负载自适应的批量隐写研究

高瞻瞻[1]，汤光明[1]，代家铭[2]
（1. 信息工程大学，郑州 450004；2. 解放军 65012 部队，沈阳 110000）

摘　要：为了进一步提高批量隐写的安全性，针对以往自适应批量隐写方案的不足设计了一种简单可靠的新方案。首先借助基于随机森林的集成分类器确定当前隐写分析技术下图像的隐写容量，在此基础上确定各个图像嵌入的信息量，方案最大程度地利用了载体资源，并通过对秘密信息进行分割分组进一步增强了安全性。实验结果表明，隐写容量计算中的安全判定方法准确性高，在保证低漏检率的同时避免了较高的虚警率，且运行时间更短。

关键词：自适应批量隐写；集成分类器；隐写容量；隐写安全
中图分类号：TP309.2　　**文献标志码**：A　　**文章编号**：1001-3695(2013)09-2780-04
doi:10.3969/j.issn.1001-3695.2013.09.055

Research on payload adaptive batch steganography

GAO Zhan-zhan[1], TANG Guang-ming[1], DAI Jia-ming[2]
(1. *Information Engineering University*, *Zhengzhou* 450004, *China*; 2. 65012 *Troop of PLA*, *Shenyang* 110000, *China*)

Abstract: In order to improve the security of batch steganography, this paper proposed a novel adaptive batch steganography system. First, it estimated embedding capacity of images under current steganalysis techniques by bringing ensemble classifier based on random forest into steganalysis, then distributed message bits to the images according to their embedding capacities. The new system could make full use of the images and improve batch steganography security further by dividing the message into groups. Experimental results show that the new security definition decreases the total detection error and running time, having a good balance between missed detection rate and false alarm rate.

Key words: adaptive batch steganography; ensemble classifier; embedding capacity; steganography security

0　引言

隐写是一种将秘密信息存储到图像、音频、视频等常见载体中以实现隐蔽通信的技术。隐写容量是指在安全的前提下载体嵌入信息量的上限。隐写容量较小一直是隐写技术面临的主要问题之一，因此，有必要把秘密信息拆分为几部分，再分别实施嵌入。这种方法最早由 Ker[1] 提出，被称为批量隐写（batch steganography）。

关于批量隐写的研究主要集中在消息分配方式和批量隐写总容量的计算两个方面。Ker 在文献[2]中对批量隐写进行了简化处理，在此基础上指出隐写者的最佳消息分配策略是平均分配秘密消息。Sajedi[3] 将隐写容量视为图像的一种性质，提出自适应批量隐写（adaptive batch steganography, ABS）的概念。在确定图像隐写容量的基础上，ABS 向每幅图像嵌入不多于其容量的秘密消息，不仅保证了载密体的安全，而且由于秘密消息的不平均分配进一步增大了隐写检测的难度。隐写总容量的计算方面，文献[4,5]在各自的前提假设下证明了批量隐写总的隐写容量与载体数目的平方根成正比。

本文在文献[3]的基础上提出了一种安全可行的自适应批量隐写方案，为了解决单个图像隐写容量的计算问题，利用基于随机森林的集成分类器综合多个特征集，通过合理地确定相关参数和门限定义了隐写安全，并给出了隐写容量的确定流程，所做工作有利于批量隐写安全性的进一步提高。

1　自适应批量隐写方案

自适应批量隐写的实现要解决三个问题：确定实施嵌入的载体数量、确定各个载体嵌入的信息量以及保证接收方获得正确的载密体次序。本文提出的自适应批量隐写方案如图 1 所示。方案的基本思想是按载体图像隐写容量的比值来分配秘密消息并最大限度地利用载体容量，该方案减小了载体的使用数，节约了信道资源，不仅解决了上述三个问题，而且通过对秘密消息的分割分组进一步提高了自适应批量隐写的安全性。其中隐写容量的计算将在后两章详细阐述，秘密消息分割重分组的过程如下：

a）计算并简化载体图像的隐写容量间的比值，如选取了三个载体，它们的隐写容量分别为 10000 bit、30000 bit、50000 bit，简化后即为 1∶3∶5。

