

基于 AHP-SVM 的信息系统风险评估^①

鲁华栋, 闫兵, 岳小冰

(河南工业职业技术学院 电子信息工程系, 南阳 473000)

摘要: 风险评估保证信息系统安全的一种重要技术, 信息系统风险受到众多影响, 具有动态性和时变性, 为了提高信息系统风险评估的精度, 提出一种层次分析法和支持向量机的信息系统风险评估模型(AHP-SVM). 首先对当前信息系统风险研究现状进行分析, 并根据专家构建评估指标体系, 然后利用层次分析法对评估指标权重进行估计, 并根据支持向量机建立信息系统风险的评估模型, 最后利用信息系统风险数据对评估结果的可靠性进行测试. 结果表明, 本文模型可以合理确定评估指标的权重, 可以描述信息系统风险因子与期望值间的变化关系, 获得了理想的信息系统风险评估结果, 具有一定的实际应用价值.

关键词: 系统风险; 影响因素; 指标权重; 层次分析法

Risk Evaluation of Information System Security Based on Analytic Hierarchy Process and Support Vector Machine

LU Hua-Dong, YAN Bing, YUE Xiao-Bing

(Henan Polytechnic Institute, Nanyang 473000, China)

Abstract: Risk assessment is an important technology to ensure the security of information system. Information system risk assessment is subject to a number of effects and has dynamic and changeable, in order to improve the accuracy of information system risk assessment, this paper puts forward a information system risk assessment model based on hierarchical analysis method and neural network. At first, the current situation of the information system risk are analyzed, and this paper builds the evaluation index system according to the experts, and then the analytic hierarchy process is used to estimate the weights of evaluation indexes, and SVM neural network is used to establish the information system risk assessment model, at last, the reliability of the evaluation results is tested by using the risk data of information system. The results show that the proposed model can reasonably determine the weight of evaluation index, it can describe relationship between risk factors and expected value of information system, obtains ideal assessment results, and so it has certain practical value.

Key words: system risk; influencing factor; index weight; analytic hierarchy process

1 引言

近些年, 信息技术的不断发展和成熟, 已经在各个领域得到了深入应用, 涌现了了大量的信息系统^[1,2]. 信息系统给人们带来便利的同时, 一些病毒、木马程序对其安全性进行破坏, 使得信息系统存在一些安全隐患, 风险评估可以了解信息系统的风险状态和安全态势, 因此风险评估成为当前信息系统安全领域待解决的一个难题^[3].

信息系统风险评估一直受到了广大高校和教育管理部门的重视, 最早采用人工方式的评价方法, 评价结果主观性较强、不科学, 且评价精度低, 不能反映教师信息系统风险的真实水平^[4]. 随后, 出现了基于专家系统的信息系统风险评估系统, 一定程度上提高了信息系统风险评估精度, 但信息系统风险评估是一个多目标、多层次的评价问题, 涉及指标和内容多, 且专家系统评价指标选择受到个人偏好、学识水平等因

① 基金项目:河南省科技攻关项目(132102210208)

收稿时间:2016-01-12;收到修改稿时间:2016-05-30 [doi:10.15888/j.cnki.csa.005546]

素影响,难以排除人为因素带来的误差,导致评价结果主观性较强,评价结果可信度低^[5].近些年,随着信息技术的迅速发展,出现基于多元线性回归、偏最小二乘、支持向量机、神经网络等信息系统风险评估方法^[6-8].多元线性回归、偏最小二乘是基于线性关系进行评价,无法对非线性问题进行很好的逼近,信息系统风险各评价指标之间以及与评价结果间是一种复杂非线性关系,因此线性评价方法应用范围受限.人工神经网络是一种基于“大数定理”的学习算法,要求训练样本大,易出现过拟合缺陷^[9].支持向量机是一种专门针对小样本的学习算法,泛化性能要优于神经网络.信息系统因子的重要性十分关键,即权重,当前通常采用人工方式确定,导致权重确定的合理性受到质疑,信息系统安全风险评估结果不一完全合理^[10-12].层次分析法(AHP)是一种定性和定量相结合的分析方法,其将与决策有关的元素分解为多个层次,可以用于确定评价指标的权重^[13].

