基于局部随机性的 YASS 隐写分析方法

雷 雨 杨晓元 潘晓中 郭敦陶

(武警工程学院电子技术系网络与信息安全武警部队重点实验室 西安 710086)

摘 要 提出了一种针对 YASS(Yet Another Steganographic Scheme)隐写算法的专用检测方法. 根据 B 块的大小,分析 8×8 数据嵌入子块的位置,并提取子块 DCT 低频系数局部随机序列的 4 个统计量作为特征,结合 Fisher 线性分类器进行分类. 针对 YASS 及其改进算法的实验表明:此方法能有效检测 YASS 隐写得到的含密图像,相比现有方法,其检测率有明显提高.

关键词 隐写分析;YASS;局部随机性;Fisher 线性分类器

中图法分类号 TP391 **DOI 号**: 10.3724/SP.J.1016.2010.01997

YASS Steganalysis Based on Local Randomness

LEI Yu YANG Xiao-Yuan PAN Xiao-Zhong GUO Dun-Tao

(*Key Laboratory of Network & Information Security of APF, Engineering College of APF, Xi'an* 710086)

Abstract In the paper, a new steganalysis scheme for attacking Yet Another Steganographic Scheme (YASS) is proposed. Based on the size of B-blocks, the locations of the 8×8 embedding sub-blocks are determined. The features are calculated from four statistics of local randomness sequences of the low frequency coefficients in the 8×8 embedding sub-blocks and Fisher linear discriminant is exploited to classify. The experiments on YASS steganography and its further extensions show the method can detect stego images reliably, and the detection accuracy exceeds those of its closest competitors obviously.

Keywords steganalysis; YASS; local randomness; Fisher linear discriminant

1 引 言

隐写分析(steganalysis)是隐写(steganography)技术的反过程,主要是利用信息的嵌入会引起载体数据分布特性或统计特性的改变这一隐写技术固有的弱点,分析各种可能的载体信息,旨在分析、检测和提取隐藏在载体数据中的秘密信息. 它可分为专用隐写分析和通用隐写分析两类. 前者主要是针对某种特定的隐写算法,其效果较好,但是灵活性和可扩展性较差. 后者并不针对某一特定的隐写方法,而是通过构建通用的检测器,对各种隐写算法所产生的含密图像进行检测,其通用性较好,但对某一特定方法的检测效果可能较差.

现有的图像隐写算法按照秘密信息的嵌入位置可分为空域隐写算法和压缩域隐写算法. 由于 JPEG 图像格式的流行,压缩域隐写算法主要集中在 JPEG 压缩域,其可分为 3 类[1]:基于 DCT 量化系数的隐写算法、基于附加信息的隐写算法和基于替换域的隐写算法. 其中第 1 类算法最为常见,其利

收稿日期:2010-08-22.**雷 雨**,男,1987 年生,硕士研究生,主要研究方向为隐写与隐写分析技术. E-mail: ly1a2b3c@163.com. **杨晓元**,男,1959 年生,硕士,教授,主要研究领域为密码学、信息隐藏. **潘晓中**,男,1964 年生,硕士,教授,主要研究领域为信息隐藏. **郭敦陶**,男,1979 年生,硕士,助教,主要研究方向为信息隐藏.

用 DCT 量化系数的冗余来嵌入秘密消息，如 F5[2] 和 MB(Model-Based)算法[3]；第 2 类算法是指隐写过程不能在 JPEG 图像上直接完成，需要借助附加信息，如 PQ(Perturbed Quantization)算法[4]；第 3 类算法是指先在替换域(如 DCT 域、小波域等)中鲁棒地嵌入秘密消息，然后再压缩成 JPEG 图像，由于 JPEG 压缩留下的痕迹很好地掩盖了秘密消息嵌入留下的痕迹，所以该类方法具有高度的隐藏性，其中最具代表性的是 YASS 及其改进算法[5-6].

