对 RS 攻击的分析及抗 RS 攻击的隐写算法*

吕述望, 陈庆元, 刘振华, 翟 卫东
(中国科学技术大学研究生院//信息安全国家重点实验室, 北京 100039)

摘 要: 基于位图的隐写分析方法较有效的检测算法是 RS 法, 其检测率达到 0.05 bpp。通过对 RS 法的分析, 提出一种针对该方法的隐写算法, 使隐藏率达到 0.2 bpp。

关键词: 隐写术; 隐写分析; RS 攻击
中图分类号: TP391 文献标识码: A 文章编号: 0529-6579 (2004) S2-0005-05

隐写术是一种隐蔽通信技术。隐写术的目的是避免引起对被隐藏消息传送行为的怀疑, 也就是赋予信息隐形性。对一个隐写算法评价的主要指标是其安全性。隐写算法的安全性取决于载体, 嵌入秘密消息的量, 嵌入的算法三方面的综合。

在密码算法的设计中, 其安全性一方面取决于对算法的理论分析, 另一方面取决于该密码算法抗已有的密码分析方法的能力。在实践中, 后者比前者更为设计者所重视。将这种抵抗攻击能力的设计思想应用于隐写算法的设计, 将对隐写算法的安全性的评价更具体, 且容易量化。我们借鉴密码算法的一些设计准则, 提出隐写算法的设计原则如下: ① 对载体特性进行充分的分析; ② 结合载体特性估计秘密消息的嵌入量; ③ 秘密消息的提取不需要载体数据; ④ 充分掌握当前的隐写分析算法, 并对其检测原理深入分析; ⑤ 依据第 1、2、3 条, 设计隐写算法。以隐写算法公开为前提, 使其安全性建立在载体特性, 嵌入量和隐写密钥三方面上。

目前基于位图的隐写分析方法较有效的检测算法是 RS 攻击算法, 其检测率达到 0.05 bpp。本文通过对 RS 攻击的分析, 运用以上提出的设计原则, 提出一种针对该方法的隐写算法, 使隐藏率达到 0.2 bpp。

1 RS 攻击

Fridrich[1,2] 提出了一种无损数据嵌入 (lossless data embedding) 方法。基于此方法, 她提出了 RS 隐写分析法 (RS steganalysis)[1] 用于检测灰度或彩色图像中的 LSB 嵌入。对 LSB 的随机化会降低 LSB 位平面的无损容量 (lossless capacity), 而且 LSB 嵌入的位平面数的不同对无损容量的影响也是不同的。因此可以认为无损容量是 LSB 位平面随机度的测度。对于大多数图像, LSB 平面基本上是随机的。但是无损容量反映了这样的性质: LSB 平面与其他比特平面是相关的。这种关系是非线性的, 无损容量是这种关系的精确的测度。

将图像分割为不相交的组, 每个组由 n 个相邻的像素构成 $G=(x_1, \cdots, x_n)$, 若图像为 8 位灰度图, 像素值 x_i 属于 $P=\{0, 1, \cdots, 255\}$。定义一个判决函数 f, 对于任意组 G, 判决函数均满足 $f(x_1, \cdots, x_n) \in R$。利用判决函数可以得到每一组的像素的随机程度。$G$ 中随机度越大即噪声越大, f 的值就越大。例如可以取 f 如下:

$$f(G) = f(x_1, x_2, \cdots, x_n) = \sum_{i=1}^{n-1} |x_{i+1} - x_i|$$

这里判决函数不是惟一的。定义可逆算子 F, F 是由 128 个 2 一轮换构成的置换,

$$F=(0,1)(2,3), \cdots, (252,253)(254,255),$$
$$F^2(x) = F(F(x)) = x, x \in P$$

定义 $F_1 = F$, $F_{-1} = (-1, 0)(1, 2), \cdots, (253, 254)(255, 256)$, 则 $F_{-1}(x) = F_1(x+1) - 1, x \in P$。利用判决函数 f 和置换 F 定义三种类型的像素组 R, S 和 U:

