

基于 POI 的出租车停靠位置选择算法研究^①

徐焕君^② 余 靖^③ 苑 帅 张玉停

(燕山大学信息科学与工程学院 秦皇岛 066004)

摘 要 基于兴趣点(POI),采用图论中判定节点重要度的方法,根据节点在路网中的基本信息,建立一种判定节点重要度的通用节点重要度(GNI)算法。针对不同场所出租车需求量不同的问题,建立了一种不同模式下的出租车按需分配呼叫模型。以 GNI 算法为基础,获得基于 POI 的专用节点重要度(DNI)算法。采用杠杆原理,同时类比多边形获取“重心”的方法,分析 DNI 算法下节点重要度排名结果,获得出租车最佳停靠位置,通过实验验证了该 DNI 算法的实用性及有效性。

关键词 兴趣点(POI);节点重要度;通用节点重要度(GNI)算法;呼叫模型;专用节点重要度(DNI)算法

0 引 言

随着互联网技术的快速发展,传统出租车行业面临着网约车的竞争,其中一个重要原因就是出租车等待位置以及资源分配不合理而造成的打车难问题。如何合理选择出租车停靠位置属于城市交通设施建设的一个主要问题,也是城市规划^[1]的重要依据。

城市兴趣点(point of interest, POI)数据涵盖了广泛的地理空间位置属性,相关学者使用 POI 为评价城市布局合理性提供了新的研究方法。文献[2]基于地理信息系统(geographic information system, GIS)和 POI 数据,研究了上海市的城市中心体系,与上海市现行城市总体规划进行对比,并提出发展建议。文献[3]基于 POI 数据对苏南 4 个城市的生活便利度进行评价,为公共服务设施空间分布及其便利度评价提供了新的方法。文献[4]基于住宅 POI 数据,考虑城市公园绿地对周边住宅价格的影响,构建了住宅市场特征价格模型并结合弹性与边

际价格分析,量化了公园绿地对住宅价格的影响。

基于 POI 的出租车停靠位置是根据节点重要度来选择的,用于判定网络节点重要度的方法本质来源于图论^[5]。目前的研究均是仅以一种类型的场所作为研究对象,如公交站点、地铁站点、铁路站点等,而没有综合考虑多类型场所。判定重要节点的方法除了度数法^[6]和介数法^[7]之外,考虑到整个网络的全局特性,又将贴中度中心性^[8]、介数中心性^[9]等也作为评价指标加入算法中。文献[10]提出一种基于熵的复杂网络节点重要性自适应判定方法,但采用的判定指标有一定的局限性。文献[11]提出了一种基于生成树数目的节点删除法,但节点删除可能会造成网络不连通,从而使判定结果不准确。文献[12]提出一种基于全网平均等效最短路径数的网络抗毁度评价模型,但没有完全体现出网络拓扑结构的差异性。网页排名(PageRank)算法^[13-14]不仅可以预测网络流量、未来链路等,还可以识别重要节点,且其计算复杂度较低。此外,文献[15,16]提出了融合节点自身重要度及连边重要度的改进加权节点收缩法,主要用于判定供应链网

① 国家自然科学基金(61872311,61973261)资助项目。

② 女,1995年生,硕士生;研究方向:复杂网络,空间数据查询;E-mail: xuhuanjun_ysu@163.com。

③ 通信作者,E-mail: xyj@ysu.edu.cn。

(收稿日期:2021-03-04)

络节点间的重要度。文献[17]利用A*算法,将复杂网络理论与动态需求特征相结合,对公交站点进行重要度判定,解决了定制公交站点及路径动态优化问题。文献[13-17]主要是针对单一类型节点判定重要度,具有一定的局限性。

本文在综合考虑节点在路网^[18]中基本信息的前提下,建立了一种在不同模式下出租车按需分配呼叫模型,给出适用于无权节点和加权节点的重要度判定方法,并以 POI 数据集为参考,对节点重要度进行排名,根据排名结果,依照杠杆原理和多边形求“重心”方法,找出最适合的出租车停靠位置,并实现可视化。

1 系统模型

本节首先从城市道路网络的结构特性出发,基于选取的 POI 处产生的呼叫出租车次数的不同提出 4 种模式。将城市中对于出租车需求量大的场所抽象为网络中的节点,将不同 POI 之间的运行线路抽象为节点间的连接边,基于此得到城市道路拓朴网络是一个复杂网络。该网络形成的拓朴图是一个无向图,数学表达式为 $G = (V, E)$ 是一个有序的二元组。其中 $V = (v_1, v_2, \dots, v_m)$ 称为点集,其元素称为顶点或节点; $E = (e_1, e_2, \dots, e_n)$ 称为边集,其元素称为无向边,简称边。

图 G 的邻接矩阵 $A = [a_{ij}]$, 其中,

$$a_{ij} = \begin{cases} 1 & \text{节点 } i \text{ 与节点 } j \text{ 有边相连} \\ 0 & \text{节点 } i \text{ 与节点 } j \text{ 无边相连} \end{cases}$$

1.1 呼叫模型

1.1.1 加权节点的权

(1) 节点类型共有 8 种,分别为车站 a、住宅区 b、学校(高校)c、医院 d、景区 e、酒店 f、餐馆 g 和购物广场 h。

(2) 时间参数共有 4 种,分别为早上 mor、中午 noon、下午 aft 和晚上 eve。

1.1.2 4 种出租车呼叫模式

模式 1 totalSame。不同节点处产生出租车呼叫次数相同,呼叫次数是一个有范围的随机数,随机数的产生受时间参数控制。

模式 2 typeSame。相同类型的节点产生出租车呼叫次数相同,不同类型的节点之间产生的呼叫次数不考虑是否会相同。呼叫次数是一个有范围的随机数,随机数的产生受时间参数控制。

模式 3 random。不同节点处产生出租车呼叫次数是随机的,随机数的产生受时间参数控制。

TotalSame、typeSame、random 模式随机数选择方案如表 1 所示。

表 1 totalSame、typeSame、random 模式随机数选择方案

时间参数	随机数范围
mor	[0, 20]
noon	[0, 10]
aft	[0, 18]
eve	[0, 15]

