

基于 CRITIC 权与灰色关联的隐写分析算法综合评估

王 磊, 高茂庭

(上海海事大学 信息工程学院, 上海 201306)

摘 要: 在基于灰色关联分析的隐写分析算法评估过程中, 针对权值分配不合理的问题, 采用标准间重要性相关性方法确定各指标的初始权重, 结合决策情况对待主动修改的指标权重进行增减, 并将这些增减的权重平均分配给被动变化的指标, 从而实现对权重的调整, 得到各指标的最终权值。运用灰色关联分析法计算各个待评价对象与最优序列之间的加权灰色关联度, 并依据加权关联度的大小完成对隐写分析算法的评估。仿真结果表明, 与其他隐写分析算法相比, 该方法能更加有效地评估各类隐写分析算法的优劣。

关键词: 隐写分析; 算法性能评估; 标准间重要性相关性权; 灰色关联分析; 权值分配

中文引用格式: 王 磊, 高茂庭. 基于 CRITIC 权与灰色关联的隐写分析算法综合评估[J]. 计算机工程, 2017, 43(4): 154-159.

英文引用格式: Wang Lei, Gao Maoting. Comprehensive Evaluation of Steganography Analysis Algorithm Based on CRITIC Weight and Grey Relation[J]. Computer Engineering, 2017, 43(4): 154-159.

Comprehensive Evaluation of Steganography Analysis Algorithm Based on CRITIC Weight and Grey Relation

WANG Lei, GAO Maoting

(College of Information Engineering, Shanghai Maritime University, Shanghai 201306, China)

[Abstract] Aiming at the problem of unreasonable distribution of weights in the evaluation process of steganography analysis algorithm based on grey relational analysis, Criteria Importance Through Intercriteria Correlation (CRITIC) method is used to determine the initial weights of each index. The weights of the active modifying indexes are increased or decreased, and then equally distributed to the passive indexes according to the actual decision. The final weights of each index are obtained by adjusting. Grey relational analysis is used to calculate the weighted grey correlation degree between each object to be evaluated and the optimal sequence. The steganography analysis algorithm is evaluated by the correlation value. Simulation results show that the proposed method can evaluate various kinds of steganography analysis algorithm more effectively than other steganography analysis algorithms.

[Key words] steganography analysis; algorithm performance evaluation; Criteria Importance Through Intercriteria Correlation (CRITIC) weight; grey relational analysis; weight distribution

DOI: 10.3969/j.issn.1000-3428.2017.04.026

0 概述

隐写术是将秘密信息嵌入伪装载体的一门科学, 其主要目标是使隐藏数据不可检测。当隐写术被不法分子利用时, 会给社会带来严重安全威胁, 故针对隐写术进行隐写分析尤为重要。隐写分析是从海量位流中分析出隐藏信息, 对貌似正常的传输媒体信号进行检测, 判断其中是否存在秘密信息, 并找到隐蔽通信信息, 阻断隐蔽通信的信道, 找到秘密信

息的接收双方^[1]。

目前, 在隐写分析领域已经出现了大量隐写分析算法, 大多是针对某种应用场景, 使得隐写分析的某些性能指标得到提升, 而这些提升往往是以牺牲其他性能指标为代价。

因此, 衡量隐写分析算法的优劣不能单单依据某个指标的提升, 而需要综合评估。通过综合评估, 不仅可以加深对现有隐写分析算法的理解, 而且对算法的设计具有一定的指导意义。

基金项目: 上海市科委科技创新项目 (12595810200)。

作者简介: 王 磊 (1990—), 男, 硕士研究生, 主研方向为数据安全、数据挖掘、云计算; 高茂庭, 教授、博士。

收稿日期: 2016-04-13 **修回日期:** 2016-05-16 **E-mail:** wang_lei_email@163.com

目前,对隐写分析算法性能的评估主要有层次分析法^[2]、模糊综合评价法^[3]、模糊聚类分析法^[4]、理想逼近法^[5]和灰色关联分析法^[6]。层次分析法和模糊综合评价法主观性较强;模糊聚类分析不适合样本较少的情况;而理想逼近法仅仅是与最优和最差方案对比,而没有考虑非最优或最差方案对逼近程度的影响。灰色关联分析法对小样本不敏感以及计算简单的特性使其很适合对隐写分析算法性能进行综合评估^[7]。故本文采用灰色关联分析法对隐写分析算法进行评估。

