

基于粗糙集理论的交通状态区域划分算法

孙立群 谢岩峰 张良智

山东交通学院顿河学院 山东济南 250023

【摘要】针对城市交通路网区域划分问题,本文选择交通状态条件变量和决策变量,进行变量模糊化和属性编码,建立决策表,借助粗糙集理论,进行属性约简,建立了模糊规则,进而提出了基于粗糙集的交通子区划分算法。利用本算法在Nguyen-Dupuis路网中进行了模拟试验,结果证明在不同交通状态下可有效开展交通区域划分,为下一步开展区域控制提供基础保障。

【关键词】粗糙集,交通状态,区域划分

The division algorithm of traffic state regions based on coarse set theory

Sun Liqun Xie Yanfeng Zhang Liangzhi

Dunhe College, Shandong Communications University Jinan City 250023

【Abstract】In view of the problem of area division of urban traffic network, this paper selects the traffic state condition variables and decision variables, conducts variable blur and attribute coding, establishes decision table, with the help of rough set theory, attribute reduction, establishes fuzzy rules, and then puts forward the traffic sub-area division algorithm based on coarse set. Using the simulation test in Nguyen-Dupuis road network, the results prove that the traffic area division can be effectively developed in different traffic conditions, providing basic guarantee for the next regional control.

【Key words】rough set, traffic state, regional division

1.引言

在对城市道路进行控制和管理中,早已超越单点控制、干线控制阶段,发展到区域控制层面,因此区域控制近年来是城市交通领域的研究热点。但相对于单点控制和线性控制,区域控制具有明显的界限不清、控制对象不明确的先天不足,即将哪些交叉口和道路划分为控制区域,至今没有公认统一的方案。不少学者在这个问题上开展了大量研究工作。

在划分区域研究中,路口关联度关系是重要研究方向之一。关联度是定量描述两个交叉口关联程度的指标,也是评价交叉口是否归属同一区域中的重要依据。Whitson模型作为经典关联度计算方法,利用交叉口距离、路段路口交通量、行程时间等变量来构建关联度模型,但没有考虑动态变化问题。杨迪等提出改进的密度峰值聚类算法,结合相邻交叉口之间的关联度对路网区域进行划分。侯军军等人首先对整个城市路网进行网格化划分,然后计算每个网格的交通指数,得到样本特征矩阵;接着利用k-means聚类算法进行聚类,并对其中奇异网格标签加以修正。

城市道路交通状态是瞬息万变的,且交通状态的划分具有一定的相对性,即同样的交通参数(流量、速度或密度)在不同的交通环境中,评价状态是不同的。考虑到状态划分具有一定的主观评价因素,本文引入粗糙集理论进行模糊分类评价,为交通管理决策者提供多维度的划分标准,更全面开展交通管理和控制。

2.粗糙集理论

2.1 不可分辨关系与隶属函数

粗糙集理论中,知识系统可表达为 $S = \{U, A, V, f\}$ 的基本形式。其中论域 U 是有限非空集合, A 是知识属性集合; V 是属性 $a \in A$ 的值域; $f: U \rightarrow V$ 为对象的论域到属性值域的单映射。如果 A 由条件属性 C 与决策属性 D 构成,即: $C \cup D = A, C \cap D = \emptyset$, S 称作决策系统。

设 B 是 A 的子集,定义二元组

$R(B) = \{(x, y) \in U \times U \mid \forall a \in B, a(x) = a(y)\}$ 。如果 $(x, y) \in R(B)$,就称 x 与 y 是 B 不可分辨的, $R(B)$ 为 B 在 U 上的一个不可分辨关系。若 $R(B)$ 可以把论域 U 分为 k 个互不相交的等价类 $\{X_1, X_2, \dots, X_k\}$,则记 $\{X_1, X_2, \dots, X_k\} = U \mid_{R(B)}$, $\{X_i\}$ 为一个等价类,二元组 $K = (U, R)$ 构成一个近似空间。设 X 为 U 的一个子集, x 为 U 中的一个对象, $[x]_R$ 表示由 x 决定的等价类。如果集合 X 表示成等价类组成的并集,则称集合 X 是可以精确定义的;否则,集合 X 只能通过近似方式来表达。集合 X 关于 R 的下近似 $R_*(X)$ 与上近似 $R^*(X)$ 定义为