Regular Groups: $G \in R \Leftrightarrow f(F(G)) > f(G)$
Singular Groups $G \in S \Leftrightarrow f(F(G)) < f(G)$
Unusable Groups: $G \in U \Leftrightarrow f(F(G)) = f(G)$

对于 G 中不同的像素应用不同的置换, 因此引入掩码 (mask), 掩码也是 n 元数组, 每个元素

* 收稿日期: 2004-09-11
基金项目: 国家重点基础研究规划 (973) 资助项目 (TG1999035804)
作者简介: 吕述望 (1941 年生), 男, 研究员; E-mail: swlu@uste.edu.cn

可以为－1，0，1 中的任意一个。因而定义的置换 $F(G)$ 为：$(F_{M(1)}(x_1), \cdots, F_{M(n)}(x_n))$。

把掩码为 M 的 R 群（regular groups）的个数与整个图像的组数的百分比定义为 R_M，掩码为 M 的 S 群（singular groups）的个数与整个图像的组数的百分比定义为 S_M，相应的对于掩码－M，定义为 R_{-M}和 S_{-M}，且 R_M、S_M、R_{-M}和 S_{-M}满足：$R_M \cong R_{-M}$和 $S_M \cong S_{-M}$。当有秘密消息嵌入时 R_M 和 R_{-M}，S_M 和 S_{-M}的差将增大，当图像的所有像素都嵌入了秘密消息后，LSB 改变了 50%，则 $R_M \cong S_M$。

2 RS 攻击的基本原理及其分析

文献 [3] 指出自然图像的像素点与其相邻像素的差值近似服从广义高斯分布，广义高斯分布的概率密度函数为：

$$f(x) = \left[2\sigma p^{1/p}\Gamma\left(1+\frac{1}{p}\right)\right]^{-1}\exp\left(-\frac{1}{p}\left|\frac{x-a}{\sigma}\right|^{p}\right)$$

其中，$\Gamma(1+1/p)$ 为伽玛函数，$p>0$，$\sigma>0$，$x \in R$。广义高斯分布是一个函数分布族，拉普拉斯分布和高斯分布只是 $p=1$ 和 $p=2$ 时的特例。其中 p 为形状因子，对于自然图像，一般在 0.25 ~ 1.5 之间。图 1 给出了 512×512 标准灰度 Lenna 图像和它的相邻像素的差值的直方图。由图 1 可以看出，相邻像素的差值的直方图关于 $x=0$ 对称。图像的相邻像素的差值为 0，1 和－1 的概率较高，着表明 LSB 平面与其他的 7 个位平面之间的有很强的相关性。

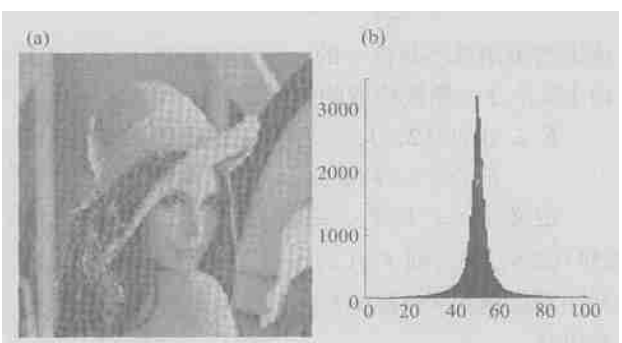

图 1 Lenna 图像（a）与相邻像素的差值（b）

Fig. 1 Lenna (a) and the right is the difference image histogram of "Lenna" (b)