在使用灰色关联分析法进行评价时,权重的确定对评价结果有很大的影响。常用权重确定法有主观法和客观法2类。主观法如层次分析法和德菲尔法^[8],主观性太强,并且过多的专家参与增加了决策者的工作量;客观法如标准离差法^[9]、变异系数法^[10]和熵权法^[11],仅仅考虑了指标内部间差异性对权重的影响,而没有考虑到指标间的影响。另外,虽然客观赋权法减少了决策者的工作量,克服了主观赋权法主观性太强的缺点,但是单纯的客观赋权法的权值仅依赖于样本,很容易受到样本本身的影响。当决策者对某指标十分敏感,但在样本中该指标上的数据差异程度过低时,单纯的客观赋权法会给该指标过低的权重,这与决策者的要求相悖。

针对以上问题,本文将主观法与客观法结合,并同时考虑指标内和指标间对权重的作用。首先运用标准间重要性相关性(Criteria Importance Through Intercriteria Correlation, CRITIC)法^[12]确定各指标的初步权重,然后结合实际将待主动修改的权重进行增减,使增减所产生的亏损由被动变化的指标平均承担,以此对初步权进行修正,确定最终的权重。

1 隐写分析算法性能指标的选取

结合实际,本文采用文献[4]提出的指标体系。文献中共给出7种指标:

1) 检出率 $P_1: P_1 = M_1/M$ 。 M_1 是隐写分析算法把实际上包含秘密信息的图像正确识别出来的个数; M 表示包含秘密信息的图像数目。

2) 漏报率 $P_2: P_2 = M_2/M$ 。 M_2 是隐写分析算法把实际上包含秘密信息的图像误认为不包含秘密信息的个数。

3) 虚报率 $P_3: P_3 = N_1/N$ 。 N_1 是隐写分析算法把实际上不包含秘密信息的图像误认为包含秘密信息的个数。 N 表示不包含秘密信息的图像数目。

4) 否定率 $P_4: P_4 = N_2/N$ 。 N_2 是隐写分析算法把实际上不包含秘密信息的图像正确划分为非秘密信息图像的个数。

5) 隐写分析误差率 $P_e: P_e = P(e|\eta_0) + P(e|\eta_1)$ 。 $P(e|\eta_0)$ 表示在待检测的图像不含秘密信息的条件下,隐写分析算法误认为该图片含有秘密信息的概率;

$P(e|\eta_1)$ 表示在待检测的图像含秘密信息的条件下,隐写分析算法误认为该图片不含有秘密信息的概率。

6) 检测时间 P_{com} : 表示隐写分析算法对单位大小的图像花费的检测时间。检测时间可由下列式子算出。

$$P_{\text{com}} = \frac{1}{L} \sum_{i=1}^L \frac{T_i}{S_i}$$

其中, S_i 是在 L 张图片中第 i 张图片的大小; T_i 代表隐写分析算法对第 i 张图片进行检测所需的时间。

7) 鲁棒性 P_5 : 隐写分析算法对有噪声的隐蔽图像的检出率。

2 相关理论

2.1 灰色关联分析

灰色关联分析是根据已有的评价标准或参考序列,通过计算样本序列与参考序列的关联度大小,判断该样本序列与参考序列的接近程度来评定该样本序列的优劣。灰色关联评价系统模型建模步骤如图1所示^[6]。

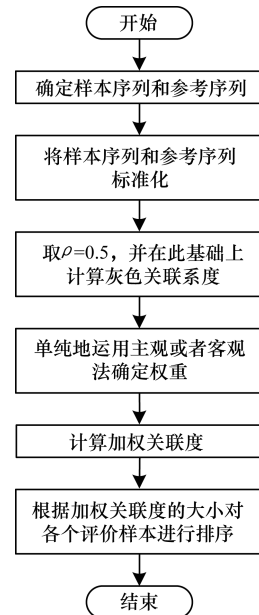


图1 灰色关联分析模型建模步骤

建模具体流程描述如下:

1) 确定样本序列和参考序列。

样本序列的选取:

