第 34 卷 第 8 期 计 算 机 工 程 2008 年 4 月
Vol.34 No.8 Computer Engineering April 2008
·安全技术· 文章编号：1000—3428(2008)08—0150—03 文献标识码：A 中图分类号：TP301.6

# 基于多种检测算法的隐写分析方法

邓诗智，刘九芬，张卫明，陈嘉勇
(解放军信息工程大学信息研究系，郑州 450002)

**摘 要：**针对隐写分析中检测隐秘信息存在的问题，综合现有的检测算法，利用贝叶斯的独立二值分类模型，提出一种隐写分析算法。对其进行分类效果分析和参数控制，并将其应用到图像空域最低有效位隐写的检测中，结果表明，该方法较大程度地降低了虚警率和漏报率，并可以通过调整参数改善分类效果。

**关键词：**隐写分析；贝叶斯分类；虚警率；漏报率；ROC 分析

# Steganalysis Approach Based on Many Detection Algorithms

DENG Shi-zhi, LIU Jiu-fen, ZHANG Wei-ming, CHEN Jia-yong
(Department of Information Research, PLA Information Engineering University, Zhengzhou 450002)

**【Abstract】** Aiming at the detection of hidden information in steganalysis, a steganalysis approach is advanced based on many methods by using Bayes classification with independent two value. Analysis of classification effect and parameters' control are given. The result shows that the approach can reduce False Detection Rate (FDR) and False Alarm Rate (FAR) and control classification effect by adjusting parameters. Applying the approach to LSB steganagraphy, experimental results show that it improves the classification effect. Its FDR is the second best and FAR cuts more than half.

**【Key words】** steganalysis; Bayes classification; false alarm rate; false detection rate; ROC analysis

## 1 概述

信息隐藏作为信息安全的新技术已成为信息安全研究领域一个最新的研究方向。数字隐写(steganography)和隐写分析(steganalysis)是信息隐藏的重要分支，前者是将秘密消息隐藏在载体中进行传送而不引起第三方怀疑，达到隐蔽通信的目的；后者是对数字隐写的攻击，即对隐秘信息的检测、提取、恢复和破坏。数字隐写隐藏了互相通信的事实，成为隐蔽通信的有效方式，但它也可能被不法分子使用，成为危害国家政治和经济安全的工具。至今，互联网上已出现 200 多个隐写软件，而且应用这些软件不需要高深的专业知识。在这种情况下，对网上恶意隐蔽通信不设防就会造成严重的安全隐患，因此，隐写分析是亟待研究的领域。

目前对隐写分析领域的研究基本上集中在对隐秘信息的检测。检测效果包括漏报率和虚警率两方面，漏报率是将载密错误判断为载体的概率，虚警率是将载体错误判断为载密的概率。漏报率和虚警率往往是矛盾的，降低虚警率的同时会提高漏报率，降低漏报率则会提高虚警率。在分类问题中最常使用的分析方法是 ROC(detector's Receiver Operating Characteristic)分析，对应的 ROC 曲线描述了判断载密的阈值、漏报率和虚警率这三者之间的关系，当虚警率等于漏报率时，对应的 ROC 曲线是一条经过原点斜率为 1 的直线，其中，AUC(Area Under Curve)即 45°对角线与横坐标围成区域的大小，为 0.5，属于随机猜测：判断正确和错误的概率均占 0.5，此时的算法检测无效，当 AUC 达到 0.85 及以上，认为检测器性能良好。

由于最低有效位(Least Significant Bit, LSB)隐写方法使用的广泛性，针对它的隐秘信息检测是隐写分析的重点，其中性能较好的算法有：RS方法[1]，SPA方法[2]，DIH方法[3]，EsLsb方法[4]，LSM方法[5]。文献[6]通过大量实验评估了Pairs[7], RS和SPA方法的可靠性，并针对这 3 种方法给出了多种改进方案，在一定程度上改进了虚警率和漏报率。它集中于改进单个算法以提高检测效果，但是单个算法的虚警率和漏报率是相互制约的，而且单个算法的检测效果也是有限的。为了进一步提高检测效果，本文综合利用已有的检测算法，提出了一种隐写分析方法，可以同时降低虚警率和漏 报率。

## 2 基于多检测算法的隐写分析算法的设计与实现

针对某种隐写分析问题，假设现有 $d$ 个已知的检测算法，每个检测算法都能给出检测对象是否载密的答案，由此得到了一个向量 $\boldsymbol{X}=(x_1, x_2, \cdots, x_d)$，其中，元素 $x_i$ 为检测算法的分类结果，载密记为 1，载体记为 0。

