

基于有权网络的区域交通子网划分方法研究

林 丹, 罗 杰

(南京邮电大学 自动化学院, 江苏 南京 210000)

摘 要:交通拥堵问题,已成为世界各国不容回避的棘手难题,引起了众多学者的关注。动态划分交通区域是提高区域交通系统整体效率的一个有效的解决方法,随着计算机技术的快速发展,复杂网络理论也有了突破性的进展。为此,在复杂网络社团划分的基础上,以交通路网的畅通特性为权重,提出了无权网络社团划分的改进算法。该算法采用路段间的车流量和路段距离作为权重的参考因素,同时结合网络中复杂度的大小,以模块度 Q 作为不同划分结果的评价标准,使得改进后的算法划分出来的社团可靠性更强。为验证提出算法的有效性和可行性,基于所编写的计算机程序,对该算法进行了仿真实验。基于仿真实验结果的改进前后的 Q 值分析对比,验证了该算法的有效性和可行性,且具有交通区域实时动态划分的潜力。

关键词:交通区域;社区结构;车流量;路段距离

中图分类号:TP301

文献标识码:A

文章编号:1673-629X(2017)09-0120-04

doi:10.3969/j.issn.1673-629X.2017.09.026

Research on Regional Transportation Sub-netting Method with Weighted Network

LIN Dan, LUO Jie

(College of Automation, Nanjing University of Posts and Telecommunications, Nanjing 210000, China)

Abstract: Traffic jam has become an inevitable problem in the world and caused a lot of attentions from scholars. Dynamic partition of traffic area is an effective solution to improve the efficiency for entire area traffic system. With the rapid development of the computer technology, the theory of complex networks has made a breakthrough. Therefore, on the basis of community structure partition on complex network, by taken the flow characteristics of traffic network as weight, an improved algorithm for no-weighted network community division is proposed. It refers to traffic flow and distance of the highway as reference factors for weight and combined with complexity of network, has enhanced the reliability of community structures by taken Modularity Q as evaluation standard for different partition result. To verify its effectiveness and feasibility, the simulation experiments are carried out based on the computer program self-compiled, which indicate that it is effective and feasible according to the contrast analysis on Q value before and after modification and has the potential of real-time dynamic division of traffic area.

Key words: traffic area; community structure; traffic flow; distance of highway

1 概 述

进入 21 世纪以来,随着全球城市化进程的高速推进,交通流的日益增大及复杂化,城市路网拥堵问题越来越严重,现有的智能交通控制策略很难提高整个区域交通系统的效率。对区域路网进行合理的划分能够改善资源的有效配置,对提高整个路网的通行效率具有重要意义。近年来,随着计算机的快速发展,以及众多学者对大规模网络数据的深入研究,推动了整个复杂网络领域研究方法的蓬勃发展。

网络在许多领域都扮演着至关重要的角色。在很多现实网络中,如物联网、社会关系网、交通路网等都具有一个相同的特性—社团结构^[1-2]。社团结构是内部连接紧密,而外部相对松散的集合,即对于大部分复杂网络都存在若干个社团,虽然每个社团内部节点之间的连接相对非常紧密,但对于各个社团之间的连接却比较稀疏^[1,3]。此现象揭示了网络的社团特性,对深入了解网络结构以及分析网络特性意义深远。

网络社团结构的划分与分级聚类以及图形分割都

收稿日期:2016-08-27

修回日期:2016-12-01

网络出版时间:2017-07-05

基金项目:江苏省自然科学基金项目(BK2011758)

作者简介:林 丹(1990-),女,硕士,研究方向为智能控制;罗 杰,博士,教授,研究方向为分布式智能控制、群体智能。

网络出版地址: <http://www.cnki.net/kcms/detail/61.1450.TP.20170705.1650.044.html>

有着密切关系。下面列举一些经典的社团划分算法:

(1) Kernighan-Lin 算法^[4]。一种试探优化算法,利用网络邻接矩阵的特征值和特征向量来进行社团划分。其本质是将网络划分为两个大小已知的社团,是一种基于贪婪算法的二分法,所以该算法具有一定的局限性。