b）将秘密消息 M 按 C_i 间比值分组切割，仍如 a）中的情况，就分别以 1 bit、3 bit 和 5 bit 为一组，如此循环，直至消息尾，最后不足的位不用补全。

c）将比特数相同的组合并，最后一组按其应有的比特数处理。

收稿日期：2012-12-10；**修回日期**：2013-01-28

作者简介：高瞻瞻（1988-），男，河北正定人，硕士研究生，主要研究方向为信息隐藏（gaozhandyx@126.com）；汤光明（1963-），女，教授，博导，主要研究方向为信息隐藏、体系对抗；代家铭（1989-），男，助理工程师，主要研究方向为信息隐藏.

将新的秘密消息组嵌入相应载体之后,还需要将简化的容量比值也存入相应载体中。这样做不仅避免了接收方计算隐写容量,也起到了明确载密体次序的作用,有利于接收方正确提取秘密消息。比值与载密体次序间的对应关系由密钥 K 决定,由双方在通信前商定。容量比值作为边信息所需的比特位十分有限,如上例就只需 $4(2^3=8)$ 位,所以选择一块未隐写区域用简单的隐写算法(如 LSB)实施嵌入即可。接收方收到载密体集后,首先提取边信息,并由容量比值推出载密体的次序,再从每幅图像中提取出秘密消息组,最后按容量比值重组得到原消息,完成隐蔽通信。

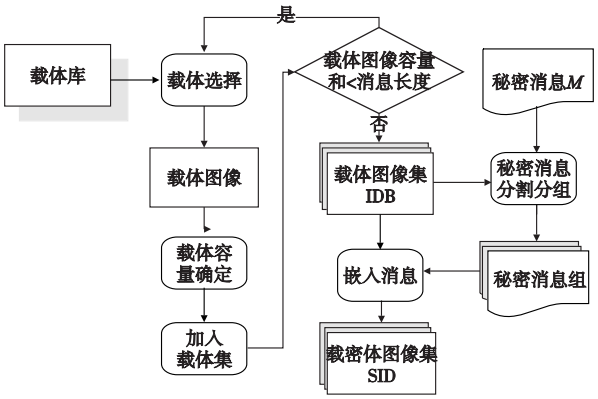

图1　批量隐写流程图

2　利用集成分类器定义隐写安全

确定合理的隐写容量是实现上述方案的关键。关于隐写安全性和隐写容量已有许多文献作过阐述,文献[6]从隐写分析的角度定义了隐写容量,指出隐写容量由隐写检测者采取的分析方法决定。文献[7]在载体和载密体元素均为独立同分布的假设下指出:如果载体和载密体的概率分布 P_c、P_s 间的相对熵至多为 ε,即

$$D(P_c\|P_s)=\int P_c\log\frac{P_c}{P_s}\leqslant\varepsilon$$

那么该隐写是 ε 安全的。如果 $\varepsilon=0$,则载密体绝对安全。文献[8]说明了在当前隐写分析技术下,单一图像的隐写容量与图像大小的平方根成正比。文献[3]则综合多种隐写分析方法的分析结果,采取最保守的策略定义了载密体的安全性,认为只要有一个分析算法检测出载密体载密,那么载密体就是不安全的。

事实上,隐写分析方对图像实施检测的过程中需要尝试多种隐写分析方法(隐写算法未知),又由于目前的通用隐写分析方法多是通过特征提取和机器学习的方式实现,因此,可以将多种隐写分析算法使用的特征集合并,在合并后的特征集上训练分类器,并以此分类器的判断结果为基础定义当前隐写分析水平下的隐写安全。

综合各特征集意味着分类器训练时特征维数的增多,对于 SVM 而言,特征维数增多会导致运算成本的急剧增大,这也正是隐写分析算法精选特征的原因所在。不过与 SVM 相比,集成分类器在处理大量训练集和高维特征向量时具有很大优势。因此,本文利用基于随机森林(random forest,RF)的集成分类器(ensemble classifier)在综合后的特征空间上进行检测。判定的基本过程如图 2 所示。

