

文章编号:1007-2780(2015)02-0326-07

一种基于权重融合的 JPEG 隐写分析方法

孙寿健*, 魏立线, 刘 佳, 苏光伟

(武警工程大学 电子技术系 网络与信息安全武警部队重点实验室, 陕西 西安 710086)

摘要:针对 JPEG 图像通用隐写检测中检测效率低、训练时间长的问題,提出一种基于集成分类器的新检测方法。算法以 CC-PEV 为特征对图像进行描述并作为隐写分析特征;然后,随机构造若干个特征子空间,用 bootstrap 方法构造图像训练子集,分别进行训练得到数个基分类器;根据基分类器的分类结果赋予基分类器不同的权重,将基分类器的结果按照其权重进行融合得到最终的结果。本文对该算法进行了测试,对它的集成性、检测准确率和训练时间进行分析。实验结果表明,相对于传统的集成方法,本文方法用自举方法构造训练集、随机方法构造子特征空间、赋予基分类器不同权重进行融合能够显著地提高算法准确率。本文方法相对于 SVM 和传统的集成分类方法,具有更高的检测率,对于特征维数更大的图像检测,具有更好的拓展性和一般适用性。

关 键 词:隐写分析;集成;分类器;权重

中图分类号:TP309 **文献标识码:**A **doi:**10.3788/YJYXS20153002.0326

A rights-based re-integration method for JPEG steganalysis

SUN Shou-jian*, WEI Li-xian, LIU Jia, SU Guang-wei

(*Key Laboratory of Network and Information Security Under the Chinese Armed Police Force, Electronic Department, Engineering College of Armed Police Force, Xi'an 710086, China*)

Abstract: Aiming at the problem of the high dimension features and complicated training in image universal steganalysis technology, this paper proposes a steganalysis method based on ensemble classifier. The method designs a ensemble classifier which combines the difference of base classifier and weighted fusion. Specifically, the algorithm uses CC-PEV as steganalysis features. Then, the method randomly structures several feature subspace, uses the bootstrap method to construct image training subsets, and train them to get a number of base classifiers. Finally, every base classifier is given weight that different from each other according to the classification result and then the final forecast is got after the combination of classification results with their respective weight. The experimental results show that using ensemble classifier for training and giving base classifiers with different weights fusion could effectively improve the accuracy of the algorithm.

Key words: steganalysis; integration; classifier; weights

收稿日期:2014-04-02;修订日期:2014-05-22.

基金项目:国家自然科学基金(No. 61379152)

* 通信联系人, E-mail: ssj1990220@163.com

1 引 言

隐写分析^[1-3](steganalysis)的目的是检测隐秘载体中秘密信息的存在性,进而估计出秘密信息的嵌入比率和嵌入位置,最终提取出秘密信息。其中,秘密信息检测是嵌入比率估计和秘密信息提取的前提,也是当前隐写分析技术研究的热点问题。它通常采用基于统计的方法,通过挖掘载体图像和隐写图像间的统计差异构造特征向量集,采用分类器进行隐写判别。因此,隐写分析实质上是多维特征空间的模式分类器,设计与特征空间相匹配的分类器对于提高算法的检测效果具有十分重要的意义。

隐写分析技术主要包括专用隐写分析和通用隐写分析两大类^[1-3],一般来说,专用隐写分析只对特定的方法有效;通用隐写分析对几种隐写方法都有效,甚至对新的未知方法也有效。专用隐写方法主要有基于 LSB 的专用隐写分析和基于 JPEG 域的专用隐写分析。通用隐写分析主要包括基于图像质量、基于统计矩和基于相邻像素相关性的隐写分析。本文提出方法属于通用隐写分析。

在文献[4]中,Fridrich 等人提出一种集成分类隐写分析方法,将模式识别领域的集成分类器运用到图像通用隐写分析中,实验证明,集成隐写分析可以在计算时间和精确度上有很大的提升。但存在以下几点可以改进的地方:(1)没有考虑基分类器之间的差异性,构造基分类器的方法比较简单,可以优化训练集的构造方法和子特征空间的构造方法;(2)集成分类策略很简单,赋予每个基分类器相同的权重。基于以上两点,本文设计了新的集成隐写分析算法,算法以 CC-PEV 为特征对图像进行描述并作为隐写分析特征;然后,随机构造若干个特征子空间,用 bootstrap 方法构造图像训练子集,分别进行训练得到数个基分类器;最后,根据基分类器的分类结果赋予基分类器不同的权重,最后将基分类器的结果按照其权重进行融合得到最终的结果。实验结果表明,赋予基分类器权重能够显著地提高算法准确率。

