

文章编号: 1000-8152(2009)10-1157-05

# 交通出行诱导的离散Stackelberg动态博弈模型及其求解算法

李文勇<sup>1,2</sup>, 陈学武<sup>1</sup>, 陆 建<sup>1</sup>

(1. 东南大学 交通学院, 江苏 南京 210096; 2. 桂林电子科技大学 机电工程学院, 广西 桂林 541004)

**摘要:** 针对交通出行诱导的实际需要, 分析了道路管理者与道路使用者的博弈策略和博弈特征. 通过道路使用者信息的模糊模型, 建立了交通出行诱导的离散动态Stackelberg博弈模型. 分析了交通出行诱导的两阶段博弈特征, 提出了基于逆向归纳法的博弈模型求解算法. 通过对模型无约束转化, 运用Monte-Carlo法对示例路网进行了计算机求解, 求解结果表明产生的交通出行诱导方案能够实现系统最优下的用户最优.

**关键词:** 出行诱导; 博弈模型; Monte-Carlo法

**中图分类号:** U491.1      **文献标识码:** A

## Discrete dynamic Stackelberg game model for traffic guidance strategy and its solving method

LI Wen-yong<sup>1,2</sup>, CHEN Xue-wu<sup>1</sup>, LU Jian<sup>1</sup>

(1. School of Transportation, Southeast University, Nanjing Jiangsu 210096, China;

2. School of Mechanical & Electrical Engineering, Guilin University of Electronic Technology, Guilin Guangxi 541004, China)

**Abstract:** In order to develop the practical actual traffic guidance, we investigate the game strategies and characteristics between the traffic controllers and the road users. The discrete Stackelberg game model, which describes the guidance strategies, is developed based on the established fuzzy model of the road user's information. After analyzing the game characteristics in two phases of traffic guidance, we put forward a solving method for this game model based on the converse induction method. Through the unconstraint model-conversion, an illustrative road-net is solved in a computer by using the Monte-Carlo method. The obtained traffic guidance scheme is optimal to the individual user as well as the whole system.

**Key words:** traffic guidance; game model; Monte-Carlo method

### 1 引言(Introduction)

交通诱导系统是一种主动式的交通控制方式, 其最大特点是通过传递情报消息引导和控制交通参与者的交通行为, 以达到交通安全、畅通、有序的控制目的. 由于交通诱导的非强制性, 使得交通出行诱导系统在交通管理者与出行者之间存在一种博弈的关系.

本文以实现系统最优与用户最优的平衡和统一为目标, 对道路管理者的交通诱导方案和出行者的路径选择进行博弈分析, 研究基于博弈理论的城市交通出行诱导策略和出行诱导方案的生成技术<sup>[1]</sup>.

### 2 交通出行诱导的博弈模型(Game model of traffic travel guidance)

#### 2.1 博弈策略分析(Game strategies analysis)

##### 1) 道路管理者的策略.

管理者的策略是对全网发布诱导信息. 诱导信息包括区域路网中每条路段上的预期行程时间和有效诱导路径集中不同诱导路径的优先级. 道路管理者为获得系统最优, 诱导信息应该全局发布, 保证道路使用者获得平等的交通信息.

##### 2) 道路使用者的策略.

道路使用者的策略是根据道路管理者的诱导信息, 选择自己“感知”阻抗最小的路径. 道路使用者的策略是对不同路径的选择概率. 所有道路使用者的策略组合构成博弈双方中参与者的策略.

#### 2.2 Stackelberg 博弈模型(Stackelberg game model)

##### 1) 道路管理者的收益(SO):

$$J_1 = [-\sum_a x_a \tau_a(x_a)]. \quad (1)$$

收稿日期: 2008-04-08; 收修改稿日期: 2009-02-22.

基金项目: 国家自然科学基金资助项目(50808050); “十一五”国家科技支撑计划资助项目(2006BAJ18B01); 广西科学研究与技术开发计划资助项目(桂科攻0719001-2).

