DOI:10.13203/j.whugis2008.10.013
第33卷 第10期 武汉大学学报·信息科学版 Vol.33 No.10
2008年10月 Geomatics and Information Science of Wuhan University Oct. 2008

文章编号:1671-8860(2008)10-1059-03 文献标志码:A

基于 HMT 和 ICA 的主动隐写分析

刘晓芹[1] 王嘉祯[1] 徐 波[1] 邓高明[1]

(1 军械工程学院计算机工程系,石家庄市和平西路97号,050003)

摘 要:提出一种基于HMT预测的ICA盲源分离主动分析方法,该方法仅仅需要得到一幅隐秘图像,并采用小波域HMT(hidden markov tree)模型预测出载体图像的一个估计,最后使用ICA盲分离来提取秘密信息和载体图像。仿真实验给出了该方法提取出的载体图像和秘密信息,并对其进行了分析。

关键词:主动隐写分析;HMT模型;ICA;仿真

中图法分类号:TP391

独立成分分析(independent component analysis,ICA)是近年发展起来的一种信号分解技术。这种技术以非高斯源信号为研究对象,在对它们作统计独立的假设条件下,对观测到的多路混合信号进行盲分离,从而较完好地分离出隐含在混合信号中的独立信源信号。

最近许多关于图像分离的论文都涉及盲分离的ICA方法[1-3]。

1 小波域 HMT 模型预测

1.1 小波域 HMT 模型

图1是小波域HMT模型[4]的四叉树结构,其中实心点表示小波系数,空心点表示小波系数所处的状态。约定如下:以一个指标来区别四叉树的不同节点,如根节点处的小波系数记为 w_1,状态记为 S_1,以 $p(i)$ 表示节点 i 的父节点。HMT模型可以描述如下:

$$\theta=\{p_{S_i}(m)\},\mu_{i,m},\sigma^2_{i,m},\varepsilon^{mr}_{i,p(i)}\}$$

式中,$p_{S_i}(m)=P(S_i=m|\theta)$ 表示小波系数 i 处于状态 m 的概率分布函数;小波系数 i 处于状态 m 时取值为 w_1 的概率 $f_{W_i|S}(w_i|s_i=m)=g(w_i;\mu_{i,m},\sigma^2_{i,m})$,服从均值、方差分别为 $\mu_{i,m}$、$\sigma^2_{i,m}$ 的高斯分布;任一节点 i 处小波系数的状态仅依赖于其父节点 $p(i)$ 处小波系数的状态;假定 W^* 的各小波系数状态已知,则小波系数的联合分布可以按下式求出:

$$f(W^*|\theta)=\prod_{1\leqslant i\leqslant N^2}\sum_{m=1}^{2}p_{S_i}(m)f_{W_i|S_i}(w_i|s_i=m) \tag{1}$$

式中,W^* 表示小波系数全体;K 为小波系数个数。

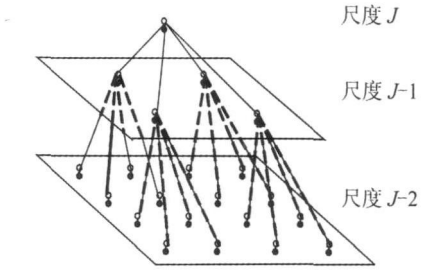

图1 HMT模型的四叉树模型

Fig.1 Four Branches Tree Structure of HMT Model

1.2 利用 HMT 模型预测原始载体图像

对于灰度图像 S,点 (i,j) $(1\leqslant i,j\leqslant N)$ 处的灰度用 $S(i,j)$ 表示。图像预测的目标是从观测到的隐写图像 Z 中估计出载体图像 S。用数学模型表述为 $Z=S+\alpha w$,其中 Z、S 和 w 分别表示观测到的隐密图像、原始载体图像和秘密信息(可以看作噪声向量),α 代表信息嵌入的强度。

通过比较HMT预测与滤波技术预测得到的载体图像相对于原始载体图像的峰值信噪比(PSNR),发现HMT预测得到的PSNR大于滤波技术得到的,从这一方面可以说明HMT预测技术要优于滤波技术。表1是分别采用HMT模型、Gaussian滤波和Wiener滤波所得到的结果。

