

基于灰色聚类理论和人工神经网络技术的 道路交通安全评价

潘艳荣*, 翟长旭, 朱顺应

(重庆交通学院 交通运输学院 重庆 400074)

摘要 在深入研究了道路交通安全的评价指标确定的基础上,提出了一种基于灰色聚类理论和神经网络技术的综合评价方法,并对神经网络输入、输出指标属性值的量化方法进行了讨论,建立了道路交通安全的BP神经网络评价模型,最后,介绍了该模型在各省道路交通安全评价中的应用实例。

关键词 灰色聚类理论;神经网络;交通安全

中图分类号 U491.31;TP183 **文献标识码** A **文章编号** 1001-716X(2005)02-0101-04

道路交通安全评价是在人类与道路交通事故做斗争的几十年上百年的过程中,经过不断完善综合道路安全研究成果而逐步成型的。对道路交通安全的评价是道路交通安全管理、决策的基础。近年来,随着交通事故的历年增加对道路交通安全方面的评价和研究一直是人们关注的焦点,国内外先后有价值函数法、层次分析法、模糊综合评价方法等。但是,人们已逐渐认识到这些评价方法势必会存在一定的人为因素,很难保证评价结果的公正性和准确性。并且,这类评价都缺乏自组织自学习能力,难以获得评价专家的经验知识和直觉思维能力,难以获得全面的综合信息。随着近年来思维科学和灰色数学等一系列学科的发展,为计算思维的建立提供了必要的基础和条件。为此本文在深入研究道路交通安全评价问题特点的基础上,根据神经网络具有较强的模式识别和自学习、自适应能力的特点,将神经网络的方法应用到交通安全评价中,建立了基于灰色聚类理论和神经网络技术的评价模型。

1 道路交通安全的灰色变权聚类评价方法

在评价道路交通安全水平时,有时不可能也没有必要在获得全部指标的统计信息后再进行评价。针对交通安全信息不完全的特点,可通过对少量已知信息的筛选、加工、延伸和扩展,运用灰色理论的聚类评价方法,将道路交通安全水平确定在某一区域内,对道路交通安全进行宏观评价。

1.1 评价指标的选取

道路交通事故的影响因素包括人(驾驶员或参与者)、车辆、道路、气候、交通管理及环境等。对交通安全指标的选取,直接或间接的反映了一个地区的交通安全水平,因此所选指标要从量上一定程度地反映这些因素质的特性,即要选取对评价目标起主要作用的指标。最后,参考我国的交通安全现状及国情的特点,确定的评价指标为:①人口事故率(人/10万人口);②车辆事故率(人/万车);③运行事故率(人/亿车公里)。

1.2 灰类及白化值的确定

由于评价指标的量纲不一致(即不可公度性),因而无法直接进行比较分析。为此,对各种评价指标,要经行无量纲化处理。针对交通安全评价指标的特点,一般采用下列公式:

$$x_{ij} = (d_{ij} - d_j^{\min}) / (d_j^{\max} - d_j^{\min}) \quad (i = 1, 2, \dots, m, j = 1, 2, \dots, n) \quad (1)$$

将经无量纲化处理的数据,采用概率统计方法确定评价标准,分析数据的累积百分频率,绘制累积百分频率曲线,在曲线上确定不同特定累积百分频率所对应的数值,作为各灰类的白化值。将交叉口的交通安全状况划分为4个级别的灰类:即优(A级)、良(B级)、中(C级)、差(D级)。选取15%、85%累积百分频率对应的点来确定优和差值,选取40%和60%累积百分频率对应的点来确定良和中值。4个

* 收稿日期 2004-03-03;修订日期 2004-04-28

作者简介 潘艳荣(1979-),女,河北唐山人,硕士,主要从事交通规划与管理方面的研究。

累积百分频率点所对应的 $\lambda_1, \lambda_2, \lambda_3, \lambda_4$ 分别为指标优、良、中、差的白化值, 确定后的各灰类白化值可在一定时间(如一年或几年)内保持稳定。

