

联网审计绩效的动态评价方法:基于 AHP 和 GM(1,1)的组合应用

陈伟¹ SMIELIAUSKAS Wally² 刘思峰³

(南京审计学院信息管理系 南京 210029)¹ (多伦多大学罗特曼管理学院 多伦多 M5S 3E6)²
(南京航空航天大学经济与管理学院 南京 210016)³

摘要 针对联网审计绩效动态评价的需要,提出了一种基于 AHP 和 GM(1,1)组合应用的联网审计绩效动态评价方法:首先通过建立针对联网审计绩效评价的 AHP 层次结构模型,构造 AHP 判断矩阵,计算各评价指标的权重,评价已实施阶段联网审计项目的绩效;然后,采用 GM(1,1)模型预测联网审计项目的潜在绩效值;最后,对联网审计绩效进行动态评价。为动态评价联网审计的绩效提供了一种有效的方法。

关键词 持续审计,联网审计,IT 绩效评价,层次分析法,GM(1,1)

中图分类号 TP399, F239.1 **文献标识码** A

Dynamic Assessment Method of Online Auditing Performance: A Combined Use between AHP and GM(1,1)

CHEN Wei¹ SMIELIAUSKAS Wally² LIU Si-feng³

(Department of Information Management, Nanjing Audit University, Nanjing 210029, China)¹

(Rotman School of Management, University of Toronto, Toronto M5S 3E6, Canada)²

(College of Economics and Management, Nanjing University of Aeronautics and Astronautics, Nanjing 210016, China)³

Abstract A dynamic assessment method for online auditing performance based on AHP (analytic hierarchy process) and GM(1,1) was proposed in this paper according to the dynamic assessment requirement of online auditing performance. In this method, the AHP hierarchy fit to the characteristics of online auditing used in China was structured. Then, pairwise comparisons of criteria were made and led to a matrix of comparisons. Then, the comparison matrices were translated into weights, and the performances of online auditing were computed. Based on the performances assessment results computed by AHP, GM(1,1) was used to forecast the future performance. Then, the performances of online auditing can be assessed dynamically according to the present and future performance. Therefore, an effective dynamic assessment method for online auditing performance was provided in this study.

Keywords Continuous auditing, Online auditing, IT performance assessment, AHP, GM(1,1)

1 引言

持续审计(Continuous Auditing, CA)是计算机辅助审计的一个重要发展方向^[1-5]。在我国,为了适应计算机辅助审计的需要,国家审计署已经成功开展了“金审工程”一期和二期的建设工作^[7]。我国目前开展的联网审计也是持续审计的一种方式。目前我国正在大力推广联网审计,比如社会保障联网审计、地税联网审计等,推广的试点从省级单位一直到市、县级单位,为研究和应用联网审计绩效评价提供了机遇。同时,为检验联网审计的实施效果,保证国家投资的正确性,联网审计绩效评价变得越来越重要。虽然目前对于联网审计实现技术和应用的研究较多^[6-8],但缺少绩效评价方面的研究与应用。研究联网审计绩效评价,有助于重新认识目前的联网审计政策和联网审计实现技术,并有助于审计机关更有效地制定联网审计实施对策。尽管在一些领域,绩效评价的研究

已引起广泛关注^[9,10],但实施联网审计是一个非常复杂的系统工程,并有着自己的特点,使得已有的一些评价指标和评价方法不完全适合联网审计的特点,不能直接将其应用于联网审计的绩效评价之中。因此,迫切需要根据联网审计的特点构建一些适合其需要的绩效评价方法。

笔者曾研究了如何采用 AHP (Analytic Hierarchy Process, 层次分析法)和 RC(Rank-Centroid, 等级法)进行联网审计绩效评价^[11-13]。但在评估联网审计实施效果的过程中,评估者对客观事物的认识遵循由浅入深的规律,而联网审计实施效果也是通过逐渐显现的方式呈现在决策者面前,评估过程中应采用全面、联系、运动的观点看问题。一般地,一项联网审计工程从实施到终结,必会经历一定的成长发展期,其潜在效应在多长时间内显现、潜在的社会效应究竟有多大等问题尚需要解决。本文中,基于新陈代谢 GM(1,1) 预测模型^[14]的优势,研究一种基于 AHP 和 GM(1,1)的联网审计

