基于 Vague 集相似度量的图像隐写系统安全性测度

欧阳春娟[1),2),3)] 李 斌[1),2)] 李 霞[1),2)] 王 娜[1),2)]

[1)](深圳大学信息工程学院 广东 深圳 518060)

[2)](深圳市现代通信与信息处理重点实验室 广东 深圳 518060)

[3)](井冈山大学电子信息工程学院 江西 吉安 343009)

摘 要 由于图像隐写所引起的各种统计特征变化是不确定的,文中将 Vague 集相似度量引入隐写系统的安全性评价中.从图像的一阶统计特征和二阶统计特征两方面,定义了基于载体数据及载密数据相关 Vague 集相似度量的隐写系统一阶和二阶安全性测度,证明了该安全性测度的有界性,对称性和一致性.在假设图像满足独立同分布的条件下,证明了所提出的两种安全性测度是等价的.结果表明,所提出的二阶安全性比一阶安全性能更好地反映隐写引起的载体统计分布变化.与确定模式下的安全性测度相比,当嵌入率低于 0.5 比特/像素,新的测度可更好地度量隐写系统安全性,因此该测度对小容量的隐写及隐写分析算法设计具有更好的指导作用.

关键词 隐写;安全性;Vague 集;相似度量;有界性

中图法分类号 TP391 **DOI 号**: 10.3724/SP.J.1016.2012.01510

A New Security Evaluation for Steganographic System Based on Vague Set Similarity Measure

OUYANG Chun-Juan[1),2),3)] LI Bin[1),2)] LI Xia[1),2)] WANG Na[1),2)]

[1)](*College of Information and Engineering, Shenzhen University, Shenzhen, Guangdong* 518060)

[2)](*Shenzhen Key Laboratory of Modern Communications and Information Processing, Shenzhen, Guangdong* 518060)

[3)](*College of Electronics and Information Engineering, Jinggangshan University, Ji'an, Jiangxi* 343009)

Abstract Since variations of statistical features caused by image steganography are indeterministic, vague set similarity measure is introduced to steganographic system for security evaluation in this paper. The first-order and the second-order security measure based on vague set similarity measure are defined to measure the similarity between cover images and stego images in the sense of the first-order and the second-order image statistical feature, respectively. The newly defined security measures are then proved to have the properties of boundedness, symmetry and unity. Finally, under the assumption that the pixels' statistical distribution is identically independent distributed, the first-order and the second-order security measure are proved to be equivalent. Simulation results show that the second-order security measure can always capture more obvious statistical changes than the first-order security measure. In addition, compared with the security measure defined under a deterministic image statistical distribution model, the newly defined security measure is more sensitive when the embedding rate is lower than 0.5 bit per pixel, thus providing better guidance for the design of steganography and steganalysis with a small embedding rate.

Keywords steganography; security; vague set; similarity measure; boundedness

收稿日期:2011-01-26;最终修改稿收到日期:2011-05-26.本课题得到国家自然科学基金(61171124,61103174,60902069,61005049)、广东省科技计划项目(2011B010200045)、广东省高校优秀青年创新人才基金(LYM10116)、深圳市重点实验室提升项目(CXB201105060068A)资助.**欧阳春娟**,女,1974年生,博士研究生,副教授,主要研究方向为信息隐藏及分析.E-mail:oycj001@163.com.**李 斌**,男,1982年生,博士,讲师,主要研究方向为信息隐藏及模式识别.**李 霞**(通信作者),女,1968年生,博士,教授,博士生导师,主要研究领域为智能优化、智能计算及应用.E-mail:lixia@szu.edu.cn.**王 娜**,女,1977年生,博士,教授,主要研究领域为智能优化和模式识别.

1 引　言

隐写和隐写分析是信息隐藏研究领域的一个重要方面[1]，作为一种新的信息安全手段，其安全性研究必然是最为重要的课题. 在研究初期，大多数文献都集中在研究新的隐写方法上，导致算法的安全性往往依赖于算法的保密，这违背了 Kerckhoffs 准则[2]. 还有一些算法盲目追求大容量，而忽略了安全性考虑. 目前，完整的隐写技术理论体系尚未建立，对隐写系统安全性的讨论仍集中于具体的隐写和隐写分析方法，通用的安全性定义或模型方面的研究成果比较少. 然而建立完善的安全性测度模型是隐写系统最基本的问题，其对于隐写及隐写分析技术的研究起着重要的指导意义.

目前，隐写系统使用较广泛的安全性评价是 Cachin[3] 在假设获得了载体数据和载密数据确定的概率分布函数，提出的相关熵的 ε-secure 指标. Wang 等人[4] 将 Cachin 的结果扩展到多元高斯变量，对 Cachin 理论中的载体数据进行数学建模，该模型只考虑了由高斯平稳过程生成的载体，对复杂的载体图像，没有给出安全性定义. Sullivan 等人[5] 考虑图像像素相关性，将图像建模为马尔可夫链(Markov Chain，MC)模型，根据载体数据和载密数据对应 MC 经验矩阵的散度距离，定义了隐写系统的安全性. 文献[6] 陈丹等人采用图像可逆的去相关变换后的相对熵，定义了隐写系统的安全性. Ambalavanan 等人[7] 根据 Markov 随机场(Markov Random Feild，MRF)模型下的条件相对熵，提出了数字图像的隐写安全性定义. MRF 与 MC 相比，包含的图像相关性信息更全面，但模型复杂度及计算量都极大. Pevný 等人[8] 采用高维特征集对图像进行建模，定义了隐写系统最大平均差异(Maximum Mean Discrepancy，MMD)安全性测度. 张湛等人[9-10] 将图像建模为 n 阶马尔可夫链(n-MC)模型，采用 Hilbert 扫描方式构建图像的 n 阶马尔可夫链，利用其经验矩阵的散度距离定义隐写系统的安全性，对阶数 n 进行调节，可调整 n-MC 模型的计算复杂度.

上述文献都是在假设获得了确定的载体数据与载密数据的某种统计分布模式下，借鉴 Shannon 信息论中的通信模型思想描述隐写系统安全性. 然而，概率分布估计问题仍是目前尚未很好解决的统计学难题. 统计学描述的是观察空间整体的概率分布，其对于单个样本的概率分布描述是没有意义的. 在隐写及隐写分析的实际应用中，无法获得无限多的载体数据样本或载密数据样本，所以也无法获得载体数据和载密数据精确的概率分布. 此外，自然图像是一个非平稳过程，隐写将引起图像各局部细节变化，图像的各种统计特征会发生不确定性变化. 因此，上述文献采用确定的信息处理方法进行隐写系统安全性评价具有一定的局限性. 本文将图像隐写过程建模为一个模糊不确定性模型，利用载体数据与载密数据相关的 Vague 集相似度量更合理地定义了隐写系统的一阶及二阶安全性测度，并证明了该安全性测度的有界性、对称性和一致性. 在假设图像满足独立同分布条件下，证明了所提出的一阶和二阶安全性测度的等价性. 仿真实验验证了本文提出的基于 Vague 集相似度量的隐写系统安全性测度对不同隐写算法的通用性，并且该测度可区分不同隐写算法在同一嵌入率下的安全性. 测量同一算法在不同嵌入率下的安全性取值表明，二阶安全性测度比一阶安全性测度能更好地反映隐写引起的载体统计特征分布变化. 此外，与确定模式下安全性测度相比较，当嵌入率较低时，Vague 集相似度量的隐写安全性测度取得了更高的反应灵敏度值，说明该测度对小嵌入量的隐写及隐写分析算法设计具有更好的指导意义. 隐写算法设计实验也表明，Vague 集相似度量的安全性测度指导设计的隐写算法抗隐写分析能力更强，因此，该测度可以更好地指导设计高安全性的隐写算法.

