

doi:10.16055/j.issn.1672-058X.2015.0012.007

诱导信息条件下车辆路径选择 ——基于有限理性模糊博弈

周代平, 李康奇, 贺琳

(重庆交通大学 交通运输学院, 重庆 400074)

摘要:诱导信息属于建议性信息,当诱导信息发布以后驾驶员并不一定接受诱导信息;在行为强化理论基础上,建立了基于有限理性模糊博弈的诱导信息条件下车辆路径选择模型,并通过仿真得出了不同初始状态下的博弈平衡结果;仿真结果表明:诱导信息的发布并不一定能改善路网交通流分布,它与路网总流量以及初始接受诱导的车辆比例密切相关。

关键词:模糊博弈;有限理性;交通诱导;自学习机制

中图分类号:U491.23 **文献标志码:**A **文章编号:**1672-058X(2015)12-0031-05

交通诱导是智能交通(ITS)的一部分,它的主要作用是将路况实时信息反馈给驾驶员,以便驾驶员选择合理的出行路径.Roth and Erev^[1],Erev^[2]认为出行者进行路径选择同时也是在不断地实践,通过多次不同的选择,出行者会总结出几次选择的路径中最佳的路径,并将其作为自己今后选择路径的依据,所以路径的选择也是一个由学习过程转化成自身经验的过程.在交通领域逐渐发展的过程中,博弈论的应用越来越广泛.Michael G.H. Bell(2000)^[3]将博弈论运用到为了选择最佳行驶路径使自己花费最小的出行者个体和为了使整个路网的收益最大的破坏者之间,建立非合作博弈模型,在保证路网安全性下,得到了 Nash 混合策略均衡解;李振龙^[4]采用演化博弈建立了驾驶员在诱导信息下的路径选择模型;鲁丛林^[5]讨论了在诱导信息条件下驾驶员的反应行为博弈模型,黄海军^[6]、秦进^[7]等通过解析的方法对诱导信息的有效性进行了评价;刘建美^[8-10]提出了驾驶员在有限理性的条件下车辆的路径选择博弈模型.但是,由于驾驶员本身的局限性,在有限理性的前提下,驾驶员的选择往往是根据自身上一次选择所获取的利益来决定自己下一次的选择,即自学习机制.在行为强化理论^[11]的基础上,以模糊数学和博弈论为工具,讨论了诱导信息条件下车辆的路径选择模型,分析给出了博弈平衡结果,以期能为交通诱导提供理论依据。

1 模型建立

1.1 模型基本假设

出行时间的长短是影响路径选择最重要的标准^[12],美国联邦总局(BPR)提出路段行程时间函数^[13]:

$$t = T \left[1 + \alpha \left(\frac{q}{c} \right)^\gamma \right] \quad (1)$$

其中: T 表示自由行驶时(交通量为零)的路段行程时间, c 为路段通行能力, q 为路段实际交通量, α, γ 为参

收稿日期:2015-06-10;修回日期:2015-08-07.

作者简介:周代平(1990-),男,湖南人,硕士,从事交通运输规划研究.E-mail: zhoudaiping2015@sina.com.

数,一般取 $\alpha=0.15, \gamma=4$.

以驾驶员的实际行驶时间作为驾驶员选择某条路径所获得的收益.考虑如图 1 所示的简单路网, A 到 B 地有 L_1, L_2 两条路径, L_1, L_2 的道路通行能力分别为 C_1, C_2 , 表示驾驶员第 k 次通过 L_1, L_2 的实际交通量, $t_{1,k}$, $t_{2,k}$ 为车辆第 k 次通过 L_1, L_2 到达 B 点的实际通行时间, t_0 为驾驶员从 A 地开往 B 地的期望时间.

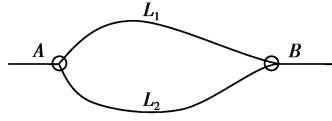


图 1 路网示意图

1.2 满意度隶属函数

满意度是指驾驶员对从 A 地开往 B 地所花实际时间的满意程度,它是个模糊的概念.取论域 $U=(0, +\infty)$, $u \in U$, 模糊集 A_1, A_2, A_3 分别表示“满意”、“一般”、“差”, 则它们的隶属函数分别为^[14]

$$A_1(u) = \begin{cases} 1 & 0 < u \leq t_0 \\ 1 - 2\left(\frac{u - t_0}{t_0}\right)^2 & t_0 < u \leq \frac{3}{2}t_0 \\ 2\left(\frac{u - \frac{3}{2}t_0}{t_0}\right)^2 & \frac{3}{2}t_0 < u \leq 2t_0 \\ 0 & 2t_0 < u < +\infty \end{cases} \quad (2)$$