为了提高信息系统风险评估的精度,提出一种层次分析法和支持向量机的信息系统风险评估模型(AHP-SVM),实验结果表明,本文模型可以合理确定评估指标的权重,获得了较优的信息系统风险评估结果.

2 层次分析法和支持向量机

2.1 层次分析法

层次分析法(AHP)是一种定性与定量相结合的分析方法,其将一个复杂系统划分为目标层、准则层和指标层,然后对各层次判断矩阵进行构建,在实际应用中,常采用 1-9 标度法确定各个因素的权重,从而得到不同层次的判断矩阵,具体如表 1 所示.

表 1 判断矩阵标度的含义

标度	含义
1	x,y同等重要
3	x比y素稍微重要
5	x比y很重要
7	x比y非常重要
9	x比y绝对重要
2, 4, 6, 8	上述两相邻判断的中值

2.2 支持向量机

支持向量机(SVM)通过找到一个最优分类超平面将所有训练样本划分两类,即

$$y_i\{\psi(x_i), \omega\} + b \geq 1, i = 1, 2, \dots, n \quad (1)$$

式中, n 表示训练样本的数量.

设 $\psi(x_i), y_i$ 与分类超平面最小距离为 $1/\|\omega\|$, 允许存在一些误分类的点, 这样式(1)变为

$$y_i\{\psi(x_i), \omega\} + b \geq 1 - \zeta_i, i = 1, 2, \dots, n \quad (2)$$

式中, $\zeta_i (i = 1, 2, \dots, n)$ 为负松弛变量, $\zeta_i = 0$ 时, 表示完全线性可分.

对于一个线性不可分问题, 那么就需要将其转为一个优化问题, 再找到最优分类超平面, 通过引入惩罚因子 $C > 0$, 则有

$$\min \psi(\omega) = \frac{1}{2} \|\omega\|^2 + C \sum_{i=1}^n \zeta_i \quad (3)$$

s.t.

$$\begin{cases} y_i\{\psi(x_i), \omega\} + b \geq 1 - \zeta_i \\ i = 1, 2, \dots, n \end{cases}$$

引入 Lagrange 算子 α_i 将式(3)转化:

$$\max W(\alpha) = \sum_{i=1}^n \alpha_i - \frac{1}{2} \sum_{i,j=1}^n \alpha_i \alpha_j y_i y_j \psi(x_i) \psi(x_j) \quad (4)$$

s.t.

$$\begin{cases} 0 \leq \alpha_i \leq C \\ \sum_{i=1}^n \alpha_i y_i = 0 \end{cases}$$

式中, $\omega = \sum_{i=1}^n \alpha_i \psi(x_i)$.

支持向量机的分类判别函数为

$$f(x) = \omega \psi(x) + b = \sum_{i \in SV} \alpha_i \psi(x_i) \psi(x) + b \quad (5)$$

特征空间的内积(核函数)为

$$K(x, y) = \sum_i \psi_i(x) \psi_j(y) \quad (6)$$

高斯核定义为

$$K(x, y) = \exp\left(-\frac{\|x - y\|^2}{2\sigma^2}\right) \quad (7)$$

式中, σ 为高斯分布宽度.

支持向量机的分类判别函数为

$$f(x) = \omega \psi(x) + b = \sum_{i \in SV} \alpha_i K(x_i, x) + b \quad (8)$$

3 AHP-SVM的信息系统风险评估模型

3.1 数学模型

信息系统评估是采用一定的理论和方法对信息系

统存在的安全和风险进行判断,并根据评估结果做出相应的防范措施,以提高信息系统的安全性和可靠性.信息系统是一个动态的变化过程,影响信息系统风险的因素众多,各因素之间互相关联,作用程度也不一致,使得信息系统风险影响因素和风险值之间是一非线性映射关系,难以建立用一个合理、精确的数学解析式.设信息系统风险评估因素为 $\{x_1, x_2, \dots, x_m\}$,信息系统风险评估的数学模型为

$$y = f(x_1, x_2, \dots, x_n) \quad (9)$$

式中, $f()$ 表示信息系统评估函数.

