

信息条件下基于预期后悔更新生成 备选路径仿真研究*

刘新全

(广西师范学院 经济管理学院, 南宁 530001)

摘要: 为了研究影响私家车驾驶者备选路径生成的因素,以预期后悔理论为基础,借助贝叶斯网络推理方法,计算了私家车驾驶者受先验知识和出行信息双重影响下的备选路径生成。通过改进的贝叶斯网络结构和参数学习程序建立了实验路网的贝叶斯网络结构,对生成的网络结构进行了参数学习,模拟了不同出行信息和先验条件下生成的备选路径,得出了驾驶者备选路径生成与驾驶者先验知识和出行信息的变化关系。

关键词: 预期后悔; 备选路径; 仿真; 私家车; 贝叶斯网络推理

中图分类号: TP391 **文献标志码:** A **文章编号:** 1001-3695(2012)01-0152-03

doi:10.3969/j.issn.1001-3695.2012.01.043

Research on simulation of generation of candidate routes based on updated anticipated regret under condition of information

LIU Xin-quan

(College of Economics & Management, Guangxi Teachers Education University, Nanning 530001, China)

Abstract: In order to study the factors to affect the generation of candidate routes, this paper introduced the Bayesian network inference method to solve the generation of candidate routes under the influence of pre-knowledge and travel information based on anticipated regret. It updated the programs of Bayesian network structure and parameter learning to establish the experiment route network and parameter learning. The generated routes were simulated and explained under the condition of different travel information and pre-experiences. It presented the relationship between the generation of candidate routes and pre-experience and travel information.

Key words: anticipate regret; candidate routes; simulation; private cars; Bayesian network inference

城市群物流效率的提高对实现城市群区域经济快速发展起到了重要作用,而影响物流配送效率的关键是车辆路径问题,研究车辆路径不仅能够提高物流企业的经济效益,而且对改善整个城市群网络效率具有重要意义。对城市群道路网络产生较大影响的一个主要因素是私家车驾驶者的路径选择,其出行行为会对物流的配送方案产生影响。

以往文献对私家车驾驶者出行信息的分析主要集中在驾驶者根据出行信息选择效用最大的路径或出行方式,也有文献分析了出行信息对出行终点选择的影响^[1],而出行信息对驾驶者预期后悔的影响研究较少,基于信息的预期后悔更新对驾驶者备选路径生成的影响研究还未见到。因此,为了分析基于后悔理论的混合效用模型对驾驶者出行选择行为的影响进而对物流配送方案生成的影响,有必要分析出行知识和出行信息对预期后悔的影响,进而对备选路径生成的影响。

1 预期后悔理论

Bell^[2]提出的预期后悔理论就是将决策者的情感心理因素与动机因素合并到期望效用的结构中,用来评价决策者对未来所选选项结果的预期反应。因此预期的情绪将影响效用函数,对决策者的决策行为产生影响,这种影响反映在决策结果

选项的改变上。预期后悔表现了决策者的决策目的不是为了追求决策结果最优,而是避免产生使自己后悔的结果。Zeelenberg等人^[3]认为预期后悔使决策者表现出的行为不是风险规避,而是后悔规避,决策者进行决策时对选择项的期望是预期后悔最小,而不是风险最小。相关文献对预期后悔的研究集中在消费者进行商品选择^[4-6]、金融投资^[7,8]方面,并取得了较好的预测效果;预期后悔理论在交通领域的应用较少,还没有相关文献对预期后悔理论在卫星城驾驶者的出行行为方面进行研究。

由于驾驶者在经历从卫星城到中心城的行程中,可能会遇到交通拥堵影响行程时间,并且对出现拥堵的路段和时间无法准确判断。因此驾驶者出行前会对选择相关路段出行可能引起的后悔进行估计,即形成预期后悔。

