对JSTEG隐写系统的提取攻击

马 宁,肖和立,张卫明,刘文芬,刘九芬

(信息工程大学 信息工程学院,河南 郑州 450002)

摘要:本文基于JPEG图象的连续JSTEG隐写术的统计检测方法,提出一种简单高效的提取方法。实验结果表明,利用这种方法可以得到秘密消息起点和终点的点估计和区间估计,从而可以精确地提取秘密消息。

关键词:信息隐藏;隐写术;提取攻击

中图分类号:TN919.81 **文献标识码**:A **文章编号**:1671-0673(2005)01-0006-04

Extracting Attack to Sequential JSTEG Stegosystems

MA Ning, XIAO He-li, ZHANG Wei-ming, LIU Wen-fen, LIU Jiu-fen

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

Abstract: In this paper, we propose a simple and effective extracting technique based on statistical detecting method of sequential JSTEG for JPEG images. By this approach we can give the point estimations and interval estimations of the begin-point and end-point of the hiding message, and then extract the hidden message accurately.

Key words: information hiding; steganography; extracting attack

1 引言

隐写术是通过将消息嵌入载体,如数字图象,来隐藏秘密通信的存在。载体图象经过微小的改动将秘密消息隐藏其中,就得到载密图象。对于隐写系统最重要的要求是不可检测性:载密图象与载体图象在统计上是不可分的。

事实上,密码分析者更加关心如何将隐藏的消息提取出来。目前关于提取攻击的文献还很少,Chandramouli针对基于扩频通信隐写术的一种特殊情况给出了提取攻击方法[3]。Fridrich等对于基于密钥的隐写算法F5和outguess,利用卡方检验给出了一种提取攻击算法[4]。

本文讨论了对于连续Jsteg隐写系统的提取攻击。对于连续的嵌入算法,提取攻击就是判断嵌入消息的起点和终点。确定起点和终点可以用相同的方法,所以假设消息的起点在系数的第一个可嵌比特位,因此提取攻击可以归结为对消息长度的估计。一些传统的隐写分析方法[6,7,9]不仅可以检测出消息的存在还可以给出对消息长度的估计,但是它们给出的估计对于提取消息还不够精确。本文给的提取攻击方法,简记LR方法,它基于一种统计检测方法。本方法可以准确估计出嵌入消息的起点和终点,并且简单易行、计算高效。

2 JPEG-JSTEG隐写算法和统计隐写分析算法

2.1 JPEG-JSTEG隐写算法

JPEG-JSTEG算法是一种典型的用JPEG格式图象作为载体的隐写算法,由D.Upham提出[8]。离散余弦变换(DCT)系数经过量化后,JPEG-JSTEG

收稿日期:2004-08-30

基金项目:国家自然科学基金重点项目(60172067);国家自然科学基金项目(60133020)。

作者简介:马 宁(1981-),女,山东鱼台人,信息工程大学硕士研究生,主要研究方向为信息安全和信息隐藏。

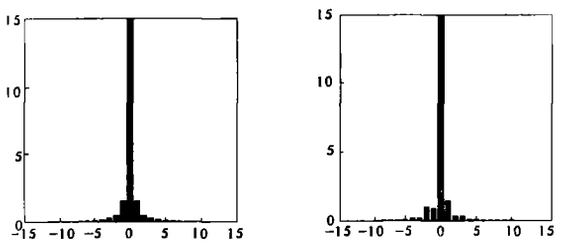
跳过 DC(直流)系数及取值为 0 和 1 的量化系数,用秘密消息比特替换量化系数的最低比特位(LSB)。

$C^N=(c_1,c_2,\cdots,c_N)$表示除去 0,1 和 DC(直流)系数载体量化后的 DCT 系数集合,$M^n=(m_1,m_2,\cdots,m_n)$表示待嵌入的秘密消息比特。选取 C^N 的子集 $S^n=(s_1,s_2,\cdots,s_n)$,对于 S^n 的所有元素 s_i $(1\leqslant i\leqslant n)$的 LSB,用 $m_i(1\leqslant i\leqslant n)$替换。