到稿日期:2011-08-08 返修日期:2011-10-20 本文受国家自然科学基金项目(70971068, 70701018),教育部人文社会科学研究(08JC630045)项目,江苏省高校“青蓝工程”中青年学术带头人项目(苏教师(2010)27号),江苏省高校优势学科建设工程项目(审计科学与技术)资助。
陈伟(1976—),男,博士后,副教授,主要研究方向为 IT 审计, E-mail: chenweich@nau.edu.cn; SMIELIAUSKAS Wally 男,博士,教授,博士生导师;刘思峰 男,博士,教授,博士生导师。

绩效动态评价方法。

2 评价指标及其权重确定方法

2.1 评价指标的确定

对联网审计的绩效进行评价,评价指标的选取至关重要。基于我国联网审计的实现原理^[5-7],采用专家调查法和文献分析法,分析了影响联网审计绩效的因素,其主要内容包括:(1)基于成本-效益的视角分析影响联网审计绩效的主要因素;(2)基于整个联网审计系统风险控制的视角分析影响联网审计绩效的主要因素;(3)其它因素,如联网审计系统设计与开发的质量等。在此基础上,综合考虑联网审计实施的一次性成本、经常性成本、有形效益、无形效益、联网审计系统的风险控制状况等,设计绩效评价指标。

2.2 建立联网审计绩效评价 AHP 层次结构模型

AHP(层次分析法)^[15]是由美国运筹学家托马斯·塞蒂(T L Saaty)在 20 世纪 70 年代中期提出的一种简便、灵活而又实用的多准则决策方法。基于 AHP 方法的原理以及所设计的联网审计绩效评价指标体系,在专家调查和数据分析的基础上,建立了联网审计绩效评价 AHP 层次结构模型,如图 1 所示。

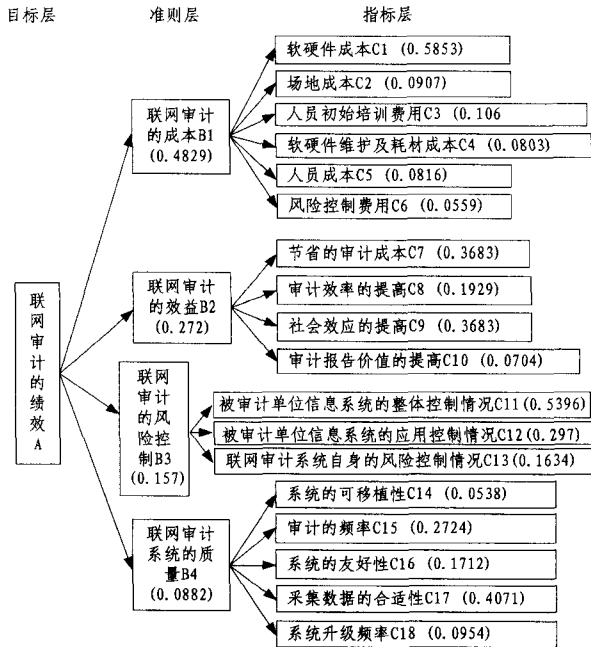


图 1 联网审计绩效评价的 AHP 层次结构模型

2.3 权重计算及一致性检验

根据建立的联网审计绩效评价层次结构模型,采用 1~9 标度法构造相应的判断矩阵如下^[12]:

$$A = \begin{bmatrix} 1 & 2 & 3 & 5 \\ 1/2 & 1 & 2 & 3 \\ 1/3 & 1/2 & 1 & 2 \\ 1/5 & 1/3 & 1/2 & 1 \end{bmatrix}$$

