

# 驾驶员因素与交通事故率的关联性

陈雪梅<sup>1</sup>, 高利<sup>1</sup>, 魏中华<sup>2</sup>, 李潜飞<sup>1</sup>

(1. 北京理工大学机械与车辆工程学院, 北京 100081; 2. 北京工业大学交通工程北京市重点实验室, 北京 100022)

**摘 要:** 驾驶员因素是导致交通事故发生的主要因素之一。综合考虑年龄、性别、累计驾驶时间、生理状况等多个驾驶员特征参数, 通过问卷调查的方式, 利用 BP 神经网络分析技术, 建立具有不同隐含层、神经元个数、作用函数的神经网络结构, 对驾驶员因素与交通事故二者之间的内在关联性进行研究, 确定最优的神经网络结构对事故率进行预测。结果表明, 采用 BP 神经网络对驾驶员交通事故率预测可行, 同时明确了易发交通事故的驾驶员子组。

**关键词:** 驾驶员; 交通事故; 神经网络; 事故率

**中图分类号:** U 491.2

**文献标识码:** A

**文章编号:** 0254-0037(2007)07-0697-05

《世界预防道路交通伤害报告》(以下简称《报告》)指出<sup>[1]</sup>, 交通事故严重威胁着人们的生命财产安全。国内外统计资料表明<sup>[2]</sup>, 80%左右的交通事故是由驾驶员因素造成的。国内学者主要针对某个或几个驾驶员因素, 比如年龄、性别、累计驾驶时间、基本生理状况与交通事故之间的关系进行了研究, 但多个因素对交通事故影响的分析及预测很少。交通安全影响因素分析最常用的方法为问卷调查和数理统计法。通过问卷调查, 建立信息数据库, 预测具有相似特点的驾驶员发生交通事故的概率。该方法的不足之处在于, 很难预测与数据库中的数据具有不同特点的驾驶员发生交通事故的可能性, 而且该方法数据量大、效率低、适应性差。自 20 世纪 90 年代以来, 人工神经网络由于具有高度非线性函数映射功能, 很强的学习、联想和容错功能, 作为一种新的数据分析技术, 正受到越来越多的关注。作者综合考虑多个驾驶员特征参数, 通过问卷调查的方式, 基于改进的 BP 神经网络, 建立了具有不同隐含层、神经元个数、作用函数的神经网络结构, 对驾驶员特征与交通事故率之间的关系进行分析, 确定了最优的神经网络结构对事故率进行预测, 明确易发交通事故的驾驶人, 预测交通事故发生的可能性, 对于改善道路交通安全管理具有重要意义。

## 1 驾驶员特征因素和研究内容

驾驶员的很多特征都会对交通安全产生影响。作者仅采用了与之密切相关的 7 个特征: 年龄、性别、教育程度、累计驾驶时间、日平均行驶里程数、基本的生理状况和心理状况:

### 1.1 年龄

很多交通研究学者对驾驶员年龄与交通事故之间的关系进行了研究。一般来说, 年轻驾驶员的交通事故率较高。25 岁以后, 随着年龄的增加, 交通事故率下降。当驾驶员的年龄大于 60 岁, 交通事故率又随着年龄的增加而增加。交通事故率较高的原因在于通常年轻驾驶员累计驾驶时间不长, 为寻求刺激经常超速行驶, 并且责任心不强。对于老年驾驶员而言, 随着年龄的增加, 对外部交通环境的感知能力开始降低, 手脚变得不太灵活, 对于新的交通环境的适应能力变差, 从而导致交通事故率增大。

### 1.2 性别

男女驾驶员的驾驶行为存在比较明显的差别。Sirkku Laapotti 和 Esko Keskinen 通过大量的调研分

收稿日期: 2006-02-28.

作者简介: 陈雪梅(1978-), 女, 山东平邑人, 讲师。

析,得出了以下很有意义的结论<sup>[3]</sup>:一般来说,男性驾驶员具有竞争性、易怒性、爱开快车、喜欢各种刺激等特点,因此男性驾驶员通常具有更多的交通事故量。同时,男性驾驶员独立性强,具有更多的驾车经验,但是统计资料表明,这些特点并没有降低交通事故量<sup>[3]</sup>。近来有的研究表明,女性驾驶员的轻微交通事故的事故率明显高于男性驾驶员,在致命交通事故中男性驾驶员的事故率却明显高于女性驾驶员<sup>[3]</sup>。国外的另一项研究表明,在出租行业,女性驾驶员的事故率是男性驾驶员的3.5倍<sup>[3]</sup>。总之,因为男女驾驶员的工作领域及参与驾驶的比率不同,所以事故率的高低也不能一概而论。

