

文章编号: 1674—8247(2024)05—0087—05

DOI:10.12098/j.issn.1674-8247.2024.05.015

## 基于 PageRank 的交路选择算法研究

王若昆<sup>1</sup> 黄克勇<sup>1</sup> 邱 鹏<sup>2</sup> 李 露<sup>1</sup>

(1. 南京铁道职业技术学院, 南京 210031;

2. 新誉庞巴迪信号系统有限公司, 江苏 常州 213166)

**摘 要:** 交路设计是行车计划编制中的重要考虑因素。本文通过对信号系统设备布置结构和设备属性的分析,以静态配置数据的形式表达分析结果,采用图机器学习方法建立用于交路选择的图模型。基于 PageRank 算法,对图模型中的节点进行重要度分析,给出了节点的重要度排序。依据节点的重要度和交路结构,实现了交路分类,给出了一种交路选择算法,为轨道交路选择提供了一种参考方案。最后,以某线路为例进行了实验验证。结果表明,算法可以有效提高交路选择效率。

**关键词:** 轨道交通;图机器学习;交路选择;PageRank

中图分类号:U283

文献标志码:A

## Study on Route Selection Algorithm Based on PageRank

WANG Ruokun<sup>1</sup> HUANG Keyong<sup>1</sup> QIU Peng<sup>2</sup> LI Lu<sup>1</sup>

(1. Nanjing Vocational Institute of Railway Technology, Nanjing 210031, China;

2. Bombardier NUG Signaling Solutions Co., Ltd., Changzhou 213166, China)

**Abstract:** Route design is a critical consideration in the compilation of rail operating schedules. This study analyzed the arrangement structure and attributes of signaling system equipment, and expressed the findings in the form of static configuration data. A graph model for rail route selection was established using graph machine learning methods. Applying the PageRank algorithm to the graph model, the importance of nodes was analyzed, yielding a ranked list of node significance. Based on node importance and route structure, a route classification was achieved, leading to the formulation of a route selection algorithm. This offers a reference solution for route selection. Then one line was taken as an example to validate the proposed algorithm, demonstrating its effectiveness in enhancing route selection efficiency.

**Key words:** rail transit; graph machine learning; route selection; PageRank

全自动运行信号系统在国内多条轨道运营,依据发改产业[2017]2000号《增强制造业核心竞争力三年行动计划》<sup>[1]</sup>的要求,全自动运行系统是轨道信号系统的核心子系统,正与中国“复兴号”等成为我国重大装备亮点,是我国制造业竞争新优势以及发展新动能。

在轨道运营过程中,交路是指列车担当运输任务的固定周转区段,即列车从起始站至终端折返站之间往返运行的线路区段<sup>[2]</sup>。交路方案决定了列车运行的区段长度、折返车站、以及按不同交路运行的列车开行对数等,是行车计划确定的基础<sup>[3]</sup>。如何根据线路不同

收稿日期:2023-03-20

作者简介:王若昆(1988-),男,讲师。

基金项目:江苏省高等学校基础科学(自然科学)研究面上项目(22KJB580008);2024年江苏高校“青蓝工程”资助项目(5002024012-RCQL01);江苏省高等学校基础科学(自然科学)研究面上项目(23KJB580012)

引文格式:王若昆,黄克勇,邱鹏,等. 基于 PageRank 的交路选择算法研究[J]. 高速铁路技术,2024,15(5):87-91.

WANG Ruokun, HUANG Keyong, QIU Peng, et al. Study on Route Selection Algorithm Based on PageRank[J]. High Speed Railway Technology, 2024, 15(5):87-91.