信息系统评估的关键是找到一个优的评估函数 $f()$,准确描述信息系统风险的变化态势,支持向量机是一种非线性分类能力优异的机器学习算法,本文选择支持向量机建立信息系统评估模型.

3.2 评估指标体系

在分析当前信息系统风险影响因素的基础上,本文构建了表 2 的信息系统风险评估指标体系.

表 2 评估指标体系

信 息 系 统 风 险 评 估 指 标 体 系	物理安全 x_1	环境安全 x_{11}	设备安全 x_{12}
		应用安全 x_{13}	管理安全 x_{14}
	运行安全 x_2	备份与恢复 x_{21}	病毒防范 x_{22}
		电磁兼容 x_{23}	操作系统安全 x_{24}
		应用系统安全 x_{25}	数据库安全 x_{26}
		通信协议安全 x_{27}	
	应用安全 x_3	身份鉴别 x_{31}	防范控制 x_{32}
		信息加密 x_{33}	抗抵赖性 x_{34}
		信息完整性校验 x_{35}	入侵检测 x_{36}
		安全审计 x_{37}	电磁泄露防护 x_{38}
	安全管理 x_4	安全管理机构 x_{41}	安全管理人员 x_{42}
		安全管理制度 x_{43}	安全技术设施 x_{44}
		密码和密钥管理 x_{45}	

4 AHP-SVM的信息系统风险评估模型

为了提高信息系统风险评估的准确性,提出一种AHP和SVM的信息系统风险评估方法.首先建立层次结构的评估指标体系,并采用层次分析法计算每一个评估指标的权重,然后采用训练样本,对自适应能力与非线性映射能力优异的SVM学习,建立信息系统风险评估模型.

4.1 AHP 确定指标权重

AHP 根据表 1 的标度描述评估指标之间的相对重

重,两个指标标度的计算公式为.

$$A_{ij} = \begin{cases} 1, & \lambda_{ij} < 1 \\ [EV_i / EV_j], & 1 \leq \lambda_{ij} \leq 9 \\ 9, & \lambda_{ij} > 9 \end{cases} \quad (10)$$

式中, EV_i 和 EV_j 为影响程度值.

对计算结果进行整理,得到层次分析法的判断矩阵如表3和表4所示.

表 3 准则层的判断矩阵

A_{ij}	x_1	x_3	x_3	x_4	权重
x_1	1	2	2	2	0.315
x_2	0.5	1	1	1	0.270
x_3	0.5	1	1	1	0.182
x_4	0.5	1	1	1	0.233

表 4 各个指标的判断矩阵

指标编号	权重	指标编号	权重	指标编号	权重
X_{11}	0.287	X_{25}	0.137	X_{36}	0.112
X_{12}	0.193	X_{26}	0.137	X_{37}	0.053
X_{13}	0.265	X_{27}	0.128	X_{38}	0.066
X_{14}	0.255	X_{31}	0.166	X_{41}	0.285
X_{21}	0.182	X_{32}	0.128	X_{42}	0.279
X_{22}	0.107	X_{33}	0.167	X_{43}	0.154
X_{23}	0.147	X_{34}	0.128	X_{44}	0.168
X_{24}	0.162	X_{35}	0.18	X_{45}	0.114

当专家对两个评价指进行判断时,如果判断矩阵的阶数 $n>2$,那么就会出现不一致现象,导致得到错误的决策结果,因此,对判断矩阵的一致性进行校验,具体如下:

$$CR = \frac{\lambda_{max} - n}{n - 1} \quad (11)$$

式中, λ_{max} 为最大特征值; n 为矩阵的阶数.