设共有 m 个样本,包含 n 个评价指标,则样本序列所组成的矩阵 X 为:

$$X = (x_{ij})_{mn}, i = 1, 2, \dots, m; j = 1, 2, \dots, n \quad (1)$$

其中, x_{ij} 表示第 i 个样本在第 j 个指标上的取值。

参考序列的选取:

通常根据已有的样本序列,从每一个评价指标中选取最优值,从而形成参考序列。对于正向指标选取最大值作为最优值,逆向指标选取最小值作为最优值。设最优序列为 $x_0 = (x_{01}, x_{02}, \dots, x_{0n})$, 则有:

$$x_0 = (\text{opt}x_{i1}, \text{opt}x_{i2}, \dots, \text{opt}x_{in}) \quad (2)$$

其中,符号 opt 表示最优, $i=1,2,\dots,m$ 。由于参考序列来自样本序列的最优值,为了便于描述,可以把参考序列当作第 0 个样本序列。

2) 样本序列和参考序列的标准化。

为了避免待评价的样本中各个指标量纲不一致而对评价结果造成的影响,在进行灰色关联分析之前,要对样本数据和参考序列进行标准化处理。比较常用的处理方法如下:

$$d_{ij} = \frac{x_{ij} - \min_k x_{kj}}{\max_k x_{kj} - \min_k x_{kj}} \quad (3)$$

或者

$$d_{ij} = \frac{\max_k x_{kj} - x_{ij}}{\max_k x_{kj} - \min_k x_{kj}} \quad (4)$$

其中, d_{ij} 表示标准化后第 i 个样本的第 j 个指标值, $i=0,1,2,\dots,m; j=1,2,\dots,n; k=1,2,\dots,m$ 。式(3)适用于值越大越好的指标;式(4)适用于值越小越好的指标。

通过式(3)、式(4)可得到标准化后的样本序列矩阵 D :

$$D = (d_{ij})_{mn}, i=1,2,\dots,m; j=1,2,\dots,n \quad (5)$$

同样,根据式(3)、式(4)可以得到标准化后参考序列 d_0 (即最优序列) 为恒定值:

$$d_0 = (d_{01}, d_{02}, \dots, d_{0n}) \equiv (1, 1, \dots, 1) \quad (6)$$

3) 求灰色关联度。

定义 Δ_{ij} 为标准化后第 i 个样本在 j 个指标上的值与第 j 个指标上最优值的偏差,则 Δ_{ij} 为:

$$\Delta_{ij} = |d_{ij} - d_{0j}|, i=1,2,\dots,m; j=1,2,\dots,n \quad (7)$$

标准化后,定义第 i 个样本在 j 个指标上的值与第 j 个最佳指标的灰色关联度 ε_{ij} 为:

$$\varepsilon_{ij} = \frac{\min_a \min_b \Delta_{ab} + \rho \max_a \max_b \Delta_{ab}}{\Delta_{ij} + \rho \max_a \max_b \Delta_{ab}} \quad (8)$$

其中, $a=1,2,\dots,m; b=1,2,\dots,n; \rho$ 为分辨系数,其取值在 0~1 之间,一般取 $\rho=0.5$,然而,分辨系数 ρ 的这种取值过于随意,缺乏客观依据。

4) 确定各个指标的权重。

可以根据应用场景,选择合适的方法。最后得出 n 个评价指标的权重向量: $W = (\omega_1, \omega_2, \dots, \omega_n)$ 。

权重的确定是灰色关联分析评价模型的关键。

5) 求加权关联度。

设 γ_i 表示第 i 个样本与最优序列的加权关联度,则 γ_i 可由式(9)求得:

$$\gamma_i = \sum_{j=1}^n \varepsilon_{ij} \omega_j, i=1,2,\dots,m \quad (9)$$

6) 根据加权关联度的大小对各个评价样本进行排序。

在式(9)中, γ_i 的值越大,表明第 i 个样本与最优序列的关联度越大;反之, γ_i 的值越小,表明第 i 个样本与最优序列的关联度越小。因此,可以根据

计算所得的关联度 γ_i 的大小,对评价对象进行排序,从而对样本序列的优劣性进行评估。

2.2 CRITIC 法

CRITIC 法是一种客观赋权法。CRITIC 法运用评价指标内部的对比强度与指标之间的冲突性的乘积来确定各指标的权值^[12]。

根据 CRITIC 法,定义 c 为评价系统中指标的信息量。那么第 j 个指标的信息量为:

$$c_j = \sigma_j \sum_{k=1}^n (1 - r_{jk}), j=1,2,\dots,n \quad (10)$$

其中, σ_j 是第 j 个指标的标准差,可以作为对比强度的度量; r_{jk} 是评价指标 j 和评价指标 k 之间的相关系数;而 $\sum_{k=1}^n (1 - r_{jk})$ 代表指标间的冲突性。

