

文章编号:1671-8879(2012)05-0091-06

# 基于分辨矩阵的交通拥堵关键 影响因素提取及解析

许 彬

(同济大学 交通运输工程学院,上海 201804)

**摘 要:**为了给规划、建设和管理不同阶段的决策者提供具备整体性和针对性的决策依据,提高决策的实效性,在全面分析提取交通拥堵影响因素集合的基础上,通过类比分析,选取处理多因素且不完备数据的粗糙集理论方法,构建基于分辨矩阵的属性约简算法模型,结合丰富的实测数据验证分析,提取出导致拥堵的关键因素,包括路段长度、路段车道数、区域出行量、出入口数及两端交叉口连接路段通行能力差。分析结果表明:交通供需平衡在交通运行中的关键作用是不容忽视的,应在设施的规划设计和需求的管控方面加强动态调整力度;分析结论对发展程度不同的城市提供确定宏观交通供需配置及微观精细一体化设计重点的理论及数据支撑。

**关键词:**交通工程;粗糙集;区分矩阵;城市路网;因素提取;规则分析

中图分类号:U491

文献标志码:A

## Extraction and analysis of critical influencing factors of traffic congestion based on discrimated matrix

XU Bin

(School of Transportation Engineering, Tongji University, Shanghai 201804, China)

**Abstract:** In order to provide a holistic and targeted basis for decision making for decision makers at different stages of planning, construction and management and to enhance decision-making efficiency, based on a comprehensive analysis of the factors affecting extraction of traffic congestion, after analogy analysis this paper selected rough set theoretical approach handling multiple factors and incomplete data and built the model on discrimated matrix attribute reduction algorithm. Combined with abundant experimental data analysis, this paper extracted critical influencing factors, including the length of sections, number of lanes, regional travel volume of sections, number of exits and entrances and capacities difference at connection sections of intersection ends. The results indicate the balance between traffic supply and demand in the traffic operation plays a key role, it is necessary to strengthen dynamic control on planning and design of the facilities and the results also provide theory and data support to determine the macroscopic traffic supply and demand allocation and micro integration design for different development levels of cities. 5 tabs, 1 fig, 16 refs.

**Key words:** traffic engineering; rough set; discrimated matrix; urban road network; factor extraction; rules analysis

收稿日期:2011-12-24

基金项目:国家高技术研究发展计划(863 计划)资助项目(2007AA11Z245)

作者简介:许 彬(1980-),男,陕西西安人,工学博士研究生,E-mail:37415560@qq.com。

0 引言

对于现代城市出现的交通问题,是与规划、建设和管理三方面的工作都息息相关的,面对目前大城市渐现的大规模拥堵现象,只在交通方面改善已不是解决问题的唯一途径及根本解决方案。首先要明确影响交通拥堵的最关键的因素是哪些?采取的措施与方法是否有力才能有的放矢和效用最大化。国外学者 Ayadh 研究表明,单向道路的比例、平均每条道路的车道数、信号交叉口密度、平均限制速度和信号交叉口比例等对路网运行状态有重要影响<sup>[1]</sup>;Bhat 采用仿真方法分析了路网特征及交通控制策略对路网运行状态的影响,其研究结果表明:单向交通道路的比例、平均每条道路的车道数、平均限制速度、平均信号周期长度对路网运行状态有重要影响<sup>[2]</sup>;林瑜等通过仿真分析,运用正交试验设计方法,分析了城市道路信号控制交叉口群系统的阻塞影响因素<sup>[3]</sup>;尹洪英等基于解释模型对路网的脆弱性进行了影响因素分析,认为道路条件、路网交通流量、自然因素和人为因素对路网的脆弱性产生影响<sup>[4]</sup>;卢树等以双向四车道以上的城市主要道路为研究对象,认为机非干扰情况、车道宽度、车流密度对行驶车速有着较大的影响<sup>[5]</sup>;戢晓峰运用粗糙集方法进行了路网瓶颈影响因素的分析,描述了相关算

法,但没有进行实证研究<sup>[6]</sup>。目前大部分研究主要集中在以交通视角、以少量交通数据分析及局部交通仿真模拟等手段进行分析。为此,本文充分考虑交通系统形成各个阶段对交通所产生的影响因素,通过选取科学的数据挖掘算法,结合目前可采集的丰富多类型数据,基于交通及相关领域先验知识,构建分析提取模型,对路网运行评价及涉及交通的各个方面的协同改善提供科学的依据。