针对以上隐写算法，人们提出了一系列专用或通用的隐写分析方法. 其中较具代表性的是几个基于 DCT 域特征提取的通用分析方法[7-8]，它们能对 F5、MB 和 PQ 算法进行有效检测，但对 YASS 算法无效. 近年来学者们对 YASS 算法进行了专门研究，Li 等[9]提出了一种 YASS 专用隐写分析方法，文中指出 YASS 算法中 8×8 数据嵌入子块的位置虽然是根据密钥随机选取的，但其随机性仅限于局部而并非全局(限定在 B 块内)，同时嵌入消息时用到的 QIM(Quantization Index Modulation)方法会造成 DCT 系数中 0 值个数的改变，作者利用这两点对 YASS 进行分析，取得了较好的效果. 刘洪等[10]对 Li 的方法进行了改进，文中指出 QIM 嵌入不仅会改变 DCT 系数中 0 值的个数，对其它系数也会造成影响，同时重新选取了 8×8 数据嵌入子块的起始位置，提取 DCT 系数的共生矩阵作为特征，相比 Li 的方法，效果有所提高. Yu 等[11]提出了一种能检测 YASS 的通用分析方法，文中指出虽然 YASS 方法能有效抵抗“校准”攻击，但其秘密消息的嵌入仍会极大地破坏图像的像素和 DCT 系数间的相关性，并分别从空域和 DCT 域提取特征，取得了较好效果. Kodovsky 等[12]沿着 Yu 的思路，分别从空域和 DCT 域提取了各自领域中效果最好的特征，融合得到 1234 维的跨域特征，其检测效果要优于 Yu 的方法，和 Li 的专用分析方法相当.

YASS 算法的核心是利用数据嵌入子块选取的随机性和 QIM 嵌入方式的鲁棒性来掩盖秘密消息嵌入留下的痕迹，以抵抗基于“校准”攻击的 JPEG 图像通用分析方法. 但其 QIM 的嵌入方式会造成载体图像局部随机性的异常，且数据嵌入子块的选取仅是局部随机. 本文利用数据嵌入子块选取的局部随机性，分析数据嵌入子块的位置，并提取子块低频系数局部随机序列的 4 个统计量作为特征，结合 Fisher 线性分类器进行分类，提出了一种针对 YASS 的专用检测方法.

2 YASS 及其改进算法的原理

YASS 是由 Solanki 等在 2007 年的信息隐藏大会上提出的[5]，它是一种能抵抗 JPEG 图像通用分析的新型隐写算法，其具体过程可归纳为以下 7 步：

1. 对要嵌入的秘密消息用具有纠删功能的 RA(Repeat Accumulate)码进行编码.

2. 将给定的图像(空域图像或者 JPEG 图像)以空域表示，然后划分其为连续而不重叠的块，块的大小为 $B\times B$，其中 $B>8$，称这些块为 B 块或者 B-block.

3. 在每个 B-block 中，根据密钥随机地选取一个 8×8 的子块，称为数据嵌入子块或者 E-block.

4. 对 E-block 进行二维 DCT 变换，所得的 DCT 系数除以对应的量化步长，量化步长由嵌入质量因子 QF_h 决定，得到未取整的量化系数.

5. 将编码后的秘密消息以 QIM 的方式嵌入到一些未取整的低频系数(zigzag 扫描后前 19 个 AC 系数)上，也叫候选嵌入系数.

6. 将嵌入数据后的系数乘以对应的量化步长，然后对 E-block 进行二维反 DCT 变换.

7. 对整幅图像进行 JPEG 压缩，其中压缩的质量因子为 QF_a，得到含密图像.

为了提高数据嵌入率，文献[5]提出增大 B-block 边长的方法，即在较大的 B-block 中选取多个 E-block，以提高整幅图像中 E-block 的数量. 例如令 $B=8n+1(n>1)$，则在一个 B-block 中能得到 n^2 个 E-block.

文献[6]中对 YASS 又进行了两方面的改进：第 1 个改进是根据 DCT 系数的方差来调节嵌入量化因子 QF_h 的选取，增加嵌入参数的随机性，以提高安全性. 第 2 个改进是利用重复嵌入的方式替换对秘密消息进行的 RA 编码，以增强嵌入数据的鲁棒性，提高数据嵌入率.

表 1 所示是本文实验用到的 YASS 算法的 6 种隐写情况. 实验中所用图像来自 UCID 图像库[13]，该库中共包括 1338 幅未经过 JPEG 压缩过的 24 位彩色 TIFF 图像，用 Matlab 转化为灰度图像后进行隐写. QF_a 固定为 75，选取 QIM 的量化步长 $\Delta=2$. 由于 Δ 较大，嵌入数据具有较高的鲁棒性，所以这里不再对秘密消息进行 RA 编码，此时库中图像仍保持了极高的嵌入成功率. 图像的嵌入率用 bpp(bit per piexl)表示，图像质量用 PSNR(Peak Signal-to-Noise Ratio)表示.