$$B_1 = \begin{bmatrix} 1 & 7 & 6 & 8 & 7 & 9 \\ 1/7 & 1 & 1 & 1 & 1 & 2 \\ 1/6 & 1 & 1 & 2 & 1 & 2 \\ 1/8 & 1 & 1/2 & 1 & 1 & 2 \\ 1/7 & 1 & 1 & 1 & 1 & 1 \\ 1/9 & 1/2 & 1/2 & 1/2 & 1 & 1 \end{bmatrix}$$

$$B_2 = \begin{bmatrix} 1 & 2 & 1 & 5 \\ 1/2 & 1 & 1/2 & 3 \\ 1 & 2 & 1 & 5 \\ 1/5 & 1/3 & 1/5 & 1 \end{bmatrix}$$

$$B_3 = \begin{bmatrix} 1 & 2 & 3 \\ 1/2 & 1 & 2 \\ 1/3 & 1/2 & 1 \end{bmatrix}$$

$$B_4 = \begin{bmatrix} 1 & 1/5 & 1/3 & 1/7 & 1/2 \\ 5 & 1 & 2 & 1/2 & 3 \\ 3 & 1/2 & 1 & 1/2 & 2 \\ 7 & 2 & 2 & 1 & 4 \\ 2 & 1/3 & 1/2 & 1/4 & 1 \end{bmatrix}$$

根据以上判断矩阵,通过计算各个判断矩阵最大特征值 λ_{max} 所对应的特征向量 W ,可以计算出相应的权重。另外,为了避免在构造各判断矩阵的过程中产生非一致性,需要计算各自的一致性比例 CR ,来检验各个判断矩阵的一致性。当 $CR < 0.10$ 时,认为判断矩阵的一致性是可以接受的,否则应对判断矩阵做适当修正。采用 Matlab 计算相应的权重及 CR 值^[12],各自权重的最终结果如图 1 所示。

3 绩效动态评价模型的建立

3.1 动态评价原理

如上所述,联网审计实施效果通过逐渐显现的方式呈现在决策者面前,评估过程中应采用全面、联系、运动的观点看问题。因此,在进行联网审计的绩效评价时,既要重视联网审计已实施阶段的绩效情况,还要考虑未来一段时间内绩效的可能变化。本文利用新陈代谢 GM(1,1) 预测模型,运用动态评价思想,预测联网审计未来一段时间内绩效的潜在发展趋势;在此基础上综合现实绩效情况和预测绩效情况,进行动态评价,从而使联网审计的绩效评价结果更真实、准确和合理,其原理如图 2 所示。另外,为了验证新陈代谢 GM(1,1) 模型的优势,还将其与新信息 GM(1,1) 模型的预测结果进行了比较。

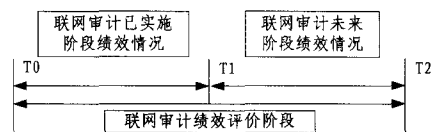


图 2 联网审计绩效的动态评价原理

3.2 基于 GM(1,1) 的潜在绩效测算方法

灰色 GM(1,1) 预测方法分析的是小样本、贫信息的数据,其中新陈代谢 GM(1,1) 模型的特点是用新信息取代旧信息。从预测角度看,新陈代谢 GM(1,1) 模型是数据不多情况下的最理想的预测模型^[14]。

定义 1(GM(1,1) 模型的原始形式) 设 $X^{(0)} = (x^{(0)}(1), x^{(0)}(2), \dots, x^{(0)}(n))$, $X^{(1)} = (x^{(1)}(1), x^{(1)}(2), \dots, x^{(1)}(n))$, 称 $x^{(0)}(k) + ax^{(1)}(k) = b$ 为 GM(1,1) 模型的原始形式。

定义 2(新信息 GM(1,1) 模型) 在 GM(1,1) 模型中,置入最新信息 $x^{(0)}(n+1)$, 称用 $X^{(0)} = (x^{(0)}(1), \dots, x^{(0)}(n), x^{(0)}(n+1))$ 建立的模型为新信息 GM(1,1)。