2 特征空间设计

校准是攻击 F5 隐写算法时提出来的一个新

概念。随后,校准成了 JPEG 通用盲隐写分析中不可或缺的一个环节,实验证明,它可以有效提高隐写分析的检测率。Jan Kodovsky 等^[5]对校准的原因及步骤进行研究。在本文算法中,采用 PEV 特征的校准特征 CC-PEV^[5],现有的实验结果表明,在特征空间中尽可能将不同类型的特征组合在一起可以有效提高检测率,PEV 特征是由几种不同类型特征组合在一起的,PEV 的特征类型和校准过程如表 1 和图 1 所示。

表 1 PEV 特征类型
Tab. 1 PEV feature type

Feature	Dimensionality
Global histogram H_i	11
Five AC histograms h_i^j	5×11
11 dual histograms g_{ij}^d	11×9
Variation V	1
Two types of blockiness B_a	2
Co-occurrence matrix C_{ij}	5×5
Markov features M_{ij}	9×9

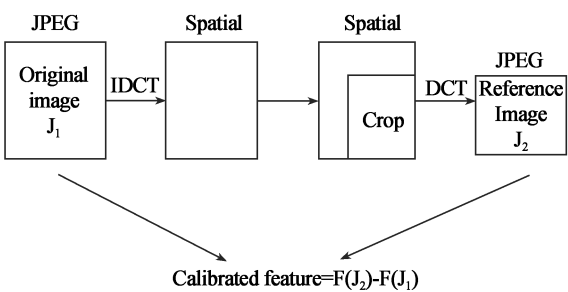


图 1 PEV 特征校准过程示意图

Fig. 1 PEV characteristics of the calibration process schematic

3 集成分类器

3.1 基分类器

SVM 分类器训练的复杂性,使得训练过程变得更长,训练复杂性更高,同时,复杂的训练过程也限制了训练特征的维数。为了提高计算效率,减小特征维数限制,本方法采用 Fisher 线性分类器作为基分类器,FLD 分类器可以不考虑特征维

数限制,同时使得每轮的训练时间减少,使得训练时间大大减少。同时,随着训练轮数的增加,最终的集成分类器的分类准确度会越来越高。

3.2 基分类器权重计算方法

考虑到基分类器样本和特征子空间的差异性,不同基分类器的对最终预测结果的影响比重也不同,所以本算法使用权重计算方法来平衡基分类器间的差异性^[6]。例如,把 N 个样本集对同一组特征重要性的判断看作是 N 个裁判的判断一样,一般情况下可以采取多数投票法来决定最后的结果。然而基于不同裁判自身的差异性,他们对最终结果影响的比重各不相同,不是简单的等权重。所以,应赋予不同基分类器不同的权重,与最终判定结果具有高度统一的基分类器应赋予更高的权重,反之,应赋予更低的权重。假设存在一个实例数为 N 的样本集,有数据集产生的 L 棵决策树,根据决策树的预测结果可获得一个 $N(L+2)$ 的矩阵,矩阵的行代表要预测的实例,矩阵的前 L 列分别代表 L 棵决策树,第 $L+1$ 列代表集成投票的结果,在前 L 列中超过半数的判定,确定为第 L 列的最终结果,第 $L+2$ 列代表样本数据集的实例类标号。则第 i 棵决策树的判定可信度可通过式(1)计算:

$$\text{TreeConfidence}_i = \frac{\sum_{j=1}^N I(\text{Tree}_{ij} = \text{Ensemble}_j)}{N} \times \text{AccEnsemble} \quad (1)$$