表 1　YASS 算法的 6 种隐写情况

图像集	QF_h	DCT 系数方差区间	B	重复嵌入的次数	嵌入成功的图像数量	图像的平均嵌入率(bpp)	图像的平均 PSNR/dB
YASS1	50	—	9	0	1298	0.159	27.99
YASS2	50	—	15	1	1311	0.077	31.93
YASS3	50	—	17	0	1283	0.179	27.52
YASS4	50-60-70	$[0,1,2,\infty)$	9	0	1303	0.138	28.73
YASS5	50-60-70	$[0,1,2,\infty)$	15	1	1323	0.069	32.62
YASS6	50-60-70	$[0,1,2,\infty)$	17	0	1292	0.154	28.27

3　针对 YASS 算法的隐写分析

3.1　分析数据嵌入子块的位置

通常情况下，对一个隐写算法进行分析时，首先要找到算法的数据嵌入域，然后在该区域内找到对数据嵌入敏感的特征，这样的特征往往分类效果较好. YASS 的数据嵌入域是由图像所有的 E-block 进行二维 DCT 变换后的系数组成，其中 E-block 的选取由密钥决定，对攻击者来讲是未知的. 但 Li 等[9]指出 E-block 位置的随机性仅限于局部而并非全局的，具体讲就是 E-block 的位置必须限定在 B-block 内，并非随机分布在整幅图像中. 虽然不能确定 E-block 的准确位置，但是可以确定一些可能放置 E-block 的位置以及一些不可能放置 E-block 的位置. 下面分 $8<B<16$ 和 $B=8n+1(n>1)$ 两种情况讨论 E-block 的位置.

(1) 当 $8<B<16$ 时的情况

记 B-block 中的元素为 $b_{s,t}$，其中 $s,t\in\{1,2,\cdots,B\}$，E-block 中的元素为 $e_{x,y}$，其中 $x,y\in\{1,2,\cdots,8\}$. 定义一个块中左上角第一个元素为这个块的始源. 在 YASS 算法中，E-block 的始源 $e_{0,0}$ 仅可能在位置 $b_{i,j}(i,j\in\{1,2,\cdots,B-7\})$ 上. 如图 1 所示，在一个大小为 9×9 的 B-block 中，E-block 的始源位置仅可能在灰色区域的 4 个元素上，而不可能出现在其它元素上.

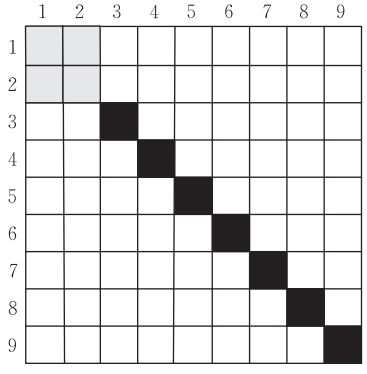

图 1　一个 9×9 的 B-block 中 E-block 的始源位置示意图

在文献[9]中，作者考虑以 B-block 主对角线上的元素 $b_{i,i}(i\in\{1,2,\cdots,B\})$ 为始源的 8×8 子块来构建可能作为 E-block 的区域 $\boldsymbol{E}_{\text{in}}$ 和不可能作为 E-block 的区域 $\boldsymbol{E}_{\text{out}}$，即将分别以灰色区域主对角线上的 $(B-7)$ 个元素为始源的 8×8 子块的并集作为 $\boldsymbol{E}_{\text{in}}$，将分别以 7 个黑色元素为始源的 8×8 子块的并集作为 $\boldsymbol{E}_{\text{out}}$. 对于含密图像，由于 QIM 的嵌入影响，$\boldsymbol{E}_{\text{in}}$ 和 $\boldsymbol{E}_{\text{out}}$ 中 8×8 子块 DCT 系数的局部随机值(具体见式(2))会出现显著差异，而对于载体图像，由于这两个区域都没有 QIM 的嵌入影响，则不会出现差异. 因此，通过 $\boldsymbol{E}_{\text{in}}$ 和 $\boldsymbol{E}_{\text{out}}$ 的选取，可以克服 E-block 位置的随机性，实现对载体图像和含密图像的分类. 本文对 Li 等的区域选择方法略作改进，将以灰色元素为始源 $(B-7)^2$ 个 8×8 子块全部作为 $\boldsymbol{E}_{\text{in}}$，从而使含密图像区域 $\boldsymbol{E}_{\text{in}}$ 和区域 $\boldsymbol{E}_{\text{out}}$ 中统计差异更加显著，既提高了分类效果，又不至于大幅增加特征提取的复杂度.

