

基于 MLP 的高速公路服务区合理规模测算模型

张艳秋^{1,2}, 陈坚^{2,3}, 杨林²

(1.招商局生态环保科技有限公司,重庆市 400067;2.重庆交通大学 交通运输学院;
3.重庆大学 建筑城规学院)

摘要:为准确测算高速公路服务区合理建设规模,提高高速公路服务质量和经营效益,结合出行者消费行为及其他交通参数,提出基于 MLP(Multi-Layer Perception)神经网络的服务区合理规模测算模型,采用 BP 算法进行训练。结果表明:模型评价参数 $rerr$ 、 $arerr$ 均在 5%以内,EC 为 0.984 5,该模型预测精度及可靠性较好。

关键词:高速公路服务区;合理规模;消费行为;MLP 神经网络

改造发展提供参考依据。

1 前言

服务区作为高速公路上的“驿站”和“窗口”,是高速公路安全、便捷、舒适出行的保障,发挥着重要作用。在中国高速公路连线成网加快发展的新时期,服务区作为高速公路系统的重要组成部分也逐渐发展起来,截至 2018 年底,中国高速公路服务区数量达到 2 850 对左右。但中国服务区在发展和建设中存在诸多问题,普遍存在服务区征地面积及建设规模与主线交通量不相符的现象,缺乏对服务区进行系统定量分析预测和建设规模可行性研究。

国内外对高速公路服务区的研究宏观上主要为服务区功能定位、规划布局,微观层面主要为服务区服务水平评价、驶入率预测等,近年来学者们对服务区规模相继展开了研究。现有对服务区规模的研究大多从服务区物理设施及驶入率入手,通过驶入率计算服务区的停车场面积,再定性分析停车场面积与服务区面积之间关系,参数取值经验性居多,对服务区合理规模与其影响因素之间定量关系的研究较少,缺少反映客观规律的模型;且现有对服务区消费行为的研究较少,服务区规模研究中也尚未考虑出行者消费行为因素,消费行为对规模的影响未得到充分重视。该文在分析服务区合理规模影响因素时充分考虑出行者消费行为及其他交通参数,构建基于 MLP 的合理规模测算模型,并用实例进行验证。可为当前服务区规划设计、升级

2 合理规模界定及影响因素

2.1 合理规模界定

服务区根据其所处地理位置、周边资源开发、吸引人群的不同,其服务区功能设置、规模大小应有所区别,服务区规模的设置与服务区经营效益、对客流的吸引力及服务水平紧密相关。服务区合理规模即在一定经济发展水平下,服务区用地面积及功能设置与主线交通量、驶入率、服务区功能定位、出行者消费需求相匹配,能最大程度上满足出行者的需求。

2.2 合理规模影响因素

根据高速公路服务区现有配置可知其总体规模由以车为服务对象的停车场、加油站、车辆维修中心,以人为对象的房建区及道路规模,以及绿化及其他规模构成。房建区及道路规模主要包括厕所、餐厅、商店、住宿休息区、附属设施等,该部分是服务区主体构成部分,是服务区运营的核心。

高速公路服务区规模的影响因素众多,包括服务区地理位置、功能定位、主线断面交通量、驶入率、车型结构、出行者消费行为等,只有明确影响服务区合理规模的主要因素,才能进行合理规模研究。根据相关研究综合分析拟定主线断面交通量、驶入率、高峰率、周转率、车型结构、消费行为作为影响合理规模的 6 大主要因素。各影响因素含义如下:

收稿日期:2019-10-06

基金项目:国家社会科学基金西部项目(编号:17XGL009)

作者简介:张艳秋,女,硕士研究生, E-mail: 1373585610@qq.com

- (1) 主线断面交通量:指一定时期内通过服务区截面的交通量(双侧),单位:(pcu/d)。
- (2) 驶入率:驶入到服务区的交通量与主线交通量的比值。
- (3) 高峰率:指在高峰时段停留的车辆数(pcu/h)与每天停放车辆数(pcu/d)的比值。
- (4) 周转率:指服务区的一个停车位在1h内被使用的次数。
- (5) 车型结构:驶入服务区大型车与小型车的比值。
- (6) 消费行为:包括同行人数、消费水平、消费需求等,消费水平和消费需求分别用实际消费金额(元)和服务区平均消费需求项(项)量化表示。