2 贝叶斯网络结构建立与参数学习

2.1 贝叶斯网络结构建立

驾驶者结合先前知识与当前出行信息对路段属性重新认识的过程,可以看做是驾驶者的一个推理过程,因此可以借助贝叶斯网络推理来完成。

贝叶斯网络结构学习是指利用训练样本集,结合合理的先验知识,确定合适的贝叶斯网络拓扑结构,从而为贝叶斯网络参数的学习创造条件^[9]。文献[9,10]分别使用贝叶斯网络化概率表达式,建立了各种出行行为之间以及出行行为与影响因素之间贝叶斯网络结构图,用于理解居民的出行决策过程,然而并未对生成路网过程分析。本文主要分析基于知识与出行信息的预期后悔更新对备选路径生成的影响,在给定路段知识即先验信息下,借助贝叶斯网络推理方法分析路网中路段属性变化情况;以此为基础分析驾驶者预期后悔的变化以及对备选路径生成的影响。图 1 为实验路网。

为了分析贝叶斯网络结构,给出如下定义:

x_a 表示路网中的路段, X 表示卫星城与中心城间的路段集合,其中 $a=1, \dots, n, x_a \in X$; T_a 表示每条路段的属性值域,表示为 $T_a = \{t_a^1, \dots, t_a^i, \dots, t_a^m\}$; 驾驶者估计的每条路段可能的行程时间与该路段的前一路段的行程时间存在一定的联系,使用贝叶斯网络描述为变量 x_a 属性值的变化依赖于直接的父节点,与其他节点是条件独立的,因此变量 x_a 对应的联合概率分布表达式为

$$P(x_1, \dots, x_a, \dots, x_n) = \prod_a P(x_a | x_1, \dots, x_{n-1}) = \prod_a P(x_a | pa(x_a)) \quad (1)$$

其中: $pa(x_a)$ 为网络结构中路段 x_a 的所有父路段的集合, $p(x_a | pa(x_a))$ 表示路段 x_a 在其父路段取某一属性值的条件概率分布。

应用贝叶斯网络结构学习方法,结合驾驶者的先验知识,通过贝叶斯网络结构学习,得出了卫星城与中心城之间以路段为变量的出行网络结构,如图 2 所示。

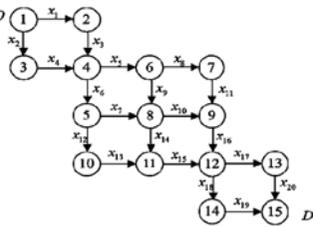


图1 实验路网

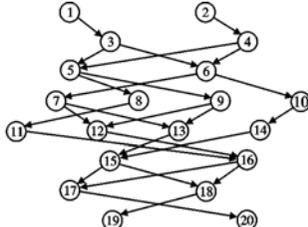


图2 以路段为变量的贝叶斯网络出行结构

2.2 基于路段的贝叶斯网络参数学习

贝叶斯网络参数学习指的是在给定贝叶斯网络拓扑结构的情况下,确定各节点处的条件概率密度。这里结合出行选择路网,将参数学习理解为:利用信息条件下给定的路段属性作为样本数据去学习网络参数的概率分布,从而更新网络中路段原有属性的先验分布,得到后验。假设根据信息得到的样本数据是完整的,那么可以使用最大似然法和最大后验概率法来进行贝叶斯网络参数学习。根据上文的定义,可知路段 x_a 为贝叶斯网络结构中的节点,假设 x_a 的取值为 $T_a = \{t_a^1, \dots, t_a^i, \dots, t_a^m\}$; $pa(x_a)$ 为节点 x_a 所有父路段构成的集合,在集合 $T_{pa(x_a)} = \{t_1^{pa(x_a)}, t_k^{pa(x_a)}, \dots, t_l^{pa(x_a)}\}$ 中取值。

根据贝叶斯网络模型,路段 x_a 对应的条件概率表示为

$$\theta_{aki} = P(x_{ai} = t_a^i | pa(x_a) = t_k^{pa(x_a)}) \quad (2)$$

其中: θ_{aki} 表示路段 x_a 的父路段取得第 k 个状态值 $t_k^{pa(x_a)}$, 路段 x_a 取值 t_a^i 时的概率,且有 $0 < \theta_{aki} < 1, \sum_{i=1}^m \theta_{aki} = 1$ 。

给定样本数据集 $Y = \{Y_1, \dots, Y_j, \dots, Y_n\}$, 样本数据集的数目为 Z , 则贝叶斯网络模型中各路段的条件概率密度表示为 h