收稿日期:2008-08-28

从表1可以看出，在嵌入长度 L 相同的情况下，HMT 模型得到的峰值信噪比要高于滤波技术得到的，并且随着嵌入长度的增加，HMT 模型的这种优势更加明显。但是，在同一预测模型下，随着嵌入长度的增加，预测效果越来越差，这是因为预测相当于噪声图像恢复，嵌入率越大，对图像质量影响越大，对预测效果的影响也越大。

表1 HMT 模型与滤波技术去噪的性能对比

Tab.1 Comparison Between Performance of HMT Model and That of Filter in Denoising

嵌入长度	PSNR		
	HMT 模型	Gaussian 滤波	Wiener 滤波
$L=10$	42.534 6	40.360 2	34.307 0
$L=100$	34.535 8	34.483 7	29.690 5
$L=1\,000$	24.321 7	19.600 3	16.827 3
$L=2\,000$	19.406 0	16.283 9	13.895 7
$L=5\,000$	16.763 5	9.941 0	13.735 9

2 主动隐写分析算法

2.1 ICA 原理及算法

ICA 盲分离[5]是把观察信号分解为相互独立的信号，其实质是寻求一种线性变换，将一组随机变量表示成一组在统计意义上相互独立的变量的线性组合。

ICA 通过求解分离矩阵 $\boldsymbol{W}$，使得 $\boldsymbol{Y}=\boldsymbol{WX}$ 与 $\boldsymbol{S}$ 对应，其中 $\boldsymbol{Y}=(Y_1,Y_2,\cdots,Y_n)^{\mathrm{T}}$。如果求解分离矩阵 $\boldsymbol{W}$ 使 $\boldsymbol{Y}$ 的各个分量尽可能地独立，那么 $\boldsymbol{Y}$ 就是 $\boldsymbol{S}$ 的最佳估计。

本文使用了 ICA 固定点迭代算法[7]，与 ICA 随机梯度算法相比，该算法有以下优点：算法迭代次数少，收敛速度快，复杂度低；克服了随机梯度算法依赖于步长因子选择的局限性；无需调整动态因子，稳定性高，具有良好的应用前景。ICA 固定点迭代法与 ICA 随机梯度法分离效果的对比见文献[6]。

2.2 主动隐写分析算法

为了便于研究，首先描述了一个线性附加的隐写算法，它适用于广泛的隐写技术。定义 $S(k)$ 为一个载体信息，$W(k)$ 为独立于载体信息的秘密信息，隐写信息可以通过下式得到：

$$Z(k)=S(k)+\alpha\omega(k),\ k=1,2,\cdots,N \quad (2)$$

式中，$S(k)$ 相邻元素的值相差不大，称其值具有连续性；$\alpha>0$，且表示信息的嵌入强度，能够根据视觉和鲁棒性等特性来调节；$\omega(k)$ 的值也具有连续性，并且当 $S(k)$ 不携带信息时，$\omega(k)$ 将等于0。本文假定 $S(k)$ 和 $\omega(k)$ 都是稳定的随机序列，且不相关，满足了 ICA 算法中 S 作为独立源信号的条件。

在这种线性附加的隐写算法的基础上，本文提出了一种基于 HMT 预测的 ICA 盲分离算法。该算法是一种主动隐写分析算法，不仅能恢复载体图像，而且能提取出秘密信息。其原理及过程如下。

1) 通过小波域 HMT 模型预测出原始载体图像 $S(k)$ 的一个估计 $Z_1(k)$。由于隐写图像 $Z(k)$ 是 $s(k)$ 和秘密信息 $\omega(k)$ 的线性组合，而小波域 HMT 模型对高斯噪声去噪来说是一个线性模型，所以去噪之后得到的 $Z_1(k)$ 仍为 $s(k)$ 和 $\omega(k)$ 的线性组合。这个估计可以看作是 $\omega(k)$ 以不同于 α 的强度 α_1 嵌入到 $S(k)$ 中得到的[0]：

$$\begin{pmatrix} Z(k) \\ Z_1(k) \end{pmatrix}=\begin{pmatrix} 1 & \alpha \\ 1 & \alpha_1 \end{pmatrix}\begin{pmatrix} S(k) \\ \omega(k) \end{pmatrix} \quad (3)$$