区段的客流特征制定合理的运行交路,既能让运能符合客流需要,又能让列车运用更为经济高效,是行车计划编制过程中的首要考虑因素。本文对目前的交路选用方案进行改进,设计了一种基于 PageRank 的交路选择算法。该算法以静态配置数据构建图模型,将进路视作边(Link),信号机视作节点(Node),利用图机器学习方法,分析图模型中节点和边的重要性,并以此作为进路分类的依据,进而对交路进行类型判断,最终给出一种交路选择方案。

1 算法基础

算法是针对交路选择而提出的一种优化方案,主要基于静态配置数据和 PageRank 算法去实现。

1.1 静态数据

控制系统的设计理念伴随着计算机技术的发展,已经从早期的具体线路数据与控制逻辑一体化设计方案,向一般通用控制逻辑与具体线路配置数据相分离的设计方案转变<sup>[4]</sup>,这种设计方案大大降低了系统在设计开发阶段的工作。线路配置数据一般包括静态配置数据和动态配置数据。静态配置数据一般指在控制系统计算中所需的固定数据,比如轨道的长度、信号机的偏移量等。静态数据量大,包含内容较多,可以完整地数据角度描述一条线路的框架结构。

常见的静态配置数据包包括 ATP 静态数据包、ATO 静态数据包、ZC 静态数据包、ATS 静态数据包和 CI 静态数据包。每一个数据包中又包含多张数据表,依据控制系统的接口要求,每一个数据表中包含不同的数据格式或子数据表,具体结构如图 1 所示。



图 1 静态数据结构图

1.2 图机器学习

图(Graph)是一种用于描述并分析有关联的实体的一种普适语言<sup>[5]</sup>。它不将实体视为一系列孤立的点,而强调实体间的相互联系,是一种较好地描述领域知识的方式。当某一领域具有丰富的关系结构,可以用关系图表示。通常图可以表示为  $G(N,E)$  的形式,不仅可以图拓扑结构进行图表示,还可以将图中不同节点在性质上的相近或不同进行图表示,例如可以将某一个联锁区内的信号机聚类为一簇,也可以将具有折返性质的信号机聚类为一簇。

图机器学习是一种利用图结构的数据来进行机器学习任务的广泛方法。图机器学习包括多种技术和算法,它们用于处理和分析图数据,其中包括图表示学习、图神经网络。图表示学习是图机器学习的重要形式,它主要关注如何将图结构中的节点、边或全图表示成低维向量<sup>[6]</sup>。这些向量表示了节点、边或全图的特

征,可以用于描述节点、边或全图之间的相似性、关系等,并用于后续的机器学习任务,如节点分类、图分类、边分类等。图表示学习有多种方法,可以根据低维向量表示,将图表示学习分为两种,一种是用人工给定规则将节点、边或全图表示为低维向量,典型代表为 DeepWalk、Node2Vec 等算法,另一种是采用神经网络的方法,通过大量样本数据对神经网络模型进行训练,自动获取一种可以表征节点、边或全图特征的函数,将这种函数应用于待处理的数据集上,获取低维向量。本文采用的是第一种方法中的 PageRank 算法对图中的节点表示成低维向量。

2 算法实现

2.1 图的表示

线路静态配置数据是信号系统的重要组成部分,也是实现信号系统与具体项目的桥梁<sup>[7]</sup>。静态配置

数据可以完整地重现信号系统各个设备、对象、元素之间的连接关系<sup>[8]</sup>。因此,本文从静态配置数据中抽取与交路有关的信息,从信号系统设备之间的拓扑关系出发,建立设备之间的拓扑关系和属性连接关系进行图模型。

可以将单轨图纸中与进路有关的信息进行图表示为:

$$G = (N, E, X) \quad (1)$$

式中, $N$ 表示图中的节点,取信号系统中的信号机作为节点; $E$ 表示图中的边,取两两信号机之间的连接关系,即进路作为边, $E \subseteq (N \times N)$ ;由于任意两进路之间可能存在一定的交叉关系,采用 $X$ 来表示进路属性, $E \subseteq R^{N1 \times S}$ ,其中 $S$ 是指每一个节点的特征数量。