各因素与服务区总体规模影响关系见图1。

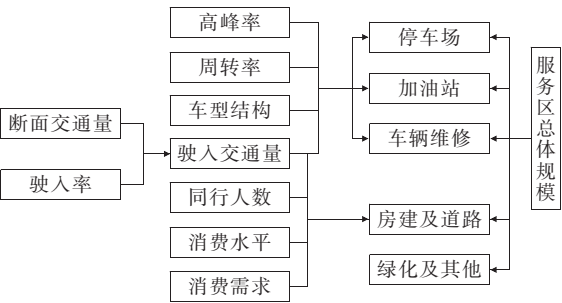


图1 因素与服务区总体规模影响关系

3 深度学习相关理论

3.1 深度学习概述

深度学习是新兴的机器学习研究领域,是当前人工智能的研究热点,旨在研究如何从数据中自动地提取多层特征表示,其核心思想是通过数据驱动方式,采用一系列的非线性变换,从原始数据中提取由低层到高层、由具体到抽象、由一般到特定语义的特征。通过从多源异构数据中进行自动特征学习,从而将不同数据映射到一个相同的隐空间,能够获得数据的统一表征。目前已在图像识别、语音识别、自然语言处理、车牌定位等领域展现出了巨大的优势。

深度学习常用模型有多层感知机(MLP)、受限玻尔兹曼机(RBM)、深度置信网络(DBN)、卷积神经网络(CNN)、循环神经网络(RNN)、自动编码器(AE)等。常用软件工具平台主要有 TensorFlow、Caffe 等,操作语言主要为 C++、Python、Matlab。深度学习常见模型及特点见表1。

3.2 MLP 神经网络

(1) 模型结构及预测原理

MLP(Multi-Layer Perception—多层感知机)属于前馈深度神经网络模型,网络结构由输入层、隐含

表1 深度学习常见模型及特点

深度模型	网络结构	训练算法	模型特点及解决问题
BP 反向传播	3 层	链式积分法	从简单神经网络到复杂神经网络的推广
多层感知机(MLP)	多层	梯度下降算法、BP	解决简单分类及预测问题
受限玻尔兹曼机(RBM)	多层	对比散度法	容易求得 BM 的概率分布;具有无监督学习能力
深度置信网络(DBN)	多层有/无向边连接	贪心逐层算法、BP	以无监督学习到的参数作为有监督学习的初始值,从而解决了 BP 的问题
卷积神经网络(CNN)	多层无向边局部连接	梯度下降算法、BP	包含卷积层和子采样层;可以接受 2D 结构的输入
简单循环网络(SRN)	3 层	梯度下降算法、BPTT	时间维度上的深层结构;上一时刻的输出是下一时刻的输入
循环神经网络(RNN)	多层	梯度下降算法、BPTT	多层的时间维度上的深层结构;能够处理序列数据
自动编码器(AE)	3 层	贪心逐层算法	通过编码器和解码器工作完成训练;通过损失函数最小化求出网络的参数

层、输出层组成,隐含层根据需要由一层或多层组成。MLP 可被看作一个有向图,由多个节点层所组成,每一层都全连接到下一层,除了输入节点,每个节点都是一个带有非线性激活函数的神经元,常采用误差反向

传播算法(BP 算法)进行训练。

模型预测的主要原理为根据历史数据输入选择的模型中进行训练,得到一个规则,当输入新的数据时这个训练好的规则可以预测得到结果,基本原理见图2。

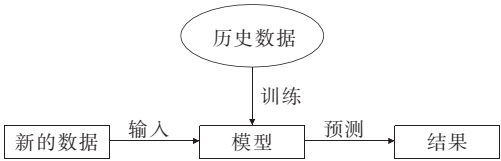


图 2 深度学习神经网络预测基本原理

(2) MLP 模型适应性

MLP 是感知机的推广,克服了感知机不能对线性不可分数据进行识别的弱点,能够逼近任意复杂的非线性函数关系,具有很强的自适应、自学习及容错能力,适合进行非线性函数拟合,可用于解决简单分类及预测问题,目前已在区域物流需求、蛋白质变性温度、短时交通流、广告点击率等方面进行了预测研究,且取得了较好的效果,服务区合理规模与各影响因素之间