1.3 灰类的白化权函数

交通安全评价 4 个灰类相应的白化权函数见式(2)至式(5)及图 1。

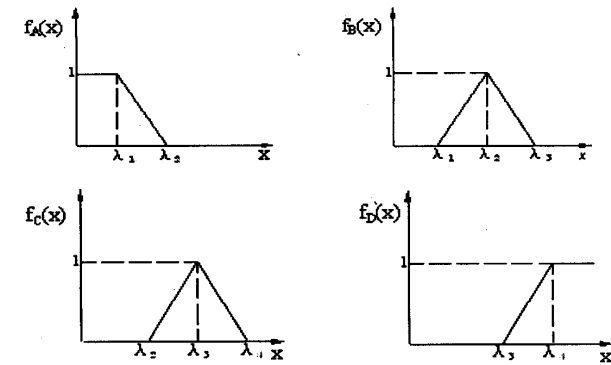


图 1

$$f_A(x) = \begin{cases} 1 & x < \lambda_1 \\ (\lambda_2 - x)(\lambda_2 - \lambda_1) & \lambda_1 \leq x \leq \lambda_2 \\ 0 & x > \lambda_2 \end{cases} \quad (2)$$

$$f_B(x) = \begin{cases} 0 & x < \lambda_1 \\ (x - \lambda_1)(\lambda_2 - \lambda_1) & \lambda_1 \leq x \leq \lambda_2 \\ (\lambda_3 - x)(\lambda_3 - \lambda_2) & \lambda_2 < x \leq \lambda_3 \\ 0 & x > \lambda_3 \end{cases} \quad (3)$$

$$f_C(x) = \begin{cases} 0 & x < \lambda_2 \\ (x - \lambda_2)(\lambda_3 - \lambda_2) & \lambda_2 \leq x \leq \lambda_3 \\ (\lambda_4 - x)(\lambda_4 - \lambda_3) & \lambda_3 < x \leq \lambda_4 \\ 0 & x > \lambda_4 \end{cases} \quad (4)$$

$$f_D(x) = \begin{cases} 0 & x < \lambda_3 \\ (x - \lambda_3)(\lambda_4 - \lambda_3) & \lambda_3 \leq x \leq \lambda_4 \\ 1 & x > \lambda_4 \end{cases} \quad (5)$$

式中 $f_A(x), f_B(x), f_C(x), f_D(x)$ 为交通安全评价指标 A 级、B 级、C 级、D 级的白化权函数。

1.4 聚类权的确定

记聚类权为 η_j^k, k 为评价灰类, η_j^k 按下式确定:

$$\eta_j^k = \lambda_j^k / \sum_{j=1}^m \lambda_j^k \quad (6)$$

式中 η_j^k ——第 j 项评价指标归入 k 种灰类的聚类权; λ_j^k ——第 j 项评价指标属于第 k 种灰类的白化值。

1.5 灰聚类分析

假设 x_{ij} 为评价对象 i 关于指标 j 的标本, 可按下式求出第 i 个评价对象对于第 k 个灰类的变权聚类系数 σ_i^k :

$$\sigma_i^k = \sum_{j=1}^m f_j^k(x_{ij}) \cdot \eta_j^k \quad (7)$$

假设 $\sigma_i^k = \max_{i \leq k \leq s} \{\sigma_i^k\}$, 则假设对象 i 属于灰类 k^* 。

2 交通安全评价中神经网络方法的实现

人工神经网络是发展极为迅速的一门边缘学科, 它是一种诸如知识表达、推理学习、联想记忆乃至复杂社会现象的统一模型, 在综合评价方面有着广泛的应用前景。根据道路交通安全评价的特点, 本文尝试将 3 层(2 个隐含层、1 个输出层)BP 神经网络结构应用于交通安全评价, 多输入多输出的神经网络结构如图 2 所示。