当 c_j 越大时,表明第 j 个指标所包含的信息比较多,对应的权重就应该越高。所以,第 j 个指标的客观权重 ω_j 应为:

$$\omega_j = c_j / \sum_{k=1}^n c_k, j=1,2,\dots,n \quad (11)$$

3 基于 CRITIC 权的灰色关联综合评估模型

采用灰色关联分析模型进行评价时,有 2 个问题需要考虑:

1) 在计算灰色关联系数 ε_{ij} 时,需要给分辨系数 ρ 进行合理赋值。

2) 灰色关联分析模型没有明确给出指标权重的确定方法,一般都会根据应用实际来确定指标权重。

为了解决以上 2 种问题,下文首先分析分辨系数 ρ 的实际意义,然后给出定量计算 ρ 的方法;针对灰色关联模型的权重,引进 CRITIC 法来确定初步权重,然后根据实际需求,对初步权重进行修改。最终提出基于 CRITIC 权的灰色关联度综合评估模型。模型建立步骤如图 2 所示。

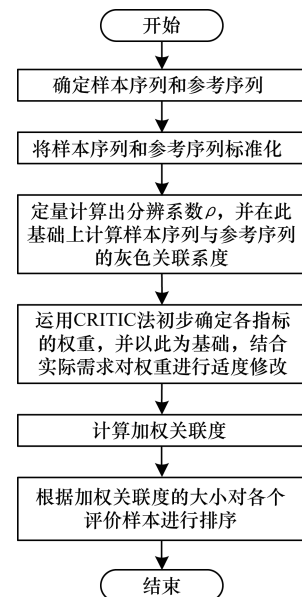


图 2 基于 CRITIC 权的灰色关联模型评估步骤

下面论述灰色关联分析模型中分辨系数的实际意义分析与计算方法,以及各指标权重的确定方法。

3.1 灰色关联模型中分辨系数的定量确定

在计算灰色关联系数 ε_{ij} 时,需要给分辨系数 ρ 进行赋值,通常 ρ 取 0.5,但缺乏客观依据。故本文对分辨系数进行详细分析,并给出客观取值,再通过式(8)计算出更科学的灰色关联度 ε'_{ij} 。

分辨系数 ρ 是 $\max_a \max_b \Delta_{ab}$ 的系数。 ρ 取值的原则为^[13]:

1)要体现出 ε_{ij} 的整体性。也就是说要做到 ε_{ij} 与 d_i, d_0 有关的同时,也要跟因素有一定的关联。

2)提高评价系统的鲁棒性。即当参考序列的值有噪声时, ρ 可以减少它对整体评价结果的影响。

一种可行的 ρ 取值规则如下:

首选,定义 $\bar{\Delta}$ 为全部偏差的均值,其计算方式如下:

$$\bar{\Delta} = \frac{1}{mn} \sum_{i=1}^m \sum_{j=1}^n \Delta_{ij} \quad (12)$$

然后,定义 $\bar{X}$ 为全部偏差均值与 $\max_a \max_b \Delta_{ab}$ 的商,即 $\bar{X} = \bar{\Delta} / \max_a \max_b \Delta_{ab}$ 。

则 ρ 应该在 $\bar{X}$ 与 $2\bar{X}$ 之间取值,即 $\bar{X} < \rho < 2\bar{X}$,且满足:

当 $\bar{X} < 1/3$ 时, $\bar{X} < \rho < 1.5\bar{X}$

当 $\bar{X} \geq 1/3$ 时, $1.5\bar{X} < \rho < 2\bar{X}$ (13)