式中 $x_a$ 和 $\tau_a$ 分别为路段 $a$ 的实时交通量和阻抗。

2) 道路使用者的收益(SUE):

$$J_2 = \left[ \sum_{rs} q_{rs} \cdot E[\min_{k \in \phi_{rs}} \{c_k^{rs}\} | \bar{C}^{rs}(\bar{x})] - \sum_a x_a \tau_a(x_a) + \sum_a \int_0^{x_a} \tau_a(\omega) d\omega \right]. \quad (2)$$

式中:  $E[\min_{k \in \phi_{rs}} \{c_k^{rs}\} | \bar{C}^{rs}(\bar{x})]$  为各路段实际阻抗为条件的感知阻抗的数学期望, 可称作“期望感知阻抗”;  $\bar{C}^{rs}(\bar{x})$  是 $(r, s)$ 间各条路径的实际阻抗组成的向量;  $c_k^{rs}$  是 $(r, s)$ 间第 $k$ 条路径的感知阻抗。  $q_{rs}$  为 $(r, s)$ 间的OD交通需求量。

道路管理者与使用者之间的博弈特征<sup>[2~4]</sup>:

1) 参与者选择行动是按先后顺序发生的。道路管理者先提供一个诱导方案, 继而道路使用者根据接受到的诱导信息, 在理性地分析后, 选择适合自己的出行路径。

2) 博弈是完美信息的。也即道路管理者发布每一种诱导信息, 道路使用者都可以获知。而道路使用者决策后选择路径的情况道路管理者也知道。

3) 博弈是完全信息的。也即道路管理者的收益是基于用户最优的, 道路使用者的收益是基于系统最优的。

### 2.3 博弈过程的逆向归纳法分析(Game process analysis based on converse induction method)

1) 道路使用者先优化自己的策略

道路管理者为了求得总的出行成本最小, 先分析系统最优下的路网状态: 路网各路段上的流量 $\{x_a^{SO}\}$ , 加载流量后路网各路段上的行程时间 $\{\tau_a^{SO}\}$ 。

道路使用者在接受诱导信息后的选择最优策略<sup>[6,7]</sup>, 即得到 $S_U^* = R_U(S_M)$ , 构成一个可行的策略域。不论道路管理者选择哪个策略, 道路使用者出行都已经是最优化的。

2) 道路管理者在前者的最优策略空间中继续优化自己的策略

道路管理者分析自己采用不同的策略时道路使用者的反应得到 $R_U(S_M)$ , 既预测到道路使用者的策略; 道路管理者选择最优策略 $S_M^*$ 使系统的收益最大。可见道路管理者利益最大化的过程, 同时也实现了道路使用者的最大的利益。

综上, 道路管理者和使用者完成了2阶段的动态博弈过程, 并最终达到共赢的局面。道路管理者可通过长期的博弈性策略发布获得经验: 道路使用者对交通诱导信息接受程度(确信度)和道路使用者的选择特性等。

## 3 离散Stackelberg动态博弈模型(Differential dynamic Stackelberg game model)

### 3.1 博弈模型的建立基础(The base of game model)

道路管理者首先给出诱导信息 $\{\tau_a^G, \forall a\}$ ,  $\tau_a^G$ 为道路管理者根据动态交通分配的结果预测给出的路段 $a$ 的行程时间。道路使用者根据管理者发布的诱导信息推算出确信度 $\lambda$ (也即对诱导信息的相信概率)。确信度是驾驶员长期对路网的行驶感知, 可用时间 $\tau$ 的隶属度函数 $P(\tau)$ 来表示。将 $\tau$ 反解可得到基于确信度 $\lambda$ 的道路使用者感知时间 $\tau$ 。

一般情况下当交通量比较小时, 各种拥挤和延误都比较小, 道路使用者会比较相信道路管理者的诱导信息, 这时确信度可取大值。随着道路上交通量的增加, 各种不确定性因素也将增多, 道路使用者对诱导信息的确信度将会逐渐减小, 但诱导信息仍是有价值的。当道路交通量达到一个临界值后, 道路使用者的选择空间将很小, 确信度也将降到固定的最小值。

式(3)给出了确信度的数学模型:

$$\lambda_i(x_i) = \alpha \times e^{-\gamma(\frac{x_i}{x_i^0})^2} \times \theta(x_i^0 - x_i) + \beta \times \theta(x_i - x_i^0), \quad x_i^0 < \sqrt{\frac{1}{2\gamma}}. \quad (3)$$

其中:  $\alpha, \gamma, \sigma$  为各等级道路需要待标定的参量,  $\beta$  为最小确信度,  $x_i$  为路段 $i$ 的交通量,  $x_i^0$  为路段 $i$ 的临界交通量。