任意一个交路都是由有方向性的进路组成的,因此从静态配置数据抽象出的图模型应是一个有向图,边的方向就是进路的方向。

通过上述方法,将线路静态配置数据中的进路信息表示为图,形式如图2所示。

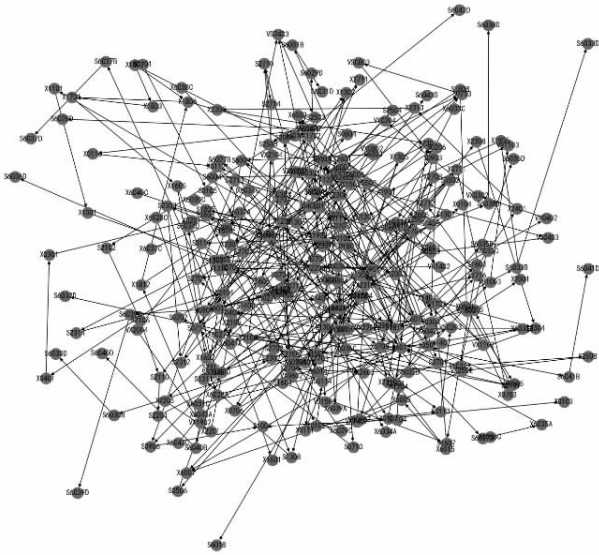


图2 进路的图表示图

由图2可知,对于任意节点 $N_i$ (信号机),均由进入连接(in\_Link)和输出连接(out\_Link)组成。

定义节点重要度为:相邻节点 $N_{near}$ 对本节点 $N_0$ 的依赖程度称节点重要度,记为 $\gamma$ 。如果任意一个节点 $N_i$ 有更多的 in\_Link,就认为 $N_0$ 的重要度高。 $\gamma$ 可以表示为:

$$\gamma = \sum_{i=0}^m \lambda_i \frac{\gamma_i}{d_i} \quad (2)$$

式中: $\gamma_i$ ——每一个 $N_{near}$ 的重要度;

$d_i$ ——每一个 $N_{near}$ 的 out\_Link 的个数;

$\lambda_i$ ——对于 $N_0$ 的相邻节点的权重值。

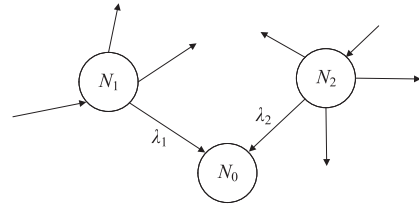


图3 节点重要度计算示意图

节点重要度计算过程如图3所示。图中 $N_0$ 表示当前节点, $N_1$ 和 $N_2$ 表示相邻节点,对于 $N_1$ 来说,有3个 out\_Link,因此 $N_1$ 对 $N_0$ 的影响即为 $\frac{\gamma_1}{3}$ ,同理, $N_2$ 对 $N_0$ 的影响为 $\frac{\gamma_2}{4}$ ,当取影响权重 $\lambda_i = 1$ 时, $N_0$ 的重要度可以表示为:

$$\gamma = \frac{\gamma_1}{3} + \frac{\gamma_2}{4} \quad (3)$$

由于需要对全图每一个节点进行重要度计算,可将当前图中所有节点重要度组成图节点重要向量 $V_\lambda$ ,且

$$V_\lambda = [\gamma_1, \gamma_2, \gamma_3, \dots, \gamma_n]^T \quad (4)$$

在计算过程中,当前节点 $N_0$ 的重要度在更新完成后,同时影响所有与之连接的相邻节点的重要度,因此在计算全图所有节点的重要度的过程中,需要对节点的重要度进行多次迭代,每次迭代之后,新的全图节点重要度向量 $V_\lambda^{i+1}$ 可以表示为:

$$V_\lambda^{i+1} = M \times V_\lambda^i \quad (5)$$

式中, $M$ 是迭代矩阵,矩阵中的元素 $m_{ij}$ 是节点 $N_i$ 对节点 $N_j$ 的重要度影响,即 $\lambda_i \frac{\gamma_i}{d_i}$ 。可以看出,如果迭代 $n$ 次,则