$$A_2(u) = \begin{cases} 0 & 0 < u \leq t_0 \\ 2\left(\frac{u - t_0}{t_0}\right)^2 & t_0 < u \leq \frac{3}{2}t_0 \\ 1 - 2\left(\frac{u - 2t_0}{t_0}\right)^2 & \frac{3}{2}t_0 < u \leq 2t_0 \\ 1 & 2t_0 < u \leq \frac{5}{2}t_0 \\ 1 - 2\left(\frac{u - \frac{5}{2}t_0}{t_0}\right)^2 & \frac{5}{2}t_0 < u \leq 3t_0 \\ 2\left(\frac{u - \frac{7}{2}t_0}{t_0}\right)^2 & 3t_0 < u \leq \frac{7}{2}t_0 \\ 0 & \frac{7}{2}t_0 < u < +\infty \end{cases} \quad (3)$$

$$A_3(u) = \begin{cases} 0 & 0 < u \leq \frac{5}{2}t_0 \\ 2\left(\frac{u - \frac{5}{2}t_0}{t_0}\right)^2 & \frac{5}{2}t_0 < u \leq 3t_0 \\ 1 - 2\left(\frac{u - \frac{7}{2}t_0}{t_0}\right)^2 & 3t_0 < u \leq \frac{7}{2}t_0 \\ 1 & \frac{7}{2}t_0 < u < +\infty \end{cases} \quad (4)$$

若 $\text{Max}(A_1(t), A_2(t), A_3(t)) = A_1(t)$, 则 $t \in A_1$, 驾驶员对实际驾驶时间 t 感觉“满意”, 若 $\text{Max}(A_1(t), A_2(t), A_3(t)) = A_2(t)$, 则 $t \in A_2$, 驾驶员对实际驾驶时间 t 感觉“一般”, 若 $\text{Max}(A_1(t), A_2(t), A_3(t)) = A_3(t)$, 则 $t \in A_3$, 驾驶员对实际驾驶时间 t 感觉“差”。

1.3 模型建立

Skinner 的“操作条件反射”理论认为,人或动物为了达到某种目的,会采取一定的行为作用于环境,当这种行为的后果对他有利时,这种行为就会在以后重复出现;不利时,这种行为就减弱或消失.人们可以用这种正强化或负强化的办法来影响行为的后果,从而修正其行为^[15].在行为强化理论的自学习机制下,驾驶员的第 $k+1$ 次是否接受诱导取决于驾驶员的第 k 次接受(或不接受)诱导所获得的收益;当驾驶员对第 k 次接受(不接受)诱导所获得的收益 t_0 为满意时,则第 $k+1$ 次继续选择接受(不接受)诱导,当驾驶员对第 k 次接受(不接受)诱导所获得的收益 t_0 满意度为差时,则第 $k+1$ 次选择不接受(接受)诱导,当驾驶员第 k 次接受(不接受)诱导所获得的收益 t_0 满意度为一般时,则驾驶员第 $k+1$ 次将会有 ξ 概率选择不接受(接受)诱导.设每次参与博弈的车辆总数一定,即 $q_{y,k} + q_{n,k} = Q$, Q 为参与博弈的车辆总数.根据驾驶员路径选择的实际情况,可将驾驶员分为两大类 q_y (接受诱导的车辆)和 q_n (不接受诱导的车辆),设 $q_{y,k}$ 为第 k 次接受诱导的车辆总数, $q_{n,k}$ 为第 k 次不接受诱导的车辆总数, $k=1, 2, 3, \dots$, 则此博弈模型的战略式表达如下:

1) 局中人:

$$G_i \in (q_y, q_n), i = 1, 2, 3, \dots, Q$$

2) 局中人的策略集:

$$S_i \in \{\text{接受诱导}(Y), \text{拒绝诱导}(N)\}, i = 1, 2, 3, \dots, Q$$

3) 局中人的收益函数:

$$U_i = T(1 + \delta(\frac{q_{j,k}}{c_\delta})^\gamma), i = 1, 2, \dots, Q; j = y, n; \delta = 1, 2$$

2 模型求解算法

Step 0: $h=h+1, v=1, m=0.1$, 给 $C_1, C_2(C_1 > C_2), T, \xi$ 赋值(定值);

Step 1: 初始化, 给出模型的初始状态, $Q=q_0, q_{y,1} = m * Q, q_{n,1} = Q - q_{y,1}, fabu(1) = g, (fabu(i) = g$ 表示第 i 次诱导信息建议车辆走路径 $L_g, g=1, 2, m$ 表示初始接受诱导的车辆比例);