其中: TreeConfidence_i 表示第 i 棵树的可信度, Tree_{ij} 表示第 i 棵树对第 j 个实例的预测结果, Ensemble_j 表示对第 j 个实例的集成预测结果。 AccEnsemble 表示的是集成预测准确率,即 Ensemble 与 Original 的一致性程度, $I(\text{Tree}_{ij} = \text{Ensemble}_j)$ 是示性函数,用来表示单棵树预测结果与集成结果是否相同。由于每棵树的 AccEnsemble 都是相同的, TreeConfidence_i 与 AccEnsemble 相乘后所获得排序结果是一样的。之所以仍然需要加入这一影响因素,是为了缩小权重间的绝对差距,同时展现集成整体效果对于特征重要性度量的影响。

然后,根据计算出来的决策树的可信度计算出决策树的权重值,即基分类器的权重值。决策树的权重值可以计算如下:

$$W_i = \frac{\text{TreeConfidence}_i}{\sum_{i=1}^N \text{TreeConfidence}_i} \quad (2)$$

3.3 集成分类器

为了进一步提高分类器分类精度,设计集成分类器来进行隐写判别。集成分类器主要包括集成分类器主要包括特征子空间构造、基分类器训练和集成策略 3 个部分^[7],其基本思想是首先构造原始特征空间的随机特征子空间,然后构造具有良好差异性的训练子集,并分别进行训练得到基分类器,最后将基分类器的结果按照一定的集成策略进行融合作为最终的分类结果。

结合 CC-PEV 特征的特点,本文按如下思路设计集成分类器:随机产生 L 个具有较低维数的特征子空间^[8], L 为基分类器个数,为原始特征空间的维数且,利用 bootstrap 方法构造图像的训练样本子集,分别按 FLD 线性分类器进行训练得到基分类器,再将所有基分类器的分类结果按权重计算法进行融合作为集成分类器的分类结果。具体步骤如算法 1 所示,其中, N^{trn} 为图像训练集, N^{tst} 为测试图像集。

Step1 分类器参数优化。设置适合的基分类器个数 L 和特征子空间维数 d_{sub} ;

Step2 基分类器 $\forall l \in \{1, \dots, L\}$, 进行以下 4 步:

(a) 随机选取 d_{sub} 维特征组成特征子空间 D_l , $D_l \subset \{1, \dots, d\} \mid |D_l| = d_{\text{sub}} < d$;

(b) 利用 bootstrap 可重复抽样方法构造包含 N^{trn} 个图像样本的训练子集 N_l^{trn} ;

(c) 利用 N_l^{trn} 中的特征训练每一个基分类器 B_l , $X_l = \{x_m^{(D_l)}, \bar{x}_m^{(D_l)}\}_{m \in N_l^{\text{trn}}}$, 得到特征矢量 v_l 和临界值 T_l

(d) $\forall y \in y_{\text{tst}}$, 预测第 l 次结果:

$$B_l(y^{(D_l)}) \triangleq \begin{cases} 1 & \text{当 } v_l^T y^{(D_l)} > T_l \\ 0 & \text{其他} \end{cases} \quad (3)$$

Step3 基分类器权值的训练。利用训练好的基分类器对新的训练集进行训练,将多数投票的方法确定集成结果,按照公式(1)(2)对每个基分类器进行权值的计算。

Step4 基分类器集成。利用权重计算法将 L 个基分类器的分类结果进行融合,作为集成分类器最终的分类结果:

$$B(y) \triangleq \begin{cases} 1 & \text{当 } \sum_{l=1}^L W_l B_l(y^{(D_l)}) > L/2 \\ 0 & \text{当 } \sum_{l=1}^L W_l B_l(y^{(D_l)}) < L/2 \\ \text{随机} & \text{其他} \end{cases} \quad (4)$$

