

# 居民出行产生量 BP 神经网络预测方法

冯树民, 慈玉生

(哈尔滨工业大学 交通科学与工程学院, 哈尔滨 150090, zlyfsm2000@sina.com)

**摘要:** 居民出行产生量预测是交通需求分析的重要内容之一, 预测结果是确定各类城市交通设施发展规模及布局规划的重要依据. 通过分析神经网络的作用机理和居民出行产生量的影响因素, 建立了居民出行产生量预测的四层 BP 神经网络模型, 以土地利用作为输入神经元, 以交通区居民出行产生量作为输出单元, 以赣州市城市综合交通规划交通调查数据对模型进行了标定与检验, 并与出行次数法和回归分析法进行了比较, 结果表明 BP 神经网络模型具有较高的预测精度.

**关键词:** 出行产生; BP 神经网络; 预测方法

**中图分类号:** U491.3      **文献标志码:** A      **文章编号:** 0367-6234(2010)10-1624-04

## A forecast method for trip production based on BP neural network

FENG Shu-min, CI Yu-sheng

(School of Transportation Science and Technology, Harbin Institute of Technology, Harbin 150090, China, zlyfsm2000@sina.com)

**Abstract:** Trip production forecast is one of key components of traffic demand analysis, which directly determines the scale and layout of different urban traffic facilities. The mechanism of artificial neural network (ANN) and influential factors of trip production were analyzed. therefore, a four-layer back-propagation neural network (BP neural network) model was set up to forecast the trip production, in which the input neurons are land-uses of different traffic zones and the output is trip production. Meanwhile, the model was calibrated and testified with traffic survey data from the urban integrative transportation planning of Ganzhou city. Furthermore, the results were compared with those obtained from trip production rate method and multiple linear regression method, It is showed that forecast precision of the BP neural network is relatively high.

**Key words:** trip production; BP neural network; forecast method

居民出行生成预测是城市交通规划交通需求预测“四阶段”法的重要内容. 居民出行生成预测作为交通需求的前提和基础, 预测结果直接影响到其他各阶段的预测, 从而影响交通设施规划, 在交通需求预测中有举足轻重的作用. 居民出行生成预测应考虑到交通发生源的空间布局关系和总体的利用情况, 一般按区域进行预测.

目前, 比较常见的预测方法为出行率法、回归分析法和交叉分类法, 研究人员也开展了相关方法的研究. 最早开展居民出行生成交叉分类分析、

回归分析的是 Ricardo<sup>[1]</sup>、Hayfield 等<sup>[2]</sup>, 并以实际调查数据进行了验证; 之后, Michael 等<sup>[3]</sup>研究了出行目的分类的居民出行生成预测技术, Marchal<sup>[4]</sup>研究了一种基于交通与土地利用仿真的出行生成方法, John 等<sup>[5]</sup>研究了私人信息在交通生成预测的应用, Yao 等<sup>[6]</sup>研究了目的地吸引的居民出行生成量预测模型, Yang 等<sup>[7]</sup>研究了基于离散活动行为的通勤者出行生成模型; 近年, 复杂网络理论、多重分类分析技术及最优化等方法与居民出行生成预测相结合也开展了较多研究, 如 Lu 等<sup>[8]</sup>研究了基于复杂网络理论的居民出行发生量与吸引量的特征分析, Cristian 等<sup>[9]</sup>建立了基于多重分类分析方法的居民出行发生量模型, Vrtic 等<sup>[10]</sup>研究了居民出行生成、分布及方式选择的二

收稿日期: 2009-10-15.

基金项目: 黑龙江省教育厅资助项目(11541295).

作者简介: 冯树民(1973—), 男, 博士, 副教授.