(2) 当 $B=8n+1(n>1)$ 时的情况

在文献[9]中，作者提出了一种增大数据嵌入率的改进策略，即在较大的 B-block 中选取多个 E-block. 当 $B=8n+1(n>1)$ 时，在一个 B-block 中能选取 n^2 个 E-block. 如图 2 所示为一个 17×17 的 B-block 中 E-block 的始源位置，它的区域 $\boldsymbol{E}_{\text{in}}$ 和区域 $\boldsymbol{E}_{\text{out}}$ 的选取和 $8<B<16$ 的情况类似. 本文将分别以灰色区域主对角线上的 $4n^2$ 个元素为始源的 8×8 子块的并集作为 $\boldsymbol{E}_{\text{in}}$，将分别以 $6n^2$ 个黑色元素为始源的 8×8 子块的并集作为 $\boldsymbol{E}_{\text{out}}$.

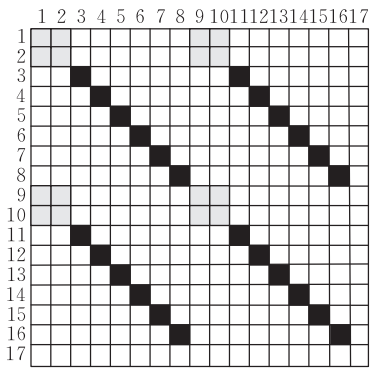

图 2　一个 17×17 的 B-block 中 E-block 的始源位置示意图

3.2 分析 QIM 嵌入方式对数据嵌入子块的影响

YASS 算法对含密图像的影响主要来自于 QIM 嵌入. 经过 QIM 嵌入得到的载体系数相比只经过量化而没有数据嵌入的载体系数有更高的局部随机性[14].

QIM 嵌入方式的基本过程是根据二值数据 0 或 1,选择相应的量化器. 定义载体系数为 x,一个步长为 Δ 的标准量化器为 $Q(x)=round(x/\Delta)\cdot\Delta$,其中 $round(\cdot)$ 表示四舍五入取整. 使用 $Q(x)$ 可以产生两个抖动量化器 $Q_s(x)$:

$$Q_s(x)=Q(x-d_s)+d_s,\ s=0,1,$$

其中 $d_0=-\Delta/4$, $d_1=\Delta/4$.

图 3 所示是 $Q_s(x)$ 的数轴表示. 分别用○和×表示量化器 Q_0 和 Q_1 的重构点集,两个量化器重构点间的最小距离为 $\Delta/2$. 若 $Q_s(x)$ 的抖动范围小于 $\Delta/4$,则嵌入数据可以被无差错地估计.

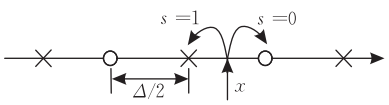

图 3 $Q_s(x)$的数轴表示

在 YASS 中,使用量化器 Q_0 和 Q_1 在未取整的 DCT 低频系数上嵌入数据,公式如式(1)所示,其中 y 表示嵌入后的载体系数.

$$y=\begin{cases}Q_0(x)=round(x/\Delta+1/4)\cdot\Delta-\Delta/4, & s=0\\ Q_1(x)=round(x/\Delta-1/4)\cdot\Delta+\Delta/4, & s=1\end{cases} \tag{1}$$