既然 FLD 分类器是一个经典的分类器,本方法只对集成分类器部分进行描述,对此不再赘述。

第 l 个分类器在集合 $\{x_i^{(D_l)}, \bar{x}_i^{(D_l)} | i \in N_l^b\}$ 上进行训练, $D_l \subset \{1, \dots, d\}$, $|D_l| = d_{\text{sub}} < d$ 是随机生成的特征子空间。 N_l^b 是在训练集 $\{1, \dots, N^{\text{trn}}\}$ 上形成的一个自举集合,且 $|N_l^b| = N^{\text{trn}}$ 。 每一个基分类器都用一般化特征向量表示。

$$v_l = (S_w + \lambda I)^{-1} (\mu - \bar{\mu}), \quad (5)$$

$\mu, \bar{\mu} \in R^{d_{\text{sub}}}$ 代表载体图像和隐写图像的特征值。

$$\mu = \frac{1}{N^{\text{trn}}} \sum_{m \in N_l^b} x_m^{(D_l)}, \bar{\mu} = \frac{1}{N^{\text{trn}}} \sum_{m \in N_l^b} \bar{x}_m^{(D_l)}, \quad (6)$$

$$S_w = \sum_m (x_m^{(D_l)} - \mu)(x_m^{(D_l)} - \mu)^T + (\bar{x}_m^{(D_l)} - \bar{\mu})(\bar{x}_m^{(D_l)} - \bar{\mu})^T, \quad (7)$$

其中: S_w 是无类别分散矩阵, λ 是使矩阵 $S_w + \lambda I$ 为正数的稳定参数,来避免 S_w 有可能出现异常数值的问题。

对于一个测试特征 $y \in y_{\text{test}}$, 第 l 个基分类器用相对应的映射 $v_l^T y^{(D_l)}$ 和对应的临界值 T_l 相比较得到它的预测结果 $B_l(y^{(D_l)})$ 。 以此类推, 得到全部 L 个基分类器的预测结果, 最终, 按照不同的权重值进行集成, 得到最终的预测结果 $B(y)$ 。

当 $B(y)$ 的取值为 1 时, 将 y 判定为隐写图像; 取值为 0 时, 判定为隐写图像。 虽然每个基分类器方法比较简单, 分类准确度不是很高, 但随着分类轮数 L 的增加集成之后分类效果明显提高, 同时缩短了训练时间。 同时, 算法 1 具有很好的通用性, 把第二步的基分类器改成 SVM 等复杂分类器也可以集成新的分类器, 具有一定的泛化性。

3.4 分类器参数优化

由于基分类器的泛化性很弱, 分类效果不明显, 构造检测准确度高、差异性大的集成分类器成为关键。 而集成分类器的最重要的两个参数 L 和 d_{sub} 的设定成为关键。 对此, 本人利用 1 000 幅载体图像 $c_i (i = 1, \dots, N)$ 及其对应的隐写图像 $s_i (i = 1, \dots, N)$ 作为测试样本对参数 L 和 d_{sub} 进行优化。 在训练集的每个基分类器最终调整为使最终检测错误率 P_E 最小^[9]:

$$P_E = \min_{P_{\text{FA}}} \frac{1}{2} (P_{\text{FA}} + P_{\text{MD}}), \quad (8)$$

其中: P_{FA} 和 P_{MD} 分别代表虚警率和误报率。

在本算法中, 由于前面基分类器使用的训练集是 bootstrap 自举方法得到的训练集, 每个训练样本都会收集大约 0.37N 的预测来检测训练样本的错误率。 把这些用于检测的来源于训练集之外(out of bag)的样本集得到的错误预测称为 OOB 错误^[4]。

$$E_{\text{OOB}}^{(n,d)} = \frac{1}{2N} \sum_{m=1}^{N^n} (B^{(n,d)}(X_m) + 1 - B^{(n,d)}(\bar{X}_m)). \quad (9)$$

由图 2 可以看出: (1) 随着基分类器个数 L 的增加, 错误检测率 P_E 逐渐减小并趋于一个稳定值, 超过一定的个数后, 错误检测率 P_E 几乎不再变化。 这是因为一开始随着基分类器的个数增加, 集成结果越来越准确, 错误检测率就随之降低。 当基分类器到达一定数量时, 基分类器之间差异的重复程度比较大, 集成结果准确率趋于稳定。 (2) 当 d_{sub} 较小时, 随着子空间维数 d_{sub} 的增大, 错误检测率 P_E 随之显著降低, d_{sub} 到达一定值后继续增加, 错误检测率 P_E 反而随之升高。 这是因为 d_{sub} 较小时, 特征间差异性不大, d_{sub} 变大可以有效提高特征差异性从而降低错误检测率, d_{sub} 到达一定数量继续增加后, 特征间的差异性有很多重复部分, 反而降低了特征差异性, 从而使错误检测率升高。