针对每个检测算法，根据 ROC 分析，得到对应的虚警率 $\alpha_i$ 和漏报率 $\beta_i$，因此，

$$p_i = Pr[x_i = 1 | \omega_1] = \alpha_i$$

$$q_i = Pr[x_i = 1 | \omega_2] = 1 - \beta_i$$

其中，$Pr$ 表示求概率；$\omega_1$ 和 $\omega_2$ 表示类别：载体和载密；$p_i$ 和 $q_i$ 分别是在实际情况下为 $\omega_1$ 和 $\omega_2$ 时 $x_i$ =1 的概率。利用贝叶斯的独立二值分类模型设计算法如下：

假设向量 $\boldsymbol{X}$ 中的元素 $x_i$ 相互独立，条件概率 $P(\boldsymbol{X} | \omega_i)$ 表

**基金项目：**国家自然科学基金资助项目(60473022)；河南省自然科学基金资助项目(0511011300)

**作者简介：**邓诗智(1984－)，男，硕士研究生，主研方向：信息隐藏，隐写分析；刘九芬，副教授；张卫明，讲师、博士；陈嘉勇，硕士研究生

**收稿日期：**2007-05-10 **E-mail：**shizhi_deng@126.com

示如下：

$$P(\boldsymbol{X}\mid\omega_1)=\prod_{i=1}^{d}p_i^{x_i}(1-p_i)^{1-x_i}$$

$$P(\boldsymbol{X}\mid\omega_2)=\prod_{i=1}^{d}q_i^{x_i}(1-q_i)^{1-x_i}$$

设 $P(\omega_1)$ 和 $P(\omega_2)$ 分别为载体和载密的先验概率，则有

$$\frac{P(\boldsymbol{X}\mid\omega_1)P(\omega_1)}{P(\boldsymbol{X}\mid\omega_2)P(\omega_2)}=\frac{P(\omega_1)}{P(\omega_2)}\prod_{i=1}^{d}\left(\frac{p_i}{q_i}\right)^{x_i}\left(\frac{1-p_i}{1-q_i}\right)^{1-x_i}$$

对其两边求自然对数，可得线性判别函数：

$$g(\boldsymbol{X})=\sum_{i=1}^{d}\left[x_i\ln\frac{p_i}{q_i}+(1-x_i)\ln\frac{1-p_i}{1-q_i}\right]+\ln\frac{P(\omega_1)}{P(\omega_2)}$$

化简后为

$$g(\boldsymbol{X})=\sum_{i=1}^{d}\omega_i x_i+\omega_0 \tag{1}$$

其中，

$$\omega_i=\ln\frac{p_i(1-q_i)}{q_i(1-p_i)}=\ln\frac{\alpha_i\beta_i}{(1-\alpha_i)(1-\beta_i)},\quad i=1,2,\cdots,d \tag{2}$$

$$\omega_0=\sum_{i=1}^{d}\ln\frac{1-p_i}{1-q_i}+\ln\frac{P(\omega_1)}{P(\omega_2)}=\sum_{i=1}^{d}\ln\frac{1-\alpha_i}{\beta_i}+\ln\frac{P(\omega_1)}{P(\omega_2)} \tag{3}$$

如果 $g(\boldsymbol{X})>0$，则判断为载体 $\omega_1$；否则判断为载密 $\omega_2$。

## 3 基于多种算法的隐写分析算法的检测效果分析

本文算法的检测效果包含 2 个方面：虚警率 $\alpha$ 和漏报率 $\beta$。为便于观察，假设所有检测算法的虚警率均相等设为 $\alpha_0$，设特征向量为 1 的数目为 $N$，设 $N_0$ 为满足 $g(\boldsymbol{X})\le 0$ 向量元素为 1 的最小数目，$P(\omega_1)$ 和 $P(\omega_2)$ 分别为载体和载密的先验概率，根据第 2 节的分析，$\alpha$ 和 $\beta$ 可以写成如下形式：

$$\alpha=P(\omega_1)\int_{g(\boldsymbol{X})\le 0}\alpha_0^{N}(1-\alpha_0)^{d-N}\mathrm{d}\boldsymbol{X}=P(\omega_1)\sum_{N=N_0}^{d}\alpha_0^{N}(1-\alpha_0)^{d-N} \tag{4}$$

$$\beta=P(\omega_2)\int_{g(\boldsymbol{X})>0}\beta_0^{d-N}(1-\beta_0)^{N}\mathrm{d}\boldsymbol{X}=P(\omega_2)\sum_{N=0}^{N_0}\beta_0^{d-N}(1-\beta_0)^{N}=P(\omega_2)\sum_{N=N_0}^{d}\beta_0^{N}(1-\beta_0)^{d-N} \tag{5}$$