模式 4 human。人工干涉节点处出租车呼叫次数,指定被选择节点处的呼叫次数为一个较大的随机数,其他节点处的呼叫次数的产生方式与 totalSame 模式一致。

Human 模式随机数选择方案如表 2 所示。

表 2 human 模式随机数选择方案

时间参数	人工选择节点随机数	其余节点随机数
mor	[50, 60]	[0, 20]
noon	[35, 40]	[0, 10]
aft	[40, 50]	[0, 18]
eve	[50, 55]	[0, 15]

1.2 相关定义

定义 1 节点邻域 (neighborhood) 与闭邻域 (closed neighborhood)。设无向图 $G = \langle V, E \rangle$, $v_i, v_j \in V$, 若 $\exists e_t \in E$, 使得 $e_t = (v_i, v_j)$, 则称 v_i 与 v_j 是相邻的, 即 $a_{ij} = 1$ 。节点 v_i 邻域为 $\delta_i = \{v_j \mid v_j \in V, a_{ij} = 1, j = 1, 2, \dots, n\}$ 。节点 v_i 闭邻域为 $\overline{\delta_i} = \{v_i\} \cup \delta_i$ 。

1.2.1 节点加权系数

本文中节点的权值为节点处出租车的随机呼叫次数,该指标的提出是考虑到判定节点重要度时除了拓朴图本身结构属性之外,还受到 POI 即节点处

权值大小的影响。在其他条件相同的前提下,节点处的权值越大,节点越重要,说明该指标的贡献度越高。根据各 POI 处对出租车的随机呼叫次数,提出节点加权系数为不同节点处的出租车呼叫权重,也称为节点贡献度。

定义 2 局部加权系数(local weighting factor, LWF)。节点的局部加权系数为节点闭邻域中的权值和,节点 v_i 处的出租车呼叫次数为 c_{v_i} , 节点 v_i 邻域的出租车呼叫次数为 c_{δ_i} , 则节点 v_i 的局部加权系数 lwf_{v_i} 可以定义为

$$lwf_{v_i} = c_{v_i} + c_{\delta_i} \quad (1)$$

为了简化计算,缩小量值,将上述局部加权系数的值进行归一化,使结果范围在 $[0, 1]$ 之间。使用的归一化方法如式(2)所示。

$$x_{\text{norm}} = \frac{x - x_{\min}}{x_{\max} - x_{\min}} \quad (2)$$

定义 3 全局加权系数(global weighting factor, GWF)。节点的全局加权系数为节点 v_i 处的权值与整个网络图中所有节点的权值和之比,节点 v_i 的全局加权系数 gwf_{v_i} 为

$$gwf_{v_i} = \frac{c_{v_i}}{\sum_{v_i \in V} c_{v_i}} \quad (3)$$

1.2.2 局部重要度

定义 4 节点的度(degree)及其闭邻域的度(degree of closed neighborhood, CND)。节点 v_i 的度等于和 v_i 相关联的边的条数,又称关联度,记作 $d_G(v_i)$ 。节点 v_i 闭邻域的度 cnd_{v_i} 为

$$cnd_{v_i} = d_G(v_i) + d_G(\delta_i) \quad (4)$$

式中, $d_G(\delta_i)$ 为节点 v_i 的邻域集合中所有节点的度的累加和。

Sigmoid 函数也叫 Logistic 函数,在信息科学中,由于其单增以及反函数单增等性质, Sigmoid 函数常被用作神经网络的激活函数,将变量映射到 $0 \sim 1$ 之间。

$$S(x) = \frac{1}{1 + e^{-x}} \quad (5)$$

为了将节点闭邻域的度的值映射到 $0 \sim 1$ 之间,对原 Sigmoid 函数进行一个变形。

$$s(x) = \frac{2}{1 + e^{-x}} - 1 \quad (6)$$

其中 x 的取值范围为 $[0, +\infty)$, $s(x)$ 的取值范围为 $(0, 1)$ 。

定义 5 节点邻域度密度(neighborhood degree density, NDD)。将节点 v_i 邻域集合中节点的度数之和与节点 v_i 的度的比值定义为节点 v_i 的邻域度密度,代表着节点的邻域关系。节点 v_i 的邻域度密度 ndd_{v_i} 为

$$ndd_{v_i} = \frac{d_G(\delta_i)}{d_G(v_i)} \quad (7)$$

定义 6 基于度的点密度(dot density, DOTD)考虑了节点闭邻域的度与节点邻域集合的度密度之间的关系,其定义如式(8)所示。

$$dotd_{v_i} = \frac{cnd_{v_i}}{ndd_{v_i}} = \frac{d_G(v_i) + d_G(\delta_i)}{d_G(\delta_i)/d_G(v_i)} \quad (8)$$

定义 7 节点的局部集聚系数(agglomeration coefficient, AGGC)。节点 v_i 的邻居节点之间所存在

的边数 $\frac{\sum_{j, k \in \delta_i} a(j, k)}{2}$ 与可能存在的最大边数之比。节点 v_i 的局部集聚系数 $aggc_{v_i}$ 为

$$aggc_{v_i} = \frac{\sum_{j, k \in \delta_i} a(j, k)/2}{d_G(\delta_i) \cdot d_G(\delta_i - 1)/2} = \frac{\sum_{j, k \in \delta_i} a(j, k)}{d_G(\delta_i) \cdot d_G(\delta_i - 1)} \quad (9)$$