通过以上方法,就可以更合理地给出分辨系数 ρ 的值,然后利用式(8)计算出新的灰色关联度 ε'_{ij} 。

3.2 指标权重的确定

针对单纯使用主观赋权法或客观赋权法的不足,本文引入 CRITIC 法,不仅考虑到指标内部变异程度对指标权重的影响,而且考虑了指标间冲突性对权重的作用。首先运用 CRITIC 初步确定各指标的初步权重,然后结合应用实际,对初步结果进行修正,得到最终的权重。

1)通过改进的 CRITIC 法确定初步权重

按照 CRITIC 方法,对比强度由标准差来确定,但如果各个指标下数据的均值不同时,对比强度使用标准差法就变得不太合理。

因此,本文对 CRITIC 方法进行了改进,采用标准差与平均数的比值即变异系数来衡量对比强度。消除各指标间均值不相等所造成的误差,则标准化后的矩阵 D 中第 j 列数据的对比强度由单纯的标准差 σ 修改为变异系数 v_j :

$$v_j = \sigma_j / \mu_j, j = 1, 2, \dots, n \quad (14)$$

其中, σ_j 为第 j 列数据的标准差; μ_j 为第 j 列数据的均值。

根据式(9)、式(13)可得,改进后,第 j 个评价指标所包含的信息量 c'_j 可表示为:

$$c'_j = v_j \sum_{k=1}^n (1 - r_{jk}), j = 1, 2, \dots, n \quad (15)$$

再根据式(11)、式(15)可得,改进后的 CRITIC 权 ω'_j 可表示为:

$$\omega'_j = c'_j / \sum_{k=1}^n c'_k, j = 1, 2, \dots, n \quad (16)$$

2)结合实际应用需求确定最终权重

在实际应用中,决策者可以根据应用实际将上一步得到的客观权重的某些指标(定义为主动指标)的权重进行适度修改,为了保证修改后总权重之和为 1 的特性,需要对未被关注的指标(定义为被动指标)权重统一进行扩大或缩小。

为此,将待主动修改的指标进行修改,把增加或减少的部分平均分配到被动修改的指标上去。

设 n 个指标中,待主动修改的指标个数为 s , 它们的下标分别为 $pos_1, pos_2, \dots, pos_s$, 被动变化的指标个数为 t , 它们的下标分别为 $neg_1, neg_2, \dots, neg_t$, $s + t = n$, 且有 $\{\omega'_1, \omega'_2, \dots, \omega'_n\} = \{\omega'_{pos}\} \cup \{\omega'_{neg}\}$ 。

则由式(16)所得的权重分 2 种:主动修改权重和被动变化权重。

对于主动修改的权重,其修改后的权重 ω''_{pos_i} 为:

$$\omega''_{pos_i} = \omega'_{pos_i} + \delta_i, i = 1, 2, \dots, s \quad (17)$$

其中, ω'_{pos_i} 为适度修改前下标为 pos_i 的指标权重; δ_i 为决策者适度调整幅度,当它为正值时,表明要增大下标为 pos_i 的指标权重;当它为负值时,表明要减小该下标所对应的权重。

对于被动变化的指标,其修改后的权重 ω''_{neg_k} 为:

$$\omega''_{neg_k} = \frac{1 - \sum_{i=1}^s \omega''_{pos_i}}{\sum_{j=1}^t \omega'_{neg_j}} \omega'_{neg_k}, k = 1, 2, \dots, t \quad (18)$$

其中, ω'_{neg_k} 为适度修改前下标为 neg_k 的指标权重。

最后,由式(17)和式(18)可以得到最终的权重向量 $\omega'' = (\omega''_1, \omega''_2, \dots, \omega''_n)$, 易得最终的权重向量集合 $\{\omega''_1, \omega''_2, \dots, \omega''_n\} = \{\omega''_{pos}\} \cup \{\omega''_{neg}\}$ 。

4 应用实例

4.1 实验数据的获取

为了便于对评价结果进行比较,本文关于各种隐写分析算法的性能指标均来自文献[4]。

共选取了 7 种类型的指标:检出率(P_1),漏报率(P_2),虚报率(P_3),否定率(P_4),检测误差(P_e),检测复杂度(P_{com})和鲁棒性(P_5)。并分别选取 5 种常见的 LSB 隐写分析算法:二次密写估计法(F_1)^[14],经典 RS 攻击法(F_2)^[15],基于差分直方图转移系数度量法(F_3)^[16],梯度能量翻转法(Gradient Energy-Flipping Rate, GEFR)(F_4)^[17]和点对分析法(F_5)^[18]作为待评价的样本。