### 3.1 检测效果、算法数目与单个算法误判率的制约关系

以虚警率为例，设虚警率为 $\alpha$，算法数目为 $d$，单个算法虚警率为 $\alpha_0$，$P(\omega_1)$ 为载体的先验概率。虚警率控制为 $\alpha<\alpha_0/K$（$K$ 为正实数)的充分条件如下：

根据式(4)，有

$$\alpha<P(\omega_1)\sum_{N=1}^{d}\alpha_0^{N}(1-\alpha_0)^{d-N}=\alpha_0\cdot P(\omega_1)\left[(1-\alpha_0)^{d}-\alpha_0^{d}\right]/(1-2\alpha_0)<\alpha_0\cdot P(\omega_1)(1-\alpha_0)^{d}/(1-2\alpha_0) \tag{6}$$

要使 $\alpha<\alpha_0/K$，则只须

$$\alpha_0\cdot P(\omega_1)(1-\alpha_0)^{d}/(1-2\alpha_0)<\alpha_0/K$$

即算法数目 $d$ 满足

$$d>\ln\left(\frac{1-2\alpha_0}{KP(\omega_1)}\right)\Big/\ln(1-\alpha_0) \tag{7}$$

将式(7)转化，得到 $\alpha$ , $d$ 和 $\alpha_0$ 三者之间的制约关系为

$$2\alpha_0+KP(\omega_1)(1-\alpha_0)^{d}<1 \tag{8}$$

对于漏报率，根据式(5)作类似分析，可得到类似结论。

### 3.2 检测效果的一个近似估计

仍以虚警率为例，由于 $\alpha_0<1/2$ 且一般控制在 0.2 以内，因此式(4)中的积分项随着 $N$ 的增加迅速减小，以至于 $N=1,2$ 时的值之和与 $N=1,2,\cdots,d$ 的总和相当，分别记为 $S_{12}$ 和 $S$。为了进一步观察两者的差距，分以下 2 种情况分析：

(1)固定算法数目 $d=7$，当 $\alpha_0$ 为 0~0.2 时，$S_{12}$ 与 $S$ 的差距控制在 0.08 以内，且 $S_{12}$ 占 $S$ 的比例随着 $\alpha_0$ 的增大而减小，$\alpha_0$ 达到最大 0.2 时比例达到最小 0.938。因此当 $\alpha_0\le 0.2$ 时，$S_{12}$ 代替 $S$ 产生的误差占 $S$ 的比例至多为 0.062，误差具体数值至多为 0.08×0.062=0.00496。

(2)固定 $\alpha_0=0.15$，当 $d$ 为 2~20 时，$S_{12}$ 占 $S$ 的比例随着 $d$ 的增大几乎保持不变，且近似等于 1。随着 $d$ 的继续增加，这个比例基本维持在 1，即固定 $\alpha_0$ 时，$d$ 的增加导致误差增加很小，近似为 0。

根据 $\alpha_0$ 和 $d$ 对积分项 $\alpha_0^{N}(1-\alpha_0)^{d-N}$ 的影响分析，得出：当 $\alpha_0\le 0.2$ 时，本文算法的虚警率 $\alpha$ 主要取决于$N$=$N_0$和 $N_0+1$ 时的值之和，记为 $\mathrm{sum}(N_0,N_0+1)$，其中，$N_0=\min\{N\mid g(\boldsymbol{X})\le 0\}$；误差在 $0.00496P(\omega_1)$ 以内，$P(\omega_1)<1$。

一般情况下，即检测算法效果相差不很大时，有 $d\ge 3$，$2\le N_0\le d-1$，$\alpha_0\le 1/2$。得到：

$$\alpha\approx\mathrm{sum}(N_0,N_0+1)=P(\omega_1)(1-\alpha_0)^{d-N_0}\alpha_0^{N_0}+P(\omega_1)(1-\alpha_0)^{d-N_0-1}\alpha_0^{N_0+1}\le P(\omega_1)(1-\alpha_0)^{d-1-N_0}\alpha_0^{N_0}\le P(\omega_1)\alpha_0^{2}\le\alpha_0/2 \tag{9}$$