### 1.3 教育程度

教育程度是影响交通安全的另一个重要因素。具有不同教育程度的驾驶员对某些事物和情境的认知广度和理解深度存在明显的差别。

### 1.4 累计驾驶时间

累计驾驶时间的长短对于驾驶员安全行车具有重要的意义。累计驾驶时间长的驾驶员在实际行车过程中积累了丰富的行车经验,这些经验不仅可以帮助驾驶员对行车环境进行准确的判断,还可以缓解驾驶员遇到突发事件时的紧张情绪,相应地降低了危险情景所具有的危险度。统计数据表明,累计驾驶时间在3a以内的驾驶员具有相对较少的驾驶经验,交通事故率通常很高。一般而言,累计驾驶时间在3a以上的驾驶员的交通事故率会随着累计驾驶时间的增加而降低。

### 1.5 日平均行驶里程数

一般而言,行驶里程数越大,行车时间越长,驾驶员越容易疲劳,对事物的认知能力下降。行驶里程数与驾驶时间的比值就是驾驶员通常采用的行驶车速。每年交通部、公安部的事故资料显示,高速行驶严重威胁道路交通安全,因此,作者考虑了日平均行驶里程来弥补单纯考虑累计驾驶时间的不足。仔细观察会注意到,有些驾驶员性格比较鲁莽,行驶车速很快,虽然累计行驶时间不长,但日平均行驶里程很大,很容易发生事故。对该指标进行研究很有价值。

### 1.6 驾驶员基本的生理状况

驾驶员行车时的生理状况对行车安全有着重要的影响。日本的 Toshiyuki Yaguchi 和 Hazime Inagaki 等人<sup>[4,5]</sup>在研究中指出,疲劳驾车是导致交通事故发生的重要原因。疾病和药物对驾驶行为和交通安全的影响也不容忽视。人体机能受到疾病因素的直接影响或间接侵入而生病时,人的协调能力降低,导致行车不利因素直线上升<sup>[6]</sup>。有些药物直接作用于中枢神经系统或间接影响脑功能,使人体各部位协调失控、反应迟钝。

### 1.7 驾驶员的心理状况

过去几十年,国内外的专家学者对驾驶员生理因素对交通安全的影响进行了大量的研究工作。驾驶员行为不仅受到自身生理条件因素的影响,还受到驾驶员心理因素的影响。在驾驶员信息感知、判断、反应过程中驾驶员的心理反应起着重要的作用<sup>[7]</sup>。加强对驾驶员心理的研究有助于对驾驶机理的深层理解及掌握,有利于从根本上降低驾驶员的驾驶错误,改善交通安全状况。

## 2 研究方法

### 2.1 研究样本

通过问卷调查的方式对13个城市的1723名驾驶员进行了问卷调查。依据问卷调查的结果将驾驶员分为127个子组。为了统计的准确和快速,问卷调查结果采用编码的处理方式,参见表1。每个输入变量

的每个分组分别用 1~4 之间的数字表示(比如性别中男性用 1 表示, 女性用 2 表示). 子组按照性别、年龄、累计驾驶时间、教育程度、日平均行驶里程、生理状况、心理状况的顺序进行编码. 比如, 一个男性(性别编码为 1)驾驶员、25 岁(年龄编码为 1)、驾龄为 4 a(累计驾驶时间编码为 2)、具有高中文化程度(教育程度编码为 3), 他每天的行驶里程为 70 km(行驶里程编码为 2)、生理状况不好(编码为 1)、心理状况好(编码为 2). 该驾驶员所归属的子组编号为 1123212.

表 1 驾驶员特征子组编号  
Table 1 Subgroups with different characteristics

驾驶员特征 子组编号	驾驶员特征						
	性别	年龄	累计驾驶时间/a	教育程度	日平均行驶里程/km	生理状况	心理状况
1	男	<28	0~3	小学及以下	<20	不好	不好
2	女	28~60	3~8	初中	20~150	好	好
3		>60	8~11	高中及以上	>150		
4			>11				

### 2.2 神经网络的选择

驾驶员行为受各种因素的影响, 具有不同特征的驾驶员对道路交通安全状况的影响是复杂和非线性的, 很难用精确的驾驶员特征与交通事故间的关系模型研究各因素对交通安全的影响作用. 人工神经网络由于具有高度非线性函数映射功能<sup>[8]</sup>, 作者利用改进的 BP 神经网络对驾驶员特征和交通事故率之间的关系进行研究.