$$V_\lambda^{i+1} = M \times M, \dots, M \times V_\lambda^i \quad (6)$$

其中, $M$ 矩阵会出现 $n$ 次。通过多次迭代, $V_\lambda$ 最终收敛到一个稳定值,如节点数量过多,迭代次数可能会比较大,这时可以在不影响效果的情况下,设置一个阈值 $\zeta$ ,使得:

$$|V_\lambda^{n+1} - V_\lambda^n| < \zeta \quad (7)$$

即认为是稳定的,阈值 $\zeta$ 是一个超参数,需要根据项目的实际情况来确定。

采用上述方法,可以将所有的信号机按照重要性进行排序,实际上这是从信号设备的拓扑结构上对信号机的重要性进行了排序,而不仅仅要考虑图纸的拓扑结构,还需要考虑项目中节点本身的属性,比如某线路中所有的虚拟信号机均不作为交路的始端信号机,所以在设计中加入了属性向量 $X$ ,可以将所有需要的节点属性量化之后存储在 $X$ 中;同时,在迭代过程中,

每一个  $N_{near}$  对  $N_0$  的影响都设置了  $\lambda$  进行限制,使得相邻的不同节点对中心节点的影响有权重之分,例如可以让与  $N_0$  在同一股道的节点对其的影响较大,而跨股道有侧向道岔的节点对其影响较小等。属性向量  $X$  和影响因子  $\lambda$ , 均应视为超参数,需要根据实际项目进行调整。当影响因子  $\lambda = 1$  时,在不考虑属性向量  $X$  的情况下,该线路中所有节点的重要度排序如图 4 所示。

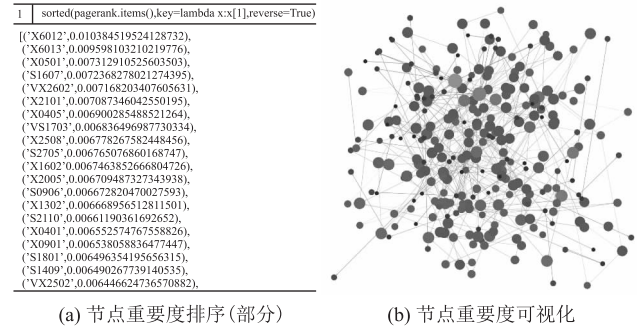


图4 节点重要度排序图

其中,图 4(a)给出了所有信号机归一化之后的权重排序,图 4(b)将所有的信号机以点的形式进行可视化展示,点越大代表权重越高,点之间的连接代表进路。

## 2.2 交路的分类

进路是指列车由一个地点到另一个地点所运行的径路,对于信号系统而言,进路就是从始端信号机开始到同方向的下一架信号机之间的径路。因此,可以使用进路始端信号机的重要度来衡量进路的重要度,即:

$$\gamma_{route} = \gamma_{route\_startSig} \quad (8)$$

在有向图中,  $E$  的重要度就是  $E$  的源节点的重要度,即

$$\gamma_E = \gamma_{E\_startNode} \quad (9)$$

可以将所有的  $E$  按照  $\gamma_E$  进行排序,形成集合  $Route_E$ ,  $Route_E$  是后续为交路进行分类的重要依据。

运营中的交路  $L$  实际上是由多个进路组成的一段闭合路径,也就是说交路是一个进路的有限集合,即:

$$L = \{Route_1, Route_2, \dots, Route_n \mid P(Route_1, Route_2, \dots, Route_n)\} \quad (10)$$

其中,  $P(Route_1, Route_2, \dots, Route_n)$  是指交路中所有  $Route_1, Route_2, \dots, Route_n$  均需满足首尾相连接的原则,即  $Route_i$  的终端节点就是  $Route_{i+1}$  的源节点。重要度的定义是根据信号设备的拓扑结构来进行衡量的,因此对于重要度高的  $E$ , 即进路,会多次被不同交路征用。交路中每一条进路的创建时间  $Route_{Time}$  是不同的,这取决于该进路的相关设备转换时间,比如进路中道岔的转换时间,进路关联的保护区段中道岔的转换

