

于晓虹, 楼文高, 洪赢政. 城市大型商场火灾风险评价的投影寻踪模型与应用[J]. 灾害学, 2017, 32(4): 17-22. [YU Xiaohong, LOU Wengao and HONG Yingzheng. Fire Risk Comprehensive Evaluation of the Large Urban Building based on Projection Pursuit Clustering Model and its Application[J]. Journal of Catastrophology, 2017, 32(4): 17-22. doi: 10.3969/j.issn.1000-811X.2017.04.004]

# 城市大型商场火灾风险评价的投影寻踪模型与应用\*

于晓虹<sup>1</sup>, 楼文高<sup>2</sup>, 洪赢政<sup>3</sup>

(1. 上海商学院 东方财富传媒与管理学院, 上海 200235; 2. 上海商学院, 上海 200235 ;  
3. 公安部上海消防研究所, 上海 200438)

**摘 要:** 根据城市大型商场火灾风险评价指标体系和 5 级风险评价标准, 采用群搜索最优化方法, 建立了投影寻踪分类模型, 对 15 个待判定样本进行综合评价, 得到了可靠、有效的评价结果。再应用基于信息熵权重的 TOPSIS 和灰色关联法进行评价, 结果与投影寻踪模型分类结果一致。分析了现有的火灾风险神经网络评价模型文献存在的主要问题。研究表明: 投影寻踪分类模型能很好地适用于城市大型商场火灾风险的综合评价与分类研究, 数学意义清晰, 后续应用方便。

**关键词:** 火灾; 风险评价; 投影寻踪分类模型; 城市大型商场; 评价与分类; 群搜索

**中图分类号:** TP183; X913.4 **文献标志码:** A **文章编号:** 1000-811X(2017)04-0017-06

doi: 10.3969/j.issn.1000-811X.2017.04.004

随着城镇化加速发展和人民生活水平不断提高, 城市大型商场(包括酒店、宾馆、地铁等)越来越多, 同时, 火灾发生数量、人员和财产损失也日益严重。因此, 对城市大型商场火灾风险进行可靠、有效的评估已成为学界和政府部门的热点和难点问题。对火灾风险开展评估研究, 国外始于 1970 年代, 国内始于 20 世纪末, 已提出和采用多种综合方法(模型)<sup>[1-3]</sup>, 还有物元可拓法(MEM)<sup>[4]</sup>、突变理论法(CT)<sup>[5]</sup>和支持向量机模型(SVM)。文献<sup>[2-3, 6-16]</sup>把上述常规方法的计算结果作为神经网络(NN)模型的训练样本, 建立了各种组合 NN 模型。此外, 还有 GRA-集对分析<sup>[17]</sup>、主成分分析(PCA)-BPNN 等组合模型<sup>[18]</sup>。上述常规方法都必须采用其他方法确定各个指标的权重, 各种方法的合理性迄今也没有定论, 而这本身就是综合评价要解决的核心问题之一。因此, 如果权重不合理, 其评价结果就难言合理了, 基于这些结果建立 NN 模型的合理性就难以保证了。况且, 建立 NN 模型还必须满足二个基本条件<sup>[19-22]</sup>: ①训练样本数量必须多于网络连接权重个数, 一般要求 3~5 倍以上; ②必须采用检验样本实时监控训练过程以避免发生“过训练”现象。令人遗憾的是, 文献<sup>[2-3, 6-16]</sup>都不满足上述条件, 因此, 建立的 NN 模型都不具有泛化能力和实用价值。文献<sup>[1]</sup>针对具有区间评价标准的问题,

通过在区间内随机生成 50 个样本的方法建立了回归型 PPC 模型, 取得了较好效果, 但不适用于没有区间评价标准的情况, 样本权重也很接近, 有待进一步改进。因此, 城市大型商场火灾风险评估问题虽然进行了有益的探索性研究, 但还远没有得到真正解决。

Friedman 等提出的可用于高维、非线性、非正态分布数据建模的投影寻踪分类(PPC)技术<sup>[23-25]</sup>, 在综合评价领域获得了广泛应用。为此, 本文在没有区间评价标准的情况下, 尝试直接根据评价标准样本建立 PPC 模型, 取得了更合理和有效的结果, 提供了一种新的大型商场火灾风险评估方法。