2 确定模式下图像隐写系统一阶及二阶统计分布安全性测度

2.1 图像隐写系统一阶统计分布安全性测度

文献[3]的隐写安全性测度定义为：载体数据构成集合 C，假设载体图像样本集合 c 和载密图像样本集合 s 分别为 C 上的一个随机变量，且具有确定的概率分布函数，分别为 P_c，P_s. 采用 P_c 和 P_s 的相对熵(relative entropy)定义图像隐写系统的安全性. 从图像一阶统计特征角度，该定义可以理解如下：

载体图像样本集合为

$$\begin{bmatrix} x \\ P_c(x) \end{bmatrix} = \begin{bmatrix} x_1 & x_2 & \cdots & x_n \\ p_c(x_1) & p_c(x_2) & \cdots & p_c(x_n) \end{bmatrix};$$

载密图像样本集合为

$$\begin{bmatrix} x \\ P_s(x) \end{bmatrix}=\begin{bmatrix} x_1 & x_2 & \cdots & x_n \\ p_s(x_1) & p_s(x_2) & \cdots & p_s(x_n) \end{bmatrix},$$

其中，$x_1,x_2,\cdots,x_n$ 为图像像素值，所有像素值构成集合 X；$p_c(x_1)$，$p_c(x_2)$，$\cdots$，$p_c(x_n)$ 及 $p_s(x_1)$，$p_s(x_2)$，$\cdots$，$p_s(x_n)$ 为对应的载体图像与载密图像像素的概率分布. 定义 $P_c(x)$ 和 $P_s(x)$ 两个集合的相对熵（也称为 Kullback-Leibler 距离）为隐写系统的安全性测度.

$$D(P_c \parallel P_s)=\sum_{x\in X}P_c(x)\log\frac{P_c(x)}{P_s(x)} \tag{1}$$

可证明 $D(P_c \parallel P_s)\geq 0$. 当 $D(P_c \parallel P_s)=0$，称此隐写系统是绝对安全的；当 $0\leq D(P_c \parallel P_s)\leq \varepsilon$，称此隐写系统是 ε-安全的. 由于很难估计高维数据的概率分布函数，该安全性测度在实际应用中通常采用独立同分布等简单模型来估计概率分布函数.

2.2 图像隐写系统二阶统计分布安全性测度

为了更好地捕捉图像像素之间的相关性，文献[5]通过计算图像马尔可夫链模型二阶统计分布的散度距离，定义了图像隐写系统的二阶统计分布安全性. 设 C 和 S 分别为按某种给定的扫描方式（如行、列、zig-zag、Hilbert 等方式）扫描载体图像和载密图像得到的一阶马尔可夫链，$\boldsymbol{M}_c$、$\boldsymbol{M}_s$ 分别为 C 和 S 的经验矩阵. 定义 $\boldsymbol{M}_c$、$\boldsymbol{M}_s$ 的散度距离为图像隐写前后的二阶统计分布的改变，即二阶统计分布安全性测度.

$$D(\boldsymbol{M}_c,\boldsymbol{M}_s)=\sum_{i,j\in R}M^c_{ij}\log\left(\frac{M^c_{ij}}{\sum\limits_j M^c_{ij}}\frac{\sum\limits_j M^s_{ij}}{M^s_{ij}}\right) \tag{2}$$

其中 $\dfrac{M^c_{ij}}{\sum\limits_j M^c_{ij}}$，$\dfrac{M^s_{ij}}{\sum\limits_j M^s_{ij}}$ 分别代表载体图像和载密图像中从像素 i 到像素 j 的转移概率. 像素 $i,j\in R$，$[0,R]$ 为图像像素的取值范围. 可证明 $D(\boldsymbol{M}_c,\boldsymbol{M}_s)\geq 0$. 当且仅当 $\boldsymbol{M}_c=\boldsymbol{M}_s$ 时，$D(\boldsymbol{M}_c,\boldsymbol{M}_s)=0$，此时隐写系统是绝对安全的.

上述图像隐写系统一阶统计分布和二阶统计分布的安全性测度，分别从图像的一阶统计特征（如一维直方图特征），二阶统计特征（如差分直方图、二维直方图特征）两方面，对隐写系统的安全性给出了定义. 其对隐写系统的安全性评价具有一定的指导作用. 但这些安全性定义必须基于一个假设：假设无限的载体数据与载密数据的概率分布是可以明确得到的，其概率转移矩阵是确定的. 然而，图像表示的客观世界中自然景物的视觉信息非常丰富，包括灰度、颜色、纹理、形状等. 对图像数据处理形式也是多种多样的. 所以，图像隐写操作会引起图像各方面变化，同时存在很多不确定的因素（如在传输过程中受到噪声影响）. 隐写后的图像统计特征会产生动态变化，针对某一具体的图像像素值进行统计，它们隶属于某区域的隶属程度往往都是不确定的. 因此，本文将图像隐写过程建模为一个模糊不确定性模型，采用载体数据和载密数据相关 Vague 集相似度量可更合理地定义图像隐写系统的安全性.

3 Vague 集相关知识

模糊集方法只能处理确定的隶属和非隶属信息，对未确定的隶属信息无效[11]. Gau 和 Buehrer[12] 于 1993 年提出了 Vague 集理论，它将模糊集理论中的隶属函数单值扩充为一个区间. Vague 集理论采用真假隶属函数定义了元素对模糊概念的属于与不属于的程度和证据，与模糊集相比，具有更强的表达不确定性能力.

定义 1[12-13]. 设 $X=\{x_1,x_2,\cdots,x_n\}$ 是一个论域，$V(x)$ 表示 X 上的所有 Vague 集的集合，$V(x)$ 中任意的 Vague 集 A 可用一个真隶属函数 t_A 和一个假隶属函数 f_A 表示，$t_A:X\to[0,1]$，$f_A:X\to[0,1]$，其中 $t_A(x_i)$ 为由支持 x_i 的证据所导出的肯定隶属度的下界，$f_A(x_i)$ 则是由反对 x_i 的证据所导出的否定隶属度的下界，且 $t_A(x_i)+f_A(x_i)\leq 1$. 元素 x_i 在 Vague 集 A 中的隶属度被区间 $[0,1]$ 的一子区间 $[t_A(x),1-f_A(x)]$ 所界定，称该区间为 x_i 在 A 中的 Vague 值，记为 $V_A(x_i)$.

对于 Vague 集 A，当 X 离散时，记为

$$A=\sum_{i=1}^{n}[t_A(x_i),1-f_A(x_i)]/x_i,\ x_i\in X.$$

定义 2[14-15]. 假定 $X=\{x_1,x_2,\cdots,x_n\}$，A,B 为 X 上的两个 Vague 集，

$$E(A)=-\frac{1}{n\ln 2}\sum_{i=1}^{n}[t_A(x_i)\ln t_A(x_i)+f_A(x_i)\ln f_A(x_i)] \tag{3}$$

为 Vague 集 A 的熵（entropy）.

$$E_B(A)=-\sum_{i=1}^{n}[t_B(x_i)\ln t_A(x_i)+f_B(x_i)\ln f_A(x_i)] \tag{4}$$

为 Vague 集 A 关于 Vague 集 B 的偏熵（partial entropy）.

Vague 集合之间的相似度量可以很好地表示两个集合的匹配性[14-16]. 在实际的图像隐写系统安全性研究中,根本无法得到无限载体图像数据样本和载密图像数据样本精确的概率分布. 由于图像样本的有限性,图像内部不同景物之间存在明显的差异性,同一景物内部灰度、颜色、纹理具有的同一性,图像隐写引起这些性质改变的不确定性,导致载体图像数据样本集合和载密图像数据样本集合的概率分布等统计数据是不确定的. 因此,对于载体数据样本集合和载密数据样本集合的概率分布等统计数据采用离散的 Vague 集合来描述更合理.

4　基于 Vague 集相似度量的图像隐写系统安全性测度

对于隐写系统,安全性是其首要指标,包括不可检测性和不可感知性两方面要求. 在保证载体样本与载密样本之间满足一定的感知失真限制下,一个隐写系统可描述如下:

称系统 $\Sigma=(F,G,M,C,K,S)$ 为隐写系统,其中 C 为载体数据样本集合,M 为消息源,S 为载密数据样本集合,K 为密钥源,满足

(1) $F: C\times M\times K\to S$ 为隐写映射;

(2) $G: S\times K\to M$ 为提取映射;

(3) 对任取 $c\in C, m\in M, k\in K$, $G(F(c,m,k),k)=m$.