RS 攻击是基于自然图像的像素点与其相邻像素的差值近似服从广义高斯分布特性，对 LSB 嵌入的隐写算法进行检测攻击。

下面我们将对 RS 攻击做进一步的分析。为了简化叙述，假设 $G=(x_1, x_2)$。

2.1 证明 $R_M \cong R_{-M}$ 和 $S_M \cong S_{-M}$

（1）取 $M=(0,1)$，则

$$F_M = (F_0, F_1), F_{-M} = (F_0, F_{-1})$$

表 1 $M=(0,1)$ 时，G 的属性

Tab. 1 The property of G with $M=(0,1)$

情况	$G=(x_1, x_2)$	$M=(0,1)$ 时 G 的属性	$-M=(0,-1)$ 时 G 的属性
1	$\lvert x_1-x_2\rvert=0$	$G \in R$	$G \in R$
2	$\lvert x_1-x_2\rvert=1, x_1 \neq x_2 \oplus 1$	$G \in R$	$G \in S$
3	$\lvert x_1-x_2\rvert>1, x_1<x_2, x_2$ 为偶数	$G \in R$	$G \in S$
4	$\lvert x_1-x_2\rvert>1, x_1>x_2, x_2$ 为奇数	$G \in R$	$G \in S$
5	$\lvert x_1-x_2\rvert=1, x_1=x_2 \oplus 1$	$G \in S$	$G \in R$
6	$\lvert x_1-x_2\rvert>1, x_1<x_2, x_2$ 为奇数	$G \in S$	$G \in R$
7	$\lvert x_1-x_2\rvert>1, x_1>x_2, x_2$ 为偶数	$G \in S$	$G \in R$

一般地，对于一幅图像，$G=(x_1, x_2)$ 的各种情况出现的概率，由对称性可知，情况 2 与情况 5 出现的概率基本相等，情况 3 与情况 6 出现的概率基本相等，情况 4 与情况 7 出现的概率基本相等，因此 $R_M \cong R_{-M}$ 和 $S_M \cong S_{-M}$。

（2）取 $M=(1,0)$，则

$$F_M = (F_1, F_0), F_{-M} = (F_{-1}, F_0)$$

表 2 $M=(1,0)$ 时，G 的属性

Tab. 2 The property of G with $M=(1,0)$

情况	$G=(x_1, x_2)$	$M=(1,0)$ 时 G 的属性	$-M=(-1,0)$ 时 G 的属性
1	$\lvert x_1-x_2\rvert=0$	$G \in R$	$G \in R$
2	$\lvert x_1-x_2\rvert=1, x_1 \neq x_2 \oplus 1$	$G \in R$	$G \in S$
3	$\lvert x_1-x_2\rvert>1, x_1<x_2, x1$ 为奇数	$G \in R$	$G \in S$
4	$\lvert x_1-x_2\rvert>1, x_1>x_2, x_1$ 为偶数	$G \in R$	$G \in S$
5	$\lvert x_1-x_2\rvert=1, x_1=x_2 \oplus 1$	$G \in S$	$G \in R$
6	$\lvert x_1-x_2\rvert>1, x_1<x_2, x_1$ 为偶数	$G \in S$	$G \in R$
7	$\lvert x_1-x_2\rvert>1, x_1>x_2, x_1$ 为奇数	$G \in S$	$G \in R$

一般地，对于一幅图像，$G=(x_1, x_2)$ 的各种情况出现的概率，由对称性可知，情况 2 与情况 5 出现的概率基本相等，情况 3 与情况 6 出现的概率基本相等，情况 4 与情况 7 出现的概率基本相等，因此 $R_M \cong R_{-M}$ 和 $S_M \cong S_{-M}$。

（3）取 $M=(1,1)$，$F_M=(F_1, F_1)$，

$$F_{-M} = (F_{-1}, F_{-1})$$

表 3　$M=(1,1)$ 时，G 的属性

Tab. 3　The property of G with $M=(1,1)$

情况	$G=(x_1, x_2)$	$M=(1,1)$ 时 G 的属性	$-M=(-1,-1)$ 时 G 的属性
1	$\lvert x_1-x_2\rvert$ 为偶数	$G\in U$	$G\in U$
2	$\lvert x_1-x_2\rvert=1$, $x1\neq x_2\oplus 1$	$G\in R$	$G\in U$
3	$\lvert x_1-x_2\rvert=1$, $x_1=x_2\oplus 1$	$G\in U$	$G\in R$
4	$\lvert x_1-x_2\rvert>1$ 为奇数，$x_1<x_2$，x_2 为奇数	$G\in S$	$G\in R$
5	$\lvert x_1-x_2\rvert>1$ 为奇数，$x_1<x_2$，x_2 为偶数	$G\in R$	$G\in S$
6	$\lvert x_1-x_2\rvert>1$ 为奇数，$x_1>x_2$，x_1 为奇数	$G\in S$	$G\in R$
7	$\lvert x_1-x_2\rvert>1$ 为奇数，$x_1>x_2$，x_1 为偶数	$G\in R$	$G\in S$