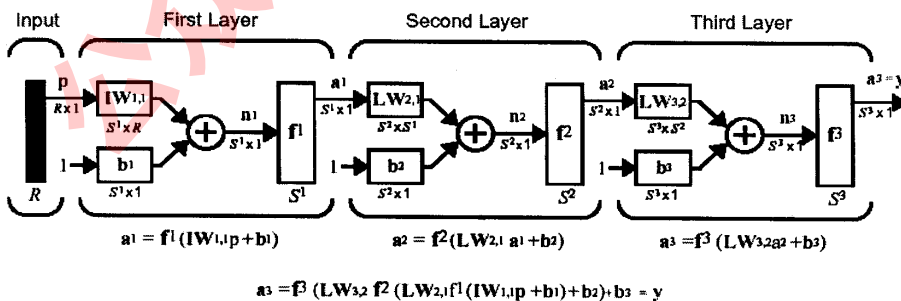


图 2 3 层 BP 神经网络结构图

2.1 神经元传递函数

隐层神经元传递函数为 S 型, 输出神经元为线型, 可以逼近任何函数。

S_1 型: $\text{Tansig}(x)$ 型 $f(x) = \text{tansig}(x) = (e^x - e^{-x}) / (e^x + e^{-x})$ (8)

S_2 型: $\text{Logsig}(x)$ 型 $f(x) = \text{logsig}(x) = 1 / (1 + e^{-x})$ (9)

Purelin 型 $f(x) = \text{purelin}(x) = x$ (10)

2.2 BP 神经网络的学习

网络学习指导思想是, 对网络权重(w_{ij}, T_{ij})进行修正, 使误差函数(E)沿负梯度方向下降。输入节点输入 x_i , 隐节点输出 y_i , 输出节点输出 O_i , 输入节点到隐节点权重为 w_{ij} , 隐节点到输出节点的权重为

T_{ij} 输出节点期望输出 t_l .

① 隐节点的输出

$$y_i = f(\sum_j w_{ij}x_j - \theta_i) = f(\text{net}_i) \tag{11}$$

② 输出节点的计算输出

$$O_l = f(\sum_i T_{li}y_i - \theta_l) = f(\text{net}_l) \tag{12}$$

③ 输出节点的误差

$$\begin{aligned} E &= 1/2 \cdot \sum_l (t_l - O_l)^2 = 1/2 \cdot \sum_l (t_l - f(\sum_i T_{li}y_i - \theta_l))^2 \\ &= 1/2 \cdot \sum_l (t_l - f(\sum_i T_{li}(f(\sum_j w_{ij}x_j - \theta_i) - \theta_l)))^2 \end{aligned} \tag{13}$$

2.3 神经网络方法实现的步骤

在道路交通安全评价神经网络方法实现过程中,样本模式的输入、输出值的确定是关键.以评价指标值为网络的输入值,以灰色变权聚类的评价结果为输出值,在样本输入值的确定过程中,首先根据道路安全评价指标的确定原则,确定了各评价指标,以各评价指标值为网络的输入,以灰色聚类评价的结果作为网络的目标输出,即输入向量为: {人口死

亡率 车辆事故率 运行事故率 } 输出目标向量为: $\{\sigma_i^A \ \sigma_i^B \ \sigma_i^C \ \sigma_i^D\}$ 其中 $i = 1 \ 2 \ \dots \ m$ 为各评价对象; $A、B、C、D$ 为所属灰类. 设好各层的传递函数 和网络的训练函数及学习速率 l_r , 网络的迭代目标 E (误差平方和), 然后对该神经网络进行学习训练.

3 实例分析

3.1 灰色聚类评价

① 白化值 λ_j^k 的确定

根据表 1 所示的 31 个省、自治区、直辖市各评价指标值, 绘制 3 个指标的累积百分频率曲线图, 确定 4 个灰类的白化值如下:

灰类: $t = \{\text{优}(A) \ \text{良}(B) \ \text{中}(C) \ \text{差}(D)\}$

评价指标

$$\lambda_j^k = \begin{bmatrix} 0.328 & 0.470 & 0.545 & 0.708 \\ 0.125 & 0.252 & 0.335 & 0.835 \\ 0.078 & 0.168 & 0.185 & 0.570 \end{bmatrix} \begin{matrix} j = 1 \\ j = 2 \\ j = 3 \end{matrix}$$