对漏报率作完全类似分析，得到类似结论：在 $d$ 个检测算法效果相差不是极大时，有 $\beta\le\beta_0/2$。

## 4 基于多种算法的隐写分析算法的参数控制

在实际应用过程中，往往由于资源限制以及对检测效果的不同要求，需要根据实际情况着重降低虚警率或者漏报率。下面以虚警率为例进行分析，对于漏报率的分析类似。

假设在实际情况下，虚警率需要控制为 $\alpha$，根据 3.1 节，要使 $\alpha<\alpha_0/K$，其中，$\alpha_0$ 为单个算法虚警率；$K$ 为正实数，为本文算法误差控制的比例。可通过以下 2 种方法控制检测效果：

(1)固定单个算法虚警率 $\alpha_0$，控制算法数目

$$d>\ln\left(\frac{1-2\alpha_0}{KP(\omega_1)}\right)\Big/\ln(1-\alpha_0)$$

(2)固定算法数目 $d$，控制单个算法虚警率 $\alpha_0$，满足

$$2\alpha_0+KP(\omega_1)(1-\alpha_0)^{d}<1$$

在检测算法已确定的情况下，只有通过方法(2)来控制本文算法的检测结果：设已知算法的最差虚警率为 $\alpha_0$，对应 $d$ 个算法的漏报率一般均不相等，且随着 $\alpha_0$ 的减小而增大，记为 $\beta_i$，$i=1,2,\cdots,d$。根据 3.1 节，为了降低虚警率 $\alpha$，可以通过降低 $\alpha_0$ 达到，这样检测得到的虚警率 $\alpha\le\alpha_0/2$，漏报率 $\beta\le\beta_0/2$，其中，$\beta_0=\max\{\beta_i,i=1,2,\cdots,d\}$。

## 5 将本文算法应用到图像 LSB 替换隐写的检测问题

### 5.1 实验资料与实验结果

选取 5 000 幅BMP图像[8]为载体对象，应用空域随机LSB替换隐写算法，嵌入 0, 1 密文序列，其中，嵌入率为 10%，得到 500 幅载密图像，整个图像库的载体先验概率为 $P(\omega_1)=10/11$。选择 4 个针对LSB替换隐写分析算法：RS, SPA, DIH和LSM。

根据第 2 节的步骤进行检测。为了便于实验和观察，将 4 个算法的虚警率 $\alpha_i$ 均取为 0.05，检测的结果见表 1。从中可以看出，本文算法的漏报率与 4 个已知算法的最优值相差不大，而虚警率降低了 50%以上，从 0.05 降低到 0. 013 291 4，算法同时改进了虚警率和漏报率，拥有更优的虚警率和次优的漏报率。

**表 1　4 种算法和本文算法的检测效果比较**

| 算法 | 判断是否载密的阈值 | 虚警率(理论虚警率) | 漏报率(理论漏报率) |
|---|---|---|---|
| RS | 0.052 937 6 | 0.050 000 0 | 0.010 489 5 |
| SPA | 0.049 124 4 | 0.050 000 0 | 0.024 475 5 |
| DIH | 0.056 960 5 | 0.050 000 0 | 0.010 489 5 |
| LSM | 0.125 573 0 | 0.050 000 0 | 0.041 958 0 |
| 本文 | 上面四者组合 | 0.013 291 4<br>(0.000 225 0) | 0.024 475 5<br>(0.000 224 0) |

为了进一步分析本文算法实际应用时的分类效果，对已知算法仍然取相同的虚警率 $\alpha_0$，且从 0.005 依次递加到 0.100，随着已知算法虚警率的降低，观察本文算法的理论误判率和实际误判率(包括虚警率和漏报率)的差距，以及算法的实际误判率与已知算法最差误判率的一半的关系等，如图 1~图 4 所示。