一般地，对于一幅图像，$G=(x_1, x_2)$ 的各种情况出现的概率，由对称性可知，情况 2 与情况 3 出现的概率基本相等，情况 4 与情况 5 出现的概率基本相等，情况 6 与情况 7 出现的概率基本相等，因此 $R_M \backsimeq R_{-M}$ 和 $SM \backsimeq S_{-M}$。

综上所述，对任意的 $M=(m_1, m_2)$，有 $R_M \backsimeq R_{-M}$ 和 $S_M \backsimeq S_{-M}$。

2.2　嵌入秘密消息后的 R_M，R_{-M}和S_M，S_{-M}

LSB 嵌入图像的像素中，嵌入率为 p，$q=1-p$，使得像素值有可能发生改变，由于秘密消息的随机性可知，像素值发生改变的概率是 $p/2$。

若像素值是偶数，LSB=0，嵌入 0 bit，像素值不变；若像素值是偶数，LSB=0，嵌入 1 bit，像素值改变，值增加 1；若像素值是奇数，LSB=1，嵌入 0 bit，像素值改变，值减少 1；若像素值是奇数，LSB=1，嵌入 1 bit，像素值不变。

取 $M=(1,0)$ 或 $M=(0,1)$，$F_M=(F_1, F_0)$，$F_{-M}=(F_{-1}, F_0)$ 见表 1, 2，情况 1 出现的概率为 P_1，情况 2 出现的概率为 P_2，情况 5 出现的概率为 P_5，$P_2=P_5$，情况 3, 4 出现的概率为 P_{34}，情况 6, 7 出现的概率为 P_{67}，$P_{34}=P_{67}$，那么嵌入秘密消息后 G 的属性将发生改变：

情况 1：

$$P\{R\to R\}=[1/2+q^2/2]P_1,$$
$$P\{R\to S\}=[1/2-q^2/2]P_1$$

情况 5：

$$P\{S\to S\}=[1/2+q^2/2]P_5,$$
$$P\{S\to R\}=[1/2-q^2/2]P_5$$

情况 2：

$$P\{R\to R\}=[1/2+q^2/2+qp/2]P_2,$$
$$P\{R\to S\}=[p^2/2+qp/2]P_2$$

情况 3，4：

$$P\{R\to R\}=[1/2+q^2/2+qp/2]P_{34},$$
$$P\{R\to S\}=[p^2/2+qp/2]P_{34}$$

情况 6，7：

$$P\{S\to S\}=[1/2+q^2/2+qp/2]P_{67},$$
$$P\{S\to R\}=[p^2/2+qp/2]P_{67}$$

那么由表 1，2 可知 R_M 减少，R_{-M} 不变和S_M 增加，R_{-M} 不变，$R_M-S_M=q^2P_1+pqP_5$，当嵌入率 $p=1$ 时，$R_M=S_M$。取

$$M=(1,1),\ F_M=(F_1,F_1),$$
$$F_{-M}=(F_{-1},F_{-1})$$

见表 3，情况 1 中 $\lvert x_1-x_2\rvert=0$ 出现的概率为 P_{10}，$\lvert x_1-x_2\rvert\neq 0$ 出现的概率为 P_{11}，情况 2 出现的概率为 P_2，情况 3 出现的概率为 P_3，$P_2=P_3$，情况 4, 6 出现的概率为 P_{46}，情况 5, 7 出现的概率为 P_{57}，$P_{46}=P_{57}$，那么嵌入秘密消息后 G 的属性将发生改变：