Step 2: 若 $fabu(i) = 1$, 则 $q_{1,i} = q_{y,i}, q_{2,i} = q_{n,i}$, 若 $fabu(i) = 2$, 则 $q_{1,i} = q_{n,i}, q_{2,i} = q_{y,i}$;

Step 3: 根据时间行程函数(式 1)求得选择路径 L_1, L_2 的通过时间 $t_{1,i}, t_{2,i}$;

Step 4: 根据隶属度函数, 求得 $t_{1,i}, t_{2,i}$ 的时间感受隶属度, 判断驾驶员对第 i 次选择收益的满意度;

Step 5: 由第 i 次的收益满意度确定第 $i+1$ 次的选择是否接受诱导;

Step 6: 比较 $q_{y,i+1}$ 和 $q_{n,i+1}$ 的大小, 若 $q_{y,i+1} > q_{n,i+1}$, 则 $fabu(i+1) = 1$; 若 $q_{y,i+1} < q_{n,i+1}$ 则 $fabu(i+1) = 2$;

Step 7: 根据 $fabu(i+1)$ 的值确定 $q_{1,i+1}, q_{2,i+1}$; 若 $fabu(i+1) = 1$, 则 $q_{1,i+1} = q_{y,i+1}, q_{2,i+1} = q_{n,i+1}$; 若 $fabu(i+1) = 2$, 则 $q_{1,i+1} = q_{n,i+1}, q_{2,i+1} = q_{y,i+1}$;

Step 8: $i = i + 1$, 若 $i < 100$, 转 Step 2; 否则转 Step 9;

Step 9: 若 $m < 1$, $Z(h, v) = q_{1,100}, m = m + 0.1, v = v + 1$ 转 Step 1, 否则转 Step 10;

Step 10: $q_0 = q_0 + \Delta q$, 若 $q_0 < 3 * (C_1 + C_2)$, 转 Step 0, 否则结束.

3 具体算例

算例中, $C_1 = 1500, C_2 = 1000, T = 30, \xi = 0.333^{[16]}, q_0 = 1000, \Delta q = 500$; 仿真结果如图 2 所示。仿真结果表明: 在诱导信息条件下, 不同的路网车流总量经过若干次博弈后都会达到稳定平衡; 即若干次博弈后, 路径 L_1, L_2 上的量都呈现稳定状态; 路网车流总流量越接近路网总通行能力, 初始接受诱导的车辆比例 m 对博弈平衡结果影响越显著; 参与博弈车辆数远大于路网总通行能力时, 通行能力较大的路径 (L_1) 更为拥堵; 在路网总流量一定时, 存在一个初始接受诱导的车辆比例 m 使得路网交通流分配达到最佳。