## 1 建立城市大型商场火灾风险评估指标体系及其数据预处理

### 1.1 评价指标体系及其评价标准

不同学者往往根据已有经验和学识建立不同的城市大型商场火灾风险综合评价指标体系。考虑到本文研究重点是比较不同评价(建模)方法的合理性和有效性, 本文直接采用有较多数据和待评价样本支撑的、由岳海玲<sup>[2]</sup>建立的评价指标体系, 由 19 个指标和五级火灾风险评价标准组成, 待评价样本(S1~S15)如表 1 所示。

\* 收稿日期: 2017-02-22 修回日期: 2017-05-06

基金项目: 国家重点研发计划课题(2016YFC0800608)

第一作者简介: 于晓虹(1978-), 女, 安徽合肥人, 硕士, 讲师, 研究方向为会计学、商业经济与管理、投影寻踪等数据挖掘技术及其应用。E-mail: yu\_xiaohong@21cn.com

通讯作者: 洪赢政(1976-), 男, 安徽亳州人, 硕士, 副研究员, 研究方向为数字化消防单兵装备及消防安全。

E-mail: 50411261@qq.com



### 1.2 建模样本及其数据预处理

为了判定表 1 中样本 S1 ~ S15 的火灾风险等级, 并且建模时尽可能避免受这些待判定样本数据的影响, 笔者根据 5 级评价标准样本和待判定样本各指标的最大值(逆向指标为最小值)与最小值组成两个虚拟样本 S-max 与 S-min, 构建了 7 个建模样本, 即很安全、安全、较安全、不安全、很不安全(也可以简称为 I ~ V 级风险)、S-max 和 S-min 样本。为了减小各评价指标不同量纲对建模结果的不利影响,

采用去均值归一化方法  $x_{i,j} = \frac{x_{i,j}^* - \bar{x}_j}{\sigma_j}$  进行预处理,  $x_{i,j}^*$  和  $x_{i,j}$  分别为第  $i$  个样本、第  $j$  个指标归一化前后的数据,  $\bar{x}_j$  和  $\sigma_j$  分别为第  $j$  个指标原始数据的均值和标准差,  $i = 1, 2, \dots, n$  ( $n$  为样本数量),  $j = 1, 2, \dots, p$  ( $p$  为指标个数)。

## 2 建立城市大型商场火灾风险评估 PPC 模型

### 2.1 PPC 建模原理简介

Friedman 等<sup>[23]</sup>提出了一种适用于非线性、高维和非正态分布数据处理的投影寻踪分类(PPC)模型, 把高维数据投影到感兴趣的 1 ~ 3 维低维子空间上, 通过分析低维子空间上的数据规律以达到揭示高维数据结构特征的目的。PPC 建模基本思想是“使样本投影点整体上尽可能分散, 在局部上尽可能密集”, 最有效一维模型的目标函数为<sup>[23-25]</sup>:

$$Q(a) = \max(S_z \times D_z);$$

$$s. t. \sum_{j=1}^p a_j^2 = 1, 1 \geq a_j \geq -1. \quad (1)$$

式中: 样本投影值标准差  $S_z = \sqrt{\{\sum_{i=1}^n [z(i) - \bar{z}]^2\} / (n - 1)}$ , 其值越大表示投影点整体上越分散; 局部密度值  $D_z = \sum_{i=1}^n \sum_{k=1}^n (R - r_{i,k}) \cdot u(R - r_{i,k})$ , 其值越大表示投影点局部越密集; 样本投影值  $z(i) = \sum_{j=1}^p a_j x_{i,j}$  ( $\bar{z}$  是其均值);  $R$  为窗口半径;  $r_{i,k} = |z(i) - z(k)|$  表示样本  $i$  和  $k$  之间的距离;  $u(t)$  为单位阶跃函数, 当  $t \geq 0$  时取 1, 否则为 0。

一维 PPC 模型(1)的结果随  $R$  值不同而改变。楼文高等<sup>[24]</sup>指出:  $R$  的合理取值范围为  $r_{\max} / 5 \leq R \leq r_{\max} / 3$  ( $r_{\max}$  为  $r_{i,k}$  的最大值)。由于(1)式为同时含有等式和不等式约束的高维非线性最优化问题, 有很多的局部最优解, 求解该问题十分困难。本文采用群搜索算法(简称 GSO, 计算原理见文献[26])进行最优化求解。