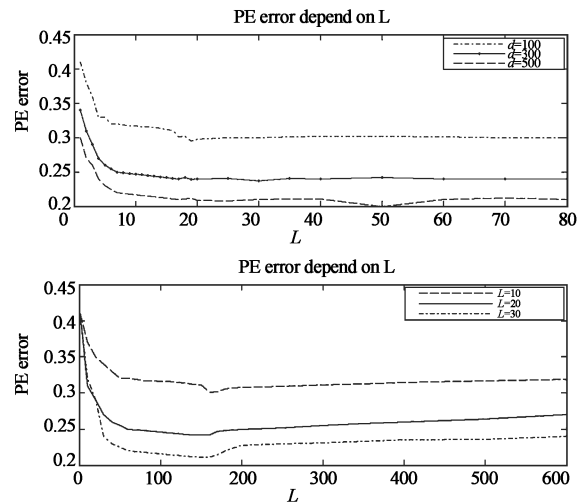


图 2 P_E 和基分类器个数 L 和子特征空间维数 d 的关系

Fig. 2 Relationship P_E and the number of base classifiers L and sub d is the dimension of the feature space

(1) L 的优化

从前面的实验曲线可以看出随着 L 的增加集成分类准确度越高,到达某固定之后结果趋于稳定,所以要找出最适合的基分类器个数 L ,使得集成分类可以在最小的计算量下达到最好的集成分类效果,在本算法中,计算 L 可以用 $E_{\text{OGB}}^{(j,d)}$ 来代替 P_E 来计算最优化的 L 。基于这样的分析,本文使用下面的思路:

$$L = \arg \min_n \begin{cases} |\min M_\mu(i) - \max M_\mu(i)| < \epsilon \\ \min M_\mu(i) = \frac{1}{\mu} \sum_{j=i-\mu+1}^i E_{\text{OGB}}^{(j,d)} \\ \max M_\mu(i) = \frac{1}{\mu} \sum_{j=i-\mu+1}^i E_{\text{OGB}}^{(j,d)} \end{cases}, \quad (10)$$

其中, $i = \{n-k, \dots, n\}$, μ, k, ϵ 都是已知给定的值,用来控制计算的精度。

(2) d_{sub} 的优化

根据上式可以求得优化的 L , 然后根据得到的 L 和上面的关系式来求最佳的 d_{sub} 。

对应于不同的基分类器数 L , 优化得到的 d_{sub} 可以会有所不同,但从前面给出的实验图可知,不同基分类器数下最佳的子特征空间维数 d_{sub} 都在很小的范围内波动,所以,在不能同时优化这两个参数的情况下,先求优化的基分类器数 L 再求优化的子空间特征维数的方法是可行的。

$$d_{\text{sub}} = \arg \min_d \begin{cases} |\min M_\mu(i) - \max M_\mu(i)| < \epsilon \\ \min M_\mu(i) = \frac{1}{\mu} \sum_{j=i-\mu+1}^i E_{\text{OGB}}^{(j,d)} \\ \max M_\mu(i) = \frac{1}{\mu} \sum_{j=i-\mu+1}^i E_{\text{OGB}}^{(j,d)} \end{cases}, \quad (11)$$

本文取得到 $\mu=5$, $\epsilon=0.05$ 得到最优化的基分类器个数 $L=20$, 子特征空间维数 $d_{\text{sub}}=210$ 。

4 实验结果与分析

4.1 实验准备

为了验证本文算法的有效性和实验效果,从 BossRank 图像库选取 1 000 幅载体图像、6 000 幅隐写图像和幅混杂图像进行实验。其中,图像大小均为 512×384 ,质量因子为 80,隐写图像采用 nsf5^[10]、mb2^[11]、Outguess^[12] 等 3 种隐写算法进行嵌入且每种算法为 2 000 幅(嵌入率分别为 10%和 50%);混杂图像包含 500 幅载体图像和