假设 $\boldsymbol{X}=\{x_i\}(i=1,2,\cdots,N)$ 是长度为 N 的载体序列,在$(-\Delta/2,\Delta/2)$上服从均匀分布. $\boldsymbol{X}_q$ 表示 $\boldsymbol{X}$ 经标准量化器 $Q(x)$ 量化后的序列,$\boldsymbol{X}_{\mathrm{QIM}}$ 表示 $\boldsymbol{X}$ 经两个抖动量化器 $Q_s(x)$ 量化后的序列,其中 $P(s=0)=P(s=1)=1/2$. 显然,$\boldsymbol{X}$ 经标准量化后,$\boldsymbol{X}_q$ 全部为 0;而经抖动量化后,$\boldsymbol{X}_{\mathrm{QIM}}$ 可能的取值有 $-3\Delta/4$,$-\Delta/4$,$\Delta/4$ 和 $3\Delta/4$,其概率函数分别为 $P\{\boldsymbol{X}_{\mathrm{QIM}}=-3\Delta/4\}=1/8$,$P\{\boldsymbol{X}_{\mathrm{QIM}}=-\Delta/4\}=3/8$,$P\{\boldsymbol{X}_{\mathrm{QIM}}=\Delta/4\}=3/8$,$P\{\boldsymbol{X}_{\mathrm{QIM}}=3\Delta/4\}=1/8$,如图 4 所示为 $\boldsymbol{X}_{\mathrm{QIM}}$ 的数轴表示. 显然,序列 $\boldsymbol{X}_{\mathrm{QIM}}$ 的随机性要大于序列 $\boldsymbol{X}_q$.

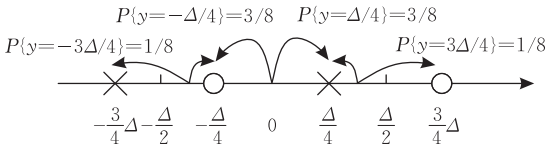

图 4 序列 $\boldsymbol{X}_{\mathrm{QIM}}$ 的数轴表示

在 YASS 中,若不对未取整的 DCT 低频系数做 QIM 操作,而只进行标准量化操作,由上文分析可知,QIM 操作后低频系数的随机性要大于标准量化操作后低频系数的随机性. 而 E-block 之间的 DCT 系数本身存在较强的相关性[15],由于 QIM 操作对低频系数引入了随机性,必然造成 DCT 系数的块间相关性变弱,而标准量化操作则不会破坏 DCT 系数的块间相关性. 由此可见,图像 E-block 的 DCT 低频系数的统计特征可用于区分载体图像和含密图像.

3.3 提取隐写分析特征

为了使 DCT 系数的块间相关性得到更准确的描述,先对选取好的 E-block 按 Hilbert 曲线进行扫描[15],得到长度为 n 的 E-block 序列. 记 $x^i_{k_1,k_2}$ 表示第 i 个块上频率(k_1,k_2)处的系数值,则 E-block 上一个 DCT 低频系数的局部随机值可用式(2)表示,其中 $4<i<n-4$,(k_1,k_2)位于候选嵌入系数内.

$$R^i_{k_1,k_2}=1-\frac{1}{9}\sum_{j=i-4}^{i+4}\theta(x^i_{k_1,k_2},x^j_{k_1,k_2}) \tag{2}$$

其中 $\theta(A,B)=\begin{cases}1, & A=B\\ 0, & \text{其它}\end{cases}$.

分别计算 E-block 序列中各个 E-block 上 DCT 低频系数的局部随机值,得到局部随机序列 $\boldsymbol{R}=\{R^i_{k_1,k_2}\}$,并计算序列 $\boldsymbol{R}$ 的均值、方差、偏度、峰度值作为特征.

当 $8<B<16$ 时,令$B=9,10,11,12,13,14,15$. 记区域 $\boldsymbol{E}_{\mathrm{in}}$ 中可能的数据嵌入位置对应的局部随机序列为 $\boldsymbol{R}^i_{\mathrm{in}}(i\in\{1,2,\cdots,(B-7)^2\})$,区域 $\boldsymbol{E}_{\mathrm{out}}$ 中不可能的数据嵌入位置对应的局部随机序列为 $\boldsymbol{R}^i_{\mathrm{out}}(i\in\{1,2,\cdots,7\})$. 分别计算 $\boldsymbol{R}^i_{\mathrm{in}}$ 和 $\boldsymbol{R}^i_{\mathrm{out}}$ 的均值、方差、偏度、峰度值,再取平均,共得到 $7\times4\times2=56$ 维特征.

当 $B=8n+1(n>1)$ 时,令 $B=17,25,33,41,49,57,65$. 同样分别计算 $\boldsymbol{R}^i_{\mathrm{in}}(i\in\{1,2,\cdots,4n^2\})$ 和 $\boldsymbol{R}^i_{\mathrm{out}}(i\in\{1,2,\cdots,6n^2\})$ 的均值、方差、偏度、峰度值,再取平均,共得到 $7\times4\times2=56$ 维特征.