道路使用者的确信度对道路管理者的收益构成很大的风险, 因此道路使用者选择多大的确信度对道路管理者来说至关重要。同时对于不同的拥挤状态,  $\lambda$ 的选择也应不同。因为道路使用者对不同的拥挤状态会有不同的心理准备。

道路使用者有了对路段的模糊估计, 便可以求出每条路径的感知阻抗:

$$\bar{C}_k^{rs} = \sum_a \bar{\tau}_a^G \delta_{a,k}^{rs}. \quad (4)$$

$\delta_{a,k}^{rs}$  为路段-路径相关变量; 如果路段 $a$ 在 $(r, s)$ 间的第 $k$ 条路径上,  $\delta_{a,k}^{rs}$  取1, 其他情况为0。

道路使用者根据用户最优原则进行路径选择, 可用Logit公式来模拟这种选择:

$$\bar{P}_k^{rs} = \frac{e^{-\theta \bar{C}_k^{rs}}}{\sum_{a \in A} e^{-\theta \bar{C}_a^{rs}}}. \quad (5)$$

得到每条路径的选择概率就可得到每条路径的流量:

$$\bar{f}_k^{rs} = q^{rs} \bar{P}_k^{rs}. \quad (6)$$

可以看出  $\sum_k \bar{P}_k^{rs} = 1$ , 即  $q^{rs} = \sum_i \bar{f}_i^{rs}$  自动满足.

道路使用者选择后, 应对路网的相关参量更新:

$$\hat{x}_i = \sum_r \sum_s \sum_k \hat{f}_k^{rs} \delta_{i,k}^{rs}. \quad (7)$$

基于BPR函数, 本文将路段行程时间简化为  $\bar{\tau} = \tau_a^0 + \beta_a \bar{x}_a$ .

### 3.2 NASH均衡下的优化求解(Optimum calculation for NASH equilibrium)

对于完全信息动态博弈的均衡解是子博弈精练的NASH均衡<sup>[8]</sup>. 可以用逆向归纳法来求解上述完全且完美信息动态博弈问题.

1) 当博弈进入第2阶段, 即由道路使用者决策的时候, 由于道路管理者在第1阶段已经选择行动  $\tau_a^G = (\tau_1^G, \dots, \tau_i^G, \dots) \in \Gamma^G$ .

道路使用者的决策问题就变成道路使用者收益的最优化问题:

$$u_U(S_M, S_U) \Leftrightarrow J_2(\tau_1^G, \dots, \tau_i^G, \dots; \hat{P}_1, \dots, \hat{P}_i, \dots). \quad (\text{PII})$$

其中:  $\tau_i$  为管理者选择发布的路段  $i$  的出行时间;  $\hat{P}_i$  为道路使用者基于对路段出行时间的感知情况对路径  $i$  的选择概率.

最优化问题为: 由  $\max_{S_U \in P} u_U(S_M, S_U)$ , 得到

$$S_U^* = R_U(S_M). \quad (8)$$

由  $S_U^* = R_U(S_M)$  可以得到道路使用者对于道路管理者的选择行动  $\tau_a^G$  作出的最优的路径选择. 由于道路管理者能和道路使用者一样求出  $S_U^* = R_U(S_M)$ , 即, 道路管理者可以预测出道路使用者对其每一个诱导策略的最优反应.

根据交通分配理论可知, 对式(PII)的优化与SUE条件是等价的. 因此  $S_U^* = R_U(S_M)$  可以得到.

2) 回到第1阶段, 道路管理者的收益:

$$J_1 = [-\sum_a \bar{x}_a \cdot \bar{\tau}_a] = J_1(\tau_1^G, \dots, \tau_i^G, \dots; \hat{P}_1, \dots, \hat{P}_i, \dots).$$

等价于

$$u_M(S_M, S_U). \quad (\text{PI})$$

道路管理者的决策即为如下最优化问题:

$$\max_{S_M \in \Gamma} u_M(S_M, S_U^*) \Leftrightarrow \max_{S_M \in \Gamma} u_M(S_M, R_U(S_M)).$$

通过求解可得到最优策略组合:

$$(S_M^*, R_U(S_M^*)). \quad (9)$$

可以看出求解得到  $S_M^*$  至关重要, 因此道路管理者的诱导信息应该慎重选择.