4.1　Vague 集相似度量的图像隐写系统一阶安全性测度

定义 3.　对隐写系统 $\Sigma=(F,G,M,C,K,S)$,其中 C 为载体信号集合,M 为消息源,S 为载密信号集合,K 为密钥源,定义 PC 和 PS 分别为载体图像数据样本集合 C 和载密图像数据样本集合 S 的概率分布集合. 所有样本的概率分布集合组成论域 $X=\{x_1,x_2,\cdots,x_n\}$, $V(x)$ 表示 X 上所有 Vague 集的集合,则 PC 和 PS 为 X 上的两个 Vague 集,

$$PC=\sum_{i=1}^{n}[t_{PC}(x_i),1-f_{PC}(x_i)]/x_i,\ x_i\in X,$$

$$PS=\sum_{i=1}^{n}[t_{PS}(x_i),1-f_{PS}(x_i)]/x_i,$$

$$x_i\in X,\ i\in[1,n],$$

$[1,n]$ 为图像像素的取值范围. 将 Vague 集 PC 和 PS 的相似度量 $T(PC,PS)$ 定义为图像隐写系统的一阶安全性测度:

$$T(PC,PS)=\frac{n\ln 2(E(PC)+E(PS))}{E_{PC}(PS)+E_{PS}(PC)}\tag{5}$$

其中 $E(PC)$, $E(PS)$ 分别为 Vague 集 PC 和 PS 的熵, $E_{PC}(PS)$ 为 Vague 集 PS 关于 PC 的偏熵, $E_{PS}(PC)$ 为 Vague 集 PC 关于 PS 的偏熵. 规定 $0\times\ln 0=0$.

当 $T(PC,PS)=1$,隐写系统是绝对安全的. 当 $T(PC,PS)=T$, $T\in(0,1)$, $\varepsilon=1-T$,则隐写系统是 ε-安全的. 当 PC 完全隶属于 X,而 PS 完全不隶属于 X,即当且仅当 $V_{PC}(x_i)=[1,1,\cdots,1]$, $V_{PS}(x_i)=[0,0,\cdots,0]$ 时, $T(PC,PS)$ 取最小值 0,此时隐写系统是绝对不安全的.

定理 1.　设 $T(PC,PS)$ 为图像隐写系统基于 Vague 集相似度量的一阶安全性测度. 则 $T(PC,PS)$ 有如下性质:

(1) 有界性. $0\le T(PC,PS)\le 1$;

(2) 对称性. $T(PC,PS)=T(PS,PC)$;

(3) 一致性. $T(PC,PS)=1\Leftrightarrow PC=PS$.

证明.　(1)

$$\begin{aligned}&E_{PC}(PS)+E_{PS}(PC)-n\ln 2(E(PC)+E(PS))\\&=-\sum_{i=1}^{n}[t_{PC}(x_i)\ln t_{PS}(x_i)+f_{PC}(x_i)\ln f_{PS}(x_i)+\\&\quad t_{PS}(x_i)\ln t_{PC}(x_i)+f_{PS}(x_i)\ln f_{PC}(x_i)]+\\&\quad\frac{n\ln 2}{n\ln 2}\sum_{i=1}^{n}[t_{PC}(x_i)\ln t_{PC}(x_i)+f_{PC}(x_i)\ln f_{PC}(x_i)+\\&\quad t_{PS}(x_i)\ln t_{PS}(x_i)+f_{PS}(x_i)\ln f_{PS}(x_i)]\\&=\sum_{i=1}^{n}\left[t_{PC}(x_i)\ln\frac{t_{PC}(x_i)}{t_{PS}(x_i)}+f_{PC}(x_i)\ln\frac{f_{PC}(x_i)}{f_{PS}(x_i)}+\right.\\&\quad\left.t_{PS}(x_i)\ln\frac{t_{PS}(x_i)}{t_{PC}(x_i)}+f_{PS}(x_i)\ln\frac{f_{PS}(x_i)}{f_{PC}(x_i)}\right].\end{aligned}$$

根据不等式: $\ln x\ge 1-\frac{1}{x}$,上式有

$$\begin{aligned}&\ge\sum_{i=1}^{n}\left[t_{PC}(x_i)\left(1-\frac{t_{PS}(x_i)}{t_{PC}(x_i)}\right)+f_{PC}(x_i)\left(1-\frac{f_{PS}(x_i)}{f_{PC}(x_i)}\right)+\right.\\&\quad\left.t_{PS}(x_i)\left(1-\frac{t_{PC}(x_i)}{t_{PS}(x_i)}\right)+f_{PS}(x_i)\left(1-\frac{f_{PC}(x_i)}{f_{PS}(x_i)}\right)\right]\\&=\sum_{i=1}^{n}[t_{PC}(x_i)-t_{PS}(x_i)+f_{PC}(x_i)-f_{PS}(x_i)+\\&\quad t_{PS}(x_i)-t_{PC}(x_i)+f_{PS}(x_i)-f_{PC}(x_i)]\\&=0,\end{aligned}$$

所以

$$E_{PC}(PS)+E_{PS}(PC)-n\ln 2(E(PC)+E(PS))\ge 0,$$

所以

$$E_{PC}(PS)+E_{PS}(PC)\ge n\ln 2(E(PC)+E(PS)),$$

又因为 $t_{PC}(x_i)$, $t_{PS}(x_i)$, $f_{PC}(x_i)$, $f_{PS}(x_i)$ 都属于

$[0,1]$,且规定 $0\times\ln 0=0$,

所以 $E(PC),E(PS),E_{PC}(PS),E_{PS}(PC)$都是非负,

所以 $0\leq T(PC,PS)\leq 1$.

(2) $T(PC,PS)=\dfrac{n\ln 2(E(PC)+E(PS))}{E_{PC}(PS)+E_{PS}(PC)}$,而

$$T(PS,PC)=\frac{n\ln 2(E(PS)+E(PC))}{E_{PS}(PC)+E_{PC}(PS)}.$$

所以 $T(PC,PS)=T(PS,PC)$.

(3) 在证明(1)中

$$E_{PC}(PS)+E_{PS}(PC)-n\ln 2(E(PC)+E(PS))$$
$$=\sum_{i=1}^{n}\left[t_{PC}(x_i)\ln\frac{t_{PC}(x_i)}{t_{PS}(x_i)}+f_{PC}(x_i)\ln\frac{f_{PC}(x_i)}{f_{PS}(x_i)}+\right.$$
$$\left.t_{PS}(x_i)\ln\frac{t_{PS}(x_i)}{t_{PC}(x_i)}+f_{PS}(x_i)\ln\frac{f_{PS}(x_i)}{f_{PC}(x_i)}\right].$$

上式等于 0 的充要条件是

$$t_{PC}(x_i)=t_{PS}(x_i),f_{PC}(x_i)=f_{PS}(x_i),$$

所以 $T(PC,PS)=1\Leftrightarrow PC=PS$. 证毕.

4.2 Vague 集相似度量的图像隐写系统二阶安全性测度

基于 Vague 集相似度量的图像隐写系统一阶安全性测度定义受到假设图像像素点是独立同分布的限制.事实上,图像各像素之间存在很强的相关性.本文进一步考虑图像像素相关性,利用图像马尔可夫链模型二阶统计分布的 Vague 集相似度量定义隐写系统的二阶安全性测度.首先,对载体图像和载密图像,按某种给定的扫描方式(如行、列、zig-zag、Hilbert 等方式)分别得到马尔可夫链序列 C 和 S,$\boldsymbol{MC}$、$\boldsymbol{MS}$ 分别为 C 和 S 的经验矩阵.同理,将经验矩阵 $\boldsymbol{MC}$、$\boldsymbol{MS}$ 定义为两个 Vague 集,可更准确地描述图像像素转移的不确定性.