接着,从 CBRI 图像库中,随机选取 100 幅大小为 1536×1024 像素的 12 位灰度图像作为样本,供

各种隐写分析算法攻击。然后从上述样本中随机选取 50 幅图像进行二值序列嵌入,从而产生隐蔽图像,且嵌入率为 0.5。最后,用 5% 的 JPEG 压缩来获取隐写分析算法的鲁棒性数据。从而得出各隐写分析算法的性能数据如表 1 所示。

表 1 5 种隐写分析算法性能的指标值

算法	$P_1/\%$	$P_2/\%$	$P_3/\%$	$P_4/\%$	$P_e/\%$	P_{com}/ms	$P_5/\%$
F_1	96.8	3.2	3.9	96.1	7.1	17.92	85.6
F_2	97.1	2.9	2.5	97.5	5.4	17.67	89.2
F_3	98.5	1.5	1.4	98.6	2.9	15.53	91.3
F_4	99.3	0.7	2.7	97.3	3.4	21.15	91.5
F_5	99.5	0.3	2.4	97.6	2.9	10.52	90.5

针对本文所提出的综合评估模型,由模型本身的特点可以看出,在进行灰色关联分析之前要去除一些相关性比较大的指标。在上述 7 种类型的指标中,检出率与漏报率之和为 1,具有很高的相关性,所以只需选择其中一个即可,本文去掉漏报率而保留检出率;同理,虚报率和否定率两者中也只需选择一个,本文去掉虚报率而保留否定率。因此,最终的指标为 5 种,分别为检出率、否定率、检测误差、检测复杂度和鲁棒性。故剔除表 1 中 P_2 列和 P_3 列,把剩余的数据作为本文模型的样本数据。

4.2 综合评估

根据 4.1 节可得样本序列所形成的矩阵 X 为:

$$X = \begin{bmatrix} 0.968 & 0.961 & 0.071 & 1.792 & 0.856 \\ 0.971 & 0.975 & 0.054 & 1.767 & 0.892 \\ 0.985 & 0.986 & 0.029 & 1.553 & 0.913 \\ 0.993 & 0.973 & 0.034 & 2.115 & 0.915 \\ 0.995 & 0.976 & 0.029 & 1.052 & 0.905 \end{bmatrix}$$

利用式(3)和式(4)将原始矩阵 X 做标准化处理得到规范化矩阵 D :

$$D = \begin{bmatrix} 0 & 0 & 0 & 0.304 & 0 \\ 0.111 & 0.56 & 0.405 & 0.327 & 0.61 \\ 0.63 & 1 & 1 & 0.529 & 0.966 \\ 0.926 & 0.48 & 0.881 & 0 & 1 \\ 1 & 0.6 & 1 & 1 & 0.831 \end{bmatrix}$$

由式(6)易知参考序列 $d_0 = (1, 1, \dots, 1)$ 。

利用式(12)、式(13)计算出分辨系数 $\rho = 0.759$ 。

结合定量计算出的分辨系数 ρ ,根据式(8)得到灰色关联度矩阵 ε' :

$$\varepsilon' = \begin{bmatrix} 0.431 & 0.431 & 0.431 & 0.522 & 0.431 \\ 0.461 & 0.633 & 0.561 & 0.53 & 0.661 \\ 0.672 & 1 & 1 & 0.617 & 0.957 \\ 0.911 & 0.593 & 0.864 & 0.431 & 1 \\ 1 & 0.655 & 1 & 1 & 0.818 \end{bmatrix}$$

由 CRITIC 法计算得到各个指标的初步权重 ω' :

$$\omega' = [0.159 \quad 0.215 \quad 0.106 \quad 0.361 \quad 0.16]$$

结合实际,假设决策者对检出率要求更为敏感,

则需要对上述客观方法得到的权值进行适度的修改。例如修改第一列检出率的值为 $0.159 + 0.2$,则根据式(17)和式(18)可得最终的权重向量 ω'' :

$$\omega'' = [0.359 \quad 0.164 \quad 0.081 \quad 0.275 \quad 0.122]$$

最后,参照式(9)可得 5 种隐写分析算法的加权关联度向量 γ' :

$$\gamma' = [0.456 \quad 0.541 \quad 0.773 \quad 0.735 \quad 0.922]$$

由以上计算结果可知 $\gamma'_5 > \gamma'_3 > \gamma'_4 > \gamma'_2 > \gamma'_1$ 。因此,决策者结合客观数据与实际的需要,最终得出这 5 种算法从优到差的顺序为 F_5, F_3, F_4, F_2, F_1 。