1 交通拥堵影响因素

中国大城市交通拥堵的根本原因是,城市功能布局失衡导致交通需求时空分布畸形化,在城市空间拓展与结构调整的同时,忽视了对城市功能布局的同步优化调整,突出表现为人口与就业分布失衡,公共服务资源配置不均衡,这是“本”;交通规划设计的不完善、常态化的交通需求管理体系缺失、交通结构失衡和智能化、系统化交通管理水平不足,这是“标”<sup>[7]</sup>。这二者之间从管理层面到理论研究分析层面的脱节是目前存在重要问题之一。

本文对目前可考虑的影响因素进行分层次的梳理总结,因素的选取综合考虑数据可获取性,从路网结构、交通供需关系及交通管理三方面选取 14 个影响因素属性。具体描述见表 1(以下属性编号与表 2、表 3 及表 4 中属性编号数字对应)。

表 1 影响因素属性描述  
Tab. 1 Properties of the affecting factors

编号	名称	描述	单位
1	车道数	反映道路规划设计及供给能力的影响	
2	路段长度	主要反映道路交通规划层面可能带来的影响	m
3	中央分隔形式	主要反映道路规划设计可能带来的影响	
4	机非分隔形式	混合交通是中国特有的交通问题,以设施层面的属性来反映它的影响	
5	所在区域	这里指的区域是大范围区域的概念,比如核心区、郊区等,反映反差较大的城市空间发展的差异	
6	所在区域出行量	反映路段所在规划小区出行量,需求产生带来的影响,反映不同区域影响程度的大小	
7	路段中有无斑马线(无信号控制)	反映慢行交通管理设施可能带来的机非冲突影响	
8	路段出入口数	反映区域单元细节规划设计可能带来的影响	
9	公交专用道	公交出行优先管理策略可能带来的影响	
10	公交站点形式	公交运营可能带来的影响	
11	用地类型	反映区域需求土地开发类型不同可能带来的影响	
12	两端交叉口连接路段通行能力差	反映路网结构及对路段通过能力的影响,为路段两端各自连接路段通行能力之和的差值	pcu/h
13	两端交叉口连接路段总数	反映路网的拓扑结构特性可能带来的影响,以路段两端交叉口连接路段总数来表征	
14	平均延误	路段的延误主要由下游交叉口延误造成,考虑数据获取性,平均延误定义取值以交叉口延误数据在评价路段上的平均值	s/m
15	决策属性(D)	以高峰小时平均车速及拥挤时间比例来判别交通状态。加入拥挤时间比例的判定标准是要避免车速平均化处理带来的问题,主要是所采用以 5 min 的某个单位时间内车速过低或过高影响拥堵判别的准确性。	

表 1 中,平均延误  $T_Y$  是评价路段交通性能及交通管理的效率和水平的参数,常用来评价路段通畅情况及排队情况。反映交通管理设施设置对路段运行带来的影响,该指标值越大,说明道路运行状况

越差,道路越拥堵;反之,该指标值越小,说明评价路段越畅通,公式为

$$T_Y = T/L$$

(1)

式中: $T$  为交叉口延误时间(s); $L$  为评价路段的长

度(m)。

决策属性是根据上海交通分析的实际状况和出行者感受确定,判别标准为:高峰时间平均车速小于 15 km/h,高峰拥堵时间比例超过 50%,判定为拥堵,取值 1,否则取值 0。

## 2 基于分辨矩阵的属性约简算法模型

### 2.1 因素重要度判别方法类分析

最常见的传统方法主要可分为从主观及客观角度出发进行因素重要程度判别进而提取的方法,主观角度是指有主观评分行为的一些方法,如层次分析法(AHP)、德尔斐法等;客观方法主要是通过对指标变量的相关系数、变异信息等进行定量分析进而得到结果,如多元统计中的主成份分析法、因子分析法、嫡权法、相关系数法、模糊数学法和神经网络方法等<sup>[8]</sup>。从实现效果方面看,普遍认为客观方法优于主观方法,这基于实际数据分析、统计分析能够消除主观性影响,但对于样本数据容量及数量要求较高;与此同时,主观方法具有容易理解、发挥专家主观能动性的优点,却存在过于依赖主观评价而对结果的实效性产生影响。