2.1　隐写安全的判定

应用基于随机森林的集成分类器的基本过程是:先随机选取出特征子集(random subspace),之后在经 Bootstrap 抽样获得的样本上进行训练,重复几次得到多个基本分类器,最后再进行集成。

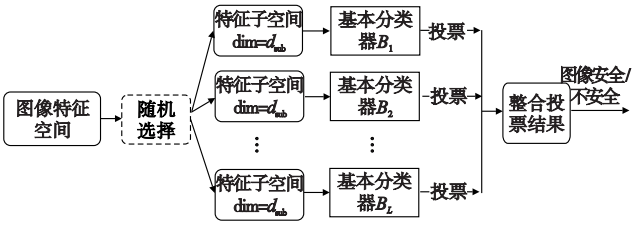

图2　隐写安全的判定

本文的基本分类器具体采用的是 Fisher 线性分类器,不仅复杂度低,也增加了分类器的不稳定性和分类器间的差异,有利于产生较好的集成效果。用 $B_l(l=1,2,\cdots,L)$ 表示一个基本分类器的分类结果,$B_l=0$ 表示分类结果认为样本安全,其值为 1 说明分类器判断样本载密。

$$B_l(y)=\begin{cases}1 & \text{when } v_l^{\mathrm{T}}y>T_l\\ 0 & \text{otherwise}\end{cases}\tag{1}$$

其中:y 表示测试样本,v_l^{T} 和 T_l 分别表示分类器训练得到的特征向量和边界值。

基本分类器形成初步判断之后,就可以根据这些结果判定样本是否安全。文献[3]实际上采用的是一票否决的方式,即 $B(y)=\max[B_l(y)]$。本文则采用投票的方式并引入门限 T,当判定样本载密的基本分类器的个数超过 $T\times L(T\in(0,1))$ 时,集成结果才认为样本不安全,最终的 $B(y)$ 由式(2)确定。

$$B(y)=\begin{cases}1 & \text{when } \sum_{l=1}^{L}B_l(y)>T\times L\\ 0 & \text{when } \sum_{l=1}^{L}B_l(y)<T\times L\\ \text{random} & \text{otherwise}\end{cases}\tag{2}$$

2.2　集成过程中的参数优化

随机森林方法中的 Bootstrap 抽样来源于 bagging 算法,它会产生袋外(out of bag, OOB)数据。OOB 数据可以用于估计 RF 的分类误差 P_E。

$$P_E=\min_{P_{\mathrm{FA}}}\frac{1}{2}(P_{\mathrm{FA}}+P_{\mathrm{MD}}(P_{\mathrm{FA}}))\tag{3}$$

其中:P_{FA}、P_{MD} 分别表示虚警率和漏检率。当有 m 个基本分类器时,这些分类器对每个训练样本平均形成 0.37 m 个判断,集成结果 $B^{(m)}(x)\in\{0,1\}$,此时 P_E 的 OOB 估计为

$$E_{\mathrm{OOB}}^{(m)}=\frac{1}{2N^{\mathrm{trn}}}\sum_{n=1}^{N^{\mathrm{trm}}}(B^{(m)}(x_n)+1-B^{(m)}(\bar{x}_n))\tag{4}$$

OOB 误差是无偏估计且可以在构建基本分类器的同时得到。因此,可以在集成过程中根据 $E_{\mathrm{OOB}}^{(m)}$ 的变化,调整特征子集维度 d_{sub} 和基本分类器个数 L 以降低 P_E,使分类器达到最好的效果[9]。

3　隐写容量的确定流程

如果一个分类器可以检测嵌入率为 0.1 的载密体,那么它对负载更高的图像也有分类能力,但实验证实这种情况下的准确率并不高。因此,本文共设计了九个集成分类器,用于对九种嵌入率(0.1,0.15,0.2,…,0.5)下的载密体进行安全判定。

通常,嵌入率越高,检测器的检测效果越好,载密体越不安全。基于这种认识,本文设计的隐写容量确定流程如图 3 所示。先令载密体的嵌入率为 0.1,并用相应的集成分类器进行检测,若视为载密体,则计算并输出隐写容量;否则就将嵌入率提高到 0.15,再用相应的分类器检测,如此循环直到分类结果判定载密体不再安全。