定义 4. 设 C 和 S 分别为按给定扫描方式得到的载体图像和载密图像的马尔可夫链,$\boldsymbol{MC}$、$\boldsymbol{MS}$ 分别为 C 和 S 的经验矩阵.m_{ij} 为经验矩阵的元素,表示从像素 i 到像素 j 的转移概率.所有经验矩阵的元素组成论域 M_{ij},$V(m)$ 表示 M_{ij} 上的所有 Vague 集的集合,则 $\boldsymbol{MC},\boldsymbol{MS}$ 为论域 M_{ij} 上的两个 Vague 集,

$$\boldsymbol{MC}=\sum_{i,j=1}^{n}[t_{\boldsymbol{MC}}(m_{ij}),1-f_{\boldsymbol{MC}}(m_{ij})]/m_{ij},\ m_{ij}\in M_{ij},$$

$$\boldsymbol{MS}=\sum_{i,j=1}^{n}[t_{\boldsymbol{MS}}(m_{ij}),1-f_{\boldsymbol{MS}}(m_{ij})]/m_{ij},\ m_{ij}\in M_{ij}.$$

将 Vague 集 $\boldsymbol{MC}$ 和 $\boldsymbol{MS}$ 的相似度量 $T(\boldsymbol{MC},\boldsymbol{MS})$ 定义为图像隐写系统的二阶安全性测度:

$$T(\boldsymbol{MC},\boldsymbol{MS})=\frac{n\ln 2(E(\boldsymbol{MC})+E(\boldsymbol{MS}))}{E_{\boldsymbol{MC}}(\boldsymbol{MS})+E_{\boldsymbol{MS}}(\boldsymbol{MC})}\tag{6}$$

其中

$$E(\boldsymbol{MC})=-\frac{1}{n\ln 2}\sum_{i,j=1}^{n}[t_{\boldsymbol{MC}}(m_{ij})\ln t_{\boldsymbol{MC}}(m_{ij})+f_{\boldsymbol{MC}}(m_{ij})\ln f_{\boldsymbol{MC}}(m_{ij})]\tag{7}$$

$$E_{\boldsymbol{MS}}(\boldsymbol{MC})=-\sum_{i,j=1}^{n}[t_{\boldsymbol{MS}}(m_{ij})\ln t_{\boldsymbol{MC}}(m_{ij})+f_{\boldsymbol{MS}}(m_{ij})\ln f_{\boldsymbol{MC}}(m_{ij})]\tag{8}$$

$E(\boldsymbol{MC})$ 为 Vague 集 $\boldsymbol{MC}$ 的熵,$E_{\boldsymbol{MS}}(\boldsymbol{MC})$ 为 Vague 集 $\boldsymbol{MC}$ 关于 Vague 集 $\boldsymbol{MS}$ 的偏熵 $i,j\in[1,n]$,$[1,n]$为图像像素的取值范围.同理可定义 $E(\boldsymbol{MS}),E_{\boldsymbol{MC}}(\boldsymbol{MS})$.规定 $0\times\ln 0=0$.

当 $T(\boldsymbol{MC},\boldsymbol{MS})=1$,隐写系统是绝对安全的.当 $T(\boldsymbol{MC},\boldsymbol{MS})=T$,$T\in(0,1)$,$\varepsilon=1-T$,系统是 ε-安全的.当 $\boldsymbol{MC}$ 完全隶属于 M_{ij},而 $\boldsymbol{MS}$ 完全不隶属于 M_{ij},即当且仅当 $V_{\boldsymbol{MC}}(m_{ij})=[1,1,\cdots,1]$,$V_{\boldsymbol{MS}}(m_{ij})=[0,0,\cdots,0]$时,$T(\boldsymbol{MC},\boldsymbol{MS})$取得最小值 0,此时,隐写系统是绝对不安全的.

定理 2. 设 $T(\boldsymbol{MC},\boldsymbol{MS})$为图像隐写系统基于 Vague 集相似度量的二阶安全性测度.则 $T(\boldsymbol{MC},\boldsymbol{MS})$有如下性质:

(1) 有界性. $0\leq T(\boldsymbol{MC},\boldsymbol{MS})\leq 1$;

(2) 对称性. $T(\boldsymbol{MC},\boldsymbol{MS})=T(\boldsymbol{MS},\boldsymbol{MC})$;

(3) 一致性. $T(\boldsymbol{MC},\boldsymbol{MS})=1\Leftrightarrow\boldsymbol{MC}=\boldsymbol{MS}$.

定理 2 的证明可参照定理 1 证明(略).

定理 3. 假设载体图像与载密图像的像素分布满足独立同分布条件,则图像隐写系统基于 Vague 集相似度量的一阶安全性测度 $T(PC,PS)$ 与二阶安全性测度 $T(\boldsymbol{MC},\boldsymbol{MS})$等价.

证明. 设载体图像与载密图像的像素分布满足独立同分布条件,m_{ij} 为马尔可夫链 MC 模型经验矩阵中的元素,i,j 为图像像素值,x_i,x_j 分别为像素值 i,j 的概率分布.则

$$t_{\boldsymbol{MC}}(m_{ij})=t_{PC}(x_ix_j)=t_{PC}(x_i)t_{PC}(x_j),$$
$$f_{\boldsymbol{MC}}(m_{ij})=f_{PC}(x_ix_j)=f_{PC}(x_i)f_{PC}(x_j),$$
$$t_{\boldsymbol{MS}}(m_{ij})=t_{PS}(x_ix_j)=t_{PS}(x_i)t_{PS}(x_j),$$
$$f_{\boldsymbol{MS}}(m_{ij})=f_{PS}(x_ix_j)=f_{PS}(x_i)f_{PS}(x_j).$$

且
$$\sum_{j=1}^{n}t_{PC}(x_j)=1,\ \sum_{j=1}^{n}f_{PC}(x_j)=1,$$
$$\sum_{j=1}^{n}t_{PS}(x_j)=1,\ \sum_{j=1}^{n}f_{PS}(x_j)=1.$$

由此可得

$$T(\boldsymbol{MC},\boldsymbol{MS})=\frac{n\ln 2(E(\boldsymbol{MC})+E(\boldsymbol{MS}))}{E_{\boldsymbol{MC}}(\boldsymbol{MS})+E_{\boldsymbol{MS}}(\boldsymbol{MC})}$$

$$=\frac{\sum_{i,j=1}^{n}[t_{\boldsymbol{MC}}(m_{ij})\ln t_{\boldsymbol{MC}}(m_{ij})+f_{\boldsymbol{MC}}(m_{ij})\ln f_{\boldsymbol{MC}}(m_{ij})+\cdots+t_{\boldsymbol{MS}}(m_{ij})\ln t_{\boldsymbol{MS}}(m_{ij})+f_{\boldsymbol{MS}}(m_{ij})\ln f_{\boldsymbol{MS}}(m_{ij})]}{\sum_{i,j=1}^{n}[t_{\boldsymbol{MC}}(m_{ij})\ln t_{\boldsymbol{MS}}(m_{ij})+f_{\boldsymbol{MC}}(m_{ij})\ln f_{\boldsymbol{MS}}(m_{ij})+\cdots+t_{\boldsymbol{MS}}(m_{ij})\ln t_{\boldsymbol{MC}}(m_{ij})+f_{\boldsymbol{MS}}(m_{ij})\ln f_{\boldsymbol{MC}}(m_{ij})]}$$

$$=\frac{\sum_{i=1}^{n}\sum_{j=1}^{n}[t_{PC}(x_ix_j)\ln t_{PC}(x_ix_j)+f_{PC}(x_ix_j)\ln f_{PC}(x_ix_j)+\cdots+t_{PS}(x_ix_j)\ln t_{PS}(x_ix_j)+f_{PS}(x_ix_j)\ln f_{PS}(x_ix_j)]}{\sum_{i=1}^{n}\sum_{j=1}^{n}[t_{PC}(x_ix_j)\ln t_{PS}(x_ix_j)+f_{PC}(x_ix_j)\ln t_{PS}(x_ix_j)+\cdots+t_{PS}(x_ix_j)\ln t_{PC}(x_ix_j)+f_{PS}(x_ix_j)\ln f_{PC}(x_ix_j)]}$$

$$=\Big(\sum_{i=1}^{n}\Big[t_{PC}(x_i)\sum_{j=1}^{n}t_{PC}(x_j)\ln\Big(t_{PC}(x_i)\sum_{j=1}^{n}t_{PC}(x_j)\Big)+\cdots+f_{PC}(x_i)\sum_{j=1}^{n}f_{PC}(x_j)\ln\Big(f_{PC}(x_i)\sum_{j=1}^{n}f_{PC}(x_j)\Big)+\cdots+$$
$$t_{PS}(x_i)\sum_{j=1}^{n}t_{PS}(x_j)\ln\Big(t_{PS}(x_i)\sum_{j=1}^{n}t_{PS}(x_j)\Big)+\cdots+f_{PS}(x_i)\sum_{j=1}^{n}f_{PS}(x_j)\ln\Big(f_{PS}(x_i)\sum_{j=1}^{n}f_{PS}(x_j)\Big)\Big]\Big)\Big/$$
$$\Big(\sum_{i=1}^{n}\Big[t_{PC}(x_i)\sum_{j=1}^{n}t_{PC}(x_j)\ln\Big(t_{PS}(x_i)\sum_{j=1}^{n}t_{PS}(x_j)\Big)+\cdots+f_{PC}(x_i)\sum_{j=1}^{n}f_{PC}(x_j)\ln(f_{PS}(x_i)\sum_{j=1}^{n}f_{PS}(x_j))+\cdots+$$
$$t_{PS}(x_i)\sum_{j=1}^{n}t_{PS}(x_j)\ln\Big(t_{PC}(x_i)\sum_{j=1}^{n}t_{PC}(x_j)\Big)+\cdots+f_{PS}(x_i)\sum_{j=1}^{n}f_{PS}(x_j)\ln\Big(f_{PC}(x_i)\sum_{j=1}^{n}f_{PC}(x_j)\Big)\Big]\Big)$$

$$=\frac{\sum_{i=1}^{n}[t_{PC}(x_i)\ln t_{PC}(x_i)+f_{PC}(x_i)\ln f_{PC}(x_i)+\cdots+t_{PS}(x_i)\ln t_{PS}(x_i)+f_{PS}(x_i)\ln f_{PS}(x_i)]}{\sum_{i=1}^{n}[t_{PC}(x_i)\ln t_{PS}(x_i)+f_{PC}(x_i)\ln f_{PS}(x_i)+\cdots+t_{PS}(x_i)\ln t_{PC}(x_i)+f_{PS}(x_i)\ln f_{PC}(x_i)]}$$

$$=\frac{n\ln 2(E(PC)+E(PS))}{E_{PC}(PS)+E_{PS}(PC)}$$

$$=T(PC,PS).$$

证毕.