$(\theta_a | Y)$, 其中^[11]:

$$\theta_a = \prod_k \theta_{ak} = \{\theta_{aki}, \dots, \theta_{aki}, \dots, \theta_{ikm}\} \quad (3)$$

研究驾驶者预期后悔更新对备选路径生成影响的过程可以分为以下几个步骤:

a) 将未知参数 θ_a 看做是随机变量的条件概率,并且对 θ_a 有一定先验知识,这种先验知识用 θ_a 的某种概率分布 $h(\theta_a)$ 表达出来,反映了在属性更新前对参数 θ_a 所拥有的信息。

b) 将样本数据集 $Y = \{Y_1, Y_2, \dots, Y_n\}$ 看做关于路段的出行信息,出行信息 Y 的边缘分布密度 $p(Y)$ 可以表示为

$$p(Y) = \int_{X^p} p(\theta_a, Y) d\theta_a = \int_{X^p} p(Y | \theta_a) h(\theta_a) d\theta_a \quad (4)$$

c) 根据贝叶斯定理,在给定出行信息 $Y = \{Y_1, Y_2, \dots, Y_n\}$ 的条件下, θ_a 的条件概率密度为

$$h(\theta_a | Y) = \frac{P(Y | \theta_a) h(\theta_a)}{P(Y)} \quad (5)$$

d) 式(5)称为 θ_a 的后验密度,利用后验分布 $h(\theta_a | Y)$ 对参数 θ_a 进行学习。对参数 θ_a 的学习使用贝叶斯方法^[12]:关于网络中路段的先验知识和得到的出行信息数据,对给定的交通出行路网来说就是借助出行信息,学习参数 θ 的所有可能取值,即计算后验分布 $h(\theta | Y)$ 。

3 参数学习算法结果分析

为了便于分析网络参数,给出贝叶斯网络结构的知识(感知路段平均行程时间)和先验信息(驾驶者感知的路段行程时间的可靠性),如表 1 和 2 所示。

表 1 驾驶者出行知识 t_a^i

a	i		a	i	
	1	2		1	2
1	12	14	11	28	31
2	10	13	12	47	52
3	13	14	13	17	20
4	12	14	14	50	53
5	17	20	15	28	32
6	40	42	16	52	54
7	19	23	17	10	14
8	26	29	18	14	17
9	65	70	19	17	20
10	36	39	20	10	13

表 2 驾驶者先验信息 θ_{aki}

a	i	k			
		1	2	3	4
1	i=1	0.8			
	i=2	0.2			
2	i=1	0.9			
	i=2	0.1			
3	i=1	0.8	0.6		
	i=2	0.2	0.4		
4	i=1	0.8	0.7		
	i=2	0.2	0.3		
5	i=1	0.9	0.8	0.6	0.2
	i=2	0.1	0.2	0.4	0.8
6	i=1	0.9	0.8	0.6	0.3
	i=2	0.1	0.2	0.4	0.7
7	i=1	0.7	0.3		
	i=2	0.5	0.5		
8	i=1	0.85	0.55		
	i=2	0.15	0.45		

续表 2

9	$i=1$	0.7	0.2		
	$i=2$	0.3	0.8		
10	$i=1$	0.75	0.3		
	$i=2$	0.25	0.7		
11	$i=1$	0.8	0.4		
	$i=2$	0.2	0.6		
12	$i=1$	0.8	0.75	0.4	0.1
	$i=2$	0.2	0.25	0.6	0.9
13	$i=1$	0.9	0.8	0.3	0.3
	$i=2$	0.1	0.2	0.7	0.7
14	$i=1$	0.5	0.3		
	$i=2$	0.5	0.7		
15	$i=1$	0.9	0.6	0.4	0.1
	$i=2$	0.1	0.4	0.6	0.9
16	$i=1$	0.8	0.7	0.5	0.3
	$i=2$	0.2	0.3	0.5	0.7
17	$i=1$	0.9	0.7	0.3	0.2
	$i=2$	0.1	0.3	0.7	0.8
18	$i=1$	0.85	0.7	0.3	0.2
	$i=2$	0.15	0.3	0.7	0.8
19	$i=1$	0.95	0.4		
	$i=2$	0.05	0.6		
20	$i=1$	0.9	0.2		
	$i=2$	0.1	0.8		