通过上述分析, 可以得到如下交通出行诱导博弈模型:

$$\begin{aligned} & \max_{S_M \in \Gamma} u_M(S_M, S_U^*), \\ & \text{s.t. } S_U = R_U(S_M), \end{aligned}$$

即

$$\begin{aligned} & \max_{\tau_i \in \Gamma} J_1(\tau_1^G, \dots, \tau_i^G, \dots; \hat{P}_1, \dots, \hat{P}_i, \dots), \\ & \text{s.t. SUE条件.} \end{aligned}$$

## 4 模型示例求解(Example calculation of model)

上述得到的博弈模型是一个有约束的非线性优化问题. 本文尝试运用Monte-Carlo法<sup>[9]</sup>并用计算机编程来求解.

### 4.1 示例分析路网(Example road-net)

示例分析路网见图1所示, 有6个节点, 8条路段构成.

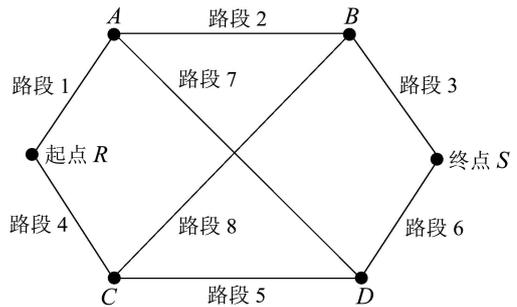


图 1 示例路网

Fig. 1 Example road-net

表 1 示例路网的初始状态量表

Table 1 Initialization parameter of example road-net

路段	自由流行程 时间 $\tau_a^0/s$	临界饱和状态下 行程时间 $\tau_a^S/s$	$\beta_a$
1	10	28	0.01
2	9	40	0.02
3	5	8	0.04
4	11	35	0.005
5	8	24	0.015
6	3	21	0.06
7	13	15	0.001
8	15	20	0.001

OD量:  $q_{RS} = 2000(\text{pcu/h})$ ; 路径的定义

$L1 : R \rightarrow A \rightarrow B \rightarrow S, L2 : R \rightarrow A \rightarrow D \rightarrow S,$

$L3 : R \rightarrow C \rightarrow B \rightarrow S, L4 : R \rightarrow C \rightarrow D \rightarrow S.$

各路段在相应交通管制措施下的行程时间( $\tau_a = \tau_a^0 + \beta_a x_a, a = 1, \dots, 8$ )见表1.

### 4.2 Stackelberg博弈模型的Monte-Carlo法求解(Calculation for Stackelberg game model based on Monte-Carlo method)

根据各路径流量与路段流量关系、总流量守恒关系、路径阻抗与路段阻抗间关系, 代入式(PI), 可得到如下数学优化模型:

$$\begin{aligned} \min_f J_1(f_1, f_2, f_3), \\ \text{s.t. 满足上述关系.} \end{aligned}$$

其最优解应满足

$$\begin{aligned} \frac{\partial J_1(f_1, f_2, f_3)}{\partial f_1} = 0, \quad \frac{\partial J_1(f_1, f_2, f_3)}{\partial f_2} = 0, \\ \frac{\partial J_1(f_1, f_2, f_3)}{\partial f_3} = 0. \end{aligned}$$

将数值代入得到线性方程组如下:

$$\begin{bmatrix} 0.3 & 0.06 & 0.23 \\ 0.06 & 0.062 & 0.03 \\ 0.23 & 0.03 & 0.232 \end{bmatrix} \begin{bmatrix} f_1^{\text{SO}} \\ f_2^{\text{SO}} \\ f_3^{\text{SO}} \end{bmatrix} = \begin{bmatrix} 318 \\ 76 \\ 291 \end{bmatrix},$$

求解可得

$$\begin{bmatrix} f_1^{\text{SO}} \\ f_2^{\text{SO}} \\ f_3^{\text{SO}} \\ f_4^{\text{SO}} \end{bmatrix} = \begin{bmatrix} 169.5 \\ 571.9 \\ 1012.3 \\ 246.3 \end{bmatrix}.$$

在制订的诱导信息

$$[\tau_1^{\text{G*}}, \tau_2^{\text{G*}}, \tau_3^{\text{G*}}, \tau_4^{\text{G*}}, \tau_5^{\text{G*}}, \tau_6^{\text{G*}}, \tau_7^{\text{G*}}, \tau_8^{\text{G*}}]^T$$