该指标考虑了节点的邻居节点之间的紧密程度。

1.2.3 全局重要度

定义 8 接近中心性(proximity centrality)。假设图中共有 m 个节点, d_{ij} 是节点 v_i 到节点 v_j 的距离(按经纬度计算), d_i 是节点 v_i 到网络中所有节点的距离的平均值。节点 v_i 的经纬度为 (lon_{v_i}, lat_{v_i}) , 节点 v_j 的经纬度为 (lon_{v_j}, lat_{v_j}) , 则两点经纬度间的距离公式为

$$a = \sin^2\left(\frac{lat_{v_j} - lat_{v_i}}{2}\right) + \cos(lat_{v_i}) \cdot \cos(lat_{v_j}) \cdot \sin^2\left(\frac{lon_{v_j} - lon_{v_i}}{2}\right) \quad (10)$$

$$c = 2 \cdot \arcsin(\sqrt{a}) \quad (11)$$

$$d_{ij} = c \cdot r \quad (12)$$

式中 r 为地球平均半径,通常选择常数值为 6371 km。

节点 v_i 的全局重要度(global importance) gli_{v_i} 为

$$gli_{v_i} = \frac{1}{d_i} = \frac{m}{\sum_{j=1}^m d_{ij}} \quad (13)$$

其中, d_i 值的相对大小在某种程度上反映了节点 v_i 在网络中的相对重要性, d_i 的值越小意味着节点 v_i 更接近其他节点, 越居于网络中心, 节点 v_i 在全局网络中越重要。

1.2.4 节点重要度

复杂网络本质上的非同质拓扑结构, 决定了网络中每个节点的重要程度是不同的。节点在复杂网络中的重要度首先取决于节点在网络中的位置, 如网络中的“末梢节点”和“非末梢节点”的重要程度显然是不一样的; 其次, 节点在网络中的重要度还取决于节点的连通能力, 该节点连接的边数越多, 节点就越重要; 除此之外, 本文中提出的节点处的权值大小也影响着节点的重要程度。

定义 9 通用节点重要度 (general node importance degree, GNID), 称

$$GNID_{v_i} = \frac{\frac{2}{1 + e^{-cnd_{v_i}}} - 1 + dotd_{v_i} + aggc_{v_i}}{3} \cdot gli_{v_i} \quad (14)$$

为节点 v_i 是通用节点时的重要度。通用节点在本文里特指无权节点, 即节点处无权值。

定义 10 专用节点重要度 (dedicated node importance degree, DNID), 称

$$DNID_{v_i} = \frac{X_{\text{norm}}(lwf_{v_i}) + gw_{v_i}}{2} \cdot \frac{\frac{2}{1 + e^{-cnd_{v_i}}} - 1 + dotd_{v_i} + aggc_{v_i}}{3} \cdot gli_{v_i} \quad (15)$$

为节点 v_i 是专用节点时的重要度。专用节点在本文里特指加权节点, 即节点处有权值。

$GNID_{v_i}$ 与 $DNID_{v_i}$ 的值越大, 节点在网络中越重要。

2 算法描述

复杂网络中节点的重要度由节点在复杂网络中的位置和在其邻域中的度以及节点处的权值大小共

同决定。

2.1 GNI 算法

本文通用节点重要度 (general node importance, GNI) 计算方法可以在常见的网络图中使用, 节点处无权值, 它考虑的是节点在网络中的基本结构, 综合节点在局部区域和全局中的信息, 设计算法过程如算法 1 所示。

算法 1 计算通用节点重要度

Input

城市兴趣点 POI 数据集拓扑图 $G = (V, E)$

Output

通用节点重要度字典 $Dict_GNI$

BEGIN

$Dict_GNI \leftarrow \text{NULL}$ // 构建通用节点重要度字典

for each node v_i in V do // 遍历图中节点

$aggc_{v_i} \leftarrow$ 使用式 (9) 计算节点 v_i 的局部聚集系数

$Sum_distance, Sum_edge \leftarrow 0, 0$

for each node v_i in V do

if $\{(v_i, v_j)\} \text{ in } E$ then // 节点 v_i 和 v_j 之间有边相连
计算 v_i 和 v_j 之间的距离, 并累加至 $Sum_distance$

$Sum_edge \leftarrow Sum_edge + 1$

end if

$d_i \leftarrow Sum_distance / Sum_edge$

end for

$GNID_{v_i} \leftarrow$ 使用式 (14) 计算节点 v_i 的通用节点重要度

$Dict_GNI \leftarrow Dict_GNI \cup \{v_i: GNID_{v_i}\}$

end for

return $Dict_GNI$

END

2.2 DNI 算法

本文专用节点重要度 (dedicated node importance, DNI) 计算方法是在 DNI 算法的基础上设计的, 它除了综合考虑节点在网络中的基本结构信息外, 还增加了节点加权系数 (贡献度) 这一指标, 更适用于本文的研究背景, 算法过程如算法 2 所示。

算法 2 计算专用节点重要度

Input

城市兴趣点 POI 数据集拓扑图 $G = (V, E)$

呼叫类型 CM

呼叫产生时间 CGT

Output

专用节点重要度字典 $Dict_DNI$

BEGIN

$Dict_DNI \leftarrow \text{NULL}$ // 构建专用节点重要度字典

$Dict_Calls \leftarrow \text{NULL}$ // 构建节点呼叫次数字典

for each node v_i in V do // 获得节点呼叫次数

call $\leftarrow$ 获得节点 v_i 在呼叫模式中对时间下的呼叫次数

$Dict_Calls \leftarrow Dict_Calls \cup (v_i:call)$ // 更新专用节点重要度字典

end for

for each node v_i in V do

$aggc_{v_i} \leftarrow$ 使用式(9)计算节点 v_i 的局部聚集系数

$Sum_distance, Sum_edge \leftarrow 0, 0$

for each node v_j in V do

if $\{(v_i, v_j)\} \in E$ then // 节点 v_i 和 v_j 之间有边相连

计算 v_i 和 v_j 之间的距离, 并累加至 $Sum_distance$

$Sum_edge \leftarrow Sum_edge + 1$

end if

$d_i \leftarrow Sum_distance / Sum_edge$

end for

$DNID_{v_i} \leftarrow$ 使用式(15)计算节点 v_i 的专用节点重要度

$Dict_DNI \leftarrow Dict_DNI \cup \{v_i:DNID_{v_i}\}$

end for

return $Dict_DNI$

END

GNI 算法适用于节点处无权值的情况下求节点重要度。DNI 算法适用于节点处有权值的情况下求节点重要度。DNI 算法更适用于本文的研究背景, 在 POI 处对出租车有需求, 即在节点处有权值的情况下判定节点重要度。