根据驾驶者的先验,结合贝叶斯网络工具箱,运用 MATLAB 7.0 编制驾驶者的参数学习程序,在 CPU 1.6 GHz,内存 2 GB 的计算机上计算得出了参数在不同出行信息量 (smpls) 和先验不同权重 ($p = priors$) 下的学习结果,如图 3~8 所示。

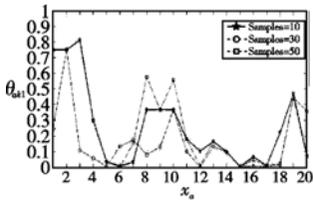


图3 $i=1, p=1$ 学习结果变化图

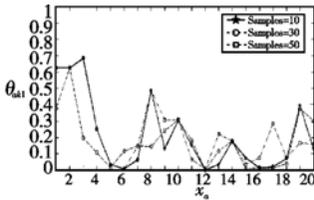


图4 $i=1, p=3$ 学习结果变化图

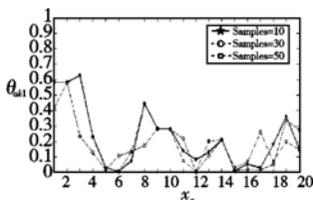


图5 $i=1, p=5$ 学习结果变化图

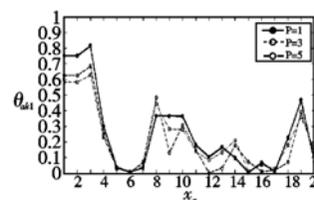


图6 $i=1, samples=10$ 属性值变化图

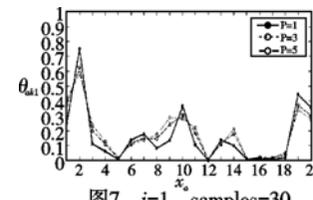


图7 $i=1, samples=30$ 属性值变化图

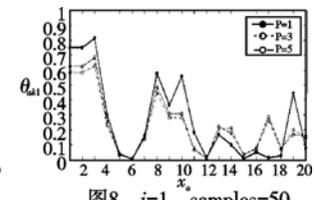


图8 $i=1, samples=50$ 属性值变化图

图 3~5 反映了出行信息与先验信息对驾驶者感知路段取得第一个属性值的肯定度的影响。从图中可以看出,当先验信息的权重值较低时,随着出行信息的增加,驾驶者选择路段第一个属性值的可能性变化较大,说明出行信息对出行者的决策影响较大;当先验权重值较高时,出行信息的变化对属性值的影响较小,说明驾驶者受先验的影响较大。

图 6~8 反映了随着先验权重的增加,驾驶者感知路段第一个属性值的变化不明显,出行信息增加时某些路段的属性值会发生变化,然而在同一数量出行信息的情况下,驾驶者对路

段属性值的感知相对较稳定。

4 基于学习结果的备选路径生成

为了便于分析生成备选路径在不同信息状态和不同先验权重情况下的变化,根据不同出行信息以及不同备选路径中路段预期后悔值的变化,绘制在先验权重不同状态下的备选路径变化图,如图 9~11 所示。