定义 3(新陈代谢 GM(1,1) 模型) 在 GM(1,1) 模型中,置入最新信息 $x^{(0)}(n+1)$, 去掉最老信息 $x^{(0)}(1)$, 称用 $X^{(0)} = (x^{(0)}(2), \dots, x^{(0)}(n), x^{(0)}(n+1))$ 建立的模型为新陈代谢 GM(1,1)。其时间响应式为

$$\begin{cases} \hat{x}^{(1)}(t+1) = (x^{(0)}(1) - \frac{b}{a})e^{-at} + \frac{b}{a} \\ \hat{x}^{(0)}(t+1) = \hat{x}^{(1)}(t+1) - \hat{x}^{(1)}(t) \end{cases}$$

式中, $t=1, 2, \dots, n$ 。

本文采用新陈代谢 GM(1,1)模型主要是考虑到原始数据有限,软硬件成本、软硬件维护及耗材成本等数据信息不断更新,且老数据的意义将不断降低(特别是评价指标中一次性成本数据)。新陈代谢 GM(1,1)模型在不断补充新数据的同时及时去掉老数据,更能够准确地预测未来的发展状况。本文在应用新陈代谢 GM(1,1)模型时,去掉图 2 中 T_0 时刻的最老绩效值,加入 T_1 时刻的最新绩效值。

3.3 绩效动态评价模型

基于以上联网审计绩效动态评价的原理,以及 AHP 的评价指标权重计算方法,设 t 为评价时间, $t \in [T_0, T_2]$, 则联网审计绩效的动态评价模型如下:

(1) 当 $T_0 < t \leq T_1$ 时,

$$P_t = \sum_{i=1}^{CB_1} w_1 w_i p_{it} + \sum_{i=CB_1+1}^{CB_1+CB_2} w_2 w_i p_{it} + \dots + \sum_{i=CB_1+CB_2+\dots+CB_{l-1}+1}^{CB_1+CB_2+\dots+CB_l} w_l w_i p_{it} \quad (1)$$

式中, w_i 为被评价对象 P_i 准则层中的各个准则的权重, $w_l \in$

$[0, 1]$; w_i 为被评价对象 P_i 指标层中的各个评价指标的权重, $w_i \in [0, 1]$, p_{it} 为被评价对象 P_i 的指标层中的各个评价指标的值, CB_l 为准则层中各准则所包含的指标的个数, $l=1, 2, \dots, S$, S 为准则层中准则的个数。

(2) 当 $T_1 < t \leq T_2$ 时,采用新陈代谢 GM(1,1)模型预测如下:

设

$$X^{(0)} = (x^{(0)}(1), x^{(0)}(2), \dots, x^{(0)}(n)) = (P_{T_0+1}, \dots, P_{T_1})$$

$$P_t = \hat{x}^{(0)}(t+1) = (1 - e^a)(x^{(0)}(1) - \frac{b}{a})e^{-at} \quad (2)$$

根据式(1)和式(2)可以计算出被评价对象各阶段的绩效评价。

4 算例与分析

4.1 数据分析及规范化处理方法

设某单位从 2005 年起实施一项联网审计项目的相关数据如表 1 所列^[11]。本算例中,对于 $C_8, \dots, C_{14}, C_{16}, C_{17}$ 这些在向被评价单位收集数据时不易准确获得的数据,其数据值用区间值表示。为了消除表 1 中联网审计绩效各项评价指标不同量纲对决策结果的影响,同时考虑提高实际评价过程的可操作性,我们采用以下常用方法对指标属性值进行规范化处理。

表 1 联网审计项目绩效评价数据

准则层	B ₁						B ₂			
指标层	C ₁	C ₂	C ₃	C ₄	C ₅	C ₆	C ₇	C ₈	C ₉	C ₁₀
评价对象/(单位)	万元	万元	万元	万元	万元	万元	万元	[1,5]	[1,5]	[1,5]
2005	526	54.7	37.8	8.62	63.5	86.5	-50.8	1	2	1
2006	0	0	0	10.58	61.5	37.5	-35.2	2	3	1
2007	0	0	0	12.83	66.5	34.8	-10.5	3	3	2
2008	0	0	0	18.51	68.5	32.6	20.7	3	4	2
2009	0	0	0	37.27	87.2	52.8	32.9	4	4	3
2010	0	0	0	25.43	74.6	28.3	38.6	4	4	3