2.3 出租车停靠位置选择算法

出租车停靠位置的选择依据周围节点的重要程度, 在多个重要程度不同的节点之间的某个位置选择出租车停靠点时, 需要考虑到停靠位置到节点的距离及节点的重要程度两个因素。本文类比多边形获取“重心”的方法, 求得最佳的出租车停靠位置。

已知节点 v_i , 其在平面坐标系中的坐标为 (x_{v_i}, y_{v_i}) , 米勒坐标系中的坐标为 (lon_{v_i}, lat_{v_i}) 。米勒坐标系下坐标转化至平面坐标系下坐标公式为

$$L = 6381372 \cdot \pi \cdot 2$$

$$H = \frac{L}{2}$$

$$x_{v_i} = \frac{lon_{v_i} \cdot \pi}{180}$$

$$y_{v_i} = \frac{lat_{v_i} \cdot \pi}{180}$$

$$y_{v_i} = 1.25 \cdot \log \cdot (\tan(0.25 \cdot \pi + 0.4 \cdot y_{v_i}))$$

$$x_{v_i} = \frac{L}{2} + \frac{L}{2\pi} \cdot x_{v_i}$$

$$y_{v_i} = \frac{H}{2} - \frac{H}{mill} \cdot y_{v_i} \quad (16)$$

节点 v_j 为节点 v_i 的邻域节点, 且其在 v_i 的邻域中重要度值最大。 o' 为节点 v_i 和节点 v_j 之间选择的出租车停靠位置, 设节点 v_i, v_j 的重要度值分别为 $DNID_{v_i}, DNID_{v_j}$, o' 点的重要度值为 $DNID_{v_i} + DNID_{v_j}$ 。节点 $v_i(x_{v_i}, y_{v_i})$ 到 $o'(x_{o'}, y_{o'})$ 的距离为 L_i , 节点 $v_j(x_{v_j}, y_{v_j})$ 到 $o'(x_{o'}, y_{o'})$ 的距离为 L_j , 节点 $v_i(x_{v_i}, y_{v_i})$ 到 $v_j(x_{v_j}, y_{v_j})$ 的距离为 L 。利用杠杆原理

$$\begin{cases} DNID_{v_i} \cdot L_i = DNID_{v_j} \cdot L_j \\ L_i + L_j = L \end{cases}$$

得:

$$\begin{cases} L_i = \frac{DNID_{v_j} \cdot L}{DNID_{v_i} + DNID_{v_j}} \\ L_j = \frac{DNID_{v_i} \cdot L}{DNID_{v_i} + DNID_{v_j}} \end{cases}$$

在平面直角坐标系下:

$$\begin{cases} (x_{o'} - x_{v_i})^2 + (y_{o'} - y_{v_i})^2 = L_i^2 \\ (x_{o'} - x_{v_j})^2 + (y_{o'} - y_{v_j})^2 = L_j^2 \end{cases} \quad (17)$$

求得 $o'(x_{o'}, y_{o'})$ 。

已知点 $o(x_o, y_o)$, 其在平面坐标系中的坐标为 (x_o, y_o) , 在米勒坐标系中的坐标为 (lon_{v_o}, lat_{v_o}) 。平面坐标系下坐标转化至米勒坐标系下坐标公式为

$$L = 6381372 \cdot \pi \cdot 2$$

$$H = \frac{L}{2}$$

$$lat_{v_o} = \frac{(\frac{H}{2} - y_o) \cdot 2 \cdot mill}{1.25H}$$

$$\begin{aligned} lat_{v_o} &= \frac{\arctan(e^{lat_{v_o}} - 0.25\pi) \cdot 180}{0.4\pi} \\ lon_{v_o} &= \frac{(x_o - L) \cdot 360}{L} \end{aligned} \quad (18)$$

式中 $mill = 2.3$, 最终结果留到小数点后 7 位。出租车停靠位置选择算法过程如算法 3 所示。

算法 3 计算出租车停靠位置

Input

专用节点重要度字典 $Dict_DNI$

城市兴趣点 POI 数据集拓扑图 $G = (V, E)$

Output

出租车停靠点位置经纬度集合 $List_TAXI$

BEGIN

$List_TAXI \leftarrow \text{NULL}$ //构建出租车停靠点位置经纬度集合

$TopNodeList \leftarrow$ 筛选出专用节点重要度排名前 5% 的节点

for each node v_i in $TopNodeList$ do

$\overline{\delta_{v_i}} \leftarrow$ 从图 G 中获得节点 v_i 的闭邻域集合

If $|\overline{\delta_{v_i}}| \geq 3$ then

v_j, v_k (从领域集合 $\overline{\delta_{v_i}}$ 中选择排名靠前的两个节点)

使用式(17)获得节点 v_i, v_j 的平衡点 o' , 并更新 o' 点的节点重要度;

使用式(17)获得节点 o', v_k 的平衡点 o , 并更新 o 点的节点重要度;

$(lon_{v_o}, lat_{v_o}) \leftarrow$ 使用式(18)获得 o 点的地理位置坐标经纬度

$List_TAXI \leftarrow List_TAXI \cup \{(lon_{v_o}, lat_{v_o})\}$

else

$v_j \leftarrow$ 从领域集合 $\overline{\delta_{v_i}}$ 中选择排名第一的节点

使用式(17)获得节点 v_i, v_j 的平衡点 o , 并更新 o 点的节点重要度

$(lon_{v_o}, lat_{v_o}) \leftarrow$ 使用式(18)获得 o 点的地理位置坐标经纬度

$List_TAXI \leftarrow List_TAXI \cup \{(lon_{v_o}, lat_{v_o})\}$

end if

end for

return $List_TAXI$

END

113 家对出租车有较大需求的 8 类场所, 分别为车站、住宅区、学校(高校)、医院、景区、酒店、餐馆和购物广场, 如图 1 和图 2 所示。这 113 家场所的选取来自当前热门 APP 推荐并且根据真实的经纬度数据(北京大学开放数据平台和高德地理信息集)与高德地图实际行驶路线构成城市道路网。城市道路网是在城镇管辖范围内由节点和各种不同功能的干道和区域性道路组成的, 是城市总体规划布局的骨架。