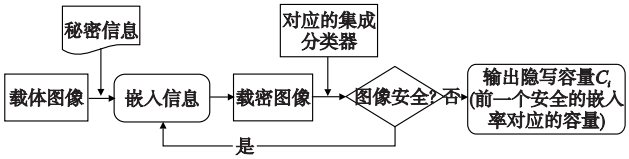

图3 隐写容量的确定流程

4 实验及结果分析

关于隐写容量,目前学术界通常采用对载体信号和隐写信号进行建模的方式进行研究,或建模为一阶随机变量,或利用高斯分布描述载体图像,抑或利用 Markov 模型刻画载体的统计分布。然而由于实际图像的复杂性、数据间的相关性,这些建模假设往往不能准确成立,从而导致基于这些假设得到的隐写容量的普适性存在问题。不仅如此,绝大多数的安全理论研究根本不能得到一个图像具体的容量值。因此,本章主要将本文方法与文献[3]的方法进行比较。

实验以 Luo 等人[10]提出的 EAI(edge adaptive image steganography)算法为例实施嵌入。实验的图像样本来自三个图像库:UCID 图像库[11],包括 1338 幅未经压缩的彩色图像;USC-SIPI 图像库[12],有灰度图和彩色图共 215 幅;Ground truth 图像库[13],包括 963 幅 JPEG 彩色图像。实验前将全部图像转换为灰度图。由于 EAI 算法是一种基于 LSB 匹配的隐写算法,为了提高安全性分析的准确性,实验综合了以下四种通用空域隐写分析方法的特征集。

a)SPAM(686D),Pevny 等人[14]将相邻像素差分值建模为马尔可夫链,提取样值的转移概率作为分类特征。

b)NIP(616D),Guan 等人[15]通过对相邻像素的灰度值进行减法运算获得邻接信息,在此基础上得到图像高阶统计特征,并利用旋转不变性降低了特征集的维度。

c)ALE(10D)[16],以灰度直方图的局部极值作为统计特征,用于检测 1 LSB 隐写算法。

d)ITF(120D),Xiong 等人[17]基于局部线性变换(LLT)将图像分解成几个细节特征子带,从 LLT 系数的直方图和共生矩阵中提取出部分特征用于隐写检测。

这四个特征集单独对 EAI 隐写进行检测时的分类效果并不理想,分类误差如图 4 所示。

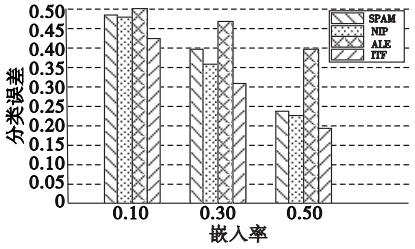

图4 单一特征集在不同嵌入率下的分类误差

4.1 分类器运行时间及分类效果

首先比较本文方法和文献[3]的方法在嵌入率为 0.5 时进行安全判定的时间,结果如表 1 所示。其中文献[3]采用的基本分类器是线性 SVM 分类器。可以看出,由于采用了合适的基本分类器和集成方式,本文方法安全判定的运行时间明显缩短。

表 1 两种方法在嵌入率为 0.5 时安全判定的运行时间 /min

判定方法	SPAM + NIP	SPAM + NIP + ALE	SPAM + NIP + ALE + ITF
本文方法	4	5	6
文献[3]方法	45	47	52

表 2 列出了文献[3]中的方法和本文方法($T=0.5$)在不同嵌入率下安全判定的分类结果。可以看出:a)与采用单个特征集时相比,本文方法取得了更好的分类效果,这说明通过集成学习和集成过程中的参数优化可以使各个基本分类器假设和目标假设之间的误差得到一定程度的抵消,提高分类的准确性;b)随着特征集数的增多,本文方法的分类误差明显降低,而文献[3]的方法提升较小,甚至出现了特征集数增多误差反而变大的现象,这说明与一票否决方式相比,本文采用的投票方式在特征集数较多的情况下更具合理性;c)随着特征集数的增多,文献[3]方法下的漏检率迅速减小,同时虚警率也变得极高,这在特征集为 SPAM + NIP + ALE + ITF 时尤为明显。