本文从 Vague 集的相似度量出发，定义了图像隐写系统的一阶和二阶安全性测度，并证明了该安全性测度满足有界性、对称性和一致性. 该安全性测度的有界性克服了图像在确定模式下文献[3]及文献[5]提出的图像隐写安全性测度取无穷大的缺陷，使隐写系统安全性测度对隐写引起载体的统计分布改变限定在一定范围内，从而对统计分布改变的反映更为明显. 当载体数据和载密数据相关的两个 Vague 集完全相同时，安全性测度取得最大值 1，此时隐写系统是绝对安全的. 当载体数据相关的 Vague 集完全地隶属于所对应的论域，而载密数据相关的 Vague 集完全不属于所对应的论域时，安全性测度取得最小值 0，此时隐写系统是绝对不安全的. 定理 3 证明了当图像分布满足独立同分布时，基于 Vague 集相似度量的图像隐写系统的一阶安全性与二阶安全性是等价的，从理论上统一了基于 Vague 集相似度量的图像隐写安全性测度.

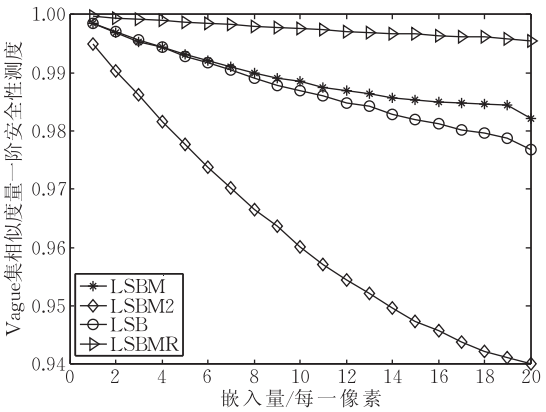
5　实验结果与讨论

由以上的理论分析及证明可知，基于 Vague 集相似度量的图像隐写系统安全性测度同确定模式下的图像隐写系统安全性测度相比更具合理性. 为了说明该测度在处理实际隐写系统安全性问题时具有以上优势，实验首先采用 Vague 集相似度量的图像隐写系统一阶和二阶安全性测度度量不同隐写算法在不同嵌入率下的安全性，验证该测度的有效性和通用性. 其次，采用本文提出的一阶和二阶安全性测度分别与文献[3]及文献[5]提出的确定模式下的安全性测度相比较，通过度量同一隐写算法在不同嵌入率下的安全性，说明 Vague 集相似度量的隐写系统安全性测度的性能优势. 最后，用两种模式的安全性测度指导设计隐写算法，通过比较隐写算法的抗分析能力，说明 Vague 集相似度量的隐写系统安全性测度具有更好的指导设计隐写算法能力.

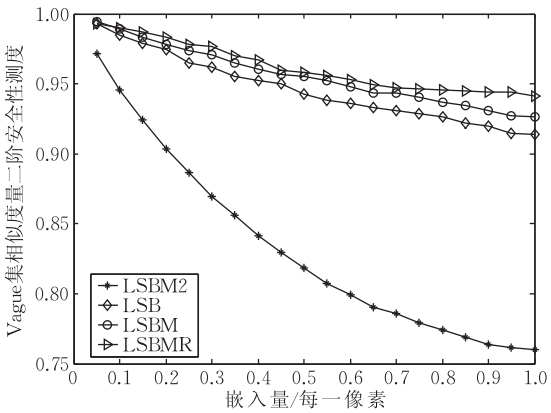
5.1　验证 Vague 集相似度量的图像隐写系统安全性测度的通用性

实验采用两组不同隐写算法，对图像进行隐写操作，比较不同隐写算法在同一嵌入率下的 Vague 集相似度量的图像隐写系统安全性. 第 1 组为空域隐写算法，包括 LSB 替换[17]、LSB±1[18]、LSB±2 和 LSB Matching Revisited[19] 4 种隐写算法（分别

记作 LSB、LSBM、LSBM2 和 LSBMR). 第 2 组隐写算法为 DCT 域隐写算法，包括 JSteg①、F3、F4 和 F5[20] 4 种隐写算法. 在第 1 组隐写算法中，采用 NRCS 图像库② 中的 1542 幅图像，将其裁剪为 1024×1024，并转化为灰度图像. 在每幅图像中嵌入相同的秘密信息，嵌入率从 0.05 bpp 增至 1 bpp，步长为 0.05 bpp，对图像进行隐写得到载密图像. 针对每种隐写算法，计算 1542 幅图像的 Vague 集相似度量的一阶和二阶安全性测度 $T(PC,PS)$ 和 $T(\boldsymbol{MC},\boldsymbol{MS})$ 在同一嵌率下的均值，如图 1 所示. 在第 2 组隐写算法中，4 种算法都是在图像分块的不为零 DCT 系数上嵌入秘密信息，数据嵌入率较低. 为了比较不同算法的安全性，选取 USC-SIPI 标准图像库③ “girl.bmp”、“lake.bmp”、“lena.bmp” 和 “baboon.bmp”4 个标准图像，在不同的图像中嵌入相同容量的秘密信息，计算这 4 幅图像分别针对不同 DCT 域隐写算法的 Vague 集相似度量隐写系统的一阶和二阶安全性测度，如表 1 所示.



(a) 一阶安全性 $T(PC,PS)$ 随嵌入率的改变图



(b) 二阶安全性 $T(\boldsymbol{MC},\boldsymbol{MS})$ 随嵌入率的改变图

图 1 不同隐写算法在不同嵌入率下的 Vague 集相似度量安全性测度

表 1 不同 DCT 域隐写算法在相同嵌入容量下的 Vague 度集相似度量的安全性测度

图像	一阶安全性测度 $T(PC,PS)$				二阶安全性测度 $T(\boldsymbol{MC},\boldsymbol{MS})$			
	Jsteg	F3	F4	F5	Jsteg	F3	F4	F5
girl.bmp	0.98219	0.9837	0.98533	0.98610	0.92327	0.92573	0.92791	0.92933
lake.bmp	0.83355	0.83427	0.83838	0.83918	0.81342	0.81475	0.81635	0.81894
lena.bmp	0.96555	0.97155	0.97239	0.97673	0.93456	0.93657	0.93678	0.93825
baboon.bmp	0.93130	0.93470	0.93673	0.93856	0.91146	0.91388	0.91659	0.91713

图 1(a)为针对第 1 组隐写算法的一阶安全性测度 $T(PC,PS)$ 均值随嵌入率增加时的变化曲线图. 由图 1(a)可知，所有的曲线都满足随着嵌入率增加，$T(PC,PS)$ 下降的趋势，即符合隐写系统随着嵌入率增加，安全性下降的规律. 从图 1(a)可得，对于空域中的 4 种隐写算法，在同样的嵌入率下，不同隐写算法的安全性是不同的. LSBMR 的 $T(PC,PS)$ 取值最大，安全性最高，其次是 LSBM、LSB，最低是 LSBM2. 其中，LSBM2 隐写的安全性与其它 3 种相比，安全性下降较大. 该实验结果与 LSBMR 的改变率为 0.375/像素，LSBM、LSB 改变率为 0.5/像素，LSBM2 改变率为 1.0/像素的理论分析得出 4 种算法的安全性排序完全相符. 图 1(b)为针对第 1 组空域隐写算法的 Vague 集相似度量的二阶安全性测度 $T(\boldsymbol{MC},\boldsymbol{MS})$ 均值随嵌入率增加时的变化曲线图，可得到类似的分析结果. 表 1 为在相同嵌入量下，DCT 域中 4 种隐写算法的 Vague 集相似度量的隐写系统一阶安全性 $T(PC,PS)$ 和二阶安全性 $T(\boldsymbol{MC},\boldsymbol{MS})$ 取值. 由表 1 可知，对于同一图像在相同的嵌入量下，F5 算法的 $T(PC,PS)$ 和 $T(\boldsymbol{MC},\boldsymbol{MS})$ 取值与其它 3 种隐写算法的安全性相比，取值最大，安全性最高，其次是 F4 算法、F3 算法，最低是 JSteg 隐写. 该实验结果与这 4 种隐写算法的理论安全性分析排序相符合. 由此可见，本文提出的基于 Vague 集相似度量的图像隐写系统安全性测度具有通用性和有效性.