(2) 基于 Laplace 图特征值的谱平分法^[5]。该算法依据 Laplace 矩阵不为零的特征值所对应的特征向量的各元素中,在社团内的节点对应的元素是近似相等的原理。但是该算法需要事先知道社团数目,对研究未知网络没有意义。

(3) GN 分裂算法^[6]。其算法原理是从网络中移除介数最大的边,把整个网络分解为社团。虽然 GN 算法对于社团划分具有较高的准确性,但是在不确定社团数目的前提下,该算法并不清楚分解到哪一步终止。针对该缺陷,一些学者提出了节点集的 GN 算法、自包含 GN 算法^[7]、极值优化算法^[8]等,但是这些算法共同的缺点就是复杂度较高。

(4) Newman 快速算法^[9-10]。该算法是通过不断合并节点来优化模块度的值,根据最优的模块度确定社团划分结果。与 GN 算法相比,该算法能保持较好的准确性,降低算法的复杂度。

上述算法均是针对无权网络提出的。由于城市交通网络是一个典型的复杂网络,交通路段之间并非完全是布尔关系,从而存在不同强度的耦合。为了衡量两点间连边的紧密程度和重要性,需要用权值来刻画不同连边的强度差异,以此构成加权网络。

有一些关于有权网络社团划分的研究,比如 COPRA^[11]和 Strength^[12]。COPRA 是基于 PAK^[13]提出的方法,在标签传播的大部分时间中,它不会收敛到恒定状态,而且划分的社团具有不确定性。Strength 利用点权和隶属度来检测重叠的社团划分,但是随着重叠的增加,Strength 的性能显著降低。

为此,将区域交通网络视为一个有权网络,提出了一种对 Newman 快速算法的改进方法,并将该算法应用于选定的交通路网,以实现交通网络的合理划分。

2 改进算法

2.1 参数定义

边权^[14]是网络用来衡量节点 i 与节点 j 共享的边的关联度大小的量,记作 w_{ij} 。 w_{ij} 的值越大,表明节点 i 与节点 j 之间的关联性越强,即两点之间的联系越紧密。反之,则表明节点 i 与节点 j 之间的关联性越弱,即两点之间的关联性越弱。

点权^[14-15]是用来描述节点在网络中的重要性,某一节点 i 的点权数据作 d_i ,而 d_i 为所有与节点 i 相连的

边权总和,即 $d_i = \sum_j w_{ij}$ 。 d_i 越大,表明该节点的权重越大,作用性越强;相反,在网络的作用就会较小。

度^[1]是拓扑图中最为基本的概念之一,某一节点 i 的度 k_i 是指与该节点相接的连边数目。节点的度是权衡一个节点重要性的指标之一。在分析网络拓扑结构时,可以基本认为某个节点的度越大,它在网络中的重要性就越高。

2.2 重新定义社团的模块度

模块度^[9]是常用的一种评估社团划分优劣的指标。它的提出是基于协调理论,主要思想是把划分社团后的网络与相应的零模型进行比较,以此判断划分的优劣。所谓一个网络对应的零模型,就是指与该网络具有某些相同的性质而在其他方面完全随机的随机图模型。在无权网络中,其表达式为^[16]:

$$Q = \sum_i (e_{ii} - a_i^2) = \text{Tre} - \|e^2\| \quad (1)$$

对于特定网络,不一样的社团划分结果对应的模块度值基本是不同的。一个特定网络划分社团,当社团的模块度最大时,表明此划分结果最优。此时的模块度为 Q_{\max} ,并且有 $0 \leq Q_{\max} \leq 1$ 。实际网络的模块度值一般都在 $0.3 \sim 0.7$ 之间。