BP 网络是一种误差后向传播网络, 一般采用 3 层网络型式, 其作用函数为非线性的 Tansig 或 logsig, 训练样本采取批处理方式:  $(x(k), y(k), k = 1, 2, \dots, n)$ ,  $N$  为样本总数, 对于任何一个训练样本, 第  $j$  个神经元的输入  $a_{pj} = \sum_0^N W_{ji} O_{pi}$ , 输出  $O_{pj} = f(a_{pj})$ , 误差  $E = \sum E_p = \sum_j (d_{pj} - O_{pj})^2$ , 权值调整公式  $\Delta W = (J^T J + UI)^{-1} J^T e$ , 其中  $J$  为误差权值的微分雅可比矩阵<sup>[8]</sup>.

BP 算法存在的一个问题是收敛速度慢. 在 BP 算法中, 连接权的调整决定于梯度和学习速率, 而且学习速率是不变的. 实际上学习速率对收敛速度的影响很大, 因此需要对它进行在线调整, 以提高收敛速度. BP 网络的另一个问题是容易陷入局部最小点而达不到预期的效果<sup>[8]</sup>. 本研究利用动量法来解决 BP 网络存在的易陷入局部最小点的问题. 使网络能跳出局部最小点, 在一定程度上加快了网络的收敛速度. 所谓动量法就是在原有权值修改公式的基础上, 再加上一个动量项, 局部最小点的跳出可以利用权值修改的惯性来实现. 权值的修改可表示为

$$\Delta W_{jk}(n) = -\eta(1 - mc)\partial E(n)/\partial W_{jk} + mc\Delta W_{jk}(n - 1)$$

式中  $mc\Delta W_{jk}(n - 1)$  即为动量项, 其中称  $mc$  为动量系数,  $0 < mc < 1$ .

利用神经网络对驾驶员特征与交通事故率之间的关系进行研究, 首先需要明确神经网络的输入和输出. 依据相似的特征把所有的驾驶员分成了不同的子组, 在本研究中把驾驶员所具有的 7 个特征作为神经网络的 7 个输入, 每个子组的故事率作为输出. 每个子组的故事率用该组中发生事故的人数与该组的总人数之比来表示. 驾驶员分为 127 个子组, 其中 100 组用作神经网络的训练数据, 27 组用作神经网络的测试数据.

### 3 驾驶员因素与交通事故率关联性研究的结果

驾驶员特征与交通事故率关系研究采用的 BP 神经网络结构如图 1 所示. 该网络由输入层、输出层和隐含层组成, 网络之间通过神经元(节点)顺序单向联接. 输入参数包括: 驾驶员子组编号、性别、年龄、累

积驾驶时间、教育程度、日平均行驶里程、生理状况、心理状况;输出参数为交通事故率. 国内外的文献指出,不同的隐层数、每层的神经元个数以及作用函数会影响神经网络的学习效果和速度,该研究中的神经网络采用不同的隐层数、神经元个数以及作用函数对驾驶员的多种特征、交通事故率进行研究,目的是找到相对最佳的一个神经网络结构,并用该神经网络预测具有一定特征的驾驶员发生交通事故的概率.

利用图1所示神经网络结构,确定具有不同特征的驾驶员与交通事故率之间的关系,并且找到了相对最优的神经网络,对驾驶员发生交通事故的可能性进行预测,并利用 MATLAB 的神经网络工具箱对数据进行分析处理,结果参见表2,图2.

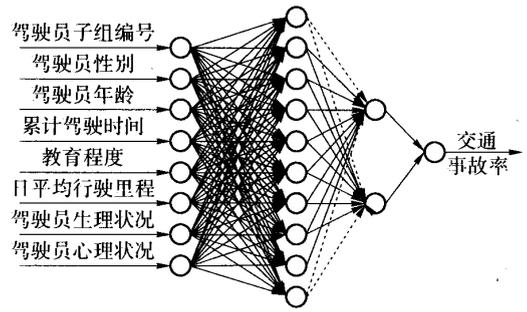
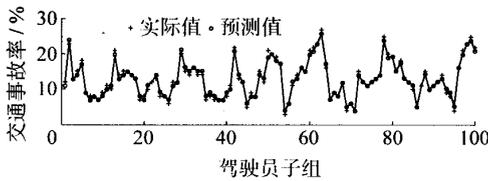


图1 驾驶员因素与交通事故率关联性研究神经网络结构  
Fig.1 Neural Network for research on the correlation between driver's factors and traffic accident rate

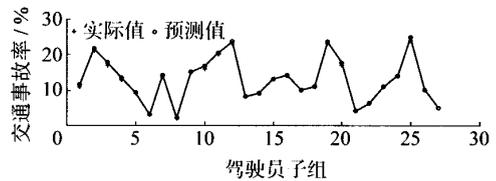
表2 不同神经网络结构的训练结果

Table 2 Neural network with different hidden layer, neuron number and transfer function