为复杂的非线性关系,MLP 通过误差反传能够发现各因素与规模之间的映射关系,故选取 MLP 进行服务区合理规模预测。

4 基于 MLP 的合理规模测算模型

(1) 模型结构及变量选取

该文选取包含一个输入层、一个隐含层和一个输出层的 3 层 MLP 模型,根据服务区合理规模影响因素分析选取因素之间相关性较弱的作为输入层变量,综合分析选取主线断面交通量、驶入率、高峰率、周转率、车型结构及消费行为中较为重要的同行人数、消费水平和消费需求 8 个因素作为 MLP 神经网络输入层变量,各变量及其量化值见表 2。

表 2 输入层变量及代表量化值

输入变量	变量描述	量化值
主线断面交通量 X_1	服务区高速公路日平均断面流量	日实际交通量值
驶入率 X_2	驶入交通量与主线断面交通量比值	实际计算值
高峰率 X_3	高峰小时交通量与日平均交通量比值	实际计算值
周转率 X_4	服务区的一个停车位在 1 h 内被使用的次数	实际计算值
车型结构 X_5	驶入服务区大型车和小型车的比值	实际计算值
同行人数 X_6	平均同一辆车一同进入服务区的人数	实际值(人)
消费水平 X_7	进入服务区车辆的平均消费金额	实际消费金额(元)
消费需求 X_8	出行者对服务区拓展功能的需求	用服务区平均消费需求项表示

隐含层神经元个数的确定通常采用试凑法确定,通过建立多个除隐含层神经元数量不同,其余条件均相同的 MLP 神经网络,比较网络迭代次数和误差精度来确定隐含层神经元个数,该文最终确定隐含层神经元个数为 8 个。输出层神经元为 1 个,即服务区合理规模,采用用地面积(m^2)表示。

故该文建立的 MLP 神经网络合理规模模型各层神经元节点分别为(8,8,1)。

(2) 样本数据预处理

为提高网络的通用性,采用归一化方法对原始样本数据进行预处理,见式(1)。

$$X_i' = 0.1 + \frac{0.8(X_i - X_{\min})}{X_{\max} - X_{\min}} \quad (1)$$

在完成 MLP 神经网络训练之后需进行逆变换,将输出数据转换为原来服务区样本的数据特征,逆变换见式(2):

$$X_i = X_{\min} + \frac{(X_i' - 0.1) \times (X_{i\max} - X_{i\min})}{0.8} \quad (2)$$

式中: X_i 为样本服务区第 i 个变量原始数据输入值;

$X_{i\max}$ 为样本服务区第 i 个变量原始数据最大值; $X_{i\min}$ 为样本服务区第 i 个变量原始数据最小值; X_i' 为变换后的样本服务区第 i 个变量数据。

(3) 模型训练

采用 BP 算法进行模型训练,首先进行神经网络的前向传播,用 $a_i^{(1)} = x_i$ 表示第一层神经元的输入值,也即激活值,此后每一层的激活值用以下步骤实现:

$$a_i^{(1)} = x_i \quad (3)$$

$$a_j^{(t+1)} = f(z_j^{(t+1)}) \quad (4)$$

$$z_j^{(t+1)} = \sum_{i=1}^n W_{ji}^{(t)} a_i^{(t)} + b_j^{(t)} \quad (5)$$

式中: x_i 为归一化后的样本服务区第一层神经元 i 节点数据输入值; $a_i^{(l)}$ 为第 l 层第 i 节点的输出值; $z_i^{(l)}$ 为第 l 层第 i 节点的激活值; $W_{ji}^{(l)}$ 为第 l 层第 i 节点与第 $l+1$ 层第 j 节点之间的联结权重参数; $b_j^{(1)}$ 为第 $l+1$ 层第 j 节点的截距项; f 为激活函数,隐层和输出层分别采用 tansig 和 logsig 激活函数。

逆向传播过程通过梯度下降法不断降低损失函数(输出值与样本值之间的误差),取网络所有输出层节

点的误差平方和作为损失函数,见式(6):

$$C(W,b)=\frac{1}{2}\sum_{i\in outputs}||y_i-a_i||^2$$

(6)

式中: y_i 为归一化后样本服务区输出层 i 节点合理规模期望值; a_i 为样本服务区输出层 i 节点合理规模输出值。

优化目标为确定 W (权值)和 b (偏置)使得损失函数 $C(W,b)$ 最小,即网络输出的值会越来越接近真实值。 W 和 b 的迭代公式如下:

$$W_{ji}^{(l)}=W_{ji}^{(l)}-\alpha\frac{\partial C(W,b)}{\partial W_{ji}^{(l)}}$$