时间,可以将  $Route_{Time}$  表示为:

$$Route_{Time} = f(D_{route}, Speed_{start}, D_{overlap}, \dots) \quad (11)$$

其中,  $D_{route}$  是指进路中的道岔;  $Speed_{start}$  是指进路入口速度;  $D_{overlap}$  是指进路的保护区段中的道岔。如果交路  $L_i$  中包含的进路集合中的进路创建时间都很长,那么当前后列车运行在不同交路时,会增加交路的创建时间。

因此,本文以  $E$  的重要度为基础,对交路进行分类。将所有  $E$  按照  $\gamma_E$  进行排序,选择其中的前  $m$  个组成重要进路集合:

$$Route_{important} = \{route \mid \gamma_E \subseteq \{Route_E^1, Route_E^2, \dots, Route_E^m\}\} \quad (12)$$

其中,  $Route_E^i$  是指集合  $Route_E$  中的第  $i$  个。以  $Route_{important}$  中的每一个元素为分类基准,进行分类,即将所有的交路分为  $m$  类,即:

$$L = L_1 \cup L_2 \cup \dots \cup L_m \quad (13)$$

其中,  $Route_E^1 \subseteq L_1$  且  $Route_E^2, \dots, Route_E^m \notin L_1$ ,  $Route_E^2 \subseteq L_2$  且  $Route_E^1, Route_E^3, \dots, Route_E^m \notin L_2$ ,  $\dots$ ,  $Route_E^m \subseteq L_m$  且  $Route_E^1, \dots, Route_E^{m-1} \notin L_m$ 。

定义平行交路集合为:当需要从某一节点  $N_i$  到另一节点  $N_j$  的交路时,由于平行进路的存在,可能有多个交路满足这一要求,这种交路集合称之为平行交路集合,记作  $L_{NiNj}$ 。

进路创建仿真平台中,节点  $N_1$  与节点  $N_7$  之间的每一条交路的创建时间如图 5 所示。

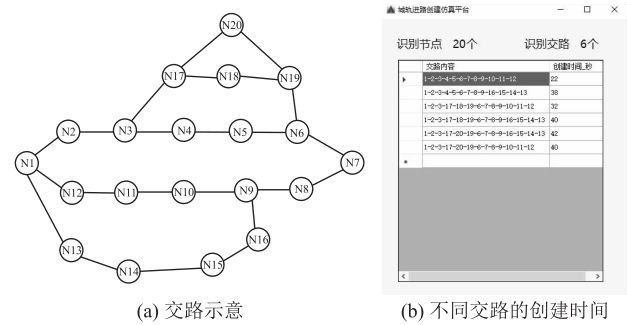


图5 不同种类交路创建时间对比图

可以发现,交路中包含  $Route_{important}$  中元素的交路,即:

$$L \cap Route_{important} \neq \varnothing \quad (14)$$

创建时间较短,反之则创建时间较长。因此在交路的选择中,从效率角度考虑,应尽量选择包含  $Route_{important}$  中元素的交路作为运行交路。

通过这种交路分类方法,可以提供两种运行策略供选择。第一,当多个交路的起点和终点相同时,可以优化选择,避免选择交路中进路集合创建时间长的交

路。第二,在交路的选择中,不仅要考虑单个列车的运行情况,更要关注整体。算法会对当前列车的前序列车交路进行分析,并在选择当前列车交路时,尽量减少对前序列车交路的影响,避免出现设备反复调转的情况,从而提高运行效率。

3 结束语

在“提质增效”的大背景下,合理的交路选择算法是提高轨道交通运行效率的有效手段。本文将信号设备拓扑结构和属性信息以静态数据的形式,通过图机器学习方法,进行节点重要度分析。依据节点的重要度,给出了一种交路选择算法,为交路选择提供了一种参考方案。

参考文献:

[1] 发改产业[2017]2000 号,增强制造业核心竞争力三年行动计划[S].  
Fa Gai Chan Ye[2017]No 2000, Three-year Action Plan to Enhance the Core Competitiveness of the Manufacturing Industry [S].  
[2] 李国芳,郭少琛,瞿何舟.城市轨道交通 Y 型交路快慢车停站方案优化[J].综合运输,2023,45(2):60-65.  
LI Guofang, GUO Shaochen, QU Hezhou. Optimization of Stop Scheme for Urban Rail Transit Based on Y-type Routing with Express and Local Train[J]. China Transportation Review, 2023, 45(2): 60-65.  
[3] 韩嘉.基于客流时空分布特性的南宁轨道交通 1 号线列车运行图交路优化[J].城市轨道交通研究,2023,26(2):51-54.

HAN Jia. Optimization of Nanning Rail Transit Line 1 Train Diagram Routing Based on Passenger Flow Spatial Temporal Distribution Characteristics[J]. Urban Mass Transit, 2023, 26(2): 51-54.  
[4] 翁顺,吴俐滢,朱宏平,等.基于模糊聚类的高速列车车速识别[J].高速铁路技术,2022,13(4):13-23,29.  
WENG Shun, WU Liying, ZHU Hongping, et al. Speed Recognition of High-speed Train Based on Fuzzy Clustering [J]. High Speed Railway Technology, 2022, 13(4): 13-23, 29.  
[5] XU Yong, Xin JIA, WANG Zhe, et al. Transit Network Models and Optimal Path Selection Algorithm for the Integrated Bus and Subway System [J]. CAAI Transactions on Intelligent Systems, 2015, 10(3):482-487.  
[6] 支瑞聪,万菲,张德政.零样本图学习综述[J].计算机系统应用,2022,31(5):1-20.  
ZHI Ruicong, WAN Fei, ZHANG Dezheng, Overview on Graph-based Zero-shot Learning [J]. Computer System and Applications, 2022, 31(5):1-20.  
[7] 汤曦,王义.基于统一坐标系的多源数据入库方法设计[J].高速铁路技术,2022,13(6):52-56.  
TANG Xi, WANG Yi. Design of a Method for Multi-source Data Warehousing Based on the Unified Coordinate System[J]. High Speed Railway Technology, 2022,13(6):52-56.  
[8] 朱莹莹,王英波.信号设备数据配置工具在 CBTC 中的研究与实现[J].铁路通信信号工程技术,2019,16(9):61-64.  
ZHU Yingying, WANG Yingbo. Research and Implementation of Signal Equipment Data Configuration Tool in CBTC System [J]. Railway Signalling & Communication Engineering, 2019, 16(9): 61-64.

(上接第 52 页)

[9] 宋盼,徐秀兰.高压脉冲轨道电路和 25Hz 轨道电路时间特性匹配解决方案[J].铁路通信信号工程技术,2019,16(7):89-92.  
SONG Pan, XU Xiulan. Solution to Time Characteristic Matching Problem of High-voltage Pulse Track Circuit and 25 Hz Phase-sensitive Track Circuit [J]. Railway Signalling & Communication Engineering, 2019, 16(7): 89-92.  
[10] 高俊明,李明兵,唐乾坤,等.400 km/h 高速铁路 ZPW-2000 轨道电路器材适应性研究[J].高速铁路技术,2022,13(1):83-

88.  
GAO Junming, LI Mingbing, TANG Qiankun, et al. Study on Adaptability of ZPW-2000 Track Circuit Equipment for 400 km/h High-speed Railway [J]. High Speed Railway Technology, 2022, 13(1): 83-88.  
[11] 马建忠.高速铁路区间轨道电路“红光带”故障处置方法与研究[J].高速铁路技术,2019,10(6):49-53.  
MA Jianzhong. Research and Treatment Method of “Red Light Band” Fault in Section Track Circuit of High Speed Railway[J]. High Speed Railway Technology, 2019, 10(6): 49-53.