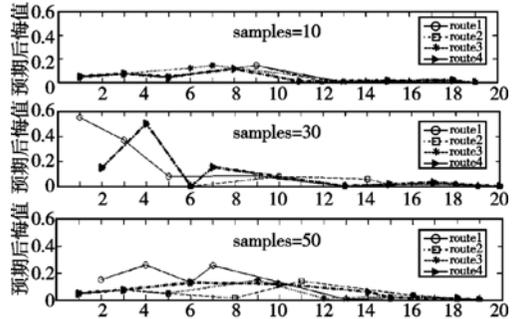


图9 $p=1$ 时备选路径变化分析

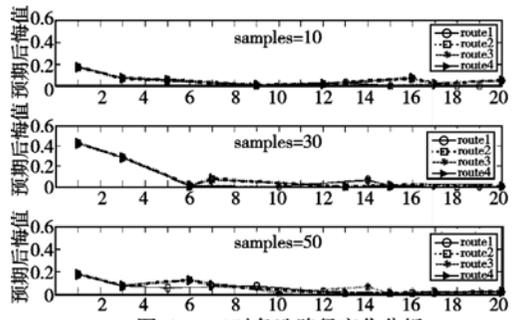


图10 $p=3$ 时备选路径变化分析

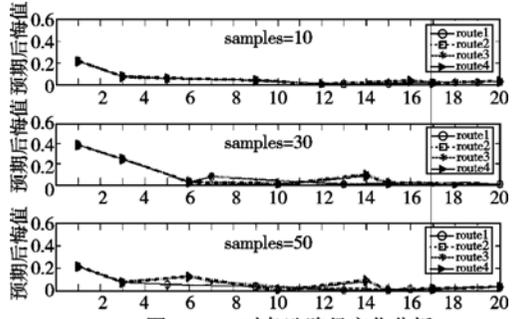


图11 $p=5$ 时备选路径变化分析

图 9、11 说明了随着先验在驾驶者路径生成中的权重逐渐增加,出行信息对驾驶者路径生成的影响力逐渐降低,此时驾驶者主要依赖先验生成备选路径集合,并且生成的备选路径差别不大;由于受到先验的影响,驾驶者对路径中路段感知的预期后悔也较低,此种情况出现在驾驶者对于备选路径十分熟悉,并且对路径中每条路段的路况也十分了解。

5 结束语

由于城市群物流配送方案的形成受到道路网络中其他车辆路径选择的影响,尤其是私家车驾驶者的出行路径选择行为。因此,本文在以往文献基础上,对私家车驾驶者出发前在信息条件下生成的备选路径进行研究,从而为进一步研究私家车驾驶者的路径选择行为奠定基础。本文从驾驶者心理角度出发,研究驾驶者对路段属性感知的变化以及这种变化对备选路径生成的影响,验证了基于预期后悔更新的备选路径生成方法更加符合实际。

4 实验结果与分析

本文采用中国科学院计算技术研究所中文自然语言处理开发平台发布的由复旦大学计算机信息与技术系国际数据库中心自然语言处理小组所提供的中文语料库,对文本分类器进行分类测试。对 581 个文本样本进行信息编码,得到 10 维文本的信息编码向量 581 个,其中 140 个作为训练样本,其余 441 个作为测试样本,在 MATLAB 环境下分别进行 BP 和 RBF 神经网络的分类算法实现;再利用 K-means 聚类方法作为 RBF 神经网络分类算法的核心思想,进行 RBF 分类,并与 BP 分类算法比较。文本分类系统的最主要的两个指标是查准率和查全率,查准率(精度)是衡量检索系统信号噪声比的一种指标,即检出的相关文献与检出的全部文献的百分比,表示为:查准率 = (检索出的相关信息/检索出的信息总量) × 100%。使用泛指性较强的检索语言(如上位类、上位主题词)能提高查全率,但查准率下降。查全率(召回率),是衡量某一检索系统从文献集中检出相关文献成功度的一项指标,即检出的相关文献与全部相关文献的百分比,表示为:查全率 = (检索出的相关信息/系统中的相关信息总量) × 100%。使用专指性较强的检索语言(如下位类、下位主题词)能提高查准率,但查全率下降。

所谓的分类正确就是指自动分类结果与人工分类结果吻合。取其中三类的统计结果如表 1 所示。

表 1 三种网络分类算法统计结果

网络	查全率			查准率			训练时间
	0.1	0.2	0.3	0.1	0.2	0.3	/s
BP	0.9492	0.9444	0.9574	1.0000	0.9714	0.9783	22.782
RBF	0.8305	0.8194	0.6170	0.8909	0.8676	0.9667	19.156
RBF+聚类	0.8136	0.7639	0.5319	0.8889	0.8871	0.8929	0.890

从表 1 中可以看出,RBF 网络分类结果不如 BP 网络的好,但使用给定样本特征值的平均值作为聚类中心的径向基神经网络分类结果较好,且训练时间很短。这证明了本文提出的算法要比传统的神经网络文本分类算法要好得多。对以上三类样本在利用结合聚类的 RBF 文本分类算法时,修改高斯函数宽度参数,不同高斯函数宽度参数对实验结果的影响如图 2