(7)

$$b_j^{(l)}=b_j^{(l)}-\alpha\frac{\partial C(W,b)}{\partial b_j^{(l)}}$$

(8)

式中: α 为学习速率,取 $(0,1]$ 。MLP 模型的训练采用 Matlab 编程实现,模型预测效果优劣选用绝对相对误差($rerr$)、平均绝对相对误差($arerr$)和均等系数(EC)进行评价。

$$rerr=\left|\frac{y_k-a_k}{y_k}\right|$$

(9)

$$arerr=\frac{1}{N}\sum_k\left|\frac{y_k-a_k}{y_k}\right|$$

(10)

$$EC=1-\frac{\sqrt{\sum_k(y_k-a_k)^2}}{\sqrt{\sum_k y_k^2}+\sqrt{\sum_k a_k^2}}$$

(11)

式中: y_k 为服务区规模实际值; a_k 为服务区合理规模预测值; EC 为均等系数,取值范围为 $(0,1)$,表示服务区合理规模预测值与实际值之间的拟合度, EC 值越大预测效果越好; EC 值大于 0.85,预测结果较好; EC 值大于 0.9,预测结果理想。

5 实例分析

运用 2017 年 8 月贵州省高速公路服务区实地调查数据进行分析。采用流量调查和问卷调查获取基础数据及出行者对服务区现有规模的满意度,以单个服务区(双侧)为对象,采用 Spss 进行数据处理、统计分析得到各服务区输入变量值及规模满意度,选取规模满意度大于 80%的服务区作为样本服务区,最后选取贵州省乌江、久长等 19 个服务区作为样本服务区,输入变量及期望合理规模原始数据与归一化后数据见表 3、4。

表 3 样本服务区输入层及期望合理规模原始数据

序号	服务区	断面交通量/ (pcu·d ⁻¹)	驶入率/ %	高峰率	周转率/ 次	车型 结构	同行 人数/人	消费 水平/元	消费 需求	合理 规模/m ²
1	乌江	30 450	25.9	0.09	2.2	0.09	5.8	132.2	2.9	53 887
2	久长	34 820	21.4	0.10	2.3	0.19	9.4	123.7	2.8	62 370
3	桐梓	25 032	21.4	0.09	1.7	0.06	6.6	141.7	3.1	31 207
4	遵义	28 045	15.9	0.11	2.1	0.13	4.0	96.5	2.6	42 522
5	水洋湾	20 980	15.1	0.10	1.8	0.28	4.2	125.5	2.7	40 010
6	怀仁	28 544	16.7	0.09	1.9	0.32	6.5	135.5	3.1	51 596
7	云峰	22 436	28.5	0.11	2.2	0.17	5.2	207.7	2.6	45 000
8	关岭	15 724	27.5	0.08	2.0	0.30	4.8	130.5	2.4	33 334
9	晴隆	14 974	24.4	0.08	2.1	0.18	4.3	176.7	3.2	45 505
10	刘官	6 857	21.6	0.13	2.5	0.42	5.8	194.3	2.3	48 933
11	红果	7 032	20.5	0.13	2.0	0.22	4.3	180.9	2.5	43 068
12	眸珠洞	10 972	24.8	0.17	1.7	0.40	7.1	62.7	2.8	39 958
13	麻江	9 332	24.7	0.17	2.4	0.16	7.3	161.3	2.7	26 373
14	凯里	7 850	12.5	0.09	1.9	0.22	3.5	72.5	2.5	10 656
15	三棵树	7 450	26.6	0.18	2.1	0.18	5.5	149.5	2.9	25 374
16	温泉	8 960	33.5	0.19	1.6	0.38	6.2	165.5	2.8	48 685
17	岩脚	11 423	21.5	0.15	2.0	0.18	4.2	168.5	3.2	41 815
18	织金洞	9 550	26.2	0.13	2.2	0.32	6.1	140.5	2.8	30 000
19	紫云	11 032	25.2	0.12	2.1	0.33	6.2	158.5	3.1	38 750