在自然、社会科学及工程技术等领域,基本都存在对不确定性因素和不完备信息的处理问题,交通领域也不例外,实际可获取的数据往往存在不精确、不完整甚至包含噪声数据等问题。经典集合论适用于描述确定性问题,却不能较好地描述模糊和不确定概念。

粗糙集(Rough Sets)理论是 Pawlak 等于 1982 年提出的一种新的处理不确定性知识的数学工具<sup>[9-10]</sup>。粗糙集理论作为一种处理不完备信息的有力工具,它可以不需任何辅助信息(如统计学中的概率分布、模糊集理论中的隶属度等),仅依据本身提供的信息就能够在保留关键信息的前提下,对数据进行化简并求得知识的最小表达,从而建立决策规则,发现给定的数据集中隐含的知识。

### 2.2 基于分辨矩阵的属性约简算法

#### 2.2.1 基本定义

基于决策表的粗糙集方法,其知识表达系统四元组定义为  $T = \langle U, A, V, f \rangle$ ,其中, $U$  为分析对象的集合即论域; $A$  为表示属性的非空集合; $V_a$  是属性  $a$  的值域; $f: u \times (A) \rightarrow V$  是一个信息函数,即对  $a \in A$ ,有  $f(x, a) \in V_a$ 。属性  $A$  由条件属性  $C$  和决策属性  $D$  构成,即  $A = C \cup D, C \cap D = \emptyset$ ,具有条件属性和决策属性的信息系统称为决策表<sup>[11-16]</sup>。

#### 2.2.2 基于分辨矩阵的属性约简算法

属性约简是粗糙集理论最重要的概念之一,约

简可分为 2 个步骤,即属性约简和属性值约简。通过属性约简得到一个最小的属性子集,并且这个子集与整个属性集具有同等的分类能力。通过属性约简,可以将原信息表的冗余属性剔除,使得条件属性的个数减少,重复信息合并,规则的泛化能力有了一定的提高。进行属性约简后的信息表被称为规则表,经过属性约简后的规则表中的每一条规则的条件项数为约简后的属性个数。

根据约简的定义,属性约简是一个迭代的过程。 $P$  要成为  $C$  相对于  $Q$  的约简,首先要测试是否满足  $P_{OS_P}(Q) = P_{OS_C}(Q)$ ,然后测试  $P$  是否相对于  $Q$  正交,  $P_{OS}()$  为约简函数。已经得到证明的是,找出一个决策表的最小约简是一个非线性难题。从约简定义着手求解约简是一个不可取的途径。经过十几年努力,人们提出了不少有效的约简算法,以寻求次优约简。针对道路及周边区域所采集数据的特征与结构形式,如样本数目有限,小类别区分高维、离散值和语义特征明显等数据特性。本文研究工作采用一种基于分辨矩阵的启发式属性约简算法,算法根据属性对不同决策数据的区分能力,逐步选出当前分辨能力最强的属性组成约简。

定义 1:令决策表  $T = \langle U, C \cup D, V, f \rangle, U = \{x_1, x_2, \dots, x_n\}$  是论域,  $a_i(x_j)$  是样本  $x_j$  在属性  $a^i$  上的取值。 $C_D(i, j)$  表示差别矩阵中第  $i$  行第  $j$  列的元素,则差别矩阵  $C_D$  可定义为

$$C_D(i, j) = \begin{cases} \{a_k \mid a_k \in p \wedge a_k(x_i) \neq a_k(x_j)\} & d(x_i) \neq d(x_j) \\ 0 & d(x_i) = d(x_j) \end{cases} \quad i, j = 1, 2, \dots, n \quad (2)$$

