘密信息的图像,却无法获得 m 序列的初始状态,由此也就无法计算 m 序列每个后续状态值,自然也就无法从被截获的图像中还原出信息隐藏时所使用的密钥信息.由此可见,文中设计的抗统计攻击的信息隐藏算法能够抵抗前面提出的攻击模式,对其中嵌入的秘密信息可以提供较高的安

全保护.

另一方面,对嵌入有秘密信息的图像防止攻击者被检测出来也是算法设计时应该要考虑的问题.文中设计的信息隐藏算法在实现时,最后对信息的修改结果和传统的基于 LSB 隐藏算法的修改结果特性上是一致的.所从抗信息隐藏检测方面而言,文中算法的安全强度和基于 LSB 隐藏算法的强度是一样的.目前对这类信息隐藏有成功的检测算法,如典型的 PoVs 算法.这类算法之所能够成功地检测到图像中是否有隐藏的秘密信息,主要原因是基于 LSB 的信息隐藏算法每次在修改最低比特位时只满足如下交换关系: $0 \leftrightarrow 1, 2 \leftrightarrow 3, \dots, 254 \leftrightarrow 255$,而没有 $1 \leftrightarrow 2, 3 \leftrightarrow 4, \dots, 253 \leftrightarrow 254$ 之间的交换关系.从而使得检测者能够通过对最低比特位的信息特点经过统计后,发现规律,判断在被检测的图像中是否嵌入有秘密信息.

为了提高文中设计的信息隐藏算法抗检测的能力,可以考虑在算法设计的过程中,将这种信息变化规律破坏,使得检测者无法进行统计规律的发现.具体的破坏过程是在算法嵌入信息的过程中:如果准备嵌入信息的像素值为 2^6 或者 2^7 ,则该比特值嵌入信息后加 1;如果准备嵌入信息的像素值为 $2^7 + 1$ 或者 $2^7 + 2^6 + 1$,则该比特值嵌入信息后减 1;其余情况按表 1 所示的方式进行处理.通过这样处理后,得到的结果图像比特值中不再有 $0 \leftrightarrow 1, 2 \leftrightarrow 3, \dots, 254 \leftrightarrow 255$ 这种单一关系,因此,攻击者采用的信息隐藏检测算法也就失效了.

表 1 抗 PoVs 算法分析的 secret 信息嵌入方法

原比特值	嵌入秘密信息后的比特值			
	$a_i = 0$		$a_i = 1$	
	$h_e(r_i) = m_i$	$h_e(r_i) \neq m_i$	$p_e(r_i) = m_i$	$p_e(r_i) \neq m_i$
$2i$	$2i$	$2i + 1$	$2i$	$2i - 1$
$2i + 1$	$2i + 2$	$2i + 1$	$2i$	$2i + 1$

5 结束语

通过分析可知,采用相同的基于 LSB 的信息隐藏算法,在同一幅图像中进行两次或两次不同的信息隐藏操作.即使所设计的 LSB 算法本身是安全的,但攻击者仍然可以通过收集所有采用同一图像嵌入秘密信息的数据,在此基础上,建立基于概率统计模型的攻击方案,并从理论分析上可以证明该攻击过程的可行性和有效性.因此,文中针对信息隐藏中这一安全缺陷,提出并设计了一个基于 m 序列的随机存储位置的高安全性信息隐藏算法.该算法即使在信息隐藏过程给泄漏的情况下,依然能够保证被嵌入的信息不被提取,且算法实现流程简单,具有较好的实用效果.

参考文献:

- [1] 刘红翼,王继军,韦月琼,等.一种基于 LSB 的数字图像信息隐藏算法[J].计算机科学,2008(1):100-102.
- [2] 张涛,平西建.基于差分直方图实现 LSB 信息伪装的可靠检测[J].软件学报,2004(1):151-158.
- [3] 肖洪华,黄永峰,朱明方.基于最低有效位和最高有效位替换法的熵理论模型[J].中山大学学报:自然科学版,2004,43(z2):42-46.
- [4] 鲁宏伟,万宝平.一种基于 BMP 图像的秘密信息隐藏算法[J].武汉理工大学学报,2006,28(6):71-77.
- [5] 王璠.BMP 图像信息隐藏新型检测方法研究[J].计算机安全,2007(9):58-62.
- [6] 吴永,刘祥伟,沈彬.图像最低有效位信息隐藏分析[J].计算机工程与设计,2007,28(2):116-121.
- [7] 蒋翠玲,林家骏,庞毅林.基于数字图像隐写容量的分析[J].微电子学与计算机,2008,25(4):72-76.
- [8] 黄继风.数字图像隐秘信息的检测与定位算法[J].微电子学与计算机,2007,24(7):98-101.

作者简介:

黄力男,(1968-),硕士,副教授.研究方向为网络安全.