为了使其适用于社团内部结构以及有权复杂网络的社团划分,重新定义式(1)中的模块度。假设网络中每个节点代表一个社团,此时有 k 个社团 (c_1, c_2, \dots, c_k) , $\mathbf{C} = (c_{ij})$ 是一个 k 维的对称矩阵,元素 c_{ij} 是社团 c_i 、 c_j 间的边权之和,代表社团内的边权占整个网络总边权的比例, $v_i = \sum_j c_{ij}$ 为矩阵各行元素总和,代表社团 i 中节点与外部节点间的边权占所有边权的比例, $\|\mathbf{C}^2\|$ 代表矩阵 \mathbf{C}^2 中的每个元素之和,因此有权网络模块度定义为:

$$Q^{\text{weight}} = \sum_i (c_{ii} - v_i^2) = \text{Tre} - \|\mathbf{C}^2\| \quad (2)$$

式(2)有一定的物理意义,即网络中社团内部的边权比例减去社团外部边权比例的期望值。 Q^{weight} 的最大值是 1,其值越大,表示社团划分结果越优。

2.3 基于 Newman 快速算法的改进

Newman 快速算法是一种自底向上进行的聚合算法,简称 FN 算法,主要步骤如下^[8]:

(1) 初始化原始网络,每个节点代表一个独立社团;最初的 e_{ij} 和 a_i 满足:

$$e_{ij} = \begin{cases} 0, & \text{其他} \\ \frac{1}{2m}, & \text{如果节点 } i \text{ 和 } j \text{ 之间有边相连} \end{cases} \quad (3)$$

$$a_i = \frac{k_i}{2m} \quad (4)$$

其中, k_i 为节点 i 的度。

(2) 对有边相连的社团依次合并,同时计算合并后模块度增量:

$$\Delta Q = e_{ij} + e_{ji} - 2a_i a_j = 2(e_{ij} - a_i a_j) \quad (5)$$

依照贪婪算法原理,每次社团合并尽量沿着使 Q 增大最多或者减少最小的方向进行。每次合并后,更新对应元素 e_{ij} ,把与 i 、 j 社团相关的行和列相加。

(3) 重复步骤(2),不停地将社团合并,直到整个网络都合并成一个社团。基于此方法,最多要合并 $n - 1$ 次。

最初的 Newman 快速算法是针对无权网络提出的,为了使其适用于加权网络,对原有算法进行改进,重新定义 e_{ij} 和 a_i 。边权代表两节点间的关联性,故

$$e_{ij} = \begin{cases} 0, & \text{其他} \\ \frac{r_{ij}}{\sum_{i,j} r_{ij}}, & \text{节点 } i \text{ 和 } j \text{ 有边相连} \end{cases} \quad (6)$$

其中, r_{ij} 为节点 i 和节点 j 间边的权值。

考虑到点权、节点的度都是衡量节点的重要指标,仅用一个作为节点重要性的评估过于片面,故由点权和度一起衡量:

$$a_i = 0.7 * \sum_j e_{ij} + 0.3 * \frac{k_i}{2m} \quad (7)$$

其中, k_i 为节点 i 的度; m 为整片区域网络的连边数; 0.7 和 0.3 为经验值。

在合并有边相连的社团的同时,需不停计算 $\Delta Q = e_{ij} + e_{ji} - 2a_i a_j = 2(e_{ij} - a_i a_j)$,每一次的合并都要沿着使 Q 增大最多或减少最小的方向。在合并后都需要更新对应元素 e_{ij} 、 a_i ,同时把与 i 、 j 社团有关的行和列相加。根据以上步骤不断合并社团,直到整个网络都成为一个社团。找到 Q 值最大时对应的社团划分结果,即为社团划分的最优解。