维约束的非集计模型,Zhou 等<sup>[11]</sup>研究了基于最优化理论及变分不等式的交通需求综合模型。

影响居民出行产生的影响因素众多且较为复杂,难以用简单的数学模型来精确预测居民出行产生量。基于这一考虑,本文建立了基于 BP 神经网络的居民出行产生量的预测模型。通过对交通小区土地利用与出行产生率之间的关系研究,发现居民出行产生率与土地利用之间具有较强的非线性关系。

## 1 居民出行产生量的影响因素

数据资料表明,居民出行产生量的主要影响因素为城市居民的社会经济特性、土地利用性质及交通区位等。

### 1.1 城市居民的社会经济特性

城市居民的社会经济特性包含交通区人口、劳动力情况、就业就学情况等特征及居民家庭收入、文化背景等等。一般认为交通区人口代表城市居民的社会经济特性。

### 1.2 土地利用特性

城市交通与土地利用相互联系、相互制约。城市土地利用对交通影响主要表现在城市土地利用对交通出行的影响,过强的土地利用常导致较高的交通需求。城市居民的出行方式、交通量基本上是土地利用空间分布的函数。土地开发结果是产生以该区为起点的新出行或吸引其他区的新出行,土地利用性质是居民出行产生量的主要影响因素。通过综合分析,本文选择商业金融用地、行政办公用地、文体医教用、及居住用地及其他用地来表征城市土地利用情况。

### 1.3 影响因素确定

居民出行生成预测包含产生预测和吸引预测两部分。一般认为,居民出行产生量与区域人口息息相关,而居民出行吸引与土地利用关系更强。

笔者通过多年的研究积累,发现居民出行产生量与土地利用具有较强的非线性关系,主要表现在区域的土地利用类型决定了该区域居民出行产生率。本文重点利用 BP 神经网络对交通小区的居民出行产生率进行研究,然后通过交通小区的居民出行产生率来预测居民出行产生量。

## 2 BP 神经网络模型

BP 神经网络已经广泛应用于控制、通信、交通等行业,BP 神经网络的优点是只要有足够的隐层和隐节点,BP 网络就可以逼近任意的非线性映射,体现了人工神经网络中最精华的部分。

### 2.1 输入、输出层设计

通过对居民出行产生影响因素的分析,确定本 BP 网络的输入层神经元个数为 5;输出层含有 1 个神经元,表示交通区的居民出行产生量。

### 2.2 结构设计

一般来说,一个 3 层 BP 网络,即单个隐层 BP 网络,即可完成任意  $n$  维输入到  $m$  维输出的映射。本文采用含有 2 个隐层的 BP 网络对来预测居民出行产生量,增加一个隐含层加强隐含层与外层连接性,该网络能够加快神经网络的学习、训练速度。

隐层神经元结合经验公式和试算法来确定,通过训练结果中误差的大小,最终确定隐层单元的个数。

### 2.3 学习算法

本 BP 网络隐含层转移函数选取双曲正切 S 型和反对数 S 型函数,输出层采用线性转移函数,Matlab 中分别利用 tansig、logsig 和 purelin 函数来实现。本文省略 BP 网络的训练算法、误差计算、权重计算等计算公式,可参考相关参考书。

同时,为了加快训练速度,避免陷入局部最小值和改善其他能力。学习算法采用带动量的批处理梯度下降法,通过 traingdm 命令实现。

### 2.4 数据归一

为了加快训练网络的收敛性,同时也是对采集的各数据单位不一致进行规划处理,一般对输入、输出数据先进行归一处理,然后对模拟的结果进行反归一处理即可得到预测结果。

在本 BP 网络中,如果不进行数据归一化处理,会出现大数“吃”掉小数的现象,从而得不到预测结果,也即不收敛,通过 mapminmax 命令实现。

### 2.5 BP 网络实现

本 BP 网络通过 MatlabR2008 实现,程序分为 4 个部分,分别为初始输入、输出,网络及参数部分,训练部分及测试部分。

## 3 模型验证

### 3.1 数据来源

数据来源于赣州市城市综合交通规划调查,如表 1 所示。输入的 19 组数据中,取奇数组为训练样本,取偶数组为测试样本。

本 BP 神经网络输入  $P_1$ 、 $P_2$ 、 $P_3$ 、 $P_4$  及  $P_5$  分别表示商业金融用地、行政办公用地、文体医教用地、居住用地及其他用地占总用地的比例;输出为交通小区出行率,用  $T$  表示。