② 聚类权的确定

表 1 31 个省份的评价指标值

地区	北京	天津	河北	山西	内蒙古	辽宁	吉林	黑龙江	上海	江苏	浙江	安徽	福建	江西	山东	河南
指标 ①	0.6397	0.7068	0.5119	0.5183	0.4238	0.6075	0.4898	0.2872	0.5125	0.5630	0.8454	0.3340	0.7648	0.4158	0.6223	0.2515
指标 ②	0.3419	0.6665	0.0659	0.3319	0.5238	0.3175	0.3104	0.6179	0.4777	0.0349	0.1238	0.3371	0.2283	0.3855	0.0118	0.1503
指标 ③	0.4902	0.8791	0.0906	0.1977	0.1298	0.3924	0.2227	0.3690	1.0000	0.0838	0.3554	0.3647	0.2177	0.3446	0.0561	0.1639
地区	湖北	湖南	广东	广西	海南	重庆	四川	贵州	云南	西藏	陕西	甘肃	青海	宁夏	新疆	/
指标 ①	0.1238	0.1420	0.8316	0.3981	0.0975	0.0000	0.2996	0.0933	0.3075	0.9274	0.2487	0.4283	0.6683	1.0000	0.8971	/
指标 ②	0.2287	0.3573	0.0000	0.1967	0.7693	0.8888	0.2073	0.8030	0.3037	1.0000	0.4149	0.8630	0.9662	0.9461	0.7514	/
指标 ③	0.0508	0.1999	0.0059	0.1432	0.1824	0.6094	0.1063	0.7190	0.0000	0.8045	0.2396	0.7562	0.9449	0.9618	0.3497	/

利用式(5), 计算的聚类权如下:

$$\eta_j^k = \begin{bmatrix} 0.6177 & 0.5281 & 0.5117 & 0.3351 \\ 0.2354 & 0.2831 & 0.3146 & 0.3951 \\ 0.1469 & 0.1888 & 0.1737 & 0.2698 \end{bmatrix}$$

③ 聚类评估值 σ_i^k 的确定

利用上述公式对聚类评估值进行计算, 确定的聚类评价系数见表 2.

表 2 聚类评价结果汇总

地点	σ_{il}^*	评价灰类	地点	σ_{il}^*	评价灰类	地点	σ_{il}^*	评价灰类
北京	$\sigma_{13}^* = 0.489$	C	安徽	$\sigma_{12\ 丑}^* = 0.4977$	B	四川	$\sigma_{(23\ 丑)}^* = 0.5786$	A
天津	$\sigma_{24}^* = 0.6367$	D	福建	$\sigma_{(13\ 丑)}^* = 0.4980$	B	贵州	$\sigma_{(24\ 丑)}^* = 0.5397$	D
河北	$\sigma_{31}^* = 0.5867$	A	江西	$\sigma_{(14\ 丑)}^* = 0.5063$	C	云南	$\sigma_{(25\ 丑)}^* = 0.4500$	C
山西	$\sigma_{42}^* = 0.6525$	B	山东	$\sigma_{(15\ 丑)}^* = 0.489$	A	西藏	$\sigma_{(26\ 丑)}^* = 1.0000$	D
内蒙古	$\sigma_{53}^* = 0.5510$	C	河南	$\sigma_{(16\ 丑)}^* = 0.6284$	A	陕西	$\sigma_{(27\ 丑)}^* = 0.5011$	C
辽宁	$\sigma_{63}^* = 0.5139$	C	湖北	$\sigma_{(17\ 丑)}^* = 0.8183$	A	甘肃	$\sigma_{(28\ 丑)}^* = 0.6170$	D
吉林	$\sigma_{72}^* = 0.7240$	B	湖南	$\sigma_{(18\ 丑)}^* = 0.5320$	B	青海	$\sigma_{(29\ 丑)}^* = 0.8272$	D
黑龙江	$\sigma_{83}^* = 0.5795$	C	广东	$\sigma_{(19\ 丑)}^* = 0.6667$	A	宁夏	$\sigma_{(30\ 丑)}^* = 1.0000$	D
上海	$\sigma_{93}^* = 0.5785$	C	广西	$\sigma_{(20\ 丑)}^* = 0.6192$	B	新疆	$\sigma_{(31\ 丑)}^* = 0.5311$	C
江苏	$\sigma_{(10\ 丑)}^* = 0.6004$	A	海南	$\sigma_{(21\ 丑)}^* = 0.4014$	A	/	/	/
浙江	$\sigma_{(11\ 丑)}^* = 0.6597$	C	重庆	$\sigma_{(22\ 丑)}^* = 0.7083$	D	/	/	/