### 2.2 城市大型商场火灾风险 PPC 建模研究

把很安全、安全、……、很不安全、S-max 和 S-min 共 7 个样本归一化后的数据导入笔者编制的基于

GSO 的 PPC 程序, 取  $R = r_{\max} / 5$ , 得到真正的全局最优解: 最佳投影向量系数  $a_{1-19} = (0.2425, -0.2406, -0.1766, 0.2173, -0.2400, -0.2436, -0.2449, -0.2350, -0.2384, -0.2327, -0.2432, 0.2275, -0.2176, -0.2274, -0.2312, 0.2189, -0.2563, -0.2380, -0.1675)$ , 各样本投影值  $z(1) \sim z(7)$  ( $-5.2517, -2.5145, 0.2962, 3.4391, 6.4624, 1.5539, -3.9856$ ),  $S_z = 4.2008, D_z = 23.3822, Q(a) = 98.2241, R = 2.3428, r_{\max} = 11.7141$ 。由前五个样本的投影值意义可知, 城市大型商场 I ~ V 级火灾风险的模型输出值分别为  $-5.2517, -2.5145, 0.2962, 3.4391$  和  $6.4624$ , 其值越大表示火灾风险越高, 越不安全。待判定样本(案例)的风险等级可以根据就近原则来判定。得到岳海玲<sup>[2]</sup>中 15 个待判定样本 S1 ~ S15 的 PPC 模型输出值分别为  $-2.2948, -1.7998, -1.4694, -1.2191, -2.2088, -0.5362, -1.2945, -0.1157, -1.9241, -1.5303, -1.3323, -1.5701, -1.5912, -1.8125, -2.0641$ 。待判定样本 S1 ~ S15(图 1 中用  $\diamond$  表示)、S-max 和 S-min(图 1 中用  $\Delta$  表示)的 PPC 模型值和 I ~ V 级火灾风险评价标准样本的模型值如图 1 所示(直线 I、II、III、IV 和 V 分别表示 I ~ V 级评价标准样本的模型值)。从图 1 可以看出, 所有待判定样本的模型值均介于 II 和 III 级之间, 其中 S6 和 S8 更接近于评价标准“较安全”水平(III 级), 其他样本都更接近于“安全”水平(II 级)。从样本 S-max、S-min 的输出值分别为 1.5539(属于 III 级风险)和  $-3.9856$ (属于 I 级风险)可知, 所有待判定样本的火灾风险不可能高于 S-max 的风险等级(III 级), 也不可能低于 S-min 的风险等级(I 级)。15 个待判定样本火灾风险大小排序为  $S8 > S6 > S4 > S7 > S11 > S3 > S10 > S13 > S12 > S2 > S14 > S9 > S15 > S5 > S1$ 。

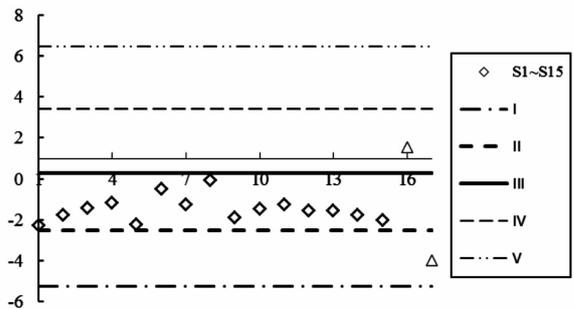


图 1 待判定样本 S1 ~ S15(横坐标 1 ~ 15, 图中用  $\diamond$  表示)以及样本 S-max 与 S-min(横坐标 16、17, 图中用  $\Delta$  表示)的输出值和 I ~ V 级风险评价标准样本的输出值(图中用直线 I ~ V 表示)

取  $R = r_{\max} / 3$  时, 得到上述相似的结果, 限于篇幅, 不再列出(备索)。

表2 典型样本 S1、S6、S8 和 S11 各个评价指标的火灾风险等级

样本	$x_1$	$x_2$	$x_3$	$x_4$	$x_5$	$x_6$	$x_7$	$x_8$	$x_9$	$x_{10}$	$x_{11}$	$x_{12}$	$x_{13}$	$x_{14}$	$x_{15}$	$x_{16}$	$x_{17}$	$x_{18}$	$x_{19}$
S3	4*	2	1	2	2	2	2	2	2	2	2	4	2	2	1~2	4	3	2	5
S6	3	2	2	4	3	2	2	2	2	2	3	2	2	2	3	3	3~4	3	4
S8	2~3	2	2~3	5	3	2	3	3	2~3	5	3	1	2	3	3~4	2~3	4	3	3
S11	3	2	2~3	3	2	2~3	2	2	2	1	3	2	2	2	2	3	3	2	4