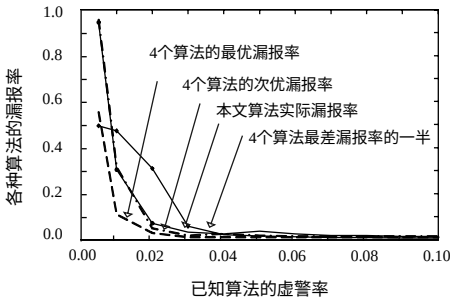


**图 1　本文算法的实际漏报率**

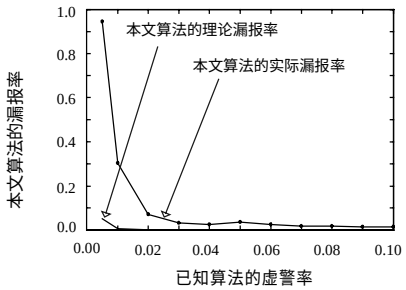


**图 2　本文算法的理论漏报率和实际漏报率**

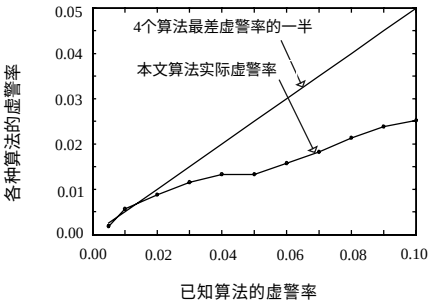


**图 3　本文算法的实际虚警率**

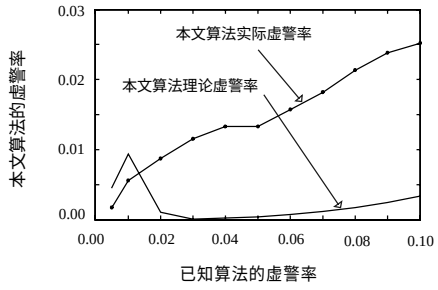


**图 4　本文算法的理论虚警率和实际虚警率**

从图 2、图 4 可以看出，本文算法的实际误判率与理论误判率具有相同的变化趋势，从图 3 可以看出，随着 4 个算法的虚警率逐渐减小，算法的实际虚警率也在减小，且均小于 4 个算法最差虚警率的一半，这与 3.1 节的分析一致，从图 1 可以看出，算法的漏报率随 4 个算法最差虚警率的减小而增加，且与 4 个算法中的次优漏报率相当，且基本小于 4 个算法中最差漏报率的一半，同样与理论的误差分析一致。

### 5.2　误差分析

从实验结果可以看出，本文算法的实际误判率与理论误判率具有相同的变化趋势，但是实际值与理论值仍有一定差距，原因为：(1)实际情况下，选取的检测算法不满足第 2 节中的独立性条件。其实这个假设比较苛刻，本文基于此假设主要是为了分析的简便。为了得到满意的实验结果，可以根据第 3 节的参数控制，比如选取更多的检测算法或者降低单个算法的误判率等方式达到。(2)单个算法的先验知识的完整性即对应 ROC 分析的准确性。这主要取决于 ROC 分析时选取的图像库是否典型，可以通过建立更加符合要求的图像库来实现，图像库的建立本身就是一项重要的工作，同时也是以后工作需要考虑的问题。

## 6　结束语

针对隐写分析中如何进一步提高检测效果的问题，本文提出了一种隐写分析算法，并分析了该算法的理论分类效果，同时给出检测效果分析，得到了分类效果的一个近似估计：理论和实际结果差不多都低于最差虚警率和漏报率的一半。应用本文算法到空域 LSB 替换隐写的检测问题，实验结果表明，实际检测效果与理论分析基本一致，并针对实际值和理论值之间的差距，进行了误差分析。

比较本文算法与单个算法：在单个算法中，虚警率和漏报率是矛盾的，不能将较优的虚警率和漏报率集中于一个算法，本文算法则在漏报率达到单个算法次优的水平下，大大降低虚警率，理论和实验结果均表明其至少降低到已知算法最差虚警率的一半，即同时改进了虚警率和漏报率，从一定程度上改善了分类效果。在实际应用时，可以通过降低已知算法最差虚警率(漏报率)来降低本文算法的虚警率(漏报率)。

值得一提的是，本文主要实验了针对空域 LSB 替换隐写的检测问题，对于其他一些隐写分析问题，只要寻找到 2 种或以上的隐写分析算法，同样可以类似地应用本文算法。

### 参考文献

[1] Fridrich J, Goljan M. Reliable Detection of LSB Steganography in Grayscale and Color Images[C]//Proc. of ACM Special Session on Multimedia Security and Watermarking. Ottawa, Canada: ACM Press, 2001.

[2] Dumitrescu S, Wu Xiaolin, Wang Zhe. Detection of LSB Steganography via Sample Pair Analysis[C]//Proc. of the 5th International Workshop on Information Hiding. [S. l.]: Springer-Verlag, 2002.

[3] Zhang Tao, Ping Xijian. A New Approach to Reliable Detection of LSB Steganography in Natural Images[J]. Signal Processing, 2003, 83(10): 2085-2093.

[4] Fridrich J, Goljan M. On Estimation of Secret Message Length in LSB Steganography in Spatial Domain[C]//Proc. of EI SPIE'04. San Jose, California, USA: [s. n.], 2004.

[5] Lu Peizhong, Luo Xiangyang, Tang Qingyang, et al. An Improved

(下转第 155 页)