3.4 Fisher 线性分类器

Fisher 线性判别(FLD)分类器是解决二类分类问题的经典分类器,具有参数设置少、计算速度快的优点. 它的基本思想是将 d 维特征空间的样本投影到一条直线上,形成一维空间,一般情况下,如果样本是线性可分的,则总能找到某个方向,使得在这个方向的直线上样本的投影能分开得最好. 在训练和测试前,需先将特征尺度化到$[-1,1]$内,这样可避免特征值过大或过小,使所有的特征都发挥作用,提高分类的准确性.

4 仿真实验及结果分析

实验使用 UCID 图像库[13],用 Matlab 将库中

的 1338 幅彩色 TIFF 图像转化为灰度 TIFF 图像后进行 YASS 隐写，得到 6 种含密图像，如表 1 所示. 同时，将转化后的灰度 TIFF 图像进行 JPEG 压缩，其中压缩的质量因子和 QF_a 相同，作为载体图像. 随机选取 800 幅载体图像和其对应的含密图像用于训练，其余嵌入成功的图像用于测试.

实验结果如表 2 所示，其中 TC 代表正常图像检测率，即载体图像被正确判断为载体图像的概率，TS 代表隐写图像检测率，即含密图像被正确判断为含密图像的概率，AR 为两者的平均值，代表最终检测正确的概率. 由于 Li 等在文献[9]中对于分块大小 $B=8n+1(n>1)$ 时未提出相应的检测方法，所以表 2 中缺少了 Li 的算法对 YASS3 和 YASS6 的检测结果. 由表 2 可知，Li 的算法和 Kodovsky 的算法的检测率相当，而本文算法的检测率较他们提高了约 4%～6%. 分析原因可知，YASS 算法对含密图像的影响主要来自于 QIM 嵌入，本文从局部随机性角度分析了 QIM 嵌入对载体系数的影响，相比较 Li 构建零值重量化系数的分析方法更加有效，提高了检测效果. 对于 Kodovsky 的算法，它对 YASS 的检测率不高主要有两个原因：(1) 算法在提取 DCT 域特征时对待测图像进行了“校准”，而 YASS 算法利用数据嵌入子块选取的随机性能很好地抵抗基于“校准”的通用隐写分析；(2) 算法的维数较高，而实验所用训练样本有限，也会对最终的检测率造成影响.

表 2　3 种算法的检测结果　　（单位：%）

图像集	Kodovsky 算法的结果			Li 算法的结果			本文算法的结果		
	TC	TS	AR	TC	TS	AR	TC	TS	AR
YASS1	91.6	91.4	91.5	97	87.4	92.2	98.2	96.8	97.5
YASS2	87	92	89.5	96.4	83	90.2	94.3	92.6	93.5
YASS3	92.4	93.8	93.1	—	—	—	99.3	95.8	97.6
YASS4	88.8	92.6	90.7	97	84	90.5	97.5	92.6	95.1
YASS5	84.2	90.6	87.4	96.2	82.6	89.4	93.8	90.6	92.2
YASS6	93.8	91.4	92.6	—	—	—	98.8	94.5	96.7

5　总结与展望

YASS 是一种能抵抗 JPEG 通用分析的新型隐写算法. 本文重点分析了 YASS 算法中 QIM 嵌入方式对载体系数的影响，提出了一种新的针对 YASS 算法的专用检测方法. 文中根据 B 块的大小，分析 8×8 数据嵌入子块的位置，并提取子块 DCT 低频系数局部随机序列的 4 个统计量作为特征，结合 Fisher 线性分类器设计了检测方法. 相比较 Li 的算法和 Kodovsky 的算法，其对 YASS 的检测率有一定提高. 但由于本文算法是专用分析方法，所以灵活性和可扩展性较差. 未来研究的重点将包括以下两个方面：(1) 找到更加有效的特征，不仅对 YASS 算法有较好的检测效果，而且对其它的压缩域隐写算法（如 F5、MB 等）和空域隐写算法（如 LSB、LSB 匹配等）也有较好的检测效果. (2) 能够进一步鉴别出各种隐写算法、估计出隐写容量和嵌入位置，甚至恢复出含密图像中的秘密信息.