表 2 两种判定方法在不同特征集和嵌入率下的分类结果

特征集	集成方法	0.1			0.3			0.5		
		P_E	P_{FA}	P_{MD}	P_E	P_{FA}	P_{MD}	P_E	P_{FA}	P_{MD}
SPAM + NIP	文献[3]方法	0.475 3	0.726 2	0.224 3	0.382 8	0.599 4	0.166 1	0.225 9	0.345 6	0.106 2
	本文方法($T=0.5$)	0.464 9	0.521 6	0.408 3	0.346 7	0.372 3	0.321 0	0.196 3	0.223 5	0.169 0
	本文方法(T取优化值)	0.464 7	0.493 7	0.435 7	0.353 1	0.494 7	0.211 4	0.248 1	0.470 1	0.026 2
SPAM + NIP + ALE	文献[3]方法	0.488 6	0.887 1	0.091 1	0.421 3	0.760 1	0.082 5	0.271 0	0.492 4	0.049 6
	本文方法($T=0.5$)	0.463 6	0.489 2	0.438 0	0.333 9	0.351 2	0.316 7	0.181 4	0.202 5	0.160 3
	本文方法(T取优化值)	0.463 6	0.489 2	0.438 0	0.345 8	0.482 0	0.209 6	0.247 2	0.469 3	0.025 2
SPAM + NIP + ALE + ITF	文献[3]方法	0.487 6	0.919 8	0.055 4	0.423 3	0.799 7	0.046 9	0.271 0	0.521 0	0.020 9
	本文方法($T=0.5$)	0.401 2	0.437 2	0.365 2	0.281 4	0.301 9	0.260 9	0.163 2	0.186 3	0.140 1
	本文方法(T取优化值)	0.401 2	0.437 2	0.365 2	0.335 0	0.493 2	0.176 7	0.250 7	0.487 7	0.013 8

4.2　参数 T 的选择

隐写安全的定义主要针对的是载密体,为了充分利用所综合的隐写分析算法(特征集),保证判定为安全的载密体不会被分析算法识别,小的漏检率是集成方法选择时的首要目标,从这点出发,门限 T 应越小越好。但正如文献[3]中一票否决的方式一样,这样做虽然利用分类器间的差异弥补了各自在漏检方面的错误却忽略了虚警错误,导致虚警被不断放大,虚警率极高,造成隐写容量的错误估算。为了在两种错误间达到很好的平衡,本文在 T 的选择策略上坚持以下两个原则:

a)漏检率越小越好。

b)虚警率高于 50% 时,由于对安全载体的误判过多,安全分析失效。

在嵌入率为 0.5,特征集为 SPAM + NIP + ALE 时,本文方法下的虚警率和漏检率随 T 的变化如图 5 所示。在上述原则的指导下,T 的值取为 0.15,此时漏检率为 0.0252,虚警率为 0.4693。同理选出了其他情况下 T 的最优取值,最终的检测结果如表 2 所示。

由表 2 可知,与文献[3]的方法相比,本文提出的隐写安全判定方法有两个优点:a)分类误差更小,且随着特征集数的增加不断减小;b)在低漏检率下不会引起过高的虚警率,即不会对安全载体形成过高误判。