$$R_*(X) = \{x \in U \mid [x]_R \subseteq X\}$$

$$R^*(X) = \{x \in U \mid [x]_R \cap X \neq \emptyset\}$$

集合 X 的边界区定义为 X 的上近似与下近似之差,即

$$BN(X) = R^*(X) - R_*(X)$$

如果 $BN(X)$ 是空集,则称 X 关于 R 是清晰的,反之,则称集合 X 为关于 R 的粗糙集。

粗糙集理论中,定义对象 x 对集合 X 的粗糙隶属函数为

$$u_X(x, R) = \frac{\text{card}(X \cap [x]_R)}{\text{card}([x]_R)}$$

式中： $[x]_R$ 为所有与 x 不可分辨的对象组成的集合； $\text{card}[x]$ 为集合 X 的基数或势。若将 x 看成一条规则，则粗糙隶属度函数的取值可认为是该规则的置信度。

2.2 决策表与约简

设集合 U 为论域， R 为 U 上等价关系。若 $\text{ind}(R) = \text{ind}(R - \{r\}) \forall r \in R$ 成立，则称 r 是可省略的，否则，称为不可省略的。如果 R 中所有关系 r 都是不可省略的，则称 R 是独立的。如果等价关系族 P 是独立的，而且 $\text{ind}(P|P = R - r, P \subseteq R) = \text{ind}(R)$ 称 P 是 R 的一个约简，记作： $\text{red}(R)$ 。约简是保证正确划分的最小等价关系集合。

决策系统可用决策表来描述。决策表是行列表达的二维表格，每行表达为一个规则或一个对象，每列表达为规则或对象的一种属性。任何一个决策系统中的对象，根据条件属性和决策属性对应关系的不同，形成不同的决策规则或对象，决策表为若干这类决策规则的集合。

决策表中的关系，有些是相关冗余的，有些是相对独立的，为简化决策表，可应用等价关系的独立性约简冗余的决策表。通过约简并找出其核属性，即基于较少的条件获取相同的决策规则。一个决策表可能同时存在几个约简，这些约简的交集定义为决策表的核。

3. 区域交通状态模型

$G(N, L)$: 一个信号控制的交通区域路网，其中节点集合 N 、链路集合 L 分别是 $N = \{1, 2, \dots, m\}$, $L = \{l_{ij} | l_{ij} = \langle i, j \rangle, i, j \in N, i, j = 1, 2, \dots, m\}$; q_{ij}^x 为从节点 i 到节点 j 各方向的交通流量， $x=1, 2, 3$ 分别表示左转、直行和右转的流量； q_{ij} 为从节点 i 到节点 j 的交通流量， C_{ij} 为从节点 i 到节点 j 各方向的交通容量； t_{ij} 为从节点 i 到节点 j 单位长度的平均行程时间， d_{ij} 是链路 l_{ij} 的长度。 s_{ij} 为节点 j 的 i 方向入口道的排队时间长度；节点 j 的总流量为 $\sum_i q_{ij}$ ，节点 j 各方向平均的排队时间长度为 $s_j = \frac{\sum_i q_{ij} s_{ij}}{\sum_i q_{ij}}$ 。

4. 粗集模糊规则的建立方法

基于粗集理论从海量数据中提取模糊规则，主要步骤如下：

(1) 属性选择。根据交通区域划分的实际需求，确定模糊控制的输入变量和输出变量，即决策表的条件属性和决策属性。区域划分，多数为更好实施区域控制，让关联度高的路段和节点划为一区。但现实中关联度高低是相对概念，且与流量、密度、速度等参数密切相关。

本文研究中选取节点的各方向平均排队时长、平均流量、流量饱和度，路段的长度、流量、平均行程速度等作为输入变量。输出变量为子区划分的定量表达，考虑到交通控制子区一般情况以节点为划分单位，即连接子区内节点的路段，自动划归该子区，连接两个不同子区节点的路段，不属于任何子区。因此输出变量为节点之间的连接关系，连接关