选取北戴河区的 POI 作为研究数据的原因:

(1) 北戴河区交通基础设施建设还在发展中, 没有市中心发达, 公交站点密度也比较小; (2) 因为旅游等需求, 人口流动量大; (3) 堵车情况不严重, 出租车市场发展潜力大。

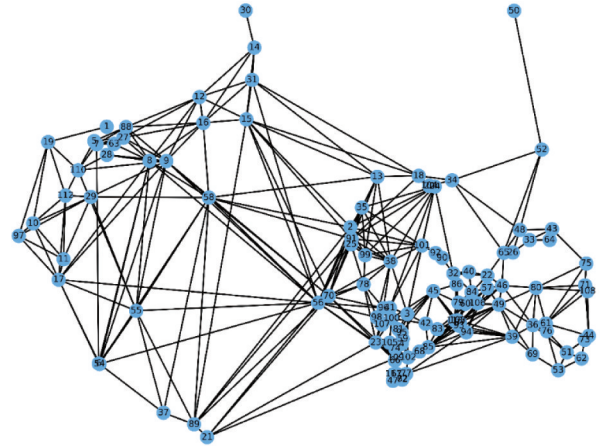


图 1 POI 及出租车可行驶路线网络图

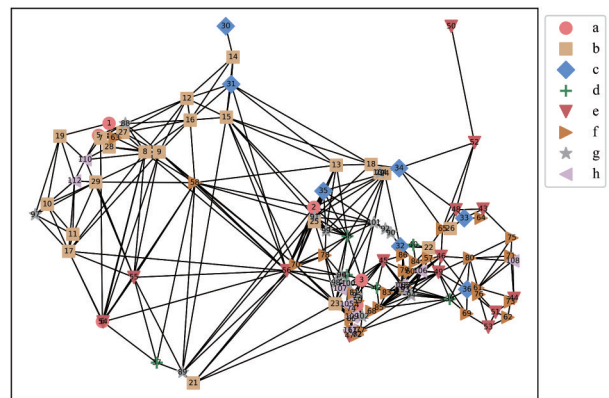


图 2 POI 按类型显示及出租车可行驶路线网络图

3 实验结果分析

3.1 数据集

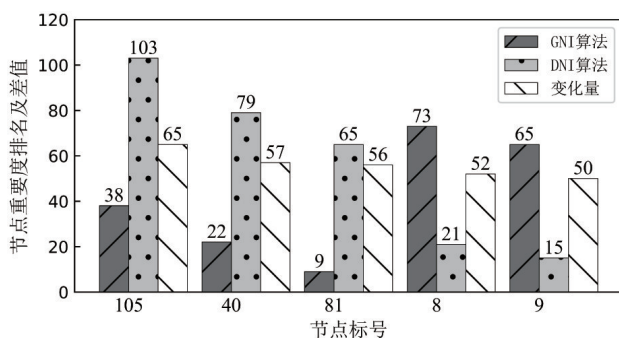
本文选取的 POI 为河北省秦皇岛市北戴河区

3.2 实验

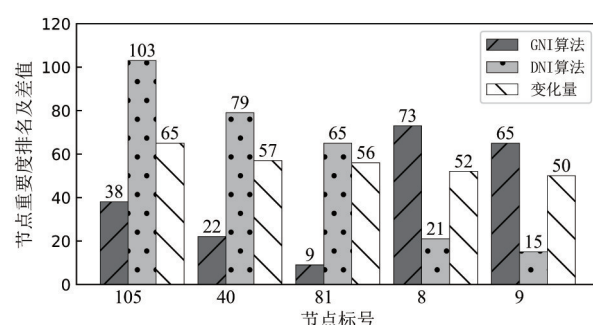
为求得北戴河区最佳的出租车停靠位置, 共进行 2 组实验。第 1 组实验根据定义 8 和定义 9 分别

在 4 种模式下以 GNI 算法和 DNI 算法求出图中 113 个节点的重要度排名,通过比较两种算法下的排名可以看到,在节点处有权值时会影响节点的重要度。第 2 组实验根据 DNI 算法下的节点重要度排名,选出排名前 5% 的节点,依据出租车停靠位置选择算法,求得最佳的出租车停靠位置。

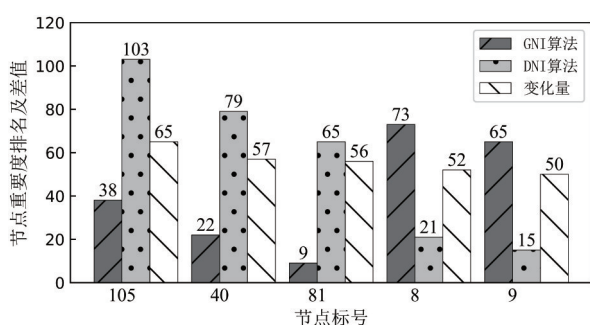
下面的各个实验是对于不同模式下 DNI 算法与 GNI 算法节点重要度变化量排名前 5 结果对比。