差别函数是从差别矩阵中构造的。方法是把  $C_D(i, j)$  中的每个属性“或”起来,然后再“与”起所有的  $C_D(i, j)$ ,形式化描述为

$$f = \bigwedge_{i,j} (\bigvee C_D(i, j)) \quad (3)$$

式中: $f$  为差别函数  $i, j = 1, \dots, |U|, |U|$  为集合  $U$  中元素的个数。

在使用吸收律化简差别函数成标准析取式后,所有的值蕴含式包含的属性就确定了信息系统的约简集合。

上述基于分辨矩阵的属性约简算法提供了一种可以从高维离散数据记录中直接获取关键属性集合的有效方法,下面将利用该算法构建分析模型及计算流程,如下页图 1 所示。

### 2.3 基于分辨矩阵的关键因素提取模型

#### 2.3.1 选取数据及影响因素

主要以确定热点区域覆盖度及数据可获取性两方面选取分析道路组成范围,选取影响因素以区域

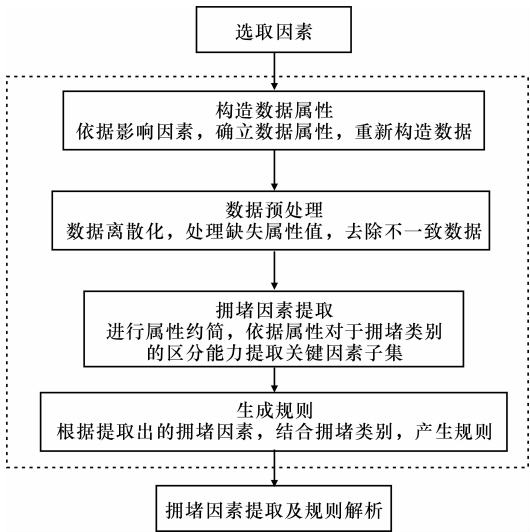


图 1 模型构建流程

Fig. 1 Model construction process

交通评价为核心目标,考虑选取有区域概念的因素。

2.3.2 数据预处理

(1)离散化标准。依据交通领域知识,相关规范及研究分析成果划分属性阈值,在此基础上对数据进行离散化。

(2)处理属性缺失值。依据数据相似性及经验,对统计数据补充或直接去除。

(3)去除不一致数据。条件属性完全一样,但决策值不同的数据记录。

2.3.3 拥堵影响因素提取约简算法构建及计算步骤

输入:信息系统  $T=(U,A)$ ,  $U$  为论域,即选取的全部道路路段,路段个数为  $n$ ,  $A=C\cup D$  是属性集合,  $C$  为条件属性集合,即  $m-1$  个拥堵影响因素,  $D$  为拥堵判别属性集合(1个)。

输出:约简  $Red$ 。

Step1 求差别矩阵  $M_{n\times n}$ ;

Step2 求  $Core_D(C)$  令  $Red = Core_D(C)$ , 令  $s = |Core_D(C)|$ ;

Step3 去掉差别矩阵中含有任意属于  $Red$  的元素,即

$$M = M - \{M(i,j) \mid a_{red} \in M(i,j), \forall a_{red} \in Red, i,j = 1,2,\dots,n\}$$

式中:  $M$  为  $M_{n\times n}$  中的任意元素。

Step4 如果  $m=\varphi$ , 那么转向第 6 步, 否则转向第 5 步;

Step5 按照差别矩阵  $M_{n\times n}$  的元素的属性的个数分类,然后按个数由小到大排序,即  $p=\{p_{t_1}, p_{t_2}, \dots, p_{t_s}\}$ , 属性个数为  $t_1, t_2, \dots, t_s$ , 其中  $t_1 < t_2 < \dots < t_s$ , 设  $P_{t_1}$  中出现次数最多的属性为  $a_{max}$ , 如果存在这样

的多个属性,那么再考察  $P_{t_2}$  中的元素,以此类推,如果到  $P_{t_s}$  中仍存在多个属性,出现次数最多,那么任选一个作为  $a_{max}$ ,  $Red = Red \cup (a_{max})$ , 转向第 3 步;

Step6 结束。

说明:  $Core()$ ,  $red()$  均为粗糙集固有函数;  $s$  为参数。

2.3.4 规则提取及分析

根据上述约简算法提取分辨能力最强的影响因素组合,并对生成的规则进行剖析。

3 实证分析

3.1 数据选取范围及基础属性数据

考虑全面性及针对性,本研究以路段为单元,选取路段应覆盖城市大部分区域,且为需求相对集中的交通走廊及干道,研究范围选取了上海市浦西的三纵三横主要交通走廊涵盖的干道及浦东部分主要干道路段。以下为选取路段基础属性值,见下页表 2 (属性编号与第 1 节中属性编号一致)。

3.2 数据离散化

数据离散化依据交通及相关领域规范、相关研究及数据本身特性。其中,延误划分参考《城市道路交通管理评价指标体系(2008 年版)》中特大城市评价标准;区域出行量的划分以上海市第四次大调查报告中出行发生量为标准进行划分。

其他属性划分见下页表 3;数据离散化结果见下页表 4。

3.3 结果及规则分析

3.3.1 因素提取

运用选取的约简算法进行属性约简,可以得到导致拥堵的关键因素组合为:

{路段长度,路段车道数,区域出行量,出入口数,两端交叉口连接路段通行能力差}。

路段的长度在路网规划阶段形成比较固定的模式,过短或过长路段对于交通运行管理会产生一定的影响,如路网规划的均匀和均衡性;路段车道数侧面反映了道路供给能力,是供需平衡在路段单元上一方面的体现;以出行量表征的区域需求特征,表明了交通需求对于交通状态的重要影响;路段沿线出入口数量的增加将带来路段行驶车辆交织可能性的增加,增加延误及安全隐患,对路段行程秩序带来一定影响,在数量上必须进行控制;两端交叉口连接路段通行能力差,反映了路网结构对于交通运行及后期管理设施设置的效果起到关键作用。

表 2 路段基础属性值  
Tab. 2 Basic attributes of road

属性编号		1	2	3	4	5	6	7	8	9	10	11	12	13	14	D
路段编号	1	343	4	3	0	3	3	0	3	1	0	4	2 500	7	0. 11	0
	2	249	4	3	0	3	3	0	1	1	0	4	500	6	0. 10	0
	3	326	4	3	1	3	3	0	2	1	0	2	0	6	0. 27	0
	4	295	4	3	0	3	3	0	2	1	0	4	0	6	0. 13	0
	5	218	4	3	0	3	3	0	3	1	0	2	0	6	0. 19	0
	6	314	5	3	1	3	4	0	1	1	1	4	6 000	6	0. 21	0
	7	184	3	0	0	2	5	0	0	0	0	2	0	4	0. 17	0
	8	249	3	3	0	2	5	0	0	0	0	2	800	4	0. 18	1
	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
	372	370	3	3	0	2	5	0	2	0	0	2	2 000	5	0. 25	1
	373	332	3	3	1	2	5	0	1	0	0	1	2 000	5	0. 32	1
	374	207	4	3	1	2	5	0	0	0	0	2	9 000	5	0. 24	1
	375	503	4	3	1	3	5	0	3	1	0	2	6 000	6	0. 37	1
	376	138	5	3	0	3	5	0	1	1	0	4	0	6	0. 15	0
	377	274	4	3	1	3	5	0	1	1	0	2	1 600	6	0. 28	0
	378	253	3	1	1	3	5	0	1	0	0	4	2 000	5	0. 24	1
	379	409	2	1	1	3	5	0	4	0	0	2	6 400	6	0. 36	1
	380	195	4	1	1	3	2	0	1	1	0	2	8 600	5	0. 08	0

表 3 属性划分及离散化标准  
Tab. 3 Attributes classification and discretizing standards

1	离散	2	离散	3	离散	4	离散	5/(万人次·d <sup>-1</sup> )	离散
0~200	1	无分隔	0	无分隔	0	中环以外	1	0~18. 75	1
200~500	2	分隔栏	1	分隔栏	1	中环内环间	2	18. 75~37. 5	2
500~800	3	分隔墩分隔墩		分隔墩		内环以内	3	>37. 5	3
>800	4	分隔带	2	分隔带	2				
				无非机动车道	3				
6	离散	7	离散	8	离散	9	离散	10	离散
无	0	0 或 1 个	1	无	0	港湾	0	娱乐休闲	1
有	1	2 个	2	有	1	路边	1	住宅	2
		>2 个	3					商业	4
								机关单位	8
11	离散	12	离散	13	离散				
0	0	0 ~ 5	1	0~0. 08	0				
0~1 300	1	6	2	0. 08~0. 14	1				
1 300~5 400	2	>6	3	>0. 14	2				
>5 400	3								

3. 3. 2 规则分析

选取有代表性的规则进行解析,选取规则见下页表 5。

从上述有代表性的约简后得到的规则,可以总结出,当路段长度小于 500 m,车道数 3 条以上,区域出行量 37. 5 万人次/d 左右,出入口数大于 2,两端交叉口连接路段通行能力差,路段交通将大都处于拥堵状态,同时在这 5 个属性值出于阈值范围两

端时,路段交通将大都处于拥堵状态。

对于提取的以上关键因素,主要体现在供需两方面,对于上海这样交通发展程度较高的城市来说,后期管理改善措施发挥效用的空间已相对较小,对于中国大多数正在快速城市化的区域来说,就需要重点在规划阶段保持供需平衡,加强需求监控管理,在确定其属性值的不同阶段,避免出现个体间的极差现象,为进入交通管理阶段预留更多的改善空间。