5.2 本文提出的隐写安全性测度与确定模式下安全性测度相比的优越性

本实验采用Vague集相似度量的隐写系统一

① Upham D. Jsteg. http://zooid.org/~paul/crypto/jsteg/

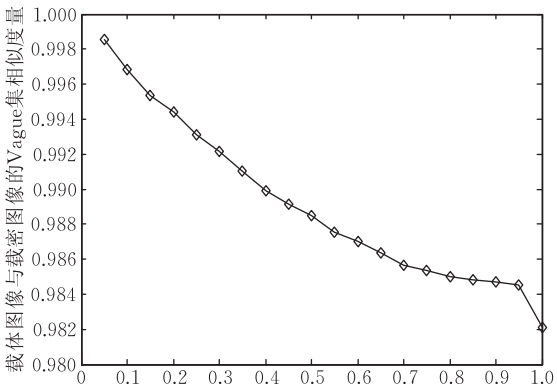
② Natural resources conservation service photo gallery[EB/OL], available: http://photogallery.nrcs.usda.gov/

③ USC-SIPI Image Database [DB/OL]. Available: http://sipi.usc.edu/services/database/index.html

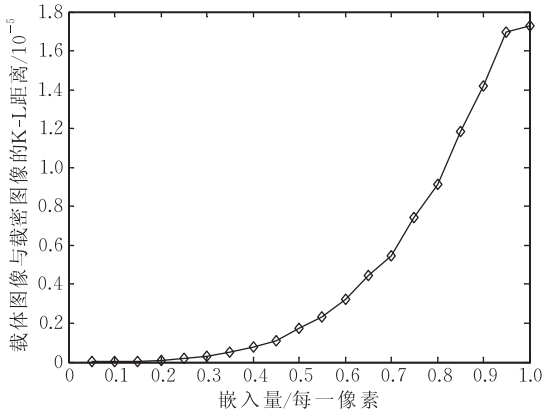
阶和二阶安全性测度分别与文献[3]在假设图像分布为独立同分布时的相对熵安全性测度，及文献[5]在确定模式下的安全性测度进行比较. 通过比较两种模式下的测度对同一隐写算法在不同嵌入率下引起的图像统计分布变化的反应灵敏度，从数学角度定量地说明模糊 Vague 集隐写系统一阶和二阶安全性测度的优越性. 实验同样采用 NRCS 图像库中的 1542 幅图像，将其转化为灰度图像，裁剪大小为 1024×1024. 嵌入率从 0.05 bpp 增至 1 bpp，步长为 0.05 bpp，采用 LSB±1 分别进行隐写嵌入. 计算在同一嵌入率下，针对 1542 幅图像的 Vague 集相似度量的一阶安全性 $T(PC,PS)$ 均值和确定模式下文献[3]所提的一阶统计分布的安全性 $D(P_c \| P_s)$ 均值. 其中图 2(a)为 $T(PC,PS)$ 均值随嵌入率增加时的变化曲线图，图 2(b)为 $D(P_c \| P_s)$ 均值随嵌入率增加时的变化曲线图. 定义 $\delta=\Delta y/y$ 为安全性测度对隐写引起变化的反应灵敏度，其中 Δy 为在一定的嵌入率变化下安全性测度值的变化量，y 为嵌入率 0～1 变化时安全性测度的总变化量. 对于图 2(a)，$\delta_T=(\Delta|T(PC,PS)|)/(|T(PC,PS)|)$，对于图 2(b)，$\delta_D=(\Delta|D(P_c \| P_s)|)/(|D(P_c \| P_s)|)$. δ 取值越大，说明该安全性测度在嵌入率变化下，对隐写所引起的统计分布反映更灵敏，从而能更精确地对隐写安全性进行度量. 由图 2 可得出以下结论：

(1) 由图 2(a)可知，$T(PC,PS)$ 的取值范围为有限区间[0,1]，随着嵌入率的增加，$T(PC,PS)$ 取值变小，隐写系统的安全性逐渐下降. 由图 2(b)可知，$D(P_c \| P_s)$ 的取值范围为无限区间$[0,+\infty)$随着嵌入率的增加，$D(P_c \| P_s)$ 取值增大，安全性降低. 两种描述相比，采用 $T(PC,PS)$ 的取值范围在有限区间中，取值范围更好控制，描述隐写系统安全性符合随着嵌入率增加，安全性下降，安全测度取值变小的单调下降规律.

(2) 由图 2(b)可知，当嵌入率从 0～0.5 变化时，$\delta_D\approx(0.2\times10^{-5})/(2\times10^{-5})=0.1$，说明当嵌入率小于 0.5 时，其反应灵敏度较低. 而由图 2(a)可知，在嵌入率从 0～0.5 的变化过程中，$\delta_T\approx(1-0.988)/(1-0.982)\approx0.67$，$\delta_T$ 约为 δ_D 的 7 倍. 且在整个嵌入率变化过程中，δ_T 反应灵敏度取值都比较一致，说明 Vague 集一阶安全性测度可以清晰地度量不同嵌入率下的隐写安全性. 而嵌入率较低时，确定模式下的隐写安全性测度 $D(P_c \| P_s)$ 不能清晰地反映隐写引起的各种统计分布变化.



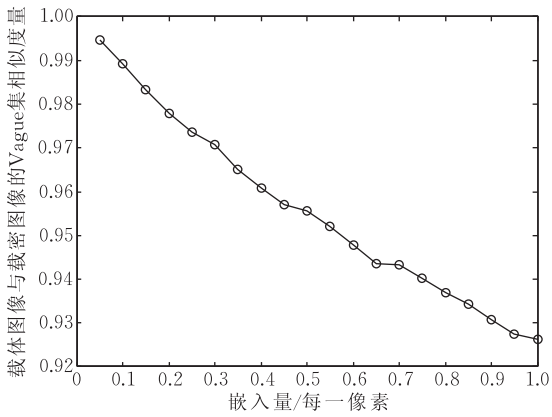
(a) $T(PC,PS)$随嵌入率增加的变化图



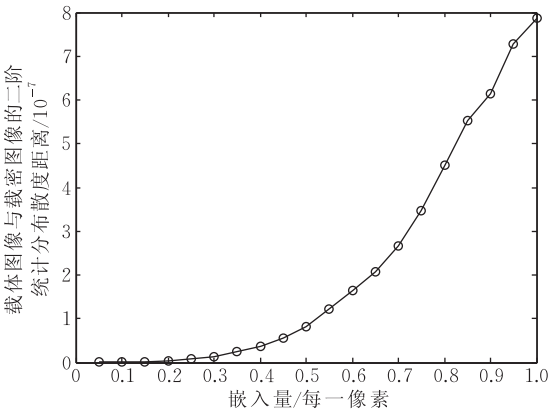
(b) $D(P_c \| P_s)$随嵌入率增加的变化图

图 2　Vague 集一阶安全性与文献[3]安全性随嵌入量增加的变化图

图 3(a)、(b)是在嵌入量从 0.05 bpp 增至 1 bpp，步长为 0.05 bpp 下，Vague 集相似度量二阶安全性测度 $T(\boldsymbol{MC},\boldsymbol{MS})$ 均值和文献[5]确定模式下的安全性测度 $D(\boldsymbol{M}_c,\boldsymbol{M}_s)$ 均值随嵌入率增加的变化曲线图. 可得出与图 2 相类似的结论. 图 2(a)和图 3(a)分别为采用 Vague 集相似度量的一阶安全性测度 $T(PC,PS)$ 和二阶安全性测度 $T(\boldsymbol{MC},\boldsymbol{MS})$，对同一种隐写算法 LSBM 在嵌入率从 0～1 变化时的安全性测度取值. 由两图可得，一阶安全性测度 $T(PC,PS)$ 取值在[0.98211,0.99855]之间，测度变化总量为 0.01644. 而二阶安全性测度 $T(\boldsymbol{MC},\boldsymbol{MS})$ 取值在[0.92617,0.99465]之间，测度变化总量为 0.06848，约为 $T(PC,PS)$ 变化总量的 4 倍. 可见，$T(\boldsymbol{MC},\boldsymbol{MS})$ 能更好地反映隐写引起的载体数据统计特征变化. 因此，二阶安全性测度 $T(\boldsymbol{MC},\boldsymbol{MS})$ 与一阶安全性测度 $T(PC,PS)$ 相比，二阶安全性测度对隐写及隐写分析算法的设计具有更好的指导作用.