3.2 神经网络的训练

由于原始样本数有限, 考虑到网络的稳定性(神

经元的个数一般小于 10), 第一个隐层单元取 5 个神经元, 第二个隐层取 3 个神经元, 输出层取 4 个神经

元.用 Matlab 程序进行神经网络训练.三层的传递函数分别为“tansig”、“tansig”、“purelin”.采用“trainlm”学习法,学习速率 l_r 取 0.1,训练了 69 步,误差平方和的目标为 0.001,迭代结果满足要求.

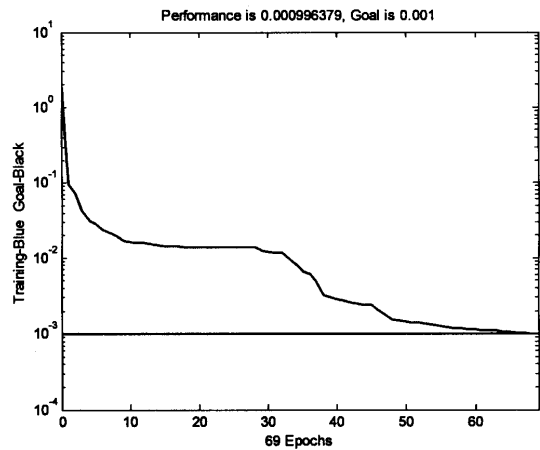


图 3 神经网络训练迭代误差减少图

训练好的网络权重与阈值为:权重为 w_1 、 w_2 、 w_3 , 阈值为 b_1 、 b_2 、 b_3 .

$$w_1 = \begin{bmatrix} 3.3088 & 2.8784 & -1.9212 \\ -0.8172 & -2.2667 & 4.1374 \\ -2.3031 & 0.4912 & 4.1688 \\ -3.1190 & -3.0588 & -9.043 \\ -4.0244 & 2.5966 & -0.0452 \end{bmatrix}$$
$$b_1 = \begin{bmatrix} 4.5270 \\ 0.6720 \\ -1.6468 \\ 1.6468 \\ -1.6584 \end{bmatrix}$$

$$w_2 = \begin{bmatrix} 1.1748 & -0.4644 & -0.5604 \\ 1.2051 & -0.7878 & 1.2683 \\ 0.5796 & -0.6352 & 0.2723 \\ 1.2518 & 0.1342 & -0.5102 \\ 0.4690 & 0.6684 & 0.5935 \\ 0.1369 & 1.3426 \\ -0.2065 & 0.0846 \\ 0.7782 & 1.5205 \\ 0.4776 & -1.2875 \\ 0.8630 & 1.4043 \end{bmatrix}$$
$$b_2 = \begin{bmatrix} -1.9316 \\ 0.9658 \\ 0 \\ 0.9658 \\ 1.9316 \end{bmatrix}$$

$$w_3 = \begin{bmatrix} -0.1342 & 0.0596 & 0.5667 \\ 0.5841 & 0.2811 & 0.3617 \\ 0.1596 & -0.5819 & -0.0778 \\ 0.5207 & 0.2404 & 0.1357 \\ 0.5884 & -0.1693 \\ 0.8816 & -0.3900 \\ 0.2057 & 0.7484 \\ -0.8995 & -0.9700 \end{bmatrix}$$
$$b_3 = \begin{bmatrix} 0.5359 \\ 0.9417 \\ 0.9802 \\ 0.5777 \end{bmatrix}$$