2.2　统计隐写分析算法

由于加密的消息可以视为伪随机序列,Westfeld[1]指出经过 JSTEG 隐藏加密消息的载密图象,它量化后的 DCT 系数值对中两个值的出现次数都趋向一致,如图 1 所示。

(a) JPEG图象“ school.jpg”



(b) 图(a)量化后DCT系数柱状图　(c) 嵌入20k比特消息后DCT系数柱状图

图 1

基于载体图象与载密图象的统计差异,Westfeld 设计了一种卡方检验判断载密图象中秘密消息的存在:

①随机样本空间的总体是所有 DCT 系数,假设把总体分成 k 类,每个观测值必落在唯一的一个类。不失一般性,仅就偶系数研究;

②嵌入消息后,第 i 类的理论期望 $h_i^*=\dfrac{|h_{2i}+h_{2i+1}|}{2}$;

③随机样本的实际频数是 $h_i^{**}=|h_{2i}|$;

④$\chi^2=\sum_{i=0}^{k}\dfrac{(h_i^{**}-h_i^*)^2}{h_i^*}$是自由度为 $k-1$ 的 χ^2 统计量;

⑤p 是当 h^{**} 和 h^* 分布相同时统计量的概率,

$$p=1-\frac{1}{2^{\frac{k-1}{2}}\Gamma\left(\frac{k-1}{2}\right)}\int_0^{\chi^2_{k-1}}e^{-\frac{x}{2}}x^{\frac{k-1}{2}-1}\mathrm{d}x。$$

这种方法一般可以用来攻击 LSB 嵌入机制的隐写术。

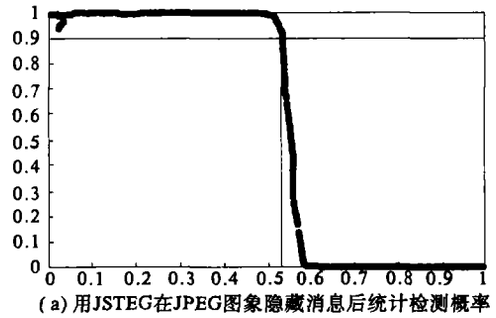
3　提取方法描述

对于连续 JSTEG,提取攻击就是估计消息起点和终点的位置。确定起点和终点可以用相同的方法,故假设消息的起点在系数的第一个可嵌比特位,因此提取攻击可以归结为对消息长度的估计。

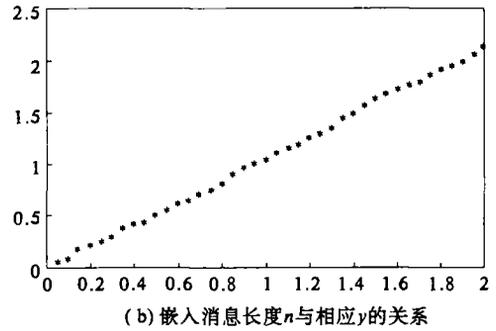
3.1　对消息长度的估计归结为线性回归问题

如图 2(a) 所示,p 值在一个小区间迅速由 1 下降到 0,假设 p 值在这一区间内是严格单调减的,因此当 p 取值 0.9 时有唯一的样本长度值与之对应,记作 y。

图 2(a) 是图 1(a) 用 JSTEG 嵌入 20kB 数据后,做所述统计检验所得图示,横轴表示样本大小,纵轴表示 p 的值,图中标示“y” 为 p 取值 0.9 时唯一对应的样本长度值。图 2(b) 是当改变嵌入消息长度 n 时,y 也相应改变,横轴表示 n 的值,纵轴表示对应的 y 的值,n 与 y 是线性关系。因此可以假设对于载体图象有线性模型 $n=ay+b,a,b\in R$。