4.3 结果分析

从评价结果可以看出, F_5 为最优的隐写分析算法, F_1 为最差的算法,这一点与文献[3,4,6]一致。但 F_2, F_3, F_4 的优劣顺序与文献[3,4,6]不同。

对于算法 F_3 和 F_4 ,从本文的实验结果来看,它们的评估得分相差不大,这与文献[3,4,6]的计算结果一致,故针对 F_3 与 F_4 的排序不做深入分析。本文模型所得评估结果与文献[3,4,6]最大的差别在于算法 F_2 与算法 F_3, F_4 的优劣顺序上。文献[3,4,6]评估结果表明 F_2 要优于 F_3, F_4 ,而本文评估结果认为 F_2 要次于 F_3, F_4 。

1) F_2 与 F_3 的对比

从表 1 可以看出, F_3 (基于差分直方图分析法)的正项指标 P_1, P_4 和 P_5 的值均大于 F_2 (RS 分析法),且 F_3 的逆向指标 P_2, P_3, P_e 和 P_{com} 的值均小于 F_2 ,也就是说, F_3 在各项指标上边均优于 F_2 ;另外 F_3 是以差分直方图间的转移系数作为 LSB 平面与图像其余比特平面之间的弱相关性度量^[16]。该算法对于空域 LSB 秘密信息有很高的识别率,并且可以计算出嵌入数据量的大小。此外该算法实现简单,计算量小,且实验结果表明,针对原始无损存储图像可以获得优于 F_2 的性能,且计算速度显著高于 F_2 。综上, F_3 算法应优于 F_2 。

2) F_2 与 F_4 的对比

F_2 的本质是根据 LSB 平面嵌入秘密信息量的增长,观察置反操作后正则群和奇异群的改变量,当秘密信息长度递增时,这种变量会增加^[15]。而 F_4 (GERF 分析法)在置反的基础上,结合梯度能量变化来进行隐写分析^[16]。因此, F_4 在检出率上会优于 F_2 。另外,为了提高算法性能, F_4 分别计算了图像水平和垂直方向的梯度能量变化,得到 2 个信息长度的估计值,然后利用上述 2 个估计长度的算术平均值,来减少原图像的 LSB 的非随机不均匀特性的影响及嵌入式消息,故鲁棒性方面, F_4 会有一定的提高,当然这也加大了 F_4 的算法复杂度。从表 1 可以看出, F_4 在检出率 P_1 和鲁棒性 P_5 方面

确实优于 F_2 ,而在算法复杂度 P_{com} 方面差于 F_2 。表1的实验数据与 F_2 与 F_4 算法本身所得到的结论吻合。另外,本文所提出的赋权法给予检出率 P_1 更多的权重,检出率明显高的 F_4 的综合评分显然要高于 F_2 。

结合1)、2)对比可知,从算法本身和算法的实验结果数据来看, F_2 要次于 F_3, F_4 ,应当排在 F_3, F_4 之后,本文算法的评价结果与之一致。

从评价方法本身来讲,文献[3]运用模糊综合评价法评价需要有大量领域专家参与。一方面受到领域专家认知水平的限制,另一方面,专家的主观想法也会对评价结果有很大的影响。文献[4]基于模糊聚类分析的方法适合样本比较多的情况,样本过少,会严重影响评价结果的准确性。且模糊聚类分析对孤立点比较敏感,所以当待评价的样本相似度太低时,也会影响算法评价结果的准确性。而本文中所提到的各种隐写分析算法本身的原理各不相同,所以样本的内聚性并不高,必然会产生孤立点样本,这也大大影响了模糊聚类分析评估结果的准确性。文献[6]虽然也采用了灰色关联分析法,但它在各指标权重的计算上,运用熵权法在一定程度上有其实用性,但熵权法仅考虑了同一指标内部间差异性对权重的影响,而没有考虑到指标间的影响。此外,计算关联度时,对于分辨系数的确定,文献[7]仅根据一般情况,将分辨系数定为0.5,缺乏客观依据,而本文利用式(11)、式(12)对分辨系数进行量化,使其值的确定更加合理。因此,本文提出的评估方法更加准确、合理。