各个判断矩阵的一致性检验结果如表5所示.

表 5 指标一致性的检验结果

准则层	x_1	x_3	x_3	x_4
λ_{max}	4	2.014	2.98	2.52
CR	0	0.0014	0.023	0.019

4.2 AHP-SVM 信息系统风险评估步骤

(1) 根据相关资料和专家经验建立信息系统风险评估指标体系.

(2) 针对具体的信息系统,收集相应的风险评估数据.

(3) 对影响指标和期望值进行归一化处理.

(4) 采用 AHP 确定评估指标体系的权重值.

(5) 确定支持向量机相关参数.

(6) 将训练样本输入到支持向量机进行学习, 建立信息系统风险评估模型.

(7) 对测试样本进行测试, 输出信息系统风险评估结果.

综合上述可知, 基于 AHP-SVM 的信息系统风险评估模型框架如图 1 所示.

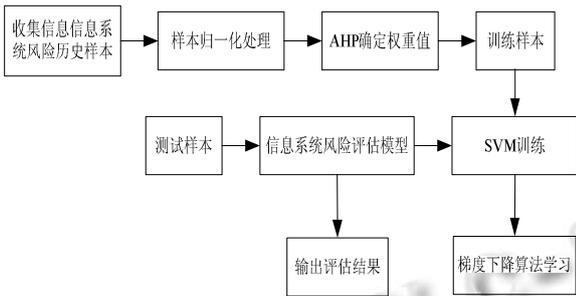


图 1 信息系统风险评估模型的框架

5 实验结果与分析

为了分析 AHP-SVM 的信息系统风险评估的可行性, 选择中国工商银行黑龙江省分行哈尔滨市南岗支行的信息系统作为实验对象, 信息系统风险值的变化曲线如图 2 所示, 风险分为 5 级, 分别采用 1、2、3、4、5 描述, 共收集到 200 个样本数据, 选择前 150 样本作为训练集建立 AHP-SVM 的信息系统风险评估模型, 其余样本数据作为测试集对 AHP-SVM 的有效性进行验证.

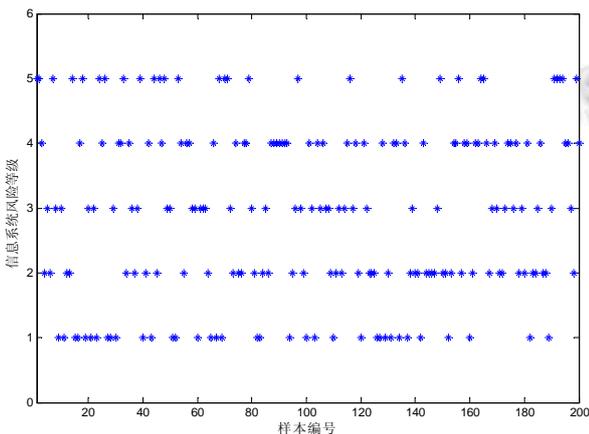


图 2 信息系统风险的样本

采用 AHP 确定评价指标的权重, 并对训练样本集和测试样本集进行相应处理, 然后采用支持向量机对信息系统训练样本集进行建模, 构建基于 AHP-SVM

的信息系统风险评估模型, 并对测试集样本进行评估, 统计实验结果具体如图 3 所示, 从图 3 可知, AHP-SVM 可以准确反映信息系统风险变化趋势, 期望值与评估值之间的误差小小, 平均评估正确率达到 95% 以上, 远远超过信息系统安全实际应用的 85%, 实验结果表明, AHP-SVM 可以应用于实际的信息系统风险评估中, 评估结果科学、可信.