注：1、2、3、4和5分别表示该指标值属于I、II、III、IV和V级风险水平，2~3表示该指标值处于II和III级的中值或很接近于中值，其他同。

### 3 结果与分析

#### 3.1 城市大型商场 15 个待判定样本火灾风险的判定及其指标的重要性分析

PPC 模型判定 15 个样本的火灾风险等级为：样本 S6 和 S8 的火灾风险为 III 级(较安全)，但 S8 的火灾风险大于 S6，其他 13 个样本的火灾风险为“安全”(II 级)。各个指标的最大权重与最小权重之比达到了 1.53，说明指标之间的重要性有较大差异，比文献[1]的结果更合理。在 19 个评价指标中， $x_{17}$  最重要，然后依次是  $x_7$ 、 $x_6$ 、 $x_{11}$ 、 $x_1$ 、 $x_2$ 、 $x_5$ 、 $x_9$ 、 $x_{18}$ 、 $x_8$ 、 $x_{10}$ 、 $x_{15}$ 、 $x_{12}$ 、 $x_{14}$ 、 $x_{16}$ 、 $x_{13}$ 、 $x_4$ 、 $x_3$  和  $x_{19}$ 。

#### 3.2 岳海玲评价结果<sup>[2]</sup>及其存在问题的分析

岳海玲<sup>[2]</sup>根据表 1 的 5 级评价标准，首先采用 GRA 对样本 S1~S14 进行评价，其结果是 I 级风险样本 1 个，II 级风险样本 8 个，III 级风险样本 3 个，IV 级风险样本 2 个和 V 级风险样本 1 个。该结果又作为 BPNN 模型训练样本的期望值，没有采用检验样本进行实时监控训练过程以避免发生“过训练”现象，建立了网络结构为 10-13-1 的 BPNN 模型，但论文没有说明是怎么从 19 个指标中筛选出 10 个评价指标的，而其训练样本的期望值却是根据 19 个指标值计算得到的，这必定会出现误差。作为典型分析案例，笔者列出了样本 S3、S6、S8 和 S11 各个评价指标的风险等级，如表 2 所示。

样本 S3 的 19 个指标中，分别有 1 个、12 个、1 个、3 个和 1 个指标值属于 I、II、III、IV 和 V 级风险，1 个指标值处于 I 级与 II 级的中值，显然，S3 仅有一个指标值属于 V 级风险，而岳海玲<sup>[2]</sup>却把其判定为 V 级风险，显然是不合理的。笔者与岳海玲<sup>[2]</sup>采用相同的 GRA 方法，得到 S3 与 I、II、III、IV 和 V 级标准样本的关联度分别为 0.115、0.621、0.146、0.201 和 0.019，应判定其为 II 级风险，与大多数单个指标的风险等级是一致的，是合理的。样本 S8 分别有 1 个、3 个、8 个、1 个和 1 个指标值处于 I、II、III、IV 和 V 级风险，有 4 个指标值处于 II 级与 III 级的中值，1

个指标值处于 III 级与 IV 级的中值。因此，S8 与 S3 相比，火灾风险应该更大，而岳海玲<sup>[2]</sup>却把 S8 判定为 III 级，明显低于 S3 的风险等级，显然是错误的。对于样本 S11，分别有 1 个、10 个、5 个和 1 个指标值处于 I、II、III 和 IV 级风险，有 2 个指标值为 II 级与 III 级的中值，仅有 1 个指标值属于 I 级风险，岳海玲<sup>[2]</sup>却把 S11 判定为 I 级风险，显然是不合理的，笔者计算得到 S11 与 5 级标准样本的关联度分别为 0.117、0.522、0.319、0.034 和 0.011，S11 显然应该是 II 级风险。从各个指标值的风险等级来看，S11 与 S3 应该基本相当，而岳海玲<sup>[2]</sup>却把他们分别判定为 I 级和 V 级风险。本文 PPC 建模结果是：样本 S3、S6、S8 和 S11 分别处于 II 级、III 级、III 级和 II 级风险，其中 S8 风险最大，S6 次之，S3 的风险最小，比较符合实际情况。