系越强，则划为同一子区趋势越高。

(2) 变量的模糊化与属性编码。在不同的条件变量输入时，输出不同的决策变量，为便于在决策表中描述，采用模糊化表达方式。如节点某方向排队时长，可采用模糊语言{不排队、短排队、中排队、长排队、超长排队}表达，并选用属性编码{0, 1, 2, 3, 4}作为属性值。节点之间的关联度划分为{低、中低、中、中高、高}5个等级，分别用{1, 2, 3, 4, 5}编码。

(3) 获取数据，构造决策表。根据系统历史统计或实测数据，利用经验推理，计算各数据隶属度，对应到模糊值与属性编码，构造决策表。

(4) 约简决策表，建立模糊规则。对决策表中的规则，冗余的部分进行合并，不相容的规则根据置信度进行仲裁后保留合理项，约简决策表，建立模糊规则。

5 基于粗集理论的交通子区划分算法

交通路口节点虽然地域范围小，但却是交通瓶颈，直接影响道路状态，并且是控制策略实施的主要场所，因此根据节点状态进行交通控制小区划分。交通状态是时空连续变量，控制小区也是动态划分。在某考察时刻交通控制子区划分算法如下：

(1) 数据预处理

将实时采集的数据进行清洗，将连续变量离散化，剔除坏值，补全缺失值，根据需要进行归一化处理。

(2) 转为属性变量，构造决策表

根据上文4中模糊规则的建立方法，将预处理数据转换为条件变量和决策变量，构造决策表。

(3) 约简决策表，获得多层次模糊规则

基于粗糙集理论进行属性约简，获得精简后的决策表；计算每条规则的置信度，根据不同的置信度可获得不同精简决策表。

(4) 建立子区划分方案

根据决策表，推导各节点分类聚合方案，形成区域划分集合。基于不同参数的权重度，可能获得不同的划分方案。

6 实例仿真

为验证上述模型和算法，本文采用 Nguyen-Dupuis 网络作为研究路网，如图1所示。对其进行仿真运行，设置初始数据参数，利用 VISSIM 工具模拟仿真，运行结果的交通状态数据进行区域划分分析。

该网络拥有12个节点，19条路段。表1分别给出了该路网各路段的车道数、长度、自由流行程时间、路段容量。其中节点1、2、3、4、8、12、13为交通发生吸引点，共14个OD对，各OD对之间的交通需求，服从均值 $q_{rs,av}$ 、方差为 ε_{rs} 的对数正态分布，设定在平峰、高峰和过渡期三

个不同时段，交通需求如表 2 所示。在每个时段内，分为 30 个时刻，每时刻产生一次服从分布的随机变量。

节点和路段流量按饱和度计，离散化后属性编码为{1: 0-0.2, 2: 0.2-0.4, 3: 0.4-0.6, 4: 0.6-0.8, 5: 0.8-1}，节点排队时长属性编码{0: 0, 1: 1-30 秒, 2: 30-60 秒, 3: 60-120 秒, 4: 120 秒以上}，路段行程速度属性编码{1: 0-5km/h, 2: 5-15km/h, 3: 15-30km/h, 4: 30-50km/h, 5: 50km/h 以上}，节点距离属性编码{1: 20 米以内, 2: 20-100 米, 3: 100-300 米, 4: 300-600 米, 5: 600 米以上}。

路网实例中十字路口采用标准的四相位控制，丁字交叉口采用三相位控制，路段长度与图中等比放大。信号周期时长最大值 $\varepsilon_{\max}=150s$ ，最小值 $\varepsilon_{\min}=30s$ ，最小绿灯时间 $g_{\min}=10s$ 。BPR 函数中 $\alpha=0.15$ ， $\beta=4$ 。

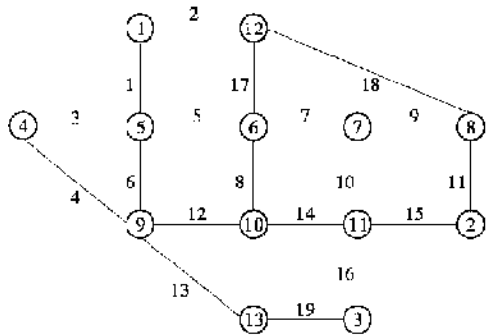


图1 Nguyen-Dupuis路网实例

根据上述设定，在 VISSIM 仿真运行至交通稳定状态。采集实时交通参数，根据前文模糊规则建立方法，生成决策表。决策表内容较多，表 3 列出了某平峰时刻与节点 6 相关的部分决策表内容。