(a) 用JSTEG在JPEG图象隐藏消息后统计检测概率



(b) 嵌入消息长度n与相应y的关系

图 2

假设有一通过 JSTEG 算法得到的载密图象,内嵌长度为 n_0 的消息,用第二部分所述方法做统计检验得到对应 p 值为 0.9 的样本长度,记做 y_0。为了估计 n_0,需要得到载体嵌入消息长度与样本长度的线性模型。将一条随机序列按照同样的算法

嵌入载密图象，嵌入长度为 n，且 $n \geqslant y_0$，即用随机序列覆盖了原来的消息，同样经过统计检验得到对应 p 值为0.9的样本长度，记做 y。若随机序列的嵌入长度取一组值，则得到相应的一组对应 p 值为0.9的样本长度。用这些数据作线性回归得到载体的线性模型 $n = ay + b$，已知 y_0，可以计算出 n_0 的估计值 $\hat{n}_0 = ay_0 + b$。

具体步骤如下：

(1) $C^N = (c_1, c_2, \cdots, c_N)$ 表示载密图象量化后DCT系数除去0,1和DC系数的集合，$M^{n_0} = (m_1, m_2, \cdots, m_{n_0})$ 表示嵌入图象的秘密消息比特。用第二部分所述方法做统计检验得到对应 p 值为0.9的样本长度，记做 y_0；

(2) 对于随机序列 $M_1^l = (M_{1,1}, M_{1,2}, \cdots, M_{1,l})$，将 M_1^l 的子集 $M_1^{l_i} = (M_{1,1}, M_{1,2}, \cdots, M_{1,l_i})$，$y_0 \leqslant l_i \leqslant \delta N, 1 \leqslant i \leqslant t$，通过JSTEG嵌入到载密图象，其中 δ 是控制因子。因为 $l_i \geqslant y$，所以 $M_1^{l_i}$ 覆盖了原消息 M^n。重复第二部分的统计检验，得到一组 $y_{1,i}(1 \leqslant i \leqslant t)$。记 $\boldsymbol{X} = [n_{1,1} \cdots n_{1,t}]^T$，$\boldsymbol{Y} = \begin{bmatrix} y_{1,1} & \cdots & y_{1,t} \\ 1 & \cdots & 1 \end{bmatrix}^T$，由最小二乘法作线性回归：$\begin{bmatrix} a_1 \\ b_1 \end{bmatrix} = (\boldsymbol{Y}^T\boldsymbol{Y})^{-1}\boldsymbol{Y}^T\boldsymbol{X}$，于是有 $n_0 = a_1 y_0 + b_1$。

方法实验结果

本文实验所用数字图象下载自因特网，嵌入的秘密消息是文本文件加密后所得密文。

实验步骤：① 计算JPEG图象用JSTEG隐写算法的隐藏容量；② 用JSTEG算法嵌入不同长度的消息，得到载密图象；③ 用3.1所述方法提取步骤② 所得载密图象中的秘密消息。

实验所用载体图象“school.jpg”(图1所示)用JSTEG隐写算法的隐藏容量为42334比特，表中 n_0 表示隐藏消息的长度，n_0 取值为4000比特(约占隐藏容量的10%)，16000比特(约占隐藏容量的40%)。

表1 消息长度为4000比特，用3.1所述方法提取消息(比特)

j	$n_0 = 4000, y_0 = 4870$				
	1	2	3	4	5
$n_{1,j}$	4900	5200	5500	5800	6100
$y_{1,j}$	6005	6495	6965	7090	7340
j	$n_0 = 4000, y_0 = 4870$				
	6	7	8	9	10
$n_{1,j}$	6400	6700	7000	7300	7600
$y_{1,j}$	7775	7995	8210	9040	9035

表2 消息长度为16000比特，用3.1所述方法提取消息(比特)

j	$n_0 = 16000, y_0 = 18100$				
	1	2	3	4	5
$n_{1,j}$	18200	18500	18800	19100	19400
$y_{1,j}$	20840	10900	21420	21890	22360
j	$n_0 = 16000, y_0 = 18100$				
	6	7	8	9	10
$n_{1,j}$	19700	20000	20300	20600	20600
$y_{1,j}$	22660	23110	23210	23550	23280