5 结束语

本文提出一种基于改进CRITIC权和灰色关联分析的隐写分析评估算法,并结合实例证明了算法的有效性。该算法不仅解决了主观评价法对领域专家过度依赖并造成主观性太强的问题,而且克服了传统客观评价法如标准差法、变异系数法和熵权法等无法考虑属性之间的影响。同时,在用改进的CRITIC权得到初步的权值后,结合决策者的现实需要进行合理的修改,使得所得到的最终权值克服了主观评价法的盲目性,解决了单一客观赋权法对数据本身的高敏感性。评估结果表明,本文算法的评价结果更加准确、合理。下一步将会在隐写分析算法评估的指标选取方面进行深入研究。

参考文献

- [1] 葛秀慧,田浩. 隐写分析原理与应用[M]. 北京:清华大学出版社,2014.
- [2] 欧继山,李薇,朱婷婷,等. 基于灰色层次分析法的隐写分析算法评估[J]. 计算机与数字工程,2015,43(6):967-971.
- [3] 张伟,郭捷,陈克非. 隐写分析算法的模糊综合评估[J]. 计算机工程,2006,32(21):141-144.
- [4] 龚劬,郭吉强. 基于FCM算法的隐写分析算法综合评估[J]. 计算机工程,2009,35(4):175-176.
- [5] 李开达,张涛,平西建,等. 基于组合赋权及TOPSIS的隐写分析算法综合评估[J]. 应用科学学报,2012,30(4):335-342.
- [6] 张秋余,张燕,袁占亭. 基于熵权与灰色关联度的隐写分析算法评估[J]. 计算机工程,2011,37(7):148-150.
- [7] 王静. 航运上市公司的经营业绩评价——基于CRITIC法的灰色关联分析[J]. 经济与管理,2012(2):94-96.
- [8] So H J, Bonk C J. Examining the Roles of Blended Learning Approaches in Computer-supported Collaborative Learning Environments: A Delphi Study [J]. Educational Technology & Society, 2010, 13(3):189-200.
- [9] 迟国泰,李刚,程砚秋. 基于AHP-标准离差的人的全面评价模型及其实证研究[J]. 管理学报,2010,7(2):301-310.
- [10] 庄平,李延喜. 基于G1-变异系数法的企业投资风险评价模型与实证研究[J]. 软科学,2011,25(10):107-112.
- [11] 章穗,张梅,迟国泰. 基于熵权法的科学技术评价模型及其实证研究[J]. 管理学报,2010,7(1):34-42.
- [12] Kazan H, Ozdemir O. Financial Performance Assessment of Large Scale Conglomerates via Topsis and Critic Methods [J]. International Journal of Management & Sustainability, 2014, 3(4):203-224.
- [13] 吕锋. 灰色系统关联度之分辨系数的研究[J]. 系统工程理论与实践,1997,17(6):49-54.
- [14] 胡烨雯,张新鹏,王朔中. 一种可估计最低位平面嵌入信息量的密写分析方法[J]. 上海交通大学学报,2004,38(z1):122-124.
- [15] 葛秀慧,田浩. 隐写分析原理与应用[M]. 北京:清华大学出版社,2014.
- [16] 张涛,平西建. 基于差分直方图实现LSB信息伪装的可靠检测[J]. 软件学报,2004,15(1):151-158.
- [17] Li Zhi, Sui Aifen, Yang Yiman. A LSB Steganography Detection Algorithm[C]//Proceedings of the 14th IEEE Conference on Personal, Indoor and Mobile Radio Communications. Washington D. C., USA: IEEE Press, 2003:2780-2783.
- [18] Dumitrescu S, Wu Xiaolin, Memon N. On Steganalysis of Random LSB Embedding in Continuous-tone Images[C]//Proceedings of International Conference on Image Processing. Washington D. C., USA: IEEE Press, 2002:641-644.

编辑 顾逸斐