下, 道路使用者的反应为

$$\frac{e^{-\theta C_i^{\text{G*}}}}{\sum_{k=1}^4 e^{-\theta C_k^{\text{G*}}}} = \frac{f_i^{\text{SO}}}{q}.$$

将 $e^{-\theta C_i^{\text{G*}}} = \lambda_i$ 进行变量代换, 上述方程组可转化为矩阵形式. 由此, 可以解出

$$[\tau_1^{\text{G*}}, \tau_2^{\text{G*}}, \tau_3^{\text{G*}}, \tau_4^{\text{G*}}, \tau_5^{\text{G*}}, \tau_6^{\text{G*}}, \tau_7^{\text{G*}}, \tau_8^{\text{G*}}]^T.$$

由于限制诱导策略发布的约束条件太少, 该方程组有无穷解, 因此诱导策略也很多. 因此通过计算机编程计算任取一组, 如

$$[27.04 \text{ s}, 39.31 \text{ s}, 5.85 \text{ s}, 34.08 \text{ s}, 15 \text{ s}, 20 \text{ s}, 13 \text{ s}, 15 \text{ s}]^T.$$

其对应每条路段的预期最优流量也可计算出来.

上述分析只是给出总体规律下的博弈结果, 考虑道路使用者对发布诱导信息的确信度 $\lambda$ , 道路管理者应发布的诱导时间 $\tau'_i$ 为:  $\tau'_i = \tau_i^{\text{G*}}(1 - \ln \lambda)$ . 表2给出了示例路网道路管理者应发布的诱导信息. 示例计算中本文仅给出当 $\lambda$ 已经确定后的示意性的结果. 其中常发性拥挤,  $\lambda$ 取0.9, 偶发性拥挤,  $\lambda$ 取0.7.

在道路使用者随机用户最优的策略下, 通过博弈道路管理者最终达到系统最优. 其路网参数见表3. 图2给出了NASH均衡流量与用户最优流量的对比图.

表2 道路管理者诱导信息表

Table 2 Guidance information table of traffic management

路段	1	2	3	4	5	6	7	8
推荐时间/s	27.04	39.31	6.47	34.08	22.66	20.00	13.00	15.00
拥挤类型	正常	正常	交通拥堵	正常	交通事故	正常	正常	正常

表3 均衡状态下路网参量表(流量单位: pcu/h)

Table 3 Equilibrium parameter table of example road-net

路段	1	2	3	4	5	6	7	8
NASH均衡流量	741.4	169.5	1181.8	1258.6	246.3	818.2	571.9	1012.3
均衡时间/s	17.414	12.390	52.270	17.293	11.695	52.092	13.572	16.012
用户最优流量	1028.6	565.6	846.4	971.4	690.6	1153.6	463	280.8
用户最优时间/s	10	9	5	11	8	3	13	15

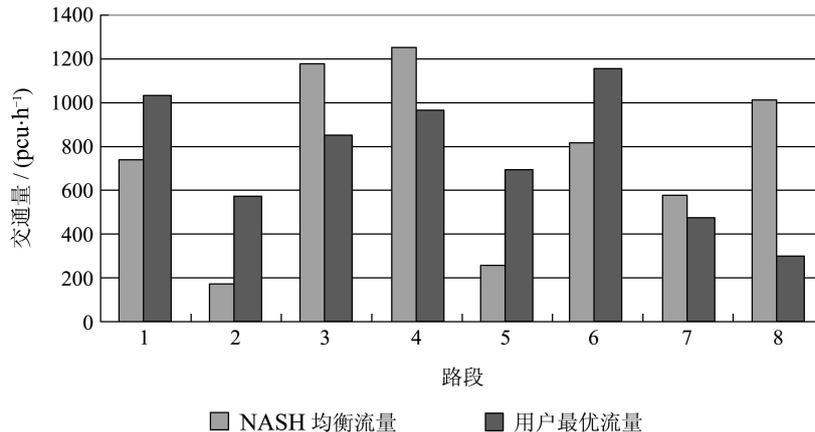


图 2 用户最优流量与NASH均衡流量对比图

Fig. 2 Contrast figure about user optimum flow and NASH equilibrium flow

## 5 结束语(Conclusion)

本文主要通过对道路管理者的交通诱导方案和出行者的路径选择进行博弈分析,建立了交通出行诱导离散Stackelberg动态博弈模型. 并采用逆向归纳法进行模型求解,得到用于网络发布的诱导策略,从而使路网流量实现NASH均衡. 论文研究成果可提高交通出行诱导系统的诱导效率,同时也可交通管理者制订诱导分流方案提供理论与技术指导.