所示。

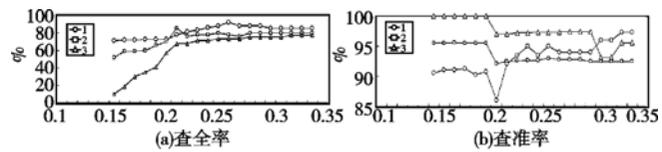


图 2 不同高斯函数宽度参数对结果的影响

5 结束语

本文以神经网络为基础,借鉴了聚类算法的思想,采用样本中心作为 RBF 分类算法的核心,并与 BP 神经网络分类算法进行了比较。从实验结果可以看出,在收敛速度和分类效果上,结合聚类的 RBF 文本分类算法好于 BP 神经网络分类算法,充分体现了改进后 RBF 分类算法的简洁和时效性。径向基函数的宽度参数会影响分类的准确度和实验误差,查全率和查准率不能同时提高,随着宽度参数的增大,误差会变小。结合了聚类思想的 RBF 文本分类算法有一定的理论研究价值和实际应用前景。

参考文献:

- [1] 包剑,冀明,冯军. 基于模糊支持向量机的文本分类[J]. 辽宁工程技术大学学报:自然科学版,2010,29(5):974-977.
- [2] 刘金岭,严云洋. 基于上下文的短信文本分类方法[J]. 计算机工程,2011,10(2):511-514.
- [3] 史晶蕊,郑玉明,韩希. 神经网络在文本分类中的应用[J]. 计算机应用研究,2005,22(10):218-221.
- [4] DEBOLE F, SCBASTIANI F. An analysis of the relative hardness of recuters-21578 subsets [J]. Journal of the American Society for Information Science and Technology,2004,56(6):584-596
- [5] AHN B S, CHO S S, KIM C. The integrated methodology of rough set theory and artificial neural network for business failure prediction [J]. Expert Systems with Applications, 2000,18(2):65-74.
- [6] 熊忠阳,黎刚,陈小莉,等. 文本分类中词语权重计算机方法的改进与应用[J]. 计算机工程与应用,2008,44(5):187-190.
- [7] 张玉芳,彭时名,吕佳. 基于文本分类 TFIDF 方法的改进与应用 [J]. 计算机工程,2006,32(9):76-78.

(上接第 154 页)

参考文献:

- [1] 李志纯,黄海军. 先进的旅行者信息系统对出行者选择行为的影响研究[J]. 公路交通科技,2005,22(2):95-99.
- [2] BELL D E. Regret in decision-making under uncertainty[J]. Operations Research,1982,30(5):961-981.
- [3] ZEELENBERG M, PIETERS R. Comparing service delivery to what might have been: behavioral responses to disappointment and regret [J]. Journal of Service Research,1999,2(1):86-97.
- [4] COOKE A D J, MEYVIS T, SCHWARTZ A. Avoiding future regret in purchase-timing decisions[J]. Journal of Consumer Research, 2001,27(4):447-459.
- [5] SIMONSON I. The influence of anticipating regret and responsibility on purchase decisions[J]. Journal of Consumer Research,1992,19(1):105-118.
- [6] ZEELENBERG M, BEATTIE J. Consequences of regret aversion 2: additional evidence for effects of feedback on decision making [J]. Or-

ganizational Behavior and Human Decision Processes,1997,72(1):63-78.

- [7] SHEFRIN H, STATMN M. The disposition to sell winners too early and ride losers too long [J]. Theory and Evidence Journal of Finance,1985,40(3):777-790.
- [8] 李新路. 后悔厌恶心理对投资者行为影响的实证分析[J]. 河北经贸大学学报,2006,27(2):40-44.
- [9] TORRES F J, HUBER M. Learning a causal model from household survey data using a Bayesian belief network [J]. Transportation Research Board Annual Meeting,2003,1836:29-36.
- [10] JANSSENS D, WETS G, BRIJS T, et al. Improving the performance of a multi-agent rule-based model for activity pattern decisions using Bayesian networks [J]. Journal of the Transportation Research Board,2004,1894(1):75-83.
- [11] 李晓毅,徐兆棣,孙笑微. 贝叶斯网络的参数学习研究[J]. 沈阳农业大学学报,2007,38(1):125-128.
- [12] 林元烈,梁宗霞. 随机数学引论[M]. 北京:清华大学出版社,2003:331-343.