(续表)

准则层	B ₃			B ₄				
指标层	C ₁₁	C ₁₂	C ₁₃	C ₁₄	C ₁₅	C ₁₆	C ₁₇	C ₁₈
评价对象/(单位)	[1,5]	[1,5]	[1,5]	[1,5]	次/每月	[1,5]	[1,5]	次/每年
2005	4	4	4	3	10	3	3	1
2006	4	4	4	3	15	3	4	2
2007	4	4	5	4	30	4	5	5
2008	5	5	5	4	30	5	5	6
2009	5	5	5	4	30	5	5	4
2010	5	5	5	4	30	5	5	3

(1) 对于用数值表示的数据,指标属性分为效益型和成本型。成本型是指属性值越小越好的属性,效益型属性是指属性值越大越好的属性。对于成本型指标,采用以下规范化处理方法:

$$f_{ij} = S(a_{ij}) = \frac{\min(a_{ij})}{a_{ij}}, a_{ij} > 0$$

$$f_{ij} = S(a_{ij}) = \frac{\max(a_{ij}) - a_{ij}}{\max(a_{ij}) - \min(a_{ij})}, a_{ij} \leq 0$$

对于效益型指标,采用以下规范化处理方法:

$$f_{ij} = S(a_{ij}) = \frac{a_{ij}}{\max(a_{ij})}, a_{ij} > 0$$

$$f_{ij} = S(a_{ij}) = \frac{a_{ij} - \min(a_{ij})}{\max(a_{ij}) - \min(a_{ij})}, a_{ij} \leq 0$$

式中, $f_{ij} (i=1, 2, \dots, n; j=1, 2, \dots, m)$ 为决策者对评价指标 $C_i (i=1, 2, \dots, n)$ 的属性值 $a_{ij} (i=1, 2, \dots, n; j=1, 2, \dots, m)$ 的

无量纲化值, $S(a_{ij})$ 为指标 a_{ij} 无量纲化的标准函数, $\max(a_{ij})$ 和 $\min(a_{ij})$ 分别为评价指标 C_i 的最大值和最小值。

(2) 针对本算例中用区间值表示的数据的特点,我们采用如表 2 所列的转换方法。表 2 中的对应关系也可以根据不同的评价对象进行调整,从而更准确地完成对不同联网审计项目的绩效评价。

表 2 评价等级区间值和规范处理值的对应关系定义

评价等级区间值	规范处理值
1	0.4
2	0.6
3	0.8
4	0.9
5	1

4.2 绩效动态评价结果及分析

根据对联网审计绩效评价指标的分析, C_1, \dots, C_6 为成本

型指标, C_7, C_{15}, C_{18} 为效益型指标。 $C_8, \dots, C_{14}, C_{16}, C_{17}$ 为用区间值表示的数据。对表 1 中的数据进行规范化处理。根据联网审计的绩效评价模型式(1)计算被评价对象 2005—2010 年的绩效评价价值,其结果如表 3 所列。

表 3 实际绩效值和模拟值

评价年份	实际绩效评价 价值	新陈代谢 GM(1,1) 模拟值	新信息 GM(1,1) 模拟值
2005	0.3698	/	0.3698
2006	0.8102	0.8102	0.8318
2007	0.8696	0.8815	0.8613
2008	0.9231	0.9027	0.8919
2009	0.9187	0.9243	0.9236
2010	0.9437	0.9465	0.9564

根据联网审计的绩效评价模型式(2),置入 2010 年最新绩效值,去掉 2005 年最老绩效值,采用灰色新陈代谢 GM(1,1)模型测算该联网审计项目未来 5 年(2011—2015 年)的潜在绩效值,其结果如表 4 所列。同时,计算出的模拟值(2006—2010 年)如表 3 所列。