嵌入载密图象的随机序列长度 $n_{1,1}$ 和相应的 $y_{1,1}, 1 \leqslant i \leqslant 10$，(见表1、表2)。用表中数据作线性回归得到线性模型 $n_1 = a_1 y_0 + b_1$，各参数及消息长度估计值见表3。

表3 线性模型参数及消息长度的估计值

n_0(bits)	y_0(bits)	a_1	b_1	n_1(bits)
4000	4870	0.88916	-503.19	3861
16000	18100	0.83523	861	15994

3.2 误差分析

用不同的随机序列重复3.1所述方法所得载体的线性模型也不同，因此得到不同的对消息长度 n_0 的估计值，记作 $n_i, i \geqslant 1$，并且经检验这些估计值 $n_i, i \geqslant 1$ 与 n_0 来自同一正态分布。序列选取的随机性导致了估计值的随机误差。因此，可以用上述正态分布的期望作为 n_0 的点估计，并且可以得到 n_0 的置信水平为 $1-\alpha$ 的区间估计。所以应取一组随机序列 $M_j^l = (M_{j,1}, M_{j,2}, \cdots, M_{j,l}), 1 \leqslant j \leqslant r$，在3.1所述方法后添加以下步骤：

(3) 对于序列 $M_j^l = (M_{j,1}, M_{j,2}, \cdots, M_{j,l})$，$2 \leqslant j \leqslant r$，重复3.1中步骤(2)，得到 n_0 的一组估计值 $n_i; 2 \leqslant j \leqslant r$；

(4) 假设 $n_1, n_2, \cdots, n_r$ 是来自同一正态分布 $N(mean, \sigma^2)$ 的样本，参数 $mean$ 和 σ^2 的极大似然估计为：

$$mean = \frac{1}{t}\sum_{i=1}^{t} n_i,$$

$$\sigma^2 = \frac{1}{t}\sum_{i=1}^{t}(n_i - mean)^2;$$

(5) 因为 $n_0 \sim N(mean, \sigma^2)$，$mean$ 是 n_0 的点估计；由于 $\frac{n_0 - mean}{\sigma^2}$ 是标准正态分布，给定置信水平 $1-\alpha$，有 $p = \left\{\left|\frac{n_0 - mean}{\sigma^2}\right| \leqslant \mu_\alpha\right\} = 1 - \alpha$ 可得到 μ_α，因此随机区间 $[mean - \mu_\alpha\sigma^2, mean + \sigma^2]$ 是 n_0 的置信水平为 $1-\alpha$ 的区间估计。

因为本提取攻击的重点是线性回归(linear regression)，简记为“LR提取攻击”。

方法实验结果如下:

实验用图"school.jpg"(图 1),JSTEG 隐藏容量 42334 比特,选取 12 条随机序列,隐藏消息长度 n_0 取值为 4000 比特(约占隐藏容量的 10%)、16000 比特(约占隐藏容量的 40%)、30000 比特(约占隐藏容量的 70%)。$n_i(1 \leqslant i \leqslant 12)$ 表示用第 i 条序列得到的估计值。区间估计的置信水平是 0.95。

表 4　选取 12 条随机序列,用 3.2 所述方法提取消息　(比特)

n_0	n_1	n_2	n_3	n_4	n_5	n_6	n_7
4000	3784	3520	4546	4443	4246	4112	3861
16000	15994	16718	16023	16067	15304	16871	15987
30000	29835	30462	30371	30371	29685	30593	29884

n_0	n_8	n_9	n_{10}	n_{11}	n_{12}	*mean*	*interval*
4000	3914	3743	4144	4120	3975	4034	[3846,4222]
16000	16498	16551	15804	15757	15584	16097	[15795,16398]
30000	30258	30237	29976	29921	29950	30127	[29945,30309]