从式(3)中可以看出，提取 $S(k)$ 和 $w(k)$ 就是一个盲源分离的问题[0]，可以把 $\begin{pmatrix} 1 & \alpha \\ 1 & \alpha_1 \end{pmatrix}$ 看作 ICA 分析中的矩阵 $\boldsymbol{A}$，而把 $\begin{pmatrix} S(k) \\ \omega(k) \end{pmatrix}$ 看作 $\boldsymbol{S}$，$\begin{pmatrix} Z(k) \\ Z_1(k) \end{pmatrix}$ 看作 $\boldsymbol{X}$。

2) 通过 ICA 原理及 ICA 固定点算法，可以得到 $\boldsymbol{S}$ 的最佳估计，即求得 $\begin{pmatrix} S(k) \\ \omega(k) \end{pmatrix}$ 的最佳估计，从而分离出 $\omega(k)$ 和 $S(k)$。

3 实验结果及分析

在仿真实验中，采用了著名的 Lenna 图像作为载体，嵌入信息是服从 $N(0,5^2)$ 的一个随机序列，参数 $\alpha=0.1$，使用式(2)在 DWT 系数上嵌入秘密信息 $\omega(k)$。因为 DWT 是非高斯分布的，所以满足了 ICA 中载体信息和秘密信息独立的条件。如果 $\omega(k)$ 的长度为 L，选用最大的 L 个 DWT 系数嵌入。本文分别对 $L=10$、100、1 000、2 000、5 000 进行了仿真实验。

3.1 载体信息的估计

仿真结果产生了载体图像的一个估计。估计的输出结果是 DWT 系数，原始载体图像的估计可以通过 DWT 逆变换直接得到。

限于篇幅，图2仅给出了嵌入信息长度 $L=100$ 时原始载体图像的估计，估计图像与原始载体图像的峰值信噪比(PSNR)为 35.315 1 dB。超过 30 dB 的提取图像被认为是可以接受的结

(a) 原始载体图像　(b) 提取的载体图像

图 2　原始载体图像及其估计

Fig.2　Cover Image and Its Estimate

果,所以本文提取载体信息的方法是可取的。

3.2　嵌入信息的估计

仿真结果同时产生了秘密信息的一个估计,包括嵌入信息长度估计和嵌入信息序列估计。通过分析大量的实验结果,发现 ICA 分离的嵌入信息是一个 $M\times N$ 的矩阵,这个矩阵的很大一部分元素值接近于 0,而其余较大的元素几乎都是以 3～5 个连续值的块形式出现。这是因为在某一个 DWT 系数上嵌入秘密信息时,周围的几个 DWT 系数受到影响,从而出现与嵌入系数相同方向的变化。本文通过计算每一个块的平均值,产生嵌入信息序列的估计,而其长度估计即为块的个数。限于篇幅,本文仅给出了长度 $L=10$、$L=100$ 时嵌入信息与其估计的对比图,分别如图 3 所示。由于提取的信息序列的组合方式不能够确定(提取算法的一个缺点),所以提取序列的每一个值与嵌入信息最近的值对应,但不重复。表 2 为嵌入信息的长度估计错误率,这个概率是在多次实验结果平均的基础上得到的。从图 3 中可以看出,提取的秘密信息与嵌入信息非常接近,验证了方法的有效性。

表 2　嵌入信息长度估计错误率

Tab.2　Ratio of Incorrect Estimated Length of Embedded Secret Messages

长度	10	100	1 000	2 000	5 000
错误率/(%)	1	5	3	4	4

仿真实验结果表明,这种盲分离方法得到的载体图像峰值信噪比均大于 30 dB,分离出的秘密信息与嵌入信息非常接近,表明该方法能够正确地分离载体图像和秘密信息,验证了方法的有效性。但是对于秘密信息的提取有一个缺点,提取信息序列的组合方式不能确定,还需要进一步的努力。