表 4 GM(1,1)预测出的绩效值

预测年份	新陈代谢 GM(1,1)	新信息 GM(1,1)
2011	0.9692	0.9903
2012	0.9924	1.0255
2013	1.0162	1.0619
2014	1.0406	1.0996
2015	1.0655	1.1386

其中,发展系数 $a = -0.0237$, 灰色作用量 $b = 0.8519$, 平均相对误差 $\Delta = \frac{1}{5} \sum_{i=2}^6 \Delta_i = 1.1287\%$ 。

另外,为了突显新陈代谢 GM(1,1)模型的优势,置入 2010 年最新绩效值,采用新信息 GM(1,1)模型测算该联网审计项目的模拟值(2005—2010 年)和未来 5 年(2011—2015 年)的潜在绩效值,其结果分别如表 3 和表 4 所列。

其中,发展系数 $a = -0.0349$, 灰色作用量 $b = 0.8045$, 平均相对误差 $\Delta = \frac{1}{6} \sum_{i=1}^6 \Delta_i = 1.781\%$ 。根据上述计算结果,该被评价对象 2005—2015 年的绩效动态评价结果如图 3 所示。

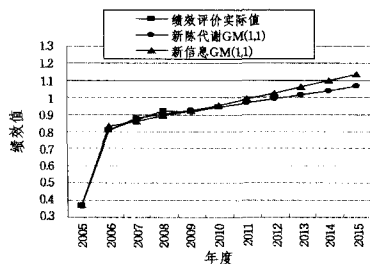


图 3 绩效动态评价结果

根据动态评价结果可以看出:

(1)该联网审计项目的绩效值在未来 5 年呈稳定上升趋势。在今后的过程中,该单位应继续提高联网审计的执行力度,并不断采用新方法(加入新信息、去掉老信息)测算其未来的绩效发展趋势。当发现未来的绩效水平出现下降趋势时,应该提前查找相应的原因,改进不足,必要情况下终止联网审计项目的实施。

(2)由于软硬件设备经过一定时期的使用后,在某一时期需较大幅度的更新,如表 1 中,2009 年各成本型指标值增加较多。尽管 2009 年的绩效评价值(0.9187)有所下降,但该联

网审计项目的绩效值在未来 5 年仍呈稳定上升趋势。因此,动态评价方法更真实和合理。

(3)新陈代谢 GM(1,1)模型的模拟结果和预测结果较为合理,新陈代谢 GM(1,1)模型优于新信息 GM(1,1)模型。

结束语 IT 绩效审计是目前审计领域研究的一个前沿和热点问题,联网审计是 IT 审计研究的一个重要方向。预测和评价联网审计实施效果的动态发展态势,建立联网审计的多阶段动态评价方法对研究联网审计绩效评价具有重要意义。本文利用新陈代谢 GM(1,1)预测模型,运用动态评价思想,对联网审计未来一段时间内绩效的潜在发展趋势进行预测。在此基础上,综合现实绩效情况和预测绩效情况进行动态评价,这样不但能更合理地评价联网审计的绩效,而且评价结果也更有应用价值。另外,通过本文的研究,相关建议如下:

(1)本文采用 AHP 计算各评价指标的权重,在实际应用中,由于 AHP 和 RC 两种方法在计算权重时的不足,可根据情况选择采用 AHP 和 RC 组合应用的方法计算各评价指标的权重^[13]。在此基础上,再建立联网审计的多阶段的动态评价方法。

(2)在实际应用中,也可以根据实际情况,先利用新陈代谢 GM(1,1)预测模型,运用动态评价思想,对联网审计各评价指标数据的潜在发展趋势进行预测。在此基础上,再综合现实数据和预测数据计算绩效值进行动态评价。