表 1 输入输出数据

分区	P1	P2	P3	P4	P5	T	人口	产生量
1	0.03	0.07	0.06	0.29	0.54	2.82	7 386	20 815
2	0.33	0.06	0.11	0.30	0.21	1.86	13 168	24 544
3	0.07	0.08	0.00	0.80	0.05	2.40	17 861	42 864
4	0.06	0.01	0.38	0.49	0.07	2.24	119 754	268 375
5	0.16	0.06	0.03	0.58	0.18	2.42	13 595	32 929
6	0.00	0.00	0.01	0.71	0.27	2.18	12 979	28 284
7	0.06	0.02	0.01	0.44	0.48	3.55	5 742	20 392
8	0.00	0.07	0.00	0.53	0.39	2.29	5 146	11 796
9	0.06	0.07	0.35	0.35	0.17	3.34	29 458	98 383
10	0.11	0.00	0.05	0.70	0.14	1.73	11 357	19 593
11	0.17	0.03	0.00	0.68	0.12	2.81	10 017	28 101
12	0.11	0.15	0.17	0.51	0.06	1.85	15 162	27 986
13	0.20	0.05	0.08	0.63	0.04	2.40	13 816	33 152
14	0.08	0.06	0.08	0.74	0.04	2.29	7 627	17 479
15	0.42	0.03	0.14	0.32	0.09	3.02	11 902	35 904
16	0.09	0.04	0.02	0.60	0.25	1.98	17 631	34 897
17	0.17	0.05	0.08	0.58	0.12	2.77	20 077	55 698
18	0.08	0.01	0.03	0.21	0.67	1.95	14 059	27 368
19	0.17	0.00	0.00	0.61	0.22	3.20	12 984	41 568

### 3.2 模型比较

#### 3.2.1 出行次数法

通过对所有数据的统计,可知人口总数为 359 721 人,出行产生量为 870 128 人次,平均出行次数为 2.42 次。

以此平均出行次数计算偶数组的出行产生量,见表 2。

#### 3.2.2 回归分析法

以所有组数据的人口数作为变量,以出行产生量作为函数,得到一元回归模型,可以看出相关性较强。

$$y = 2.250 5x + 3 189, R^2 = 0.975 3.$$

以该函数来计算偶数组的出行产生量,见表 2。

#### 3.2.3 BP 神经网络法

以奇数组数据进行训练,以偶数组数据进行测试. 通过以上训练、测试,得到测试结果,训练误差曲线见图 1, BP 网络测试结果见图 2 和表 2。

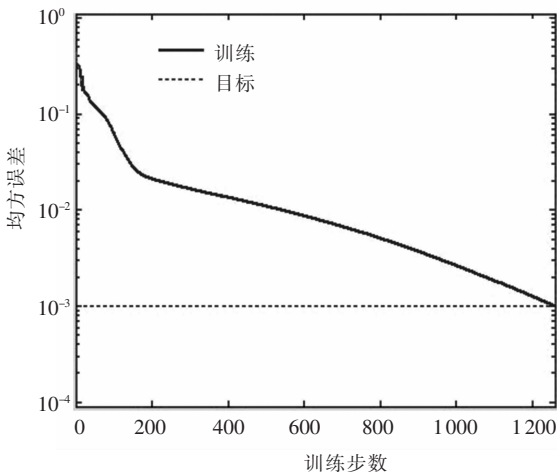


图 1 训练误差曲线

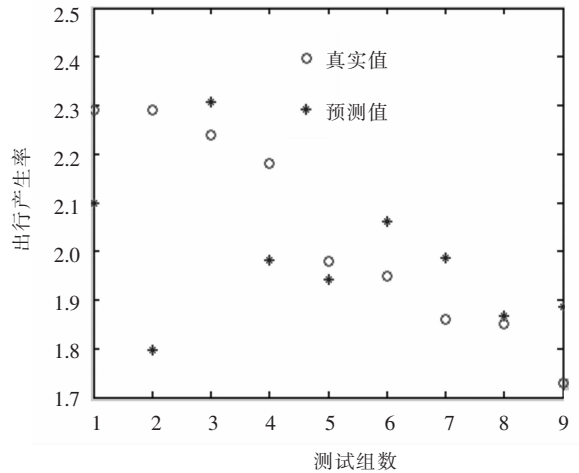


图 2 测试结果

### 3.3 结果分析

通过对 3 种方法计算结果比较可知：

在进行多元线性回归分析时,也是利用奇数组数据进行回归,相关系数  $R^2$  达到 0.99,表明出行产生量与交通区人口具有较强的相关性. 即便如此,利用该回归模型进行偶数组数据的出行产生量预测,预测结果仍不甚理想. 而利用出行次数法进行预测时,由于只是与人口有关,因此,预测结果误差也最大。