情况 1 中

$$\lvert x_1-x_2\rvert=0:\ P\{U\to U\}=P_{10},$$

情况 1 中

$$\lvert x_1-x_2\rvert\neq 0:\ P\{U\to U\}=[1/2+q^2/2]P_{11},$$
$$P\{U\to S\}=[p^2/4+qp/2]P_{11},$$
$$P\{U\to R\}=[p^2/4+qp/2]P_{11},$$

情况 2：

$$P\{R\to R\}=[q^2/2+pq+q^2/4]P_2,$$
$$P\{R\to U\}=[pq+q^2/2]P_2,$$
$$P\{R\to S\}=[q^2/4]P_2,$$

情况 3：

$$P\{U\to U\}=P_3,$$

情况 4，6：

$$P\{S\to S\}=[q^2/2+pq+q^2/4]P_{46},$$
$$P\{S\to U\}=[pq+q^2/2]P_{46},$$
$$P\{S\to R\}=[q^2/4]P_{46},$$

情况 5，7：

$$P\{R\to R\}=[q^2/2+pq+q^2/4]P_{57},$$
$$P\{R\to U\}=[pq+q^2/2]P_{57},$$
$$P\{R\to S\}=[q^2/4]P_{57},$$

那么由表 3 可知 R_M 减少，R_{-M} 增加和S_M 增加，R_{-M} 减少，U_M 和 U_{-M} 不变。$R_M-S_M=(q^2/2+pq)P_2$，当嵌入率 $p=1$ 时，$R_M=S_M$。

综上所述，对任意的 $M=(m_1, m_2)$，LSB 嵌入

图像的像素中，嵌入率为 p 时，R_M 减少，R_{-M} 增加和 S_M 增加，R_{-M} 减少，且 R_M 和 S_M 的差与 p 有关。

3 抗 RS 攻击的隐写算法

通过第 3 节对 RS 攻击的分析，我们发现 RS 攻击对 LSB 随机嵌入比较敏感，因此抗 RS 攻击的方法一般认为有两个：一个是不使用 LSB 随机嵌入，另一个是针对 RS 攻击的特性，选择合适的像素区，用 LSB 随机嵌入。我们选择后一种，即仍采用 LSB 嵌入，但选择合适的像素进行嵌入。

将载体图像分割为不相交的组，每个组由 4 个相邻的像素构成 $G=(x_1, x_2, x_3, x_4)$，掩码 $M=(0, 1, 1, 0)$，则变换为

$$F_M(G)=(F_0(x_1), F_1(x_2), F_1(x_3), F_0(x_4)),$$

$$F_{-M}(G)=(F_0(x_1), F_{-1}(x_2), F_{-1}(x_3), F_0(x_4))$$

判决函数为

$$f(G)=f(x_1, x_2, x_3, x_4)=|x_1-x_2|+|x_2-x_3|+|x_3-x_4|$$

那么 G 在由 F_M, F_{-M} 作用后分类，有 2 种情况：

(1) 对于 F_M，F_{-M}，各自的分类不同；

(2) 对于 F_M，F_{-M} 为同类，即均为 R 类，S 类，U 类。

对于情况（1）：不在 G 中嵌入秘密消息，则在含隐图像中相应的 G 仍属于情况（1）。

对于情况（2）：若均为 R 类，嵌入秘密消息后有 3 种情况：① 不同类；②同类，$R\rightarrow R$；③同类，$R\rightarrow U$。前两种情况不嵌入秘密消息，则 G 在含隐图像中属性不变，若第 3 种情况，则嵌入秘密消息，G 在含隐图像中属性变为同类为 U。

若均为 U 类，嵌入秘密消息后有 3 种情况：① 不同类；②同类，$U\rightarrow R$；③同类，$U\rightarrow U$。情况①不嵌入秘密消息，则在含隐图像中属性不变；情况②，嵌入秘密消息，但嵌入的秘密消息在下一次嵌入中，可能被再次嵌入，G 在含隐图像中属性变为同类为 R；情况③，嵌入秘密消息，G 在含隐图像中属性不变。