(a) $T(\boldsymbol{MC}, \boldsymbol{MS})$随嵌入率增加的变化图



(b) $D(\boldsymbol{M}_c, \boldsymbol{M}_s)$随嵌入率增加的变化图

图 3 Vague 集二阶安全性与文献[5]安全性随嵌入量增加的变化图

5.3 两种模式下的安全性测度指导设计隐写算法比较实验

本实验采用 Vague 集相似度量的隐写系统一阶和二阶安全性测度分别与文献[3]及文献[5]在确定模式下的安全性测度来指导设计隐写算法. 通过比较两种模式的安全性测度指导设计的隐写算法的抗隐写分析能力,来说明本文提出的模糊 Vague 集隐写系统安全性测度对隐写算法设计具有更好的指导作用. 在 LSB±1 隐写(记作 LSBM)算法中,当秘密信息与载体像素的最低有效位相同时,不更改载体图像像素,当两者不相同时,对载体像素进行随机加减 1 得到载密图像. 本实验采用粒子群算法[21]优化隐写嵌入过程中需要修改像素的加减 1 序列. 通过对图像不同的加减 1,尽量保持图像的统计特征,提高隐写算法的安全性. 粒子的初解为 0,1 序列,其中 1 代表加 1,0 代表减 1,序列的长度为隐写过程中需进行加减 1 的像素个数. 在优化过程中,首先以文献[3]中确定模式下的安全性测度 $D(P_c \parallel P_s)$为优化目标,该目标函数取值越小,隐写图像的安全性越好,得到以确定模式下的一阶统计分布的安全性测度为指导的改进 LSBM 隐写算法(记作 PSO_LSBMD1). 再以 Vague 集相似度量的一阶安全测度 $T(PC, PS)$为目标函数进行优化,该目标函数取值越大,隐写图像的安全性越好,得到以 Vague 集相似度量的一阶安全性测度为指导的改进 LSBM 隐写算法(记作 PSO_LSBMT1). 针对 NRCS 图像库中的 1542 幅图像,全部转化为灰度图像,裁剪大小为 512×512,粒子数为 30,迭代次数为 20,用以上 3 种算法在嵌入率为 1.0 bpp 下得到加密图库. 采用文献[22]中(该文献以一维质心和二维质心特征进行隐写分析)的隐写分析算法,对以上 3 种隐写算法得到的加密图库及原图库提取特征进行隐写分析. 用 Fisher 作为分类器,其中 500 幅图像为训练集,其余的图像为测试集. 实验得到的 ROC 曲线(Receiver Operating Characteristic Curve)图如图 4 所示. 由图可知,PSO_LSBMT1 和 PSO_LSBMD1 两种隐写算法与 LSBM 算法相比,具有更低的 AUC(Area under ROC Curve)值. 说明采用两种安全性测度均可指导设计安全性更高的隐写算法. 同时,PSO_LSBMT1 获得了比 PSO_LSBMD1 更小的 AUC 值为 0.6634,可见,PSO_LSBMT1 具有更强的抵抗隐写分析能力. 因此,采用本文提出的 Vague 集相似度量的隐写安全性测度与确定模式下的安全性测度相比,指导设计高安全性的隐写算法能力更强. 分析其原因,是由于采用 Vague 集相似度量的隐写安全性测度,充分考虑了图像隐写引起的不确定性,用模糊 Vague 集相似度量来计算载体数据与载密数据相关统计量之间的差距. Vague 集相似度量实际上是用分段的方式进行度量两者的差距,所以比用确定方式度量两者的差距更精确.

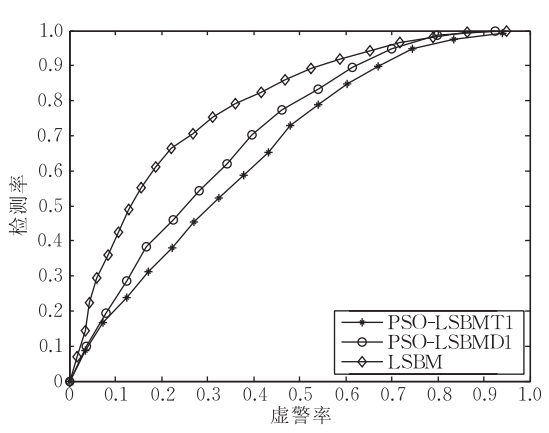


图 4 Vague 集相似度量的一阶安全性与文献[3]的安全性测度指导设计的隐写算法 ROC 曲线图

采用文献[5]中确定模式下的安全性测度 $D(\boldsymbol{M}_c, \boldsymbol{M}_s)$和 Vague 集相似度量的二阶安全性测度

$T(\boldsymbol{MC},\boldsymbol{MS})$按以上方式指导设计改进的 LSBM 算法，分别记为 PSO_LSBMD2 和 PSO_LSBMT2. 图 5 为这 3 种算法按以上相同的实验设计得到的隐写分析 ROC 曲线图，可得到与图 4 相似的实验结论.

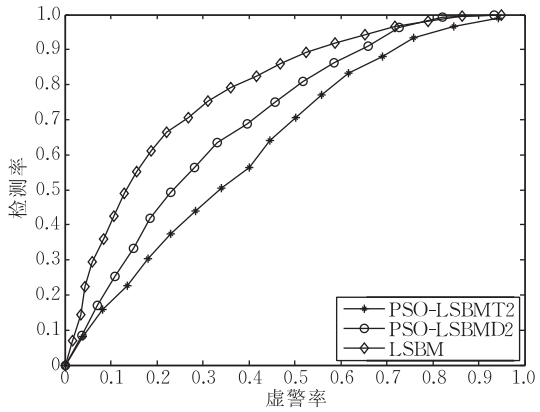


图 5　Vague 集相似度量的二阶安全性与文献[5]的安全性测度指导设计的隐写算法 ROC 曲线图

6　结束语

Vague 集理论作为一种新的不确定性智能信息处理方法，具有比 Fuzzy 集更强的数据表达能力. Vague 集之间的相似度量表示了两集合的匹配程度. 本文根据图像数据的复杂性与隐写引起图像统计特征变化的不确定性，在假设图像为独立同分布和马尔可夫链模型下，分别定义了基于 Vague 集相似度量的隐写系统一阶和二阶安全性，并证明了该安全性测度的有界性、对称性和一致性. 在假设图像为独立同分布条件下，证明了 Vague 集相似度量的隐写系统一阶和二阶安全性测度的等价性. 通过仿真实验，验证了基于 Vague 集相似度量的隐写系统安全性测度对不同的隐写算法均可得到随着嵌入率增大，隐写系统安全性下降的结论，该测度同时可区分不同隐写算法在同一嵌入率下的安全性. 通过度量同一隐写算法在不同嵌入率下的安全性可得，二阶安全性测度比一阶安全性测度能更好地反映隐写引起的载体数据统计特征变化. 此外，与同条件确定模式下的安全性测度相比，在嵌入率低于 0.5 bpp 时，Vague 集相似度量的隐写系统安全性测度反应灵敏度远高于确定模式下的反应灵敏度，说明该测度对小嵌入量的隐写及隐写分析算法具有更好的指导作用. 隐写算法设计实验进一步表明，Vague 集相似度量的隐写安全性测度指导的隐写算法抗分析能力更强，说明了该测度对隐写算法设计的指导能力更强.