应用文献 11 的方法对决策表进行约减，并计算置信度，表 4 是对表 3 进行约减的结果。根据约简后的决策表，对研究问题的所有交通节点进行子区划分。原则上按照关联度高低顺序，依次进行划分；当出现矛盾结果时，关联度低的关

表3 某平峰时刻与节点6相关的部分决策表

节点 i 饱和度	节点 i 排队时长	路段 ij 距离	路段 ij 饱和度	路段 ij 行程速度	节点 j 饱和度	节点 j 排队时长	节点 ij 关联度
5: 2	5: 1				6: 2	6: 1	5, 6: 2
		5, 6: 4	5, 6: 2	5, 6: 4			5, 6: 2
6: 2	6: 1				7: 2	7: 1	6, 7: 2
		6, 7: 4	6, 7: 2	6, 7: 4			6, 7: 2
6: 2	6: 1				10: 2	10: 1	6, 10: 2
		6, 10: 4	6, 10: 3	6, 10: 3			6, 10: 3
6: 2	6: 1				12: 2	12: 1	6, 12: 2
		6, 12: 4	6, 12: 2	6, 12: 4			6, 12: 2

表4 表3的约简

节点 i 饱和度	节点 i 排队时长	路段 ij 距离	路段 ij 饱和度	路段 ij 行程速度	节点 j 饱和度	节点 j 排队时长	节点 ij 关联度
		4	2	4			2
2	1				2	1	2
>=2			3		>=2		3
>=2				3	>=2		3

系，服从关联度低的划分方案。

表1 Nguyen-Dupuis网络中各路段的相关参数值

路段编号 a	长度 (米)	t_{a0} (分)	车道数	C_a (辆/h)
1	400	5	3	3200
2	500	6	2	2400
3	480	8	2	2500
4	600	12	2	2200
5	500	6	3	3100
6	360	7	3	3200
7	480	5	2	2400
8	360	7	3	3300
9	500	9	2	2500
10	360	7	2	2300
11	360	7	2	2200
12	500	6	2	2300
13	700	9	3	3200
14	480	5	3	3500
15	500	9	3	3300
16	500	5	2	2500
17	400	5	2	2300
18	1080	15	2	2400
19	480	5	3	3300

表2 各OD对交通需求的对数正态分布参数

OD 对	平峰期		高峰期	
	均值	方差	均值	方差
$q_{rs,av}$	ε_{rs}	$q_{rs,av}$	ε_{rs}	
1-2、2-1、4-3、3-4、4-8、8-4、12-13、2-12、3-1、12-3	4200	0.1	7500	0.12
1-3、3-1、4-2、2-4、13-12、12-2	4500	0.1	7200	0.12