3.3 神经网络仿真

利用上面训练好的神经网络,选取 5 个省份进行仿真,学习的结果为:

表 3 检验样本的神经网络输出

省份	神经网络的实际输出				训练灰类	期望输出				实际灰类
北京	0.0247	0.2697	0.4819	0.2440	C	0.0000	0.2718	0.4890	0.2391	C
天津	-0.0253	0.0063	0.3504	0.6888	D	0.0000	0.0000	0.3112	0.6367	D
河北	0.5863	0.1645	0.2719	-0.0150	A	0.5867	0.1564	0.2569	0.0000	A
山西	0.0219	0.6650	0.3178	-0.0416	B	0.0374	0.6525	0.3101	0.0000	B
内蒙古	0.1449	0.3983	0.5034	-0.0032	C	0.1742	0.4751	0.5510	0.0000	C

从表 3BP 神经网络的训练结果可以看出,期望输出和实际输出存有一定误差,但网络的训练结果的灰类和灰色聚类理论计算的灰类完全一致,可以看出该训练好的神经网络模型有一定的推广性,可用于各地区的道路交通安全评价.

3.4 与传统道路交通安全评价方法优势对比分析

从而可以看出该方法的优势主要体现在:避免了确定指标权重时的主观性,并通过对给定样本模式的学习,获取评价专家的经验、知识、主观判断及

对目标重要性的倾向,当需对样本模式以外的对象系统做出综合评价时,该方法便可再现评价专家的经验、知识和直觉思维,从而实现定性分析与定量分析的有效结合,旨在降低评价过程中的随机性和评价专家主观上的不确定性和认识上的模糊性,从而较好地保证了评价的客观公正性.同时,该方法完全用计算机来实现,提高评价的科学性,具备了较高的问题求解效率.此外,模型具有较强的推广价值,特别是当数据样本更为丰富时,(下转 119 页)

build road bands and bridges along the rivers in the cities ; it specially demonstrate the necessity of avoiding the newbuilding bridges crossing river and Binjiang road to do a harm to the harbor and the water way.

Key words : harbor ; waterway ; Binjiang road ; the bridges crossing river

.....

(上接 104 页)

模型的精度还将有很大的提高空间.

4 结 语

虽然令人满意的预测结果证明了道路交通安全神经网络评价模型的有效性 ,但也存在一些不足 ,主要表现在训练精度较低以及网络不太稳定等问题 .原因可能在于用于网络训练的样本有限 .综上所述 ,这种基于灰色聚类评价理论和 BP 神经网络技术的评价模型 ,吸取了灰色变权聚类评价理论和 BP 神经网络技术的优点 ,使评价结果更具有说服力 .

参考文献 :

[1] 刘小明 ,任福田 .论道路交通安全[M].北京 :人民交通出版社 ,2001 :1-5 ,19-20.

[2] 胡守仁 .神经网络应用技术[M].长沙 :国防科技大学出版社 ,1993 :77-82.

[3] 周 伟 ,张生瑞 ,高行山 ,等 .基于模糊理论和神经网络技术的公路网综合评价方法研究[J].中国公路学报 ,1997 (10) :76-83.

[4] 刘思峰 ,郭天榜 ,党耀国 ,等 .灰色系统理论及应用[M].北京 :科学出版社 ,1991 :78-102.

[5] 裴玉龙 .道路交通事故成因分析及预防对策研究[D].东南大学 ,2002 :63-80.

An evaluation of the road traffic safety based on the grey cluster and neural network

PAN Yan-rong , ZHAI Chang-xu , ZHU Shun-ying

(School of Traffic & Transportation , Chongqing Jiaotong University , Chongqing 400074 , China)

Abstract : On the basis of studying the road traffic safety indexes , this paper bring forward a comprehensive evaluation method based on the grey cluster theory and neural network technology . Then , discuss the normalization methods of attribute value of input and output index of BP neural network . At last , the practical example of this method in comprehensive evaluation in all the provinces is given.

Key words : grey cluster theory ; neural network ; road traffic safety