综上所述，在含隐图像中只有同为 U 类的 G 中可能含有秘密消息。

嵌入算法：

对载体图像的每个 G 计算是否为情况（2），将所有符合条件的 G 记为一个集合 S。由随机序列（由密钥 K 生成）控制选取嵌入集合 E。对 E 中元素依次嵌入秘密消息，若嵌入后符合嵌入条件，则嵌入，否则使用下一个元素。每个元素为一个 G，即 4 个像素，每个像素最多可嵌入 3 比特。

提取算法：

对含隐图像的每个 G 计算是否为情况（2），将所有符合条件的 G 记为一个集合 S。由随机序列（由密钥 K 生成）控制选取嵌入集合 E。对 E 中元素依次判决是否属于 U 类，若是则提取秘密消息。

该算法的优点是：能够抗 RS 攻击，同时嵌入量比较大（平均最大嵌入量为 0.2 bpp）。缺点是：嵌入量与载体图像的选择相关，同时也与秘密消息有关，嵌入量不易预测，嵌入秘密消息时需要试嵌入。

4 对算法进行 RS 攻击的实验结果

本节我们对上节给出的隐写算法进行 RS 攻击。所选取的载体图像是 432 幅数码相机拍摄的 24 位彩色人物风景照片。其中 53 幅为 640×480，11 幅为 729×503，27 幅为 882×603，341 幅为 1 280×1 024。在选取的 G 上每像素嵌入 3 bit。平均最大嵌入率为 0.240 8 bpp。

嵌入秘密消息后，对含隐图像进行 RS 攻击：

载体图像分割与隐写算法相同时，被检测到 4 个，被检测率 0.93%。载体图像分割与隐写算法不相同，仅向右错一列开始分割时，被检测到 13 个，被检测率 3.01%。载体图像分割与隐写算法不相同，向下错一行，向右错一列开始分割时，被检测到 32 个，被检测率 7.41%。载体图像分割与隐写算法不相同，仅向下错一行开始分割时，被检测到 18 个，被检测率 4.17%。由上述 4 种攻击可知，该算法能够比较好地抵抗 RS 攻击，但对第 3 种攻击稍为敏感。

5 结 语

隐写嵌入算法是以载体特征的研究为基础的，当前的嵌入算法主要是替代信息位，其好处是提取秘密消息不需要原始载体，其弊端是对载体的统计特性影响大。RS 攻击算法就是利用这一原理进行设计的。为了抵抗这类攻击，常用的方法是适当的选取嵌入区，从而降低嵌入量。是否有其他方法能够替代它，而同时又具有它原来的优点，并避免它的缺点？这有待于我们进一步的研究。

参考文献：

[1] FRIDRICH J, GOLJAN M. Practical steganalysis of digital images-state of the art[C]. Security and Watermarking of Multimedia Contents Ⅳ, California, 2002, 4675.

[2] FRIDRICH J, GOLJAN M, DU R. Distortion-free data embedding // Lecture Notes in Computer Science[C]. Berlin: Springer-Verlag, 2001, 2137.

[3] 翟卫东, 刘振华, 陈庆元, 等. 应用于信息隐藏检测中的几个统计学结论 //信息隐藏全国学术研讨会(CIHW2002)论文集[C]. 北京: 机械工业出版社, 2002: 53—67.

Studying RS Steganalysis and a Steganographic Algorithm which Can Resist this Steganalysis

LÜShu wang, CHEN Qing yuan, LIU Zhen hua, ZHAI Wei dong

(State Key Laboratory of Information Security //
Graduate School of University of Science &Technology of China, Beijing 100039, China)

Abstract: At present, RS steganalysis is more effective steganalysis based on bitmap, it indicates that the detecting rate is less than 0.05 bits per sample. With studying RS steganalysis, a steganographic algorithm which can resist this steganalysis was showed, it's safe bit rate is less than 0.2 bits per sample.

Key words: steganography; steganalysis; RS steganalysis