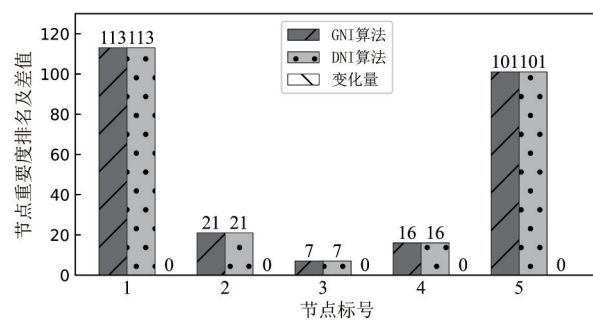
(a) mor; callcount = 12



(b) noon; callcount = 7



(c) aft; callcount = 1

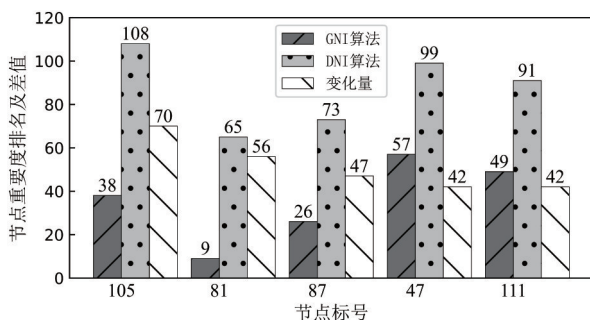


(d) eve; callcount = 0

图 3 totalSame 模式

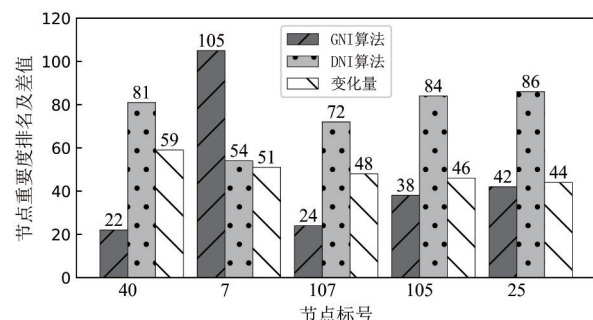
(2) typeSame 模式

如图 4 所示,该模式下,相同类型的节点处产生的呼叫次数相同。时间参数为 mor 时,8 种类型节



(a) mor

点的呼叫次数分别为 5、3、7、7、0、4、1、0;时间参数为 noon 时,8 种类型节点的呼叫次数分别为 9、0、5、0、2、3、1、6;时间参数为 aft 时,8 种类型节点的呼叫



(b) noon

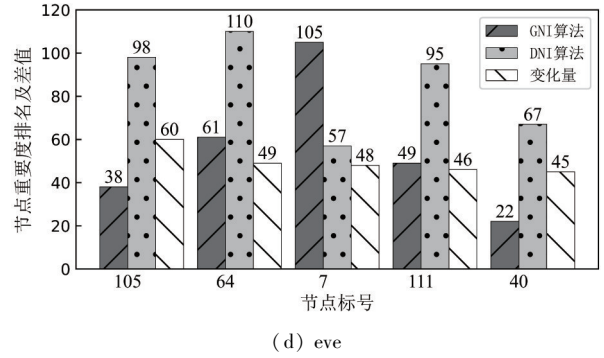
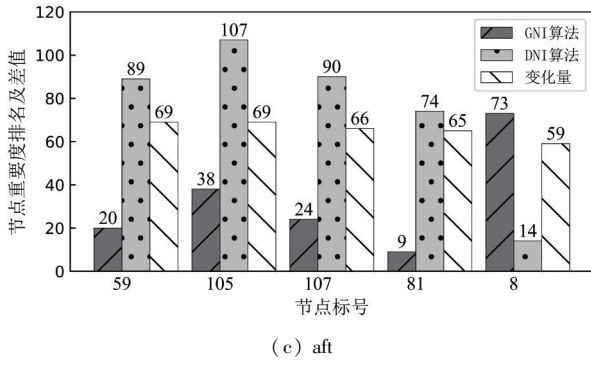


图4 typeSame 模式

次数分别为 6、9、16、7、15、9、4、5;时间参数为 eve 时,8 种类型节点的呼叫次数分别为 12、3、2、6、1、4、5、1。由实验结果可知,typeSame 模式下,节点重要度排名在 DNI 算法下与 GNI 算法下有明显变化。

(3) random 模式

如图5所示,该模式下,不同节点处产生的呼叫次数是在时间参数的控制下随机生成的。由实验结果可知,random 模式下,节点重要度排名结果在 DNI 算法下与 GNI 算法下有明显变化。

(4) human 模式

如图6所示,该模式下,人工选择标号为 50 的

节点,此节点处产生一个较大的随机数,而其余节点产生的随机数方式与 totalSame 模式一致。时间参数为 mor 时,{50 号节点:58,其他节点:14};时间参数为 noon 时,{50 号节点:37,其他节点:0};时间参数为 aft 时,{50 号节点:45,其他节点:2};时间参数为 eve 时,{50 号节点:50;其他节点:5}。