(3)在今后的工作中,基于灰色关联模型研究影响联网审计绩效评价的因素也将是联网审计绩效评价的一个研究重点,作者将另文讨论^[16]。

参考文献

- [1] 陈伟,张金城, Qiu R. 计算机辅助审计技术(CAATs)研究综述[J]. 计算机科学, 2007, 34(10): 290-294
- [2] Groomer S M, Murthy U S. Continuous auditing of database applications: an embedded audit module approach[J]. Journal of Information Systems, 1989, 3(2): 53-69
- [3] Alexander K, Ephraim F S, Miklos A V. Continuous online auditing: a program of research[J]. Journal of Information Systems, 1999, 13(2): 87-103
- [4] Rezaee Z, Sharbatoghlie A, Elam R, et al. Continuous auditing: building automated auditing capability[J]. Auditing: A Journal of Practice and Theory, 2002, 21(1): 147-163
- [5] Chen W, Zhang J C, Jiang Y Q. One continuous auditing practice in China: data-oriented online auditing (DOOA) [A]// The 7th IFIP International Conference on e-Business, e-Services, and e-Society (I3E2007) [C]. Boston: Springer, 2007: 521-528
- [6] 陈伟, Qiu R, 刘思峰. 持续审计(CA)研究综述[J]. 小型微型计算机系统, 2008, 29(9): 1755-1760
- [7] 国家 863 计划审计署课题组. 计算机审计数据采集与处理技术研究报告[M]. 北京: 清华大学出版社, 2006
- [8] 陈伟, Qiu R, 刘思峰. 一种基于数据匹配技术的审计证据获取方法[J]. 计算机科学, 2008, 35(8): 183-187, 194
- [9] Wei C C. Evaluating the performance of an ERP system based on the knowledge of ERP implementation objectives[J]. The International Journal of Advanced Manufacturing Technology, 2008, 39(1/2): 168-181
- [10] Sueyoshi T, Shang J, Chiang W C. A decision support framework for internal audit prioritization in a rental car company: A com-

bined use between DEA and AHP [J]. European Journal of Operational Research, 2009, 199(1): 219-231

- [11] 陈伟. 一种基于等级法的联网审计绩效评价方法[J]. 计算机科学, 2010, 37(11): 111-116
- [12] 陈伟. 一种基于 AHP 的联网审计绩效评价方法[J]. 审计与经济研究, 2011, 26(5): 47-52
- [13] 陈伟, Smieliauskas W. 联网审计技术与绩效评价[M]. 北京:

清华大学出版社, 2012

- [14] 刘思峰, 党耀国, 方志耕, 等. 灰色系统理论及其应用(第五版)[M]. 北京: 科学出版社, 2010
- [15] Satty T L. The Analytic Hierarchy Process [M]. New York: McGraw-Hill, 1980
- [16] 陈伟, Smieliauskas W, 刘思峰. 联网审计绩效评价影响因素的灰色关联分析[J]. 计算机应用研究, 2012, 29(2): 435-437

(上接第 177 页)

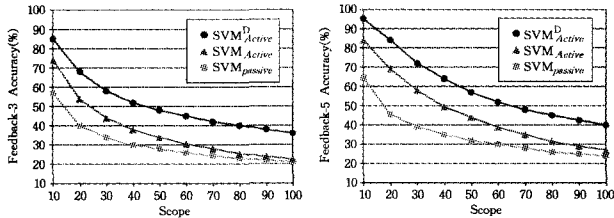
像, 其中颜色特征由 125 维的颜色直方图和 6 维的颜色矩构成, 而纹理特征则利用 3 层的离散小波变换之后在 10 个子带上分别取均值和方差构成一个 20 维的特征。参与比较的各种方法均使用 RBF 核的 SVM 分类器, 相关参数由交叉验证决定。在每一轮相关反馈中, 各种方法均选取相同数目的图像进行标注。

图 4 给出了分别在 3 轮和 5 轮相关反馈之后返回的前 10~100 个结果的准确率, 横坐标 x 表示前 x 个返回结果, 纵坐标表示准确率; 图 5 给出了前 30 和前 50 个返回结果范围内准确率随相关反馈轮次变化的趋势, 横坐标 x 对应于第 x 次相关反馈, 纵坐标表示准确率。

本进行标注, 从而保证了所选取样本对于模型改进的价值, 最终利用有限的样本标注量获得尽可能大的分类性能提升。实验结果表明, 相对于传统方法, 动态可行域划分算法能够显著提高所选取样本的信息量, 对于减轻人工标注负担、改善分类性能具有重要意义。