参　考　文　献

[1] Kodovsk J, Fridrich J. Influence of embedding strategies on security of steganographic methods in the JPEG domain// Proceedings of the SPIE Electronic Imaging, Security, Forensics, Steganography, and Watermarking of Multimedia Contents X. San Jose: SPIE, 2008: 1-13

[2] Westfeld A. High capacity despite better steganalysis (F5-a steganographic algorithm)//Proceedings of the 4th International Workshop on Information Hiding. Heidelberg: Springer-Verlag, 2001: 289-302

[3] Sallee P. Model-based methods for steganography//Proceedings of the International Workshop on Digital Watermarking. Heidelberg: Springer-Verlag, 2004: 154-167

[4] Fridrich J, Goljan M, Lisonek P, Soukal D. Writing on wet paper//Proceedings of the SPIE, Electronic Imaging, Security, Steganography and Watermarking of Multimedia Contents VII. San Jose: SPIE, 2005: 328-340

[5] Solanki K, Sarkar A, Manjunath B S. YASS: Yet another steganographic scheme that resists blind steganalysis//Proceedings of the 9th International Workshop on Information Hiding. France: Springer-Verlag, 2007: 16-31

[6] Sarkar A, Solanki K, Manjunath B S. Further Study on YASS: Steganography based on randomized embedding to resist blind steganalysis//Proceedings of the SPIE Security, Steganography, and Watermarking of Multimedia Contents X. San Jose: SPIE, 2008: 16-31

[7] Fridrich J. Feature-based steganalysis for JPEG images and its implications for future design of steganographic schemes// Proceedings of 6th International Workshop on Information Hiding. Heidelberg: Springer-Verlag, 2004: 67-81

[8] Pevny T, Fridrich J. Merging Markov and DCT features for

multi-class JPEG steganalysis//Proceedings of the SPIE Electronic Imaging, Security, Steganography, and Watermarking of Multimedia Contents IX. San Jose: SPIE, 2007: 3-4

[9] Li B, Shi Y Q, Huang J W. Steganalysis of YASS//Proceedings of the 10th ACM Multimedia & Security Workshop. Oxford: ACM, 2008: 139-148

[10] Liu Hong, Wang Jian-Jun. A new steganalysis method of YASS. Journal of Fudan University (Natural Science), 2009, 48(4): 443-450(in Chinese)
(刘洪，王建军. 一种新的 YASS 的隐写分析方法. 复旦大学(自然科学版), 2009, 48(4): 443-450)

[11] Yu X Y, Babaguchi N. Breaking the YASS algorithm via pixel and DCT coefficients analysis//Proceedings of the 19th International Conference on Pattern Recognition. Tampa, 2008: 1-4

[12] Kodovsky J, Pevny T, Fridrich J. Modern steganalysis can detect YASS//Proceedings of the SPIE Electronic Imaging, Media Forensics and Security XII. San Jose: SPIE, 2010: 1-11

[13] Schaefer G, Stich M. UCID — An uncompressed colour image database//Proceedings of the Storage and Retrieval Methods and Applications for Multimedia. San Jose, 2004: 472-480

[14] Malik H. Steganalysis of QIM steganography using irregularity measure//Proceedings of the 10th ACM Multimedia & Security Workshop. Oxford: ACM, 2008: 149-158

[15] Westfeld A. Generic adoption of spatial steganalysis to transformed domain//Proceedings of the 10th International Workshop on Information Hiding. USA: Springer-Verlag, 2008: 161-177

LEI Yu, born in 1987, M. S. candidate. His research interests include steganography and steganalysis.

YANG Xiao-Yuan, born in 1959, M. S., professor. His research interests include cryptography and information hiding.

PAN Xiao-Zhong, born in 1964, M. S., professor. His research interests focus on information hiding.

GUO Dun-Tao, born in 1979, M. S., assistant. His research interests focus on information hiding.

Background

Steganography is a typical application of information hiding. The main purpose of steganography is to embed secret data into a cover media for covert communication. In contrast to steganography, steganalysis is to detect the existence of hidden data. To detect secret data in network is crucial to the national security. So steganalysis technique is an important part of network security. The kinds of steganographic methods have appeared in the past ten years. In these methods, YASS (yet another steganographic scheme) is a newly developed JPEG steganographic method. It can successfully resist many universal steganalysis methods. There are some specific and universal steganalysis methods for YASS. However, the detection rate of these methods is not high. Therefore, a effective steganalysis method for YASS is necessary.

Since studying for Ms. D., the author has researched the area of steganalysis techniques following his tutor. He and his colleague have completed the project of Image Universal Detection System between October 2008 and June 2010. In last years, he and his colleague have co-authored 4 papers in these areas.