由实验结果可知,human 模式下,节点重要度排名在 DNI 算法下与 GNI 算法下有明显变化。

3.2.2 出租车停靠位置选择

本文节点重要度排名方法选取 DNI 算法,以 human 模式下,设定时间参数等于 mor 时为例,生成

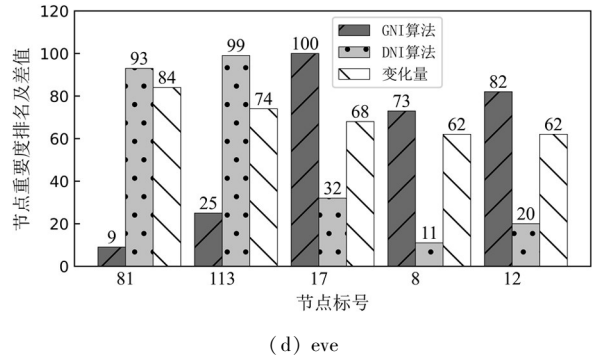
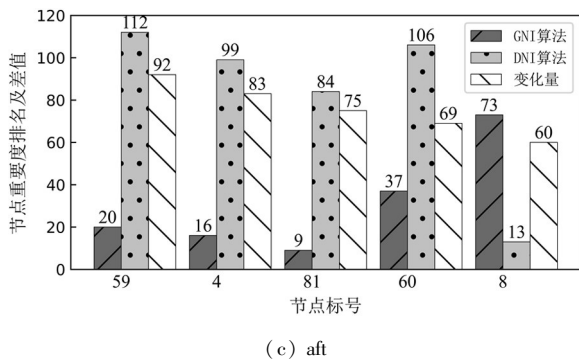
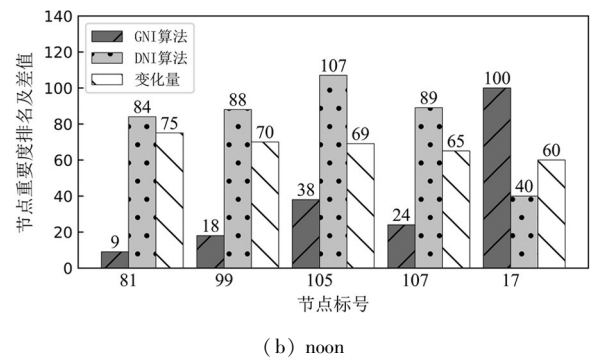
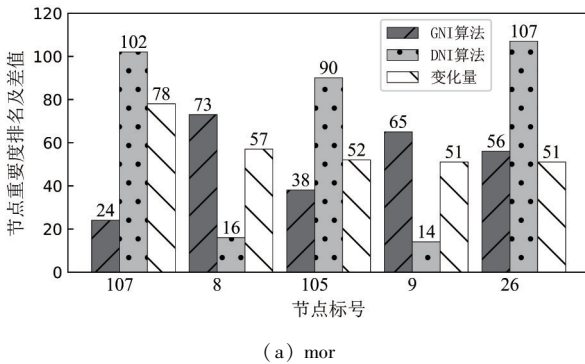


图5 random 模式

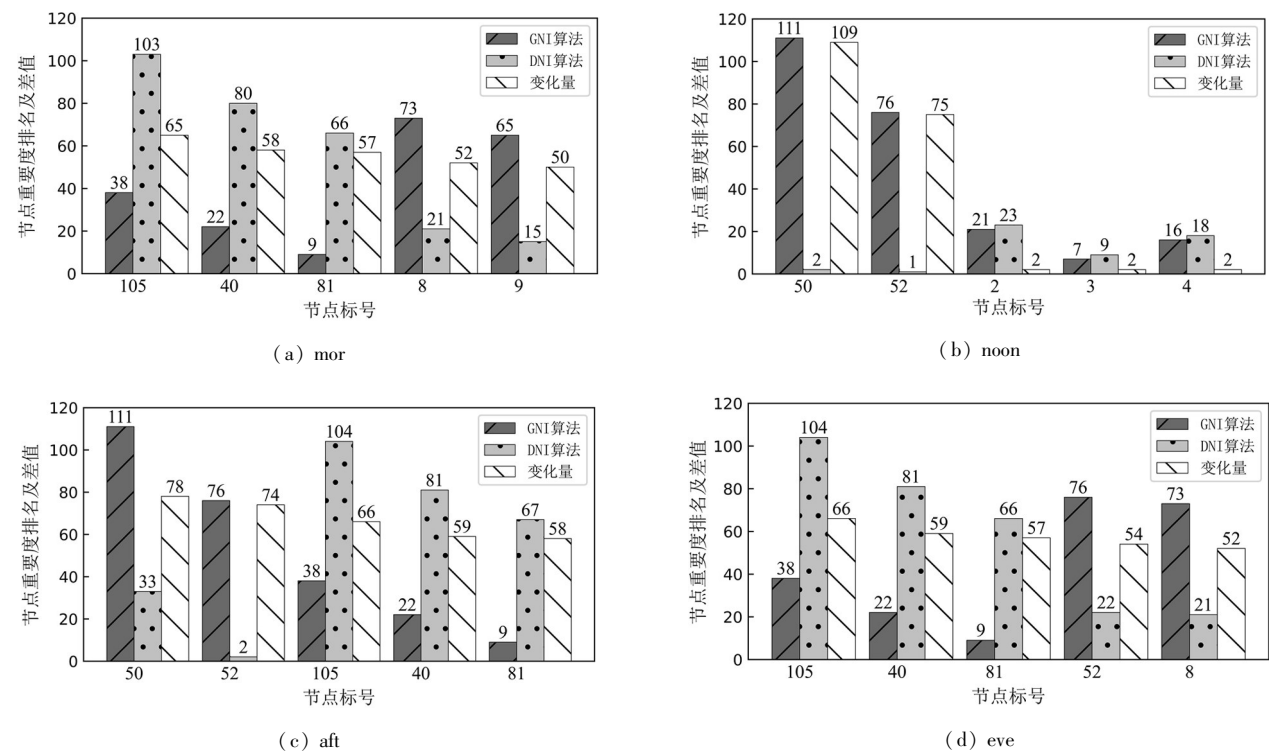


图 6 human 模式

出租车停靠位置并将其结果可视化。

根据出租车停靠位置选择算法,排名前 5% 的节点的出租车停靠位置如图 7 所示(黑色粗加号)。从表 3 及图 7 中可知,排名第 2 的节点与排名第 4 的节点分别在其闭邻域中找最合适的出租车停靠位置时,得到的结果一致。因为标号为 56 的节点与标号为 70 的节点互为闭邻域的关系,在通过出租车停靠位置选择算法计算时,是会出现结果一致的可能性的。而在该模式下,标号为 50 的节点处随机数为 56,此时节点加权系数对 50 号节点的重要度排名及最终的出租车停靠位置都有影响。