结果表明,该 BP 网络训练至 1 259 步时收敛,且根据误差大小,隐含层神经元个数为 8 时,误差最小,网络的实际输出与期望输出最为接近。

由 BP 网络的输出结果可以看出,实际输出与目标输出之间的差距不大,平均误差为 7%,较出行次数法 20% 和回归分析法的 25% 更为接近实际结果。

为了提高该网络预测结果的精度,在实际交通需求预测时,应对交通小区进行细化,从而能够

更为实际体现出行产生量与人口、土地利用之间的非线性关系。

表 2 预测结果比较

序号	出行产生	回归方法		出行次数法		BP 网络	
		结果	误差	结果	误差	结果	误差
1	11 796	14 768	0.25	12 453	0.06	10 807	0.08
2	17 479	20 350	0.16	18 457	0.06	13 729	0.21
3	268 375	272 636	0.02	289 805	0.08	275 434	0.03
4	28 284	32 392	0.15	31 409	0.11	25 698	0.09
5	34 897	42 859	0.23	42 667	0.22	34 204	0.02
6	27 368	34 822	0.27	34 023	0.24	28 962	0.06
7	24 544	32 817	0.34	31 867	0.30	26 073	0.06
8	27 986	37 304	0.33	36 692	0.31	28 201	0.01
9	19 593	28 742	0.47	27 484	0.40	21 351	0.09
平均	51 147	57 410	0.25	58 317	0.20	51 607	0.07

## 4 结 语

通过对居民出行产生量影响因素的分析可知,居民出行产生量与交通区人口、土地利用具有较强的非线性关系。

本文建立了基于人口、土地利用的四层 BP 神经网络模型,并通过赣州市交通调查数据进行了训练和测试,得到了 BP 神经网络的预测出行产生量结果.结果表明,利用 BP 神经网络进行出行产生量预测较出行次数法和回归分析法更为可靠,可以应用于规划目标年的居民出行产生量的预测。

## 参考文献:

- [1] RICARDO D, WILLIAM M. An empirical comparison of disaggregate category and regression trip generation analysis techniques [J]. Transportation, 1977, 6: 287 - 307.
- [2] HAYFIELD C P, STOKER R B. The geographical stability of a typical trip production model: applications of national and local data in four urban areas [J]. Transportation, 1978, 7: 211 - 224.
- [3] MICHAEL D A, JUSTIN P O. Evaluation of two trip generation techniques for small area travel models [J]. Journal of Urban Planning and Development, 2002, 128 (2): 77 - 88.
- [4] MARCHAL F. A trip generation method for time - dependent large - scale simulations of transport and land - use [J]. Networks and Spatial Economics, 2005, 5: 179 - 192.

- [5] JOHN S M, LESTER A H. Assessing the utility of private information in transportation planning studies: a case study of trip generation analysis [J]. Socio-Economic Planning Sciences, 2006, 40: 94 - 118.
- [6] YAO Liya, GUAN Hongzhi. Trip generation model based on destination attractiveness [J]. Tsinghua Science and Technology, 2008, 13(5): 632 - 635.
- [7] YANG Min, WANG Wei. Discrete activity-based trip generation model for commuters [C] // Proceedings of International Conference on Traffic and Transportation Studies. Nanning: [s. n.], 2008: 188 - 196.
- [8] LU Huapu, SHI Ye. Characteristics of trip generation and attraction based on complex network theory [C] // Proceedings of International Conference on Transportation Engineering. Chengdu: [s. n.], 2007: 1334 - 1340.
- [9] CRISTIAN A G, ALAN T. Multiple classification analysis in trip production models [J]. Transport Police, 2007, 14: 514 - 522.
- [10] VRTIC M, FROHLICH P. Two-dimensionally constrained disaggregate trip generation, distribution and mode choice model: theory and application for a swiss national model [J]. Transportation Research Part A, 2007, 41: 857 - 873.
- [11] ZHOU Zhong, ANTHONY C, WONG S C. Alternative formulation of a combined trip generation, trip distribution, modal split, and trip assignment model [J]. European Journal of Operational Research, 2009, 198: 129 - 138.

(编辑 赵丽莹)