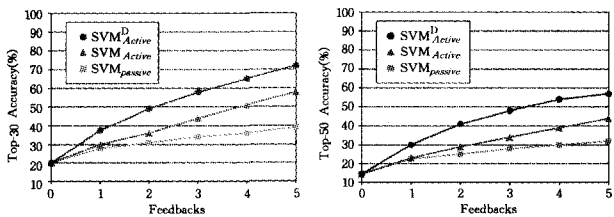
参考文献

- [1] Zhu X. Semi-supervised learning literature survey [R]. Wisconsin; Computer Sciences, University of Wisconsin-Madison, 2008
- [2] Cohn A, Ghahramani Z, Jordan M I. Active learning with statistical models [J]. Journal of Artificial Intelligence Research, 1996, 4(1): 129-145
- [3] McCallum A, Nigam K. Employing EM in pool-based active learning for text classification [C] // Proceeding of the 15th International Conference on Machine Learning. San Francisco: Morgan Kaufmann, 1998: 350-358
- [4] Burges J C A. A tutorial on support vector machines for pattern recognition [J]. Data Mining and Knowledge Discovery, 1998, 2(2): 121-167
- [5] Mitchell T. Generalization as search [J]. Artificial Intelligence, 1982, 18(2): 203-226
- [6] Tong S, Koller D. Support vector machine active learning with applications to text classification [J]. The Journal of Machine Learning Research, 2000, 2(1): 45-66
- [7] Tong S, Chang E. Support vector machine active learning for image retrieval [C] // Proceedings of the 9th ACM International Conference on Multimedia. New York: ACM, 2001: 107-118
- [8] Zhang X, Cheng J, Lu H, et al. Weighted co-SVM for image retrieval with MVB strategy [C] // Proceedings of 2009 IEEE International Conference on Image Processing. San Antonio, TX: Signal Processing Society, 2007: 517-520
- [9] Smeulders A W M, Worring M, Santini S, et al. Content-based image retrieval at the end of the early years [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2000, 22(12): 1349-1380
- [10] Lew M S, Sebe N, Djeraba C, et al. Content-based multimedia information retrieval: state of the art and challenges [J]. ACM Transactions on Multimedia Computing, Communications, and Applications, 2006, 2(1): 1-19
- [11] Rui Y, Huang T S, Ortega M, et al. Relevance feedback: a power tool for interactive content-based image retrieval [J]. IEEE Transactions on Circuits and Systems for Video Technology, 1998, 8(5): 644-655
- [12] Zhou X S, Huang T S. Relevance feedback in image retrieval: a comprehensive review [J]. Multimedia Systems, 2003, 8(6): 536-544



(a) 3 轮相关反馈之后的准确率 (b) 5 轮相关反馈之后的准确率

图 4 前 n ($n=10, 20, \dots, 100$) 个返回结果的准确率



(a) 前 30 个返回结果的准确率 (b) 前 50 个返回结果的准确率

图 5 经过 k ($k=0, 1, \dots, 5$) 轮相关反馈之后的准确率

实验结果表明, 主动学习(曲线●、▲)优于非主动学习(曲线■), 说明通过有针对性地主动选取样本进行标注, 能够更加有效地改进分类性能; 主动学习中, 动态可行域划分算法(曲线●)优于传统批量采样方法(曲线▲), 说明在主动学习中选择性采样方法起到了至关重要的作用, 动态可行域划分算法通过综合利用当前分类模型和先前标注样本两方面信息对采样过程进行动态优化, 进一步提高了所选取样本的信息量, 从而显著地提高分类精度。

结束语 本文深入分析了 SVM 分类模型下的主动学习的特点, 针对传统批量采样方法的不足, 提出了动态可行域划分算法。根据特征空间与参数空间的对偶关系, 特征空间中对样本的标注等价于参数空间中对可行域的划分。算法采用动态批量采样模式, 通过综合利用当前分类模型和先前标注样本两方面信息, 动态选取具有更优的可行域划分能力的样