因此，岳海玲<sup>[2]</sup>采用 GRA 的评价结果存在明显错误，以此作为 BPNN 模型训练样本的期望值，其结果必然也是错误的。而且，岳海玲<sup>[2]</sup>建立 BPNN 模型时，训练样本数量(14 个)少于评价指标个数(19 个)，是不可能建立可靠的 BPNN 非线性模型的，也不可能泛化能力和实用价值。

#### 3.3 采用基于信息熵权重的 TOPSIS 法和 GRA 法进行综合评价

基于信息熵权重的 TOPSIS 法和 GRA 法的计算原理分别见文献[27-28]，本文仅给出计算结果。各指标的信息熵权重  $\omega_{1-19} = (0.0499, 0.0519, 0.0441, 0.0647, 0.0492, 0.0499, 0.0643, 0.0499, 0.0592, 0.0596, 0.0643, 0.0499, 0.0518, 0.0489, 0.0541, 0.0461, 0.0540, 0.0499, 0.0380)$ ，最大权重与最小权重之比与 PPC 模型的结果比较一致。求得 15 个待判定样本与最优方案(火灾风险最小)的相对贴近度  $C_{1-15} = (0.1496, 0.1829, 0.2251, 0.2728, 0.1349, 0.3595, 0.2388, 0.4407, 0.1776, 0.2478, 0.2614, 0.2070, 0.2211, 0.1937, 0.2132)$ ，S-max、S-min 的相对贴近度分别为 0.6578 和 0.0293，I~V 级火灾风险评价标准样本的相对贴近度分别为 0.0095、0.1236、0.4650、0.8887、1.0000。因此，可以判定所有样本的风险等级：① S6 和 S8 为 III 级风险，其他待判定样本为 II 级风

表 3 所有待判定样本、S-max 和 S-min 与 I、II、III、IV 和 V 级火灾风险标准样本的关联度

	S1	S2	S3	S4	S5	S6	S7	S8	S9	S10	S11	S12	S13	S14	S15	S-max	S-min
I	0.704	0.670	0.669	0.638	0.689	0.593	0.631	0.570	0.678	0.645	0.638	0.658	0.665	0.654	0.699	0.491	0.818
II	0.765	0.811	0.819	0.762	0.716	0.748	0.728	0.696	0.772	0.819	0.781	0.742	0.771	0.760	0.792	0.601	0.773
III	0.692	0.718	0.697	0.729	0.692	0.789	0.716	0.799	0.728	0.752	0.764	0.733	0.735	0.723	0.682	0.787	0.573
IV	0.565	0.584	0.613	0.624	0.556	0.650	0.623	0.663	0.577	0.598	0.609	0.583	0.594	0.589	0.581	0.780	0.477
V	0.479	0.494	0.516	0.519	0.464	0.536	0.513	0.553	0.487	0.505	0.511	0.483	0.497	0.493	0.491	0.629	0.415

险; ②样本 S-max 和 S-min 分别为 III 级和 I 级风险; ③15 个待判定样本的火灾风险大小排序为 S8 > S6 > S4 > S11 > S10 > S7 > S3 > S13 > S15 > S12 > S14 > S2 > S9 > S1 > S5, 该结果与两种 PPC 模型的排序结果基本一致, 与岳海玲<sup>[2]</sup>的结果相差很大。

应用 GRA 法求得了 15 个待判定样本与 I、II、III、IV 和 V 级风险的关联度, 如表 3 所示。可以判定样本 S6 和 S8 为 III 级风险, 其他样本为 II 级风险, 样本 S-max 和 S-min 分别为 III 级和 I 级风险。因此, 笔者采用 GRA 法的判定结果与 PPC 模型以及 TOPSIS 法的结果完全一致, 与岳海玲<sup>[2]</sup>的结果相差很大。

### 3.4 应用 NN 进行火灾风险评价研究文献存在的问题分析

采用 NN 模型进行火灾风险评价的文献如表 4 所示。如上所述, 建立可靠的 NN 模型必须满足上述二个基本条件, 文献[29]也明确指出: 必须将样本分成训练样本、检验样本和测试样本, 训练样本数量应该达到网络连接权重的 10 倍以上, 用检验样本实时监控训练过程, 并在出现“过训练”现象时停止训练。显然, 文献[4-15, 17]的训练样本数量都少于网络连接权重个数, 还都没有采用检验样本实时监控训练过程, 必定会发生“过训练”现象, 导致 NN 模型没有泛化能力和实用价值。同时, NN 建模具有显著的“垃圾进去, 垃圾出来”(garbage in, garbage out)特征<sup>[19, 21, 23]</sup>, 如果前置评价方法的结果不合理, 则后续 NN 建模就毫无意义。