表 4 离散化数据结果  
Tab. 4 Discreting data results

属性编号		1	2	3	4	5	6	7	8	9	10	11	12	13	14	D
路段 编号	1	2	2	2	0	3	3	0	3	1	0	4	2	3	1	0
	2	2	2	2	0	3	3	0	1	1	0	4	1	2	1	0
	3	2	2	2	1	3	3	0	2	1	0	2	0	2	2	0
	4	2	2	2	0	3	3	0	2	1	0	4	0	2	1	0
	5	2	2	2	0	3	3	0	3	1	0	2	0	2	2	0
	6	2	3	2	1	3	3	0	1	1	1	4	3	2	2	0
	7	1	3	0	0	2	1	0	1	0	0	2	0	1	2	0
	8	2	3	2	0	2	1	0	1	0	0	2	1	1	2	1
	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
	372	2	3	2	0	2	1	0	2	0	0	2	2	1	2	1
	373	2	3	2	1	2	1	0	1	0	0	1	2	1	2	1
	374	2	2	2	1	2	1	0	1	0	0	2	3	1	2	1
	375	3	2	2	1	3	2	0	3	1	0	2	3	2	2	1
	376	1	3	2	0	3	3	0	1	1	0	4	0	2	2	0
	377	2	2	2	1	3	3	0	1	1	0	2	2	2	2	0
	378	2	3	1	1	3	3	0	1	0	0	4	2	1	2	1
	379	2	1	1	1	3	3	0	3	0	0	2	3	2	2	1
	380	1	2	1	1	3	2	0	1	1	0	2	3	1	0	0

表 5 选取规则  
Tab. 5 Section rules

路段长度	路段车道数	区域出行量	出入口数	两端交叉口连接路段通行能力差	拥堵判别
3	1	2	2	3	1
1	2	1	1	1	0
2	3	2	3	2	1
2	2	2	1	3	1
2	2	3	1	1	1
2	2	2	1	0	0
2	3	1	1	1	0
2	1	2	1	1	0
1	3	2	1	2	1
2	2	2	1	0	0

4 结 语

(1)全面分析了导致交通拥堵的各个层面的因素,进而筛选出最重要的影响因素,为评价指标计算要素的选取制定提供依据;选取了科学适用的粗糙集分辨矩阵约简算法,选取较大范围的数据进行实际论证,提取出了拥堵影响因素组合,并对约简后得到的规则进行分析,得出了较为合理的定量描述的拥堵判别标准,为构建状态评价指标体系及阈值分析提供参考,为规划、建设和管理各个层面的管理决策提供科学依据。

(2)道路条件和交通需求对于大城市的拥堵影响程度大,说明城市交通必须在规划和先导管理阶段制定合理的控管措施和规则,这才是交通治理之本。

参考文献:  
References:

[ 1 ] Ayadh M T. Influence of the city geometric features on the two fluid model parameters[D]. Virginia: Virginia Polytechnic Institute and State University,1986.

[ 2 ] Bhat S C S. Effects of geometric and control features on network traffic;a simulation study[D]. Arlington: University of Texas at Arlington,1994.

[ 3 ] 林 瑜,杨晓光,马莹莹. 交通阻塞影响因素分析的正交试验设计方法[J]. 系统工程,2005,23(10):39-43. LIN Yu,YANG Xiao-guang,MA Ying-ying. An analysis of factors causing congestion with the application of orthogonal experimental design method [J]. Systems Engineering,2005,23(10):39-43. (in Chinese)

[ 4 ] 尹洪英,徐丽群,权小锋. 基于解释结构模型的路网脆弱性影响因素分析[J]. 软科学,2010,24(10):122-126. YIN Hong-ying,XU Li-qun,QUAN Xiao-feng. Research on influencing factors of road network's vulnerability based on interpretive structural model [J]. Soft Science,2010,24(10):122-126. (in Chinese)

[ 5 ] 卢 树,陆化普,唐忠华,等. 城市道路行程车速影响因素分析[J]. 公路交通科技,2003,20(1):89-92. LU Shu,LU Hua-pu,TANG Zhong-hua,et al. Analysis of influencing factors on urban travel speed[J]. Journal of Highway and Transportation Research and Development,2003,20(1):89-92. (in Chinese)