神经网络模型	隐含层数	隐含层神经元个数	作用函数	相对误差/%
ANN1	1	12	Tansig	7.3
ANN2	1	18	Tansig	6.9
ANN3	2	12-2	Tansig-logsig	5.1
ANN4	2	18-2	Tansig-logsig	3.3
ANN5	2	18-2	Tansig-purelin	3.7



(a) 神经网络训练结果



(b) 神经网络测试

图2 驾驶员因素与交通事故率关联性研究神经网络结果

Fig.2 Neural network results for the correlation between driver's factors and traffic accident rate

从表2可以看出,当神经网络具有相同隐含层时,隐含层的神经元个数越多,预测值与真实值之间的相对误差越小.从表2的结果还可以得到,神经网络ANN4的训练效果最好.本研究采用神经网络4对测试数据进行测试.图2(a)给出了具有2个隐含层、作用函数分别为Tansig-logsig的神经网络4的训练结果,图2(b)给出了利用选定的最优神经网络模型对27组测试数据进行测试后驾驶员事故率的真实值、神经网络预测值及二者之间的误差.

从图2(b)可以看出,经过训练的神经网络可以很好地预测驾驶员的交通事故率.对以上所有的数据进行统计,预测准确度的最小值为95.7%,最大值为98.7%.但是应该注意到本研究对每个参数的划分很粗糙,这样导致对驾驶员交通事故率的预测有一定偏差.下一步对其进行更加合理的划分,选用更加有效的神经网络结构,可以获得更准确的结果.

仔细分析图2(a)、(b)可以得出,发生事故最多的子组为第63,85,99和125组,它们分别对应的子组编号为1112212,1111221,1211112,2112121.分析表明,男性、年轻、累计驾驶时间短的驾驶员在生理状况或心理状况不良的情况下易发交通事故.

## 4 结论

1) 综合考虑包括年龄、性别、累计驾驶时间、教育背景、日平均行驶里程数、生理状况、心理状况等多个驾驶员特征参数, 利用改进的 BP 神经网络分析, 把神经网络技术用于驾驶员因素、交通事故率内在关联性研究是有效的, 预测的准确率在 95% 以上。

2) 明确易发交通事故的驾驶员子组的特点是: 男性、年轻、累计驾驶时间短、生理状况或心理状况不良。

### 参考文献:

- [1] 世界卫生组织, 世界银行. 世界预防道路交通伤害报告(WHO 翻译版)[M]. 刘光远, 译. 北京: 人民卫生出版社, 2004.
- [2] 巴布可夫. 道路条件与交通安全[M]. 景天然, 译. 上海: 同济大学出版社, 1990.
- [3] SIRKKU Laapotti, ESKO Keskinen. Differences in fatal loss-of-control accidents between young male and female drivers[J]. *Accident Analysis and Prevent*, 1998, 30(4): 435-442.
- [4] SAROJ K L, ASHLEY C. A critical review of the psychophysiology of driver fatigue[J]. *Biological Psychology*, 2001, 55(3): 173-194.
- [5] IVAN D B. Prospects for technological countermeasures against driver fatigue[J]. *Accident Analysis & Prevention*, 1997, 29(4): 525-531.
- [6] 张宏, 王力, 李蔚. 机动车驾驶员身体素质对交通事故的影响[J]. *职业与健康*, 2001, 17(8): 9-11.  
ZHANG Hong, WANG Li, LI Wei. Driver physical quality effect on traffic accident[J]. *Occupation and Health*, 2001, 17(8): 9-11. (in Chinese)
- [7] LUCAS J. Drivers' psychological and physical reactions after motor vehicle accidents[J]. *Transportation Research Part F: Traffic Psychology and Behaviour*, 2003, 6(3): 135-145.
- [8] 周继成. 人工神经网络——第六代计算机的实现[M]. 北京: 科学普及出版社, 1993.

## Correlation Between Driver's Factors and Traffic Accident Rate

CHEN Xue-mei<sup>1</sup>, GAO Li<sup>1</sup>, WEI Zhong-hua<sup>2</sup>, LI Qian-fei<sup>1</sup>

(1. School of Mechanical and vehicular Engineering, Beijing Institute of Technology, Beijing 100081, China;

2. Traffic Engineering Key Lab of Beijing, Beijing University of Technology, Beijing 100022, China)

**Abstract:** Driver's factors are very important for the traffic safety. Questionnaire and BP neural network are used, the neural network structure with different hidden layer, neuron number and transfer function is established considering many parameters, such as age, gender, accumulative driving time, physiological conditions, etc. The optimal neural network structure is obtained to predict the traffic accident rate. The results show that the neural network is available for the prediction of accident rate and the driver subgroups with accident proneness are identified.

**Key words:** driver; traffic accident; neural network; accident rate