上述 3 种隐写算法图像各 200 幅(嵌入率分别为 10%和 50%)。为了衡量算法的性能和效率,本算法分别从集成性能、准确率和训练时间 3 个方面来衡量算法。

4.2 集成分类器对检测性能的影响

本方法中,使用多个基分类器集成作为最终的预测结果,其中所使用的基分类器为 FLD 线性分类器,下面来比较在这几种隐写图像下的单个分类器和集成分类器下的检测率。从表 2 中可以看出,集成分类器相对于单个分类器的检测结果都有不同程度的提高,不同嵌入率的隐写图像,其检测结果准确率都能提高,这说明,集成分类器能有效提高算法的性能。原因如下,相对于单个分类器,集成分类器通过构造更多的差异性使预测结果更加准确,同时使用权重赋值方法使得最终的融合结果体现不同基分类器对最终结果影响的大小不同,更具有准确性。同时通过比较发现,相对于低嵌入率的隐写图像,高嵌入率的隐写图像的检测率具有更大的提升幅度,这是由于高嵌入率的隐写图像和原始载体图像的差异性更大,其所对应的特征差异性更大,更容易被基分类器检测出来,从而使集成分类器精度更高。

表 2 单个分类器和集成分类器的检测结果

Tab. 2 Detection results of single classifier and ensemble classifier

Steganographic algorithm	Embedding rate/%	Single classifier results/%	Ensemble classifier results/%
nsf5	10	74.2	75.5
	50	81.7	87.6
mb2	10	75.3	86.4
	50	82.8	89.1
Outguess	10	74.5	74.2
	50	80.1	86.7

4.3 不同方法检测准确率比较

为了进一步评价本算法的性能,将本文方法同支持向量机 SVM^[10] 和文献[4]所用方法相比较;SVM 先将训练集中的图像特征进行训练找出最优分类面,然后对待检测图像进行隐写类别判定;文献[4]所用方法为集成分类方法,训练多个

基分类器,最后用简单的多数投票法来判定图像类别。3 种方法的检测结果如表 3 所示。

表 3 三种分类方法的检测结果

Tab. 3 Detection results of three different methods

Steganographic algorithm	Embedding Rate/%	SVM's results/%	Reference[4]'s results /%	Proposed's results/%
nsf5	10	70.6	77.3	78.5
	50	74.4	85.8	86.6
mb2	10	70.9	78.2	80.4
	50	78.5	87.7	88.1
Outguess	10	68.6	73.8	73.2
	50	76.2	81.1	82.7

从表 3 可以看出,本文算法比其他两种方法的检测准确率都要高。具体来讲,相对于 SVM 分类器^[13],本文方法明显优于前者,这是因为本文分类器充分考虑到特征子空间的差异性和基分类器间的差异性,随机构造特征子空间,使用自举方法选取训练样本,保证了样本和特征的一般性,采用集成方法更精确的预测待测样本的类别;相对于文献[4]所用方法,本文方法的准确率有小幅度的提升,主要是由于在基分类器融合过程中,文献[4]只是用很简单的多数投票法,没有考虑到不同基分类器对最终结果影响的不同,而本文所用方法为权重赋值法,即给不同的基分类器赋予不同的权值,从而得到更准确的分类结果。

4.4 不同方法训练时间比较

训练时间是衡量算法性能的又一指标,一个好的算法在兼顾准确率的同时,能够使训练时间尽量减到最小,统计不同方法在不同大小训练集下的训练时间,对比结果如表 4 所示。

表 4 三种方法的训练时间

Tab. 4 Training time of three different methods

N _{trn}	SVM's time/min	Reference[4]'s time/min	Proposed's time/min
1000	5	<1	<1
2000	10	1	1
3000	14	1.5	2
4000	20	2	2.5
5000	26	<3	3

由表 4 可以看出,相对于 SVM 分类器,本文所用方法能够大大缩减训练时间,训练效率得到明显提升。这是因为本文所用方法的基分类器相对简单,训练单个基分类器时间较短,从而使所用时间大大减少。相对于文献[4]所用方法,本文所用方法的集成策略相对要复杂,所以训练时间会比文献[4]略微增加,但通过表 3 可知,在几乎同样的训练时间里,本文检测正确率要高一些。