一阶马尔可夫链模型中任一像素只受前一像素的影响，考虑了图像像素间的相关性信息，但由于自然图像每一像素点至少与其邻域内像素有着较强的相关性. 因此，使用一阶马尔可夫链模型仍将丢失一些像素相关信息，当分析者利用更复杂的模型提取统计特征或分析计算能力特别强时，采用 Vague 集相似度量的图像隐写系统一阶和二阶安全性测度将会具有一定的局限性. n 阶马尔可夫链模型可较全面地描述数字图像相关性信息，所包含图像相邻像素间的相关性信息量更全面. 如何定义 n-MC 模型下的 Vague 集相似度量的隐写系统安全性测度，以及利用 Vague 集相似度量的隐写系统安全性测度指导设计高安全性的隐写算法及高性能的隐写分析算法，均可作为下一步研究方向.

参　考　文　献

[1] Wang Shuo-Zhong, Zhang Xin-Peng, Zhang Wei-Ming. Recent advances in image-based steganalysis research. Chinese Journal of Computers, 2009, 32(7): 1247-1263(in Chinese)
(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. 计算机学报，2009，32(7)：1247-1263)

[2] Wang Yu-Min, Liu Jian-Wei. The Security of Communication Network——Theory and Technology. Xi'an: Xidian University Press, 1999(in Chinese)
(王育民，刘建伟. 通信网的安全——理论与技术. 西安：西安电子科技大学出版社，1999)

[3] Cachin C. An information-theoretic model for steganography. Information and Computation, 2004, 192(1): 41-56

[4] Wang Y, Moulin P. Steganalysis of block-structured stego text//Proceedings of the SPIE: Security, Steganography and Watermarking of Multimedia Contents VI. Bellingham, WA, USA, 2004, 5306: 477-488

[5] Sullivan K, Madhow U, Chandrasekaran S et al. Steganalysis for Markov cover data with applications to images. IEEE Transactions on Information Forensics and Security, 2006, 1(2): 275-287

[6] Chen Dan, Wang Yu-Min. The steganography security definition based on image model. Journal of Harbin Institute of Technology, 2006, 38(Supplement): 901-904(in Chinese)
(陈丹，王育民. 基于图像模型的掩密安全性定义. 哈尔滨工业大学学报，2006，38(增刊)：901-904)

[7] Ambalavanan A, Chandramouli R. A Bayesian image steganalysis approach to estimate the embedded secret message length//Proceedings of the 7th Workshop on Multimedia and Security. New York, USA, 2005: 33-38

[8] Pevný T, Fridrich J. Benchmarking for steganography//Proceedings of the 10th Information Hiding International Workshop. Santa Barbara, CA, USA, 2008: 251-267

[9] Zhang Zhan, Qu Fang, Liu Guang-Jie et al. A novel security evaluation method for digital image steganography based on high-order Markov Chain model. Information and Control, 2010, 39(4): 455-461(in Chinese)
(张湛, 瞿芳, 刘光杰等. 基于高阶 Markov 链模型的数字图像隐写安全性评估方法. 信息与控制, 2010, 39(4): 455-461)

[10] Zhang Zhan, Liu Guang-Jie, Wang Jun-Wen et al. Steganalysis of spread spectrum image steganography based on high-order Markov chain mode. Acta Electronica Sinica, 2010, 38(11): 2578-2584(in Chinese)
(张湛, 刘光杰, 王俊文等. 基于图像高阶 Markov 链模型的扩频隐写分析. 电子学报, 2010, 38(11): 2578-2584)

[11] Cao Wen-Ming, Wang Rui. The Study of Sensor Networks Coverage by Fuzzy Information Processing. Beijing: Electronic Industry Press, 2010(in Chinese)
(曹文明, 王瑞. 传感器网络覆盖定位模糊信息处理方法. 北京: 电子工业出版社, 2010)

[12] Gau W L, Buehrer D J. Vague sets. IEEE Transactions on Systems, Man and Cybernetics Society, 1993, 23(2): 610-614

[13] Zhang Cheng-Yi, Dang Ping-An, Li Dong-Ya. Note on "fuzzy entropy of Vague sets and its construction method". Computer Applications and Software, 2004, 21(5): 27-28

[14] Li Fan, Xu Zhang-Yan. Measures of similarity between Vague sets. Journal of Software, 2001, 12(6): 922-927(in Chinese)
(李凡, 徐章艳. Vague 集之间的相似度量. 软件学报, 2001, 12(6): 922-927)

[15] Quan Shuang-Yan. Similarity measures between Vague sets based on information. Computer Engineering and Applications, 2007, 43(25): 87-89(in Chinese)
(权双燕. 信息意义下的 Vague 集的相似度量. 计算机工程与应用, 2007, 43(25): 87-89)

[16] Li Yan-Hong, Olson David L, Qin Zheng. Similarity measures between intuitionistic fuzzy (Vague) sets: A comparative analysis. Pattern Recognition Letters, 2007, 28(2): 278-285

[17] Petitcolas F A P, Anderson R J, Kuhn M G et al. Information hiding — A survey. Proceedings of the IEEE, 1999, 87(7): 1062-1078

[18] Sharp T. An implementation of key-based digital signal steganography//Proceedings of the 4th Information Hiding Workshop. Pittsburgh, PA, USA, 2001, 2137: 13-26

[19] Mielikainen J. LSB matching revisited. Signal Processing Letters, 2006, 13(5): 285-287

[20] Westfeld A. F5—A steganographic algorithm: High capacity despite better steganalysis//Proceedings of the 4th Information Hiding Workshop. Pittsburgh, PA, USA, 2001, 2137: 289-302

[21] Shi Y, Eberhart R. Empirical study of particle swarm optimization//Proceedings of the International Conference on Evolutionary Computation. Washington, USA, 1999: 1945-1950

[22] Ker A D. Steganalysis of LSB matching in grayscale images. IEEE Signal Processing Letters, 2005, 12(6): 441-444

OUYANG Chun-Juan, born in 1974, Ph. D. candidate, associate professor. Her research interests focus on image steganography and steganalysis.

LI Bin, born in 1982, Ph. D., lecturer. His research interests focus on image steganography and pattern recognition.

LI Xia, born in 1968, Ph. D., professor, Ph. D. supervisor. Her research interests include intelligence optimization and intelligence computing and application.

WANG Na, born in 1977, Ph. D., professor. Her research interests include intelligence optimization and pattern recognition.

Background

Steganography is the art and science of secret communication in such a way that the presence of a message cannot be detected, while steganalysis tries to reveal the presence of the hidden messages. The security is a basic issue for the steganographic system. Therefore the study of the security measure becomes one of the hotspots in the research field of information security. So far, most of existing evaluation systems for steganographic security are founded upon the assumption that image statistical distribution is deterministic. However, it is impractical to obtain infinite data samples, which means an exact probability distribution of the cover and stego images cannot be derived. As a result, we propose to consider the steganography as a fuzzy process. The vague set is more effective in revealing the indeterminacy when compared to the

fuzzy set. Thus an evaluation system for steganographic security is introduced in this paper to measure the similarity between cover images and stego images. Firstly, the first-order and the second-order security measure were defined in terms of the vague set similarity measure between the distributions of the cover images and the stego images. Secondly, the newly defined security measures are proved to have the properties of boundedness, symmetry and unity. Particularly, the property of boundedness limits the value of the security measure in the range of [0,1], which is an absent characteristic in other evaluation systems with a deterministic image data statistical distribution model. Besides, when a steganographic system satisfies the property of unity, it is perfectly secure. Lastly, under the assumption that the pixels statistical distribution is independent identical distributed, the first-order and the second-order security measure are proved to be equivalent. Simulation results show the effectiveness of the proposed security evaluation system for evaluating different steganographic algorithms. When measuring the security under the same embedding rate, the second-order security measure can always reflect more obvious statistical changes than the first-order security measure. The new security measure is more sensitive than other measures with the deterministic image statistical model when the embedding rate is low. Experimental results also indicate that the proposed security measure may shed a light on designing steganography with high security level.

This work is supported by the National Natural Science Foundation of China (Grant Nos. 61171124, 61103174, 60902069, 61005049), Science Technology Planning Project of Guangdong (Grant No. 2011B010200045). Foundation for Distinguished Young Talents in Higher Education of Guangdong (Grant No. LYM10116), and the Upgrade Projects of Shenzhen Key Laboratory (Grant No. CXB201105060068A).