3 区域交通子网划分

区域交通的子区划分,对研究交通网络、提高路网的整体通行效率具有现实意义^[17-18]。为了考察改进算法在加权网络^[19-20]社团结构划分的有效性,选取南京仙林地区的一块交通区域,运用改进算法对其进行划分。为了表征相邻路口间的关联性,在路口 i 和路口 j 之间引入关联度 r_{ij} 。关联度借用牛顿万有引力定律的形式,定义把两个路口间的吸引力正比于两路口间的车流量,反比于两路口间距离的平方。表达如下:

$$CI = \frac{V}{D^2} \quad (8)$$

式(8)中, V 为两路口间的车流量,单位为 1 000 辆/时; D 为两路口间路段的距离,单位为 km。

该关联度表示:相邻路口间的距离越小,路口间的关联性越强,相邻路口的流量越大,路口间的联系越强。

3.1 路网描述

对图 1 中的主干线进行研究,即玄武大道、文澜路、文苑路、仙林大道、学海路、仙境路、九乡和西路等路段,该区域路网中共有 16 个信号控制路口和 23 条路段。为了便于分析和对比,对该区域路网中的各信号路口进行编号,见图 2。



图 1 南京仙林某块区域地图

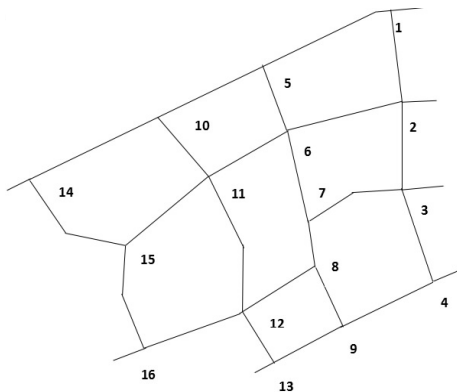


图 2 南京仙林某区域点线图

3.2 算法验证

在无权网络中,应用 Newman 快速算法对图 2 所示的区域路网进行社团划分,结果如图 3 所示。

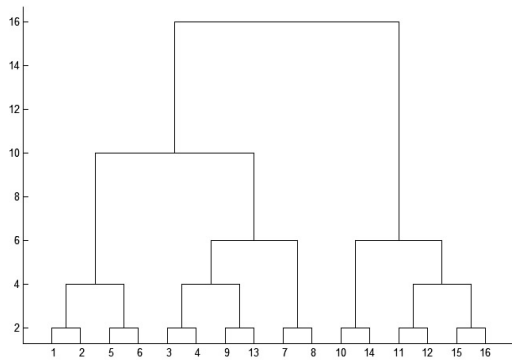


图 3 Newman 快速算法的划分结果

求得 Q 的最大值为 0.397 0。此时最优的划分结果为:1、2、5、6 为一个社团,3、4、9、13、7、8 为一个社团,10、14、11、12、15、16 为一个社团。通过这样的划分结果,发现无权网络不能很好地表现路段间的关联性。

通过百度地图得到每条路段的距离,在互联网上查询到某一天 16:00-17:00 时段的车流量,运用改进算法对该交通路网进行区域划分,结果如图 4 所示。

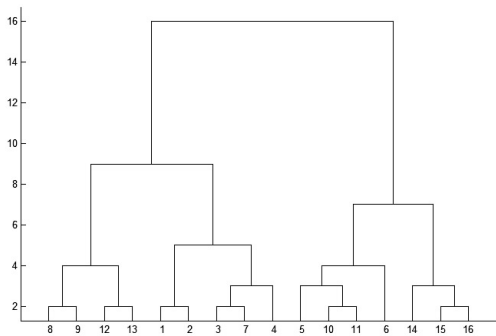


图 4 改进算法的划分结果

运算过程中, Q 的值依次为: $-0.072\ 6, 0.010\ 6, 0.075\ 6, 0.137\ 2, 0.184\ 6, 0.219\ 9, 0.262\ 8, 0.296\ 4, 0.326\ 4, 0.354\ 4, 0.380\ 2, 0.406\ 5, 0.414\ 9, 0.408\ 5, 0.353\ 8$ 。在 Q 值达到峰值 $0.414\ 9$ 后开始呈下降趋势。由此可见,当 8、9、12、13 为一个社团,1、2、3、7、4 为一个社团,5、10、11、6 为一个社团,14、15、16 为一个社团时,得到了最好的社团划分效果。