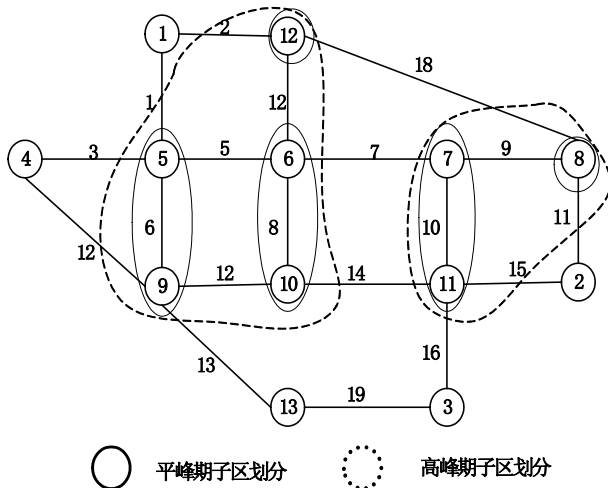


图3 在不同交通状态时的交通控制子区划分

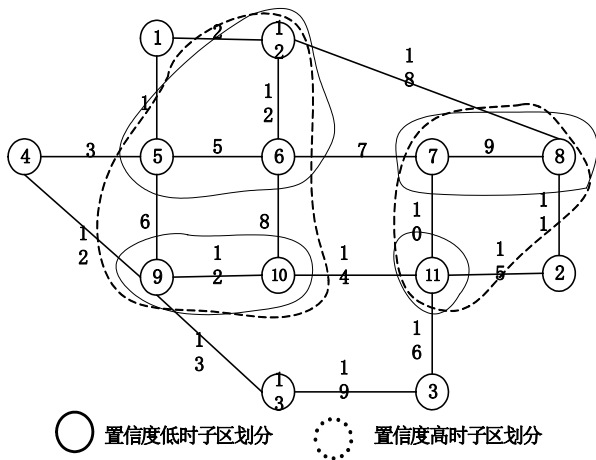


图4 不同置信度的子区划分

在不同的交通状态下,节点之间的关联度差异较大,因此子区划分的结果完全不同,如图3给出了在平峰期和高峰期 Nguyen-Dupuis 路网的不同划分结果。平峰时,交通流量小,饱和度低,节点间关联度小,子区划分多;高峰时,交通流量大,饱和度高,节点间关联度大,子区划分少。在选用不同置信度时,子区划分结果也不相同,图4给出了置信度分别为0.8和0.6时的子区划分情况。

7 结论和讨论

在进行交通区域控制的研究中,将整个城市交通网络划分为若个不同控制区域分别实施控制是简化复杂交通问题的常见思路,而进行区域分解,让关联程度更高的路口路段聚集在同一区域。但关联度高低本身没有统一有效的规则,本文基于粗糙集理论,提出了模糊规则的建立思路,将经验性路口-路段关联关系纳入规则体系,进而建立了交通子区划分方法。并通过典型路网实例分析了该方法的有效性。

本文的实例分析表明,不同置信度的选择对交通子区划分有较大的影响,作为积极的方面,通过设置不同的置信度可全面评价交通状态,分析不同区域划分对全局控制的效能影响。但也存在明显不足,置信度的选择不佳可能导致完全不同的区域划分和控制效果。

论文对提出的区域划分方法进行了初步的分析和评价,下一步将在更多不同的维度和更大的路网范围中对算法有效性和制约性进行深入的数据实验,另外如何选择合适的模糊规则约简方法,也是后一步开展的工作。

参考文献

- [1]美国运输部联邦公路局. 交通控制系统手册[M].北京:人民交通出版社,1987.
 - [2]杨迪,徐文瑜,王鹏.基于改进密度峰值聚类的路网划分方法[J].计算机应用研究,2023,12:3578-3583.
 - [3]侯军军,龙佰超,王洪钰,肖建力.基于交通指数聚类的路网区域动态划分[J].上海理工大学学报,2021,43(4):360-367.
 - [4]刘端阳,潘策.基于Neut的城市路网交通子区划分方法[J].浙江工业大学学报,2019,47(4):436-441.
 - [5]Lan Huihui, Wu Xianyu. Research on key technology of signal control subarea partition based on correlation degree analysis[J]. Mathematical Problems in Engineering, 2020: article ID 1879503
 - [6]郭红莉,刘晓雯,种潇敏,曲卫东.基于AP算法的路网同质区域划分方法研究[J],计算机技术与发展,2021,11:136-141.
 - [7]郑黎黎,杨帆,孙宝凤.基于GN算法的城市路网区域划分方法研究[J].重庆交通大学学报:自然科学版,2020,39(4):6-10.
 - [8]王春雷,钱勇生,基于双向并行灾变粒子群算法的区域交通控制[J],计算机工程与应用,2007,43(34):229-232.
 - [9]M. Abdulaal, L.J. LeBlanc: CONTINUOUS EQUILIBRIUM NETWORK DESIGN MODELS. Transportation Research Part B, 1979, 13: 19-32.
 - [10]张飞舟,曹学军,孙敏.基于多智能体的城市交通集成控制系统设计[J],北京大学学报(自然科学版),2008,44(2):289-292.
 - [11]刘清.Rough集及Rough推理[M].北京:科学出版社,2002.
 - [12]Niu X Z, Zhu J H, Wu C Q, et al. On a clustering-based mining approach for spatially and temporally integrated traffic sub-area division[J]. Engineering Applications of Artificial Intelligence, 2020, 96: 103932.
- 基金项目:山东省高等学校科技计划项目(No.J10LG52);交通运输部科技项目(No.2012-319-817-230)