由表 4 可知对消息长度的点估计误差是几十比特,随机区间长度为 400 比特左右以 0.95 的概率包含消息的终点。这个结果对于密码分析是有意义的。

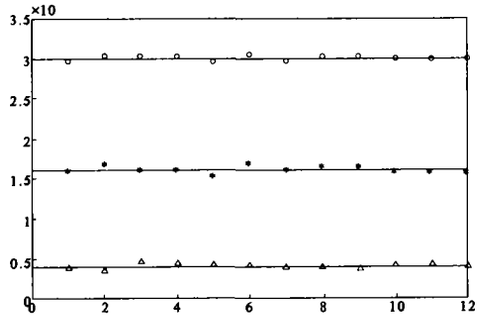


图 3　用 12 条随机序列估计隐藏消息的长度

图 3 中实线表示实际消息的长度,离散点表示每条序列对不同长度消息的点估计值。

4　结论

一些传统的隐写分析方法,如 RS[6],基本集方法(SPA)[7] 和基于差分直方图的方法(DIH)[9],在变换域 LSB 隐写系统不仅可以检测出消息的存在还可以给出对消息长度的估计。但是一般它们只能给出嵌入率,对于密码分析还不够精确。嵌入不同消息长度(隐藏容量的 10%,隐藏容量的 40%,隐藏容量的 70%),本文提取攻击方法(仅列出点估计值)与上述 3 种攻击方法攻击效果对比如表 5 所示。由表 5 看出传统的检测算法在估计长度方面误差较之本文提取攻击方法偏大。

表 5　攻击效果比较　(比特)

n_0	RS	PS	DIH	LR
4000	2104	953	2523	4034
16000	9620	9250	9965	16097
30000	30989	29614	27884	30127

信息隐藏技术在因特网的应用日益广泛,加密的消息可以隐藏在一些载体内在因特网上传播。对于密码分析者,将加密的消息提取出来是首先要面对的问题,而传统隐写分析的检测方法不能解决。本文针对连续的 JSTEG 隐藏算法提出了"LR 提取攻击方法",对于用 JSTEG 隐藏信息的 JPEG 图象可以精确的提出其中的秘密消息。我们今后还将对随机嵌入消息的各种隐写系统的提取攻击作研究。

参考文献:

[1] A Westfeld, A Pfitzmann. Attacks on Steganographic Systems [A]. Proceedings of the third International Workshop on Information Hiding[C]. Dresden, 1999, 8:61 - 76.

[2] A Westfeld. F5-A Steganographic Algorithm: High Capacity Despite Better Steganalysis[A]. Proceedings of the 4th International Workshop on Information Hiding[C]. Pittsburgh, PA, 2001, 4:289 - 302.

[3] R Chandramouli. A mathematical approach to steganalysis [A]. Proceedings of SPIE Security and Watermarking of Multimedia Contents IV[C]. San Jose. CA, 2002, 4:14 - 25.

[4] Fridrich J and Goljan M. Searching for the Stego-key[A]. Proceedings of EISPIE[C]. San Jose, CA, 2004, 1: 288 - 302.

[5] Fridrich J, Goljan M, Du R. Reliable detection of LSB steganography in grayscale and color images[A]. Proceedings of ACM Workshop on Multimedia and Security[C]. Ottawa, Canada, 2001:27 - 30.

[6] Fridrich J, Goljan M, Du R. Detecting LSB steganography in color and gray-scale images[J]. IEEE Multimedia, 2001, 8 (4):22 - 28.

[7] S Dumitrescu, Xiaolin Wu. Steganalysis of LSB embedding in multimedia signals[A]. Proceedings of IEEE International Conference on Multimedia and Expo[C]. Tokyo: 2002, 8:581 - 584.

[8] JPEG-JSTEG-V4[EB/OL]. http://www.funet.fi/pub/crypt/steganography/jpeg-jsteg-v4.diff.gz, May, 2003.

[9] 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,2004,15(1):151 - 158.