## 4 结束语

城市大型商场火灾风险综合评价是当前城市管理中的重要课题之一。研究探索出一种简捷、实用、有效的火灾风险综合评价方法具有重要的理论意义和实践价值。根据已有的火灾风险评价指标体系和五级风险评价标准, 本文尝试建立火灾风险综合评价的 PPC 模型, 火灾风险得分与评价指标之间呈线性关系, 数学意义清晰, 模型结构

表 4 现有采用 NN 建模文献的具体情况

文献序号	拓扑结构	隐层节点个数确定方法	训练样本	训练样本期望值*
2	10-13-1	没有说明	14	GRA
3	43-15-1	根据公式	15	FAHP-FCM
6**	7-6-1	没有说明	26	CM
8	8-20-1	没有说明	25	GRA
9	8-20-3	根据公式	20	模糊数学法
10	6-13-1	根据公式	40	CM
11	21-8-1	根据公式	10	AHP
13	4-9-1	根据公式	20	CM
15	8-4-1	少于样本数	15	灰色关联法
16	15-25-1	没有说明	50	Delphi

注: \* 是指 NN 模型训练样本的期望输出值是用这些方法确定的。\*\* 其他文献没有说明拓扑结构, 故略。

简单, 后续应用便捷。通过计算各个新样本的投影值, 就可以很方便地判定其火灾风险等级, 不受其他因素的影响。对 15 个待判定样本的评价研究表明, PPC 模型得到了可靠、有效的评价结果。

采用基于信息熵法权重的 TOPSIS 法和 GRA 法进行对比研究, 15 个待判定样本的评定风险等级与 PPC 模型的结果相同。现有文献都不满足 BPNN 建模的二个基本条件, 样本数据太少, 没有采用检验样本实时监控训练过程很容易发生“过训练”现象, 导致 NN 模型没有泛化能力和实用价值。

## 参考文献:

- [1] 吴瑶, 冯国珍, 楼文高. 投影寻踪法在高层建筑火灾风险评估中的应用[J]. 灾害学, 2016, 31(3): 196-201.
- [2] 岳海玲. 基于遗传神经网络的地下商场火灾风险评价方法[D]. 西安: 西安科技大学, 2006.
- [3] 李姗姗. 大型商业综合体火灾风险性分析[D]. 西安: 西安科技大学, 2015.
- [4] 雷中英, 陈微. 基于物元可拓方法的高层建筑防火系统安全评估[J]. 长江大学学报(自然科学版), 2009, 6(2): 104-107.
- [5] 丁元春, 翁发禄, 唐敏康. 基于突变理论的高层建筑火灾安全综合评价[J]. 建筑科学, 2012, 28(1): 22-25.