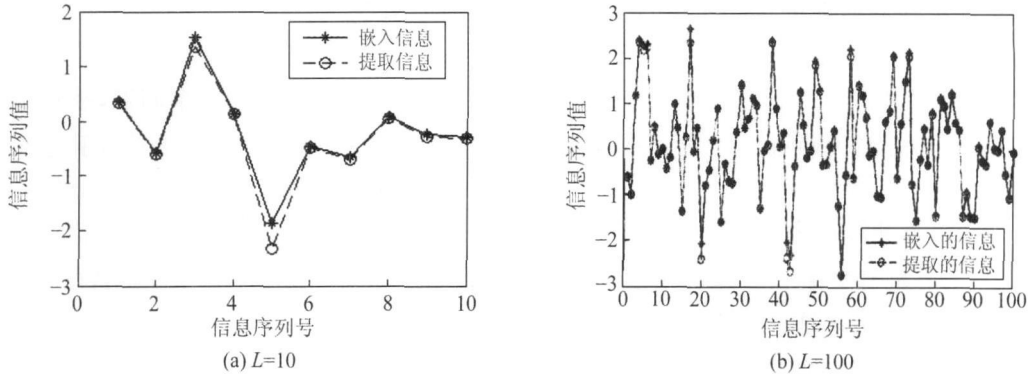

图 3　嵌入信息及其估计

Fig.3　Secret Messages Embed and Its Estimate

参 考 文 献

[1] 何小海. 图像通信[M]. 西安:西安电子科技大学出版社, 2005

[2] Sun J, Xu W B. A Global Search Strategy of Quantum Behaved Particle Swarm Optimization [C]. IEEE Conference on Cybernetics and Intelligent Systems, New York, 2004

[3] Chandramoouli R. A Mathematical Framework for Active Steganalysis[J]. Special Issue on Multimedia Security, ACM Multimedia System Journal manuscript, 2003, 9(3): 301-311

[4] Romberg J K, Choi H, Baraniuk R G. Bayesian Tree-Structured Image Modeling Using Wavelet-Domain Hidden Markov Models[J]. IEEE Transactions on Image Processing, 2001, 10(7): 1 056-1 068

[5] 王毅,齐华,郝重阳. 一种基于独立分量分析的模糊图像盲分离算法[J]. 计算机应用, 2006,26(10): 2 366-2 371

[6] 张金霞. 基于 ICA Fixed-Point 算法的信号图像分析[J]. 青海大学学报, 2005, 23(5): 75-77

第一作者简介:刘晓芹,博士生,主要研究方向为信息安全。
E-mail:liuxiaoqin1121@163.com

(下转第 1066 页)

[M]. 西安:西安电子科技大学出版社,2002

网络安全。
E-mail:pauljli@sina.com

第一作者简介:李建,博士生,主要研究方向为信息安全、计算机

Software Download Framework Based on Trust Measurement

LI Jian[1,3] *HE Yongzhong*[2] *XU Kaiyong*[1]

(1 Institute of Electronic Technology, Information Engineering University, 12 East Shangchengd Road, Zhengzhou 450004, China)

(2 School of Computer and Information Technology, Beijing Jiaotong University, Xizhimenwai, Shangyuanchun, Beijing 100044, China)

(3 GuangXi Military Area Headquarters, Zhiwuyuan Road, Nanning 530021, China)

Abstract: According to the fact that there are some secure faults in software download schemes due to the lack of trust measurement in software download for mobile equipments, making use of trust mechanism of identity management and secure characteristics such as integrity verification, protect storage, domain isolation, access control and remote platform verifying trusted computing possess, software download service model based on trust measurement of trusted identity management architecture has been put forward, software download service process has been designed, software download service protocol has been studied, the security performance of the protocol has been analyzed, the security performance comparison between several software download schemes has been made, the result shows that the scheme improve security performance of software download effectively.

Key words: trust measurement; trusted computing; identity management

About the first author: LI Jian, Ph.D candidate. His main interest is security of information and computer network.
E-mail: pauljli@sina.com

(上接第1061页)

Active Steganalysis for Stego-image Based on HMT and ICA

LIU Xiaoqin[1] *WANG Jiazhen*[1] *XU Bo*[1] *DENG Gaoming*[1]

(1 Dept. of Computer Engineering, Ordnance Engineering College, 97 West Heping Road, Shijiazhuang 050003, China)

Abstract: Two copies of stego-image are needed in the active steganalysis proposed by Chandramouli, as an improvement, a method of active steganalysis is proposed in which only one copy of stego-image is needed and the practicability is enhanced. An estimate of a cover image by wavelet-domain hidden Markov tree(HMT) is obtained; and the estimate is regarded as another stego-image which is different in the embed ratio. Then cover image and the secret message are separated by using the independent component analysis(ICA). Simulation experiments validate the advantage of HMT model, and give the results of blind separation.

Key words: active steganalysis; HMT model; independent component analysis; simulation

About the first author: LIU Xiaoqin, Ph. D candidate, she is mainly engaged in research on information security.
E-mail: liuxiaoqin1121@163.com