表 4 归一化后的样本服务区输入层及期望合理规模数据

序号	服务区	断面交通量	驶入率	高峰率	周转率	车型结构	同行人数	消费水平	消费需求	合理规模
1	乌江	0.775 0	0.610 5	0.172 7	0.633 3	0.166 7	0.411 9	0.483 4	0.633 3	0.768 8
2	久长	0.900 0	0.439 0	0.245 5	0.722 2	0.388 9	0.900 0	0.436 6	0.544 4	0.900 0
3	桐梓	0.620 0	0.439 0	0.172 7	0.188 9	0.100 0	0.520 3	0.535 9	0.811 1	0.417 9
4	遵义	0.706 2	0.115 2	0.318 2	0.544 4	0.255 6	0.167 8	0.286 5	0.366 7	0.593 0
5	水洋湾	0.504 0	0.199 0	0.245 5	0.277 8	0.588 9	0.194 9	0.446 5	0.455 6	0.554 1
6	怀仁	0.720 4	0.260 0	0.172 7	0.366 7	0.677 8	0.506 8	0.501 7	0.811 1	0.733 3
7	云峰	0.545 7	0.709 5	0.318 2	0.633 3	0.344 4	0.330 5	0.900 0	0.366 7	0.631 3
8	关岭	0.353 7	0.671 4	0.100 0	0.455 6	0.633 3	0.276 3	0.474 1	0.188 9	0.450 8
9	晴隆	0.332 2	0.553 3	0.100 0	0.544 4	0.366 7	0.208 5	0.729 0	0.900 0	0.639 1
10	刘官	0.100 0	0.446 7	0.463 6	0.900 0	0.900 0	0.411 9	0.826 1	0.100 0	0.692 1
11	红果	0.105 0	0.404 8	0.463 6	0.455 6	0.455 6	0.208 5	0.752 1	0.277 8	0.601 4
12	眸珠洞	0.217 7	0.568 6	0.754 5	0.188 9	0.855 6	0.588 1	0.100 0	0.544 4	0.522 4
13	麻江	0.170 8	0.564 8	0.754 5	0.811 1	0.322 2	0.615 3	0.644 0	0.455 6	0.343 1
14	凯里	0.128 4	0.100 0	0.172 7	0.366 7	0.455 6	0.100 0	0.154 1	0.277 8	0.100 0
15	三棵树	0.117 0	0.637 1	0.827 3	0.544 4	0.366 7	0.371 2	0.578 9	0.633 3	0.327 7
16	温泉	0.131 6	0.900 0	0.900 0	0.100 0	0.811 1	0.466 1	0.667 2	0.366 7	0.688 3
17	岩脚	0.230 6	0.442 9	0.609 1	0.455 6	0.366 7	0.140 7	0.683 7	0.900 0	0.582 0
18	织金洞	0.177 0	0.621 9	0.463 6	0.633 3	0.677 8	0.452 5	0.529 2	0.544 4	0.399 2
19	紫云	0.219 4	0.583 8	0.390 9	0.544 4	0.700 0	0.466 1	0.628 6	0.811 1	0.534 6

抽取前 15 组数据作为训练样本,其余 4 组作为测试样本,MLP 神经网络期望目标值设为 0.01,学习速率为 0.001,最大迭代次数为 1 000,采用 Matlab 编程实现该搭建好的神经网络的训练,通过调试参数使网络达到训练精度要求,训练拟合曲线见图 3,训练形成稳定的网络预测模型后,将测试样本进行仿真检验,得到仿真拟合曲线见图 4,将测试输出结果反归一化后即 为 服 务 区 合 理 规 模 预 测 值 ,并 计 算 得 到 模 型 预 测 效 果 评 价 参 数 ,见 表 5。

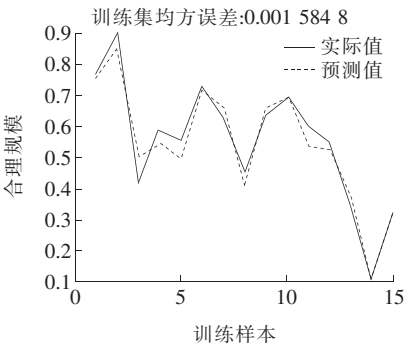


图 3 训练集训练拟合曲线

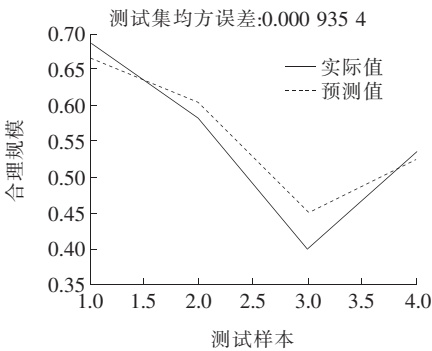


图 4 测试集仿真拟合曲线

表 5 测试样本服务区合理规模测算结果及评价

样本序号	合理规模期望值/m ²	合理规模预测值/m ²	绝对相对误差 (rerr)/%
16	48 685	47 334	2.77
17	41 815	43 346	3.66
18	30 000	31 339	4.46
19	38 750	38 181	1.47