这样的社团划分结合了复杂网络度的特性以及路网特有的距离、流通量特性,通过构建出边权和点权物理模型,使得划分出来的结果可信度更高、更合理。由于车流量随着时间的变化而不断改变,可以获取不同时段的车流量对交通路网进行动态划分。根据新的划分结果重新分配路网资源,以此解决同一时段有的路口拥堵、有的路口空闲等资源分配不合理的问题。

4 结束语

对区域路网的合理划分可以提高交通资源的利用率。由于交通路网间点与点的连接并非单纯的布尔关系,基于现有的 Newman 快速算法对无权网络的划分方法,提出一种改进算法,并验证了该算法的有效性及其可行性。由于网络结构是由网络中的各种属性共同决定,故研究某个网络功能与某种网络结构之间的关系,只能得到一个更为合理的结论,而很难得出一个完全充分的结果。计算机对大量数据处理能力的提升,促进了复杂网络研究的蓬勃发展,而对社团结构的深入探究也推进了计算机图形学的发展。

参考文献:

[1] 汪小帆,李翔,陈关荣. 复杂网络理论及其应用[M]. 北京:清华大学出版社,2006.

[2] 王浩,李国欢,姚宏亮,等. 基于影响力计算模型的股票网络社团划分方法[J]. 计算机研究与发展,2014,51(10): 2137-2147.

[3] Karer B, Newman M E J. Stochastic block models and com-

munity structure in networks[J]. Physical Review E,2011,83 (1):016107.

[4] Kernighan B W, Lin S. An efficient heuristic procedure for partitioning graphs[J]. Bell System Technical Journal, 1970, 49(1):291-307.

[5] Pothen A, Simon H D, Liou K P. Partitioning sparse materices with eigenvectors of graphs[J]. SIAM Journal on Matrix Analysis Applications, 1990, 11(3):430-452.

[6] Givran M, Newman M E J. Community structure in social and biological networks[EB/OL]. [2008-12-07]. <http://www.biomedsearch.com/nih/Community-structure-in-social-biological/12060727.html>.

[7] Radicchi F, Castellano C, Cecconi F, et al. Defining and identifying communities in networks[J]. Eur. Phys. J. B, 2004, 101(9):2658-2663.

[8] Duch J, Arenas A. Community detection in complex networks using extremal optimization[J]. Physical Review E, 2005, 72 (2):027104.

[9] Newman M E J, Girvan M. Finding and evaluating community structure in networks[J]. Physical Review E, 2004, 69(2): 026113.

[10] Newman M E J. Fast algorithm for detecting community structure in networks[J]. Physical Review E, 2004, 69(6): 066133.

[11] Gregory S. Finding overlapping communities in networks by label propagation[J]. New Journal of Physics, 2010, 12(10): 103018.

[12] Chen D, Shang M, Lv Z, et al. Detecting overlapping communities of weighted networks via a local algorithm[J]. Physica A Statistical Mechanics & Its Applications, 2010, 389(19):4177-4187.

[13] Raghavan U N, Albert R, Kumara S. Near linear time algorithm to detect community structures in large-scale networks[J]. Physical Review E, 2007, 76(3):036106.

[14] Li M, Fan Y, Chen J, et al. Weighted networks of scientific communication: the measurement and topological role of weight[J]. Physica A Statistical Mechanics & Its Applications, 2005, 350(2-4):643-656.

[15] 王天成,刘真真,李天明,等. 复杂网络社团结构划分方法及其应用[J]. 信息通信, 2015(8):43-45.

[16] 陈宁宁. 信号控制子区动态划分及区域自适应协调控制研究[D]. 广州:中山大学,2010.

[17] 王秀凤,马英红. 基于加权网络模块强度的社团划分[J]. 计算机应用研究, 2013, 30(3):695-698.

[18] 刘传建. 复杂网络中的社团结构划分及分析应用[D]. 济南:山东大学,2014.

[19] 赫南,李德毅,淦文燕,等. 复杂网络中重要性节点发掘综述[J]. 计算机科学, 2007, 34(12):1-5.

[20] 安娜,谢福鼎,张永,等. 一种基于 GN 算法的文本概念聚类新方法[J]. 计算机工程与应用, 2008, 44(14):142-144.