按图 5 所示结果,利用确定了最优门限值的集成分类器判定载密体的安全性,本文得到了样本库中部分图像的隐写容量,结果如表 3 所示。

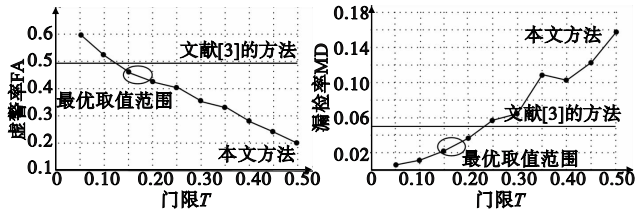

图5　虚警率和漏检率随门限T的变化

表 3　部分图像的隐写容量

20000 bit	40000 bit	60000 bit
		

5　结束语

本文设计了一种新的自适应批量隐写方案,得出了一种利用基于随机森林的集成分类器估算图像隐写容量的方法。该估算方法具有一定的通用性,能够用于预测一些新的隐写嵌入算法下的隐写容量。实验结果表明,与相关文献相比,本文方法运行时间短、分类误差低,同时避免了过高的虚警率。通过实验还发现,使用本文提出的隐写容量确定方法时,特征集数的增多虽然增大了运行代价,但在效果上只会使隐写安全的判定更加准确,进而得到更使人信服的隐写容量,而不会有其他不利影响。因而在实际应用时,可以选择多种较好的分析方法的特征集来确定图像的隐写容量,从而实现更加安全的自适应批量隐写。

参考文献:

[1] KER A D. Batch steganography and pooled steganalysis[C]//Proc of the 8th Information Hiding Workshop. Berlin: Springer, 2006: 165-281.

[2] KER A D. Perturbation hiding and the batch steganography problem [C]//Proc of the 10th Information Hiding Workshop. Berlin: Springer, 2008: 45-59.

[3] SAJEDI H. Adaptive batch steganography considering image embedding capacity[J]. Optical Engineering, 2009, 48(8): 087002.

[4] DONG Yan-shi, HAN Ke-song. Boosting SVM classifiers by ensemble[C]//Proc of the 14th International Conference on World Wide Web. New York: ACM Press, 2005: 1072-1073.

[5] KER A D. Steganographic strategies for a square distortion function [C]//Proc of SPIE 6819, Security, Forensics, Steganography and Watermarking of Multimedia Contents X. [S. l.]: SPIE, 2008: 0301-0313.

[6] CHANDRAMOULI R, MEMON N D. Steganography capacity: a steganalysis perspective[C]//Proc of SPIE 5020, Security and Watermarking of Multimedia Contents V. 2003: 173-177.

[7] CACHIN C. An information-theoretic model for steganography [C]// Proc of the 2nd International Workshop on Information Hiding. Berlin: Springer, 1998: 306-318.

[8] KER A D. A capacity result for batch steganography[J]. IEEE Signal Processing Letters, 2007, 14(8): 525-528.

[9] KODOVSKY J, FRIDRICH J, HOLUB V. Ensemble classifiers for steganalysis of digital media[J]. IEEE Trans on Information Forensics and Security, 2012, 7(2): 432-444.

[10] LUO Wei-qi, HUANG Fang-jun, HUNAG Ji-wu. Edge adaptive image steganography based on LSB matching revisited[J]. IEEE Trans on Information Forensics and Security, 2010, 5(2): 201-214.

[11] SCHAEFER G, STICH M. UCID: an uncompressed color image database[C]//Proc of SPIE 5307, Storage and Retrieval Methods and Applications for Multimedia. 2003: 472-480.

[12] Department of Electrical Engineering, University of Southern California. USC-SIPI image database [DB/OL]. http://sipi.usc.edu/services/database/index.html.

[13] Department of Computer Science and Engineering, University of Washington. Ground truth database [DB/OL]. http://www.cs.washington.edu/research/imagedatabase.

[14] PEVNY T, BAS P, FRIDRICH J. Steganalysis by subtractive pixel adjacency matrix[J]. IEEE Trans on Information Forensics and Security, 2010, 5(2): 215-224.

[15] GUAN Qing-xiao, DONG Jing, TAN Tie-niu. An effective image steganalysis method based on neighborhood information of pixels [C]//Proc of the 18th IEEE International Conference on Image Processing. 2011: 2721-2724.

[16] CANCELLI G, DOERR G, COX I J, *et al*. Detection of ±1 LSB steganography based on the amplitude of histogram local extrema [C]//Proc of the 15th IEEE International Conference on Image Processing. 2008: 1288-1291.

[17] XIONG Gang, PING Xi-jian, ZHANG Tao, *et al*. Image textural features for steganalysis of spatial domain steganography[J]. Journal of Electronic Imaging, 2012, 21(3): 033015.