平均绝对相对误差(arerr):3.09%,均等系数 EC:0.984 5

由表5预测结果可知:4对服务区合理规模预测结果的绝对相对误差、最大绝对相对误差、平均绝对相对误差均在5%以内,均等系数为0.984 5,大于0.9,表明该合理规模期望值与预测值拟合度较高,预测结果理想,从而验证了所构建的基于MLP的合理规模测算模型的准确性和可行性。

6 结语

该文在分析高速公路服务区规模测算现有方法的基础上,结合高速公路服务区转型升级新要求,在研究服务区合理规模影响因素时充分考虑出行者消费行为及其他交通参数,构建基于MLP的合理规模测算模型,并用实际数据进行模型验证,结果表明该训练模型预测精度达到98%,可靠性较高,为服务区规划设计及升级改造提供了参考依据。该文所提出的合理规模测算模型亦可应用到其他方面的预测。

参考文献:

- [1] 付建广,周伟,王元庆.高速公路沿线服务区布局规划研究[J].中国公路学报,2001(S1).
- [2] 周智涛,潘兵宏,邓兵.高速公路服务设施合理间距研究[J].中外公路,2014(6).
- [3] 崔洪军,邢小高,梁国凤,等.基于司乘人员需求特征的服务设施间距和功能配置研究[J].重庆交通大学学报(自然科学版),2014(2).
- [4] 高建平,肖英洁,兰北章,等.高速公路开放式服务区功能定位方法[J].长安大学学报(自然科学版),2015(5).
- [5] 王建伟,汤毅.高速公路服务区驶入率的交通势测度模型[J].中国公路学报,2008(5).
- [6] 刘嘉,沈小燕,刘浩学.基于BP神经网络的高速公路服务区车辆驶入率预测[J].公路,2012(6).
- [7] 刘亚非,陈军,刘生涛.陕西省高速公路服务区驶入率调查与预测研究[J].公路工程,2013(4).
- [8] Chen Y, Wang J W. A Study on Forecast Method of Pause Rate in Expressway Service Area[C]. CICTP, 2012.
- [9] Shen X Y, Wang L H, Liu H X, et al. Estimation of the Percentage of Mainline Traffic Entering Rest Area Based on BP Neural Network[J]. Journal of Applied Sciences, 2013, 13(14): 2 632—2 638.
- [10] Kim T, Won J, Lee S, Estimating Parking Lot Space for Freeway Service Area[C]. Proceedings of the 2003 Korea Planner Association Conference, 2003: 609—620.
- [11] Baek S, Kim C. A Research on Parking Demand of Freeway Service Area[C]. Proceedings of the 2006 Korean Society of Civil Engineers Conference, 2006: 3 573—3 576.
- [12] Choi Y, Baek S. A Study on Proper Size of an Expressway Service Area[J]. Journal of Korean Society of Transportation, 2009(27): 7—18.
- [13] 陈传德,邢利霞.高速公路服务区合理规模的确定[J].公路,2010(11).
- [14] 王殊,徐英俊.基于排队论的高速公路服务区停车位合理规模的计算[J].公路交通科技,2016(3).
- [15] Koo C, Hong T, Kim J. A Decision Support System for Determining the Optimal Size of a New Expressway Service Area: Focused on the Profitability[J]. Decision Support Systems, 2014, 67(1): 9—20.
- [16] 苏清华,汪晓红,王建伟,等.基于使用者需求的高速公路服务区功能配置和规模研究[J].公路交通科技(应用技术版),2016(8).
- [17] 欧志霖.基于A—CBR的高速公路服务区最优规模研究[J].中外公路,2016(6).
- [18] 卢超飞.高速公路服务区服务设施的合理配置研究[D].北京交通大学硕士学位论文,2017.
- [19] 田雨波.混合神经网络技术[M].北京:科学出版社,2009.
- [20] 陈彩霞.基于改进粒子群——小波神经网络的预测模型及其应用研究[D].华中师范大学硕士学位论文,2011.