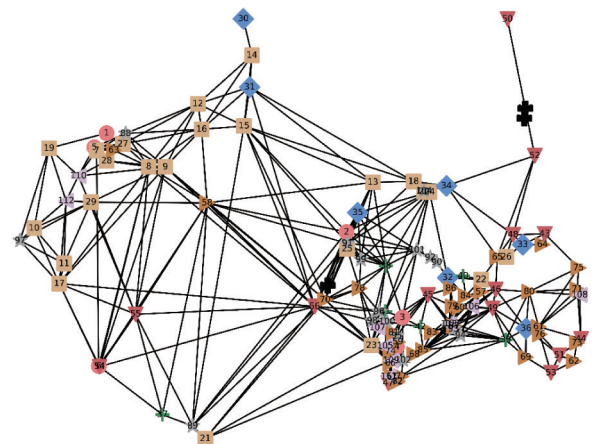


图 7 出租车停靠位置在网络图中的可视化

表 3 human 模式下(时间参数为 mor)出租车停靠位置选择方案

排名	节点标号	节点处随机数	节点重要度	生成出租车停靠位置经纬度
1	52	1	0.148877	(119.5217146, 39.853491)
2	56	1	0.117975	(119.4737954, 39.8284906)
3	50	56	0.113028	(119.5219501, 39.8549661)
4	70	1	0.107622	(119.4737954, 39.8284906)
5	91	1	0.100263	(119.5219501, 39.8549661)

本文以北戴河区 113 个 POI 进行实例分析,得出以下结论。

(1) POI 处的呼叫次数大小,即节点处随机数的权值大小,会影响节点重要度的排名结果。

(2) 节点的权值与节点在图中的基本属性之间并无相关性,进一步说明单个指标具有不同侧重点,而综合指标更能从物理拓扑结构及呼叫次数属性方面体现节点的重要性。

(3) 出租车停靠位置的选取与 POI 处的重要性相关,停靠位置会更靠近较为重要的 POI。

4 结 论

本文选取北戴河区 113 个 POI,考虑了其实际地理位置信息与出租车行驶路线,提出了 POI 处出租车呼叫的次数作为新指标,建立了一种具有不同模式的出租车按需分配呼叫模型。该模型更适合于判定加权节点的重要度,同时基于节点的重要度排名,给出了较好的出租车停靠位置。下一步的研究方向会将出租车的实际行驶距离以及行驶道路上需要经过的红绿灯数等指标考虑进去,以得到更好的出租车停靠位置。

参考文献

- [1] 骆莉,陈素平. 分析城市道路网规划布局及优化[J]. 美与时代(城市版), 2018(1): 38-39
- [2] 施歌,江南,姚恋秋. 基于 GIS 和兴趣点(POI)数据的城市中心体系识别方法研究——以上海市为例[J]. 现代测绘, 2017, 40(6): 27-30
- [3] 卢中辉,陈其龙,任启龙,等. 基于 POI 数据的苏南城市生活便利度评价[J]. 淮阴师范学院学报(自然科学版), 2020, 19(3): 233-238
- [4] 刘雅轩,陈彤. 基于 POI 数据的乌鲁木齐市城市公园绿地对周边住宅价格的影响研究[J]. 干旱区资源与环境, 2020, 34(11): 36-43
- [5] Kathryn L F. Introduction to graph theory[J]. *Networks*, 1997, 30(1): 73-73
- [6] Callaway D S, Newman M E, Strogatz S H, et al. Network robustness and fragility: percolation on random graphs[J]. *Physical Review Letters*, 2000, 85(25): 5468-5471
- [7] Freeman L C. A set of measures of centrality based on betweenness[J]. *Sociometry*, 1977, 40(1): 35-41
- [8] Li G S, Li M, Wang J X, et al. United neighborhood closeness centrality and orthology for predicting essential proteins[J]. *IEEE/ACM Transactions on Computational Biology and Bioinformatics*, 2020, 17(4): 1451-1458
- [9] Riondato M, Upfal E. ABRA: approximating betweenness centrality in static and dynamic graphs with rademacher averages[J]. *ACM Transactions on Knowledge Discovery from Data*, 2018, 12(5): 1-38
- [10] Zhou A, Sun Q B, Yang G Y, et al. An Entropy-based self-adaptive node importance evaluation method for complex networks[J]. *Complexity*, 2020(10): 1-13
- [11] 陈勇,胡爱群,胡骏,等. 通信网中最重要节点的确定方法[J]. 高技术通讯, 2004, 14(1): 21-24
- [12] 饶育萍,林竞羽,周东方. 网络抗毁度和节点重要性评价方法[J]. 计算机工程, 2009, 35(6): 14-16
- [13] 任晓龙,吕琳媛. 网络重要节点排序方法综述[J]. 科学通报, 2014, 59(13): 1175-1197
- [14] 库珊,刘钊. 基于 PageRank 与 HITS 的改进算法的网页排名优化[J]. 武汉科技大学学报, 2019, 42(2): 155-160
- [15] 张旭,袁旭梅,袁继革. 基于加权改进节点收缩法的供应链网络节点重要度评估[J]. 计算机应用研究, 2017, 34(12): 3801-3805
- [16] 刘娜,沈江,于鲲鹏,等. 基于改进节点收缩法的加权供应链网络节点重要度评估[J]. 天津大学学报(自然科学与工程技术版), 2018, 51(10): 1056-1064
- [17] 魏长钦,王伟智. 随机需求下定制公交站点及路径动态优化模型[J]. 福州大学学报(自然科学版), 2020, 48(1): 98-104
- [18] 谈进辉. 东莞市路网复杂网络特性及鲁棒性研究[J]. 现代交通技术, 2019, 16(4): 59-62

Research on the algorithm of choosing taxi stops based on POI

Xu Huanjun, Yu Jing, Yuan Shuai, Zhang Yuting

(School of Information Science and Engineering, Yanshan University, Qinhuangdao 066004)

Abstract

Based on point of interest (POI) and adopting the method of determining the importance of nodes in graph theory, a general node importance (GNI) algorithm for determining the importance of nodes is established according to the basic information of the nodes in the road network. Aiming at the problem of the different demand for taxis in different places, an on-demand taxi call model with different modes is established. Based on the GNI algorithm, a dedicated node importance (DNI) algorithm based on POI is obtained. Using the principle of leverage and making an analogy with the method of obtaining the 'center of gravity' for polygons, the ranking results of node importance under the DNI algorithm are analyzed to obtain the best taxi parking stand, and the practicality and effectiveness of the DNI algorithm are verified through experiments.

Key words: point of interest (POI), node importance, general node importance (GNI) algorithm, call model, dedicated node importance (DNI) algorithm