5 结 论

图像通用隐写分析往往存在特征空间维数高、训练检测复杂的问题,通过分析隐写分析算法训练中基分类器之间的差异性,提出采用权重融合的集成策略用于集成隐写分析中,结果表明,用自举方法构造训练集、随机方法构造子特征空间、赋予基分类器不同权重进行融合能够显著地提高算法准确率,实验验证了本文算法的有效性,比之前的方法具有更高的检测准确率。本文中的基分类器可以换成分类性能更好的分类器,这样集合的结果会更准确些,也是下一步的研究方向。

参 考 文 献:

[1] 张军,熊枫,张丹. 图像隐写分析技术综述 [J]. 计算机工程, 2013, 39(4): 165-168, 172.
Zhang J, Xiong F, Zhang D. Overview on image steganalysis technology [J]. *Computer Engineering*, 2013, 39 (4): 165-168, 172. (in Chinese)

[2] 王朔中,张新鹏,张卫明. 以数字图像为载体的隐写分析研究进展 [J]. 计算机学报, 2009, 32(7): 1247-1263.
Wang S Z, Zhang X P, Zhang W M. Recent advances in image-based steganalysis research [J]. *Chinese Journal of Computers*, 2009, 32(7): 1247-1263. (in Chinese)

[3] 韩晓东. JPEG 图像隐写分析技术研究 [D]. 郑州: 解放军信息工程大学, 2009.
Han X D. Research on steganlysis for JPEG images [D]. Zhengzhou: PLA Information Engineering University,

2009. (in Chinese)

- [4] Kodovsky J, Fridrich J, Holub V. Ensemble classifiers for steganalysis of digital media [J]. *IEEE Transactions on Information Forensics and Security*, 2012, 7(2): 432-444.
- [5] Kodovský J, Fridrich J. Calibration revisited [C]. *Proceedings of the 11th ACM Workshop on Multimedia and Security*, 2009: 63-74.
- [6] 尹华. 面向高位和不平衡数据分类的集成学习研究 [D]. 武汉: 武汉大学, 2012.
Yin H. Research of ensemble learning for high-dimensional and imbalanced data classification [D]. Wuhan: Wuhan University, 2012. (in Chinese)
- [7] Kodovský J, Fridrich J. Steganalysis of JPEG images using rich models [C]. *International Society for Optics and Photonics*, 2012: 83030A-83030A-13.
- [8] Kodovský J, Fridrich J. Steganalysis in high dimensions: Fusing classifiers built on random subspaces [C]. *International Society for Optics and Photonics*, 2011: 78800L-78800L-13.
- [9] 李开达. 隐藏信息检测综合评估及融合技术研究 [D]. 郑州: 解放军信息工程大学, 2012.
Li K D. Research on comprehensive evaluation and fusion of steganalysis techniques [D]. Zhengzhou: PLA Information Engineering University, 2012. (in Chinese)
- [10] Westfeld A. F5—a steganographic algorithm [M]. Springer Berlin Heidelberg, 2001: 289-302.
- [11] Ullerich C, Westfeld A. Weaknesses of MB2 [M]. Springer Berlin Heidelberg, 2008: 127-142.
- [12] Fridrich J, Goljan M, Hoge D. Attacking the outguess [C]. *Proceedings of the ACM Workshop on Multimedia and Security*, 2002.
- [13] 徐茂. 基于 SVM 的 JPEG 图像隐写分析研究 [D]. 长沙: 湖南大学, 2012.
Xu M. Steganalysis for JPEG Image Based on SVM [D]. Changsha: Hunan University, 2012. (in Chinese)

作者简介: 孙寿健(1990—), 男, 山东德州人, 硕士研究生, 主要从事图像信息隐藏方面的研究。E-mail: ssj1990220@163.com

魏立线(1966—), 男, 陕西户县人, 硕士研究生, 教授, 主要从事信息安全的研究。

刘佳(1982—), 男, 河南汝州人, 博士, 讲师, 主要从事信息安全的研究。

苏光伟(1984—), 男, 山东淄博人, 硕士研究生, 主要从事信息安全的研究。