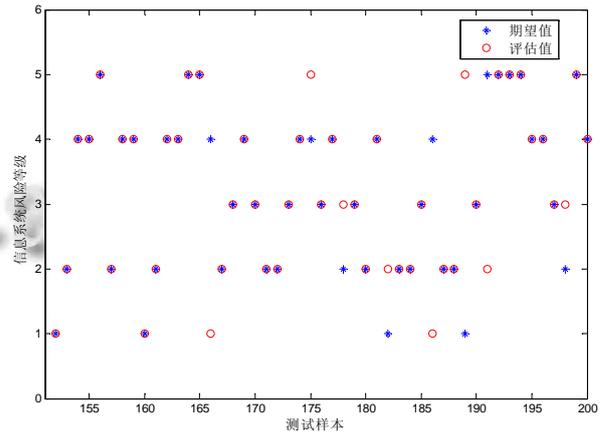


图 3 AHP-SVM 的评估性能测试结果

为了进一步分析 AHP-SVM 的优越性, 选择单一层次分析法(AHP)和支持向量机(SVM)以及经典信息系统网络评估模型^[15]进行对比实验, 它们的平均评估正确率如表 6 所示. 评估正确率和评估误差定义如下:

$$\text{评估正确率} = \frac{\text{正确评估样本数}}{\text{总样本数}} \times 100\% \quad (12)$$

$$\text{评估错误率} = \frac{\text{错误评估样本数}}{\text{总样本数}} \times 100\% \quad (13)$$

AHP-SVM 的信息系统风险评估准确性要好于其它模型, 具有明显的优势, 为信息系统管理员提供有价值的信息.

表 6 AHP-SVM 与其它模型性能对比

模型	评估正确率(%)	评估错误率(%)
AHP	88.58	11.42
SVM	92.14	7.86
文献[15]	93.35	6.65
AHP-SVM	95.80	4.2

6 结论

为了更好的保证信息系统安全, 以提高信息系统风险评估的准确性为目标, 提出一种层次分析和支持向量机的信息系统风险评估模型. 仿真测试结果表明,

本文模型是一种精度高、结果可靠的信息系统安全风险评估模型,在信息系统安全管理中具有广泛的应用前景.

参考文献

- 1 王祯学,周安民,方勇,等.信息系统安全风险估计与控制理论.北京:科学出版社,2011.
- 2 郭璇.信息系统安全风险估计方法和技术研究.电子技术与软件工程,2016,4:213-214.
- 3 陈颂,王光伟,刘欣宇,等.信息系统安全风险估计研究.通信技术,2012,1:128-130.
- 4 刘佳,徐赐文.信息安全风险评估方法的比较分析.中央民族大学学报(自然科学版),2012,(2):91-96.
- 5 牛红惠,刘凌霞.支持向量机在信息安全风险评估中应用研究.计算机仿真,2011,28(6):117-120.
- 6 付沙.一种基于信息熵的信息系统安全风险分析方法.情报科学,2013,6:38-42.
- 7 赵刚,刘换.基于多层次模糊综合评判及熵权理论的实用风险评估.清华大学学报(自然科学版),2012,(10):1382-1387.
- 8 郑雷雷,宋丽华,郭锐,等.故障树分析法在信息安全风险评估中的应用,2011,38(10):106-109.
- 9 王帆,霍明奎,王晓婷.基于模糊灰度的信息系统安全风险估计与对策.情报科学,2014,32(1):110-114.
- 10 罗辉,潘平,李换双.基于量子支持向量机的信息安全风险评估方法.贵州大学学报(自然科学版),2013,30(1):74-78.
- 11 慧,党德鹏.基于 RBF 模糊支持向量机的信息安全风险评估.计算机工程与设计,2011,32(6):2113-2115.
- 12 付钰,吴晓平,叶清,彭熙.基于模糊集与熵权理论的信息系统安全风险评价研究.电子学报,2010,38(7):1489-1494.
- 13 刘加伶,付明明,冯欣,张红.最优组合赋权法在信息安全风险评估中的应用.重庆理工大学学报(自然科学),2016,30(3):87-93.