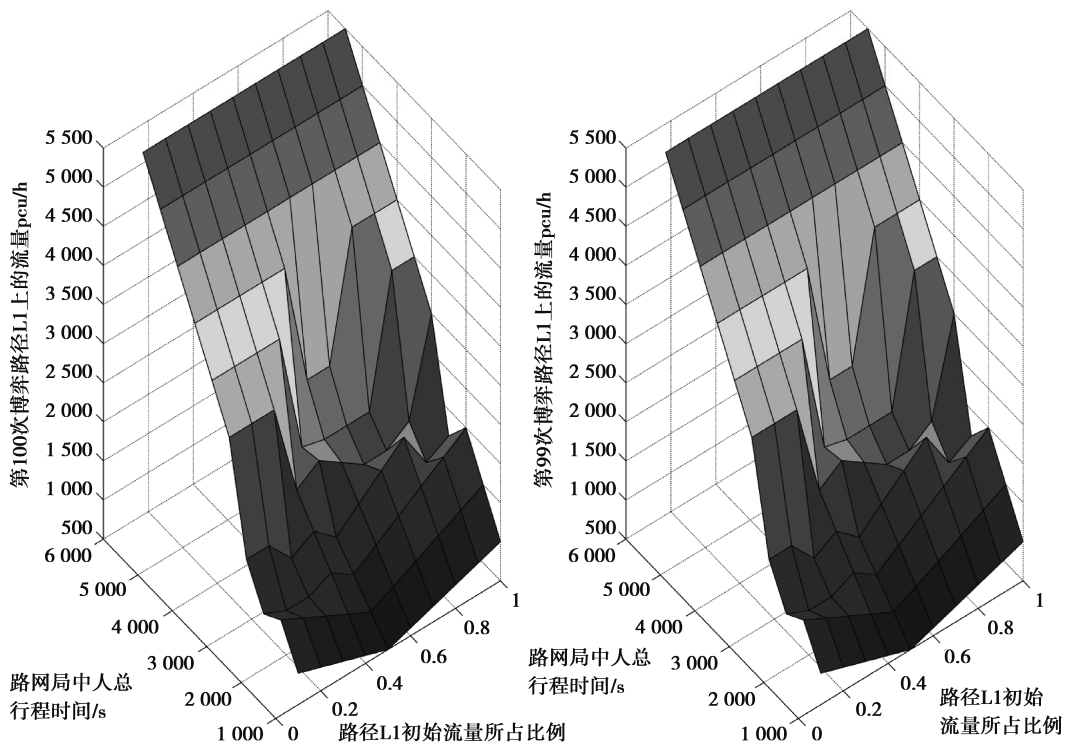


图 2 仿真结果

4 结 语

在考虑驾驶员的有限理性的前提下, 以行为强化理论为基础, 建立了基于自学习机制的诱导信息条件下的车辆路径选择模型, 并给出了算法和具体算例。仿真结果表明: 诱导信息的发布并不一定能改善路网交通流分布, 它与路网总流量以及初始接受诱导的车辆比例密切相关。

参考文献:

- [1] ROTH A E, EREV I. Learning in Extensive-form Games; Experimental Data and Simple Dynamic Models in Intermediate Term[J]. Nobel Symposium. Games and Economic Behavior, 1995(8):164-212
- [2] EREV I, BEREBY, ROTH A. The Effect of Adding a Constant to All payoffs; Experimental Investigation and Implications for Reinforcement Learning Models[J]. Journal of Economic Behavior and Organization, 1999, 39:111-128
- [3] BELL M G. A Game Theory Approach to Measuring the Performance Reliability of Transport Networks[J]. Transportation Research Part B, 2000, 34(6):533-545
- [4] 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
- [5] 鲁丛林. 诱导条件下的驾驶员反应行为的博弈模型[J]. 交通运输系统工程与信息, 2005, 5(1):58-61
- [6] 黄海军, 吴文洋. 交通信息对交通行为影响的评价模型[J]. 系统工程理论与实践, 2002, 22(10):81-83
- [7] 秦进, 黎新华. 交通信息的有效性研究[J]. 公路交通科技, 2005, 22(2):104-107
- [8] 刘建美, 马寿峰. 基于有限理性的个体出行路径选择进化博弈分析[J]. 控制与决策, 2009, 24(10):1450-1454
- [9] LIU J M, MA S F. A Dimension-reduced Method of Sensitivity Analysis for Stochastic User Equilibrium Assignment Model[J]. Applied Mathematical Modelling, 2010(34):325-333
- [10] LIU J M, MA S F. Algorithms of Game Models on Individual Travel Behavior[C] // The 8th International IEEE Conference of Chinese logistics and transportation professionals, Chengdu, China, 2008
- [11] SKINNER. The Behavior of Organisms An Experimental Analysis[M]. New York: Appleton Century Company, 1938:110-150
- [12] OUTRAM V. E, THOMPSON. Driver Route Choice[C]. Proceedings-PTRC Annual Meeting, London, 1977
- [13] 杨佩坤, 钱林波. 交通分配中路段行程时间函数研究[J]. 同济大学学报:自然科学版, 1994, 22(1):41-44
- [14] 杨纶标, 高英仪, 凌卫新. 模糊数学原理及应用[M]. 广州:华南理工大学出版社, 2011
- [15] SKINNER B F. The Behavior of Organisms; An Experimental Analysis[M]. New York: Appleton-Century Company, 1938
- [16] 周元峰. 基于信息的驾驶员路径选择行为及动态诱导模型研究[D]. 北京:北京交通大学, 2007

Research on the Model of Vehicle Routing Choice Based on the Condition of the Bounded Rationality Fuzzy Game with Inducing Information

ZHOU Dai-ping, LI Kang-qi, HE Lin

(School of Transportation and Traffic, Chongqing Jiaotong University, Chongqing 400074, china)

Abstract: The inducing information belongs to recommended information, so drivers do not necessarily accept the induction after the inducing information released. This paper, based on the theory of reinforcement, establishes the model of vehicle routing choice based on the condition of fuzzy game with inducing information and obtains the result of balance of the game under the different initial state through simulation. The simulation results show that the release of the induction information does not necessarily improve the road network traffic flow distribution, which is closely related to the proportion of the total flow of the road network and the initial acceptance of the induction vehicles.

Key words: fuzzy game; bounded rationality; traffic induction; self-learning mechanism