- [6] 徐敏, 陈国良, 周心权. 高层建筑火灾风险的神经网络评价[J]. 湘潭矿业学院学报, 2003, 2003(3): 69-72.
- [7] 伍爱友, 施式亮, 王从陆. 基于神经网络和遗传算法的城市火灾风险评价模型[J]. 中国安全科学学报, 2006, 16(11): 108-111.
- [8] 孙宁. 基于 BP 神经网络的高层建筑火灾危险性评估方法研究[D]. 西安: 西安科技大学, 2006.
- [9] 岳海玲, 李树刚, 郭子东, 等. 基于 RBF 神经网络的高层建筑火灾风险评价[J]. 火灾科学, 2006, 15(2): 77-81.
- [10] Xia D. Fire risk evaluation model of high-rise buildings based on multilevel BP neural network[C]//2007 Fourth International Conference on Fuzzy Systems and Knowledge Discovery (FSKD 2007), Haikou, Hainan, 2007: 24-27.
- [11] 于作明, 王小文. 基于 BP 神经网络的商场建筑火灾风险评价[J]. 沈阳航空工业学院学报, 2007, 24(1): 72-74.
- [12] 宋译, 肖国清, 何利文. 基于人工神经网络理论的建筑物火灾安全评价研究[J]. 中国安全科学学报, 2008, 18(4): 61-67.
- [13] 钟凯. 基于 BP 神经网络的高层建筑防火评估[D]. 呼和浩特: 内蒙古科技大学, 2012.
- [14] 高臣勇. 高层民用建筑火灾风险评估指标体系对比研究[J]. 武警学院学报, 2013, 29(6): 35-38.
- [15] 李柯萱, 邵必林, 刘浩. 基于灰关联和神经网络的大型商场火灾风险评价方法[J]. 安全与环境学报, 2013, 13(1): 254-258.
- [16] 周瑾. 基于神经网络的建筑物火灾风险评价[J]. 中国安全生产科学技术, 2013, 9(10): 177-182.
- [17] 陈骥, 邹树梁, 唐德文, 等. 基于灰色关联度的集对分析方法在高层建筑火灾危险性评价中的应用[J]. 中国安全生产科学技术, 2015, 11(5): 136-141.
- [18] 牛发阳, 段美栋, 王建波, 等. 基于 PCA-FPP-BP 神经网络的高层建筑火灾安全评价[J]. 工业安全与环保, 2016, 42(7): 26-29.
- [19] StatSoft Inc. Electronic Statistics Textbook [EB/OL]. [2013] (<http://www.statsoft.com/textbook>).
- [20] Zhang G, Patuwo B E, Hu M Y. Forecasting with artificial neural networks: The state of the art[J]. International Journal of Forecasting, 1998, 14(1): 35-62.
- [21] 楼文高, 乔龙. 基于神经网络的金融风险预警模型及其实证研究[J]. 金融论坛, 2011(11): 52-61.
- [22] 楼文高. 活性污泥工艺神经网络模拟与控制研究[D]. 上海: 同济大学, 2005.
- [23] Friedman J H, Tukey J W. A Projection Pursuit Algorithm for Exploratory Data Analysis [J]. IEEE Transactions on Computers, 1974, C-23(9): 881-890.
- [24] 楼文高, 乔龙. 投影寻踪聚类建模理论的新探索与实证研究[J]. 数理统计与管理, 2015, 34(1): 47-58.
- [25] 楼文高, 熊聘, 冯国珍, 等. 影响投影寻踪聚类建模的关键因素分析与实证研究[J/OL]. 数据统计与管理, 1-21. (2017-07-14). <http://kns.cnki.net/kcms/detail/11.2242.01.20170714.1738.014.html>. DoI: 10.13860/i.cnki.stti.20170717-002.
- [26] 张雯雯, 刘华艳. 改进的群搜索优化算法在 MATLAB 中的实现[J]. 电脑与信息技术, 2010, 18(3): 44-46.
- [27] 楼文高, 王广雷, 冯国珍. 旅游安全预警 TOPSIS 评价研究及其应用[J]. 旅游学刊, 2013, 28(4): 66-75.
- [28] 唐启义. DPS 数据处理系统: 实验设计、统计分析及数据挖掘[M]. 北京: 科学出版社, 2010.
- [29] Marsland S. Machine Learning: An Algorithmic Perspective[M]. New York: Chapman & Hall/CRC, 2014.

## Fire Risk Comprehensive Evaluation of the Large Urban Building based on Projection Pursuit Clustering Model and its Application

YU Xiaohong<sup>1</sup>, LOU Wengao<sup>2</sup> and HONG Yingzheng<sup>3</sup>

(1. Faculty of Eastmoney Media and Management, Shanghai Business School, Shanghai 200235, China;

2. Shanghai Business School, Shanghai 200235, China;

3. Shanghai Fire Research Institute of Ministry of Public Security, Shanghai 200438, China)

**Abstract:** Based upon the fire risk comprehensive evaluation (FRCE) index system of the urban shopping malls (USM) and its five-class evaluation criteria, the projection pursuit clustering (PPC) model for FRCE of the USM were established by applying group search optimization (GSO) in this paper. The reliable and effective results of unknown fifteen samples were obtained by applying the PPC model. The results applying PPC model are good agreement with the results using the TOPSIS (Technique for Order Preference by Similarity to Ideal Solution) and gray relation analysis (GRA) based on information-entropy weights. The shortcomings of the published literatures on FRCE using neural networks are discussed and analyzed, and these established BPNN models are unreliable, non-generalized and useless. The case study shows that the PPC model is robust and suitable for FRCE of the USM. The PPC models possess the characteristics of clear concept and convenient application.

**Key words:** fire risk; projection pursuit clustering model; urban large building; evaluation and clustering; group search optimization(GSO)