本文尝试使用Monte-Carlo法来对建立的博弈模型进行求解,当变量多时,这种方法的效率较低. 因此,寻找求解速度适当的具有全局寻优能力的数值求解算法是本文后续应该重点研究的方向

## 参考文献(References):

- [1] 肖条军. 博弈论及其应用[M]. 上海: 上海三联书店, 2003. (XIAO Tiaojun. *The Theory and Application of Games*[M]. Shanghai: ShangHai Sanlian bookstore Press, 2003.)
- [2] RAMADURAI, GITAKRISHNAN, UKKUSURI, et al. *Dynamic traffic equilibrium: theoretical and experimental network game results in the single-bottleneck model-TRB(07-3390)*[C]. Washington, DC: The Transportation Research Board(TRB), 2007.
- [3] 王亮, 马寿峰, 李艳君. 对交通控制与诱导结合研究的评述[J]. 系统工程, 2004, 22(2): 56 - 62. (WANG Liang, MA Shoufeng, LI Yanjun. Discussion on the combination of traffic control with vehicles route guidance[J]. *Systems Engineering*, 2004, 22(2): 56 - 62.)
- [4] 李振龙. 基于Stackelberg 博弈的动态用户最优配流和信号控制[J]. 控制理论与应用, 2005, 22(3): 353 - 358. (LI Zhenlong. Dynamic user optimal assignment and signal control based on the Stackelberg game[J]. *Control Theory & Applications*, 2005, 22(3): 353 - 358.)
- [5] 马寿峰, 卜军峰, 张安训. 交通诱导中系统最优与用户最优的博弈协调[J]. 系统工程学报, 2005, 20(1): 30 - 37. (MA Shoufeng, BU Junfeng, ZHANG Anxun. Game based coordination method between system optimum and user equilibrium in route guidance system[J]. *Journal of Systems Engineering*, 2005, 20(1): 30 - 37.)
- [6] 李振龙. 诱导条件下驾驶员路径选择行为的演化博弈分析[J]. 交通运输系统工程与信息, 2003, 3(2): 23 - 27. (LI Zhenlong. A Study of route choice behavior of drivers based on the evolutionary game under the condition of traffic flow guidance[J]. *Communication and Transportation Systems Engineering and Information*, 2003, 3(2): 23 - 27.)
- [7] 徐建闽, 许伦辉, 撒元功. 交叉口有交通信号控制时用户最优动态配流模型[J]. 控制理论与应用, 2000, 17(1): 117 - 120. (XU Jianmin, XU Lunhui, SA Yuangong. Dynamic user-optimal assignment model of road network with traffic signal control at its intersections[J]. *Control Theory & Applications*, 2000, 17(1): 117 - 120.)
- [8] PARK DONGJOO, KIM HANSOO, HONG SEOKKI, et al. *Location-based dynamic route guidance system of Korea: system design, algorithms, and initial results-TRB(07-1074)*[C]. Washington, DC: The Transportation Research Board(TRB), 2007.
- [9] 于永生, 刁联旺. 区间数多属性决策的Monte Carlo方法[J]. 东南大学学报(自然科学版), 2008, 38(1): 187 - 190. (YU Yongsheng, DIAO Lianwang. Monte Carlo method for interval multi-attribute decision-making[J]. *Journal of Southeast University(Natural Science Edition)*, 2008, 38(1): 187 - 190.)

## 作者简介:

**李文勇** (1976—), 男, 桂林电子科技大学副教授, 博士, 研究方向为交通信息工程及控制, E-mail: traffic@guet.edu.cn;

**陈学武** (1968—), 女, 博士, 教授, 博士生导师, 研究方向为交通运输规划与管理, E-mail: chenxuewu@seu.edu.cn;

**陆建** (1972—), 男, 博士, 教授, 博士生导师, 研究方向为交通运输规划与管理, E-mail: lujian1972@seu.edu.cn.